

Full Length Research Paper

***In silico* characterization of beta-galactosidase using computational tools**

Gangadhar. C. Gouripur¹, Rohit. B. Kaliwal² and Basappa. B. Kaliwal^{1,3*}

¹Department of Studies and Research in Biotechnology and Microbiology, Karnatak University, Dharwad - 580 003, India.

²Department of MCA, Vishwvaraya Technological University, P.G Centre Regional Office, Gulbarga- 585106, India.

³Department of Studies in Biotechnology and Microbiology, Karnatak University, Dharwad, Karnataka, India.

Received 21 February, 2015; Accepted 25 August, 2015

β -galactosidase (EC.3.2.1.23) is an important enzyme, mainly used in the preparation of lactose hydrolyzed milk suitable for people with lactose intolerance. It is essential to understand the structural and functional aspects of various β -galactosidase produced from different sources. The present work deals with the use of bioinformatics to describe the physiochemical, functional and structural properties of β -galactosidase enzymes on *Bacillus sp.* selected from the gene bank of NCBI. The grand average hydropathy (GRAVY) and low range of AFY63015.1 value indicates the possibility of better interaction with water and instability index were computed to characterise YP_004205251.1, ZP_10511829.1, BAL72724.1, AFY63015.1, NP_242888.1 stating that they are stable and disulfide bridges, CYS_REC recognizes the presence of 38 cysteine residues in β -galactosidase sequences and predicted most probable SS bond patterns of pairs in YP_004981461 and 1AFY63015.1. The self-optimized prediction method (SOPM) was used to predict the secondary structure. The SOPM results indicated the presence of alpha helix is more dominated in sequences AFY63015.1 and YP_004981461.1. Overall this represents *in silico* analysis of sequence, structural and functional information of β -galactosidase of *Bacillus* species.

Key words: *Bacillus sp.*, β -galactosidase, *in silico* analysis, physico-chemical characterization, proteomics tools.

INTRODUCTION

Lactase also known as β -galactosidase (E.C 3.2.1.23) is an enzyme that hydrolyzes lactose (abundant disaccharide found in milk) into glucose and galactose and has a potential importance in the dairy industry (Voget et al., 1994; Domingues et al., 2005; Gouripur et al., 2013). β -Galactosidase has tremendous potential in research and application in various fields like food,

pharmaceutical, bioremediation, biosensor, diagnosis and treatment of disorders. It is used in the preparation of lactose hydrolyzed milk suitable for lactose intolerant people. (Kaur et al., 2006; Patil et al., 2011).

The commercial enzymes used for lactose hydrolysis are β -galactosidase obtained from various origins (Jurado et al., 2002). Sources of β -galactosidase include plants,

*Corresponding author. E-mail: b_kaliwal@yahoo.com. Tel: 08192208444. Fax: 08192263087.

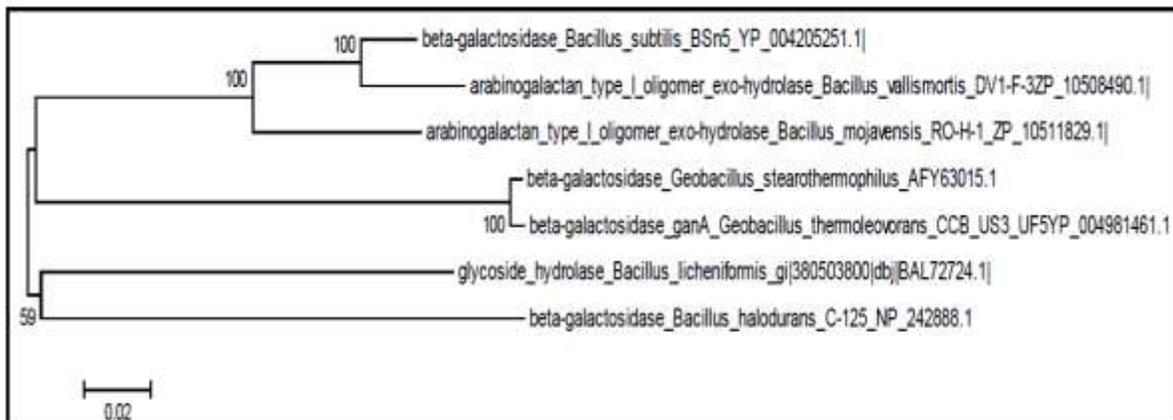


Figure 1. Phylogenetic tree β -galactosidases for *Bacillus* sequence. Distance options according to the kimura 2-parameter model and clustering with the neighbor-joining were performed using the software package mega (Molecular Evolutionary Genetics Analyses) ver. 5.1. Bootstrap percentage (50%) based on 1000 replications are given at branch points. Bar 0.02.

animals, bacteria, yeasts (intracellular enzyme), fungi and moulds(extracellular enzyme). Among these sources, bacteria are preferred because of their simplicity in fermentation with optimum activity and good stability (Sani et al., 1999). Some strains have been proved to have probiotic activity enhancing the digestion of lactose (Vinderola and Reinheimer, 2003).

The lactose, a disaccharide specifically known as 4-O- β -D-galactopyranosyl-D-glucose is found exclusively in milk. The nutritional value of lactose is limited due to the fact that a large portion, approximately 50% of the world's inhabitants lacks this enzyme and cannot utilize lactose, therefore, developing lactose maldigestion or intolerance (Vasiljevic and Jelen, 2002). This, however, creates a potential market for the application of β -galactosidase. The current share of food enzymes is 37% of total enzyme sales corresponding to 720 million dollars in the year 2004. This value has increased to 863 million dollars by the year 2009, increasing the demand for the discovery of new species, producing enzymes such as β -galactosidase with novel characteristics, which will be of great value to the enzyme industry for different applications (Cortes et al., 2005). β -galactosidase differ in their physicochemical properties, structures, specific activities, thermostability and yields, thus providing a great deal of choices in their potential usages. Therefore, the present investigation deals with the *in silico* analysis and characterization of β -galactosidase from *Bacillus* species.

