International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 5 Issue: 5 1220 - 1226

Computational based Structural, Functional and Phylogenetic Analyses Of 3-Deoxy-D-Arabino-Heptulosonate 7-Phosphate (DAHP) Synthase of Corynebacterium Spp

Krishnendu Pramanik[1], Indrani Ghosh[1], Anumita Sarkar[1,2], Tushar Kanti Maiti[1]*
[1]Microbiology Laboratory, Department of Botany, The University of Burdwan, 713104, W.B., India
[2]Government General Degree College, Singur, 712409, W.B., India
*Corresponding author: tkmbu@yahoo.co.in

Abstract—The pathway of aromatic amino acid biosynthesis was initiated by the enzyme DAHP synthase. In this branched biosynthetic pathway phenylalanine, tyrosine and tryptophan are produced. This enzyme was repressed and inhibited by three said end product of aromatic amino acids. The importance of this enzyme is microbial fermentation of aromatic amino acid production. The present study envisaged the *in silico* analysis of this protein (DAHP synthase). The structural, functional and phylogenetic studies of this protein of *Corynebacteriumglutamicum* have been determined and it was observed that the protein of interest is a thermostable, acidic protein having molecular weight ranging in between 38.67 to 52.32 kDa.

Keywords- *Corynebacterium; DAHP synthase; in silico study; aromatic amino acid*

*****

I. INTRODUCTION

The production of aromatic amino acids viz.phenylalanine, tyrosine, and tryptophan takes place by a common pathway [1]using DAHP synthase. At first in this pathway two products of carbohydrate metabolism i.e. phosphor-enol-pyruvate and erythrose 4-phosphate are condensed which gives rise to DAHP (3-deoxy-D-arabino-heptulosonate-7-phosphate) using the enzyme DAHP synthase and cyclized.Then it undergoes a number of reactions to produce Chorismate through shikimate pathway. In next phase from Chorismate the aromatic amino acids are produced through a series of diverging individual pathways [2].Moreover, this is an important pathway of plant where shikimate is produced and which is precursor of secondary metabolite of plants. These 3 amino acids act as the precursor of wide range of secondary metabolites which are very important for plant growth and human health.

The phenylalanine and tryptophan are essential amino acids for human. The demand of essential amino acid production is gradually increased day by day; particularly demand of phenylalanine production is increased due to production of low-calorie "aspartame"- an artificial sweetener and dipeptide of aspartic acid phenylalanine[3]. It is 200 times sweeter than Sucrose [4] and used for diabetic patient. Microbial production of amino acid is now common practice as it is cost effective and produced always in L-form is needed for human being. The aromatic amino acid biosynthesis is repressed and inhibited by the three amino acids. So, it is necessary to know the structural, functional insights of the DAHP synthase for overproduction of amino acid.

There are different roles of different amino acids. Particularly Tryptophan contributes to the nutritional properties of many foods (plant-based) along with some other amino acids like methionine lysine and threonine [5].Aromatic amino acids and there regulation are greatly used in bacteria because of their use in food and also in drug industries. In bacteria aromatic amino acids are the source of protein. The aromatic amino acids are also industrially used for the production aspartame and for the production of monoamine neurotransmitters serotonin, norepinephrine, epinephrine, and dopamine (anti-Parkinson's disease drug named L-dopa), in the peripheral and central nervous systems of many mammalian species [5, 6].

C. glutamicum are a group of bacteria which are characterised mainly by the production of glutamic acid. But last two decades are using to produce various types of amino acids.*C. glutamicum*, is a non-pathogenic bacterium. It is Gram-positive soil bacterium with its non-sporulating nature. They are facultative anaerobic. This bacterial species has great importance because of mainly two reasons. First, itis used for large scale production of various amino acids including aromatic amino acids, predominantly L-glutamate (produce 2.1 million tons/year) and L-lysine (produce 1.1 million tons/year) [7]. Second, it serves as a model for the Corynebacterineae, which is under the Actinomycetales order that includes many important pathogens (particularly *Mycobacterium tuberculosis*).

In the present study, attempts have made to characterize the *Corynebacterium* DAHP synthase protein by computational based tools. Apart from structural and functional studies, phylogenetic studies have also been approached to determine the evolutionary relatedness among the different species of *Corynebacterium* as well as different strains of *C. glutamicum* in terms of their protein and related cDNA.

II. MATERIALS AND METHODS

A. Sequence retrieval and cDNA acquisition

There were16 DAHP synthase amino acid sequences of different *Corynebacterium* species and 28 amino acid sequences of different strains of *C. glutamicum* taken in fasta format from NCBI database (https://www.ncbi.nlm.nih.gov/). All the sequences were reverse translated by reverse translate tool (www.expasy.org/tools) to acquire cDNA.These

sequences were used for further *in silico* studies these sequences were used.

B. Phylogenetic analysis

From the retrieved amino acid sequences of DAHP synthase and the corresponding reverse translates total four phylogenetic trees were produced. Here two trees of protein sequences for different species of *Corynebacterium* and two for cDNA of different strains of *C. glutamicum.* These phylogenetic trees were constructed using MEGA 7 (version 7.0.18) software [8] for comparing the evolutionary relatedness of the taxa. In both cases the phylogenetic tree was made by using Minimum Evolution method.In case of all four trees the evolutionary history was inferred using the Minimum Evolution method [9]. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analysed[10].

