Biotechnology and Bioprocess Engineering **2005**, 10: 418-424

Systemic Optimization of Microalgae for Bioactive Compound Production

Jeong-Dong Kim and Choul-Gyun Lee*

Institute of Industrial Biotechnology, Department of Biological Engineering, Inha University, Inchon 402-751, Korea

Abstract The complexity of the biological system/biological systems has been fascinating and challenging for a long time. With the advent of mathematical tools with various omics technology, systems biology was born and is already ubiquitous in every area of biology and biotechnology. Microalgal biotechnology is no exception in this new trend. As tens of microalgal genomes become publicly available on the Internet, vast amounts of data from genomics, transcriptomics, and proteomics are reported everyday. Though there has not yet been enough data gathered on microalgal metabolomics, the *in silico* models for relatively simple cyanobacteria or for organelles, such as chloroplasts, will appear presently. With the help of systems biology, a more in-depth understanding of microalgae will be possible. Consequently, most industrially-interested microalgae can be metabolically redesigned/reconfigured as cell factories. Microalgae will be served as the hosts in white biotechnology.

Keywords: cyanobacteria, microalgae, systems biology, optimization, *in silico*, metabolic reconstruction

INTRODUCTION

System-level approaches of biological systems started in the early 1990s [1], though the first concept of system-level understanding was reported much earlier [2]. This holistic approach to understanding and simulating biological systems became a hot issue in the first part of the 21st century [3,4], forming a new discipline referred to as systems biology. Systems biology is the global and integrative study of biological systems and is linked to various kinds of high-throughput omics that have made remarkable contributions to understand biological systems and helping understand the cells as systems [5].

Over the past decades, our knowledge of biological systems and living creatures has dramatically increased. Additionally, the host of information on cell components and compartments (such as organelles and tissues) has produced to molecular levels (such as genes and proteins). The reductionist approach of the 20th century has achieved many cornerstones in bioscience and medical areas. For example, Mendel's work on heredity was based on this reductionist approach while his contemporaries failed to deduce any rules due to the complex nature of heredity.

However, this reductionist approach often cannot see the trees in the forest. Understanding the cells in a gene-to-gene or gene to protein relationship between genes and reactions or genes and phenotypes with minimal or no interaction between other genes and components has some limitations. (i) This reductionist approach only works for an isolated system with minimal or no connections to other components. The same protein from the same gene can have a variety of functions depending on the network state or connectivities [6]. (ii) A component with a great number of connectivity tends to be more vital for the cells to function properly, which cannot be determined or predicted using the reductionist approach. (iii) The essence of life can be fully understood only when the concentrations, the flux distributions, the reactions, the connectivities, and the network state of all the components such as genes, proteins, products, metabolites and all the other molecules in the cell are known both quantitatively and qualitatively.

On the other hand, systems biology in the 21st century utilizes the integrative or holistic approach, using all the available data from various techniques and high-end computing powers with novel statistical analysis to reconstruct the relationships and networks for all the existing components in a multi-level, multi-variant manner. It was these new omics techniques, mathematical methods and computing powers that have hampered the development of systems biology, even though the idea has been around for quite some time. The reconstruction of the model would be extremely hard and the proper simulation or prediction for validation would not be possible without the recent technological developments.

Systems biology investigates the behavior, structure, and connectivity of all elements in a particular biological system of a functioning state. In this study, a mathematical model of this system will be reconstructed based on all

***Corresponding author**
Tel: +82-32-860-7518 Fax: +82-32-872-4046
e-mail: leecg@inha.ac.kr

available data on biological systems of interest. The developed model will allow hypotheses about the characteristics and behavior of the system. These hypotheses can, in turn, be verified by both *in silico* and *in vitro* experiments. The *in silico* simulation will produce new data that will permit verification of the hypothesis. Parallel *in vitro* experiments will also generate new biological data, which will also help validate the hypotheses developed. The iterative integrative process will generate new biological information. The goal of systems biology is to deduce a natural conclusion or to understand cells as a system. When one has large amounts of data from a particular reaction or system, one can list the data one item at a time. This will indicate the most precise data points but would be extremely difficult to see the trends or to predict the data outside the obtained range. However, if a mathematical model of the system can be obtained, the model may not represent the precise values, but would definitely be the simplest way to express the essence of the results. Furthermore, it would be easy to see the trends and predict other situations through extrapolation. Systems biology performs the same action for a biological system of interest. The amount of data from various omics techniques and conventional wet experiments can be overwhelming even with a supercomputer to deduce trends. Having an *in silico* model using systems biology would give biologists and biotechnologists of the 21st century the flexibility and insight to understand the dynamic interactions of biological systems.

