

Bayl

June 2003

TO: U.S. Department of Energy
Office of Science

RE: **Special Report**

PROJECT TITLE: An integrative approach to energy, carbon, and redox metabolism in the cyanobacterium *Synechocystis* sp. PCC 6803

PRINCIPAL INVESTIGATOR: Ross Overbeek
(630)-567-7677
ross@.....

RECIPIENT
ORGANIZATION: Integrated Genomics, Inc.
2201 West Campbell Park Drive
Chicago, Illinois 60612

DOE AWARD NUMBER: DE-FG02-02ER15349

PROJECT PERIOD: from 09/30/2002 through 09/29/2004

BUDGET PERIOD
COVERED BY THIS REPORT: from 09/30/2002 through 09/29/2003

DOE Patent Clearance Granted

Mark P. Dvorscak
Mark P. Dvorscak
(630) 252-2393
E-mail: mark.dvorscak@ch.doe.gov
Office of Intellectual Property Law
DOE Chicago Operations Office

09/16/2004
Date

Signature of Principal Investigator

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

Special Report

The accompanying Progress Report covers the period from 9/30/2002 to 9/29/2003 for the above-referenced grant. Please note that this Report, originally scheduled as a 9-month progress report, establishes that Integrated Genomics, Inc. ("IG") has successfully completed the entire initial-year goals three months early, by July 2003. The early completion of the first-year requirements was facilitated by management changes at IG.

Specifically, in March 2003, ownership changes at IG resulted in replacement of the company's top-level management. Concurrent with those changes, Ross Overbeek (the named PI) resigned from the company. In light of Dr. Overbeek's resignation the company has decided not to continue work on *Synechocystis* 6803.

In order to meet grant obligations to DOE, Dr. Overbeek and IG intensified efforts to complete the first-year goals for delivery by July 2003. Accordingly, the percentage of work-time by all grant participants in May and June of 2003 has been significantly increased: Dr. Gerdes devoted 100% of her time to grant related tasks; while Dr. Ivanova, V. Joukov and T. Walunas each devoted from 60% to 80% of their time to the grant. An additional software designer, Yuri Grechkin, was also added to the project team. These efforts have allowed us to accomplish our first-year goals--i.e. qualitative metabolic reconstruction and compilation of a formally-encoded reaction network of *Synechocystis* 6803. See accompanying Progress Report for the details

June 2003

TO: U.S. Department of Energy
Office of Science

RE: **Progress Report**

PROJECT TITLE: An integrative approach to energy, carbon, and redox
metabolism in the cyanobacterium *Synechocystis* sp. PCC 6803

PRINCIPAL INVESTIGATOR: Ross Overbeek
(630)-567-7677

RECIPIENT
ORGANIZATION: Integrated Genomics, Inc.
2201 West Campbell Park Drive
Chicago, Illinois 60612

DOE AWARD NUMBER: DE-FG02-02ER15349

PROJECT PERIOD: from 09/30/2002 through 09/29/2004

BUDGET PERIOD
COVERED BY THIS REPORT: from 09/30/2002 through 09/29/2003

Signature of Principal Investigator

Svetlana Gerdes (co-PI)

NOTE:

This project is a part of a collaborative research effort involving three other groups lead by:
Willem F.J. Vermaas (Arizona State University) coordinating PI
Julian Whitelegge and Kym Faull (University of California, Los Angeles)
Robert W. Roberson (Arizona State University)

Progress Report

Integrated Genomics, Inc. (IG) has begun work on this GTL Project in October 2002. The main objectives for the first year were:

- I. To produce a detailed metabolic reconstruction of *Synechocystis* sp. PCC 6803 especially in the interrelated areas of photosynthesis, respiration, and central carbon metabolism to support a more complete understanding and modeling of this organism. Provide a web-based access to the intermediate results to our collaborators for joint curation and refinement, and for projection and interpretation of experimental data. Ultimately deliver these results to the whole scientific community via specialized web-based resource and publications.
- II. To provide detailed bioinformatic analysis of selected functional systems related to carbon and energy generation and utilization, and of the corresponding pathways, functional roles and individual genes in order to support wet lab experimental projects of our collaborators on this grant.

The completion of these tasks initially planned for a year (from 9/30/2002 to 9/29/2003) was achieved ahead of schedule - by July 2003. See accompanying Special Report for explanation.

Specific aim I.

I.1. Qualitative metabolic reconstruction of Synechocystis sp. PCC 6803. A draft metabolic reconstruction of *Synechocystis* sp. PCC 6803 has been automatically generated in ERGOTM based on comparative analysis of several hundred microbial genomes. The draft reconstruction was further subjected to detailed manual curation by a team of bioinformaticians, focusing on the functional systems related to carbon generation and utilization.

