Biocatalysis and Agricultural Biotechnology 22 (2019) 101404

Contents lists available at ScienceDirect

Biocatalysis and Agricultural Biotechnology

journal homepage: http://www.elsevier.com/locate/bab

Insight into the structural configuration of metagenomically derived lipase from diverse extreme environment

Rajesh Kumar Sahoo[a], A. Swaroop Sanket[b], Mahendra Gaur[a], Aradhana Das[a], Enketeswara Subudhi[a,*]

[a] Centre for Biotechnology, School of Pharmaceutical Science, Siksha 'O' Anusandhan (Deemed to be University), Campus-2, Kalinga Nagar, Ghatikia, Bhubaneswar, 751003, Odisha, India
[b] Microbiology Department, ICMR-Regional Medical Research Center, NALCO Nagar, Chandrasekharpur, Bhubaneswar, 751023, Odisha, India

ARTICLE INFO

Keywords:
Lipase
Physico-chemical characteristics
Extended helices
Homology modelling
Ramachandran plot
VERIFY3D

ABSTRACT

The present study aims at finding the structural similarity of five metagenomically derived lipase protein sequences from diverse extreme environmental conditions i.e. hot springs, low-temperature regions, anaerobic lagoons and marine source. Using PROTPARAM server, the analysis of physicochemical characteristics of all lipases were found to be neutral in pH and are hydrophilic in nature except clone RK-Lip-479 and clone JKP01. The molecular weight varied from 31.2 kDa to 46.4 kDa while amino acid residues varied from 293 (clone FosC12) to 417 (clone JKP01). The instability index of all the proteins ranged from 22 to 35, which indicate their stability under given conditions except clone h1Lip1 lipase with higher instability index of 44.72 indicating less stable. Determination of the secondary structure of all proteins by SOPMA tool revealed the dominance of α-helices ranging between 27.34% (clone h1Lip1 lipase) to 45.05% (clone FosC12 lipase) and coils ranged between 27.3% (clone FosC12 lipase) to 34.53% (clone JKP01 lipase). 3D structures of all lipase proteins deduced using SWISS-MODEL automated server were found to be monomeric. The structures were retrieved and visualized using PyMol v0.99. The QMEAN values were within the range of 0.8–1.5 suggesting the models to be of suitable. The outcome of RAMPAGE and VERIFY3D servers for the models obtained were also observed to be of much suitable. The insight into the structural configurations of proteins revealed as above may help in varied applicability of metagenomics lipases.

1. Introduction

Lipases belonging to hydrolase class of enzymes catalyze triglycerides by releasing fatty acids and glycerol (Messaoudi et al., 2011). Lipases can be obtained from different domains of living organisms and widely exist in microorganisms belonging to varied habitats. The enzymatic classification of family I bacterial "true" lipases is based on a sequence comparison of their amino acids and the physicochemical and biological properties, which is further identified into 11 subfamilies (Messaoudi et al., 2010). Although lipases belong to many different protein families, they have the same architecture, the α/β-hydrolase fold (Ollis et al., 1990). These enzymes have multitude of hydrolytic actions on several carboxyl esters substrates at the water–lipid interface in non-aqueous environment; esterification, transesterification, and inter-esterification (Jaeger et al., 1999; Sahoo et al., 2017; Villeneuve et al., 2000). This versatility makes lipases "the enzymes of choice" for potential applications in the food, dairy, pharmaceuticals, detergents, textile, paper industry, cosmetic and biodiesel (Sahoo et al., 2017) as well as in the synthesis of fine chemicals and new polymeric materials (Hasan et al., 2006). Their discovery from varied sources can be of biotechnological importance (Ken Ugo et al., 2017). The multifunctional properties of lipase enzymes are due their diverse structure, thus these enzymes exhibit high stability to alkaline conditions, adaptability to cold, tolerant to various organic solvents, metals and detergents and could be of great interest for various industrial applications (Sahoo et al., 2017). Moreover, these enzymes can be used as chiral catalysts for stereoselective conversion of a variety of amines and primary and secondary alcohols for production of cosmetics and pharmaceuticals (Jaeger and Eggert, 2002; Khan and Jithesh, 2012). The catalytic triad of lipases remains in the order of Serine-Asparagine-Histidine residues in

* Corresponding author.
E-mail addresses: rajeshkumarsahoo@soa.ac.in (R.K. Sahoo), swaroop.sanket@gmail.com (A.S. Sanket), mahendragaur@soa.ac.in (M. Gaur), aradhanadas@soa.ac.in (A. Das), enketeswarasubudhi@soa.ac.in (E. Subudhi).

https://doi.org/10.1016/j.bcab.2019.101404
Received 3 August 2019; Received in revised form 20 September 2019; Accepted 26 October 2019
Available online 31 October 2019
1878-8181/

Table 1
Lipase sequences retrieved from NCBI database.