What are Microalgae?

Unlike most taxonomy terms, the term "algae" does not describe a particular branch of evolution, but instead refers to all the photosynthetic organisms that are neither plant nor bacteria. This means algae are all the photosynthetic organisms that have chlorophyll *a* in a thallus form [7]. Simpler forms of algae that are either single cells or undifferentiated multicellular are referred to as microalgae. If the multicellular algae differentiate in reproduction and other functions (as in kelps and other seaweeds), they are macroalgae. Euglena and similar genera are free-swimming, single-celled forms that contain chlorophyll. They are also able, under certain conditions, to ingest food in an animal-like manner. The green algae include most freshwater forms. The green slimes found in stagnant water and the green film found on the bark of trees is green algae. The more complex brown algae and red algae are primarily saltwater forms. The green color of the chlorophyll is masked by the presence of other pigments. Most algae are classified in the kingdom Protista using a 5 kingdom classification or in the Eukarya domain in 3 kingdom system. Blue-green algae (cyanobacteria) have been grouped with other prokaryotes in the Monera kingdom or Bacteria domain (Microalgae can be considered as the most primitive form of plants). While the photosynthesis mechanism in microalgae is similar to that of higher plants, they are generally more efficient converters of solar energy because of their simple cellular structure. In addition, because the cells grow in aqueous suspension, they have more efficient access to water, CO_2, and nutrients.

Why Microalgae Now?

Microalgae are primary producers for food chain ecosystems and are considered as a low utility group. However, as land resources are being depleted, microalgae are now treated as substitute resources or new sources for various biological materials. The number of microalgae species has been estimated at between 22,000 and 26,000 [8], most of which are yet to be described or defined. But some of the recent estimations suggest that there could be well over 100,000 species of microalgae. The high variety of species in various habitats is one of the main reasons why microalgae are drawing renewed interest from researchers. Microalgae can be used to produce a wide range of primary and secondary metabolites, such as proteins, lipids, carbohydrates, carotenoids, vitamins, and other biologically active compounds [9-11]. Besides, some of the compounds from microalgae show different characteristics from their counterparts from land organisms [12]. Due to this vast potential, commercial use of algal cultures spans approximately 50 years with various applications to get the compounds from algae. However, unlike the rapid and significant advances in the biotechnological use of bacteria, yeast and mammalian cells, algal biotechnology and the development of photoautotrophic cultures has progressed more slowly in spite of their recognized utility [13]. The recent progress in large-scale cultivation techniques such as efficient photobioreactor (PBR) technology spreads the interests on microalgal cultures for various bioactive products and health food.

Systems Biology to Understand the Complexity of Life

There are several important features and merits of systems biology. First, system structures and system dynamics of microalgae (or any other biological systems) will be well understood when using the system-level approaches. Second, a system-level understanding of metabolic pathways, physiology and malfunction will provide a method to control the state of cells, thus producing primary and secondary metabolites. Finally, the applications of system-level approaches in biotechnology provide the possibility to design new biological systems with the desired properties that do not exist in nature [14].

A pictorial roadmap of systems biology is shown in Fig. 1. As discussed earlier, systems biology involves a multi-level, multi-variant approach, including the integration of vast amounts of data from bioinformatics, genomics, transcriptomics, proteomics, metabolomics, *etc.* as represented in Fig. 1. These data will be the basis of the metabolic models and separate models for regulatory networks and signaling networks should be combined into the metabolic model. Then, the iterative integrative process will fine-tune the model to give a new method for white biotechnology. Systems biology techniques, together with engineering optimization, will generate cell factories that are efficiently optimized to produce any metabolite in the cell.

