Whole-cell simulation: a grand challenge of the 21st century

Masaru Tomita

Study of the cell will never be complete unless its dynamic behavior is understood. The complex behavior of the cell cannot be determined or predicted unless a computer model of the cell is constructed and computer simulation is undertaken. Rapid accumulation of biological data from genome, proteome, transcriptome and metabolome projects can bring us to the point where it is no longer purely speculative to discuss how to construct virtual cells *in silico*. This article describes attempts to construct whole cell models. The E-CELL project has completed a couple of virtual cell models, and computer simulations have revealed some biological surprises.

Suppose that a certain organism's genome has been completely sequenced. Then suppose that structures and functions of all its gene products have been thoroughly identified. Suppose further that a giant map of the entire metabolic pathways has been drawn flawlessly. Then what? Would we have conquered the cell? The answer is clearly 'no' because the overall 'behavior' of the cell would still not be understood.

To say that we understand the overall behavior of the cell, we must be able to answer questions such as: 'How would the cell behave if we change the environment, for example, by adding or decreasing a certain substance?' and 'What is the result if a certain gene gets knocked out or over-expressed?' Slightly more sophisticated questions include: 'What gene needs to be inserted for the cell to behave in such a way' and 'What is the ideal culture medium in which to maximize the cell's ability to do such a thing?'

There is no doubt that computer simulation is required to understand dynamic behaviors of cellular metabolism. Virtual experiments in computers, often called '*in silico*' experiments, are not just useful but indispensable for biology because of the large amount of data generated from the genome, proteome, transcriptome and metabolome projects.

'Customized medicine' based on single nucleotide polymorphisms (SNPs) in which an appropriate dose of medicine is selected and administered to a specific patient based on his/her SNP data, is not an exception. When a large amount of SNP information from many different individuals is collected, the susceptibility of the patient to a certain drug can be statistically predicted by analysing a specific locus of a specific gene in his/her genome for SNPs. This statistical method is useful when a polymorphism of one specific gene causes a polymorphism of the phenotype (monogenic), in which case, analysis of the genotype can lead to a precise conclusion, without *in silico* experiments, as to what phenotype or physical trait that person has. It is not so easy, however, to understand a phenotype in which multiple genes are involved (polygenic).

For example, suppose it has been shown statistically that a susceptibility to a certain drug is high (+) when a person possesses a type A gene; the susceptibility is also high (+) when possessing a type B gene at another locus. When a person has both A and B gene types, however, there is no guarantee that his/her susceptibility to that drug will be twice as high (+/+); it is possible that his/her susceptibility would be low (–). Furthermore, suppose having a type C gene at a different locus is statistically known to result in low susceptibility (–). Then, what will happen when a patient has A and C gene types? It is possible, for example, that the susceptibility is high (+) when the person only has A and C gene types but very low (–/–) if the person has a combination of B and C gene types, and high (+) again with combination of A, B and C, and so on. This non-monotonic nature of the genetic system is the reason why statistical analyses alone are limited in understanding cellular behaviors. Computer models and *in silico* experiments are necessary to understand and predict phenotypes of the cell, especially when they are polygenic phenotypes. After all, most biological and pathological phenomena in which the pharmaceutical industry has a great interest, such as cancer and allergy, are polygenic.

It is still an open question as to whether or not it is feasible to construct a computer model of a whole living cell that is sufficiently sophisticated to predict answers to the types of questions mentioned above. It is thought that the task of whole-cell modeling is too difficult to be achieved in the foreseeable future. Although no attempts were made towards whole-cell modeling until the late 1990s, the importance of computer simulation of cellular metabolisms has been suggested and emphasized since the 1980s. Cellular processes that are important subsystems of the cell have been modeled and simulated by many different groups: the regulation of gene expression[1–5], cell cycle[6,7], signal transduction[8] and metabolic pathways[9–12]. However, although these models made significant contributions to the development of *in silico* biology, the programs were only able to handle specific subsystems, and it was difficult to combine different subsystem models into one single-cell model.

