ORIGINAL ARTICLE

CrossMark

Computational-based structural, functional and phylogenetic analysis of *Enterobacter* phytases

Krishnendu Pramanik[1] · Shreyasi Kundu[1] · Sandipan Banerjee[1] · Pallab Kumar Ghosh[1] · Tushar Kanti Maiti[1]

Abstract

Myo-inositol hexakisphosphate phosphohydrolases (i.e., phytases) are known to be a very important enzyme responsible for solubilization of insoluble phosphates. In the present study, *Enterobacter* phytases have characterized by different phylogenetic, structural and functional parameters using some standard bio-computational tools. Results showed that majority of the *Enterobacter* phytases are acidic in nature as most of the isoelectric points were under 7.0. The aliphatic indices predicted for the selected proteins were below 40 indicating their thermostable nature. The average molecular weight of the proteins was 48 kDa. The lower values of GRAVY of the said proteins implied that they have better interactions with water. Secondary structure prediction revealed that alpha-helical content was highest among the other forms such as sheets, coils, etc. Moreover, the predicted 3D structure of *Enterobacter* phytases divulged that the proteins consisted of four monomeric polypeptide chains i.e., it was a tetrameric protein. The predicted tertiary model of *E. aerogenes* (A0A0M3HCJ2) was deposited in Protein Model Database (Acc. No.: PM0080561) for further utilization after a thorough quality check from QMEAN and SAVES server. Functional analysis supported their classification as histidine acid phosphatases. Besides, multiple sequence alignment revealed that "DG–DP–LG" was the most highly conserved residues within the *Enterobacter* phytases. Thus, the present study will be useful in selecting suitable phytase-producing microbe exclusively for using in the animal food industry as a food additive.

Keywords *Enterobacter* spp. · Myo-inositol hexakisphosphate phosphohydrolases · In silico analysis · Histidine phosphatase superfamily · Acidic phytases · Thermostable

Abbreviations

UniProtKB	The Universal Protein Knowledgebase
MEGA	Molecular evolutionary genetics analysis
MW	Molecular weight
pI	Isoelectric point
EC	Extinction coefficient
AI	Aliphatic index
II	Instability index
GRAVY	Grand average of hydropathicities
QMEAN	Qualitative Model Energy ANalysis
SAVES	The structure analysis and verification server

✉ Tushar Kanti Maiti
tkmbu@yahoo.co.in

1   Microbiology Laboratory, Department of Botany, The University of Burdwan, Purba Bardhaman, West Bengal 713104, India

Introduction

Phytic acid is an indigestible organic form of phosphorus, which is found in cereal grains, legume plants, oilseeds, nuts, tubers and different types of seeds (Nakashima et al. 2007; Greiner and Konietzny 2016). Myo-inositol hexakisphosphate phosphohydrolases (i.e., phytases) belong to a subclass of phosphatase enzymes which break down the phytic acid by hydrolyzing the phosphate residues from phytic acid (myo-inositol hexakisphosphate). There are several types of phytases, such as 3-phytase (EC 3.1.3.8), 4-phytase (EC 3.1.3.26) and 5-phytase (EC 3.1.3.72) which are involved in the liberation of P moiety from phytic acid at the position C3, C4 and C5, respectively. These enzymes have a molecular mass between 40 and 70 kDa (Kumar et al. 2010; Afinah et al. 2010).

Ruminant animal utilized phytate through the action of phytases produced by microbial flora in the rumen. The anaerobic gut fungi and bacteria found in the ruminants are responsible for the phytate hydrolysis within the rumen.

The hydrolyzed inorganic phosphate from the phytate is utilized by both the microflora and the ruminant host (Yanke et al. 1998). However, the situation is quite different with monogastric animals such as chicken, swine, fish and birds are not able to utilize phytase phosphorus because they lack phytase-degrading enzymes or lack the insufficient amount of phytases in their digestive tracts (Gupta et al. 2015; Dahiya 2016). Hence, phytate in grains passed through the gastrointestinal tract without being digested properly (Niño-Gómez et al. 2017). Since phosphorus is an essential requirement for bone formation, deficiency of this mineral would cause of osteoporosis, tooth decay, narrow jaws, mental retardation, rickets, etc. (Berlyne et al. 1973). Phytic acid can inhibit the nutrient absorption by the intestine and also prevent the different types of ion uptake such as Zn^{2+}, Ca^{2+}, Mg^{2+}, Cu^{2+} and digestive enzymes such as pepsin, trypsin, amylase, etc. (Gupta et al. 2015). Here, phytic acid acts as an anti-nutritional agent (Escobin-Mopera et al. 2012; El-Toukhy et al. 2013). The phytase catalyzes the hydrolysis of phytic acid and converts the organic and insoluble form of phosphorus to more soluble form and ultimately improve and facilitate intestinal absorption (Wise 1982). Importance of the phytase protein can be evidenced by many earlier workers (Konietzny et al. 1994; Mullaney et al. 2000; He and Honeycutt 2001; Lei et al. 2007, 2013; Rodriguez et al. 2000; Tazisong et al. 2008; Yao et al. 2012; He et al. 2013; Gontia-Mishra and Tiwari 2013; Zając et al. 2018).

Phytases have been reported in many fungi, e.g., *Aspergillus niger* (George et al. 2007), *Aspergillus fumigatus* (Zhang et al. 2007), *Peniophora lycii* (George et al. 2007), *Penicillium simplicissimum* (Tseng et al. 2000), *Candida krusei* (Quan et al. 2002), *Debaryomyces castellii* (Ragon et al. 2009) and *Fusarium oxysporum* (Gontia-Mishra et al. 2014). Moreover, a number of gram-positive- and gram-negative bacteria, e.g., *Bacillus subtilis* (Kerovuo et al. 1998; El-Toukhy et al. 2013), *Bacillus licheniformis* (Kumar et al. 2014), *Pseudomonas* spp. (Richardson and Hadobas 1997), *Klebsiella pneumoniae* (Escobin-Mopera et al. 2012), *Klebsiella terrigena* (Greiner et al. 1997), *Enterobacter* sp. (Kalsi et al. 2016), *Serratia* sp. (Zhang et al. 2011; Kalsi et al. 2016), and *Yersinia mollaretii* (Shivange et al. 2012) have also been reported earlier to possess this important enzyme.

Phytase enzyme has many beneficial roles in different industries, agriculture, poultry farm, etc. In poultry farm, phytase is used for overcoming the phosphorus deficiency (Elkhalil et al. 2007). Natuphos® and Ronozyme® are the two commercially available phytases derived from *Aspergillus*, one of the most abundant extracellular producers of phytase (Casey and Walsh 2004). This phytase was actually a recombinant phytase produced by the expression of the phyA gene of *Aspergillus ficuum* in *Aspergillus niger*, first produced in 1994 (Casey and Walsh 2004). However, few studies reported that bacterial phytases were also promising

for commercial phytase production (Konietzny and Greiner 2004; Kalsi et al. 2016; Niño-Gómez et al. 2017). In fact, bacterial phytases are reported to be more beneficial over the fungal phytases due to higher substrate specificity, better catalytic capacity and enhanced resistance to proteolysis (Konietzny and Greiner 2004). In addition, phytase-producing bacteria are known to involve in plant growth promotion also (Singh et al. 2014).

