

In Silico Construction of the Carbon Fixation Pathway in *Synechococcus* sp. WH8102

Phuongan Dam^{1,2}, Zhengchang Su^{1,2}, Victor Olman¹ and Ying Xu^{1,2}

¹*Department of Biochemistry and Molecular Biology, University of Georgia at Athens, and*

²*Institute of Computational Biology, Oak Ridge National Laboratory*

phd@csbl.bmb.uga.edu, zhx@csbl.bmb.uga.edu, olman@csbl.bmb.uga.edu and
xyn@csbl.bmb.uga.edu

Abstract

*Because the carbon fixation pathway plays an essential role in the primary production and natural carbon recycling process, and because the genome of the marine cyanobacterial *Synechococcus* sp. WH8102 (SYNWH8102) was recently sequenced, SYNWH8102 was chosen to further our understanding of the interaction and regulation of the carbon fixation pathway at the molecular level. In this abstract, we present the predicted carbon fixation pathway in SYNWH8102 as a result of our recently developed computational protocol for inference of regulatory and signaling pathways. The results of our pathway prediction include: (a) Major components of the carbon fixation pathway reported in the literature are present in SYNWH8102. (b) Approximately, 48 new candidates are added into the network from the results of the pathway expansion step. (c) Additionally, our in-house motif finding program, CUBIC, found several motifs that are present in the promoter regions of multiple genes involved in this pathway, suggesting that these genes are transcriptionally co-regulated.*

1. Introduction

Currently, very little is known about the particular pathway in this bacterial strain although this pathway is partially studied in related organisms including *Synechococcus* sp. PCC7942, *Synechocystis* sp. PCC6803 and the marine *Synechococcus* sp. PCC7002 [1-4]. From previous studies, it is generally known that the carbon fixation pathway in cyanobacteria has the following key components. (a) A two-component sensor/ regulator system, *regA/ regB*, was reported to be critical for the expression of multiple carbon fixation genes. (b) Another transcription regulator,

cbbR, is also shown to modulate the expression of genes in the pathway. (c) Various CO₂ and HCO₃⁻ transporters are reported to be responsible for low and high affinity flux of the substrates. (d) Carbon fixation is catalyzed inside the carboxysome. (e) Enzymes involving in the pathway include rubisco, carbonic anhydrase and enzymes functioning in the Calvin-Benson cycle [1-4].

The pathway reconstruction protocol developed in our laboratory was used to build the carbon fixation pathway in SYNWH8102 [5]. Using multiple data mining tools, our pathway model was expanded, refined and evaluated after the initial mapping of multiple partial pathways in related organisms to SYNWH8102.

2. Method

The protocol for pathway modeling was previously reported [5]. In short, this procedure consists of the following key components: (a) construction of template pathways from related microbial organisms and mapping the initial templates to the target genome, (b) refinement and expansion of the mapped pathway and (c) validation and evaluation of the pathway models using experimental data or other information. During the refinement and expansion step, data mining tools including operon prediction [6], phylogenetic profile analysis, inference of protein-protein interactions, and binding site prediction program were used [7].

3. Results and discussions

3.1. Construction of a template pathway

Because the carbon fixation pathway in several organisms has been partially studied, we used these

partial pathways to build the template pathway. The initial template includes 20 genes that play various roles in the carbon fixation pathway from transporters of bicarbonate, enzymes in the pathway to transcriptional regulator, *cbbR*. From orthologous mapping, 18 orthologs were found in SYNWH8102, and this set of original orthologs was used for further expansion and refinement of the model. The result is at www.csbl.bmb.uga/~phd/carbonfixation.html.

3.2. Pathway expansion and refinement

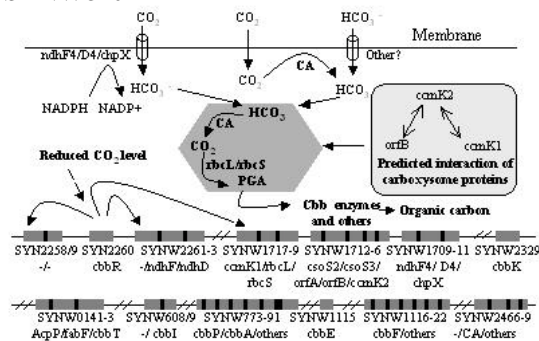
Based on our operon prediction, phylogenetic profile analysis and inference of protein-protein interactions, 48 new candidates are suggested to be functionally involved in the carbon fixation pathway. Among those, SYNW162, 280 and 2331 are more interesting because they have similar phylogenetic profile to original orthologs, and are predicted to interact with original orthologs in the pathway. The result is at www.csbl.bmb.uga/~phd/carbonfixation.html.