MATERIALS AND METHODS

Sequence retrieval

The β -galactosidase sequences were obtained from NCBI and retrieved from Swiss-Prot, a public domain protein database

(Boeckmann et al., 2003). A total number of seven sequences were retrieved from Swiss-Prot by random selection. Protein sequences of β -galactosidase from *Bacillus* sp. were retrieved in FASTA format and used for further analysis. The phylogenetic tree was constructed using the aligned sequences by the neighbor-joining (NJ) method using kimura 2-parameter distances in the MEGA beta 5.1 software (Tamura et al., 2011). The method of analyses and statistics used for phylogeny test was by bootstrap, with its bootstrap replication number 1000 as shown in the Figure 1.

Physico-chemical characterization

The physicochemical properties were calculated from the primary structure of β -galactosidase. Molecular weight, theoretical isoelectric point (pI), total number of positive and negative residues, extinction coefficient (Gill and Von Hippel, 1989), instability index (Guruprasad et al., 1990), aliphatic index (Ikai et al., 1980) and grand average hydropathy (GRAVY) (Kyte and Doolittle, 1982) were computed using the ExPASy's Prot-Param (Gasteiger et al., 2005) (<http://us.expasy.org/tools/protparam.html>) prediction server.

Secondary structure prediction

Self-optimized prediction method (SOPM) was employed for calculating the secondary structural features of the selected target protein sequences considered for this study (Table 4). The identification of transmembrane regions of a protein was identified by server SOSUI (Classification and Secondary Structure Prediction of Membrane Proteins). It represents the transmembrane regions identified for β -galactosidase proteins (Hirokawa et al., 1998).

Functional characterization

The predicted transmembrane helices were visualized and analyzed using the self-optimized prediction method (SOPM) (Geourjon and Deleage, 1995). The SOPM is used to improve the success rate in the prediction of the secondary structure of proteins. SOPM parameters such as Window width-17, Similarity threshold-7,

Table 1. β -galactosidase sequences retrieved from Swiss-Prot database for *Bacillus* Species.

Accession number	Length	Description	Organism's
YP_004205251.1	686	Beta-galactosidase	<i>Bacillus subtilis</i> BSn5
ZP_10508490.1	686	Arabinogalactan type I oligomer exo-hydrolase	<i>Bacillus vallismortis</i> DV1-F-3J
ZP_10511829.1	686	Arabinogalactan type I oligomer exo-hydrolase	<i>Bacillus mojavensis</i> RO-H-1
BAL72724.1	684	Glycoside hydrolase	<i>Bacillus licheniformis</i>
AFY63015.1	686	Beta-galactosidase	<i>Geobacillus stearothermophilus</i>
YP_004981461.1	686	Beta-galactosidase ganA	<i>Geobacillus thermoleovorans</i> CCB_US3_UF5
NP_242888.1	689	Beta-galactosidase	<i>Bacillus halodurans</i> C-125

Number of states-4 were computed, and the computational methods were applied for determining disulphide bonds. Disulphide bonds are important in determining the functional linkages, so, S-S bonds were analyzed using the primary protein sequence data with the help of CYS_REC. Motifs in the considered sequences were scanned using Motif Search. SOSUI server was used to predict the transmembrane tendency of the proteins. Kyte and Doolittle mean hydrophobicity score was calculated, and the plot was obtained using Kyte and Doolittle method, keeping a window size of 9 (Hirokawa et al., 1998).

RESULTS AND DISCUSSION

In silico analysis of β -galactosidase sequences from *Bacillus* sp. BLAST searches of the NCBI database using β -galactosidase sequences as queries showed that 7 other β -galactosidases share 75-100% identity with Phylogenetic tree β -galactosidases for *Bacillus* sequence. Distance options according to the kimura 2-parameter model and clustering with the neighbor-joining were performed using the software package mega (Molecular Evolutionary Genetics Analyses) ver. 5.1. Bootstrap percentage (50%) based on 1000 replications are given at branch points (Bar 0.02) (Figure 1) (Tamura et al., 2011). The sequence differences among β -galactosidases suggest differences in their enzymatic properties and biological functions.

The present results (Table 1) showed that the 7 β -galactosidase sequences of *Bacillus* sp. were retrieved from SWISS-PROT (Bairoch and Apweiler, 2000). The primary structure was analysed, and different parameters computed using ExPasy ProtParam tool was tabulated (Tables 2 and 3). The results suggest that β -galactosidase sequences are mostly hydrophobic, and their hydrophobic nature is due to the presence of high content of non-polar residues (Sivakumar et al., 2007). The average molecular weight of β -galactosidase is 79709.0 Dalton (Table 3). The isoelectric point (pI) is the pH at which the charge covers the surface of the protein, but the net charge of the protein is zero. At pI, proteins are stable and compact. The computed pI value of all β -

galactosidase sequences had pI < 7 indicating that β -galactosidase is acidic in nature. Amino acid composition determines the fundamental properties of the enzyme while the amino acid composition of xylanase sequences the isoelectric point (pI) values of all protein sequences are acidic in nature (Neelima et al., 2009). The computed isoelectric point (pI) will be useful for developing buffer systems for purification of the recombinant proteins by the isoelectric focusing method (Gasteiger et al., 2005).

Although ExPasy's ProtParam computes the extinction coefficient (EC) for a range of (276, 278, 279, 280 and 282 nm) wavelength, 280 nm is favoured because β -galactosidase absorb this wavelength strongly thus interference from other substances in proteins can be minimised. The EC of β -galactosidase at 280 nm ranges from 150940 to 162525 M⁻¹ cm⁻¹ with respect to the concentration of Cys, Trp and Tyr (Table 3). The high EC value of YP_004981461.1 indicates the presence of high concentration of Cys, Trp and Tyr. The computed EC values will help in the quantitative study of protein-protein and protein-ligand interactions (Gill and Von Hippel, 1989). A protein whose instability index is smaller than 40 are predicted as stable, and a value above 40 predicts that the protein is unstable. The instability index existing at ExPasy's ProtParam classifies Stability and instability (Sivakumar, 2010). β -galactosidase for the following sequence were YP_004205251.1, ZP_10511829.1, BAL72724.1, AFY63015.1, NP_242888.1 which indicates that they are stable.

