In Silico Modeling and Characterization of Fas D Protein from *Rhodococcus fascians* Associated with Pathogenecity

Juri Saikia , Rituparna Sarma , D.K. Sharma

Abstract— *Rhodococcus fascians* infects a wide range of plants, initiating the formation of leafy galls that consist of shoot amplification and shoot growth inhibition.To provoke symptoms, *R. fascians* strain D188 requires pathogenicity genes that are located on a linear plasmid, pFiD188. The *fas* genes are essential for virulence and constitute an operon that encodes, among other functions, a cytokinin synthase gene i.e., Fas D. Loss or mutation in Fas D gene lead to a complete loss of pathogenicity. Structural and functional characterization of Fas D of Rhodococcusfascians is of interest while its structure remains unknown. Thus a homology molecular model of Fas D was constructed for gleaning possible structural insights.

Index Terms— cytokinine synthase gene, fas D, homology modeling, leafy gall, mutation, pathogenecity, *Rhodococcus fascians*,

1 INTRODUCTION

Rhodococcus fascians is a Gram positive bacterial phytopathogen that causes leafy gall (Baker 1950) disease in both dicotyledonous and monocotyledonous hosts, commonly afflicts tobacco (*Nicotiana*) plants, (Baker 1950 ; Bradbury 1986) and is a well-adapted epiphyte which provokes leafy gall formation through secretion of signal molecules that interfere with the hormone balance of the host (Vereecke, Danny, et al.2002).Infection of dicotyledonous plants can result in the local proliferation of meristematic tissue, leading to galls that are covered with leaflets (Cornelis et al., 2001). On monocotyledonous plants, such as lilies, R. fascians provokes severe malformations of the bulbs and the formation of long side shoots (Crespiet al.,1992) resulting in abnormal plants that are unfit for commercial use (Crespiet al.,1994). Infection of tobacco seedlings with *R. fascians* strongly inhibits growth, accompanied by arrested root development, thickening and stunting of the hypocotyl, and inhibition of leaf formation (Danny and Vereeckeetal.,2002).In laboratory R. fascians strain D188 genes involved in pathogenicity were shown to be located on a large, conjugative, linear, fasciation-inducing plasmid (pFiD188) (Elia et al.,1984).Random mutagenesis of pFiD188 led to the identification of three virulence loci,vizhyp, Att and fas, of which the best characterized is the essential fas locus (Faivre and Amiot 1967).This locus consists of an operon of six genes,(fasA-fasF) that encodes the cytokinin biosynthetic machinery (Gasteigeret al.,2003). It is very essential to study the fas D gene, the main virulence factors in R. fascians,and for that the 3D structure of the Fas D gene product isopentenyltransferase(Ipt) should be available.In order to a structure-based virtual screening exercise it is necessary to have the 3D structure of the protein. Most commonly the structure of the proteins has been determined by experimental techniques such as X-ray crystallographyor NMR. For proteins, if the structure is not available, one can resort to the techniques of protein structure prediction.(Maes.et al., 2001. Temmerman.etal.,2000).Currently the 3D structure of isopentenyl transferase (Ipt) from Rhodococcus fascians D188 is not available in the Protein Data Bank (PDB). Hence protein modeling of isopentenyl transferase from Rhodococcus fascians D188 can be carried out using the Modeller software.

All the authors are affiliated to “Bioinformatics Infrastructure Facility (BIF), Gauhati University, Guwahati -781014, Assam, India
Corresponding author: Juri Saikia.Tel- +91-9706049775
Email. Id- jurijiwanee@gmail.com

International Journal of Scientific & Engineering Research, Volume 6, Issue□k□ , È à-2015
ISSN 2229-5518

2 Methodology

2.1 Sequence retrieval

The complete protein sequences of iso- pentenyl transferase (Ipt) from Rhodococcus fascians D188 was downloaded from NCBI for this study.

2.2 Template selection

A suitable template (PDB ID:2ZE5 _A with resolution 2.3 Å) was selected using BlastP (Protein Blast) against the PDB (Protein databank) database for the Fas D query sequence with calculated DOPE Score.