After using CUBIC [7] to search for common motifs in the promoter regions of *cbbR* orthologs of ten cyanobacteria genomes, the motifs were ranked based on the informational content score and the resemblance of the motifs to the characteristics of LysR members because CbbR is a member of the LysR family. The most relevant motif contains an inverted repeat of 3 nucleotides connected by a spacer of 13 nucleotides. The same motif was also identified in the upstream regions of 36 other SYNWH8102 orfs. Many of them are homologs of genes involving in photosynthesis or transporting inorganic molecules.

The result pathway is shown in Figure 1. For clarification purpose, SYNWH8102 orfs added to the pathway during the expansion and refinement step were omitted. However, these orfs are listed at www.csbl.bmb.uga/~phd/carbonfixation.html. Compared to the template pathway, the predicted pathway also contains similar components. (a) The transcriptional regulator, CbbR, is predicted to regulate gene transcription of the carboxysome protein CcmK1, two critical enzymes: rubisco and phosphoribulokinase, and two other operons. In *Synechocystis*, *cbbR* homolog, *sll1594*, was reported to affect transcription of an *ndh* operon, supporting our result of the *ndh* operon (SYNW2261-63) being regulated by CbbR [8]. (b) The putative CO₂/HCO₃⁻ transporters (SYNW1709-11) is probably responsible for low affinity flux of substrates, and is not regulated by CbbR. (c) Protein components that form the carboxysome are predicted to interact through our protein-protein interaction map. (d) All enzymes that involve in the conversion of inorganic CO₂ into

organic carbon are present in SYNWH8102. However, the *regA/regB* system was not found in SYNWH8102, suggesting a different mode of transcriptional regulation present in SYNWH8102.

Figure 1. Predicted carbon fixation pathway in SYNWH8102



4. Acknowledgment

The work was supported by DOE office of Biological/Environmental Research, Genome to Life Project, "Carbon Sequestration in *Synechococcus* sp.: From Molecular Machines to Hierarchical Modeling."

5. References

1. Badger MR, Price GD: "CO₂ concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution". *J Exp Bot* 2003, 54(383):609-622.
2. Gibson JL, Tabita FR: "The molecular regulation of the reductive pentose phosphate pathway in Proteobacteria and Cyanobacteria". *Arch Microbiol* 1996, 166(3):141-150.
3. Tabita FR: "Molecular and cellular regulation of autotrophic carbon dioxide fixation in microorganisms". *Microbiol Rev* 1988, 52(2):155-189.
4. Gibson JL, Tabita FR: "Localization and mapping of CO₂ fixation genes within two gene clusters in *Rhodobacter sphaeroides*". *J Bacteriol* 1988, 170(5):2153-2158.
5. Su Z, P.Dam, Chen X, Olman V, T.Jiang, B.Palenik, Y.Xu: "Computational Inference of Regulatory Pathways in Microbes: An application to phosphorus assimilation pathways in *Synechococcus* WH8102". *Genome Informatics* 2003, 14:3-13.
6. Chen X, Su Z, Dam P, Palenik B, Xu Y, Jiang T: "Operon prediction by comparative genomics: an application to the *Synechococcus* sp. WH8102 genome". *Nucleic Acids Res* 2004, 32(7):2147-2157.
7. Olman V, Xu D, Xu Y: "Identification of regulatory binding sites using minimum spanning trees". *Pac Symp Biocomput* 2003:327-338.
8. Figge RM, Cassier-Chauvat C, Chauvat F, Cerff R: "Characterization and analysis of an NAD(P)H dehydrogenase transcriptional regulator critical for the survival of cyanobacteria facing inorganic carbon starvation and osmotic stress". *Mol Microbiol* 2001, 39(2):455-468.