S. No.	Accession Number	Microorganism	Source
1	KR232658	Uncultured bacterium clone RK-Lip-479 lipase	Hot spring (High temperature active, >45 °C)
2	DQ118648	Uncultured bacterium clone h1Lip1 lipase	Marine sediment (Low temperature active, <40 °C)
3	FJ392756	Uncultured bacterium clone JKP01 lipase	Hot spring (High temperature active, >64 °C)
4	KC152848	Uncultured bacterium clone 29-A lipase	Marine sponge, South China Sea (Low temperature active, <30 °C)
5	JF417979	Uncultured bacterium clone FosC12 lipase	Anaerobic Lagoon (Low temperature active, <30 °C)

amino acid sequences of α/β hydrolases and forms the basis for functionality of lipases. Several lipases have been observed to contain the consensus sequence of Gly-Xaa-Ser-Xaa-Gly (GXSXG) with variable X amino acid (Xaa) residue (Kim et al., 1997).

Bioinformatics approaches provide an insight to understand structural as well as physicochemical properties of proteins (Satpathy et al., 2016). In order to understand the molecular organization and function of lipases, deduction of the 3-dimensional structure of the protein by homology modelling is valuable as the structure for large number of proteins are still undetermined by crystallography. Molecular modelling based on homologous sequences remains the most accurate approach to select suitable lipases for wide varieties of applications (Ginalski, 2006).

In the present study, physicochemical aspects of lipase proteins of metagenomic origin were studied. Properties such as isoelectric point, extinction coefficient, instability index, aliphatic index and grand average hydropathy were taken into account.

2. Materials and methods

2.1. Data retrieval

Translated protein sequences of five metagenomic derived lipases i.e. DQ118648, FJ392756, KC152848, JF417979 and KR232658 (submitted by us) (Table 1) were retrieved from NCBI protein database having accession numbers AAZ67909, ACJ07039, AGE44121, AEK97793 and AKQ44144 respectively.

2.2. Physicochemical characterization of protein sequence

The sequences were subjected to PROTPARAM server (Wilkins et al., 1999) for analysis of molecular weight, amino acid length, aliphatic index, instability index, grand average hydropathy, isoelectric point, positively charged residues, negatively charged residues and extinction coefficient.

2.3. Analysis of secondary structures

The prediction and analysis of secondary structures of proteins generally reveal several important aspects of a protein. The linear sequence of amino acids was used for prediction of secondary structure and determined by the pattern of hydrogen bonding. The secondary structures were the iterative organization in 3-D space of conjugative amino acid residues in long peptide chain. The retrieved protein sequences were subjected to SOPMA server (Geourjon and Deléage, 1995) for analysis of secondary structures.

2.4. Modelling and validation

The 3D structures of all proteins were deduced using the online SWISS-MODEL server using the automated mode (Biasini et al., 2014; Tamboli et al., 2015). The SWISS model server uses the GROMOS96 force field to minimize the energy during model building. The deduced models were further subjected to deduce the Ramachandran plot using RAMPAGE server (Lovell et al., 2003). VERIFY3D servers were used to verify the built models (Eisenberg et al., 1997). PyMol v2.1 (Delano, 2002) was used to visualize the models in PDB format.

3. Results & discussion

3.1. Physicochemical characterization of protein sequences

Computation of physicochemical properties of lipases was done by PROTPARAM server (Wilkins et al., 1999). The results for all the physicochemical characteristics considered in the present study are

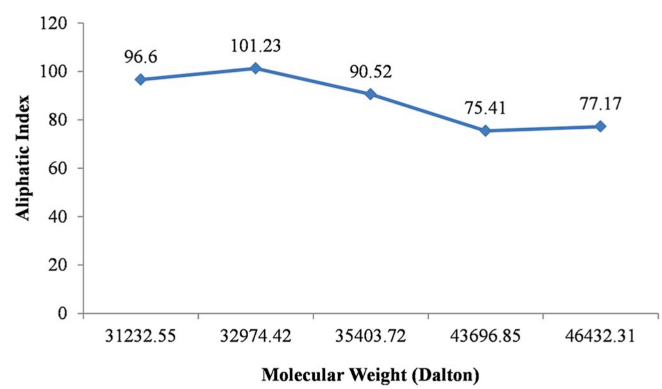


Fig. 1. Correlation between molecular weight (Daltons) of the proteins and the respective aliphatic indices.

Table 3
Proportions of different secondary structure predicted using SOPMA server in all five lipase.

Accession No.	Helix (%)	Extended Helices (%)	Beta (%)	Coil (%)
KR232658	31.01	23.56	14.18	31.25
DQ118648	41.54	12.92	12	33.54
FJ392756	27.34	23.98	14.15	34.53
KC152848	37.25	18.63	11.44	32.68
JF417979	4505	14.33	13.31	27.3

Table 2
Physicochemical characteristics of metagenomic lipases.

