Ab initio modelling and function prediction of the protein Q9RR74 of *Deinococcus radiodurans*

Mahimn Dave
Department of Biotechnology, Christ Campus, Rajkot, Gujarat, India
Mahimndave2001@gmail.com

1. Introduction

Deinococcus radiodurans is a polyextremophile, gram-positive, and non-sporulating bacterium well known for its tolerance to high radiation exposure and highly robust DNA repair mechanisms that help to thrive in adverse stress conditions. Survivability can pertain to various factors such as complex cell wall composition, unique structural features of the genome, robust anti-oxidant systems to combat ROS and DNA repair mechanisms. Several studies reveal molecular insights for the extremophilic nature of *D. radiodurans*. Because of the incapacity to synthesize NAD and use nucleosides and TCA cycle products as carbon sources, nucleotides as NAD precursors and amino acids as TCA cycle product derivatives get accumulated which by acting as ROS scavengers protects the proteins from oxidative damage (Ghosal et al., 2005). This is mediated by complex formation of divalent manganese (Mn+2) ion which tends to get accumulate during oxidative stress (Daly et al., 2010). PolyP are polyphosphate granules found in the nucleoid region of *D. radiodurans* (Eltsov & Dubochet, 2005). PolyP can be used as an energy source for ATP synthesis, as a phosphorylating agent for several biomolecules, for bacterial Ca^{+2}-induced transformation, for pathogen virulence, and as a growth regulator (Slade & Radman, 2011). *D. radiodurans* has a one-of-a-kind genome that has numerous genes with putative DNA-repair functions, and several unexpected traits have been identified and provisionally characterised as a result of the genome analysis of *D. radiodurans*; also radiation resistance trait is attributed to the redundancy of DNA repair genes in its genome (Sadraeian & Molaee, 2009).

D. radiodurans also expresses toluene dioxygenase responsible for oxidizing toluene, chlorobenzene, 3,4-di-chloro-1-butene and indole and Minton et.al. (1998) showed how engineering recombinant *D. radiodurans* can enhance their capability to degrade organopollutants in radiation-contaminated environment (Lange,

Wackett, Minton, & Daly, 1998). Rao et.al. in 2006 constructed a recombinant R1 strain of *D.radiodurans* by incorporating a gene encoding a nonspecific acid phosphatase (phoN), obtained from a local isolate of *Salmonella enterica* serovar Typhi and efficiently precipitated over 90% of the uranium from uranyl nitrate solution within 6 h (Appukuttan, Rao, & Apte, 2006). Despite Deinococci's strong resistance to many stressors including radiation, desiccation, oxidising agents, and DNA intercalating agents, the most well-studied strain, *D. radiodurans* R1 (DR1), is very vulnerable to some radionuclides and heavy metals like Cd, Co, and Pb (Ruggiero et al., 2005). However, Archana and Chaturvedi (2014) engineered DR1 strain for cytosolic expression of synthetic phytochelatin and bacterial metallothionein genes that enhanced tolerance and bioaccumulation of cadmium (Chaturvedi & Archana, 2014).

Proteomic analysis of *Deinococcus radiodurans* revealed changes in protein expression involved in inorganic ion transport and metabolism, nucleotide transport and metabolism, transcription, translation, replication, recombination and repair, post-translational modification, protein turnover, chaperones out of which“conserved hypothetical protein” with gene no. DR_2623 (UniProt ID : Q9RR74) was found to be upregulated, of which no detailed information regarding function and structure is available in UniProt and in PDB (Zhang et al., 2005).

In this work, considering the lack of structural and functional characterisation of the conserved hypothetical protein Q9RR74, 3D structure is modeled, and evaluated and its function is predicted

2. Methods

2.1 Modeling of Q9RR74 using ROBETTA tool

The server employs a structure prediction method that generates a model for input protein sequence regardless of sequence homology to known structures and Robetta uses comparative modelling to create models for domains that have similarity to proteins with known structure and the Rosetta de novo structure prediction method to create models for domains that don't have any sequence homology.