On the other hand, ZP_10508490.1 and YP_004981461.1 are slightly unstable as their instability index is <40 (Table 3). This result was almost similar to the instability index, which gives clues about the stability of a protein *in vitro* can be calculated. All the considered sequences were classified as stable with a value ranging from 13.57 to 37.23 as a value > 40 indicates an unstable protein (Guruprasad et al., 1990).

The aliphatic index (AI) which is defined as the relative volume of a protein occupied by aliphatic side chains is regarded as a positive factor for the increase of thermal

Table 2. Amino acid compositions (in %) of β -galactosidase for *Bacillus* species.

Amino acid composition	YP_004205251.1 (%)	ZP_10508490.1 (%)	ZP_1051182.1 (%)	BAL72724.1 (%)	AFY63015.1 (%)	YP_004981461.1 (%)	NP_242888.1 (%)
Ala	5.8	6.1	6.1	6.3	6.3	6.1	5.5
Arg	4.7	5	4.5	4.8	6	5.8	4.2
Asn	3.9	4.2	4.5	4.5	4.1	4.1	4.8
Asp	7.3	6.7	6.9	6.6	6.3	6.3	6.8
Cys	0.7	0.7	0.7	0.7	1	1	0.6
Gln	4.5	4.2	4.5	4.1	3.9	4.1	3.8
Glu	5.7	6.4	6	6.6	7.1	7.3	6.4
Gly	5.4	5.4	5.5	5.6	6.3	6.3	5.4
His	4.5	4.5	4.2	3.9	3.9	3.9	4.1
Ile	5.4	5.7	5.8	5	4.8	4.8	4.9
Leu	8.6	8.6	8.9	8.2	8.6	8.6	8.7
Lys	5.5	5.2	5.8	5.8	5.1	5.1	6
Met	2.3	2.5	2.3	2.8	2.2	2.2	2.3
Phe	4.2	4.1	4.5	4.8	4.5	4.7	4.5
Pro	4.5	4.2	4.2	4.7	4.1	4.1	4.4
Ser	6.4	6.6	6.3	6	4.8	5	6.8
Thr	6.3	6	6	6.3	5.8	5.8	6.8
Trp	2.8	2.8	2.8	2.9	2.9	2.9	2.9
Tyr	4.7	4.8	4.5	4.7	5.1	5.1	4.8
Val	6.7	6.3	5.8	5.7	7.1	6.9	6.4
Pyl	0	0	0	0	0	0	0
Sec	0	0	0	0	0	0	0

stability of globular proteins. Aliphatic index for the β -galactosidase sequences (Table 3) ranged from 74.14 to 80.45. The very high aliphatic index of all β -galactosidase sequences indicates that these β -galactosidases may be stable for a wide temperature range. This result was almost similar with an aliphatic index of antifreeze proteins that ranged from 57.89 to 125.23 among sequences of different varieties (Sivakumar et al., 2007). The GRAVY value for a peptide or protein was calculated as the sum of hydropathy values

of all the amino acids, divided by the number of residues in the sequence. GRAVY indices of β -galactosidase ranged from -0.517 to -0.452. This low range of AFY63015.1 value indicates the possibility of better interaction with water. Similarly, compared GRAVY value of tyrosinases ranged from -0.660 to -0.191. The very low GRAVY index of tyrosinases infers that these tyrosinases could result in a better interaction with water (Sivakumar et al., 2007). The secondary structure indicates whether a given amino acid

lies in a helix, strand or coil. The secondary structure features as predicted using self-optimized prediction method and is represented in Table 4. The results reveal that alpha helix dominated among secondary structure elements followed by, extended strand, beta turns and random coils while an extended strand outnumbered random coils in between secondary structure elements. The secondary structural elements consequence prediction result for protein's alpha helix is more dominated in

Table 3. Parameters computed using ExPASy's ProtParam tool of β -galactosidase for *Bacillus* species.

Gene ID	YP_004205251.1	ZP_10508490.1	ZP_10511829.1	BAL72724.1	AFY63015.1	YP_004981461.1	NP_242888.1
Hydrophobicity	-0.458163	-0.462682	-0.452332	-0.51652	-0.456413	-0.472011	-0.490276
A.A.	686	686	686	684	686	686	689
Mol.Wt	79190.4	79312.5	79203.5	79227.4	79643.1	79709	79581.6
pI	5.89	5.86	5.94	5.89	5.95	5.85	5.73
- R	89	90	88	90	92	93	91
+ R	70	70	71	73	76	75	70
EC	152430	153920	150940	157930	162525	162525	159420
II	37.78	40.15	37.36	37.56	39.25	40.76	36.9
AI	79.85	80.01	80.45	74.14	79.29	78.29	77.24
GRAVY	-0.458	-0.463	-0.452	-0.517	-0.456	-0.472	-0.49
TNA	11017	11026	11029	10997	11079	11078	11059
Formula	C ₃₅₇₆ H ₅₄₁₉ N ₉₅₉ O ₁₀₄₂ S ₂₁	C ₃₅₇₆ H ₅₄₂₁ N ₉₆₃ O ₁₀₄₄ S ₂₂	C ₃₅₈₀ H ₅₄₃₀ N ₉₅₈ O ₁₀₄₀ S ₂₁	C ₃₅₇₉ H ₅₃₉₈ N ₉₅₆ O ₁₀₄₀ S ₂₄	C ₃₆₀₃ H ₅₄₄₇ N ₉₇₃ O ₁₀₃₄ S ₂₂	C ₃₆₀₆ H ₅₄₄₁ N ₉₇₁ O ₁₀₃₈ S ₂₂	C ₃₅₉₇ H ₅₄₃₂ N ₉₅₂ O ₁₀₅₈ S ₂₀

A.A, Amino Acid; Mol.Wt, molecular weight; pI, isoelectric point; -R, number of negative charged residues (asp+Glu); +R, number of positive charged residues (Arg+Lys); EC, extinction coefficient at 280 nm; II, Instability Index; AI, aliphatic Index; GRAVY, grand average hydropathicity; TNA, total number of atoms.

Table 4. Secondary structure elements of Beta-galactosidase sequences for *Bacillus* species.