C. Physiochemical characterisation

To determine the physiochemical properties of the recovered sequences a web based server was used. The amino acid residue (AAR), Isoelectric point (PI) molecular weight (MW), instability indices (II),aliphatic indices (AI), Extinction coefficient (EC), and grand average of hydropathicity (GRAVY) werederived from Expasy ProtParam Tool [11]. Isoelectric point referred to the pH at which net charge is zero.Extinction coefficient is the quantitative study of protein-ligand and protein-protein interaction. Instability Index means the stability of proteins. Aliphatic Indices referred to the relative volume of protein occupied by aliphatic side chain. GRAVY can be defined as the sum of all hydropathy values of all amino acids divided by number of residues in a sequence [12, 13].

D. Secondary structure prediction

To predict the secondary structure and to count the number of secondary components viz. helices sheets and turns, three servers are used named – CFSSP: Chou and Fasman secondary structure prediction server (www.biogem.org/tool/chou-fasman)[14, 15] PSIPRED v3.3 (http://bioinf.cs.ucl.ac.uk/psipred/).

E. Homology modelling and evaluation

*C. glutamicum*SCgG2 was chosen as the representative of the 28 strains of *C. glutamicum*. It was used to predict the tertiary structure of DAHP synthase protein.Here the online EXPASY Swiss Model (ProMod3 Version 1.0.2.) workspace[16] was used to get the 3D models of *C.glutamicum* SCgG2 (**AGN21760.1**).Here the Q-MEAN score was obtained from Swiss Model and the build models were generated in SAVES Server (http://services.mbi.ucla.edu/SAVES/).It gives the idea about the quality of built model. Ramachandran plot was obtained using the pdb file in SAVES Server (http://services.mbi.ucla.edu/SAVES/Ramachandran/)to find the energetically allowed regions related to backbone dihedral angles of amino acid residues in DAHP synthase protein structure.

F. Functional analysis

To distinguish the conserved domains of the DAHP synthase protein, NCBI database (http://www.ncbi.nlm.nih.gov/cdd/) was mainly observed. However, the result acquired from NCBI database was also supported by some additional tools named motif search (http://www.genome.jp/tools/motif/), and InterProScan (http://www.ebi.ac.uk/interpro/). Functionally interacting partners of DAHP synthase protein were predicted by STRING (http://string-db.org/) analysis(Version 10.0).

III. RESULTS AND DISCUSSIONS

A. Sequence retrieval

Retrieval of amino acid sequences from various authentic databases e.g. UniProtKB, NCBI etc. for Bioinformatics study was practiced previously by many authors. Verma et al. [17] reported computational based structure-function study on *Bacillus*phytases while Pramanik et al [12, 13] worked on *Mesorhizobium* ACC deaminase and *Pseudomonas* alkaline phosphatse emphasizing the structure-function as well as phylogenetic studies using several bio-computational tools.

B. Phylogenetic analysis

In case of the phylogenetic tree of DAHP synthase protein of different species of *Corynebacterium* it was found that the *C. glutamicum* SCgG2 (accession no AGN21760.1) clustered with *C.casei* LMG S-19264(AHI20479.1) having the similarity 100%. In case of the tree on reverse translate of DAHP synthase protein of different species of *Corynebacterium* a cluster was formed between *C. glutamicum* SCgG2 (protein accession no. AGN21760.1) and *C.casei* LMG S-19264 (protein accession no. AHI20479.1) and showed 100% similarity (Fig. 1, Fig. 2).

However, the phylogram showing DAHP synthase protein of different strains of *C. glutamicum* the *C. glutamicum* SCgG2 (AGN21760.1) cluster with *C. glutamicum* SCgG1(AGN18737.1), *C. glutamicum* S9114 (EGV40089.1), *C. glutamicum* ZL-6 (ANR65063.1) and*C. glutamicum*Z188 (EPP41090.1) showed 64% similarity.In case of tree showing reverse translate of DAHP synthase protein of different strains of *C. glutamicum*, the *C. glutamicum* SCgG2 (protein accession no. AGN21760.1) cluster with *C. glutamicum* S9114 (protein accession no.EGV40089.1), *C. glutamicum* Z188 (protein accession no.EPP41090.1), *C. glutamicum* ZL-6(protein accession no.ANR65063.1), *C. glutamicum* SCgG1 (protein accession no.AGN18737.1), *C. glutamicum* ATCC 13032 (protein accession no.BAB98383.1), *C. glutamicum* MB001 (protein accession no. AGT04988.1) and showed 100% similarity. Phylogenetic comparison of protein sequence and their corresponding cDNA was studied previously [12, 13] to find if there exists any correlationbetween the said two parameters(Fig. 3, Fig. 4).

