

# Computational Construction of Nitrogen Assimilation Pathway in Cyanobacteria *Synechococcus* sp. WH8102

Z. Su<sup>1</sup>, P. Dam<sup>1</sup>, X. Chen<sup>2</sup>, V. Olman<sup>1</sup>, T. Jiang<sup>2</sup> and Y. Xu<sup>1</sup>

<sup>1</sup>Department of Biochemistry and Molecular Biology, University of Georgia, and Institute of Computational Biology, Oak Ridge National Laboratory; <sup>2</sup>Department of Computer Science and Engineering, University of California at Riverside

[xyn@bmb.uga.edu](mailto:xyn@bmb.uga.edu)

## Abstract

*One of the challenging problems in post-genomic era is to computationally infer the regulatory pathways of a less-studied organism. We have previously developed a computational protocol for inference of regulatory and signaling pathways in less-studied microbes, through mining high-throughput biological data of various types. Using a similar protocol we have constructed the nitrogen assimilation pathway in cyanobacteria *Synechococcus* sp. WH8102. A considerable amount of high confident knowledge of this pathway were derived.*

## 1. Introduction

*Synechococcus* WH8102 (WH8102) is a major cyanobacterial strain living in a wide range of oceanographic areas, which contributes to a great portion of global CO<sub>2</sub> fixation [2]. Therefore, a better understanding of this organism might help solve some environmental problems. Unfortunately, our current knowledge of this organism is very limited, because little experiments have been carried out on it, even though its genome has recently been sequenced [1]. With the availability of many high throughput data as well as the huge experimental data in the literature from the related organisms, however, it is now possible to gain a quick understanding of this organism by pure computational studies. We have previously developed a computational protocol for inferring pathways by mining various types of data [3]. We now applied the protocol to constructing the nitrogen assimilation pathway in this organism, which is highly related to the CO<sub>2</sub> fixation mechanism in cyanobacteria.

## 2. Methods

**2.1. Mapping known nitrogen assimilation pathways in the other closely related organisms to the WH8102 genome.** We first constructed nitrogen assimilation pathways in six closely related cyanobacteria: *Anabaena* sp PCC 7120, *Synechocystis* sp PCC 6803, *Synechococcus* sp PPC 7492,

*Synechococcus* sp PPC 8801, *Synechococcus* sp WH8705 and *Synechococcus* sp WH7002, since this pathway in these organisms is relatively well studied experimentally. We then mapped these template pathways into WH8102 genome by PMAP program (described in these proceedings), and created a primary nitrogen assimilation pathway of WH8102.

**2.2. Expansion of the primary pathway.** We used three sources of data to expand the above-constructed primary pathway. (1) We recruited proteins into the primary pathway by predicting physical interaction partners of the proteins in the primary pathway. The rationale of the recruitment is simply the rule of "guilty by association". (2) We added proteins to the pathway by phylogenetic profile analysis, even though we do not know the exact positions of these added proteins in the pathway. (3) The DNA binding domain of the global nitrogen control regulator *ntcA* are highly conserved among cyanobacteria, suggesting that the *ntcA* binding sites are also conserved among these species. The intergenic regions of orthologous genes that are known being regulated by *ntcA* were used to find the putative consensus *ntcA* binding sites and the associated profiles. Putative *ntcA* binding sites were then predicted by scanning the intergenic regions of all ORFs of WH8102 with these profiles. Genes with a high score of binding site in their regulatory region are predicted to be members of *ntcA* regulon and thus are involved in the pathway.

## 3. Results and Discussion

1. The primary pathway constructed for WH8102 contains the essential components required for up taking and metabolizing ammonium, nitrate/nitrite, and urea, suggesting that it could utilize these compounds as nitrogen source. However, it lacks the nitrogen fixation machinery.

2. We have identified a pseudo palindromic motif GTAnnnnnnnTAC (Figure 1) as the putative consensus *ntcA* binding site. Scanning the regulatory regions of WH8102 using the profile associated with this motif, predicted more than 200 genes having a highly similar motif in their regulator regions. Protein-protein interaction predictions and phylogenetic profile analysis recruited 18, and 257 proteins to the primary pathway, respectively.

3. A candidate set of proteins that might involved in the coupling between CO<sub>2</sub> fixation and nitrogen assimilation regulation was suggested based on the protein-protein interactions predictions and phylogenetic profile analysis. These proteins are sought by biologists for decades, narrowing down the candidates to an amenable number might provide a great chance for experimentalists to solve this mystery quickly.

