Advances in Life Science and Technology
ISSN 2224-7181 (Paper) ISSN 2225-062X (Online)
Vol 4, 2012

Categorizing Phenomenal features of α-Amylase (*Bacillus species*) using Bioinformatic Tools

Pradeep N V[1], Anupama[2*], Ankitha K[3], Pooja J[4].
1, 2 Assistant Professor, Dept of Biotechnology, Ballari Institute of Technology and Management, Bellary, India.
3, 4 B.E Scholar, Dept of Biotechnology, Ballari Institute of Technology and Management, Bellary, India.
* E-mail of the corresponding author: akkushekhar@gmail.com

Abstract

α- Amylase is an industrially important enzyme which hydrolyses starch and glycogen and yields glucose and maltose. α- amylase is used in ethanol production to break starches in grains into fermentable sugars, food industry, pharmaceutical industries, and detergents and in the tanning industries. An attempt has been made to use computational tools to characterize various parameters of α-amylase. The physicochemical properties were computed using ProtParam tool. It was found that the molecular weight ranged between 52911.4-60557.1Da, the pI of all the α-amylase was found to be basic in nature, the stability index of all the enzymes was found to be less than 40 and the grand hydropathy (GRAVY) was found to be negative indicating stability of the enzymes. The secondary structures of enzymes were predicted by SOPMA tool. The random coils predominated the other conformations. Multiple sequence alignment and the phylogenetic analysis were carried out by CLC workbench. Neighbour joining algorithm was used for constructing a phylogenetic tree. The 3D structures of the enzymes were predicted using ESyPred3D server. The above mentioned tools help the biochemist, researchers and students to use these tools for analyzing the properties of proteins.

Keywords: α-Amylase, ProtParam, SOPMA, CLC workbench, ESyPred3D.

1. Introduction

Enzymes are substances present in the cells of living organisms in minute amounts and are capable of speeding up chemical reactions (associated with life processes), without themselves being altered after the reaction. They accelerate the velocity of biochemical reactions (Sajitha *et al*., 2011).

The global market for industrial enzymes estimated as $2 billion in 2004 and expected to rise at an average annual growth rate of 3.3% (Sujata, 2010**).** Among them Amylases are the most important enzymes and are of great significance for biotechnology, constituting a class of industrial enzymes having approximately 25% of the world enzyme market (Reddy, 2003). The history of amylases began in 1811 when the first starch degrading enzyme was discovered by Kirchhoff (Rani, 2003). α- Amylases (E.C.3.2.1.1) are enzymes that catalyses the hydrolysis of internal α-1, 4-glycosidic linkages in starch into low molecular weight products, such as glucose, maltose and maltotriose units. They can be obtained from several sources, such as plants, animals and microorganisms. Some of the microbial enzymes are produced in large quantities for industrial purposes (Bailey and ollis, 1977). The major advantage of using microorganisms for the production of amylases is the economical bulk production capacity and the fact that microbes are easy to manipulate to obtain enzymes of desired characteristics (Souza and Magalhães, 2010).

Each application of α-amylase requires unique properties with respect to specificity, stability, temperature and pH dependence (Nagwa, 2010). α-amylases are used in brewing and fermentation industries for the conversion of starch to fermentable sugars (Chi *et al*., 1995; Farid *et al*., 2002), in the textile industry for designing textiles and in the laundry industry in a mixture with protease and lipase to launder clothes, in the

paper industry for sizing, and in the food industry for preparation of sweet syrups, to increase diastase content of flour, and for the removal of starch in jelly production. After the addition of α amylase in bread-baking process increased the bread's volume and kept its softness longer (Sajitha *et al.,* 2011). Among various extracellular enzymes, a-amylase ranks first in terms of commercial exploitation. Spectrum of applications of a-amylase has widened in many sectors such as clinical, medicinal and analytical chemistry (Hema *et al.,* 2006).

Bacterial amylase are resistant to high temperatures, they are active in wide pH ranges and also they can manage de-sizing in a short time. Because of the above properties bacterial amylases are widely used in de-sizing process (Sarjkaya *et al*., 1989). Among bacteria, *Bacillus* sp. is widely used for thermo stable α-amylase production to meet industrial needs *B. subtilis, B. stearothermophilus, B. licheniformis* and *B. amyloliquefaciens* are known to be good producers of a-amylase and these have been widely used for commercial production (Sivaramakrishnan *et al.,* 2003).