Apart from isolating and characterizing various types of phytases, extensive computational investigations of microbial phytases have been worked out to discover many unexplored features lying within the sequences (Kumar et al. 2012, 2014; Kumar and Agrawal 2014; Gontia-Mishra et al. 2014; Mathew et al. 2014; Verma et al. 2016; Ebrahimi et al. 2016; Niño-Gómez et al. 2017; Pramanik et al. 2018). The outcomes of these studies are biotechnologically advantageous to utilize them in agricultural and industrial perspectives.

This study describes the use of various bio-computational tools to perform the phylogenetic, structural and functional analyses of the phytase enzyme of *Enterobacter* spp. to understand its distinctive properties essential for its industrial application.

Materials and methods

Sequence retrieval and phylogenetic analysis

UniProtKB server is the knowledge base of Swiss Institute of Bioinformatics and a comprehensive source of protein sequences and annotation database (Apweiler et al. 2004). This proteomics-based source server (http://www.uniprot.org) was used for retrieval of phytase sequences of different spp. of *Enterobacter* (Gram-negative-, rod-shaped bacteria). For this, *Enterobacter aerogenes* GN = ASV18_08060 (Protein Acc. no.: A0A0M3HCJ2) was first selected as the query sequence and BLAST search was done. A total of 16 sequences of different spp. of *Enterobacter* were selected (based on highest sequence identity, lowest E-value, maximum query coverage and bit score) and downloaded from UniProtKB in FASTA format for progressing to the computational investigation. Furthermore, the sequences were analyzed phylogenetically to decipher the evolutionary relationship among the selected proteins of interest by using MEGA7 software (Kumar et al. 2016). Percent similarity index within the sequences was calculated by Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/).

Physiochemical characterization

ExPASy—ProtParam is a tool which is used for the calculation of different physical and chemical parameters for a

protein (Gasteiger et al. 2005). This tool (http://www.espas y.org/protparam) was used to calculate physiochemical properties of retrieved sequences. Various computed parameters like amino acid composition, molecular weight (MW), isoelectric point (pI), molar extinction coefficient (EC), aliphatic index (AI), instability index (II) and grand average of hydropathicities (GRAVY) were calculated using this server.

Secondary structure prediction

Secondary structure (alpha-helix, 3_{10} helix, pi helix, beta bridge, extended strand, beta-turn, bend region, random coil etc.) of a protein refers to the interaction of H-bond donor and acceptor residues of a polypeptide chain. Practically, secondary structural arrangements can be accurately done by nuclear magnetic resonance (NMR) spectroscopy (Meiler and Baker 2003). But some useful and easy computational tools used for in silico prediction of secondary structures of the protein. Here, secondary arrangements were predicted by SOPMA web-based server (Geourjon and Deleage 1995). For this, 16 protein sequences of different *Enterobacter* spp. were analyzed.

Homology protein modeling, its evaluation and submission

Homology 3D protein modeling of selected 16 sequences was performed using Swiss model workspace (Biasini et al. 2014) selecting its suitable and best-matched template. Evaluation of the predicted protein model was done in both QMEAN (Benkert et al. 2009) and SAVES server (http://services.mbi. ucla.edu/SAVES/). Based on the evaluation report the best-built model was finally submitted to Protein Model Database (PMDB) (https://bioinformatics.cineca.it/PMDB/) and the PMDB ID was obtained.

Functional analysis and protein–protein interaction

To find out the functional motifs and superfamily of the protein sequence, Motif Finder (http://www.genome.jp/tools/motif/) was used for the analyses. Highly conserved domains among the sequences were analyzed by Clustal Omega (https://www. ebi.ac.uk/Tools/msa/clustalo/). To know the interaction of *Enterobacter* phytase with other closely associated proteins STRING v10.0 (http://string-db.org/) server was used and functional protein association network was generated.

Results and discussion

Sequence retrieval and phylogenetic analysis

Amino acid sequences of total 16 bacterial strains (*Enterobacter kobei* GN = ASV01_18150, *Enterobacter* sp. GN02225 GN = YA53_05950, *Enterobacter* sp. GN02186 GN = YA50_01625, *Enterobacter* sp. 35730 GN = SS33_10495, *Enterobacter asburiae* GN = ABF78_15640, *Enterobacter* sp. 35699 GN = SS37_20000, *Enterobacter* sp. 50588862 GN = APT89_19470, *Enterobacter hormaechei* GN = ASU65_02620, *Enterobacter* sp. GN02174 GN = YA49_02765, *Enterobacter* sp. 35027 GN = SS00_18385, *Enterobacter* sp. GN02616 GN = ABR35_11890, *Enterobacter* sp. 42324 GN = SS44_19905, *Enterobacter sakazakii* GN = AFK64_08585, *Enterobacter agglomerans, Enterobacter cancerogenus* GN = NH00_23800, *Enterobacter aerogenes* GN = ASV18_08060) were phylogenetically analyzed to predict their evolutionary interrelationship (Fig. 1). *Enterobacter aerogenes* (A0A0M3HCJ2) clustered only with *E. cancerogenus* (A0A0A3YJT6) and their closest neighbour is *E. agglomerans* (F2VRZ7) present in the same clade. Rest of the 13 taxa were in a separate clade showing interrelationships within them. This might be correlated with the percent similarity index of the phytase sequences of the strains with the selected query phytase protein, A0A0M3HCJ2 (Fig. 2). It was found (from the percent similarity index in Clustal Omega) that A0A0M-3HCJ2 showed the highest similarity with A0A0A3YJT6 (68.65%) and F2VRZ7 was next to it (66.43%). These percent similarity indices were also manifested by the closest clustering pattern of these three sequences (Fig. 1). Interestingly, rest 13 sequences showed percent similarity in between 19.46–24.11% with the A0A0M3HCJ2 sequence (Fig. 2). This was the principle reason behind forming a separate clade of 13 taxa from the rest three sequences in the phylogenetic tree of phytase proteins (Fig. 1). From this analysis, it can be interpreted that there might be a correlation among the selected taxa based on their protein sequences. The similar phylogenetic analysis was also made by some earlier workers to decipher the evolutionary significance of different taxa solely based on their protein sequences (Verma et al. 2016; Pramanik et al. 2017a, b, 2018).

Physicochemical characterization

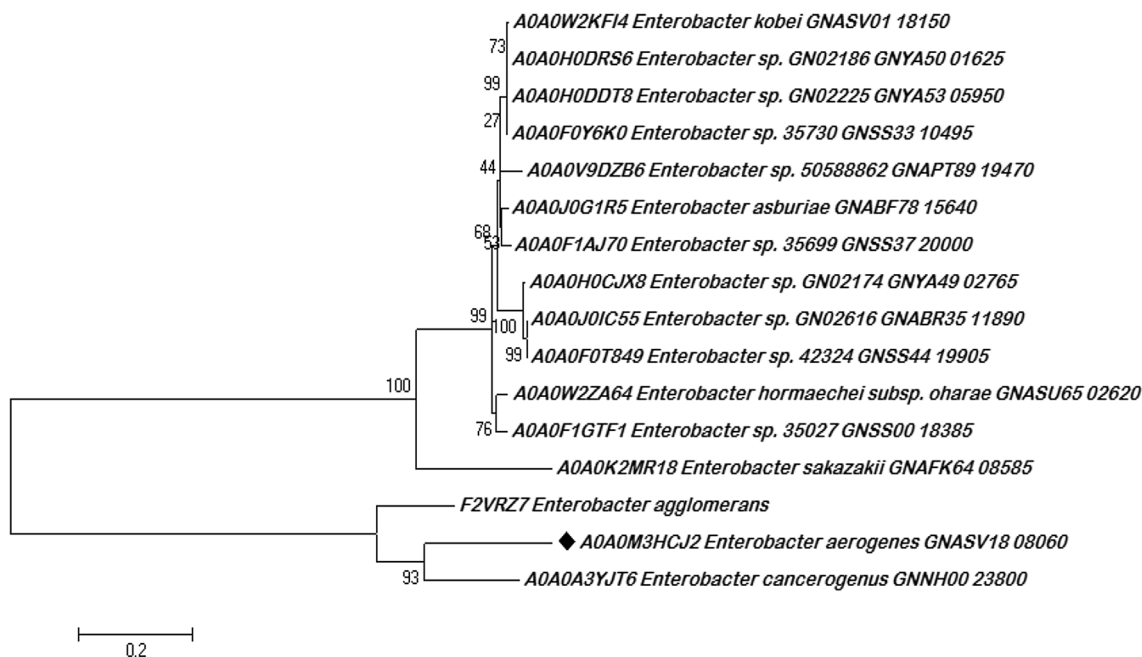Computational-based analysis about the physicochemical behavior of the proteins gives a theoretical overview of