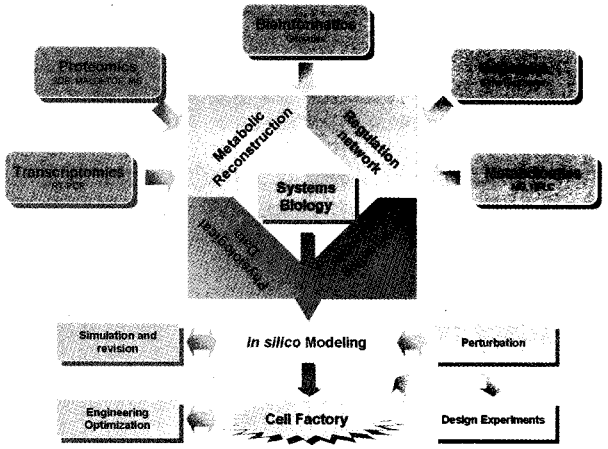

Fig. 1. Schematic diagram of systems biology to bring genomes to life and then to cell factories.

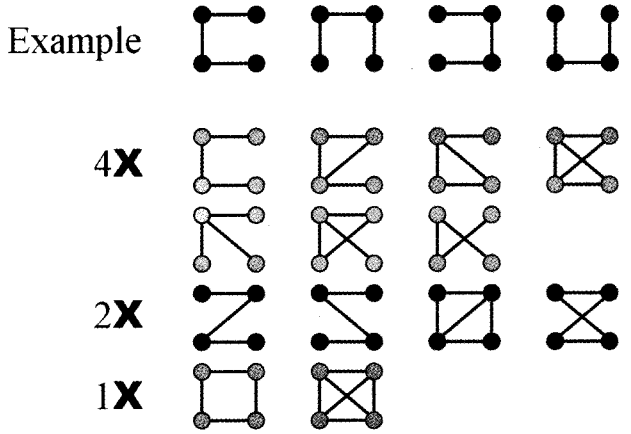

Fig. 2. Pictorial representation of various network states. There are 38 ways to connect just 4 dots and the number of possible network states increases exponentially as the number of nodes increases.

For example, assume that there are only four components in the cell as in Fig. 2. Each way of connecting all components is a unique network state in the overall system. Since the four components in Fig. 2 are different from each other, identical connection patterns in different orientations indicate a different state (show different characteristics), as shown in the top line of Fig. 2.

How many ways are there to connect all four components? Or, how many different states can a biological system have with a mere four components? The diagrams in the second and the third lines can be rotated to give different network states, allowing/resulting in 4 times the number of the diagrams (denoted with 4× in Fig. 2). Likewise, the states in the fourth line have another configuration with the same shape, but with a rotation of 90 degrees. Amazingly, there are 38 ways to make systems with just these four components. This number will increase exponentially if the characteristics of the connection are considered. Any state of the network can have an infinite number of possibilities through various factors: the flux between the components, the effect of environmental parameters such as temperature and pH, the concentration or the availability of components, the existence of different alleles for the same component, the loss of activity, and the mutation of a component. Consequently, one of the major goals of systems biology used in white biotechnology is to find the best possible network state among all possible network states to exploit the biological system's potential.

Genomic Researches for Microalgae

Microalgae can be found everywhere, both in terrestrial and aquatic habitats (freshwater and marine) and have extensively been studied, most recently due to the potentials of secondary metabolites. These can be used as pharmaceuticals, nutraceuticals (including dietary supplements), fine chemicals and other bioactive compounds. A large number of microalgae that have industrial potentials or that are model strains with vast amounts of research data are either completely sequenced already or being sequenced.

As of August 2005, 275 complete genomes were publicly available and nearly 1,500 genome projects are undergoing research/study in major sequencing centers [15]. Table 1 lists some databases that have genome sequences of microalgae. Among these, only a handful of microalgal genomes are available on the Internet. They are mostly cyanobacterial sequences with relatively smaller