Computational packages and online servers are the current tools used in the protein sequence analysis and characterization. The physicochemical and the structural properties of the proteins are well understood with the use of computational tools. The statistics about a protein sequence such as number of amino acid, sequence length, and the physico-chemical properties of a proteins such as molecular weight, atomic composition, extinction coefficient, GRAVY, aliphatic index, instability index, etc. can be computed by ProtParam (Sivakumar *et al.,* 2007*)*. The secondary structure prediction, sequence similarity, evolutionary relationship and the 3D structure of various proteins can be computed using ESyPred 3D server.

2. Materials and Methods

2.1 Softwares

Windows operating system, ExPASy-ProtParam tool, SOPMA, CLC work bench, EsyPred 3D server.

2.2 Sequence retrieval

Sequences of α-amylases from different bacillus species were retrieved from National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov) having the accession numbers AAA22194.1, ABY86223.1, CAA01355.1, ADE44086.1, AAK00598.1.

2.3 ProtParam

ProtParam (http://www.expasy.org/tools/protparam.html) computes various physicochemical properties that can be deduced from a protein sequence. No additional information is required about the protein under consideration (Walker, 2005). The properties include theoretical isoelectric point (pI), molecular weight, total number of positive and negative residues, extinction coefficient, instability index, aliphatic index and grand average hydropathy (GRAVY) were computed using the Expasy's ProtParam server (Neelima, 2009).

2.4 SOPMA

SOPMA is Self Optimized Prediction Method of Alignment and is used for prediction of secondary structure of proteins (Prashant *et al*., 2010). Self optimized prediction method is based on the homologue method (Mugilan *et al.,* 2010). This method calculates the percentage of α-helix, β-sheets, turns, random coils and extended strands. SOPMA is neural network based method; global sequence prediction may be done by this sequence method. SOPMA is employed for calculating the secondary structural features of the selected protein sequences which are considered for this study. The protein sequences of enzymes in FASTA format were imported and submitted to SOPMA server. The results appeared with secondary structure and percentage of secondary structure of proteins (Prashant *et al.,* 2010).

2.5 CLC work bench

CLC free workbench is integrative, interactive free software assisting in most of the bioinformatics analysis.

The sequences were then subjected for multiple sequence alignment with gap open cost 10.0 and gap open extension cost 1.0. The resulting alignment was used for the construction of evolutionary tree with Neighbour Joining algorithm (Ashokan *et al*., 2011).

2.6 ESyPred 3D

Three-dimensional protein structure is an important source of information to better understand the function of protein, its interactions with other compounds (ligands, proteins, DNA, etc). The 3D protein structure can be predicted according to three main categories of methods (1) Homology or comparative modeling; (2) Fold recognition (predicting global fold of protein); (3) *ab initio* techniques (trying to model the 3D structure of proteins using only the sequence and a force field). In this study ESyPred 3D server was used to predict the 3D structures of α-amylases. This program (ESyPred3D) implements the four steps of the homology modeling approach (1) Databank searching to identify the structural homology, (2) Target-template alignment, (3) Model building and optimization and (4) Model evaluation (Lambert *et al*., 2002).

3. Results and Discussions

3.1 PROTPARAM

The M.Wt of α- Amylases with accession numbers AAA22194.1, ABY86223.1, CAA01355.1, ADE44086.1, AAK00598.1 were found to be 52911.4, 58309.4, 58492.1, 58378.9, 60557.1 Da respectively. The molecular weight of bacillus megaterium (AAK00598.1) is highest with 60557.1 Da. The average molecular weight of α-amylase is calculated as 57729.78 Da. Similar study was conducted by Sivakumar *et al* (2007) for antifreeze proteins.

Isoelectric point (pI) is the pH at which net charge existing on the protein is zero. The pI value of for *Bacillus licheniformis* α-amylase (CAA01355.1) is highest having the value 6.26 and lowest for *Bacillus cereus* (ABY86223.1) with 5.32. The pI for the protein sequences ranges from 5.32-6.26 which indicate all the considered α-amylase sequences are acidic in nature (Vinita, 2011; Neelima *et al.,* 2009; Sivakumar *et al.,* 2007).