2.2 Evaluation of the predicted structure

2.2.1 Verify3D

It helps in validating the corroboration between 3D structure of the protein and its own amino acid sequence by converting the

3D environment of each residue into 1D profile and comparing it to known structures (Bowie, Lüthy, & Eisenberg, 1991).

2.2.2 ERRAT

ERRAT employs a statistical method to identify non-bonded atomic interactions within the predicted model by computing an error function for a 9-residue windows through the protein sequence and compares it with non-bonded atomic interactions of high-quality structures (Colovos & Yeates, 1993).

2.2.3 PROCHECK

It analyses residue-by-residue geometry and overall structure geometry to assess the stereochemical quality of the predicted structure (Laskowski, MacArthur, Moss, & Thornton, 1993).

2.3 Function prediction of Q9RR74 using the PredictProtein server

PredictProtein is a web-based tool that helps to predict the structure and function of a protein. Users just need to give an input sequence, and results are obtained from database comparisons and prediction methods. The output contains likely homologous protein, secondary structure regions, PROSITE motifs, and other aspects of protein structure such as solvent accessibility to different residues and functional annotation (Rost, Yachdav, & Liu, 2004).

3. Results

3.1 Modeled 3D structure of Q9RR74 using ROBETTA server

ROBETTA modelled five structures out of which the best was selected with the least angstrom error estimate, with a predicted GDT of 0.88. GDT of 1.0 shows good structure. In Figure 1(a) image of the best-modelled structure of protein Q9RR74 by ROBETTA and in figure (b) angstrom error estimate of each residue is given.

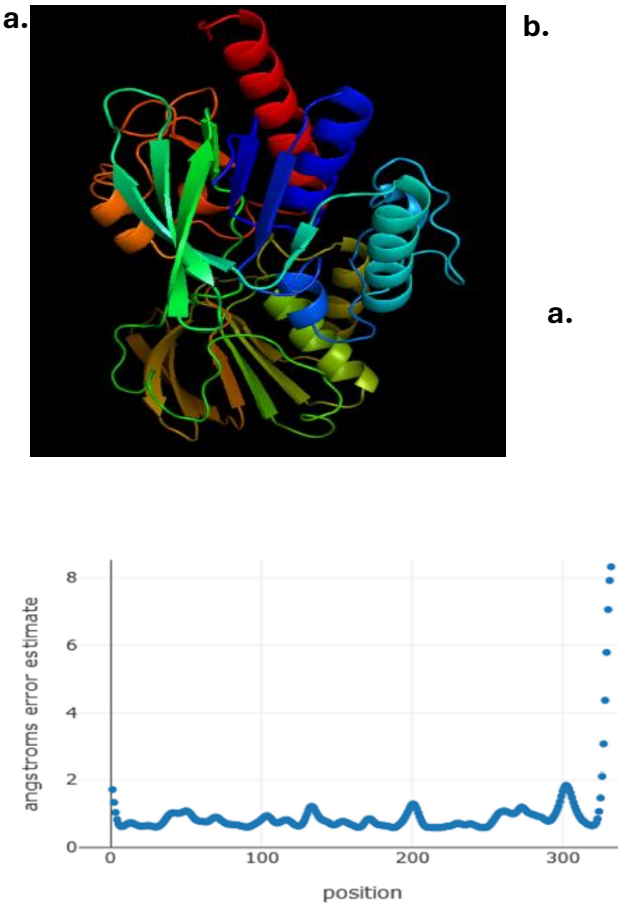
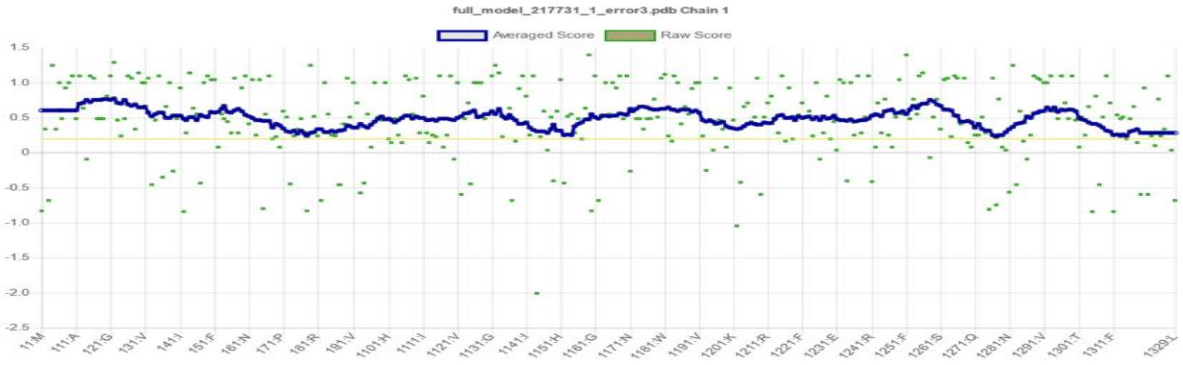