ISSN: 2321-8169
1220 - 1226

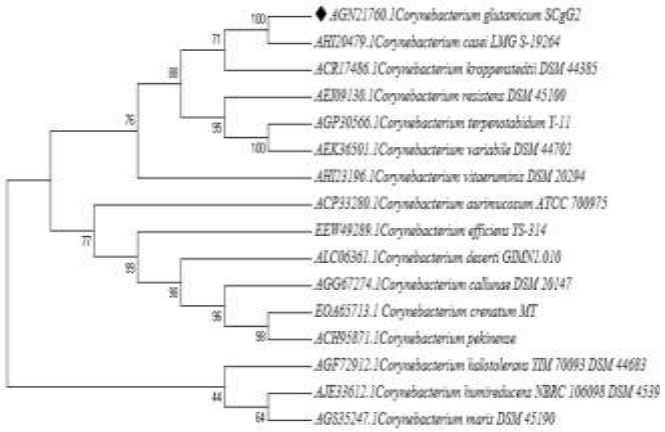

Figure 1. Phylogram of DAHP synthase protein of 16 different *Corynebacterium* sp. using Equal input method of Minimum Evolution method.

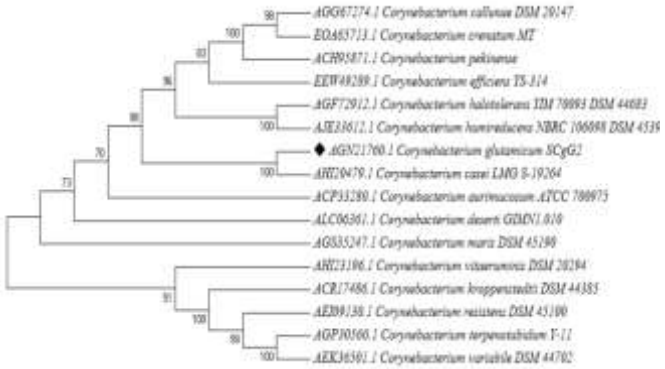

Figure 2. Phylogram of reverse translate of DAHP synthase protein of 16 different *Corynebacterium* sp. using Kimura 2-parameter method of Minimum Evolution method.

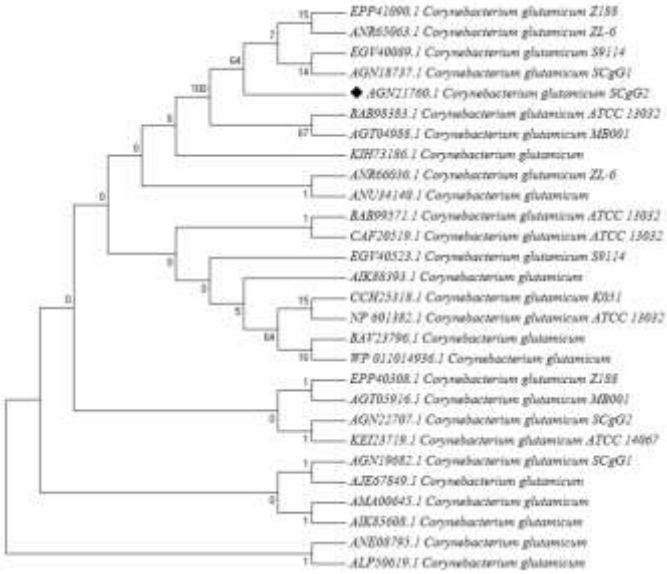

Figure 3. Phylogram of DAHP synthase protein of 28 different strains of *Corynebacteriumglutamicum* using Poisson correction method of minimum evolution method.

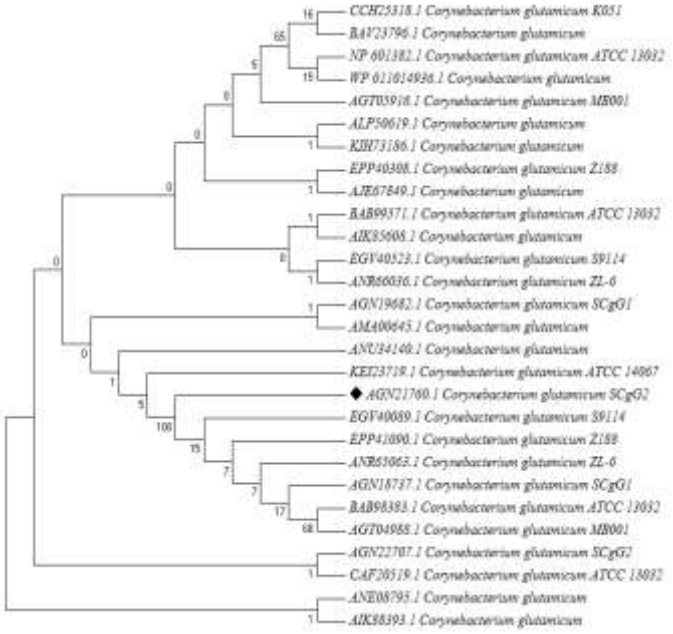

Figure 4. Phylogram of reverse translate of DAHP synthase protein of 28 different strains of *Corynebacteriumglutamicum* using Kimura 2-parameter method of Minimum Evolution method.