The Instability index is a measure of proteins, used to determine whether it will be stable in a test tube. If the index is less than 40, then it is stable in the test tube. If it is greater than it is not stable. Instability index relies upon the occurrence of certain dipeptides along the length of the protein to distinguish between the unstable and stable protein (Vinita, 2011). All the considered sequences have instability index less than 40 which indicates their stability.

The aliphatic index refers to the relative volume of a protein that is occupied by aliphatic side chains and contributes to the increased thermo stability of protein (Vinita, 2011). Higher aliphatic index indicates greater stability of protein. The aliphatic index is higher for *Bacillus subtilis* (AAA22194.1) with 70.42 and is lower for *Bacillus amyloliquefaciens* (ADE44086.1) with 66.77.
The GRAVY value for a peptide or protein is calculated as the sum of hydropathy values of all the amino acids, divided by the number of residues in the sequence (Misaki *et al.,* 2009). Grand average of hydropathicity (GRAVY) index indicates the solubility of proteins: a positive GRAVY value designates it to be hydrophobic nature whereas a negative GRAVY value indicates more surface accessibility of the protein to interact with water (Vinita, 2011). The grand hydropathy number for α-amylases in this study varies from -0.487 to -0.606 indicating more surface accessibility of the protein to interact with water.

The physicochemical properties of the enzymes under study are tabulated in table 1 (M. Wt -Molecular weight; pI -Isoelectric point; II -Instability index; AI -Aliphatic index; GRAVY -Grand Average Hydropathy).

3.2 SOPMA

The secondary structure evaluation of protein is carried out with the help of Self Optimized Prediction Method with Alignment (SOPMA). The secondary structure indicates whether a given amino acid lies in a helix, strand or coil. Secondary structural features are predicted using SOPMA is represented in Table 2. The random helix predominates with α-helix contributing to the stability of the secondary structures (Sivakumar *et al.,* 2007: Ashokan *et al*., 2011). Secondary structures infer that all the α- Amylases are rich in random coils varying from 43.19 to 39.84%. They have less number of β-turns.

3.3 CLC Workbench

Multiple Sequence Alignment

Multiple sequence Alignment was carried out to calculate the taxonomical variation in α-amylase from different *Bacillus* species (Figure 1). α- Amylases from the organisms *B. cereus, B. megaterium, B. licheniformis, B. amyloliquefaciens* show high degree of similarity whereas *B. subtilis* show less similarity. The similar studies were conducted by Ashokan *et al.,* (2011) on catalase and by Sivakumar *et al.,* (2007) on antifreeze proteins.

Phylogenetic Analysis

Phylogenetics is the study of evolutionary relationship between the organisms that are derived from common ancestors. The phylogenetic tree was obtained by neighbor joining algorithm (Figure 2). The similar studies were conducted by Ashokan *et al.,* (2011) on catalase.

3.4 ESyPred 3D

ESyPred3D is an automated homology modeling method. It is used to obtain the 3D structure of proteins. When the sequences of the organisms were subjected to ESyPred3D server the three-dimensional structure of proteins were obtained. Similar studies were carried out by Lambert (2002).

4. Conclusions

Five different sources of α-amylase have been chosen for the study of physicochemical parameters, secondary structure prediction, multiple sequence alignment, phylogenetic analysis and three-dimensional structure prediction of α- Amylases by using computational tools and servers. By using ProtParam tool the physicochemical characters were analyzed. Physicochemical characters revealed that the molecular weight of α- Amylase from *Bacillus megaterium* (AAK00598.1) is highest with 60557.1 Da. The molecular weight is lowest of α- Amylase from *Bacillus subtilis* (AAA22194.1) with 52911.4 Da. The average molecular weight of α-amylase was found to be 57729.78 Da. The pI of all the organisms were found to be acidic in nature. All the considered sequences have instability index less than 40, indicating that they are stable. All the five organisms have negative GRAVY value which indicates more surface accessibility of the protein to interact with water. The secondary structures of proteins were predicted using SOPMA tool, it was seen that the percentage of random coils were more and it predominated the other conformations and the least percentage of conformation was seen in β-turns. Multiple sequence alignment and phylogenetic analysis were conducted by using CLC workbench which showed that the organisms *Bacillus cereus, Bacillus megaterium, Bacillus licheniformis, Bacillus amyloliquefaciens* shows high degree of similarity whereas *Bacillus subtilis* show less similarity. The three-dimensional structures of proteins were obtained by ESyPred3D server. This helps researchers and students to use these tools and characterize the properties of enzymes and proteins and also to conduct secondary structure prediction, multiple sequence alignment,

phylogenetic analysis and the 3D structures.

