View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Scholarworks@UTRGV Univ. of Texas RioGrande Valley

University of Texas Rio Grande Valley

ScholarWorks @ UTRGV

School of Medicine Publications and Presentations

School of Medicine

2017

In silico study of RxLR effectors of Phytophthora infestans HP-10-31, A2 mating type potato late blight pathogen

Youvika Singh

Virupaksh U. Patil

Anupam Dhasmana
The University of Texas Rio Grande Valley

Swarup K. Chakraborty

Pradeep K. Shukla

See next page for additional authors

Follow this and additional works at: https://scholarworks.utrgv.edu/som_pub

 Part of the Medicine and Health Sciences Commons

Recommended Citation

Singh, Youvika; Patil, Virupaksh U.; Dhasmana, Anupam; Chakraborty, Swarup K.; Shukla, Pradeep K.; and Rawat, Shashi, "In silico study of RxLR effectors of Phytophthora infestans HP-10-31, A2 mating type potato late blight pathogen" (2017). *School of Medicine Publications and Presentations*. 158.
https://scholarworks.utrgv.edu/som_pub/158

This Article is brought to you for free and open access by the School of Medicine at ScholarWorks @ UTRGV. It has been accepted for inclusion in School of Medicine Publications and Presentations by an authorized administrator of ScholarWorks @ UTRGV. For more information, please contact justin.white@utrgv.edu, william.flores01@utrgv.edu.

Authors

Youvika Singh, Virupaksh U. Patil, Anupam Dhasmana, Swarup K. Chakraborty, Pradeep K. Shukla, and Shashi Rawat

This article is available at ScholarWorks @ UTRGV: https://scholarworks.utrgv.edu/som_pub/158

International Journal of Advanced Biotechnology and Research (IJBR)
ISSN 0976-2612, Online ISSN 2278–599X,
Vol-8, Issue-2, 2017, pp398-404
http://www.bipublication.com

Research Article

In silico study of *RxLR* effectors of *Phytophthora infestans* HP-10-31, A2 mating type potato late blight pathogen

Youvika Singh[1], Virupaksh U. Patil[1], Anupam Dhasmana[2], Swarup K. Chakbraborty[1], Pradeep K. Shukla[3] and Shashi Rawat[1*]

[1]Central Potato Research Institute, Shimla, Himachal Pradesh. India
[2]Research Cell, Amity University, Lucknow, Uttar Pradesh. India
[3]Department of Biological Science, School of Basic Sciences, SHIATS, Allahabad. India
*Corresponding Author:Email- shashi29@yahoo.com, Tel: +91-9418669929; Fax- +91-177-2624460

ABSTRACT:
Phytophthora infestans is one of the most compelling plant pathogen among the scientific community throughout the world. It is the causative agent of potato late blight and responsible for tremendous economic loss worldwide. Pathogenic effector proteins are instrumental in modulating host immunity and disease resistance has been a major concern. In *P. infestans*, a class of cytoplasmic effectors recognized as RxLR is characterized by highly conserved region and abet in parasitic colonization by modifying the host defense system. We have sequenced an Indian strain of *P. infestans* HP-10-31 genome and identified several RxLR motif-containing genes.In this study we selected two *RxLR* effector genes named contig15921_2 and contig06738_6 from this A2 mating type strain. We used I-TASSER server to generate three-dimensional structure and observe the Nicotinamide adenine dinucleotide and S-adenosyl-L-homocysteine conserve domains. Our *in silico* study reveals the binding properties of these proteins are favorable with corresponding ligands. This study gives insight into the interaction between putative RxLR effector proteins with its ligand that further aid our understanding of host-pathogen interaction and help in designing new agents to combat the agro pathogenicity.

Keywords:*Phytophthora*, effectoromics, motif, RxLR, ligand binding, NAD, SAH

[I] INTRODUCTION

Late Blight caused by *Phytophthora infestans* in potato and tomato is the most devastating disease which account for huge economic loss worldwide [1]. The annual economics loss caused by the disease is estimated to over 6.7 billion globally [2]. Moreover, estimates predict that fungicide sprays contribute to $1 million investment each year [3]. Therefore Late Blight is not only threatens food security but also hazardous to environment.