2.3 Homology model

The homology model of Fas D (Figure 1) was developed using Modeller (Sanchez R &Sali1997)

2.4 Efficiency test of protein homology modeling structures

The protein homology structure are evaluated using twoonline software; ERRAT(Colovos 1993) and RAMPAGE(Lovell et al.,2002). ERRAT is a proteinstructure verification algorithm. ERRAT runs by statisticalanalysis of non-bonded interactions between different typesof atom. It generates a single output plot showing the errorvalue to the residue window. By statistical data comparisonwith highly evaluated structures, it generates the error valuesto yield the confidence limits. This is extremely beneficial totest the homology model reliability (ERRAT v2.0).RAMPAGE is an online server which designs a Ramachandranplot from the input data by plotting phi (φ)versus psi (Ψ) dihedral angles of each residue. The plot isdivided into three distinct regions: allowed, disallowed andfavored regions based on density dependent plotting of the residues.

2.5 Cavity prediction and characterization

The active site prediction server (Singh T et al.2011) and dogsitescorer server (Volkamer et.al., 2012)were used for the calculation of pockets and cavities respectively. The server outputs data with cavities for PDB (protein databank) input files. The analysis shows 7 pockets and 35 cavities in the Fas D homology model. The cavity residue stretch and volume in FasD were shown in Table 1.

2.6 Sequence analysis

ClC Genomics was used to analyze the sequence.

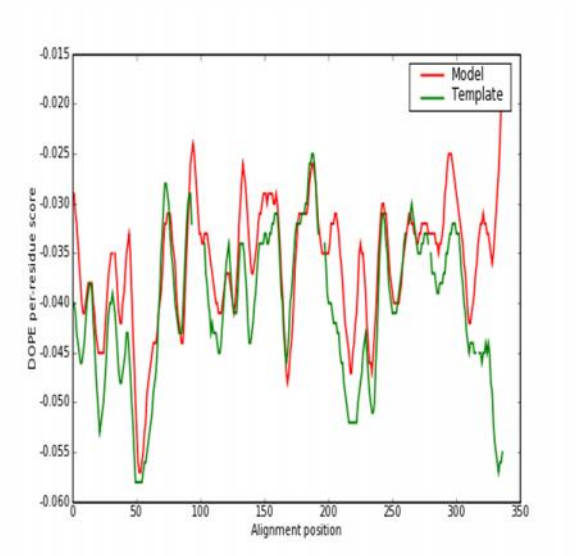

Figure 2: DOPE profile of the model with the template is shown as a function of residue position number.

3 Results and discussion

3.4 Sequence retrieval and analysis:

The retrieved protein sequence of Fas D gene (Accession No YP_007878707.1) was found 255 aa in length and 28.035 Kd. in molecular weight.The Count of hydrophobic and hydrophilic residues in the protein shows hydrophobic residue 135 and Hydrophilic 59 0.231 Other 61 0.239 and Isoelectric point of 5.65, which indicated that the proteins is acidic. Aliphatic Index-92.196, which suggested that it had great stability over different temperature ranges. Fas D has Negatively Charged residues 31 ,Positively Charged residues 24 0.094 and nutral 200 0.784. Half life of the protein is 10 hours. The half-life of different proteins can vary from minutes to days. As it shows 10 hours that means it can resist degradation when isolated in-vitro condition(Vogel , & Marcotte 2012). The protein has 11 alpha chain and 11Beta strand.

International Journal of Scientific & Engineering Research, Volume¤Ú , È à-2015
ISSN 2229-5518

3.5 Homology modeling and evaluation:

Modeller 9V8 predicted the 3D structures of the proteins and the results were available in PDB format. The Modeller server satisfactorily predicted the protein structure, Isopentenyl transferase (Ipt) using best score orthologous template.

The MODELLER predicted homology structures of Isopentenyl transferase (Ipt) from R. fascians D188 was accessed by ERRAT and RAMPAGE, evaluation servers. The output of the ERRAT assessment showed scored 74.494 % and when assessed in RAMPAGE it was scored 94.5% . From these assessments it can be assumed that the structure is reliable.