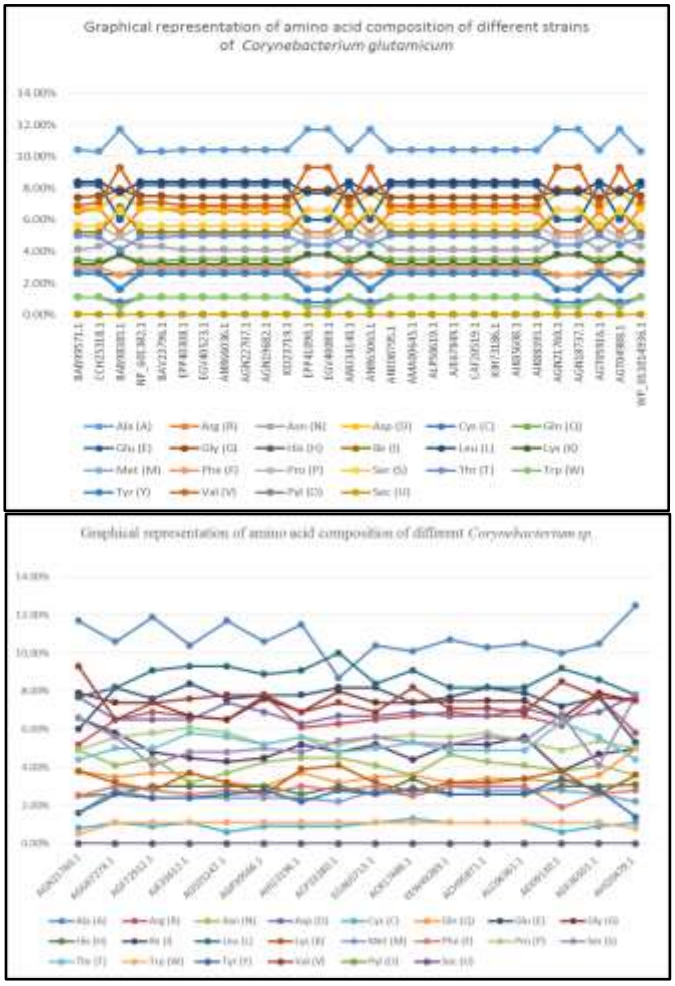

Figure 5. Graphical representation of amino acid composition.

C. Physiochemical characterization

In case of different species (Table 1), amino acid residues calculated ranging from 360 to 475. Isoelectric point of them rangein between 4.93 to 5.29. It indicated that all of the proteins are acidic in nature. Molecular weight of the proteins wasin between 38.67 to 52.32 kDa. Nine proteins have instability indices above 40, so these are unstable in nature. Others are stable. The aliphatic indices are above 82.5. Lower range of GRAVY indicates the better interaction of protein with water.

Table 1. Physicochemical properties of different species of *Corynebacterium*.

DAHP SYNTHASE PROTEIN ACC. NO.	AAR	PI	MW (KDa)	II	AI	EC	GRAVY
AGN21760.1	**366**	**4.93**	**39.1**	**34.59**	**94.10**	**20065**	**-0.157**
AGG67274.1	462	4.97	50.88	42.54	84.31	45630	-0.321
AGF72912.1	462	5.16	50.7	41.67	87.27	44140	-O.301
AJE33612.1	462	5.15	51.04	45.13	86.39	44140	-0.343
AGS35247.1	462	4.99	50.80	37.35	87.47	45505	-0.295
AGP30566.1	462	5.18	51.00	40.04	85.54	47120	-0.358
AHI23196.1	462	5.13	50.68	40.29	87.27	42650	-0.293
ACP33280.1	462	5.14	51.13	37.80	87.40	47120	-0.374
EOA65713.1	462	4.93	50.90	39.63	83.66	45630	-0.335
ACR17486.1	475	5.11	52.32	45.15	86.46	48735	-0.296
EEW49289.1	466	5.09	51.31	40.13	83.15	45630	-0.336
ACH95871.1	466	4.97	51.31	38.84	82.73	45630	-0.353
ALC06361.1	466	4.97	51.34	38.89	83.99	45630	-0.338
AEI09130.1	469	5.27	51.62	40.47	85.48	51465	-0.331
AEK36501.1	466	5.18	51.47	38.62	84.81	47120	-0.366
AHI20479.1	360	5.29	38.67	47.28	84.08	24200	-0.280

In case of different strains (Table 2), amino acid residues calculated ranging from 366 to 466. Isoelectric point of them ranging between 4.93 to 5.02.Its indicates that the proteins are of acidic in nature. Molecular weight of the proteins was between 39.1 to 51.3 kDa. All of the proteins having instability indices under 40, which indicates that the proteins are stable in nature. The aliphatic indices of all proteins are significantly higher which may take as a positive factor for increased thermostability of globular protein. Extinction Coefficients are taken assuming all pairs of Cys residues form Cysteines .Extinction Coefficient mean the amount of light absorbed by protein at a particular wavelength. Lower range of GRAVY indicates the better interaction of protein with water.

Table 2. Physicochemical properties of different strains of *Corynebacteriumglutamicum*.

