4-2016

The NMR Solution Structure and Function of RPA3313: A Hypothetical Protein from *R. palustris*

Austin J. Lowe
University of Nebraska-Lincoln, aj.lowe2123@yahoo.com

Jonathan Catazaro
University of Nebraska-Lincoln, jcatazaro@huskers.unl.edu

Cheryl Arrowsmith
University of Toronto

Robert Powers
University of Nebraska-Lincoln, rpowers3@unl.edu

The NMR Solution Structure and Function of RPA3313: A Hypothetical Protein from *R. palustris*

Austin J. Lowe*, Jonathan Catazaro*, Cheryl Arrowsmith and Robert Powers

Department of Chemistry, University of Nebraska-Lincoln, Lincoln, NE

Introduction

Protein function elucidation often relies heavily on amino acid sequence analysis and other bioinformatics approaches. The reliance is further extended to structure homology modeling for ligand docking and protein-protein interaction mapping. However, sequence analysis of RP3313 exposes a large, unannotated class of hypothetical proteins mostly from the *alphaproteobacteria* order (Figure 1). In the absence of sequence and structure information, further functional elucidation of this class of proteins has been significantly hindered. A high quality NMR structure of RP3313 reveals that the protein forms an α/β roll with a possible phosphate binding loop between the first β-strand and the N-terminus of the α-helix. This fold is indicative of a protein involved in the handling or synthesis of nucleic acid chains. Results of a mass spectrometry proteomic analysis strongly point toward interaction with the ribosome and its subunits. The combined structural and proteomic analyses suggest that RP3313 by itself or in a larger complex may assist in the transportation of substrates to the ribosome for further processing.

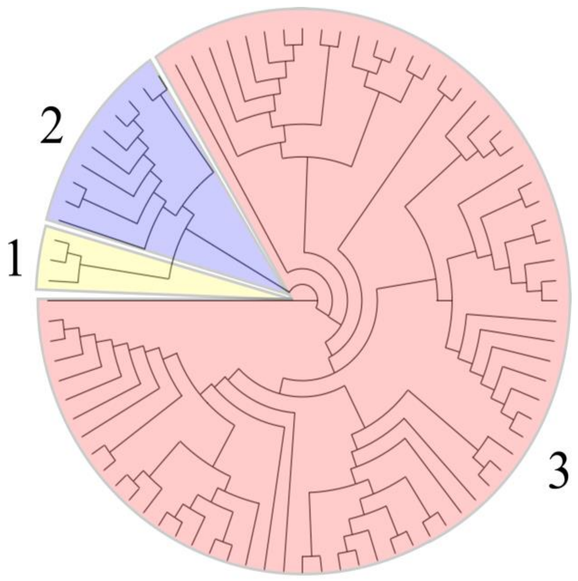
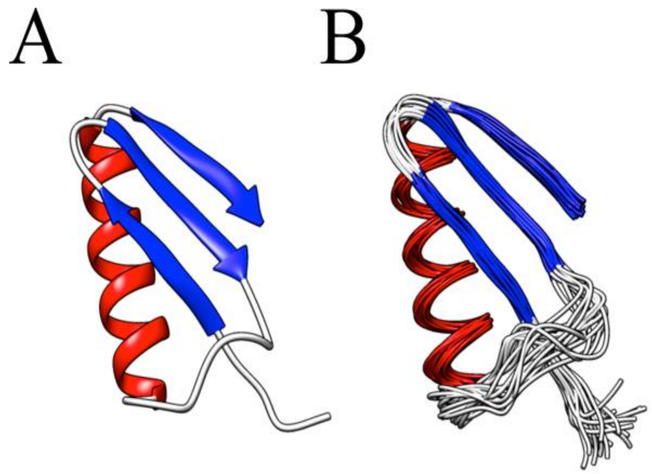


Figure 1. A neighbor-join tree of the protein BLAST results of RPA3313 against non-redundant protein sequences. All of the sequence hits belong the *alphaproteobacteria* order and have identities >32%. Groups 1, 2, and 3 are dominated by the genera *Beijerinckia*, *Rhodopseudomonas*, and *Bradyrhizobium*, respectively. The tree highlights the fact that this protein is a member of an unannotated and a structurally uncharacterized class of proteins.

Results



C

PSVS Z-score	Ensemble Score
Verify3D	-1.93
ProsaII	-0.58
Procheck (phi-psi)	0.12
Procheck (all)	-0.71
MolProbity	-0.45
Ramachandran space	
most favored	98.30%
allowed	1.70%
disallowed	0.00%

Figure 2. (A) The average structure of the ensemble. The structures are colored according to secondary structure: red for α-helix, blue for β-strand, and white for loops and disordered regions. (B) An ensemble of the 20 best water-refined structures of RPA3313.. (C) A table of results compiled from PSVS[2] analysis.