Accession No.	MW (Dalton)	AA	AI	II	GRAVY	pI	+R	-R	EC ($M^{-1}cm^{-1}$)
JF417979	31232.60	293	101.23	22.52	+0.075	5.98	21	26	28880
KC152848	32974.40	306	96.60	27.62	+0.039	5.77	21	26	25120
DQ118648	35403.70	325	90.52	44.72	+0.001	6.55	28	30	57660
KR232658	46353.21	416	77.84	33.77	−0.290	6.50	40	43	83560
FJ392756	46432.30	417	77.17	34.43	−0.337	6.64	41	43	87905

(MW: molecular weight; AA: amino acid residues; AI: aliphatic index; II: instability index; GRAVY: grand average of hydropathy; pI: isoelectric point; +R: positively charged residues; –R: negatively charged residues; EC: extinction coefficient).

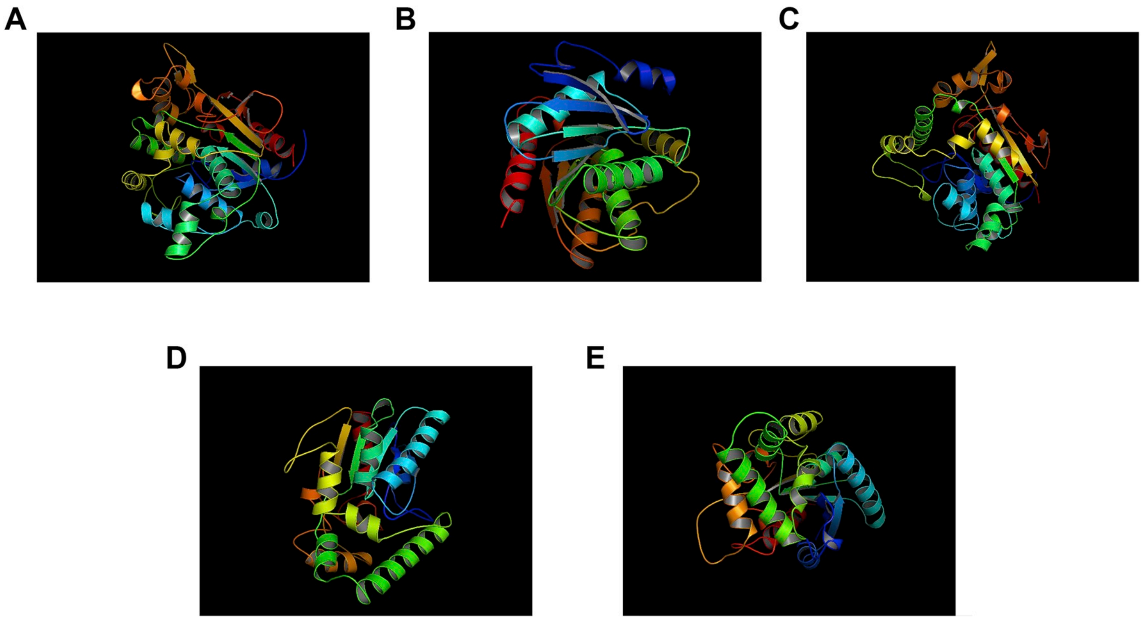


Fig. 2. 3-D models of a) Uncultured bacterium clone RK-Lip-479 lipase **b)** Uncultured bacterium clone h1Lip1 lipase **c)** Uncultured bacterium clone JKP01 lipase **d)** Uncultured bacterium clone 29-A lipase **e)** Uncultured bacterium clone FosC12 lipase.

Table 4
Model summary.

Accession No.	Sequence Identity	QMEAN SCORE
KR232658	86.45%	0.87
DQ118648	73.12%	0.94
FJ392756	66.30%	1.12
KC152848	89.93%	1.02
JF417979	76.10%	0.72

given in Table 2. The isoelectric point (pI) of all proteins ranged from 5.77 to 6.64 suggesting the proteins to have slightly acidic near-neutral for both positively and negatively charged residues. The extinction coefficient (EC) values were in the range of 25120–87905 M^{-1} cm^{-1} suggesting a weak to strong absorbance of light during interaction with substrates. The clone JKP01 lipase showed the highest values signifying the presence of high amount of trp, tyr and cys residues. The values of instability index were in the stability range except clone h1Lip1 lipase which has an instability index (II) of 44.72 and thus considered to be unstable. The grand averages of hydropathy index (GRAVY) of all lipases under study were positive except for uncultured bacterium clone JKP01 lipase and RK-Lip-479 lipase. This directs the interpretation that RK-Lip-479 lipase and JKP01 lipase interact less with water whereas others are more hydrophilic. The correlation between higher value of aliphatic index (AI) of a protein with higher thermo-stability has been interpreted earlier (Ikai, 1980). In the present study, all the protein