DAHP SYNTHASE PROTEIN ACC. NO.	AAR	PI	MW (KDa)	II	AI	EC	GRAVY
BAB99571.1	462	4.93	50.9	39.63	83.66	45630	-0.335
CCH25318.1	466	4.97	51.3	38.25	83.56	45630	-0.341
BAB98383.1	366	4.93	39.1	34.36	95.16	20065	-0.143
NP_601382.1	466	4.97	51.3	38.25	83.56	45630	-0.341
BAV23796.1	466	4.97	51.3	38.25	83.56	45630	-0.341
EPP40308.1	392	5.02	42.9	37.57	82.45	34630	-0.292
EGV40523.1	462	4.93	50.9	39.63	83.66	45630	-0.335
ANR66036.1	462	4.93	50.9	39.63	83.66	45630	-0.335
AGN22707.1	462	4.93	50.9	39.63	83.66	45630	-0.335
AGN19682.1	462	4.93	50.9	39.63	83.66	45630	-0.335
KEI23719.1	462	4.93	50.9	39.63	83.66	45630	-0.335
EPP41090.1	366	4.93	39.1	34.59	94.10	20065	-0.157
EGV40089.1	366	4.93	39.1	34.59	94.10	20065	-0.157
ANU34140.1	462	4.93	50.9	39.63	83.66	45630	-0.335
ANR65063.1	366	4.93	39.1	34.59	94.10	20065	-0.157
ANE08795.1	462	4.93	50.9	39.63	83.66	45630	-0.335
AMA00645.1	462	4.93	50.9	39.63	83.66	45630	-0.335
ALP50619.1	462	4.93	50.9	39.63	83.66	45630	-0.335
AJE67849.1	462	4.93	50.9	39.63	83.66	45630	-0.335
CAF20519.1	462	4.93	50.9	39.63	83.66	45630	-0.335
KIH73186.1	462	4.93	50.9	39.63	83.66	45630	-0.335
AIK85608.1	462	4.93	50.9	39.63	83.66	45630	-0.335
AIK88393.1	462	4.93	50.9	39.63	83.66	45630	-0.335
AGN21760.1	**366**	**4.93**	**39.1**	**34.59**	**94.10**	**20065**	**-0.157**
AGN18737.1	366	4.93	39.1	34.59	94.10	20065	-0.157
AGT05916.1	462	4.93	50.9	39.63	83.66	45630	-0.335
AGT04988.1	366	4.93	39.1	34.36	95.16	20065	-0.143
WP_011014936.1	466	4.97	51.3	38.25	83.56	45630	-0.341

Note. AAR= The amino acid compositions, PI= Isoelectric point, MW= Molecular weight, II= Instability indices, AI= Aliphatic indices, EC= Extinction coefficient and GRAVY= Grand average of hydropathicity.

D. Secondary structure analysis

From the web server it was found that the percentage of three classes of secondary arrangements was: Helices: 72.4%; sheets: 54.4%; and turns: 14.5%. It indicated helices>sheets>turns in the protein. Secondary arrangements (Fig. 6) indicate that the protein is not unfolded indicating stability of protein. There is no presence of disordered protein binding site.Moreover, it has found that a larger fraction of the residues of thermophiles are in the α-helical conformation which helps to withstand the high temperature [12]. Here α-helical conformation is also higher which indicate the thermostable nature of the protein. Similar practices of prediction of secondary structural elements was reported by some earlier workers [12, 13, 17].

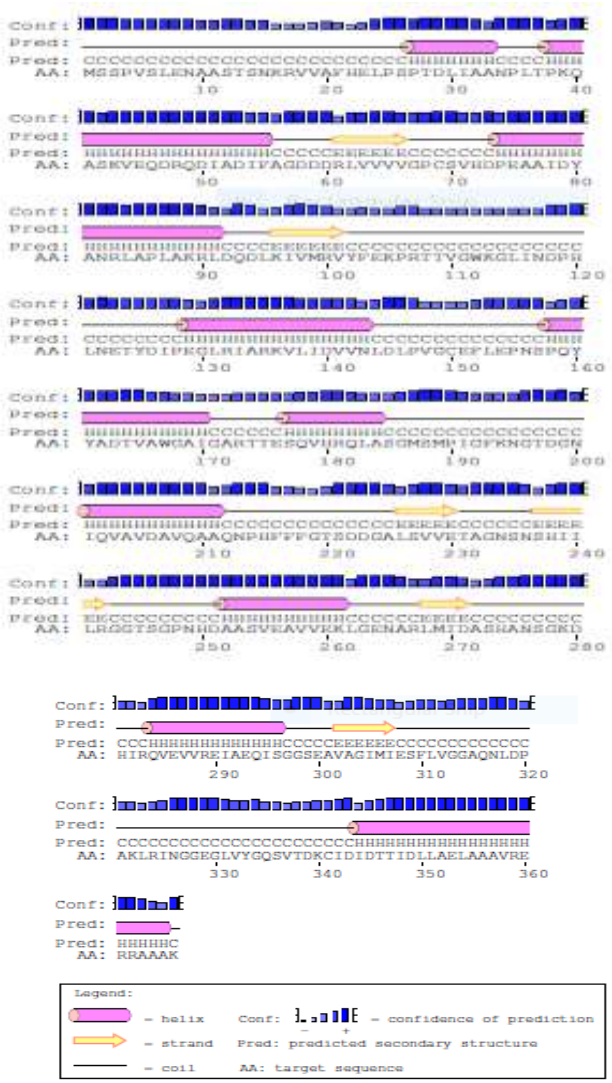

Figure 6. Predicted secondary structure of selected protein.