Fig. 1 Phylogenetic classification of phytase proteins from 16 different *Enterobacter* spp
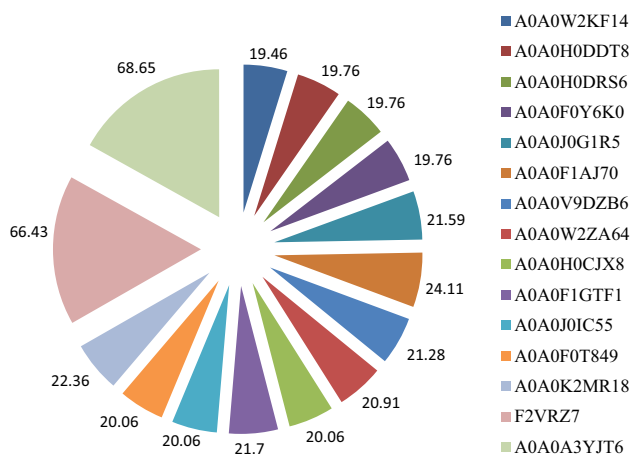

Fig. 2 Calculation of percent similarity index of the selected phytase proteins with the query protein (A0A0M3HCJ2)

the nature of the protein. In this study, all the 16 phytase sequences were physicochemically characterized by a number of bio-computational tools (Table 1; Fig. 3). The analysis revealed that the isoelectric point ranges from 5.03 to 8.90 which is said to be a wide range. The isoelectric point of the proteins below the neutral pH can be said to be acidic and above it can be regarded as alkaline. Here, 13 strains showed acid phytases and rest 3 phytases were basic/alkaline in nature. Acid phytases are evident in the work of Zhang et al. (2011) while basic one was reported by Tran et al. (2011) while working on other bacterial phytases. Acid phytases were also evident from

the in silico analysis of *Aspergillus niger* (Niño-Gómez et al. 2017). Except for the four strains (A0A0W2KF14, F2VRZ7, A0A0A3YJT6 and A0A0M3HCJ2), all strains showed instability index below 40 (Table 1). Thus, the majority of the proteins can be said as stable. Besides, higher ranges of aliphatic index prove the protein as thermostable in nature (Ikai 1980). Thermostable phytase from *Enterobacter* sp. was reported by Kalsi et al. (2016) which corroborated the present work. Besides higher aliphatic indices were also reported in *Aspergillus niger* (Niño-Gómez et al. 2017) while working on the prediction study of both 3-phytase A and 3-phytase B. Assuming all pairs of Cys residues form cystines, the molar extinction coefficients were calculated in water at 280 nm. However, EC ranges from 35,535 to 81,275 M^{-1} cm^{-1} (Table 1). The average molecular weight of the proteins is around 48 kDa. The GRAVY, calculated from ExPASy showed very low (Table 1) which implies that the proteins have better interactions with water. This interpretation was also corroborated by the work of Mathew et al. (2014), Verma et al. (2016). Moreover, the strains showed the difference in composition of amino acid residues of their respective proteins as presented graphically in Fig. 3.

Secondary structure prediction

The predicted secondary arrangements of different *Enterobacter* spp. revealed mainly four types of secondary elements which were alpha helices, random coils, extended strands and beta turns (Table 2). The alpha-helical content was the

Table 1 Physicochemical features of 16 *Enterobacter* spp. with their respective protein and gene accession numbers

Name of the bacteria	Protein Acc.no.	pI	II	AI	EC	MW	GRAVY
Enterobacter kobei GN=ASV01_18150	A0A0W2KF14	8.90	46.39	87.67	70,275	46.54	− 0.300
Enterobacter sp. GN02225 GN=YA53_05950	A0A0H0DDT8	5.80	22.35	86.15	37,025	48.61	− 0.330
Enterobacter sp. GN02186 GN=YA50_01625	A0A0H0DRS6	5.80	22.35	86.15	37,025	48.61	− 0.330
Enterobacter sp. 35,730 GN=SS33_10495	A0A0F0Y6K0	5.98	22.35	86.15	37,025	48.63	− 0.337
Enterobacter asburiae GN=ABF78_15640	A0A0J0G1R5	5.41	22.80	86.79	37,025	48.54	− 0.300
Enterobacter sp. 35,699 GN=SS37_20000	A0A0F1AJ70	5.51	21.62	87.04	37,025	48.54	− 0.314
Enterobacter sp. 50,588,862 GN=APT89_19470	A0A0V9DZB6	5.70	24.71	87.46	35,535	48.61	− 0.298
Enterobacter hormaechei GN=ASU65_02620	A0A0W2ZA64	5.41	19.74	88.74	37,025	48.63	− 0.287
Enterobacter sp. GN02174 GN=YA49_02765	A0A0H0CJX8	5.40	21.60	89.40	37,025	48.63	− 0.291
Enterobacter sp. 35,027 GN=SS00_18385	A0A0F1GTF1	5.79	23.90	90.07	35,535	48.49	− 0.255
Enterobacter sp. GN02616 GN=ABR35_11890	A0A0J0IC55	5.89	21.85	88.32	37,025	48.65	− 0.304
Enterobacter sp. 42,324 GN=SS44_19905	A0A0F0T849	5.89	21.85	88.32	37,025	48.65	− 0.304
Enterobacter sakazakii GN=AFK64_08585	A0A0K2MR18	5.03	31.44	89.58	38,515	48.84	− 0.338
Enterobacter agglomerans	F2VRZ7	8.42	46.84	92.22	80,245	46.30	− 0.238
Enterobacter cancerogenus GN=NH00_23800	A0A0A3YJT6	6.31	45.44	83.89	81,275	46.78	− 0.260
Enterobacter aerogenes GN=ASV18_08060	A0A0M3HCJ2	8.90	46.39	87.67	70,275	46.54	− 0.300