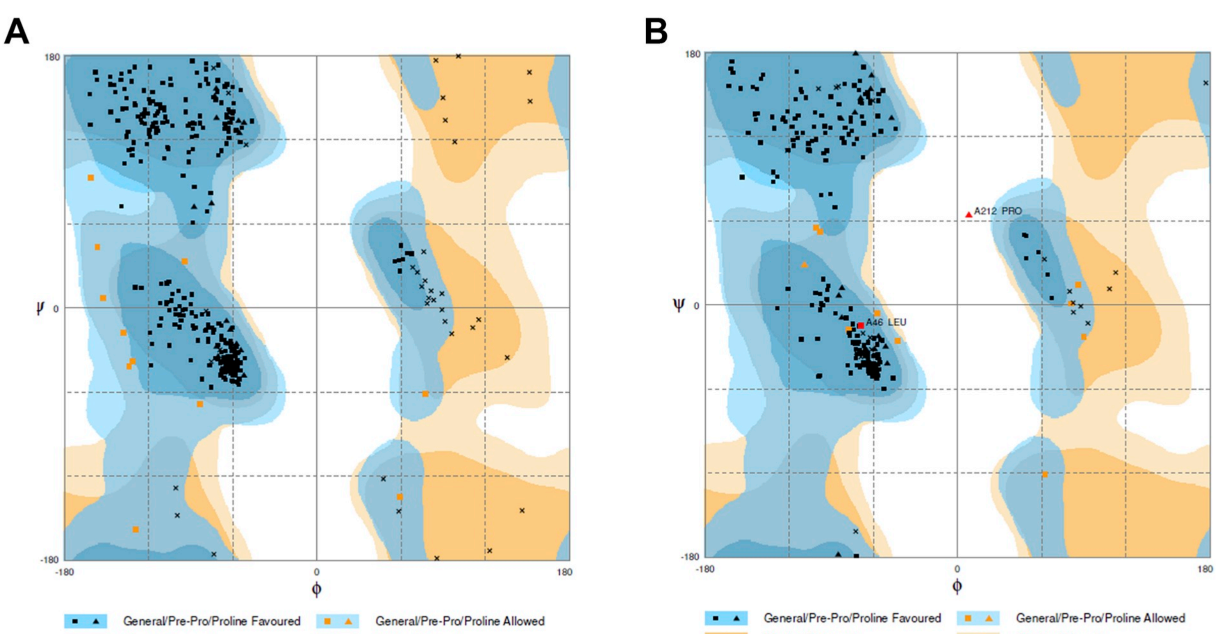


Fig. 3. Ramachandran plots for a) Uncultured bacterium clone RK-Lip-479 lipase **b)** Uncultured bacterium clone h1Lip1 lipase **c)** Uncultured bacterium clone JKP01 lipase **d)** Uncultured bacterium clone 29-A lipase **e)** Uncultured bacterium clone FosC12 lipase.