References

Ashokan K V, Mundaganur D S and Mundaganur Y D. (2011), "Catalase: Phylogenetic characterization to explore protein cluster", *Journal of research in Bioinformatics*, **1**, 001-008.

Bailey J. E and Ollis D. F. (1986), "Biochemical engineering fundamentals", *McGraw-Hill*, 2nd edition.

Christophe L, Nadia L, Xavier D B and Eric D. (2002), "ESyPred3D: Prediction of proteins 3D structures", *Bioinformatics*, **18** (9), 1250-1256.

Hema A, Ujjval T and Kamlesh P. (2006), "Alpha Amylase Production by *Bacillus cereus* MTCC 1305 Using Solid-State Fermentation", *Food Technol. Biotechnol*, **44** (2), 241–245.

Misaki W, Henry K, Kaddu D M, Joseph O, Wilson B. (2008), "Proteomics of Marburg and Ebola glycoproteins: Insights into their physicochemical similarities and Irregularities", *African Journal of Biotechnology*, **8** (17), 4025-4031.

Mugilan A, Ajitha M C, Devi and Thinagar. (2010), "In silico Secondary Structure Prediction Method (Kalasalingam University Structure Prediction Method) using Comparative Analysis", *Trends in Bioinformatics*, **3** (1), 11-19.

Nagwa M S, Maha A A, Abeer M, Fatma S, and Nouf A. (2010), "Screening of Microorganisms Isolated from some Enviro-Agro-Industrial Wastes in Saudi Arabia for Amylase Production", *Journal of American Science*, **6** (10), 926-939.

Neelima A, Amit K B, Srilaxmi M, Upadhyayula S M. (2009), "Comparative characterization of commercially important xylanase enzymes", *Bioinformation*, **3** (10), 446-453.

Prashanth V T, Uddhav S C, Madura S M, Vishal P D and Renuka R K. (2010), "Secondary Structure Prediction and Phylogenetic Analysis of salt Tolerant Proteins", *Global Journal of molecular Sciences*, **5**(1), 30-36.

Rani G, Paresh G, Harapriya M, Vineet K G, Bhavna C. (2003), "Microbial a-amylases: a biotechnological perspective", *Process Biochemistry*, **38** , 1599-/1616.

Reddy N S, Annapoorna N and Sambasiva K R S. (2003), "An overview of the microbial α-amylase family", *African Journal of Biotechnology*, **2** (12), 645-648.

Sajitha N, Vasanthabharathi V, Lakshminarayanan R and Jayalakshmi S. (2011), "Amylase from an Estuarine Bacillus megaterium", *Current Research Journal of Biological Sciences*, **3** (2), 110-115.

Sarjkaya E and Cirakoglu C. (1989), "Investigation of α-amylase production from Bacillus Subtillis in different media.Commun Fac", *Sci. Univ. Ank. Serie C V*, **7**, 31-37.

Sivakumar K, Balaji S and Gangaradhakrishnan. (2007), "In silico characterization of antifreeze proteins using computational tools and servers", *J. Chem. Sci*, **119** (5), 571-579.

Sivaramakrishnan, Dhanya G, Kesavan M N, Carlos R S and Ashok P. (2006), "α-Amylases from Microbial Sources – An Overview on Recent Developments", *Food Technol. Biotechnol*, **44** (2), 173–184.

Souza P M and Magalhães P O. (2010), "Application of Microbial α-Amylase in Industry – A Review", *Brazilian Journal of Microbiology*. **41**, 850-861.

Sujata D B. (2010), "Screening of nutritional components for α-Amylase Production in Submerged fermentation by bacteria isolated from soil using Plackettburman Design", *International Journal of Pharmacy and Pharmaceutical Sciences*, **2** (1), 93-98.

Vinita H. (2011), "Physiochemical, Functional and Structural Characterization of Wheat Germin Using In silico Methods", *Current Research Journal of Biological Sciences*, **3** (1), 35-41.

Walker J M. (2005), "The Proteomics Protocols Hand Book", *Human Press* (Chap.25).