E. Homology modeling, evaluation and submision

The protein models derived from SWISS MODEL (Fig. 7) workspace were evaluated by Q-MEAN and SAVES SERVER (Fig. 7).The Z-score acquired from both the servers were compared. Z-score is the absolute quality of a model which is related to the reference structures solved by X-ray Crystallography obtained from Q-MEAN Z-score [18].

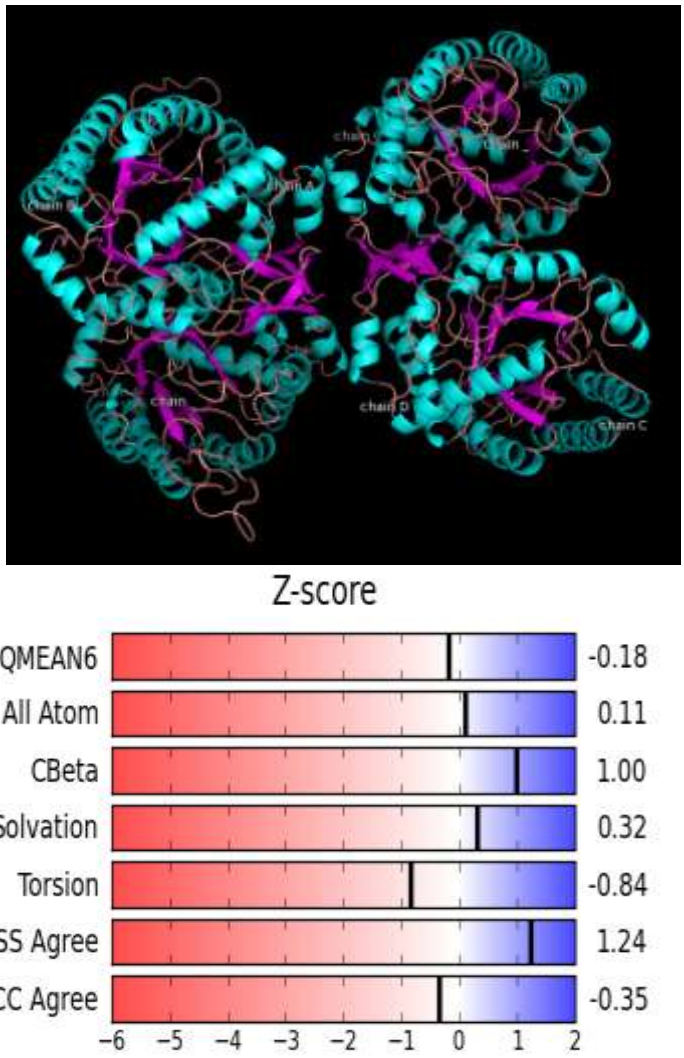

Figure 7. Tertiary protein model with Z-score analysis result

Again, Z-score or standard score is a score which is normalized to mean 0 and standard deviation 1 [18]. Here, the Z-score mean and standard deviation which were obtained from SAVES SERVER (Fig. 8) were 0.536 and 1.416 .From the Ramachandran plot (Fig. 9) it was found that 88.6% residues are in most favored regions, 10.4% residues were in additional allowed regions and 0.7% residues in generously allowed regions and in disallowed region there are 0.3% residues. All of these values are somewhat close to the expected statistics [19].

F. Functional analysis

DAHP synthase catalyzes the condensation of phosphoenolpyruvate (PEP) and erythrose 4-phosphate (E4P) to form DAHP. It is the first step of shikimate pathway which leads to the biosynthesis of aromatic amino acids, vitamin K, folic acid and ubiquinone. The commercial use of *C. glutamicum* is the aromatic amino acid production [20]. Mainly one type of conserved domain was found for DAHP synthase protein of *C. glutamicum* SCgG2 (AGN21760.1) which belongs to DAHP_synth_1 super family. Its accession number is cl17225.

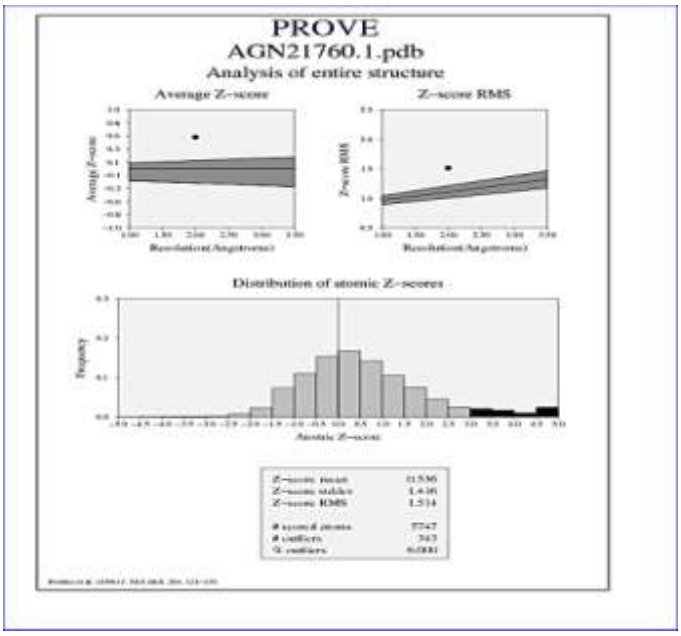

Figure 8.Z-score of the built model from SAVES server.