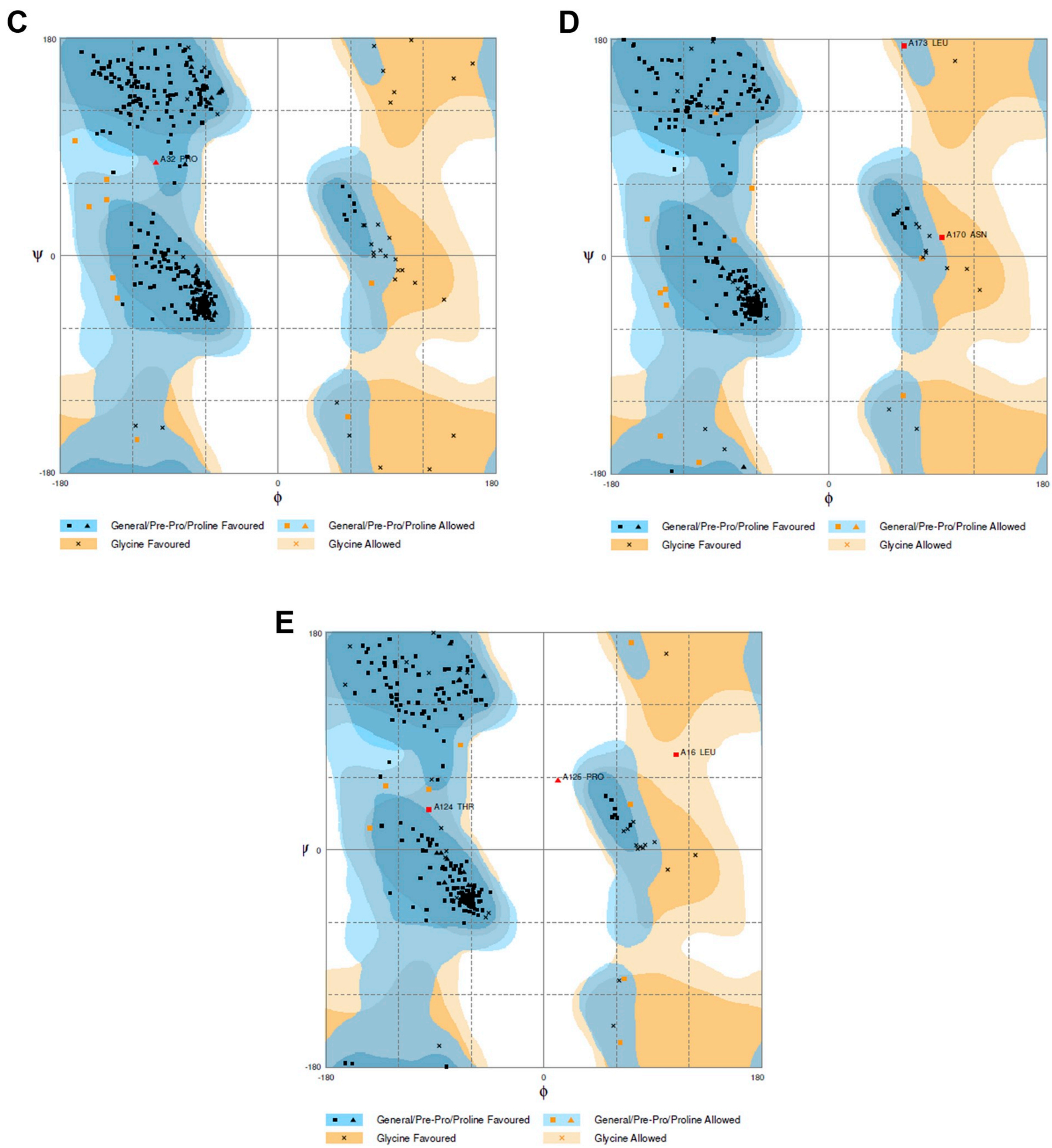


Fig. 3. (*continued*).

Table 5
Number of residues in favored region, allowed region, outlier region.

Accession	Number of residues in favored region	Number of residues in allowed region	Number of residues in outlier region
KR232658	97.1	2.9	0.0
DQ118648	95.6	3.6	0.0
FJ392756	97.4	2.3	0.3
KC152848	95.4	3.9	0.7
JF417979	96.2	2.8	1.0

sequences have aliphatic indices above 77 which definitely prove the stability of these proteins in extreme conditions. The aliphatic index varied from ~77 to ~102; hence, we tried to find any significant correlation between molecular weight and the aliphatic index of all the protein sequences (Ikai, 1980). We observed an inverse correlation between molecular weight and aliphatic index (Fig. 1), the reason being the influence, presence of bulky amino acids, hydrogen bonding and single/multiple mutations (Sharma et al., 2012).

3.2. Analysis of secondary structures

The secondary structure reveals important aspects of proteins. Secondary structural elements of lipases were calculated using SOPMA tool. Table 3 shows the percentage of α-helices, extended β-sheets, β-sheets and coils observed in the metagenomic lipases. Alpha helix content (45.05%) was greatest in clone FosC12 lipase whereas the lowest (27.34%) was found in clone JKP01 lipase. The lower content of α -helix in uncultured bacterium clone JKP01 lipase may be due to high content of proline and glycine, which are regarded as helix breakers. The high