Secondary structure	YP_004205251.1 (%)	ZP_10508490.1 (%)	ZP_10511829.1 (%)	BAL72724.1 (%)	AFY63015.1 (%)	YP_004981461.1 (%)	NP_242888.1 (%)
Alpha helix	38.05	39.07	38.05	37.87	39.21	39.21	37.01
310 helix	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Pi helix	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Beta bridge	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Extended	20.70	21.28	21.28	18.42	20.55	19.83	19.74
Beta turn	11.52	10.50	10.06	9.94	10.35	10.35	9.87
Bend region	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Random coil	29.74	29.15	30.61	33.77	29.88	30.61	33.38
Ambiguous state	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Other states	0.00	0.00	0.00	0.00	0.00	0.00	0.00

sequences AFY63015.1 and YP_004981461.1.

This result was almost similar with the cyclooxygenases that has rich alanine content.

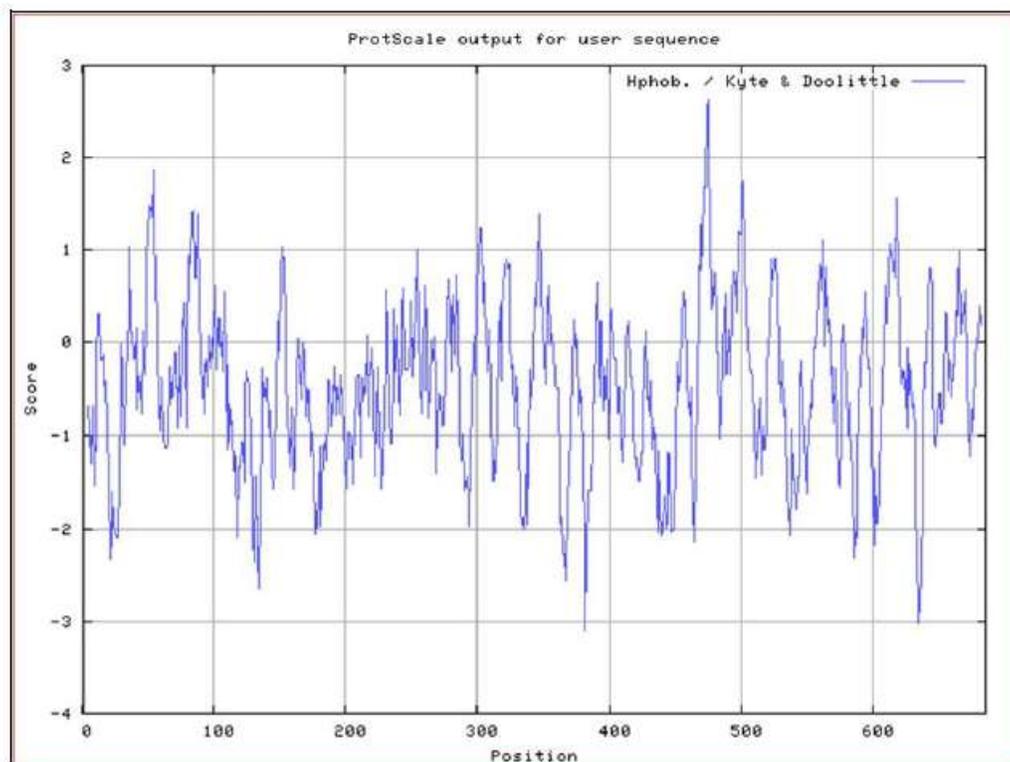


Figure 2. Kyte and Doolittle hydrophathy plot for β -galactosidase for *Bacillus subtilis* BSn5.

Although cyclooxygenases have mixed secondary structure, that is α -helices, β -strands and coils, the α -helices are dominant features in the protein structure. The very high coil structural content of cyclooxygenases is due to the rich content of more flexible glycine and hydrophobic proline amino acids. Proline has a special property of creating kinks in polypeptide chains and disrupting ordered secondary structure (Combet et al., 2000).

A set of conserved amino acid residues located in the district that provides clues to the functions is termed as a motif. Motifs predicted using Motif search found that all protein ID contained EF-hand calcium-binding domain motif. The average length of the motif predicted was 20 starts and 32 ends. Motifs could be predicted for all protein sequences. Besides all other physicochemical characterization, functional characterization of β -galactosidase proteins were also performed, including transmembrane (TM) region identification and prediction of disulphide bonding pairs. The SOSUI server performed the identification of transmembrane helices with their corresponding length and differentiates membrane proteins from stable proteins. SOSUI distinguishes between the membrane and soluble proteins; it further predicts the transmembrane helices from amino acid sequences quickly with high precision. β -Galactosidase protein by SOSUI server and all others β -galactosidase

were predicted to be soluble proteins (Hirokawa, et al., 1998).

CYS_REC identifies the positions of cysteines, the total number of cysteines present and computes the most probable S-S bond pattern of the pairs in a protein sequence. Possible disulphide bond pairing and patterns with probability were predicted by CYS_REC from primary sequence and S-S bonds were identified. The tool CYS_REC recognizes the presence of 38 cysteine residues in β -galactosidase sequences and predicted that the most probable SS bond patterns of pairs in YP_004981461 and 1AFY63015.1 were cysteine residues. Similarly, CYS-REC tool performing the prediction of S-S bonding states of cysteines and their location in proteins were also reported (Sivakumar et al., 2007).

The presence of cysteine residue disulphide bonds in corresponding β -galactosidase proteins is shown (Table 5). The predicted transmembrane region was found to be rich in hydrophobic amino acids in which many points lie above the zero baseline, and a clear peak was observed in a plot that indicates the plausible transmembrane region (Figures 2 to 8). The hydrophobicity score of BAL72724 β -galactosidase was -2.778 minimum and 2.778 maximum; a range (Table 6) as disulphide bridges play a significant role in determining the thermal stability of enzymes (Hirokawa et al., 1998).