Briefly, this work included: i) refinement of automatic functional annotations, ii) detailed bioinformatics analysis of "unknown" ORFs using unique ERGO software tools, as well as publicly available programs in order to adequately predict their functions, iii) building a collection of metabolic and non-metabolic pathways expected to be present in *Synechocystis* sp. PCC 6803 based on analysis its genome sequence, iv) manual connection of ORFs with the corresponding pathways (see (3) for detailed description of metabolic reconstruction technology utilized at IG).

The following results have been produced to date:

- In the course of this work functional roles were suggested for a much higher number of *Synechocystis* 6803 ORFs (64%, or 2101 out of 3264 total ORFs, see Table 1), than it was previously available for the cyanobacterial community through Cyanobase (<http://www.kazusa.or.jp/cyanobase/Synechocystis/>), its main genomics resource, where only 51% (1673 out of 3264) ORFs have specific functional annotations;
- Dozens of new pathways were added to ERGO collection (see Table 1), including composition of cyanobacterial respiratory and photosynthetic membrane complexes, inorganic carbon concentration and fixation mechanisms, several pigment and cofactor biosynthetic pathways (phycobilins, carotenoids, chlorophyll, tocopherols, etc.);

- A complete draft has been obtained for the following functional systems: central carbon metabolism, CO₂ fixation, respiratory and photosynthetic electron transport, biosynthesis of photosynthetic pigments, amino acid biosynthesis, major vitamins and cofactor biosynthesis, and membrane transport;
- The up-to-date status of the reconstruction is shown in Table 1;

The results of the *Synechocystis* 6803 genome analysis performed by Integrated Genomics are available to the whole scientific community for browsing, analysis and display through ERGO Light platform. This resource is freely available to the research community at light (<http://www.ergo-light.com/ERGO/>) and currently includes 7 genomes: *Bacillus cereus* ATCC 14579, *Bacillus anthracis* A2012 (public TIGR genome), *Bacillus thuringiensis israelensis* ATCC 35646, *Brucella melitensis* 16M, *Fusobacterium nucleatum* ATCC 25586, and *Fusobacterium nucleatum vincentii* ATCC 49256. ERGOTM Light will be one of the major conduits at IG for the distribution of data obtained from publicly funded projects.

Features of ERGOTM Light include:

- Access to IG metabolic and non-metabolic pathway database;
- A user-friendly graphical interface
- Manually curated annotations
- Chromosomal clustering data
- Detailed information about all genes
- Connectivity to other public databases, including KEGG, NCBI, SwissProt.

The qualitative metabolic reconstruction in its current state has been made publicly available at: http://www.ergo-light.com/RN/CY_Tree.html (user ID: PCC6803, password: DOE_GTL). This specialized resource is providing a hierarchical presentation of metabolic systems and pathways hyperlinked to ERGO Light at the level of pathways and individual genes (for additional information and examples, see Appendix A)

Further improvement and updates of the reconstruction will be performed during the second year of the project. We expect the input from our collaborators as well as from the broader scientific community to aid in further improvement of the quality of ORF calling, functional annotations, and reaction network in *Synechocystis* 6803 reconstruction. For instance, as a part of the current project an effort to improve predictions of ORF starting codons in *Synechocystis* 6803 has been initiated by Ross Overbeek in collaboration with Dr. Lee Ann McCue (Harvard University, Cambridge, MA).

Table 1. The current status of the *Synechocystis* sp. PCC 6803 metabolic reconstruction

Category	Counts: % of Total:	
ORFs total	3,264	100.00%
ORFs with assigned function	2,101	64.37
ORFs without assigned function	1,163	35.63
ORFs without function or similarity	228	6.95

ORFs without function, with similarity	935	28.65
ORFs in asserted pathways	1,634	50.06
ORFs not in asserted pathways	1,630	49.94
ORFs with assigned function but no pathway	468	14.34
ORFs in the functional overview	1,606	49.20
Functions assigned	1,428	100.00%
Functions assigned, hypothetical	19	1.33
Functions assigned, connected to asserted pathways	1,049	73.46
Functions assigned, not connected to asserted pathways	379	26.54
Functions missing from asserted pathways	121	
Functions with no sequence	12	
Pathways asserted total	1,037	100.00

I.2. Converting metabolic reconstruction into a mathematical model of Synechocystis sp. PCC 6803 metabolism. A qualitative metabolic reconstruction above is a first mandatory step towards mathematical modeling of cellular metabolism. Next step is development of formally encoded reaction network suitable for modeling. However, the following problems are common for most available network encodings (both at IG and in any public integrations), which preclude a possibility of building useful quantitative models:

1. redundancy and imprecision of compound names (synonyms and typos);
2. inconsistent formula representation (such as salt versus free acid etc);
3. lack of convention in treating generic compound names (such as "alcohol");
4. non-balanced reactions (lost reactants, wrong stoichiometry, etc);
5. undefined direction/reversibility;
6. abundance of irrelevant reactions, that are not supported by reconstruction;
7. lost relevant reactions that are supported experimentally and/or by reconstruction;
8. inconsistency in enzyme names, functions, and E.C. numbers;
9. inconsistency in encoding of complex reactions (polymerization, degradation, etc).

