Structural Analysis and Physico-chemical Properties of Protease from Streptomyces Canus using Computational Approach

Shivam Sharma[1], Apoorva Sharma[2], Neetu Jabalia[*3] and Nidhee Chaudhary[4]

[1,2,3,4]Amity Institute of Biotechnology Amity University Uttar Pradesh, Sec-125, Noida
E-mail: [1]shivamsharma823@gmail.com, [2]apoorva.sharma216@gmail.com,
[3]njabalia@amity.edu, [4]apoorva.sharma216@gmail.com

Abstract—*Streptomyces canus are Gram-positive bacteria belonging to family Streptomycetaceae. It has major medicinal utilities and can be used for producing anti-bacterial, anti-parasitic, anti-fungal drugs and immunosuppressants. In recent years, Streptomyces has come up as a use for retrieval of heterogeneous expression of protein, in biotechnology. This study involved computational approach for the determination of structural properties of the protease. The physcio-chemical properties prediction was done by using Protparam. The secondary structure prediction has been carried with GOR4 prediction tool, predicting the alpha-beta helices of the protease sequences. 3-D structure of the protease has been predicted using homology modeling methods. The structural studies will provide a better insight to the functional analysis.*

Keywords*: Streptomyces canus, protease , homology modeling, GOR.*

1. INTRODUCTION

Protease occurs naturally in all organisms and is an essential constituent for all the existing live forms. They act as an important industrial enzyme occupying for about 60% of total enzyme market. Proteases are group of enzymes which catalyse the cleavage of peptide bonds [1]. They are also generally referred to proteolytic enzymes or proteinases. They are widely distributed in plants animals and microorganisms. Many proteases have been isolated from latex , fruits and seeds. They have importance in both commercial and physiological fields. Protease is an enzyme which is widely used in detergents, leather, waste management and silver recovery [2] The genera *Streptomyces canus* belonging to class actinobacteria are colonial and aquatic organisms [3]. Actinobacteria being the most prepotent bacterial phylum, comprise Streptomyces which is one of the largest bacterial genera [4]. Some actinobacterias and Streptomyces play an important role in the biotic buffering of soils [5]. Actinobacteria, particularly Streptomyces species are producers of metabolites that have various medicinal uses, such as antivirals, immonomodifiers, antifungal [4] enzyme inhibitors and antibacterials [6]. The species being utilitarian in agriculture also, as it instigates the production of insecticides, fungicides and promoters of growth for animals and plants [7] The aim of this study is to characterize the sequence of protease from *Streptomyces canus* by using *in-silico* techniques which will be valuable to understand the structural features and will raise the prospects of its academic or commercial use.

2. MATERIALS AND METHODS

Sequence analysis- The protease sequence from *Streptomyces canus* was retrieved from the NCBI [8]. The retrieved sequence was in FASTA format and used for further analysis.

Functional characterization- The presence of motifs in the protease sequence were identified by using MotifScan [9]. The subcellular location of protease was done by using TargetP [10].

Physico-chemical Properties- For physico-chemical characterization of protease sequence from *Streptomyces canus*, isoelectric point (pI), molecular weight, total number of positive and negative residues, extinction coefficient[11], instability index[12], aliphatic index[13] and grand average hydropathicity were computed using Expasy ProtParam server [14].

Secondary structure prediction- The secondary structure analysis of protease sequence was based only on knowledge of its primary structure. The secondary structure feature of the protease sequence from *Streptomyces canus* was identified by GOR IV.

Tertiary structure prediction-In order to generate 3-D model, homology modeling method was applied. Swiss-Modeler program was used to execute the modeling of 3D structure of the sequence.

Structure Validation and Evaluation- The modeled structure was further validated by Procheck [15]. Energy

minimization of the protease structure from *Streptomyces canus* was done by Swiss PDB Viewer. Rampage server was used to find the accuracy of generated 3-D model [16].