pI isoelectric point, *MW* molecular weight, *II* instability index, *AI* aliphatic index, *EC* extinction coefficient, *GRAVY* grand average of hydropathicity
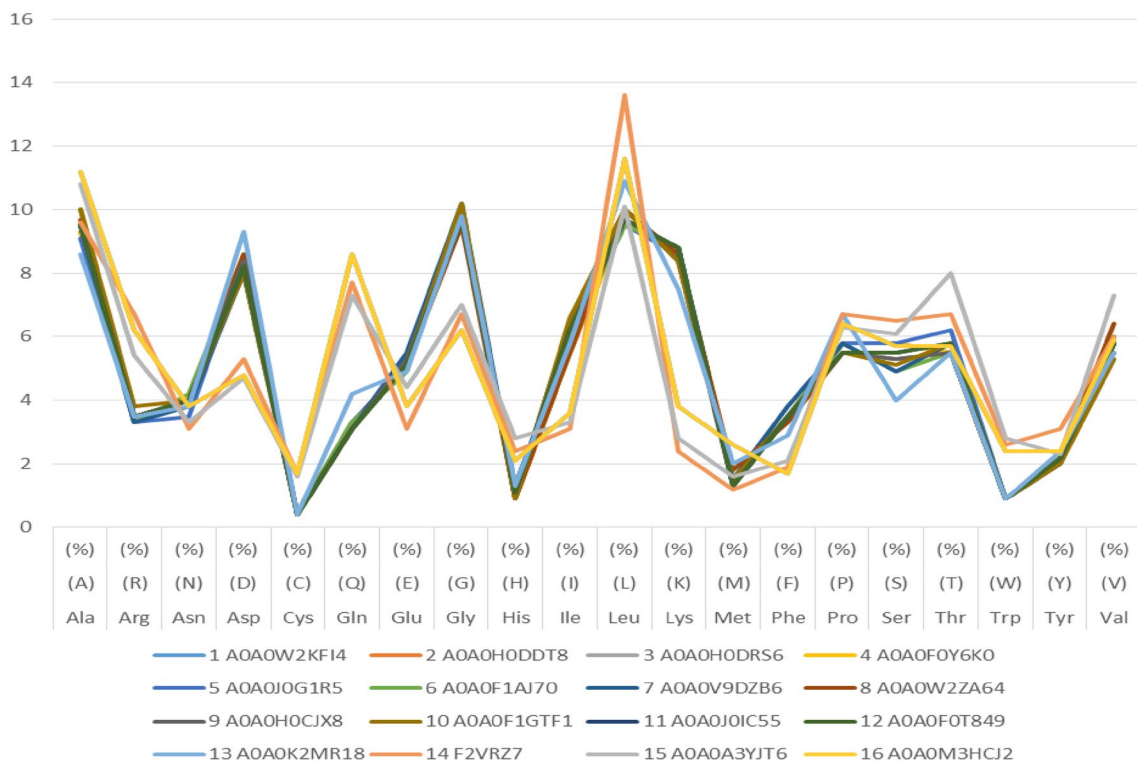

Fig. 3 Graphical overview of composition difference of amino acids among the selected 16 strains of *Enterobacter* spp

highest (42.76%) as found in case of *E. aerogenes* (A0A0M-3HCJ2) among all of the *Enterobacter* spp. (Table 2, Suppl. 1). Niño-Gómez et al. (2017) also showed that 43 and 38%

α-helices were found in the computational investigation of *Aspergillus niger* 3-phytase A and 3-phytase B, respectively. This indicated the thermostable nature of the protein

Table 2 Comparison of secondary and tertiary protein structures among different spp. of *Enterobacter*

Serial no.	Name of the *Enterobacter* spp.	Protein Acc. No.	Contents of principal secondary structural elements				Quality assessment of tertiary structure		
			Alpha-helix (%)	Extended strand (%)	Beta-turn (%)	Coil (%)	QMEAN score*	SAVES ERRAT overall quality factor	Amino acids in favoured region in Ramachandran plot (%)
1	*Enterobacter kobei* GN=ASV01_18150	A0A0W2KF14	14.38	30.53	11.95	43.14	−6.46 👎	85.641	83.5
2	*Enterobacter* sp. GN02225 GN=YA53_05950	A0A0H0DDT8	15.27	29.87	11.95	42.92	−6.46 👎	85.641	83.5
3	*Enterobacter* sp. GN02186 GN=YA50_01625	A0A0H0DRS6	15.27	29.87	11.95	42.92	−6.46 👎	85.641	83.5
4	*Enterobacter* sp. 35,730 GN=SS33_10495	A0A0F0Y6K0	15.27	29.87	12.39	42.48	−6.46 👎	85.641	83.5
5	*Enterobacter asburiae* GN=ABF78_15640	A0A0J0G1R5	14.82	32.30	12.83	40.04	−7.02 👎	62.944	83.5
6	*Enterobacter* sp. 35,699 GN=SS37_20000	A0A0F1AJ70	16.59	30.75	12.83	39.82	−7.37 👎	72.959	84.8
7	*Enterobacter* sp. 50,588,862 GN=APT89_19470	A0A0V9DZB6	14.60	32.30	14.16	38.94	−7.71 👎	82.673	83.5
8	*Enterobacter hormaechei* GN=ASU65_02620	A0A0W2ZA64	17.48	30.31	12.39	39.82	−7.04 👎	87.864	85.7
9	*Enterobacter* sp. GN02174 GN=YA49_02765	A0A0H0CJX8	17.92	30.53	12.39	39.16	−6.25 👎	77.885	82.6
10	*Enterobacter* sp. 35,027 GN=SS00_18385	A0A0F1GTF1	17.48	31.19	12.61	38.72	−7.38 👎	72.959	84.8
11	*Enterobacter* sp. GN02616 GN=ABR35_11890	A0A0J0IC55	18.58	29.65	12.39	39.38	−6.66 👎	78.218	82.2
12	*Enterobacter* sp. 42,324 GN=SS44_19905	A0A0F0T849	18.58	29.65	12.39	39.38	−6.67 👎	78.818	82.6
13	*Enterobacter sakazakii* GN=AFK64_08585	A0A0K2MR18	18.18	26.16	11.75	43.90	−6.11 👎	71.707	85.6
14	*Enterobacter agglomerans*	F2VRZ7	39.95	15.55	9.57	34.93	0.32 👍	96.026	96.1
15	*Enterobacter cancerogenus* GN=NH00_23800	A0A0A3YJT6	41.22	13.82	8.20	36.77	−6.67 👎	87.080	94.9
16	*Enterobacter aerogenes* GN=ASV18_08060	A0A0M3HCJ2	42.76	11.16	8.79	37.29	0.25 👍	96.141	97.9

*👍 indicates acceptable QMEAN score; 👎 indicates not-acceptable QMEAN score

as Kumar et al. (1999) indicated that α-helical conformations are abundant in case of thermophiles to withstand high temperatures. These types of structural findings were also performed by a number of bioinformatics researchers. A similar prediction of secondary structures of the residual interactions of *Aspergillus fumigatus* phytase was done by

Zhang et al. (2007) and *Bacillus* phytases were done by Verma et al. (2016).