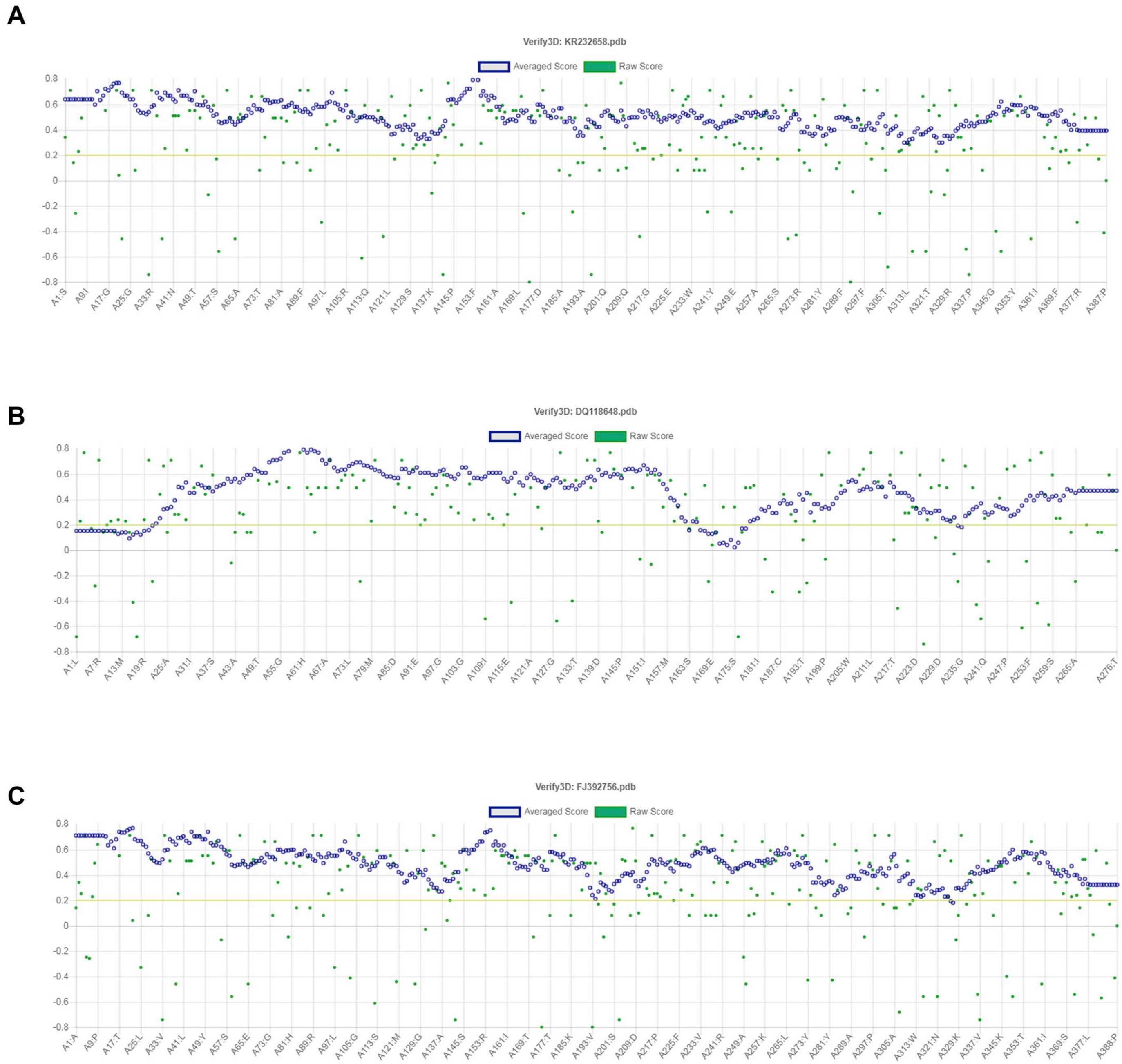


Fig. 4. VERIFY_3D plots for a) Uncultured bacterium clone RK-Lip-479 lipase **b)** Uncultured bacterium clone h1Lip1 lipase **c)** Uncultured bacterium clone JKP01 lipase **d)** Uncultured bacterium clone 29-A lipase **e)** Uncultured bacterium clone FosC12 lipase.

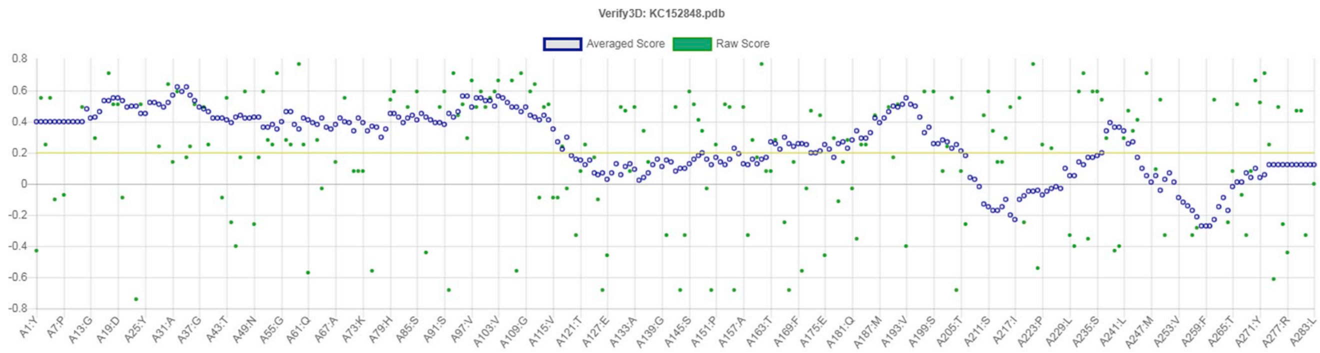
percentage of α-helices in uncultured bacterium clone FosC12 lipase (45.05%) and uncultured bacterium clone h1Lip1 lipase (41.54%) is attributed to the presence of high content of methionine, alanine, leucine, glutamate and lysine. In case of β-sheets, the presence of high content of aromatic amino acid residues tyrosine, tryptophan and phenylalanine and the C^{β}-branched amino acids isoleucine, valine and threonine are attributed. Only 12% of β-sheets were observed in uncultured bacterium clone h1Lip1 lipase and 14.18% for clone JKP01 lipase. However further studies are needed to confirm the role of such amino acids in regulating the secondary structure of amino acids.

3.3. Modelling and validation

Homology modelling requires a suitable template having more than 25% sequence identity. All the lipase proteins (target) under consideration had a similarity more than 35% of sequence identity with the templates. 3D structures of all lipase proteins were deduced using SWISS-MODEL automated serve. The SWISS-MODEL uses the query sequence against the template repository by aligning thereby performing structural superimposition and comparison. Based on the QMEAN scores, quality of models were selected from SWISS-MODEL and structures were retrieved and visualized using PyMol v0.99 (Fig. 2 a, b, c, d, e). All the models obtained were monomeric. The QMEAN values varied from 0.8 to 1.5 suggesting the models to be of best-fitting (Table 4).