3. RESULTS AND DISCUSSIONS

Table 1 shows the amino acid composition in the protease sequence from *Streptomyces canus.*

Table 1: Amino acid composition of protease from *Streptomyces canus.*

Name of the amino acid	No. of amino acids	%Composition
Ala	61	13.6
Arg	44	9.8
Asn	6	1.3
Asp	24	5.4
Cys	4	0.9
Gln	6	1.3
Glu	20	4.5
His	6	1.3
Ile	11	2.5
Leu	28	6.3
Lys	5	1.1
Met	6	1.3
Phe	11	2.5
Pro	20	4.5
Ser	34	7.6
Thr	40	8.9
Trp	6	1.3
Tyr	5	1.1
Val	40	8.9
Pyl	0	0
Sec	0	0

A set of conserved amino acid residues located in the vicinity that provide clues to the functions is termed as motif. Motif prediction was done by using MotifScan (Table 2).

Table 2: Motifs determined from the protease sequence of *Streptomyces canus*

Accession no.	Motif	Description	Start	End
WP_059296777	PKC_PHOSPHO_SITE	Protein kinase C phosphorylation	63	65
			96	98
			113	115
			171	173
			365	367
	RGD	Cell attachment sequence	137	139
			219	221
	TYTR_PHOSPHO_SITEASN_GLYCOSYLATION	Tyrosine kinase phosphorylation site	430	438
	CK2_PHOSPHO_SITE	Casein kinase II phosphorylation site	103	106
			143	146
			238	241
			291	294
			431	434
	MYRISTYL	N-myristoylation site	20	25
			62	67
			122	127
			144	149
			152	157
			168	173
			225	230
			261	266
			280	285
			302	307
			327	332
			336	341
			369	374
			393	398
			418	423
	ASN_GLYCOSYLATION	N-glycosylation site	192	195

Subcellular location of protease was from *Streptomyces canus* identified by TargetP web server (Table 3).

Table 3: Subcellular location of protease sequence from *Streptomyces canus*

Name	WP_059296777
Len	447
mTP	0.810
SP	0.228
Other	0.023
Loc	M
RC	3

Isoelectric point (pI) is the pH at which the surface of protein is having a charge but the total charge of protein is 0. At pI proteins are stable and compact. The value of isoelectric point (pI) of the protein from *Streptomyces canus* was 9.22 which indicates that it is alkaline. The computed isoelectric point will be useful for developing buffer system when the enzyme is to be purified in solution by isoelectric focusing method. The molecular weight determined was 45558.8 . The extinction coefficient of protease was determined by the Expasy's Protparam. Extinction coefficient of protein at 280nm is $40700M^{-1}cm^{-1}$, assuming all pairs of Cys residues form cystines and extinction coefficient came out to be 40450 when all pairs of Cys residues are reduced. The computed protease concentrations and extinction coefficient will be useful in the quantitative study of various interactions in the solution. The estimated half life of the protease is 30 hours in mammalian reticulocytes(in vitro), >20 hours in yeast(in vivo), >10 hours in *E.coli (*in vivo). The instability index (I) provides an estimate of the stability of protease in a test tube. A protease sequence whose instability index is smaller than 40 is predicted as stable while the value above 40 predicts that the protein is unstable. The instability index (I) is computed to be 34.37. This classifies the protein as stable. The total number of negatively charged residues (Asp+Glu) is 44 and the total number of positively charged residues (Arg+Lys) is also 49. The aliphatic index (AI), defined as the relative volume of a protein occupied by aliphatic side chains, is regarded as-- a

positive factor for that leads to the increase of thermal stability of globular proteins. Aliphatic index for the protease sequence from *Streptomyces canus* was 73.62. The high aliphatic index of the sequence indicated that the protease will be stable over a wide temperature range. The Grand Average hydropathy (GRAVY) value for a protein is calculated as the total of hydropathy values of all the amino acids, and divided by the total number of residues present in the protease sequence. GRAVY index for this sequence is -0.165. This low range of value indicated better interaction with water [17] (Table 4).

Table 4: Physico-chemical properties of protease sequence from *Streptomyces canus*

Properties	**WP_059296777**
No. of amino acids	447
Molecular Weight	45558.8
Pi	9.22
Positively charged residues	49
Negatively charged residues	44
Extension coefficients	40700
Instability index	34.37
Aliphatic index	73.62
GRAVY	-0.165

Secondary structural features were predicted by GORIV. The conformational entropy associated with random coil significantly leads to stabilization and protein folding. Proline, which has a high content in the protein, has special property that creates kinks in polypeptide chains and disrupts ordered secondary structure and might have contributed to the high content of random coil structure (Fig. 1).