Homology protein modeling, its evaluation and submission

Although homology modeling was performed for all 16 proteins (Suppl. 1), *Enterobacter aerogenes* (A0A0M-3HCJ2) was selected as representative species (Fig. 4) to elucidate phytase protein structure of *Enterobacter* spp. based on QMEAN score, an overall quality factor from SAVES server and Ramachandran plot (Table 2, Suppl. 2). The predicted 3D protein modeling of *Enterobacter aerogenes* (A0A0M3HCJ2) divulged that the protein consisted of four monomeric polypeptide chains, i.e., it was a tetrameric protein (Fig. 4a). This structural finding was also corroborated by some earlier reports. Ragon et al. (2009) reported tetrameric phytase in fungal species—*Debaryomyces castellii* CBS 2923 and Shivange et al. (2012) reported in the bacterium—*Yersinia mollaretii*. Helix, sheet and loops are differentially coloured and presented in Fig. 4b. Besides distinct disulfide bridges were demarked as green spherical (s) in Fig. 4b. These disulfide bridges which are formed by the oxidation of thiol groups of the cystine residues (Trivedi et al. 2009) are one of the factors for the stability of a protein (Cheng et al. 2007). After the construction of 3D (.pdb) model, the evaluation and quality estimation of the model was performed (Table 2,

Suppl. 2–3) and Ramachandran plot was built to show the positions allocated for each amino acid residues (Table 2, Suppl. 2). Analysis of Ramachandran plot for the PDB structure of *E. aerogenes* (A0A0M3HCJ2) protein showed that 97.9% residues were present in the most favoured region (Table 2, Suppl. 2). Presence of more than 90% residues in the favoured region of Ramachandran plot is the characteristics of the good quality model (Yadav et al. 2013). QMEAN4, QMEAN6 and Z-score calculated were 0.25, −0.26 and <1, respectively (Table 2, Suppl. 2). The desirable QMEAN scores and Z-score should be within 0–1 (Berman et al. 2000) and <1 Benkert et al. (2009), respectively, in comparison with a non-redundant set of PDB structures to obtain a high-quality model. Moreover, the overall quality factor by the SAVES ERRAT was 96.141% (i.e., >95%) (Suppl. 3). A good, high-resolution structure should score 95% or higher as an overall quality factor (Benkert et al. 2009). However, comprehensive evaluation of the predicted model proves to be a higher resolution model (> 3 Å) as determined by both QMEAN (Table 2, Suppl. 2) and SAVES server (Table 2, Suppl. 3). A similar type of model validation was also conducted by Pramanik et al. (2018) while working with *Klebsiella* phytases. Finally, the model (in .pdb format) was deposited in PMDB database and its accession number obtained is PM0080561 and the said model can be retrieved for any further investigations in future.
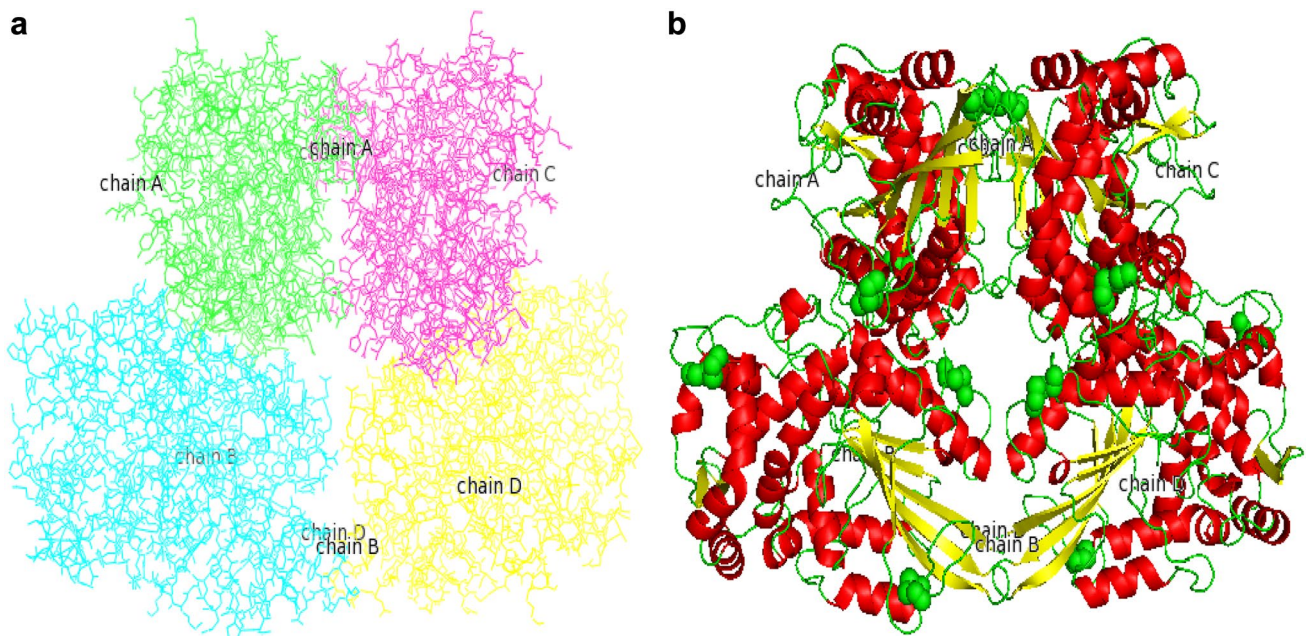
Fig. 4 Predicted 3D model structure of phytase of *Enterobacter aerogenes* (A0A0M3HCJ2) viewed by PyMol: **a** showing four distinct chains of the protein. **b** Tertiary structure showing prominent secondary elements and disulfides (red = helix, yellow = sheet, green = loop, green balls = disulfides)