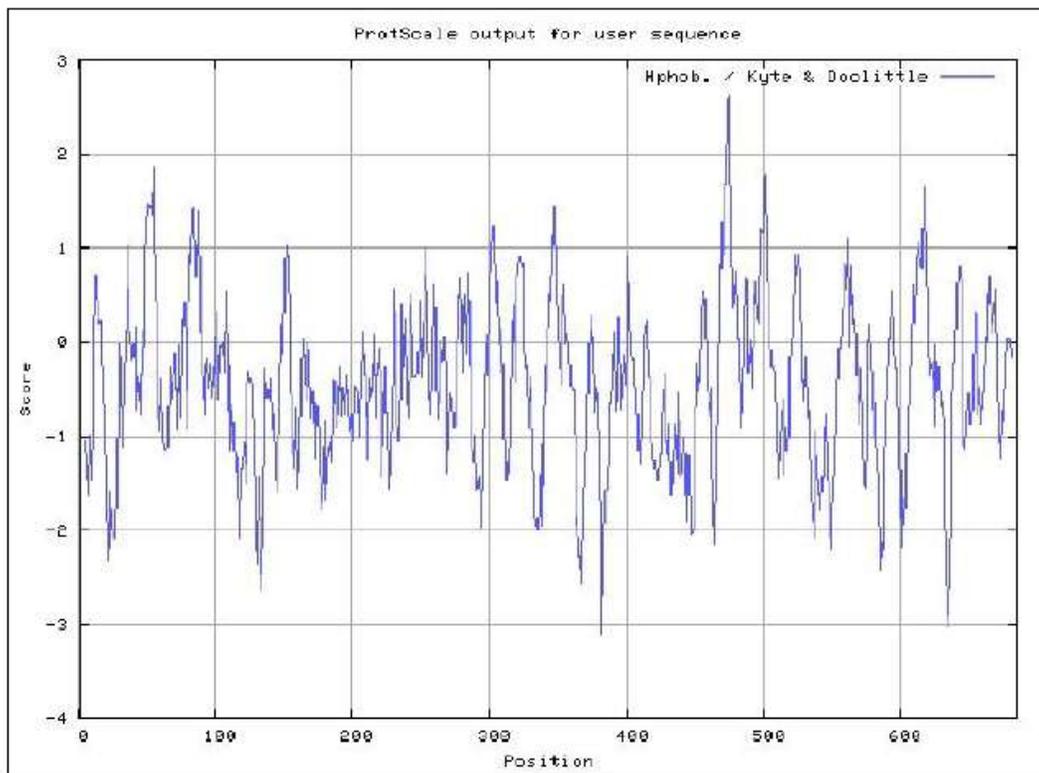


Figure 3. Kyte and Doolittle hydropathy plot for β -galactosidase for *Bacillus vallismortis* DV1-F-3.

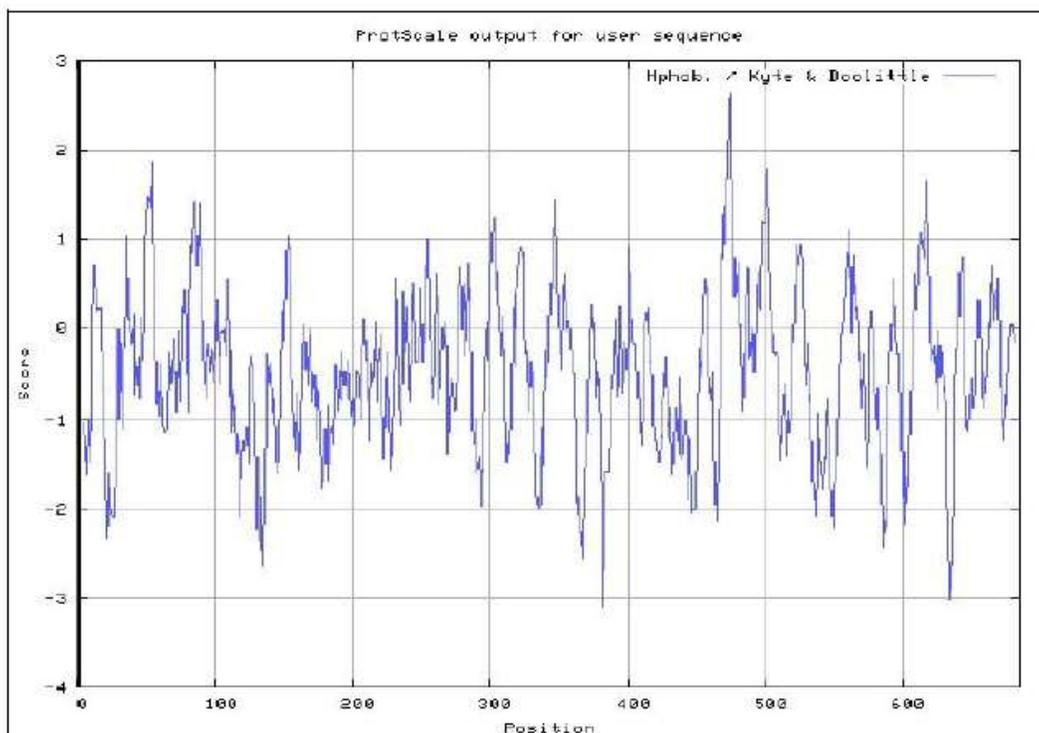


Figure 4. Kyte and Doolittle hydropathy plot for β -galactosidase for *Bacillus mojavensis* RO-H-1.