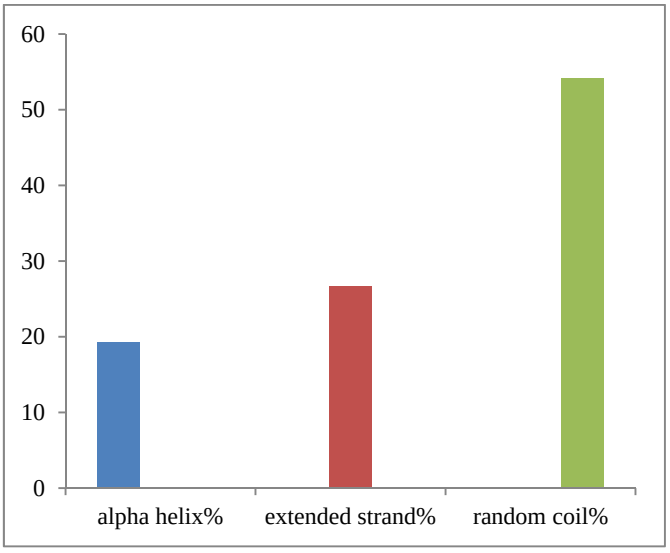

Fig. 1: Graphical representation of secondary structure composition of protease from Streptomyces canus

After selection of potential template for the enzymes, 3-D models were generated by the use of Swiss-modeler program (Table 5).

Table 5: 3-D model generated from protease sequence of *Streptomyces canus*

Accession no.	**WP_059296777**
Modelled Residue range	174-362
Template	2sfa.1.A
Sequence identity(%)	50.27
Total energy	-5833.330 KJ/mole

Fig. 2 showed the generated 3-D structure of protease sequence from accession number "WP_059296777" by using Swissmodellar.

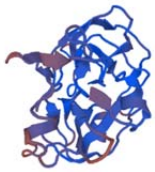

Fig. 2: Ribbon representation of protease from Streptomyces canus

Validation of the three dimensional protein model is the final step of protein modeling. It evaluates the significance and accuracy by using Ramachandran plot in Procheck web-based server Fig. 3.

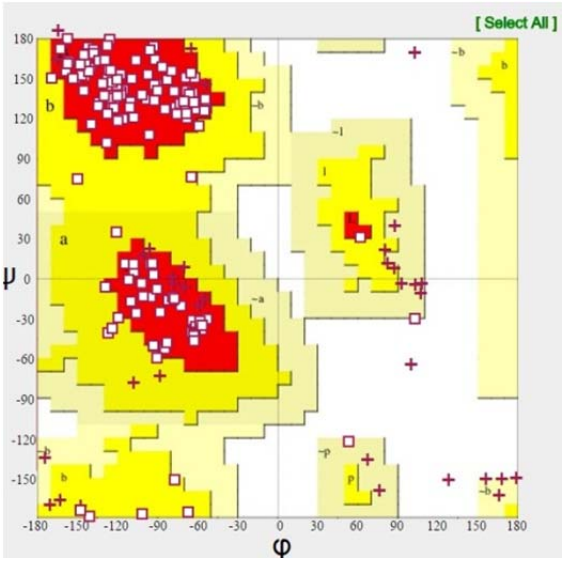

Fig. 3: Ramachandan Plot of protease sequence from *Streptomyces canus*

Model evaluation of 3-D modeled structure of protease from *Streptomyces canus* was by Rampage program. It revealed 95.2 % residues were in favored regions.

Table 6: Rampage evaluation results of protease showing the accuracy of our model.