Fig.1. (a) Modeled structure of Q9RR74 (b) Angstrom error estimate plot of the modelled structure of Q9RR74.

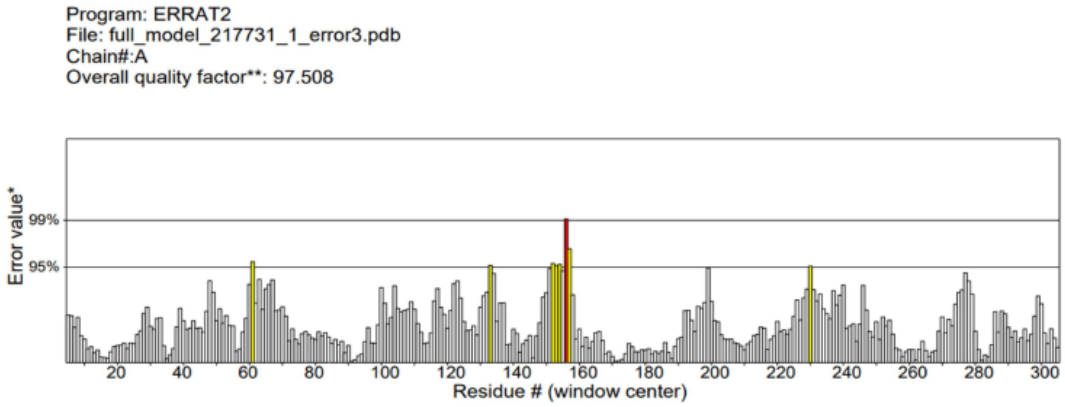
3.2 Structural evaluation of the modelled structure

On structural quality assessment of the model, Verify3D showed that 100.00% of the residues have averaged 3D-1D score >= 0.2 which shows the compatibility of the 3D structure of the predicted protein model by ROBETTA with its amino acid sequence (Fig.2a). Overall quality factor was 97.5078 according to ERRAT (Fig.2b). Ramachandran Plot of model3 showed that 90.4% of total residues fall under most favored region according to PROCHECK (Fig.2c).

In Fig.2c, A, B, L are the residues falling falling under most favoured region and that constitutes of 253 residues or about 90.4% of total residues. Residues falling in additional allowed region is shown by a, b, l, p and that constitutes of 25 residues or about 8.9% of total residues. There are no residues falling under generously allowed region and 2 residues or about 0.7% residues fall under disallowed region. Glycine residues are shown in triangles.



b.



c.