Like other plant pathogens, *P. infestans* too secretes the armory of effector proteins that obstruct host physiology and defense system to enable successful parasitic colonization [4-6]. However, plant species possess resistance (R) genes which detect these effectors then termed avirulence (AVR) proteins and activate Effector-Triggered Immunity against the pathogen. Based on their target site, two classes of effector proteins are recognized in *Phytophthora*: apoplastic and cytoplasmic. Apoplastic effectors are secreted into plant extracellular space where they interact with extracellular host molecule and counter-defense by inhibiting host enzymes, triggering hypersensitive response. On the other hand,

cytoplasmic effectors translocated inside plant cell through specialized structure (haustoria) and target different subcellular compartments[7]. In oomycetes, most attention has been given to RxLR effectors (cytoplasmic) which are identified as Avr genes having RxLR motif (R:arginine, x:any amino acid, L:leucine and R:arginine) [8, 9]. RxLR proteins are modular proteins characterized by two main functional domain N-terminal domain encompassing signal peptide RxLR motif involved in secretion and targeting effectors into plant host cell respectively and C-terminal domain which is involved in effector function [4, 10, 11]. Oomycetes sequencing projects and genomic approaches have yielded several effector proteins. The number of RxLR genes identified in *P. infestans* was 563 [12] is almost double that in *P. sojae* (335) and *P. ramorum* (309) [13]that suggest the importance and complex organization of RxLR amino acid domain in these organisms.

Despite refinements in recent approaches towards molecular and genomics advancements, yet very little is known about the molecular pathogenicity of *P. infestans*.Knowledge of tertiary structure comes handy in functional prediction and interaction of target protein with other compounds. Understanding molecular basis of pathogenesis and structural properties would further deepen our knowledgein effector virulence and plant pathogenesis. We have identified several number of RxLR effectors from genome sequencing project of Indian isolate of *P. infestans* HP-10-31 A2 mating type (unpublished). In present study two RxLR effectors: contig15921_2 and contig06738_6 are selected for *in silico* study since there are no earlier studies reported on these proteins. Three-dimensional structure of these effectors was generated by I-TASSER. Further we employed various bioinformatics tools to study physicochemical and binding properties. We also find that S-adenosyl-L-homocysteine [12] and Nicotinamide Adenine Dinucleotide (NAD) may serve as possible ligand for contig06738_6 and contig15921_2 respectively. Our docking studies shows that this ligand has high affinity with the RxLR proteins. We believe that this effector trafficking into plant host facilitate the disease establishment. Using bioinformatics approach could further leads to the development of effective agro management strategies against Late Blight control.

[II] MATERIALS AND METHODS

2.1.Sequence information

The amino acid sequence of RxLR protein was obtained from sequencing data of *P. infestans* HP-10-31 (NCBI Accession number LYVM00000000). The protein sequence of contig06738_6 and contig15921_2 is presented in figure 1 A & B.

2.2. Physicochemical analysis

Physicochemical data were generated using ProtParam tool (http://web.expasy.org/protparam/) [14] of Expasy Proteomic server. FASTA format of the protein sequence were used as input material. Various physico-chemical parameters such as molecular weight, isoelectric point (pI), total number of positively and negatively charged residues, extinction coefficient [15], instability index [16], aliphatic index [17] and grand average hydropathy (GRAVY) [18] were computed.

2.3. Homology modeling and model quality assessment

I-TASSER (Iterative Threading ASSEmbly Refinement) was employed for generating three-dimensional structure of contig06738_6 and contig15921_2. I-TASSER is a bioinformatics tool for protein structure and function prediction [19]. It generate 3D model from multiple threading alignment and iterative structural assembly whose accuracy can be measured in terms of *C-score*. Out of five models generated, the final model was taken on the basis of highest *C-score* and TM value and further structural refinement was performed. Stability and quality of the selected models were analyzed by Ramachandran plot using online RAMPAGE server [20]. Ligand for both effector proteins were predicted by using COFACTOR server [21]. It also predicts the possible binding sites to the target protein.

2.4. Docking Studies

2.4.1. Preparation of Receptor

The structure of proteins involved in present study is author's own sequence and modelled RxLR effectors of Late Blight pathogen *P. infestans* HP-10-31. Chimera was used for energy minimization, removal of steric collision with the steepest descent steps 1000, steepest descent size 0.02 Å, conjugated gradient steps 1000, and the conjugate gradient step size 0.02 Å for the conjugate gradient minimization [22, 23].