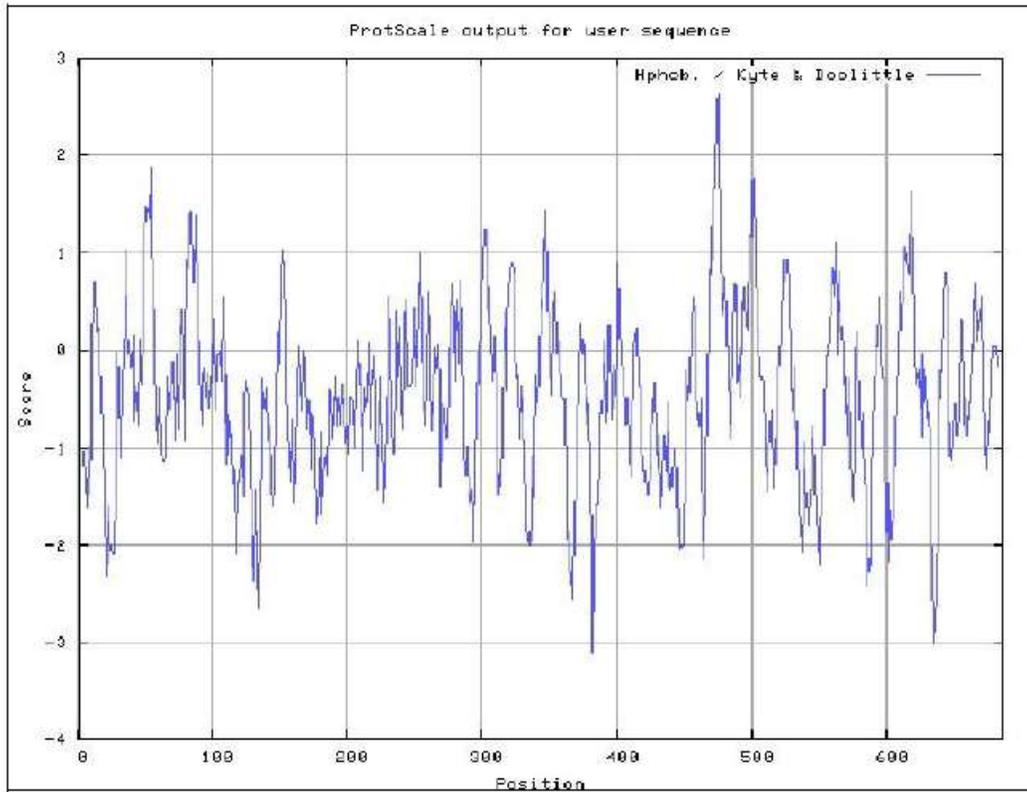


Figure 5. Kyte and Doolittle hydropathy plot for β -galactosidase for *Bacillus licheniformis*.

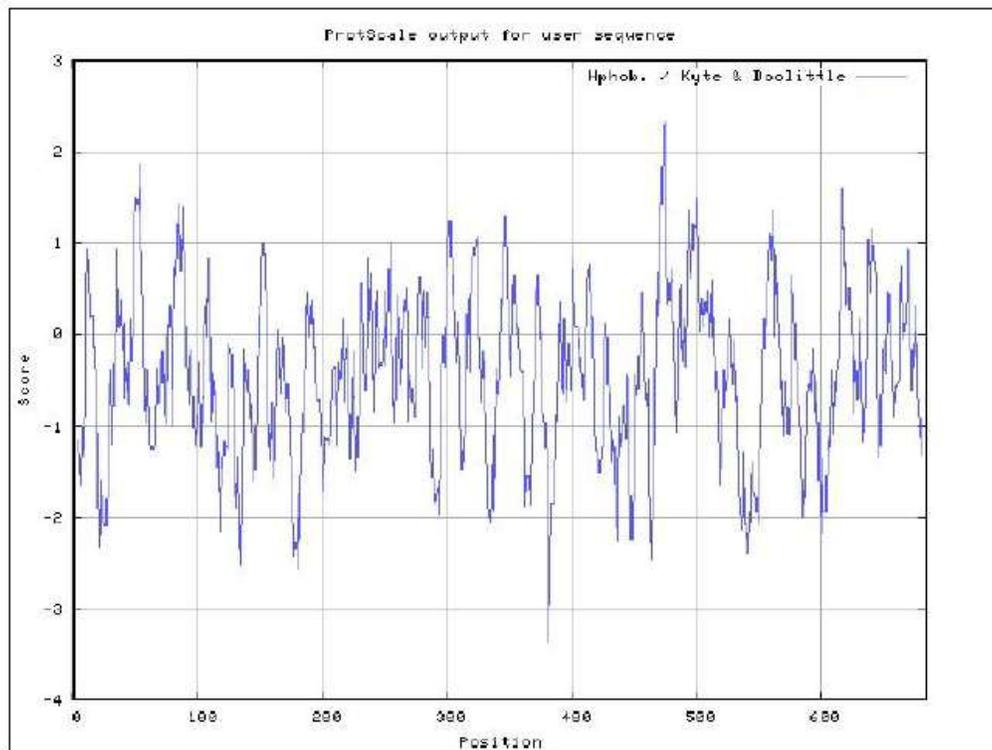


Figure 6. Kyte and Doolittle hydropathy plot for β -galactosidase for *Geobacillus stearothermophilus*.

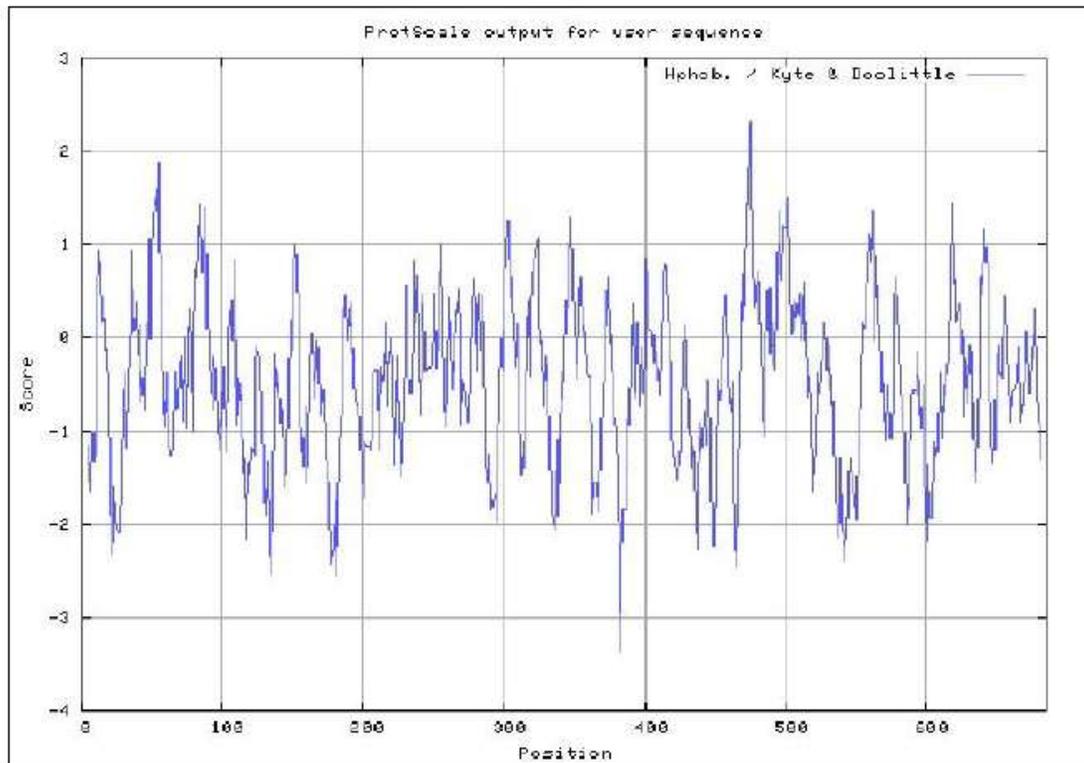


Figure 7. Kyte and Doolittle hydropathy plot for β -galactosidase for *Geobacillus thermoleovorans.CCB_US3_UF5*