Functional analysis and protein–protein interaction

From the functional analyses of the proteins, mainly two conserved motifs, i.e., His_Phos_1 and His_Phos_2 were found (Suppl. 4) both of which belong to the Histidine phosphatase superfamily (Suppl. 4). This superfamily is a vast functionally diverse group of proteins consisting of two clades sharing a finite sequence similarity (Rigden 2008). The multiple sequence alignment among all the 16 phytase protein sequences identified several fully conserved amino acid residues (marked as * in Fig. 5) among which "DG–DP–LG" sequence was found to be the highly conserved sequences (Fig. 5). Similarly, Niño-Gómez et al. (2017) while working with 3-Phytase A and 3-Phytase B from *Aspergillus niger* found that RHGXRXP-HD were the highly conserved sequences as revealed from the multiple sequence alignment study. In addition, the protein–protein interaction network produced through STRING server depicted that the query protein interacts with ten different proteins directly or
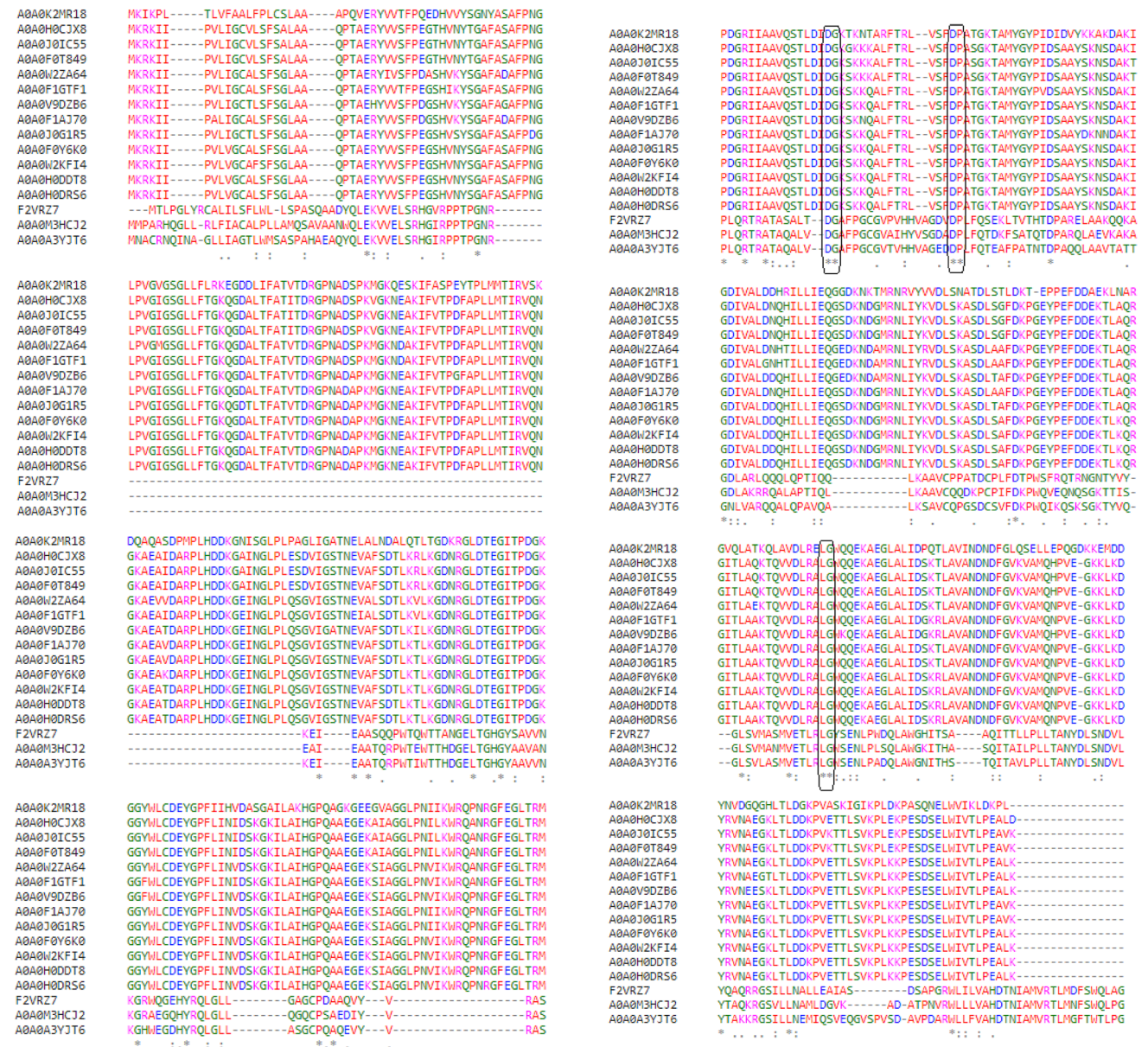

Fig. 5 Multiple sequence alignment of 16 phytase sequences of *Enterobacter* spp. showing highly conserved amino acid residues. An * (asterisk) indicates single, fully conserved residue, a : (colon) indicates conservation between groups of strongly similar properties and a. (period) indicates conservation between groups of weakly similar properties. Therefore, the hierarchy of conservation using these symbols is * (identical) > : (colon) >. (period). Highlighted area in the sequence alignment indicates highly conserved sequences

```
A0A0K2MR18    --------------------------------------------------
A0A0H0CJX8    --------------------------------------------------
A0A0J0IC55    --------------------------------------------------
A0A0F0T849    --------------------------------------------------
A0A0W2ZA64    --------------------------------------------------
A0A0F1GTF1    --------------------------------------------------
A0A0V9DZB6    --------------------------------------------------
A0A0F1AJ70    --------------------------------------------------
A0A0J0G1R5    --------------------------------------------------
A0A0F0Y6K0    --------------------------------------------------
A0A0W2KFI4    --------------------------------------------------
A0A0H0DDT8    --------------------------------------------------
A0A0H0DRS6    --------------------------------------------------
F2VRZ7        YSRGNIPPGSSLVLERWRDTRSGKRFLRIYFQAQSLDGIRQLQPLDDKHPLLRQEWHQPD
A0A0M3HCJ2    YSRGNIPPGSSLVIERWRNSKNGERYLRVYFQAQSLDDLRRLQMPDQQHPMLQQEWRQPD
A0A0A3YJT6    YTRGNIPPGSSLVFERWRDKQSGERWLHVYFQAQSLDDLRRLQSVDQAHPLLREEWRTPD


A0A0K2MR18    ------------------------------
A0A0H0CJX8    ------------------------------
A0A0J0IC55    ------------------------------
A0A0F0T849    ------------------------------
A0A0W2ZA64    ------------------------------
A0A0F1GTF1    ------------------------------
A0A0V9DZB6    ------------------------------
A0A0F1AJ70    ------------------------------
A0A0J0G1R5    ------------------------------
A0A0F0Y6K0    ------------------------------
A0A0W2KFI4    ------------------------------
A0A0H0DDT8    ------------------------------
A0A0H0DRS6    ------------------------------
F2VRZ7        CRVTDVGLLCPYQSTLTQLRKNLDNSAVLPVSVILP
A0A0M3HCJ2    CQQTSVGTLCPYQSALEALSKQIDKSAVPPVNMVL-
A0A0A3YJT6    CRVTEVGTLCPMASTLAKFGQQLDPTAITPISYAP-
```
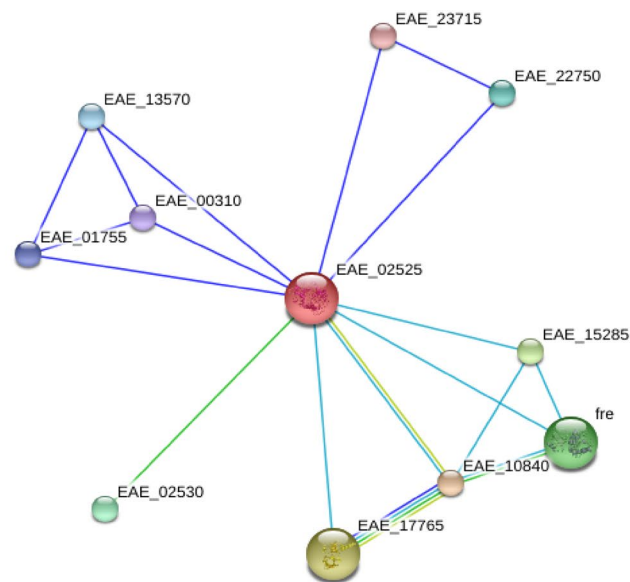
Fig. 5 (continued)

indirectly (Fig. 6). Occupying the central position, the query protein interacts with bifunctional riboflavin kinase/FMN adenylyltransferase, alpha subunit of riboflavin synthase, NAD(P)H-dependent FMN reductase, FMN reductase, major facilitator superfamily protein, spore coat U domain-containing protein, lipoprotein, TonB-dependent receptor, a hypothetical protein of 117 amino acid residues and putative universal stress family protein as shown in Fig. 6. STRING analysis is very important in terms of system-level understanding of cellular processes in a computer-based way. The interacting network can be used for filtering and assessing functional genomics data as well to provide an instinctive platform for annotating evolutionary properties of proteins in addition to structure–function aspects (Schwartz et al. 2008). Exploring this type of STRING generated predicted interaction networks can suggest important directions for future experimental research, e.g., prediction of the possible pathway of the protein of interest. Similar protein–protein interaction study was previously worked out by Goñi et al. (2008), Gao et al. (2012), Guney and Oliva (2012), Quan et al. (2014), Zhang et al. (2016), Pramanik et al. (2017a, b, 2018) etc.