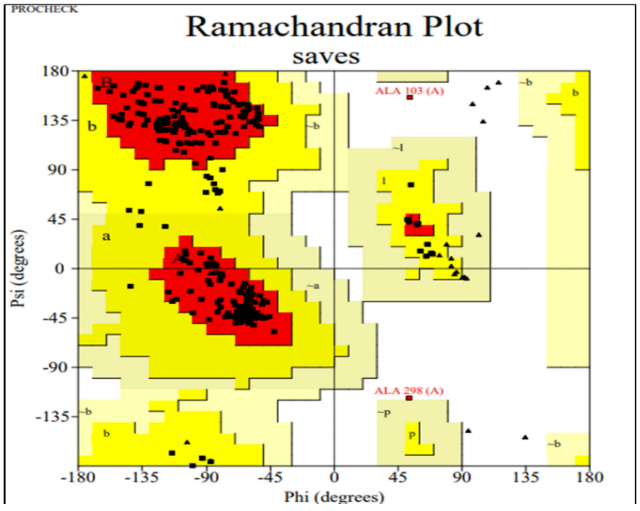


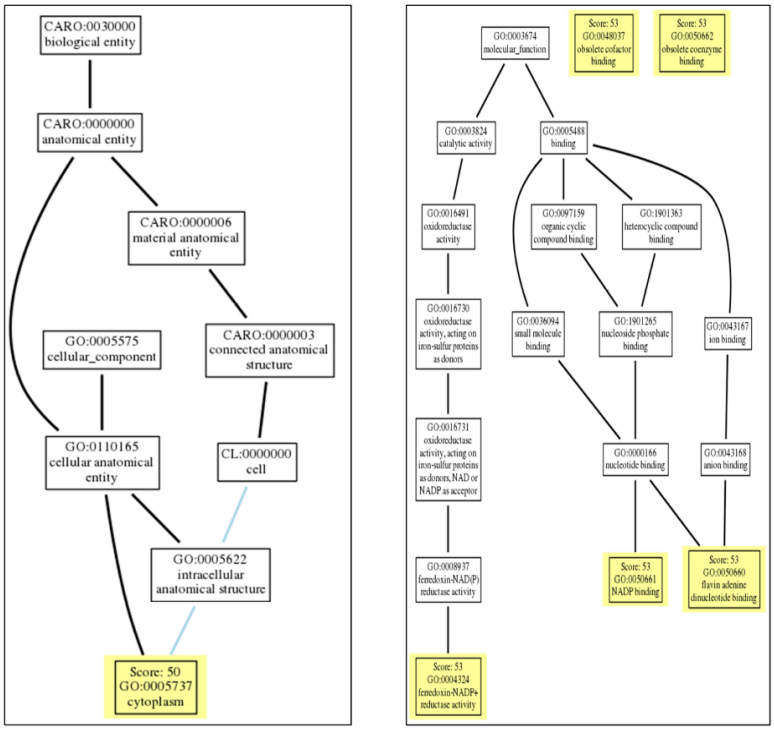
Fig.2. (a) Verify3D results showing the compatibility of the 3D structure of the predicted protein model by ROBETTA with its amino acid sequence (b)ERRAT results depicting the quality factor of the modelled structure (c) Ramachandran plot of the modelled structure.

Functional annotation using Predict Protein gave results in terms of three ontologies: Molecular Function Ontology, Biological Process Ontology and Cellular Component Ontology.

Each predicted term is associated with a reliability (0-100), indicating the distance of the protein used for annotation transfer. In the below table, results of function prediction of protein Q9RR74 is tabulated in fig3.a

The results of PredictProtein depicts the probable the function of protein Q9RR74 which is that it might be involved in biological oxidation-reduction process occurring in cytoplasm of Deinococcus radiodurans and probable molecular function might involve Ferredoxin-NADP+ reductase activity and Flavin adenine dinucleotide binding.