When Ramachandran plot was plotted, most of the amino acids of proteins under consideration were in the favorable region (>95%) whereas only number of residues in allowed region ranged from 2.3 to 3.9% and number of residues in outlier region ranged from 0.0 to 1.0%

D



E

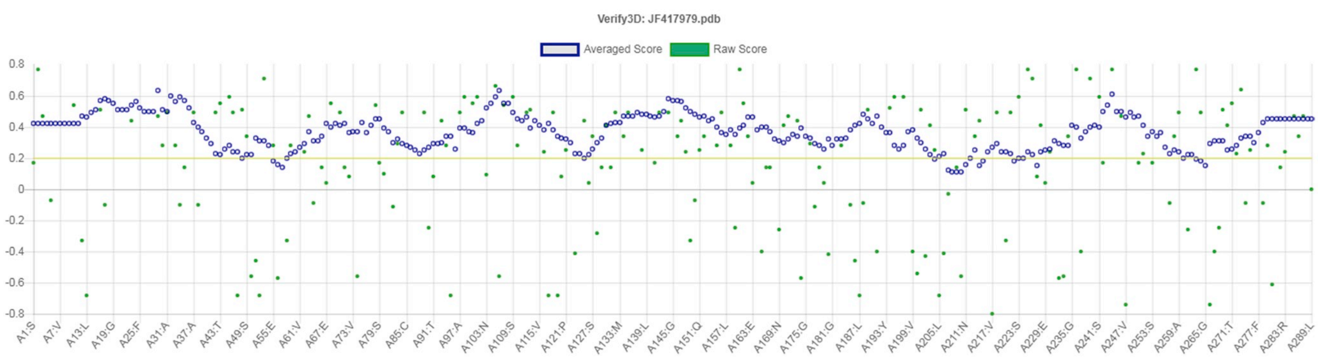


Fig. 4. (*continued*).

(Fig. 3 a, b, c, d, e). This suggested overall good quality of protein structures (Table 5).

The secondary structure of proteins under consideration showed the presence of α-helices and extended helices with higher values that of β-sheets and turns. It is considered that α-helices are the most common protein structural element of transmembrane proteins. The α-Helices remain packed in the membranes while leaving no polar groups exposed to water. The three-phase behavior of stiff-soft-stiff tangent modulus in α-Helices helps the proteins sustain their activities under extreme conditions. It is also evident from the present study that all proteins belong to extreme conditions and are rich in α-Helices (Ackbarow et al., 2007). All the models were also verified by VERIFY3D server (Fig. 4 a, b, c, d, e). Results indicated that all models had an average of ~80% residues with score ≥ 0.2 in the 2D/1D profile suggesting it to be of good quality models. Moya-Salazar et al. (2019) modelled lipase proteins of *Mucor circinelloides, F. circinelloides* 1006PhL, and *Rhizopus oryzae* and validated the models using ERRAT and VERIFY 3D with ~89% accuracy (most favored regions) for the proteins. The characteristics and 3D-models obtained in this present study exhibit potential for further insights taking applicability of lipase proteins of metagenomic origin into consideration. The outcomes also suggest that lipase proteins from extremophilic sources have higher functionality as also observed by Sahoo et al. (2017).

4. Conclusion

Metagenomics is to be considered an alternative and reliable approach for deciphering the microbial diversity circumventing in vitro culturing the microorganisms. Several industrially important enzymes have been isolated and characterized through metagenomic approaches and have been implemented in various industries with quality outcomes as compared to those derived from cultivable approaches. Lipases of metagenomic origin have been reported to poses better physicochemical characteristics and protein structures. The study of physicochemical characteristics like; pI, MW, amino acid residues, GRAVY, extinction coefficient etc. provide an insight into the structural organization of amino acids and peptides. This study can conclude that understanding structural and functional aspects of lipases of metagenomic origin will be of utmost importance as lipases are widely used in industries. RAMAPAGE and VERIFY3D mediated verification of the constructed models also further provide a deep insight into the folding behavior of proteins. The predicted 3D structures will be helpful for determining active sites, molecular docking studies on substrates, inhibitors etc.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The financial support of Department of Biotechnology (File no. BT/PR7944/BCE/8/1035), Minstry of Science and Technology, New Delhi, India for the present work is greatly acknowledged.

References

Ackbarow, T., Chen, X., Keten, S., Buehler, M.J., 2007. Hierarchies, multiple energy barriers, and robustness govern the fracture mechanics of alpha-helical and beta-sheet protein domains. Proc. Natl. Acad. Sci. U. S. A 104, 16410–16415. https://doi.org/10.1073/pnas.0705759104.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T.G., Bertoni, M., Bordoli, L., Schwede, T., 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res. https://doi.org/10.1093/nar/gku340.

Delano, W.L., 2002. The PyMOL Molecular Graphics System. Schrodinger. https://pymol.org/2/.

Eisenberg, D., Lüthy, R., Bowie, J.U., 1997. VERIFY3D: assessment of protein models with three-dimensional profiles. Methods Enzymol. https://doi.org/10.1016/S0076-6879(97)77022-8.

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.

Ginalski, K., 2006. Comparative modelling for protein structure prediction. Curr. Opin. Struct. Biol. 16, 172–177. https://doi.org/10.1016/j.sbi.2006.02.003.

Hasan, F., Shah, A.A., Hameed, A., 2006. Industrial applications of microbial lipases. abdulhameed hameed - Academia.edu. Enzym. Microb. Technol. 39, 235–251.