Hence, this elaborative study will be very helpful in selecting commercial phytases in future derived from *E. aerogenes* as potential phytases in animal feed additive. In addition, direct application of *E. aerogenes*, in cultivated fields will help crops to grow better by soluble P uptake released due to phytase activity of the strain.

Conclusion

In silico characterization of *Enterobacter* phytases revealed that the 48 kDa proteins were tetrameric, thermostable and acidic in nature belonging to the histidine phosphatase superfamily. This types of thermostable, acidic phytases derived from *E. aerogenes* might be beneficial in terms of application in various industrial fields such as monogastric animals, poultry farms, etc., by solubilized phosphorus supplement in their food essential for their normal growth and development. Besides phytase containing *E. aerogenes* can be directly applied in the agricultural field to meet up the phosphorus deficiency in crop plants. Hence, this study will help researchers to understand the essential structure–function properties of *Enterobacter* phytases.

Fig. 6 Protein–protein interaction map including the predicted interacting proteins with phytase protein (A0A0M3HCJ2)

Compliance with ethical standards

Conflict of interest Authors declare no conflict of interest.

References

Afinah S, Yazid AM, Anis Shobirin MH et al (2010) Phytase: application in food industry. Int Food Res J 17:13–21

Apweiler R, Bairoch A, Wu CH et al (2004) UniProt: the Universal Protein knowledgebase. Nucleic Acids Res 32:D115–D119. https://doi.org/10.1093/nar/gkh131

Benkert P, Künzli M, Schwede T (2009) QMEAN server for protein model quality estimation. Nucleic Acids Res 37:510–514. https://doi.org/10.1093/nar/gkp322

Berlyne GM, Ben-Ari J, Nord E et al (1973) Bedouin osteomalacia caused by high phytic unleavened bread. Am J Clin Nutr 26:910–911

Berman HM, Westbrook J, Feng Z et al (2000) The Protein Data Bank. Nucl Acids Res 28:235–242. https://doi.org/10.1093/nar/28.1.235

Biasini M, Bienert S, Waterhouse A et al (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res 42:252–258. https://doi.org/10.1093/nar/gku340

Casey A, Walsh G (2004) Identification and characterization of a phytase of potential commercial interest. J Biotechnol 110(3):313–322. https://doi.org/10.1016/j.jbiotec.2004.03.001

Cheng C, Wong K-B, Lim BL (2007) The effect of disulfide bond on the conformational stability and catalytic activity of beta-propeller phytase. Protein Pept Lett 14:175–183. https://doi.org/10.2174/092986607779816069

Dahiya S (2016) Industrial application of phytases. Int J Appl Res 2:95–98

Ebrahimi S, Jamei R, Ghasemian A et al (2016) In Silico comparison of disulfide-bearing and disulfide-free phytases among microorganisms. J Pharm Nutr Sci 7:24–34

Elkhalil EAI, Männer K, Borriss R et al (2007) In vitro and in vivo characteristics of bacterial phytases and their efficacy in broiler chickens. Br Poult Sci 48:64–70. https://doi.org/10.1080/00071660601148195

El-Toukhy NMK, Youssef AS, Mikhail MGM (2013) Isolation, purification and characterization of phytase from Bacillus subtilis MJA. African J Biotechnol 12:2957–2967. https://doi.org/10.5897/AJB2013.12304

Escobin-Mopera L, Ohtani M, Sekiguchi S et al (2012) Purification and characterization of phytase from Klebsiella pneumoniae 9-3B. J Biosci Bioeng 113:562–567. https://doi.org/10.1016/j.jbiosc.2011.12.010

Gao P, Wang Q-P, Chen L et al (2012) Prediction of human genes' regulatory functions based on protein–protein interaction network. Protein Pept Lett 19:910–916

Gasteiger E, Hoogland C, Gattiker A et al (2005) Protein identification and analysis tools on the ExPASy Server. Proteom Protoc Handb 571–607. https://doi.org/10.1385/1592598900

George TS, Simpson RJ, Gregory PJ et al (2007) Differential interaction of Aspergillus niger and Peniophora lycii phytases with soil particles affects the hydrolysis of inositol phosphates. Soil Biol Biochem 39:793–803

Geourjon C, Deléage G (1995) SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput Appl Biosci 11:681–684. https://doi.org/10.1093/bioinformatics/11.6.681

Goñi J, Esteban FJ, de Mendizábal NV et al (2008) A computational analysis of protein–protein interaction networks in neurodegenerative diseases. BMC Syst Biol 2:52. https://doi.org/10.1186/1752-0509-2-52

Gontia-Mishra I, Tiwari S (2013) Molecular characterization and comparative phylogenetic analysis of phytases from fungi with their prospective applications. Food Technol Biotechnol 51:313–326

Gontia-Mishra I, Singh VK, Tripathi N et al (2014) Computational identification, homology modelling and docking analysis of phytase protein from Fusarium oxysporum. Biologia 69:1283–1294

Greiner R, Konietzny U (2016) Phytase for food application. Food Technol Biotechnol 44:125–140

Guney E, Oliva B (2012) Exploiting protein–protein interaction networks for genome-wide disease-gene prioritization. PLoS One. https://doi.org/10.1371/journal.pone.0043557

Gupta RK, Gangoliya SS, Singh NK (2015) Reduction of phytic acid and enhancement of bioavailable micronutrients in food grains. J Food Sci Technol 52:676–684. https://doi.org/10.1007/s13197-013-0978-y

He Z, Honeycutt CW (2001) Enzymatic characterization of organic phosphorus in animal manure. J Environ Qual 30:1685–1692

He Z, Zhong J, Cheng HN (2013) Conformational change of metal phytates: solid state 1D ^{13}C and 2D ^{1}H–^{13}C NMR spectroscopic investigations. J Food Agric Environ 11:965–970

Kalsi HK, Singh R, Dhaliwal HS et al (2016) Phytases from Enterobacter and Serratia species with desirable characteristics for food and feed applications. 3 Biotech 6:64. https://doi.org/10.1007/s13205-016-0378-x

Kerovuo J, Lauraeus M, Nurminen P et al (1998) Isolation, characterization, molecular gene cloning, and sequencing of a novel phytase from Bacillus subtilis. Appl Environ Mcrobiol 64:2079–2085

Konietzny U, Greiner R (2004) Bacterial phytase: potential application, in vivo function and regulation of its synthesis. Braz J Microbiol 35:11–18. https://doi.org/10.1590/S1517-83822004000100002

Konietzny U, Greiner R, Jany K-D (1994) Purification and characterization of a phytase from spelt. J Food Biochem 18:165–183. https://doi.org/10.1111/j.1745-4514.1994.tb00495.x

Kumar V, Agrawal S (2014) An insight into protein sequences of PTP-like cysteine phytases. Nusant Biosci 6:97–101

Kumar S, Tsai C-J, Nussinov R (1999) Factors enhancing protein thermostability. Protein Eng 13:179–191. https://doi.org/10.1093/protein/13.3.179

Kumar V, Sinha AK, Makkar HPS et al (2010) Dietary roles of phytate and phytase in human nutrition: a review. Food Chem 120:945–959. https://doi.org/10.1016/j.foodchem.2009.11.052

Kumar V, Singh G, Verma AK et al (2012) In silico characterization of histidine acid phytase sequences. Enzyme Res 2012:845465. https://doi.org/10.1155/2012/845465