From MOTIF FINDER (Fig. 10) of DAHP synthase protein of *C. glutamicum* SCgG2 (AGN21760.1) it was found (Fig.10) that there were two motifs viz. DAHP_synth_1 (PF00793, DAHP synthetase I family; Independent E-value- 9.1e-103) and Glyco_transf_56 (PF07429, 4-alpha-L-fucosyltransferase glycosyltransferase group 56; Independent E-value- 0.026).

From STRING server it was found that the query proteins and first shell of interactorswas phospho-2-dehydro-3-deoxyheptonate aldolase; Stereospecific condensation of phosphoenolpyruvate (PEP) and D-erythrose-4-phosphate (E4P) giving rise to 3-deoxy-D- arabino-heptulosonate-7-phosphate (DAHP) (366 aa) (aroF). It's predicted functional interacting partners were (Fig. 11):

- 3-dehydroquinate synthase (365 aa)-(aroB),
- phospho-2-dehydro-3-deoxyheptonate aldolase (462 aa)-(cg2391),
- prephenate dehydrogenase (346 aa)-(tyrA),
- prephenatedehydratase (315 aa)-(pheA),
- undecaprenyl pyrophosphate synthase (256aa)-(cg1130).
- hypothetical protein (173 aa)-(cg1131)
- 3-deoxy-D-arabinoheptulosonate-7-phosphate synthase (466aa)-(NCgl2098)
- 3-phosphoshikimate 1-carboxyvinyltransferase (430aa)-(aroA)
- shikimate kinase (185 aa)-(cg1828) and
- shikimate kinase; shikimate kinase; Catalyzes the specific phosphorylation of the 3-hydroxyl group of shikimic acid using ATP as a cosubstrate (By similarity) (190 aa)-(aroK).

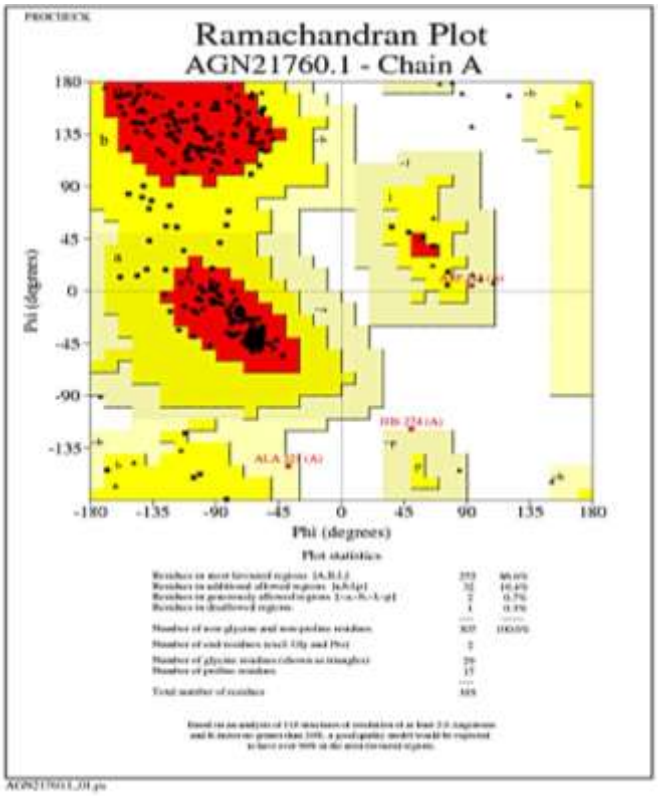

Figure 9.Ramachandran plot obtained for the built protein.

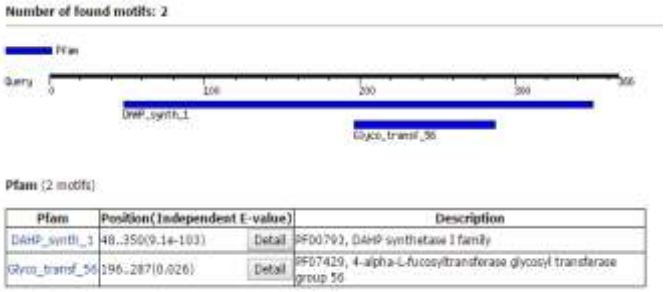

Figure 10. Motif search result.

Figure 11.STRING analysis showing interacting proteins with the AGN21760.1.