The following results have been produced to date:

- Integrated Genomics, Inc. has developed software tools to automatically convert the qualitative metabolic reconstruction into a formal computer-readable version of *Synechocystis sp. PCC 6803* reaction network.

This software also supports curation of compounds, reactions and functional roles necessary to resolve all the problems mentioned above. Appendix B illustrates the web-based user interface designed and implemented at IG (www.ergo-light.com/RN , user ID: PCC6803, password: DOE_GTL) to support curation of reaction network as collaborative effort by all participating labs;

- System-by-system curation effort was started. It includes:
 - a. manual curation of compounds, adding, reconciling and refining:
 - i. names/synonyms;
 - ii. atomic formulas;

- iii. structure representations ("smiles")
 - b. manual curation of reactions:
 - i. mass-balance;
 - ii. reversibility/direction;
 - iii. assignment with functional roles (E.C. numbers);
 - iv. hierarchical placement (pathways, systems)
 - c. manual curation of functional roles and genes:
 - i. names and E.C. numbers;
 - ii. connection to genes (reconcile annotations);
 - iii. revealing missing genes (and produce conjectures)
- Curation of the following subsystems has been largely completed: central carbon metabolism (CCM), amino acids biosynthesis (AAB), and nucleotide and cofactor biosynthesis. Statistics illustrating the current progress in this effort are summarized in Table 2.
- Reactions are connected to pathways and individual genes in ERGO Light.

Table 2. The current status of reactions/compounds curation

	Carbohydrate metabolism	Amino-acid metabolism	Coenzyme metabolism	Nucleotide metabolism
Total number of functional roles (distinct E.C. numbers)	102	105	58	34
Total number of reactions:	144	113	54	52
mass-balanced	95%	99%	78%	96%
Total number of compounds:	91	63	59	20
fully refined	90%	94%	81%	100%

Refinement of other functional systems relevant for the scope of the funded project will be continued during its second year. The produced reaction network will constitute an invaluable novel resource for our collaborators and for the broader scientific community needed to support mathematic modeling of cellular sub-networks, search for missing components, interpretation of experimental results, and deeper understanding of physiology of cyanobacteria, as well as other organisms such as plants.

Specific aim II.

The current version of the *Synechocystis* sp PCC 6803 metabolic reconstruction produced during the first year of this project already allows insight into *Synechocystis* 6803 physiology and genetics, that has proved very useful for experimental research in this organism and other cyanobacteria.

The following focused research projects were conducted at Integrated Genomics during the reported period to support wet lab experimental projects of our collaborators on this grant:

1. *Detailed reconstruction of carotenoid biosynthesis in Synechocystis 6803* was performed. This pathway has several "missing genes" (functions not connected to any sequenced gene), which are being pursued by Dr. Vermaas's lab. We suggested a candidate ORF for a "missing" lycopene cyclase (beta- and gamma-carotene synthase) - hypothetical gene RCY45362 (*slr0941*). This prediction is currently being pursued experimentally in the lab of Dr. Vermaas.
2. *Prediction of photoheterotrophic growth capability for Thermosynechococcus elongates*. The dominant mode of growth of cyanobacteria is photoautotrophy, in spite of the fact that they possess a respiratory activity and many enzymes of sugar catabolism. A thermophilic cyanobacterium capable of photoheterotrophic growth is needed for the purpose of crystal structure analysis of membrane protein complexes. Availability of the complete genome sequence makes *T. elongates* a strong candidate for this role. To facilitate experimental attempts by Dr. Vermaas's group to grow this strain under conditions other than photoautotrophic ones, we performed detailed comparative analysis of central carbon metabolism reconstructed from the complete genome sequences of several cyanobacteria with known growth capabilities. Several consistent differences were detected between obligate and facultative photoautotrophs in organization of glycolysis and tricarboxylic acid (TCA) cycle (manuscript in preparation). This analysis has shown that central metabolism of *Thermosynechococcus elongates* is close to that of facultative photoautotrophs. We predict that introduction into its genome of a gene for (S)-malate dehydrogenase (for instance, *Synechocystis 6803 slr0891*) may confer photoheterotrophic growth capacity. This prediction is currently being pursued experimentally in the lab of Dr. Vermaas.
3. *Predictions regarding fixed-carbon compounds taken up by Thermosynechococcus elongates*. The detailed comparative analysis of membrane transporters present in the genome of *T. elongates* resulted in prediction of several compounds as potential substrates for its photoheterotrophic growth. One of them, N-acetyl-D-glucosamine has been proven very useful in propagation of this organism.
5. *Reconstruction of NAD metabolism in Synechocystis 6803 and several other cyanobacterial species*. We performed *in silico* reconstruction of NAD metabolism based on the presence or absence of orthologs of all known genes involved in NAD biosynthesis, salvage, recycling, and transport in genomic sequences of cyanobacteria available to date. The resultant model was further refined and verified based on *in vitro* characterization of several key enzymes of NAD biosynthesis (NaMN adenylyltransferase, nicotinamide phosphoribosyltransferase, and NAD synthetase), and on genetic analysis of the *nadD*⁻ and *nadM*⁻ mutants in *Synechocystis 6803* (manuscript in preparation: (2)).
6. We have initiated a systematic analysis of regulatory sites in the genome of *Synechocystis 6803* based on literature data and on prediction of novel regulatory elements and the corresponding regulons, using methodology described in (1, 4). An example of this effort is