4. A few of regulatory proteins are predicted to be involved in the nitrogen assimilation pathway. They are promising targets for experimentalists to further elucidate this important pathway.

5. We have constructed a working model for the nitrogen assimilation pathway in WH8102 as shown in Figure 2.

### Acknowledgements:

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."

### References

- [1] Palenik B., Brahamsha B., Larimer F. W., Land M. and et.al, The genome of a motile marine Synechococcus, *Nature*, 2003, 424: pp. 1037-1042.
- [2] Palenik B and Dyhrman ST, Recent progress in understanding the regulation of marine primary productivity by phosphorus, In: Phosphorus in Plant Biology: Regulatory Roles in Molecular, Cellular, Organismic, and Ecosystem Processes, edited by J.D.J.P.Lynch. American Society of Plant Physiologists, 1998.
- [3] Su Z., P.Dam, X Chen, V Olman and et.al, Computational Inference of Regulatory Pathways in Microbes : An application to phosphorus assimilation pathways in *Synechococcus* WH8102, *Genome Informatics*, 2003, 14: pp. 3-13.

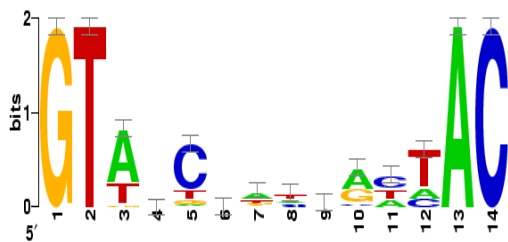


Figure 1. The predicted consensus ntcA binding site founded from the regulatory regions of orthologous genes known being regulated by ntcA in eight cyanobacteria species.

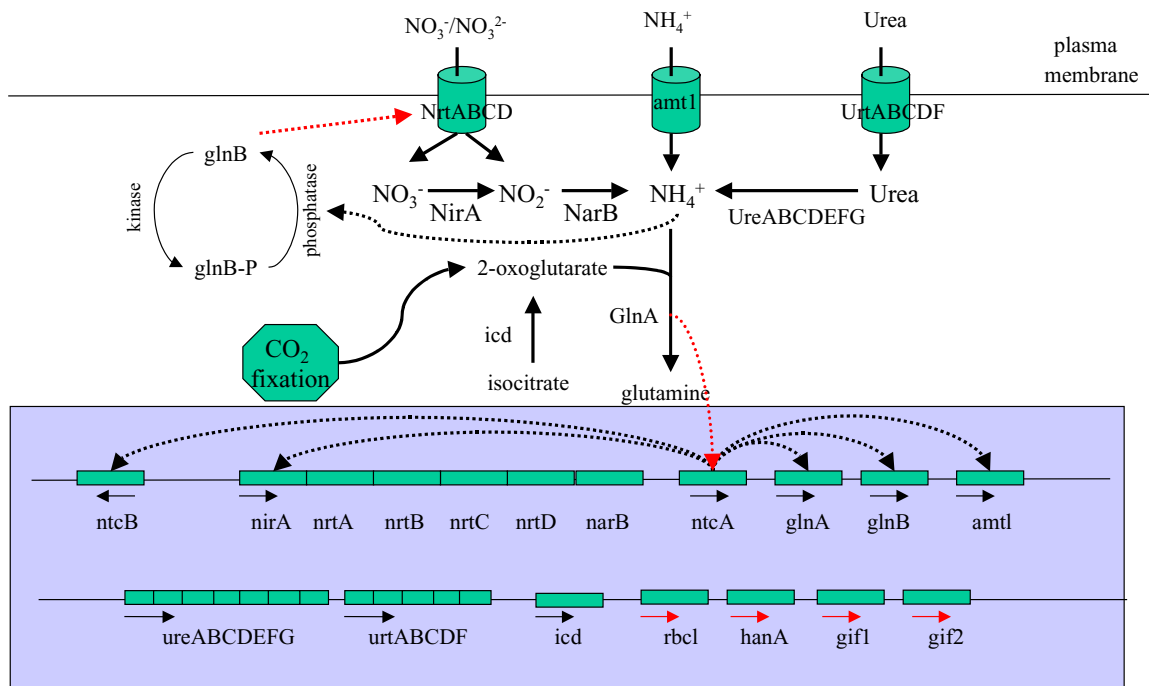


Figure 2. The working model of nitrogen assimilation pathway in WH8102