CONCLUSION

DAHP synthase is an important enzyme which catalyzes the first step of shikimate pathway. It finally produces the aromatic amino acids. From the *in silico* study of DAHP synthase it was found that the protein wasthermostable and acidic in nature having molecular weight ranging from 38.67 to 52.32 kDa. In this study secondary and tertiary models were prepared from different servers. The phylogenetic trees showed a correlation of the relationship among the corresponding reverse translates. However, the overview may help to obtain a theoretical idea about the structural and functional properties of DAHP synthase enzyme.

IV. ACKNOWLEDGEMENTS

First author is thankful to Department of Science and Technology, New Delhi, Government of India for awarding the INSPIRE fellowship (Reg. No.: IF150197).

REFERENCES

[1] H.E. Umbarger, "Amino acid biosynthesis and its regulation," Annu. Rev. Biochem., vol. 47, pp. 533-606, 1978.

[2] F. Gibson and J. Pittard, "Pathways of Biosynthesis of Aromatic Amino Acids and Vitamins and Their Control in Microorganisms," Bacteriol. Rev., pp. 465-492, 1968.

[3] A. Klausner,"Builiding of success in phenylalanine," Biotechnology, vol. 3, pp. 301-307, 1985.

[4] G.A. Chrosby,"New sweetners," CRC Crit. Food Sci., vol.7, pp. 297-323, 1976.

[5] K. Wakasa and A. Ishihara, "Metabolic engineering of the tryptophan and phenylalanine biosynthetic pathways in rice," Plant Biotech., vol. 26, pp. 523–533, 2009.

[6] P.J. Facchini, K.L. Huber-Allanach, and L.W. Tari, "Plant aromatic L-amino acid decarboxylases: evolution, biochemistry,regulation, and metabolic engineering applications," Phytochemistry, vol. 54, pp. 121–138, 2000.

[7] L. Eggeling and M. Bott, "Handbook of *Corynebacteriumglutamicum*," Taylor & Francis,Boca Raton, 2005.

[8] S. Kumar, G. Stecher, and K. Tamura, "MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets," Mol. Biol. Evol., vol. 33, pp. 1870-1874, 2016.

[9] A. Rzhetsky and M. Nei, "Minimum Evolution method: A simple method for estimating and testing minimum evolution trees," Mol. Biol. Evol., vol. 9, pp. 945-967, 1992.

[10] J. Felsenstein, "Confidence limits on phylogenies: An approach using the bootstrap," Evolution, vol. 39, pp. 783-791, 1985.

[11] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, and A. Bairoch, "Protein identification and analysis tools on the ExPASy server," In: Walker, John M. (Ed.), "The Proteomics Protocols Handbook. Humana Press," pp. 571–607, 2005.

[12] K. Pramanik, T. Soren, S. Mitra, and T.K.Maiti, "*In silico* structural and functional analysis of *Mesorhizobium* ACC deaminase," Comp. Biol. Chem., vol. 68, pp. 12–21, 2017.doi: 10.1016/j.compbiolchem.2017.02.005.

[13] K. Pramanik, P.K. Ghosh, S. Ray, A. Sarkar, S. Mitra, and T.K. Maiti, "An *in silico* structural, functional and phylogenetic analysis with three dimensional protein modeling of alkaline phosphatase enzyme of *Pseudomonas aeruginosa"*. J. Gen. Eng. Biotech., Accepted manuscript, 2017. doi: 10.1016/j.jgeb.2017.05.003.

[14] P.Y. Chou and G.D. Fasman, "Conformational parameters for amino acids in helical, b-sheet, and random coil regions calculated from proteins," Biochemistry, vol. 13, pp. 211–222, 1974a. doi:http://dx.doi.org/10.1021/bi00699a001.

[15] P.Y. Chou and G.D. Fasman, "Prediction of protein conformation," Biochemistry, vol. 13, pp. 222–245, 1974b. doi:http://dx.doi.org/10.1021/bi00699a002.

[16] M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T.G. Cassarino, M. Bertoni, L. Bordoli, and T. Schwede, . "SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information," Nucl. Acids Res. vol. 42, pp. W252–W258, 2014. doi:http://dx.doi.org/10.1093/ nar/gku340.

[17] A. Verma, V.K. Singh, and S. Gaur, "Computational based functional analysis of *Bacillus*phytases," Comput. Biol. Chem., vol. 60, pp. 53-58, 2016.http://dx.doi.org/10.1016/j.compbiolchem.2015.11.001.

[18] P. Benkert, M. Biasini, and T. Schwede, "Toward the estimation of the absolute quality of individual protein structure models," Bioinformatics, vol. 27, pp. 343–350, 2011.doi: http://dx.doi.org/10.1093/bioinformatics/btq662.

[19] S.C. Lovell, I.W. Davis, W.B. Arendall III, P.I.W. de Bakker, J.M. Word, M.G. Prisant, J.S. Richardson, and D.C. Richardson, "Structure validation by Calpha geometry: phi,psi and Cbeta deviation," Proteins: Structure, Function & Genetics., vol. 50, pp. 437-450, 2002.

[20] M. Ikeda, "Towards bacterial strains overproducing l-tryptophan and other aromatics by metabolic engineering," Appl. Microbiol. Biotechnol., vol. 69, pp. 615-626, 2006.