analysis of coregulated cobalamin (vitamin B₁₂) metabolism genes in the genomes of *Synechocystis* 6803 and several other cyanobacteria (manuscripts in preparation: 5, 6) The results of this on-going effort will be highly relevant for metabolic modeling, as well as for generating functional predictions for hypothetical cyanobacterial ORFs.

References:

1. **Gelfand, M. S., P. S. Novichkov, E. S. Novichkova, and A. A. Mironov.** 2000. Comparative analysis of regulatory patterns in bacterial genomes. *Brief Bioinform* 1:357-71.
2. **Kurnasov, O., S. Gerdes, K. Shatalin, B. Polanuyer, and A. Osterman.** 2003. Genomics of NAD biosynthesis and recycling in Cyanobacteria. manuscript in preparation.
3. **Overbeek, R., N. Larsen, T. Walunas, M. D'Souza, G. Pusch, E. J. Selkov, K. Liolios, V. Joukov, D. Kaznadzey, I. Anderson, A. Bhattacharyya, H. Burd, W. Gardner, P. Hanke, V. Kapatral, N. Mikhailova, O. Vasieva, A. Osterman, V. Vonstein, M. Fonstein, N. Ivanova, and N. Kyrpides.** 2003. The ERGO^(TM) genome analysis and discovery system. *Nucleic Acids Res.* 31:164-171.
4. **Rodionov, D. A., A. A. Mironov, and M. S. Gelfand.** 2002. Conservation of the biotin regulon and the BirA regulatory signal in eubacteria and archaea. *Genome Res* 12:1507-16.
5. **Rodionov, D. A., A. G. Vitreschak, A. A. Mironov, and M. S. Gelfand.** 2003. Comparative genomics of the vitamin B12 metabolism and regulation in prokaryotes. (manuscript in preparation).
6. **Vitreschak, A. G., D. A. Rodionov, A. A. Mironov, and M. S. Gelfand.** 2003. Conserved RNA structural element regulates vitamin B12 metabolism in bacteria. (manuscript in preparation).

Appendix A

Qualitative metabolic reconstruction of *Synechocystis* sp. PCC 6803 posted at http://www.ergo-light.com/RN/CY_Tree.html (user ID: PCC6803, password: DOE_GTL). Examples with interface screenshots.

A: Hierarchical functional overview of metabolism of *Synechocystis* sp. PCC 6803 reconstructed from its genome sequence. Arrows point to links which allow the user to access: i) large functional blocks; ii) graphic panels illustrating functional subsystems; iii) individual pathway pages; iv) individual protein pages for *Synechocystis* 6803 ORFs. The last two links connect the user directly with the corresponding pages in ERGO light.

The screenshot displays the ERGO light web interface for the metabolic reconstruction of *Synechocystis* sp. PCC 6803. The browser address bar shows http://www.ergo-light.com/RN/CY_Tree.html. The page title is "Metabolic reconstruction of *Synechocystis* sp PCC6803".

The left sidebar contains a hierarchical tree view of the metabolic reconstruction:

- ☐ CY Metabolic Reconstruction
 - ☐ Bioenergetics
 - ☐ Bacterial Photosynthesis
 - ☐ Photosynthesis and respiration in cyanobacteria
 - ☐ Electron transport in cyanobacterial thylakoid membrane
 - ☐ Bacterial cytochrome d ubiquinol oxidase
 - ☐ Cytochrome d ubiquinol oxidase subunit I (EC 1.10.3.-)
 - ☐ RCY08256 - Cytochrome d ubiquinol oxidase subunit I (EC 1.10.3.-)
 - ☐ Cytochrome d ubiquinol oxidase subunit II (EC 1.10.3.-)
 - ☐ RCY50685 - Cytochrome d ubiquinol oxidase subunit II (EC 1.10.3.-)
 - ☐ Bacterial proton-transporting ATP synthase
 - ☐ Bacterial cytochrome c oxidase aa3-caa3 type
 - ☐ Cytochrome c oxidase polypeptide I (EC 1.9.3.1)
 - ☐ RCY32756 - Cytochrome c oxidase polypeptide I (EC 1.9.3.1)
 - ☐ RCY37398 - Cytochrome c oxidase polypeptide I (EC 1.9.3.1)
 - ☐ Cytochrome c oxidase polypeptide II (EC 1.9.3.1)
 - ☐ RCY14296 - Cytochrome c oxidase polypeptide II (EC 1.9.3.1)
 - ☐ RCY15025 - Cytochrome c oxidase polypeptide II (EC 1.9.3.1)
 - ☐ Cytochrome c oxidase polypeptide III (EC 1.9.3.1)
 - ☐ Cytochrome c oxidase polypeptide IV (EC 1.9.3.1)
 - ☐ Cytochrome c oxidase polypeptide IVB (EC 1.9.3.1)
 - ☐ Cytochrome b6/f complex
 - ☐ Flavodoxin (photosystem I)
 - ☐ Flavodoxin
 - ☐ Bacterial ferredoxin
 - ☐ Ferredoxin
 - ☐ Protein composition of photosystem I
 - ☐ Photosystem I subunit VII
 - ☐ Photosystem I reaction center subunit II
 - ☐ Photosystem I reaction center subunit IV
 - ☐ Photosystem I reaction center subunit III
 - ☐ Photosystem I reaction center subunit VIII
 - ☐ Photosystem I reaction center subunit IX
 - ☐ Photosystem I reaction center subunit X
 - ☐ Photosystem I reaction center subunit XI
 - ☐ RCY41758 - Photosystem I reaction center subunit XI
 - ☐ Photosystem I reaction center subunit XII
 - ☐ Bacterial NADH-quinone oxidoreductase-complex I 1.6.5.3
 - ☐ NADH-quinone oxidoreductase chain A (EC 1.6.5.3)

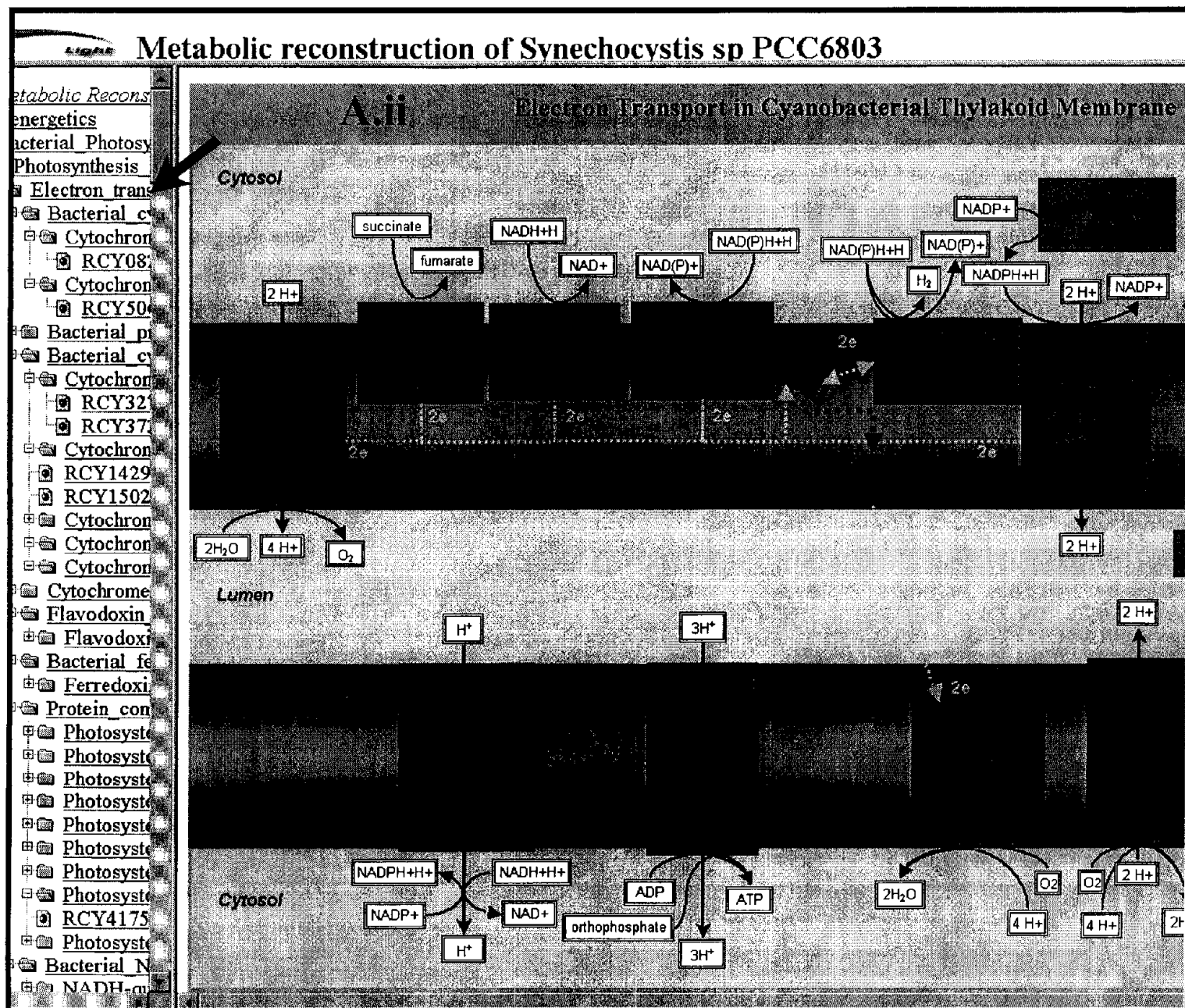
Arrows labeled i, ii, iii, and iv point to the following links in the tree view:

 - i)** Photosynthesis and respiration in cyanobacteria
 - ii)** Electron transport in cyanobacterial thylakoid membrane
 - iii)** Cytochrome c oxidase polypeptide I (EC 1.9.3.1)
 - iv)** RCY32756 - Cytochrome c oxidase polypeptide I (EC 1.9.3.1)

The right sidebar shows a graphic panel titled "PHOTOSYNTHESIS A" with a small image of a cyanobacterial cell.

A

Examples of the tree expansion to the pages listed above are provided in the corresponding panel A.ii, A.iii, A.iv



Metabolic reconstruction of *Synechocystis* sp PCC6803

Metabolic Reconstruction

energetics

Bacterial Photosynthesis

Photosynthesis and respiration in cy

Electron transport in cyanobacteria

Bacterial cytochrome d ubiquinol

Cytochrome d ubiquinol oxidase

RCY08256 - Cytochrome d ub

Cytochrome d ubiquinol oxidase

RCY50685 - Cytochrome d ub

Bacterial proton-transporting AT

Bacterial cytochrome C oxidase

Cytochrome c oxidase polypeptide

RCY32756 - Cytochrome c ox

RCY37398 - Cytochrome c ox

Cytochrome c oxidase polypeptide

RCY14296 - Cytochrome c oxid

RCY15025 - Cytochrome c oxid

Cytochrome c oxidase polypeptide

Cytochrome c oxidase polypeptide

Cytochrome c oxidase polypeptide

Cytochrome b6/f complex

Flavodoxin (photosystem I)

Flavodoxin

Bacterial ferredoxin

Ferredoxin

Protein composition of photosyst

Photosystem I subunit VII

Photosystem I reaction center sub

Photosystem I reaction center sub

Photosystem I reaction center sub

Photosystem I reaction center sub

Photosystem I reaction center sub

Photosystem I reaction center sub

RCY41758 - Photosystem I react

Photosystem I reaction center sub

Bacterial NADH-quinone oxidore

NADH-quinone oxidoreductase

Data

Query

Help

Pathway page

Switch Groups

Pathway Views

Diagram Blank

See Assertions

See Assertion

Pathway: Bacterial cytochrome C oxidase aa3-caa3 type

Pathway Name Bacterial cytochrome C oxidase aa3-caa3 type

Ref. Organism *Synechocystis* sp. PCC6803

Function Tree Display pathway in the context of sub-systems

4 Assertions

in 0 out of 0 Archaea (sorted by name)

in 4 out of 7 Bacteria (sorted by name)

in 0 out of 0 Eukaryota (sorted by name)

in 0 out of 0 Viruses (sorted by name)

Best KEGG Maps Oxidative phosphorylation

#	E.C. #	Functional Description	ORF's assigned this function
1	1.9.3.1	Cytochrome c oxidase polypeptide I	RCY32756 (ciaDII) RCY37398 (ciaD)
2		Cytochrome c oxidase polypeptide II	RCY15025 (ciaE)
3	1.9.3.1	Cytochrome c oxidase polypeptide III	RCY12508 (ciaEII) RCY34156 (ciaE)
4		Cytochrome c oxidase polypeptide IV	