Kumar V, Singh G, Sangwan P et al (2014) Cloning, sequencing and in silico analysis of β-propeller phytase Bacillus licheniformis strain PB-13. Biotechnol Res Int 2014:841353 **(1–11)**

Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol msw054. https://doi.org/10.1093/molbev/msw054

Lei XG, Porres JM, Mullaney EJ et al (2007) Phytase: source, structure and application. In: Polaina J, MacCabe AP (eds) Industrial enzymes. Springer, Dordrecht, pp 505–529

Lei XG, Weaver JD, Mullaney EJ et al (2013) Phytase, a new life for an "old" enzyme. Annu Rev Anim Biosci 1:283–309

Mathew A, Verma A, Gaur S (2014) An in-silico insight into the characteristics of β-propeller phytase. Interdiscip Sci Comput Life Sci 6:133–139. https://doi.org/10.1007/s12539-013-0010-2

Meiler J, Baker D (2003) Rapid protein fold determination using unassigned NMR data. Proc Natl Acad Sci USA 100:15404–15409. https://doi.org/10.1073/pnas.2434121100

Mullaney EJ, Daly CB, Ullah AH (2000) Advances in phytase research. Adv Appl Microbiol 47:157–199. https://doi.org/10.1016/S0065-2164(00)47004-8

Nakashima BA, McAllister TA, Sharma R et al (2007) Diversity of phytases in the rumen. Microb Ecol 53:82–88. https://doi.org/10.1007/s00248-006-9147-4

Niño-Gómez DC, Rivera-Hoyos CM, Morales-Álvarez ED et al (2017) "In silico" characterization of 3-phytase A and 3-phytase B from Aspergillus niger. Enzyme Res 2017 9746191:1–23. https://doi.org/10.1155/2017/9746191

Pramanik K, Soren T, Mitra S et al (2017a) In silico structural and functional analysis of Mesorhizobium ACC deaminase. Comp Biol Chem 68:12–21. https://doi.org/10.1016/j.compbiolchem.2017.02.005

Pramanik K, Ghosh PK, Ray S et al (2017b) An in silico structural, functional and phylogenetic analysis with three dimensional protein modeling of alkaline phosphatase enzyme of Pseudomonas aeruginosa. J Genet Eng Biotechnol 15:527–537. https://doi.org/10.1016/j.jgeb.2017.05.003

Pramanik K, Pal P, Soren T et al (2018) In silico structural, functional and phylogenetic analysis of Klebsiella phytases. Plant Biochem Biotechnol. https://doi.org/10.1007/s13562-018-0445-y **(in press)**

Quan C-S, Fan S-D, Zhang L-H et al (2002) Purification and properties of a phytase from Candida krusei WZ-001. J Biosci Bioeng 94:419–425

Quan Z, Quan Y, Wei B et al (2014) Protein–protein interaction network and mechanism analysis in ischemic stroke. Mol Med Rep 11:29–36. https://doi.org/10.3892/mmr.2014.2696

Ragon M, Hoh F, Aumelas A et al (2009) Structure of *Debaryomyces castellii* CBS 2923 phytase. Acta Crystallogr Sect F Struct Biol Cryst Commun 65:321–326. https://doi.org/10.1107/S1744309109008653

Richardson AE, Hadobas PA (1997) Soil isolates of *Pseudomonas* spp. that utilize inositol phosphates. Can J Microbiol 43:509–516. https://doi.org/10.1139/m97-073

Rigden DJ (2008) The histidine phosphatase superfamily: structure and function. Biochem J 409:333–348. https://doi.org/10.1042/BJ20071097

Rodriguez E, Wood ZA, Karplus PA et al (2000) Site-directed mutagenesis improves catalytic efficiency and thermostability of *Escherichia coli* pH 2.5 acid phosphatase/phytase expressed in *Pichia pastoris*. Arch Biochem Biophys 382:105–112

Schwartz AS, Yu J, Gardenour KR et al (2008) Cost-effective strategies for completing the interactome. Nat Methods 6:55–61

Shivange AV, Serwe A, Dennig A et al (2012) Directed evolution of a highly active *Yersinia mollaretii* phytase. Appl Microbiol Biotechnol 95:405–418. https://doi.org/10.1007/s00253-011-3756-7

Singh P, Kumar V, Agrawal S (2014) Evaluation of phytase producing bacteria for their plant growth promoting activities. Int J Microbiol 2014:426483. https://doi.org/10.1155/2014/426483

Tazisong IA, Senwo ZN, Taylor RW et al (2008) Hydrolysis of organic phosphates by commercially available phytases: biocatalytic potentials and effects of ions on their enzymatic activities. J Food Agri Environ 6:500–505

Tran TT, Hashim SO, Gaber Y et al (2011) Thermostable alkaline phytase from *Bacillus* sp. MD2: effect of divalent metals on activity and stability. J Inorg Biochem 105:1000–1007. https://doi.org/10.1016/j.jinorgbio.2011.04.005

Trivedi MV, Laurence JS, Siahaan TJ (2009) The role of thiols and disulfides in protein chemical and physical stability. Curr Protein Pept Sci 10:614–625

Tseng YH, Fang TJ, Tseng SM (2000) Isolation and characterization of a novel phytase from *Penicillium simplicissimum*. Folia Microbiol (Praha) 45:121–127

Verma A, Singh VK, Gaur S (2016) Computational based functional analysis of *Bacillus* phytases. Comput Biol Chem 60:53–58. https://doi.org/10.1016/j.compbiolchem.2015.11.001

Wise A (1982) Blood lead levels after chronic feeding to mice of lead acetate with calcium phylate in the diet. Bull Environ Contam Toxicol 29:550–553

Yadav PK, Singh G, Gautam B et al (2013) Molecular modeling, dynamics studies and virtual screening of Fructose 1, 6-biphosphate aldolase-II in community acquired-methicillin resistant *Staphylococcus aureus* (CA-MRSA). Bioinformation 9:158–164

Yanke LJ, Bae HD, Selinger LB et al (1998) Phytase activity of anaerobic ruminal bacteria. Microbiology 144:1565–1573

Yao MZ, Zhang YH, Lu WL et al (2012) Phytases: crystal structures, protein engineering and potential biotechnological applications. J Appl Microbiol 112:1–14

Zając A, Dymińska L, Lorenc J et al (2018) Syntheses, spectroscopic properties and molecular structure of silver phytate complexes-IR, UV-VIS studies and DFT calculations. J Mol Struct 1156:483–491

Zhang W, Mullaney EJ, Xin GL (2007) Adopting selected hydrogen bonding and ionic interactions from *Aspergillus fumigatus* phytase structure improves the thermostability of *Aspergillus niger* PhyA phytase. Appl Environ Microbiol 73:3069–3076. https://doi.org/10.1128/AEM.02970-06

Zhang R, Yang P, Huang H et al (2011) Two types of phytases (Histidine acid phytase and β-propeller phytase) in *Serratia* sp. TN49 from the gut of *Batocera horsfieldi* (Coleoptera) Larvae. Curr Microbiol 63:408–415. https://doi.org/10.1007/s00284-011-9995-0

Zhang Q, Zhang P, Cai Y (2016) The use of protein–protein interactions for the analysis of the associations between PM2.5 and some diseases. Biomed Res Int 2016:4895476. https://doi.org/10.1155/2016/4895476 **(1–7)**