255 no of groups, 1971 no of atoms , 2008 no of bonds ,5 beta strands , 15 alpha helix , and 18 no of turns present in the structure.

No of pockets in the predicted structure was calculated as 7 in no by using dogsite scorer.The largest pocket was found 1222.78 Å volume , 1988.09 A^2 surface. For protein structure based drug designing identification and evaluation of surface binding pockets and occluded cavities are initial steps.Size and shape characterization of active sites are very much essential for variety of applications such as automated ligand docking or *in situ* modeling.(Weisel et.al., 2007).

No special amino acids were found to be present in the structure.Valine is present in highest no 11 followed by Alanine 7 in no.Valine are not involved in protein function like catalysis as their side chains are non- reactive. , but they help in substrate recognition.(Betts et al.,2003)

The active site prediction server helped to identify the cavities present in the Fas D protein. 35 cavities are found in the structure with the residue sequence stretch. cavity point and volume cavity to locate the active sites in Fas D for potential ligand binding characterization.Table 1 (see supplementary material).

Cavities in protein surface create physiochemical properties which are required for molecular functions.

4 Conclusion

The structure was submitted to PMDB and they assigned the structure with the accession number id: PM0079293

It is of interest to characterize the structure of Isopentenyl transferase (Ipt) from Rhodococcus fascians D188.Hence, a structural model of the protein was reported with its 35 cavities identified using prediction methods.The models presented here can serve as a guide for the allocation of amino acid residues involved in each fold, which is important for further investigations on molecular mechanism of actions**.** The models also gains importance for the structural biologist and even to the computer aided pesticide designers from different angles.

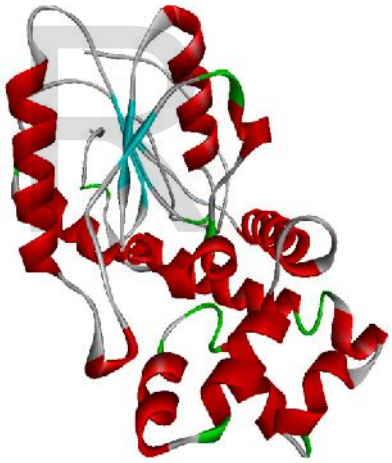

Figure 1: A homology model of Fas D is shown. The homology model was developed using modeler 9V8. (Peng Jian, &Xu Jinbo.2011**)**

ACKNOWLEDGMENT

The authors wish to thank BIF Centre Zoology Department , Gauhati University Guwahati And DBT for providing facilities in the centre.

REFERENCES

1. Baker K F. Bacterial fasciation disease of ornamental plants in California. Plant Dis Rep. 1950;34:121–126.

International Journal of Scientific & Engineering Research, Volume 6, Issue , -2015
ISSN 2229-5518