Metabolic reconstruction of Synechocystis sp PCC6803abolic Reconstructionergeticsterial Photosynthesishotosynthesis and respiration in cyElectron transport in cyanobacteriaBacterial cytochrome d ubiquinoCytochrome d ubiquinol oxidaseRCY08256 - Cytochrome d ubCytochrome d ubiquinol oxidaseRCY50685 - Cytochrome d ubBacterial proton-transporting ATBacterial cytochrome C oxidaseCytochrome c oxidase polypeptRCY32756 - Cytochrome c oxRCY37398 - Cytochrome c oxCytochrome c oxidase polypeptRCY14296 - Cytochrome c oxidRCY15025 - Cytochrome c oxidCytochrome c oxidase polypeptCytochrome c oxidase polypeptCytochrome c oxidase polypeptCytochrome c oxidase polypeptCytochrome b6/f complexFlavodoxin (photosystem I)FlavodoxinBacterial ferredoxinFerredoxinProtein composition of photosystPhotosystem I subunit VIIPhotosystem I reaction center subPhotosystem I reaction center subPhotosystem I reaction center subPhotosystem I reaction center subPhotosystem I reaction center subPhotosystem I reaction center subPhotosystem I reaction center subRCY41758 - Photosystem I reactPhotosystem I reaction center subBacterial NADH-quinone oxidoreNADH-quinone oxidoreductase**Examine RCY32756****View Annotations****Local Blast (NR) -- Protein****Local Blast (NR) -- DNA****Function Cluster****Function Couplings****Function Tree****Paralogs****Paralog Cluster****A.iv****External Tools****TMPreD****ProtScale****PSI-Blast (NR)****RPS-Blast (NR)****EBI NCBI BLAST2****ProSite****ProDom Analysis****Pfam Analysis****EBI InterPro Scan****COG Analysis****CDART Analysis****Examine Checked****Annotations****Contig Regions****Amino Acid:****Primary Information for RCY32756 Synechocystis sp. PCC6803**

Aliases	ctaDII; AP003590; pirnr[NF00424092; slr2082; tr gil1653054 ; gil16330563
Contig Location	Synechocystis -PCC-6803 from 1,540,650 to 1,54 = 3,573,470 bp
AA Residues, DNA	544 aa, 1632 bp (returns sequence)
Molecular Weight	60,766 (returns all proteins with Molecular Weigh
Iso-electric Point	8.67 (returns all proteins that have pI within ± 0.1
Function	Cytochrome c oxidase polypeptide I (EC 1.9.3.1)
EC 1.9.3.1 Links	Enzyme Commission, ExPASy, KEGG Maps

Contig Region for RCY32756**Pathway Information for RCY32756**

Status	Name
Asserted	Bacterial cytochrome ba3 oxidase
Asserted	Bacterial cytochrome cao3 oxidase
Possible	Bacterial cytochrome C oxidase aa3-caa3 type
Possible	Bacterial cytochrome caa3 type HiPIP oxidase
Possible	Bacterial cytochrome cbb3 oxidase
Possible	Mammalian terminal cytochrome aa3 oxidase 1.
Possible	Oxidative phosphorylation - Reference pathway

External Annotations for RCY32756

Appendix B

Software for collaborative curation of reactions, compounds, and enzymes within the *Synechocystis* sp. PCC 6803 metabolic reaction network posted at <http://www.ergo-light.com/RN> (user ID: PCC6803, password: DOE_GTL).

Examples with interface screenshots.

A: The hierarchically organized list of reactions related to amino acid biosynthesis in *Synechocystis* 6803, automatically generated from metabolic reconstruction in ERGO. Arrows point to links which allow the user to retrieve: i) all reactions; ii) reactions for a given node (pathway). **B:** List of reactions relevant for the functional node "amino acid metabolism" retrieved from the report. Arrows point to links, which allow the user to retrieve a summary/curation window for: i) an individual reaction; ii) an individual compound. **C:** The summary/curation window for the reaction # 2593, retrieved from the previous screen (B). **D:** The summary/curation window for the compound L-lysine, retrieved from the previous screen (B). Arrow indicates a link to smile representation of the compound (shown as inset). The list of "dangling" compounds for the node "amino acid metabolism". All shown windows allow manual editing.

Reactions for Amino acid metabolism for "*Synechocystis* sp. PCC6803"