Table 1. Major databases for microalgal genome sequences

Name	URL
CyanoBase: Genome Database for Cyanobacteria	http://www.kazusa.or.jp/cyano/
Genome Projects of Cyanobacteria and Related Organisms	http://bio.c.u-tokyo.ac.jp/labs/ikeuchi/c_genomeE.html
GOLD: Genomes OnLine Database	http://www.genomesonline.org/
DOE JGI: Joint Genome Institute	http://www.jgi.doe.gov/
NCBI Genome Project	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?DB=genomeprj
TIGR - Comprehensive Microbial Resource	http://cmr.tigr.org/
Genoscope	http://www.genoscope.cns.fr/

genome size. Up to now, at least 19 species of cyanobacteria have been fully sequenced: *Anabaena* PCC 7120, *Anabaena variabilis* ATCC 29413, *Chlorobium tepidum* TLS, *Chloroflexus aurantiacus*, *Crocosphaera watsonii* WH 8501, *Gloeobacter violaceus* PCC 7421, *Nostoc punctiforme* PCC 73102, *Nostoc* sp. PCC 7120, *Prochlorococcus marinus* SS120, *Prochlorococcus marinus* MED4, *Prochlorococcus marinus* MIT 9312, *Prochlorococcus marinus* MIT 9313, *Rhodopseudomonas palustris* CGA009, *Rhodopirellula baltica* SH 1, *Synechococcus elongatus* PCC 6301, *Synechococcus elongatus* PCC 7942, *Synechococcus* sp. WH 8102, *Synechocystis* PCC 6803, *Thermosynechococcus elongatus* BP-1, and *Trichodesmium erythraeum* IMS101. Furthermore, quite a few number of cyanobacterial strains are currently being sequenced, including *Prochloron didemni*, *Synechococcus* sp. CC 9311, *Synechococcus* sp. CC 9605, *Synechococcus* sp. CC 9902, *Synechococcus* sp. PCC 7002, *Synechococcus* sp. RCC 307, *Synechococcus* sp. WH 7803, and *Synechococcus* sp. WH 7805 [16,17]. The size of most cyanobacterial genomes ranges between 1.5 to 9 Mb usually with only 1 chromosome, which makes the genome projects of cyanobacteria relatively easy. Typical cyanobacterial genome has 1,500 to 8,000 ORFs, which give a rough ratio of 1 ORF per 1 kb.

However, for eukaryotic microalgae, no complete genome sequence is publicly available yet, though the genome sequence of *Chlamydomonas reinhardtii* is almost complete [16]. The genome of *C. reinhardtii* consist of 17 haploid chromosomes with mitochondrion and plastids and the size of it is about 100 Mb. Due to the size of the genome of eukaryotic algae, relatively few species are under being sequenced: *C. reinhardtii*, *Micromonas pusilla*, *Ostreococcus* sp. CCE 9901, *Ostreococcus tauri*, *Thalassiosira pseudonana*, and *Volvox carteri* [16,17].

As the number of available sequences increases exponentially, annotation methods using bioinformatics and other systematic analysis have developed accordingly. Based on the experience in our own lab, about 60-80% of sequenced cDNA clones of an eukaryotic microalgae could be annotated based on similarity searches by BLAST [18] even though there were very few sequences available for eukaryotic microalgae in the online database. Simultaneous explosions of online sequence databases and bioinformatics tools along with Internet accessibility make annotation process much easier than late 20th century.

Transcriptomics Research for Microalgae

Transcriptomics using DNA microarray are one of the recent technological developments widely used to analyze gene expression or transcriptional profiles despite the difficulties in quantitative comparison. Presently, there is only one commercial DNA microarray (CyanoCHIP, Takara Bio, Otsu, Japan) available for microalgae, though there are some movements for developing for *Synechocystis* PCC 6803 with 2,950 DNA fragments, which cover about 95% of the total ORFs in the cyanobacterial strain [19,20]. Most transcriptomics work reported to date, are therefore on *Synechocystis* PCC 6803 using the CyanoCHIP [21-25]. Other DNA microarrays that have been developed and used are for cyanobacterial strains. The Matsunaga Group developed the first cyanobacterial biodiversity microarray using bacterial magnetic particles [26]. DNA chips with specific purposes were also designed, such as the oligonucleotide microarray for detection and quantification of diverse genes of nitrogen cycle [27]. Microarray for *Prochlorococcus* (*P. marinus* MEDA and *P. marinus* MIT 9313) - the smallest and most abundant photosynthetic microbe in the oceans - is being developed for the analysis of gene expression patterns to understand how these microbes cope with the dilute environment of the oligotrophic oceans [16]. DNA microarrays for *Nostoc* sp. PCC 7120 and *Synechococcus* sp. WH8102 are also being considered for development [16,25].