2. Betts, M. J., & Russell, R. B. (2003). Amino acid properties and consequences of substitutions. *Bioinformatics for geneticists*, *317*, 289.
3. Bradbury, J. F. 1986. Rhodococcus Zopf 1891, p. 185-187. In J. F. Bradbury (ed.), Guide to plant pathogenic bacteria. CAB International Mycological Institute, Slough, United Kingdom.
4. Colovos C, Yeates T.O. Verification of protein structures:patterns of nonbonded atomic interactions. Protein Sci.1993;2:1511e1519.
5. Cornelis,K., T. Ritsema, J. Nijsse, M. Holsters, K. Goethals, and M. El Jaziri. 2001. The plant pathogen Rhodococcus fascians colonizes the exterior and interior of aerial parts of plants. Mol. Plant-Microbe Interact. 14:599-608. [PubMed].
6. Crespi M, Messens E, Caplan A B, Van M Montagu M, Desomer J. Fasciation induction by the phytopathogen Rhodococcus fascians depends upon a linear plasmid encoding a cytokinin synthase gene. EMBO J. 1992;11:795–804. [PMC free article] [PubMed]
7. Crespi M, Vereecke D, Temmerman W, Van Montagu M, Desomer J (1994) The fas operon of Rhodococcus fascians encodes new genes required for efficient fasciation of host plants. JBacteriol 176:2492–2501.
8. Danny Vereecke,1 Karen Cornelis,1 Wim Temmerman,1 Mondher Jaziri,2 Marc Van Montagu,1 Marcelle Holsters,1,* and Koen Goethals1Chromosomal Locus That Affects Pathogenicity of Rhodococcus fascians J Bacteriol. 2002 February; 184(4): 1112–1120. doi:10.1128/jb.184.4.1112-1120.2002.PMCID: PMC134788.
9. Elia S, Gosselé F, Vantomme R, Swings J, De Ley J. Corynebacterium fascians: phytopathogenicity and numerical analysis of phenotypic features. Phytopathol Z. 1984;110:89–105.
10. Faivre-Amiot A. Quelques observations sur la présence de Corynebacterium fascians (Tilford) Dowson dans les cultures maraichères et florales en France. Phytiatr-Phytopharm. 1967;16:165–176.
11. Gasteiger E., Gattiker A., Hoogland C., Ivanyi I., Appel R.D., Bairoch A.ExPASy: the proteomics server for in-depth protein knowledge and analysisNucleic Acids Res. 31:3784-3788(2003).
12. Goethals K, Vereecke D, Jaziri M, Van Montagu M, Holsters M (2001). "Leafy gall formation by *Rhodococcus fascians*". *Annu Rev Phytopathol***39**:27–52. doi:10.1146/annurev.phyto.39.1.27. PMID 11701858.
13. Lovell SC et al. Proteins: Structure, Functions Genetics. 2002 50: 437 [PMID: 12557186]
14. Maes T., Vereecke D, Ritsema T, Cornelis K, Thi ThuH. Ngo , Van Montagu M, Holsters M, and Goethals K. 2001. The att locus of Rhodococcus fascians strain D188 is essential for full virulence on tobacco through the production of an autoregulatory compound. Mol. Microbiol. 42:13-28. [PubMed].
15. Miller H J, Janse J D, Kamerman W, Muller P J. Recent observations of leafy gall in Liliaceae and some other families. Neth J Plant Pathol. 1980;86:55–68.
16. PengJian, XuJinbo. A multiple-template approach to protein threading. Proteins. April 2011;79(6):1930e1939.
17. Sanchez R &Sali A. PROTEINS: Structure, Functions, Genetics, 1997 1: 50 [PMID: 10940251]

International Journal of Scientific & Engineering Research, Volume 6, Issue , -2015
ISSN 2229-5518

18. Singh T et al. J Chem Inf Modeling. 2011 51: 2515 [PMID: 21877713].

19. Thimann K V, Sachs T. The role of cytokinins in the "fasciation" disease caused by Corynebacterium fascians. Am J Bot. 1966;53:731–739.

20. Topf M, Baker ML, Marti-Renom MA, Chiu W, Sali A.Refinement of protein structures by iterative comparative modeling and CryoEM density fitting. J Mol Biol. 2006;357(5):1655e1668

21. Vantomme R, Elia S, Swings J, De Ley J. Leafy gall disease on Liliaceae. Med Fac Landbouww Rijksuniv Gent. 1982;47/3:1105–1108.

22. Vereecke, D, Burssens S, Simón-Mateo C, Inzé D, Van Montagu M, Goethals K, Jaziri M. The Rhodococcus fascians-plant interaction: morphological traits and biotechnological applications. Planta. 2000;210:241–251. [PubMed].

23. Vereecke, D., Cornelis, K., Temmerman, W., Jaziri, M., Van Montagu, M., Holsters, M., & Goethals, K. (2002). Chromosomal locus that affects pathogenicity of Rhodococcus fascians. *Journal of bacteriology*, *184*(4), 1112-1120.

24. Vogel, C., &Marcotte, E. M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nature Reviews Genetics, 13(4), 227-232.

25. Volkamer, A., Kuhn, D., Rippmann, F., & Rarey, M. (2012). DoGSiteScorer: a web server for automatic binding site prediction, analysis and druggability assessment. *Bioinformatics*, *28*(15), 2074-2075.