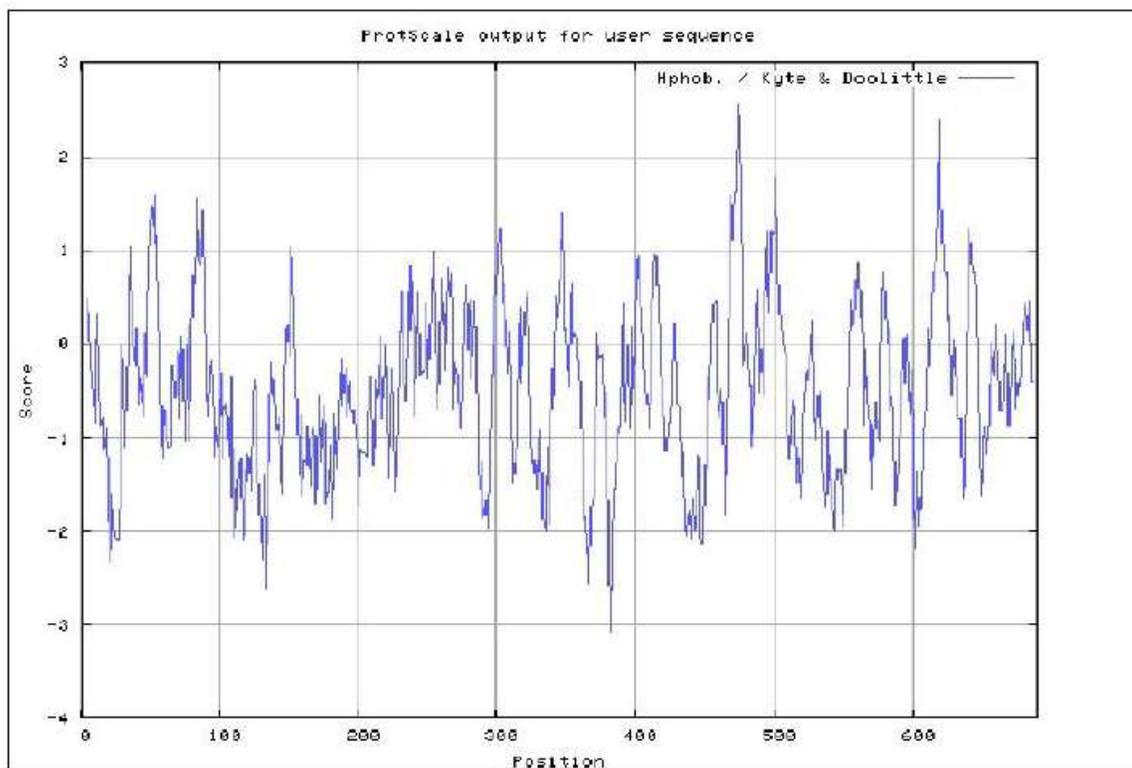


Figure 8. Kyte and Doolittle hydropathy plot for β -galactosidase for *Bacillus halodurans C-125*.

Table 5. Patterns of Cystine - Cystine binding of β -galactosidase for *Bacillus* species.

Accession number	Cystine residues	Status	Score	Probable pattern
YP_004205251.1	5	CYS 123 is not SS-bounded	Score= -17.4	The most probable pattern of pairs: 123-323, 163-168,
		CYS 163 is probably not SS-bounded	Score= -5.0	
		CYS 165 is probably not SS-bounded	Score= -2.7	
		CYS 168 is probably not SS-bounded	Score= -10.9	
		CYS 323 is not SS-bounded	Score= -16.0	
ZP_10508490.1	5	CYS 123 is not SS-bounded	Score= -17.1	The most probable pattern of pairs: 123-323, 163-168,
		CYS 163 is probably not SS-bounded	Score= -13.5	
		CYS 165 is probably not SS-bounded	Score= -2.6	
		CYS 168 is not SS-bounded	Score= -17.3	
		CYS 323 is not SS-bounded	Score= -16.6	
ZP_10511829.1	5	CYS 123 is not SS-bounded	Score= -17.4	The most probable pattern of pairs: 123-323, 163-168,
		CYS 163 is probably not SS-bounded	Score= -7.8	
		CYS 165 is probably SS-bounded	Score= 1.8	
		CYS 168 is probably not SS-bounded	Score= -9.1	
		CYS 323 is not SS-bounded	Score= -16.6	
BAL72724.1	5	CYS 120 is not SS-bounded	Score= -23.0	The most probable pattern of pairs: 160-165,
		CYS 160 is probably not SS-bounded	Score= -5.6	
		CYS 162 is probably SS-bounded	Score= 3.0	
		CYS 165 is probably not SS-bounded	Score= -5.3	
		CYS 379 is not SS-bounded	Score= -37.4	
AFY63015.1	7	CYS 123 is probably not SS-bounded	Score= -14.8	The most probable pattern of pairs: 123-163, 165-168,
		CYS 163 is probably SS-bounded	Score= 0.4	
		CYS 165 is probably SS-bounded	Score= 11.6	
		CYS 168 is probably not SS-bounded	Score= -8.4	
		CYS 323 is not SS-bounded	Score= -24.6	
		CYS 509 is not SS-bounded	Score= -16.3	
		CYS 638 is not SS-bounded	Score= -37.2	

Table 6. Hydrophobicity score and plot of β -galactosidase for *Bacillus* species.