Anupama is an M.Tech (Environmental Engineering) scholar from K.L.E's College of Engineering and technology, Belgaum, India. She obtained her Engineering Degree in Biotechnology from B.I.T.M, Bellary, India. Published three international papers related to wastewater treatment. Presently working as an Assistant Professor in Dept of Biotechnology, Ballari Institute of Technology and Management, Bellary, India. She is a member of World Environmental Organization and International Society of Environmental Information Sciences.

Pradeep N V is an M.Tech (Environmental Engineering) scholar from K.L.E's College of Engineering and technology, Belgaum, India. He obtained his Engineering Degree in Biotechnology from B.I.T.M, Bellary, India. Published four international papers related to wastewater treatment. Presently working as an Assistant Professor in Dept of Biotechnology, Ballari Institute of Technology and Management, Bellary, India. A member of World Environmental Organization and International Society of Environmental Information Sciences.

Table 1: Parameters computed using Expasy's ProtParam tool.

Organism Name	Accession No	No of amino acids	M. Wt	pI	II	AI	GRAVY
Bacillus subtilis	AAA22194.1	477	52911.4	5.33	34.55	70.42	-0.487
Bacillus cereus	ABY86223.1	513	58309.4	5.32	22.14	67.25	-0.624
Bacillus licheniformis	CAA01355.1	512	58492.1	6.26	29.44	70.12	-0.595
Bacillus amyloliquefaciens	ADE44086.1	514	58378.9	5.95	34.13	66.77	-0.620
Bacillus megaterium	AAK00598.1	533	60557.1	5.80	21.12	67.49	-0.606

Table 2: Secondary structure prediction of proteins using SOPMA tools.

Organism Name	Accession No.	No of amino acids	α –Helix (Hh) (%)	β -Turn (Tt) (%)	Extended Strands (Ee) (%)	Random Coils (%)
Bacillus subtilis	AAA22194.1	477	33.12%	5.66%	18.03%	43.19%
Bacillus cereus	ABY86223.1	513	31.19%	5.65%	21.83%	41.33%
Bacillus licheniformis	CAA01355.1	512	33.01%	5.27%	21.88%	39.84%
Bacillus amyloliquefaciens	ADE44086.1	514	29.96%	6.81%	21.01%	42.22%

Bacillus megaterium	AAK00598.1	533	30.21%	5.82%	20.83%	43.15%

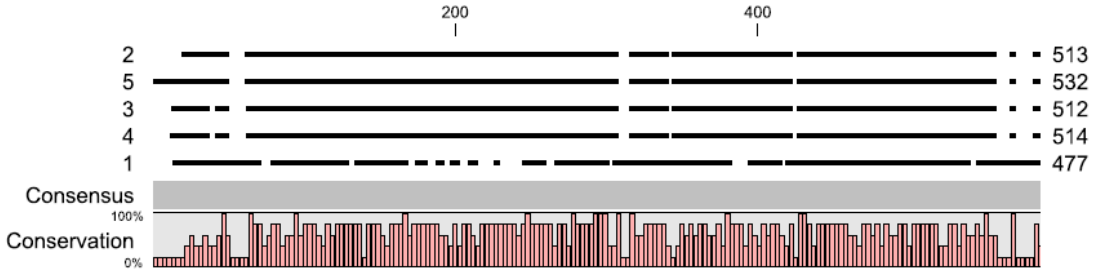

Figure 1: Result of multiple sequence alignment (1- *B. subtilis* 2-*B. cereus;* 3-*B. licheniformis; 4-B. amyloliquefaciens; 5-B. megaterium*).

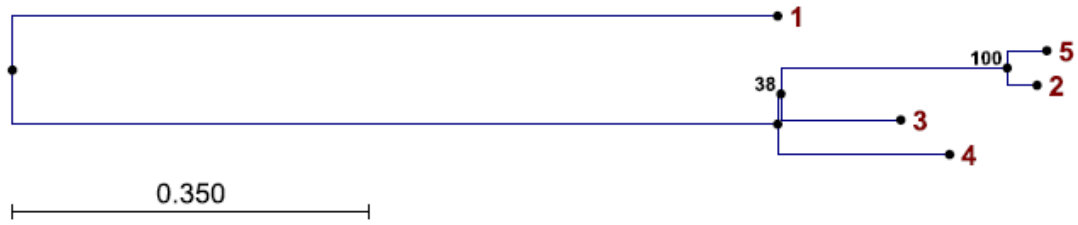

Figure 2: Result of phylogenetic tree obtained by neighbor joining method (1- *B. subtilis* 2-*B. cereus;* 3-*B. licheniformis; 4-B. amyloliquefaciens; 5-B. megaterium*).