As the sequencing of some eukaryotic microalgae is being completed, there are some reports indicating transcriptional analysis of green algae. The circadian regulation of transcription of *C. reinhardtii* was examined using a cDNA chip [28]. Transcriptional profiling of astaxanthin-induced *Haematococcus pluvialis* was performed by our research team [29]. As discussed earlier, more omics data including transcriptomics data from various DNA microarrays will help us reconstruct more vigorous *in silico* models for microalgae. In turn, this will help us to understand the microalgae themselves as systems.

Proteomics Research for Microalgae

There are only a few cyanobacterial proteomics studies, which are mostly used as two-dimensional electrophoresis (2DE). They then go through sequencing by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) techniques. Other recent techniques such as ESI-MS/MS or MALDI-TOF TOF are also widely used. Most cyanobacterial proteomics studies carried out in *Synechocystis* sp. PCC 6803 [30-35] and some reports on *Nostoc commune* DRH1, *Nostoc puntiforme* PCC 73102 and *Microcystis aeruginosa* PCC 7806 were recently added [25]. Currently, most proteomic works on microalgae are mainly focused on photosynthesis and other membrane-mediated functions [30, 31,34-38]. A cyanobacterial 2D database, Cyano2Dbase at Kazusa DNA Research Institute has completed information on 234 spots from whole cell extracts as well as 50+ spots from thylakoid membrane fraction of *Synechocystis* PCC 6803 [39]. Nevertheless, most cyanobacterial proteomics works are still in rudimentary stages. Compared to hundreds of proteomes from *Bacillus* or *Pseudomonas*, only a few cyanobacteria are represented by their proteomes [40]. More detailed review on cyanobacterial proteomic studies can be found elsewhere [25].

Even less work has been reported on eukaryotic microalgae and most of them are used for commercialized strains. Proteome analysis by SDS-PAGE of *Nannochloropsis salina*, which is an important green algae in aquaculture industry, indicated that cell protein profiles depended on the pollutant concentration [41]. It has been reported that there is a profound change in the cell

wall composition of *H. pluvialis* as green algae synthesizes astaxanthin [42]. Our lab also found that the expression profiles of the genes in carotenoid synthesis pathways and stress related reactions changed dramatically (unpublished data).

Metabolomics Research for Microalgae

As discussed above, the microalgae should be understood as systems to be used as cell factories to produce various bioactive compounds (refer to Fig. 1). The metabolic reconstruction by systems biology begins with the capture of data. Even though the major part of the available data is from the genomics, transcriptomics and proteomics, the metabolomics data is the most important factor in building successful *in silico* models. The metabolome represents the collection of all metabolites in a biological organism, which represents the end product of its gene expression. Thus, while mRNA gene expression data (transcriptomics) and proteomic analyses (proteomics) do not tell the whole story of what might be transpiring in a cell, metabolic profiling can give an instantaneous snapshot (a specific network state) of the physiology of that cell. One of the challenges of systems biology is to integrate information for proteomics, transcriptomics, and metabolomics to provide a more accurate description of different network states of living organisms.

Despite the ultimate goal of systems biology and the utmost importance of these techniques, genomics, transcriptomics, and proteomics studies of microalgae (esp. eukaryotes) are just started and few metabolomics research on microalgae has been reported. However, the recent advances in the metabolome measurement techniques enable the simultaneous measurement of a large number of metabolites, essential for metabolomic analysis. Mass spectrometry (MS) also plays a leading role. MS can be used alone for the simultaneous measurement of a range of metabolites, but when combined with liquid chromatography and gas chromatography the technique, known as LC/MS and GC/MS, is respectively much more powerful. Capillary electrophoresis combined with MS (CE/MS) is a recent tool/method that is suitable for high-throughput analysis with a small amount of samples and high resolution [4]. Rapid and accurate metabolome profiling will boost the number of successful *in silico* models in the next decade.