Accession number	Position		Score	
	Min	Max	Min	Max
YP_004205251.1	382	475	-3.111	2.622
ZP_10508490.1	382	475	-3.111	2.622
ZP_10511829.1	382	475	-3.111	2.622
BAL72724.1	541	471	-2.778	2.778
AFY63015.1	382	475	-3.378	2.322
YP_004981461.1	382	475	-3.378	2.322
NP_242888.1	382	474	-3.067	2.544

Conclusions

To obtain desirable results in an industrial application, it

needs to influence the characteristic properties of an enzyme which is a tedious task. Protein engineering techniques used to bring about this goal need a sound

knowledge about the protein both at sequence and structure level. In the present study, 7 β -galactosidase sequences were selected to determine the physicochemical properties and various protein structure levels using *in silico* techniques. Primary structure analysis revealed that most of the β -galactosidase employed in the current study was hydrophobic in nature, and three of them contain disulphide linkages. The secondary structure analysis confirmed that in most of the sequences, random coils dominated followed by an alpha helix, extended strand and beta turns. The presence of Cys residues in β -galactosidase indicates the presence of disulfide bridges which is further confirmed using CYS_REC. This study provides the insight into physicochemical properties and functions of β -galactosidase, thus aiding in formulating their uses as an individual molecule study of β -galactosidase which has shown various unknown properties. Thus, research and development of β -galactosidase finds vast applications in several industries.

Conflict of interests

The authors have not declared any conflict of interests.

ACKNOWLEDGEMENTS

Authors are profusely thankful to the Department of Biotechnology (DBT), Ministry of Science and Technology, Government of India, New Delhi, for funding the Bioinformatics Infrastructure Facility Project (BT/BI/25/001/2006 VOL II date 05-03-2012) and also the Interdisciplinary Program for Life Science Project (BT/PR/4555/INF/22/126/2010 dated 30-09-2010) and P. G Departments of Biotechnology and Microbiology Karnatak University, Dharwad, for providing the facilities for pursuing the research work at the Department.

REFERENCES

- Bairoch A, Apweiler R (2000) The SWISS-PROT protein sequence database and its supplement. TrEMBL in 2000. *Nucleic Acids Res.* 28:45-48.
- Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, Gasteiger E, Martin MJ, Michoud K, O'Donovan C, Phan I (2003). The swiss-prot protein knowledge base and its supplement trembl in 2003. *Nucleic Acids Res.* 31:365-370.
- Cortes G, Trujillo-Roldan MA, Ramirez OT, Galindo E (2005). Production of β -galactosidase by *Kluyveromyces marxianus* under oscillating dissolved oxygen tension. *Process Biochem.* 40:773-778.
- Combet C, Blanchet C, Geourjon C, Deléage G (2000). NPS@: Network Protein Sequence Analysis. *Trends Biochem. Sci.* 25:147-150.
- Domingues L, Lima N, Teixeira JA (2005). *Aspergillus niger* β -galactosidase production by yeast in a continuous high density reactor. *Process Biochem.* 40:1151-1154.
- Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD (2005). Protein Identification and Analysis Tools on the ExPASy Server in: John M. Walker (Eds.), *The Proteomics Protocols Handbook*, Humana Press. pp. 571-607.
- 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.
- Gill SC, VonHippel PH (1989). Extinction coefficient. *Anal Biochem.* 182:319-328.
- Guruprasad K, Reddy B, Pandit MW (1990). Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Prot Eng.* 4:155-164.
- Gouripur G, Kaliwal B (2013). Isolation and characterization of β -galactosidase producing *Bacillus subtilis* from milk. *World J. Pharm. Res.* 3:597-618.
- Hirokawa T, Chieng SB, Mitaku S (1998). *SOSUI*: classification and secondary structure prediction system for membrane proteins. *Bioinformatics* 14:378-379.
- Ikai AJ (1980). Thermo stability and aliphatic index of globular proteins. *J. Biochem.* 88:1895-1898.
- Jurado E, Camacho F, Luzon G, Vicaria M (2002). A new kinetic model proposed for enzymatic hydrolysis of lactose by a β -galactosidase from *Kluyveromyces fragilis*. *Enzyme Microbial. Technol.* 31:300-309.
- Kaur K, Mehmood S, Mehmood A (2006). Hypolactasia as a molecular basis of lactose intolerance. *Ind J. Biochem. Biophys.* 43:267-274.
- Kyte J, Doolittle RF (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105-132.
- Neelima A, Amit KB, Srilaxmi M, Upadhyayula M (2009). Comparative characterization of commercially important xylanase enzymes. *Bioinformation* 3(10):446-453.
- Patil MM, Mallesha KV, Bawa AS (2011). Characterization of partially purified β -galactosidase from *Bacillus Sp* MTCC-864. *Recent Res. Sci. Technol.* 3:84-87.
- Sani RK, Chakraborti S, Sobti RC, Patnaik PR, Banerjee UC (1999). Characterization and some reaction engineering aspects of thermostable extracellular β -galactosidase from a new *Bacillus* species. *Folia Microbiol.* 44:367-371.
- Sivakumar K, Balaji S, Gangaradhakrishna N (2007). In silico characterization of antifreeze proteins using computational tools and servers. *J. Chem. Sci.* 119(5):571-579.
- Sivakumar K (2010). *Biocomputation and Biomedical Informatics: Case Studies and Applications*, Edition: 1, Publisher: Medical Information Science Reference (an imprint of IGI Global), Editors: Athina Lazakidou University of Peloponnese, Greece. pp. 143-157.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28: 2731-2739.
- Vasiljevic T, Jelen P (2002). Lactose hydrolysis in milk as affected by neutralizers used for the preparation of crude β -galactosidase extracts from *Lactobacillus bulgaricus* 11842. *Innov. Food Sci. Emerging Technol.* 3:175-184.
- Vinderola CG, Reinheimer JA (2003). Lactic acid starter and probiotic bacteria: a comparative "in vitro" study of probiotic characteristics and biological barrierresistance. *Food Res. Intern.* 36:895-904.
- Voget CE, Flores MV, Faloci MM, Ertola RJ (1994). Effects of the ionic environment on the stability of *Kluyveromyces lactis* β -galactosidase. *Lebensm. Wiss Technol.* 27:324-330.