S.no.	Acc.no.	Number of Residues		
		Favored region	Allowed Region	Outlier Reg.
1.	WP_059296777	~98% expected : 178 (95.2%)	(~ 2.0 % expected) : 7(3.7%)	:2(1.1%)

4. CONCLUSION

In our study, we have characterized amino acid sequence of protease present in *Streptomyces canus* to acquire an understanding about their functional properties, physico-chemical properties and various protein structure levels by using *in-silico* techniques. The primary structure prediction reveals that it is hydrophilic and stable for wide range of temperature. Secondary structure analysis established that in most of the sequences random coils dominated among secondary structure elements which inherent in the local interactions followed by alpha helix, extended strand and beta turns. In this study, structure analysis will facilitate knowledge about functional aspects of protease in *Streptomyces canus* which will further aid in formulating their uses in academics and industries.

REFERENCES

[1] Neetu Jabalia and Nidhee chaudhary, “Heat Resistant Proteases Isolated from Medicinally Significant Spices: Partial Purification and Biochemical Characterization GSTF Journal of BioSciences”, *J Bio*, 3,2, August 2015.

[2] Sumantha, A., Sandhya, C., Szakacs, G., Soccol, C. R. and Pandey, A., “Production and partial purification of a neutral metalloprotease by fungal mixed substrate fermentation.”, Food *Technology and Biotechnology*, 2005, pp. 313-319.

[3] Servin JA, Herbold CW, Skophammer RG, Lake JA (January 2008). "Evidence excluding the root of the tree of life from the actinobacteria". Mol. Biol. Evol. 25 (1): 1–4. doi:10.1093/molbev/msm249. PMID 18003601.

[2] C.Michael Hogan. 2010. Bacteria. Encyclopedia of Earth. eds. Sidney Draggan and C.J.Cleveland, National Council for Science and the Environment, Washington DC.

[3] Ningthoujam, Debananda S.; Sanasam, Suchitra; Tamreihao, K; Nimaichand, Salam (November 2009). "Antagonistic activities of local actinomycete isolates against rice fungal pathogens". African Journal of Microbiology Research. 3 (11): 737–742.

[4] Gupte, M.; Kulkarni, P.; Ganguli, B.N. (2002). "Antifungal Antibiotics". Appl. Microbiol. Biotechnol. 58: 46–57.

[5] Mahajan, GB (2012). "Antibacterial agents from actinomycetes - a review". Frontiers in Bioscience. 4: 240–53. doi:10.2741/e373.

[6] Bressan, W (2003). "Biological control of maize seed pathogenic fungi by use of actinomycetes". Biocontrol. 48 (2): 233–240. doi:10.1023/a:1022673226324.

[7] Atta, M.A (2009). "Antimycin-A Antibiotic Biosynthesis Produced by Streptomyces Sp. AZ-AR-262: Taxonomy, Fermentation, Purification and Biological Activities". Austral. J. Basic and Appl. Sci. 3: 126–135.

[8] National Center for Biotechnology Information, U.S. National Library of Medicine 8600 Rockville Pike, Bethesda MD, 20894 USA

[9] Sigrist CJ, Cerutti L, de Castro E, Langendijk-Genevaux PS, Bulliard V, Bairoch A, Hulo N. PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Res.* 2010; **38**(Database issue):D161-6.

[10] Olof Emanuelsson, Soren Brunak, Gunnar von Heijne, Henrik Neilsen; Nature protocols, 2, 2007, pp. 953-971.

[11] Gill, S., C. and Von, Hippel, P. H.,. (1989), “ Calculation of protein extinction coefficients from amino acid sequence data” *Anal. Biochem*,1989.

[12] Ilya A. Vakser, ProteinProtein Docking: From Interaction to Interactome, Biophys J. 2014 Oct 21; 107(8): 1785–1793. doi: 10.1016/j.bpj.2014.08.033.

[13] Ikai, A.J. (1980). Thermostability and aliphatic index of globular proteins J. Biochem. 88,18951898, PMID:7462208 [PubMed indexed for MEDLINE].

[14] Gasteiger, E. and Walker, J.M. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In The Proteomics Protocols Handbook (ed. Walker, J.M.), Humana Press, New Jersey, USA, pp 571-607.

[15] https://services.mbi.ucla.edu

[16] Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl*. Crystallogr. *26*, 283-291.

[17] *Neetu Jabalia[1] , Hina Bansal[1], P.C.Mishra[2], Nidhee Chaudhary, In silico* studies on cysteine protease from *Zingiber office inale* [jpp].