Ethical, Legal and Social Issues (ELSI)

This new paradigm will eventually provide biotechnologists with powerful tools to accurately predict and ultimately manipulate the biology of organelles and cells as systems. This, in turn, will confer abilities to alter microalgal sensitivity to environmental signals and to use microalgae as cell factories. We not only may be able to engineer the entire metabolism of the microalgae to squeeze more products from the culture, but also to design *de novo* biological systems that can grow in wastewater and carbon dioxide, while producing energy and high-valued bioactive compounds. As the Kyoto Protocol will come into effect from 2008, simultaneous carbon sequestration and renewable energy production would be a huge change.

However, designing a new organism should be very careful since the new organisms may have the potential to harm the environment in some manners as some GMOs did. As with any knowledge, systems biology may enrich life by promoting knowledge and technologies that promise to enhance our health and quality of life, but at the same time deteriorating life when used improperly. Intellectual property issues may arise as well over the applications and commercialization of resources and data. To maximize the benefits of these emerging technologies while anticipating and minimizing risks, collaborations and other cross-disciplinary interactions along with rigorous discussion between biotechnologists/biologists and nonscientists will be necessary before the technology are being exploited.

What Systems Biology Can Do for Microalgal Biotechnology

The development of bioinformatics is bringing an exponential growth in the utilization of new sequences, genomes, and expression protein profiles information from cells to the whole body. In addition, molecular biology data types of cyanobacteria that are sequence types, ribosomal and phycocyanin genes, and primary and secondary metabolic genes, are being held within GenBank [17]. Systematic analysis of gene and protein expression profiles requires experimental technologies including DNA microarray technology and proteomics. The Internet makes it easy to access the accumulated molecular microbiological data and the bioinformatics tools maximize the usefulness of these data. This knowledge and new scientific research discovered using systems biology can be combined to predict interacting networks in cellular processes, including phenotypes of whole microorganisms in systems of higher complexity.

Systems biology is undoubtedly based on interdisciplinary fields such as biocatalysis, biomathematics, biocomputing, biomolecular and cellular analysis, molecular structure, *etc.*, together with conventional biology and biotechnology. In addition, a major importance needs to be placed on the establishment of methodologies and techniques that enable us to understand biological systems as only systems.

In silico models for relatively simpler organisms, such as *E. coli*, *Helicobacter pylori*, *Haemophilus influenzae*, and *Saccharomyces cerevisiae* are already available and the mode is for more complex organisms, such as human organs or human are being constructed [5,43]. Though these models have been used for a short period only for a handful or organisms or organs, enormous discoveries have already been made. This trend will only be accelerated. Currently, there is no *in silico* model for any microalgal strains. The models for relatively simple cyanobacteria or for organelles, such as chloroplasts, will appear within a year or so. With a deeper understanding of bio-

logical systems and the appropriate simulation tools and models, the increase in productivity in microalgal cultures will be dramatically raised in a relatively short period. The development of microalgal biotechnology may not have kept the pace with the biotechnological use of bacteria, yeast and mammalian cells till now, but the future of microalgal biotechnology will be as bright as any other areas in white biotechnology.

Acknowledgement This work was supported by Systems Biology Program from Inha University, for which the authors are grateful.