Ontology	**GO ID**	**GO Term**	**Reliability (%)**
Biological Process	GO:0055114	Oxidation-Reduction Process	53
Cellular Component	GO:0005737	Cytoplasm	50
Molecular Function	GO:0048037	Obsolete cofactor binding	53
	GO:0050661	NADP binding	53
	GO:0050662	Obsolete coenzyme binding	53
	GO:0004324	Ferredoxin-$NADP^+$ reductase activity	53
	GO:0050660	Flavin adenine dinucleotide binding	53



4. Discussion

Models constructed of protein by ab-initio approach reliability depends upon how precise the algorithm is to build protein from scratch without having any templates. From the results obtained, the ROBETTA tool gave the best model of protein Q9RR74 and was found to be reliable based upon structural quality assessment with an overall quality factor of 97.508 however, further confirmation is required. Structural refinement of the predicted structure would help in improving the overall quality of the structure. To evaluate the stability of the modeled structure, molecular dynamics simulation is crucial for which GROMACS can be used. PredictProtein helps in predicting probable protein function but it depends upon whether the amino acid sequence corresponds to any other known protein in the database with which it is being compared. The predicted function of protein Q9RR74 is that it might be involved in the biological oxidation-reduction process. This would help to study its role in a targeted manner in experiments for various redox reactions occurring in bioremediation processes and would further help in confirming its role.

References

Appukuttan, D., Rao, A. S., & Apte, S. K. (2006). Engineering of Deinococcus radiodurans R1 for bioprecipitation of uranium from dilute nuclear waste. *Applied and Environmental Microbiology, 72*(12), 7873-7878.

Bowie, J. U., Lüthy, R., & Eisenberg, D. (1991). A method to identify protein sequences that fold into a known three-dimensional structure. *Science, 253*(5016), 164-170.

Chaturvedi, R., & Archana, G. (2014). Cytosolic expression of synthetic phytochelatin and bacterial metallothionein genes in Deinococcus radiodurans R1 for enhanced tolerance and bioaccumulation of cadmium. *Biometals, 27*, 471-482.

Colovos, C., & Yeates, T. O. (1993). Verification of protein structures: patterns of nonbonded atomic interactions. *Protein science, 2*(9), 1511-1519.

Daly, M. J., Gaidamakova, E. K., Matrosova, V. Y., Kiang, J. G., Fukumoto, R., Lee, D.-Y., . . . Levine, R. L. (2010). Small-molecule antioxidant proteome-shields in Deinococcus radiodurans. *PloS one, 5*(9), e12570.

Eltsov, M., & Dubochet, J. (2005). Fine structure of the Deinococcus radiodurans nucleoid revealed by cryoelectron microscopy of vitreous sections. *Journal of bacteriology, 187*(23), 8047-8054.

Ghosal, D., Omelchenko, M. V., Gaidamakova, E. K., Matrosova, V. Y., Vasilenko, A., Venkateswaran, A., . . . Makarova, K. S. (2005). How radiation kills cells: survival of Deinococcus radiodurans and Shewanella oneidensis under oxidative stress. *FEMS microbiology reviews, 29*(2), 361-375.

Lange, C. C., Wackett, L. P., Minton, K. W., & Daly, M. J. (1998). Engineering a recombinant Deinococcus radiodurans for organopollutant degradation in radioactive mixed waste environments. *Nature biotechnology, 16*(10), 929-933.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of applied crystallography, 26*(2), 283-291.

Rost, B., Yachdav, G., & Liu, J. (2004). The PredictProtein server Nucleic Acids Res. *The PredictProtein server Nucleic Acids Res, 32*.

Ruggiero, C. E., Boukhalfa, H., Forsythe, J. H., Lack, J. G., Hersman, L. E., & Neu, M. P. (2005). Actinide and metal toxicity to prospective bioremediation bacteria. *Environmental Microbiology, 7*(1), 88-97.

Sadraeian, M., & Molaee, Z. (2009). *Bioinformatics analyses of*

Deinococcus radiodurans in order to waste clean up. Paper presented at the 2009 second international conference on environmental and computer science.

Slade, D., & Radman, M. (2011). Oxidative stress resistance in Deinococcus radiodurans. *Microbiology and molecular biology reviews, 75*(1), 133-191.

Zhang, C., Wei, J., Zheng, Z., Ying, N., Sheng, D., & Hua, Y. (2005). Proteomic analysis of Deinococcus radiodurans recovering from γ-irradiation. *Proteomics, 5*(1), 138-143.