2.4.2. Preparation of Ligand

Ligand file of SAH and NAD were downloaded in .mol format from chemSpider Chemical Database. These file were converted into .pdb files using Discovery Studio Visualizer (V 2.5.5) since they could not be used directly by AutoDock 4.2 tools [24]; thus they. Further the ligand was submitted for minimization using chimera version 1.5.3 using genetic algorithm steps 2000 and 0.5 grid units optimized [25].

2.4.3. Docking Studies

Docking studies were performed by AutoDock (V 4.2) [26, 27], on Microsoft Windows 7 professional Version 2002, Service pack 3, operating System on Intel (R) Core [8] 2 Duo, CPU T6500 @ 2.10 GHz, 1.19 GHz, and 2.96 GB of RAM of Dell Machine. We execute molecular docking methods followed by searching of best confirmation of RxLR effectors with their ligands on basis of binding energy. Water molecules were removed from protein structures before docking and H-atoms were added to all target proteins. Kollman united charges and saluation parameters were added to proteins. Gasteiger charge was added to ligands. Grid box was set to maximum part of proteins and ligands. The values were set 60×60×60 A . Lamarckian genetic algorithm (LGA) [28] was used for protein ligand flexible docking calculations. The LGA parameters like population size (ga_pop_size), energy evaluation (ga_num_generation), mutation rate, crossover rate and set size were set to 150, 2,500,000, 27,000, 0.02, 0.8, and 0.2 Å, respectively. The LGA runs were set to 10 runs. All conformations of proteins and ligands were obtained and analyzed for interaction and binding energy of docked structure using Discovery Studio Visualizer (V 2.5.5).

[III] RESULTS AND DISCUSSION

3.1. Physicochemical properties elucidate the extracellular characters of RxLR genes

Various physicochemical features were evaluated using ProtParam (Table 1).RxLR effector contig15921_2 showed pI >7.0, which indicate their basic nature of these proteins. Earlier studies on *Coxiela burnetii*, a Q fewer causing, obligate intracellular bacterial pathogen revealed that 55% of identified effector proteins were basic in nature which is required to counterbalance the phagolysosomal acidic environment [29, 30] .It might be predicted that similar kind of mechanism also works in pathogenesis of *P. infestans*. Stability of the protein was studied by analyzing the values for instability index, aliphatic index and grand average of hydropathicity (GRAVY) index. The thermostability of globular proteins is directly proportional to aliphatic index (AI), which is relative volume of protein occupied by aliphatic side chains. Effector proteins being studied were found to have high aliphatic index that indicate their possible thermostable property in the wide range of temperature. The instability index provides an estimate of the stability of protein in a test tube and protein whose instability index is smaller than 40 is predicted as stable, a value above 40 predicts that the protein may be unstable. Both of effector proteins are stable as they possess instability index lower than 40. The negative GRAVY indices of all studied proteins demonstrated their affinity with water as increasing GRAVY scores indicates greater hydrophobicity. These results showed that both proteins are in agreement with the properties required for extracellular function of the protein.

3.2. Three dimensional modelling and structure validation of RxLR Proteins

The quality of predicted model by I-TASSER can be assessed by highest *C-score* and TM value. On

basis of these parameters we selected first model for contig06738_6 with *C-score* of 3.88 and TM score 0.30±0.10 and for contig15921_2 with *C-score* 2.0 and TM score 0.47±0.15. The generated model was further refined and used for docking studies. Ramachandran plot shows 84.2% and 84.4% residues fall in favorable region for contig06738_6 and contig15921_2 respectively. Also 12.6% and 13.3% residues fall in allowed region for the RxLR proteins (Figure 2). Together these statistics confirmed the quality of the models is acceptable and reliable. We further used COFACTOR server to search for ligand and respective binding site for the proposed model. This server uses the combination of both global and local structural comparison algorithms to gather the biological functions of proteins from their 3D structure [21]. For our protein of interest this server predicts S-Adenosyl-L-homocysteine [12] and Nicotinamide adenine dinucleotide (NAD) ligand for contig6738_6 and contig15921_2 respectively.