REFERENCES

[1] Palsson, B. (2000) The challenges of *in silico* biology. *Nat. Biotechnol.* 18: 1147-1150.
[2] Weiner, N. (1948) *Cybernetics or Control and Communication in the Animal and the Machine*, MIT Press, Cambridge, MA, USA.
[3] Kitano, H. (2001) *Foundations of Systems Biology*, MIT Press, Cambridge, MA, USA.
[4] Tomita, M. and T. Nishioka (2005) *Metabolomics: The Frontier of Systems Biology*, Spinger-Verlag, Tokyo, Japan.
[5] Palsson, B. (2002) *In silico* biology through "omics". *Nat. Biotechnol.* 20: 649-650.
[6] Schwikowski, B., P. Uetz, and S. Fields (2000) A network of protein-protein interactions in yeast. *Nat. Biotechnol.* 18: 1257-1261.
[7] Lee, R. E. (1989) *Phycology*, Cambridge University Press, Cambridge, UK.
[8] Vonshak, A. (1997) *Spirulina platensis* (*Arthrospira*): *Physiology, Cell-biology and Biotechnology*, Taylor & Francis, London, UK.
[9] Dufosse, L., P. Galaup, A. Yaron, S. M. Arad, P. Blanc, K. N. Chidambara Murthy, and G. A. Ravishankar (2005) Microorganisms and microalgae as sources of pigments for food use: a scientific oddity or an industrial reality? *Trends Food Sci. Technol.* 16, 389-406.
[10] Hejazi, M. A. and R. H. Wijffels (2004) Milking of microalgae. *Trends Biotechnol.* 22: 189-194.
[11] Liang, S., X. Liu, F. Chen, and Z. Chen (2004) Current microalgal health food R & D activities in China. *Hydrobiologia* 512: 45-48.
[12] Borowitzka, M. A. (1992) Algal biotechnology products and processes - matching science and economics. *J. Appl. Phycol.* 4: 267-279.
[13] Chapman, D. J. and K. W. Gellenbeck (1989) A historical perspective of algal biotechnology. In: *Algal and Cyanobacterial Biotechnology* (Cresswell, R. C., T. A. V. Rees, and H. Shah, Eds.), pp. 1-27. Longman Scientific & Technical, Harlow, UK.
[14] Bailey, J. E. (1999) Lessons from metabolic engineering for functional genomics and drug discovery. *Nat. Biotechnol.* 17: 616-618.
[15] http://www.genomesonline.org/. GOLD (Genomes OnLine Database).
[16] http://www.jgi.doe.gov/. DOE Joint Genome Institute.
[17] http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?DB=genomeprj. NCBI Genome Project.
[18] Eom, H., Park, S., Lee, C.-G., and Jin, E. (2005) Gene expression profiling of an eukaryotic microalga, *Haematococcus pluvialis*. *J. Microbiol. Biotechnol.* in press.
[19] Kaneko, T., A. Tanaka, S. Sato, H. Kotani, T. Sazuka, N. Miyajima, M. Sugiura, and S. Tabata (1995) Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. I. Sequence features in the 1 Mb region from map positions 64% to 92% of the genome. *DNA Res.* 2: 153-166.
[20] Kaneko, T., S. Sato, H. Kotani, A. Tanaka, E. Asamizu, Y. Nakamura, N. Miyajima, M. Hirosawa, M. Sugiura, S. Sasamoto, T. Kimura, T. Hosouchi, A. Matsuno, A. Muraki, N. Nakazaki, K. Naruo, S. Okumura, S. Shimpo, C. Takeuchi, T. Wada, A. Watanabe, M. Yamada, M. Yasuda, and S. Tabata (1996) Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.* 3: 109-136.
[21] Domain, F., L. Houot, F. Chauvat, and C. Cassier-Chauvat (2004) Function and regulation of the cyanobacterial genes lexA, recA and ruvB: LexA is critical to the survival of cells facing inorganic carbon starvation. *Mol. Microbiol.* 53: 65-80.
[22] Kobayashi, M., T. Ishizuka, M. Katayama, M. Kanehisa, M. Bhattacharyya-Pakrasi, H. B. Pakrasi, and M. Ikeuchi (2004) Response to oxidative stress involves a novel peroxiredoxin gene in the unicellular cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Cell Physiol.* 45: 290-299.
[23] Hihara, Y., A. Kamei, M. Kanehisa, A. Kaplan, and M. Ikeuchi (2001) DNA microarray analysis of cyanobacterial gene expression during acclimation to high light. *Plant Cell* 13: 793-806.
[24] Schmitt Jr., W. A. and G. Stephanopoulos (2003) Prediction of transcriptional profiles of *Synechocystis* PCC6803 by dynamic autoregressive modeling of DNA microarray data. *Biotechnol. Bioeng.* 84: 855-863.
[25] Burja, A. M., S. Dhamwichukorn, and P. C. Wright (2003) Cyanobacterial postgenomic research and systems biology. *Trends Biotechnol.* 21: 504-511.
[26] Matsunaga, T., H. Nakayama, M. Okochi, and H. Takeyama (2001) Fluorescent detection of cyanobacterial DNA using bacterial magnetic particles on a MAG-microarray. *Biotechnol. Bioeng.* 73: 400-405.
[27] Taroncher-Oldenburg, G., E. M. Griner, C. A. Francis, and B. B. Ward (2003) Oligonucleotide microarray for the study of functional gene diversity in the nitrogen cycle in the environment. *Appl. Environ. Microbiol.* 69: 1159-1171.
[28] Kucho, K., K. Okamoto, S. Tabata, H. Fukuzawa, and M. Ishiura (2005) Identification of novel clock-controlled genes by cDNA macroarray analysis in *Chlamydomonas reinhardtii*. *Plant Mol. Biol.* 57: 889-906.
[29] Eom, H., C.-G. Lee, and E. Jin, (2005) Gene expression profile in astaxanthin-induced *Haematococcus pluvialis* using a cDNA microarray. *Planta*, submitted.
[30] Kashino, Y., W. M. Lauber, J. A. Carroll, Q. Wang, J. Whitmarsh, K. Satoh, and H. B. Pakrasi (2002) Proteomic analysis of a highly active photosystem II preparation from