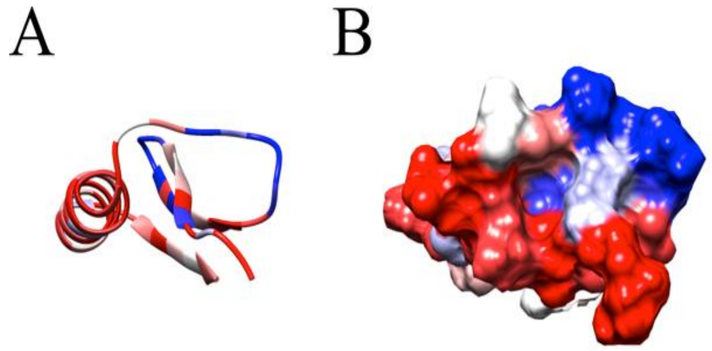


Figure 3. ConSurf Analysis. (A) ConSurf per residue scores for the structure RPA3313. Conserved residues are red, unconserved are blue. (B) ConSurf scores mapped onto the molecular surface representation. A conserved pocket is formed between the N-terminus of the α-helix and the first β-strand. Residues found to be less conserved are mostly found in the loop regions and the disordered tail (C-terminus, not shown).

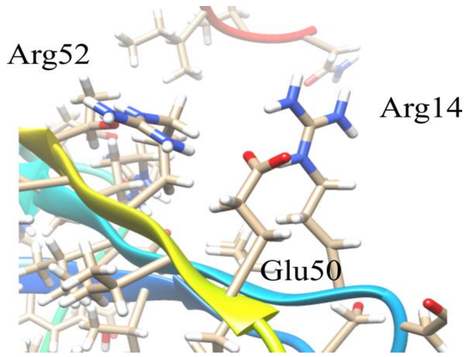


Figure 4. Arginine stacking interaction on the surface of RPA3313. Arginines 14 and 52 interact across the top of the β-sheet and are stabilized by glutamate 50. Each of the residues is evolutionarily stable indicating the occurrence of a conserved interaction.

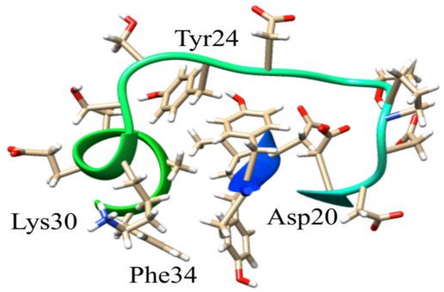


Figure 5. Surface pocket and possible ligand binding site on RPA3313. Shown are the residues with side chains that point into the pocket of the protein. The evolutionarily conserved residues are labeled. The remaining residues may also participate in the function of the pocket even though they are not highly conserved.

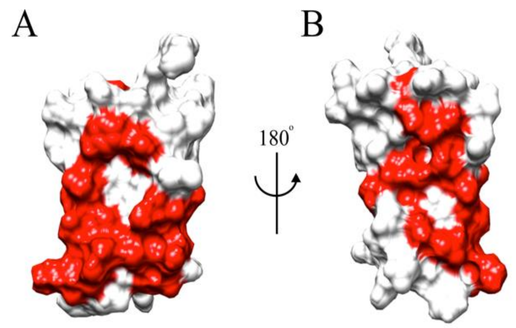


Figure 6. cons-PPISP predictions. (A) Front and (B) back orientations of RPA3313 with predicted protein interaction sites colored red.

Conclusions

Rhodopseudomonas palustris is a metabolically versatile organism which has the unique ability to grow both aerobically and anaerobically on a wide variety of carbon sources including carbon dioxide. RP3313, an unknown protein identified during the sequencing of its genome, is conserved across *alphaproteobacteria* and remains both structurally and functionally unannotated. The NMR solution structure reveals that the protein adopts a left handed α/β roll with a domain conserved binding pocket. Above the binding pocket exists a possible phosphate binding site, which is also comprised of conserved residues. *In silico* docking identifies a low energy confirmation of ATP in the pocket and ^{15}N NMR titrations confirm the binding. However, the dissociation constants suggest that ATP may not be the physiological substrate. It is likely that a substrate similar to ATP in size and chemical moieties binds well in the pocket. Crosslinking studies done with *E. coli* and *R. palustris* yielded similar results for protein-protein interactions with RP3296. In both instances, multiple ribosomal subunits were identified by MS/MS sequencing and a MASCOT database search. It appears that RP3313 can associate with the ribosome, but further functional elucidation remains to be completed. It is possible that the protein is only expressed during certain metabolic modes of growth, as this protein is not found in evolutionary distant bacterial species with more limited metabolism.

References

1. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. (1990) "Basic local alignment search tool." J. Mol. Biol. 215:403-410.
2. Bhattacharya, A., Tejero, R., Montelione, G.T. (2006) “Evaluating protein structures determined by structural genomics consortia.” Proteins: Struct. Funct. Bioinf. 66:778-795.
3. Ashkenazy H., Erez E., Martz E., Pupko T., Ben-Tal N. (2010) “ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids.” Nucl. Acids Res. 38:W529-W533.
4. Perkins, D.N., Pappin, D.J.C., Creasy, D.M., Cottrell, J.S. (1999) “Probability-based protein identification by searching sequence databases using mass spectrometry data.” Electrophoresis. 20:3551-3567.