Ikai, A., 1980. Thermostability and aliphatic index of globular proteins. J. Biochem. 88, 1895–1898. https://doi.org/10.1093/oxfordjournals.jbchem.a133168.

Jaeger, K.-E., Dijkstra, B.W., Reetz, M.T., 1999. Bacterial biocatalysts: molecular biology, three-dimensional structures, and biotechnological applications of lipases. Annu. Rev. Microbiol. 53, 315–351. https://doi.org/10.1146/annurev.micro.53.1.315.

Jaeger, K.-E., Eggert, T., 2002. Lipases for biotechnology. Curr. Opin. Biotechnol. 13, 390–397. https://doi.org/10.1016/S0958-1669(02)00341-5.

Ken Ugo, A., Vivian Amara, A., CN, I., Kenechuwku, U., 2017. Microbial lipases: a prospect for biotechnological industrial catalysis for green products: a review. Ferment. Technol. 06 https://doi.org/10.4172/2167-7972.1000144.

Khan, M., Jithesh, K., 2012. Expression and purification of organic solvent stable lipase from soil metagenomic library. World J. Microbiol. Biotechnol. 28, 2417–2424. https://doi.org/10.1007/s11274-012-1051-0.

Kim, K.K., Song, H.K., Shin, D.H., Hwang, K.Y., Sun, S.W., 1997. The crystal structure of a triacylglycerol lipase from *Pseudomonas cepacia* reveals a highly open conformation in the absence of a bound inhibitor. Structure 5, 173–185. https://doi.org/10.1016/S0969-2126(97)00177-9.

Lovell, S.C., Davis, I.W., Arendall, W.B., De Bakker, P.I.W., Word, J.M., Prisant, M.G., Richardson, J.S., Richardson, D.C., 2003. Structure validation by Cα geometry: ϕ,ψ and Cβ deviation. Proteins Struct. Funct. Genet. https://doi.org/10.1002/prot.10286.

Messaoudi, A., Belguith, H., Gram, I., Ben Hamida, J., 2010. Classification of EC 3.1.1.3 bacterial true lipases using phylogenetic analysis. Afr. J. Biotechnol. 9, 8243–8247. https://doi.org/10.5897/ajb10.721.

Messaoudi, A., Belguith, H., Ben Hamida, J., 2011. Three-dimensional structure of *Arabidopsis thaliana* lipase predicted by homology modeling method. Evol. Bioinf. 99–105. https://doi.org/10.4137/EBO.S7122.

Moya-Salazar, J., Vértiz-Osores, J., Jibaja, S., Rupa, R.A.-E.R., Alarcón-Díaz, M., Vega-Guevara, M., Flores, R.C., 2019. Fungi lipases homology modeling and molecular docking with fatty acids and tripalmitin of palm oil effluent. Arch. Org. Inorg. Chem. Sci. 4 https://doi.org/10.32474/AOICS.2019.04.000183.

Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M., Harel, M., Remington, S.J., Silman, I., Schrag, J., Sussman, J.L., Verschueren, K.H.G., Goldman, A., 1990. The α/β hydrolase fold. Protein Eng. 5, 197–211. https://doi.org/10.1093/protein/5.3.197.

Satpathy, R., Konkimalla, V.B., Ratha, J., 2016. In Silico phylogenetic analysis and molecular modelling study of 2-haloalkanoic acid dehalogenase enzymes from bacterial and fungal origin. Adv. Bioinform. https://doi.org/10.1155/2016/8701201.

Sahoo, R.K., Kumar, M., Sukla, L.B., Subudhi, E., 2017. Bioprospecting hot spring metagenome: lipase for the production of biodiesel. Environ. Sci. Pollut. Res. https://doi.org/10.1007/s11356-016-8118-7.

Sharma, P.K., Kumar, Rajender, Kumar, Rakesh, Mohammad, O., Singh, R., Kaur, J., 2012. Engineering of a metagenome derived lipase toward thermal tolerance: effect of asparagine to lysine mutation on the protein surface. Gene. https://doi.org/10.1016/j.gene.2011.09.028.

Tamboli, A.S., Rane, N.R., Patil, S.M., Biradar, S.P., Pawar, P.K., Govindwar, S.P., 2015. Physicochemical characterization, structural analysis and homology modelling of bacterial and fungal laccases using in silico methods. Netw. Model. Anal. Health Inf. Bioinf. J. 4, 17. https://doi.org/10.1007/s13721-015-0089-y.

Villeneuve, P., Muderhwa, J.M., Graille, J., Haas, M.J., 2000. Customizing lipases for biocatalysis: a survey of chemical, physical and molecular biological approaches. J. Mol. Catal. B Enzym. 9, 113–148. https://doi.org/10.1016/S1381-1177(99)00107-1.

Wilkins, M.R., Gasteiger, E., Bairoch, A., Sanchez, J.C., Williams, K.L., Appel, R.D., Hochstrasser, D.F., 1999. Protein identification and analysis tools in the ExPASy server. Methods Mol. Biol. 112, 531–552. https://doi.org/10.1385/1-59259-890-0:571.