the cyanobacterium *Synechocystis* sp. PCC 6803 reveals the presence of novel polypeptides. *Biochemistry* 41: 8004-8012.

[31] Wang, Y., J. Sun, and P. R. Chitnis (2000) Proteomic study of the peripheral proteins from thylakoid membranes of the cyanobacterium *Synechocystis* sp. PCC 6803. *Electrophoresis* 21: 1746-1754.

[32] Gan, C. S., K. F. Reardon, and P. C. Wright (2005) Comparison of protein and peptide prefractionation methods for the shotgun proteomic analysis of *Synechocystis* sp. PCC 6803. *Proteomics* 5: 2468-2478.

[33] Whitelegge, J. P., J. E. Katz, K. A. Pihakari, R. Hale, R. Aguilera, S. M. Gomez, K. F. Faull, D. Vavilin, and W. Vermaas (2004) Subtle modification of isotope ratio proteomics; an integrated strategy for expression proteomics. *Phytochemistry* 65: 1507-1515.

[34] Herranen, M., N. Battchikova, P. Zhang, A. Graf, S. Sirpio, V. Paakkarinen, and E. M. Aro (2004) Towards functional proteomics of membrane protein complexes in *Synechocystis* sp PCC 6803. *Plant Physiol.* 134: 470-481.

[35] Huang, F., I. Parmryd, F. Nilsson, A. L. Persson, H. B. Pakrasi, B. Andersson, and B. Norling (2002) Proteomics of *Synechocystis* sp. strain PCC 6803: Identification of plasma membrane proteins. *Mol. Cell. Proteom.* 1: 956-966.

[36] Schröder, W. P. and T. Kieselbach (2003) Update on chloroplast proteomics. *Photosynth. Res.* 78: 181-193.

[37] Norling, B., E. Zak, B. Andersson, and H. Pakrasi (1998) 2D-isolation of pure plasma and thylakoid membranes from the cyanobacterium *Synechocystis* sp. PCC 6803. *FEBS Lett.* 436: 189-192.

[38] Nelson, N. and A. Ben-Shem (2005) The structure of photosystem I and evolution of photosynthesis. *Bioessays* 27: 914-922.

[39] Sazuka, T., M. Yamaguchi, and O. Ohara (1999) Cyano2 Dbase updated: Linkage of 234 protein spots to corresponding genes through N-terminal microsequencing. *Electrophoresis* 20: 2160-2171.

[40] Bernal, A., U. Ear, and N. Kyrpides (2001) Genomes OnLine Database (GOLD): a monitor of genome projects worldwide. *Nucleic Acids Res.* 29: 126-127.

[41] Mohammady, N. G. D., Y. C. Chen, A. A. El-Mahdy, R. F. Mohammad, and E. D. Mohammady (2005) Temporal alterations of *Nannochloropsis salina* (Eustigmatophyceae) grown under aqueous diesel fuel stress. *J. Appl. Phycol.* 17: 161-170.

[42] Wang, S.-B., Q. Hu, M. Sommerfeld, and F. Chen (2004) Cell wall proteomics of the green alga *Haematococcus pluvialis* (Chlorophyceae). *Proteomics* 4: 692-708.

[43] Palsson, B. O. (2004) *In silico* biotechnology: Era of reconstruction and interrogation. *Curr. Opin. Biotechnol.* 15: 50-51.

[Received September 29, 2005; accepted October 12, 2005]