A

- Peptide metabolism
 - [glutathione biosynthesis](#)
- Polyamine metabolism
 - Polyamine degradation
 - [Acetylspermidine deacetylation](#)
 - [Gamma-aminobutyraldehyde dehydrogenase reaction](#)
 - Polyamine biosynthesis
 - [Norspermidine biosynthesis via carboxynorspermidine](#)
 - [L-arginine decarboxylation to agmatine](#)
 - [Putrescine biosynthesis from agmatine via agmatinase](#)
- Amino acid biosynthesis
 - L-asparagine biosynthesis
 - [Asparagine biosynthesis via Asn synthase glutamine-hydrolysing](#)
 - [L-methionine biosynthesis](#)
 - L-homocysteine biosynthesis
 - [isoleucine biosynthesis from threonine](#)
 - [threonine biosynthesis from homoserine](#)
 - [L-aspartate 4-semialdehyde--homoserine anabolism \(NADPH, NADH\) \(cytosol\)](#)
 - L-lysine biosynthesis
 - [Diaminopimelate decarboxylation to lysine](#)
 - [Diaminopimelate epimerase](#)
 - [Tetrahydrodipicolinate biosynthesis](#)
 - [aspartate--aspartate 4-semialdehyde, orthophosphate anabolism](#)
 - [L-glutamate, oxaloacetate--2-oxoglutarate, L-aspartate catabolism \(cytosol\)](#)
 - [aspartate decarboxylation via EC 4.1.1.11](#)
 - [L-ornithine, carbamoyl phosphate--L-arginine, fumarate anabolism \(ATP\) \(cytosol\)](#)
 - [L-proline biosynthesis from L-1-pyrroline 5-carboxylate](#)

Address <http://code.i-gee/RN/CGI/m.cgi>

126 reactions for Amino acid metabolism **E** **Synechocystis sp. PCC6803"**

☒ Filter out non-connectible compounds

React ID	Reaction	Balance	Functions
30	L-asparagine + water <=> L-aspartate + NH ₄ (+)	-	3.5.1.38 - GLUTAMINASE-(ASPARAGIN-)ASE, 3.5.1.1
45	7-phospho-3-deoxy-arabino-heptulosonate -> orthophosphate + 3-dehydroquinate	-	4.6.1.3 - 3-DEHYDROQUINATE SYNTHASE
81	(4-aminobutyl)guanidine + water -> putrescine + urea	-	3.5.3.11 - AGMATINASE
102	NAD(+) + water + L-phenylalanine <=> NADH + NH ₄ (+) + phenylpyruvate + H(+)	-	1.4.1.20 - PHENYLALANINE DEHYDROGENASE
118	L-histidinol + water + 2 NAD(+) <=> L-histidine + 2 NADH + 2 H(+)	-	1.1.1.23 - HISTIDINOL DEHYDROGENASE
121	acetyl-CoA + water + 3-methyl-2-oxobutanoate -> CoA + (R)-3-hydroxy-3-carboxy-4-methylpentanoate	-	4.1.3.12 - 2-ISOPROPYLMALATE SYNTHASE
2474	water + N ¹ -(1)-acetylspemidine -> acetylspemidine	C ₉ H ₂₁ N ₃ O	3.5.1.48 - ACETYLSPERMIDINE DEACETYLASE
2475	(S)-3-carboxy-4-methylpentanoate + 2-oxoisohexanoate		
2525	tryptamine -> NH ₄ (+) + indole-3-acetaldehyde	-H ₂	1.4.3.4 - AMINE OXIDASE (FLAVIN-CONTAINING)
257	L-aspartate + 4-semialdehyde + pyrrolidine -> 2,3-dihydrodipicolinate + 2 water		4.2.1.52 - DIHYDRODIPICOLINATE SYNTHASE
2593	'meso'-diaminopimelate -> L-lysine + CO ₂ (2)		4.1.1.20 - DIAMINOPIMELATE DECARBOXYLASE
2634	5-phospho-D-1-ribulose-5-phosphatase + L-glutamine <=> D-erythrose-1,6-bisphosphate + L-glutamate + water		2.4.2.-
2722	(S)-2-acetolactate + NADPH + H(+) <=> (R)-2,3-dihydroxy-3-methylbutanoate + NADP(+)		1.1.1.86 - KETO-ACID REDUCTOISOMERASE

1 reaction for Diaminopimelate decarboxylation to lysine **C**

☒ Filter out non-connectible compounds

React ID	Reaction	Balance	Functions
2593	'meso'-diaminopimelate -> L-lysine + CO ₂ (2)		4.1.1.20 - DIAMINOPIMELATE DECARBOXYLASE

Edit compound **D**

Compound Id: 222

CAS: 56-87-1

Name: L-lysine

Aliases:

- 2,6-diaminohexanoic acid
- lysine
- lysine acid
- Lys
- (S)-2,6-diaminocaproic acid
- (S)-2,6-diaminohexanoic acid

Add alias

Electric charge:

Molecular formula: C₆H₁₁4N₂O₂

Smiles: [C@@H](CCCC[NH3+])(N)C(=O)[O-]

Dangling compounds for Amino acid metabolism **E**

Substrates:

Compound	Where
D-erythrose 4-phosphate	138
N ¹ -(1)-acetylspemidine	2474
N ³ -(3)-acetylspemidine	3004
tryptamine	2525

Products:

Compound	Where
GSH	1290
spemidine	2474, 3004
L-lysine	2593
putrescine	81
indole-3-acetaldehyde	2525