26. Weisel, M., Proschak, E., & Schneider, G. (2007). PocketPicker: analysis of ligand binding-sites with shape descriptors. *Chem Cent J*, *1*(7), 1-17.

27. WimTemmerman, Danny Vereecke, Rozemarijn Dreesen, Marc Van Montagu, Marcelle Holsters,* and Koen Goethals Leafy Gall Formation Is Controlled by fasR, an AraC-Type Regulatory Gene in Rhodococcus fascians J Bacteriol. 2000 October; 182(20): 5832–5840.PMCID:PMC94707

Supplementary material:

Table 1: Predicted cavities in Fas D from *Rhodococcus fascians*

Fas D cavities with sequence	Volume(Å)	Cavity point (Å)
cavity_1_VDNIQKRTLPASEGYFMHC	1215	78.279 67.256 14.370
cavity_2_RFDLQNWVTHIYSKAEGM	1209	70.939 57.570 30.995
cavity_3_RLSFNIEMADVQWYGHPKT	1074	68.505 54.371 16.104
cavity_4_EVGLRMFATKNQSDC	931	86.467 53.331 24.435
cavity_5_RSIEFNQLHVDATYGM	907	66.863 46.223 11.156

International Journal of Scientific & Engineering Research, Volume 6, Issue , -2015
ISSN 2229-5518

cavity_6_DREFLSQNIGHAYVT	887	60.371	48.340	17.056
cavity_7_SRVYLDMANWPEGIHQKF	875	66.154	68.201	26.854
cavity_8_AWDRQSNIHKTLVGEYFMC	854	77.009	62.135	22.565
cavity_9_RPSGFTALVIEYWQNM	782	63.428	71.085	3.018
cavity_10_VDEHRYFGQAMTWLC	768	81.316	47.814	7.914
cavity_11_HQNFTDIRYGLVSAWE	763	71.939	47.466	31.861
cavity_12_QDHALVEWRPFTNMGCS	745	81.564	46.174	30.145
cavity_13_ILQNDTPAKFVRSMGYWE	739	55.377	65.760	15.186
cavity_14_FHDWPIAQKRTESYGVMCL	672	76.962	59.060	10.840
cavity_15_VTDSLMRFPYGKAENIH	654	61.480	76.357	22.764
cavity_16_EVAYSHNTLMPQIKG	645	71.755	74.640	9.003
cavity_17_KTDIVEFSLRYGMANWP	631	58.211	67.285	27.650
cavity_18_DREFQSHNGLTIAYV	618	64.612	48.773	22.866
cavity_19_GYQARMFHLCEDV	596	85.882	54.920	10.595
cavity_20_TVWQNPRIEADH	594	73.673	65.990	-1.169
cavity_21_PNLIFSRVMGTKAEYH	580	57.554	74.553	9.139
cavity_22_REDFVLSNMAWTIYGHQK	573	63.252	58.105	24.186
cavity_23_IVADLREYGQPMFTNCH	558	79.871	49.477	20.799
cavity_24_NSFVQADELWRGYPTCH	556	73.127	39.808	16.444
cavity_25_RFPLTAGIVESQWNHD	501	64.285	60.785	5.638
cavity_26_NLQIPTAFRSGVME	487	51.308	69.498	6.658
cavity_27_DERFQSHNGLTIAYV	487	59.564	49.928	23.152
cavity_28_FSPLMVAKGYEIHNTQ	476	64.050	72.167	13.187
cavity_29_NFIGSQVHDALWEPTC	435	74.736	37.798	25.719
cavity_30_QARLGFPSIETVW	369	56.482	54.881	5.621
cavity_31_ILQNDTKAFVPSRMY	351	50.803	67.283	19.230
cavity_32_RLFTAIEVQWPD	342	64.252	57.943	-0.272
cavity_33_ELQRHAVYFDG	279	72.771	50.670	4.368
cavity_34_RDVTLCWHASEN	269	86.421	41.806	16.358
cavity_35_QIDTKAFRESLVM	252	49.738	59.409	18.964