3.3. Molecular docking studies

In silico docking studies was carried out to explore orientation and binding energy of RxLR effector protein contig06738_6 and contig15921_2 with that of SAH and NAD respectively. Analysis of receptor-ligand complex was based on parameters binding energies, inhibition constant (Ki) active site residue, H-bond and their distance (Table 2). Docking studies depicted in Fig. 3 shows that active region of both contig06738_6 and contig15921_2 RxLR effectors bind to the ligand SAH and NAD respectively. The evaluated ΔG values shows that both RxLRs binding to ligands are energetically favorable (Table 2). Binding study of these ligands is important since NAD plays role in metabolism and several extracellular regulatory mechanisms including in redox reactions [31]. Another study shows that fungal elicitor activates key transcription gene specific to pathogen defense related pathways in *Petroselinum crispum* cells and shows increase enzymatic activity of S-adenosyl-L-homocysteine hydrolase enzyme activity in *in vitro* as well as *in vivo*[32]. Docking studies of small molecules in receptor binding site and estimation of binding affinity of complex is vital part of structure based drug design [33].

[IV] CONCLUSION

Albeit migration followed by selection is one of the predominating reason for genetic diversity in global population of *P. infestans*. Eventually it became a major challenge for agrobiologist to study the diversity in available field isolates of this pathogen that once caused havoc on humanity. It has been thought that effector proteins accomplish parasitism by reprograming host immune system. These key questions can be answered by combining molecular, biochemical and bioinformatics tools to understand the phenomenon of effectromics. In this study we have identified two key RxLR effector proteins in Indian strain of *P. infestans*. Our physicochemical studies revealed that these predicted proteins are of basic nature along with thermostable and hydrosoluble which indicates their extracellular nature. The three dimensional structure were predicted by I-TASSER server and best model were chosen based on high *C-score*. This server also provides the possible ligand for the refined structure based on the multiple templates it used to predict the 3D model. NAD and S-adenosyl-L-homocysteine binding site have been located in domain may provide functional insight of these effectors in pathogenicity. The docking study clearly demonstrated good interaction of two RxLR effectors contig06738_6 and contig15921_2 with SAH and NAD. The characterization of binding sites of both RxLRs opens new avenues for development and deployment of pesticide against Late Blight pathogen *P. infestans*.

ACKNOWLEDGEMENT

We thank Indian Council of Agricultural Research (ICAR), Government of India, New Delhi for funding this project. The authors also thank the

Division of Plant Protection, CPRI, Shimla for providing the *P. infestans* HP-10- 31 strain.

REFERENCES:

1. Fry, W.E. and S.B. Goodwin, *Re-emergence of potato and tomato late blight in the United States.* Plant Disease, 1997. **81**(12): p. 1349-1357.
2. Haverkort, A., et al., *Applied biotechnology to combat late blight in potato caused by Phytophthora infestans*. Potato Research, 2009. **52**(3): p. 249-264.
3. Haverkort, A.J., et al., *Societal Costs of Late Blight in Potato and Prospects of Durable Resistance Through Cisgenic Modification.* Potato Research, 2008. **51**(1): p. 47-57.
4. Birch, P.R., et al., *Trafficking arms: oomycete effectors enter host plant cells.* Trends Microbiol, 2006. **14**(1): p. 8-11.
5. Chisholm, S.T., et al., *Host-microbe interactions: shaping the evolution of the plant immune response.* Cell, 2006. **124**(4): p. 803-14.
6. Kamoun, S., *A catalogue of the effector secretome of plant pathogenic oomycetes*. Annu Rev Phytopathol, 2006. **44**: p. 41-60.
7. Whisson, S.C., et al., *A translocation signal for delivery of oomycete effector proteins into host plant cells*. Nature, 2007. **450**(7166): p. 115-8.
8. Vleeshouwers, V.G., et al., *Understanding and exploiting late blight resistance in the age of effectors*. Annu Rev Phytopathol, 2011. **49**: p. 507-31.
9. van Poppel, P.M., et al., *The Phytophthora infestans avirulence gene Avr4 encodes an RXLR-dEER effector*. Mol Plant Microbe Interact, 2008. **21**(11): p. 1460-70.
10. Morgan, W. and S. Kamoun, *RXLR effectors of plant pathogenic oomycetes.* Curr Opin Microbiol, 2007. **10**(4): p. 332-8.
11. Schornack, S., et al., *Ten things to know about oomycete effectors.* Mol Plant Pathol, 2009. **10**(6): p. 795-803.
12. Haas, B.J., et al., *Genome sequence and analysis of the Irish potato famine pathogen Phytophthora infestans.* Nature, 2009. **461**(7262): p. 393-8.
13. Tyler, B.M., et al., *Phytophthora genome sequences uncover evolutionary origins and mechanisms of pathogenesis.* Science, 2006. **313**(5791): p. 1261-6.
14. Gasteiger, E., et al., *ExPASy: The proteomics server for in-depth protein knowledge and analysis*. Nucleic Acids Res, 2003. **31**(13): p. 3784-8.
15. Gill, S.C. and P.H. von Hippel, *Calculation of protein extinction coefficients from amino acid sequence data*. Anal Biochem, 1989. **182**(2): p. 319-26.
16. Guruprasad, K., B.V. Reddy, and M.W. Pandit, *Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence*. Protein Eng, 1990. **4**(2): p. 155-61.
17. Ikai, A., *Thermostability and aliphatic index of globular proteins.* J Biochem, 1980. **88**(6): p. 1895-8.
18. Kyte, J. and R.F. Doolittle, *A simple method for displaying the hydropathic character of a protein.* J Mol Biol, 1982. **157**(1): p. 105-32.
19. Roy, A., A. Kucukural, and Y. Zhang, *I-TASSER: a unified platform for automated protein structure and function prediction.* Nat Protoc, 2010. **5**(4): p. 725-38.
20. Lovell, S.C., et al., *Structure validation by Calpha geometry: phi,psi and Cbeta deviation.* Proteins, 2003. **50**(3): p. 437-50.
21. Roy, A., J. Yang, and Y. Zhang, *COFACTOR: an accurate comparative algorithm for structure-based protein function annotation*. Nucleic Acids Res, 2012. **40**(Web Server issue): p. W471-7.
22. Wang, J., et al., *Automatic atom type and bond type perception in molecular mechanical calculations.* J Mol Graph Model, 2006. **25**(2): p. 247-60.
23. Wang, J., et al., *Development and testing of a general amber force field*. J Comput Chem, 2004. **25**(9): p. 1157-74.
24. Morris, G.M., et al., *Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4*. J Comput Aided Mol Des, 1996. **10**(4): p. 293-304.
25. Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research and analysis.* J Comput Chem, 2004. **25**(13): p. 1605-12.
26. Morris, G.M., et al., *Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function.* Journal of Computational Chemistry, 1998. **19**(14): p. 1639-1662.
27. Rarey, M., et al., *A fast flexible docking method using an incremental construction algorithm*. J Mol Biol, 1996. **261**(3): p. 470-89.
28. Goodsell, D.S., G.M. Morris, and A.J. Olson, *Automated docking of flexible ligands: applications of AutoDock.* J Mol Recognit, 1996. **9**(1): p. 1-5.

29. Seshadri, R., et al., *Complete genome sequence of the Q-fever pathogen Coxiella burnetii.* Proc Natl Acad Sci U S A, 2003. **100**(9): p. 5455-60.
30. Samoilis, G., et al., *Analysis of whole cell lysate from the intercellular bacterium Coxiella burnetii using two gel-based protein separation techniques.* J Proteome Res, 2007. **6**(8): p. 3032-41.
31. Billington, R.A., et al., *Emerging functions of extracellular pyridine nucleotides.* Mol Med, 2006. **12**(11-12): p. 324-7.
32. Kawalleck, P., et al., *Induction by fungal elicitor of S-adenosyl-L-methionine synthetase and S-adenosyl-L-homocysteine hydrolase mRNAs in cultured cells and leaves of Petroselinum crispum.* Proc Natl Acad Sci U S A, 1992. **89**(10): p. 4713-7.
33. Seeliger, D. and B.L. de Groot, *Ligand docking and binding site analysis with PyMOL and Autodock/Vina.* J Comput Aided Mol Des, 2010. **24**(5): p. 417-22.

Protein ID	Number of Amino Acids	Molecular Weight (kDa)	Theoretical pI	Asp+Arg (negative charge)	Arg+Lys (positive charge)	Aliphatic Index	Hydropathy Index
Contig06738_6	457	51.18	6.22	59	54	79.47	-0.483
Contig15921_2	543	59.37	8.25	59	61	95.06	-0.101

Table 1. Physiological features of RxLR proteins identified in *P. infestans* HP-10-31.