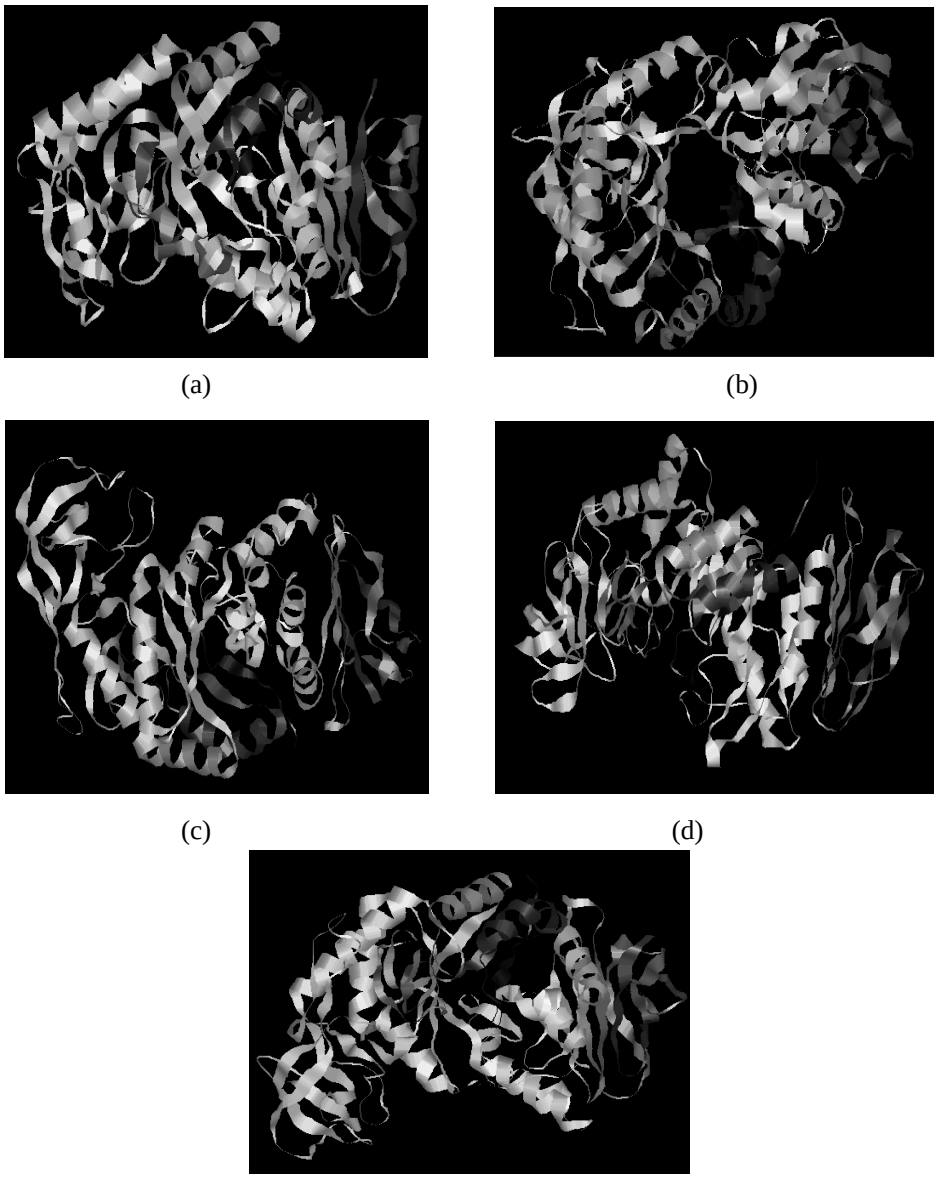

Figure 3: Three-dimensional structure of α-amylase of five different organisms is predicted by ESyPred tool a- *B. subtilis;* b- *B. cereus;* c- *B. licheniformis; d- B. amyloliquefaciens; e- B. megaterium).*