Figure 1: The energy minimized final three-dimensional structure obtained from I-TASSER of *P. infestans* HP-10-31 RxLR proteins **(A)** contig06738_6 and **(B)** contig 15921_2 with corresponding amino acid sequence. RxLR characteristic features are depicted in red color.

Protein ID	Contig06738_6	Contig15921_2
Ligand	SAH	NAD
Binding Energy	-3.92 kcal/mol	-0.33 kcal/mol
K_i (mM)	1.34	572.02
Residues	Leu8, Leu11, Ala12, Phe15, Val119, Phe124, Ile271, Gly272, Gly273, Ser274, Trp276, Gly308, Thr309, His333, Glu358	Arg2, Asn395, Asp510, Glu323, Gly397, Gly399, Gly505, Ile398, Leu319, Leu449, Lys316, Lys478, Met511, Phe320, Phe321, Phe452, Pro504, Ser396, Ser451, Thr460, Tyr474, Tyr506, Val507
H-Bonds	:SER274:OG - :UNK1:S17 :UNK1:H45 - :LEU8:O :UNK1:H42 - :GLY308:O :UNK1:H43 - :GLY272:O	3.05618 2.00555 2.4168 2.16313
Distance	:LYS316:NZ - :UNK1:O19 :TYR474:OH - :UNK1:N33 :UNK1:H66 - :GLY505:O :UNK1:H68 - :ASN395:O	2.25366 2.70817 1.59346 2.48805

Table 2. Summary of the binding results for RxLR effectors of *P. infestans* HP-10-31 with their respective ligands.

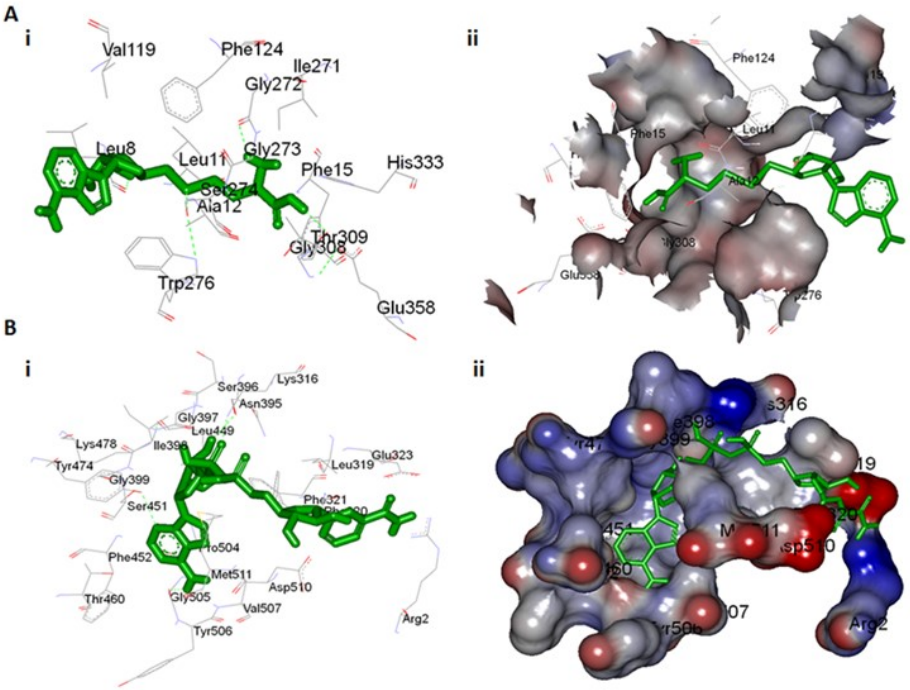

Figure 3: Interaction of RxLR effectors of *P. infestans* HP-10-31 **(A)** contig06738_6 with SAH **(B)** contig15921_2 with NAD. Both ligands are shown in green color. **A (i)** and **B (i)** correspond to 2D protein-ligand interaction. **A (ii)** and **B (ii)** correspond to three dimensional protein-ligand interaction. PyMol was used to generate the molecular surface representation.