The first cell model

To conquer and directly challenge the task of whole-cell modeling, the E-CELL Project (Ref. 13) was initiated in 1996 at the Shonan-Fujisawa Campus of Keio University (Fujisawa, Japan), following the publication of the entire genome sequence of *Mycoplasma genitalium* (http://www.tigr.org/tdb/mdb/mdbcomplete.html). *M. genitalium* has the smallest genome (580 kb) and the smallest number of genes (~480) of all living organisms currently known and its genomic sequences have been published (see http://www.tigr.org/). The size of its genome is one order of magnitude less than that of *Escherichia coli*, and thus is an ideal candidate for

Masaru Tomita
Institute for Advanced Biosciences,
Keio University,
5322 Endo, Fujisawa,
252-8520, Japan.
e-mail: mt@sfc.keio.ac.jp

 0167-7799/01/$ – see front matter PII: S0167-7799(01)01636-5

Table 1. A summary of the 127 genes of the self-surviving cell

Gene type	*Mycoplasma genitalium*	Other	Total
Glycolysis	9	0	9
Lactate fermentation	1	0	1
Phospholipid biosynthesis	4	4	8
Phosphotransferase system	2	0	2
Glycerol uptake	1	0	1
RNA polymerase	6	2	8
Amino acid metabolism	2	0	2
Ribosomal L subunit	30	0	30
Ribosomal S subunit	19	0	19
rRNA	2	0	2
tRNA	20	0	20
tRNA ligase	19	1	20
Initiation factor	4	0	4
Elongation factor	1	0	1
Protein coding genes	98	7	105
RNA coding genes	22	0	22
Total	120	7	127

The column labeled '*M. genitalium*' represents the number of genes taken from the genome of *Mycoplasma genitalium*. The column labelled 'Other' represents the number of genes that are not found in the gene list of *M. genitalium* and thus have been taken from other organisms, such as *Escherichia coli*. The 127 genes include 22 RNA coding genes.

whole-cell modeling. Approximately 80% of its 480 genes share nucleotide sequence homology with other genes of which the function could be predicted when searched for on BLAST. The functions of the remaining 20% of the genes, however, are still unknown and it was soon realized that it would be difficult to reconstruct the whole organism without knowing the functions of the remaining genes. However, it has been demonstrated in intensive gene-knock-out studies by The Institute for Genomic Research (TIGR) that many of the 480 genes are not always necessary for *M. genitalium* to survive. Therefore, in collaboration with TIGR, a minimal set of genes sufficient and necessary for its survival and for maintenance of its homeostasis were selected (Table 1) and the first hypothetical 'virtual cell' with the 127 genes was constructed (Fig. 1).

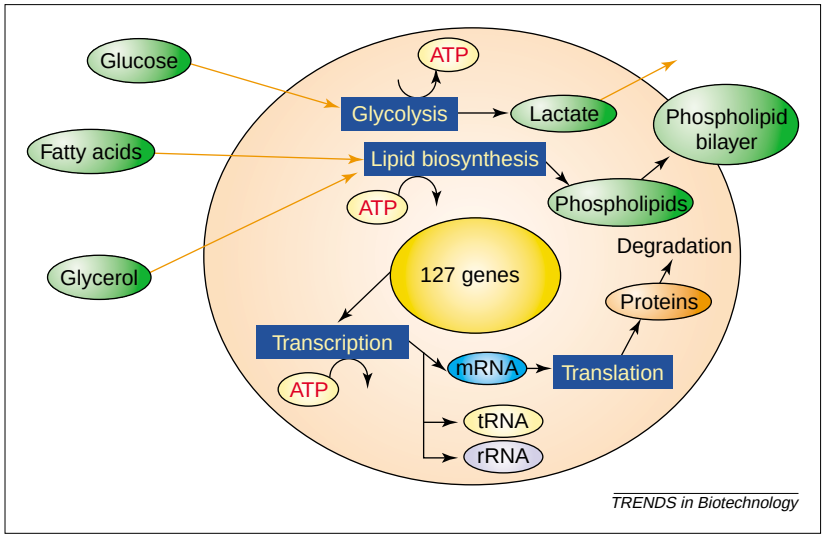

Fig. 1. The 'self-surviving cell model'. This minimal cell has 127 genes, just sufficient to maintain protein and membrane structure, by generating ATP through the glycolysis pathway.

This virtual self-surviving cell (SSC) model takes up glucose into the cytoplasm, metabolizes the glucose through the glycolysis pathway and produces ATP as an energy source. The ATP is consumed mainly for protein synthesis; the 127 genes are transcribed by RNA polymerase into mRNAs, and then translated into proteins by ribosome. Proteins are modeled to degrade spontaneously over time and so the cell has to constantly produce protein to sustain life. The membrane structure of the cell is also modeled to degrade over time; thus, the cell has a phospholipid biosynthesis pathway for biosynthesis of the cell membrane, uptaking fatty acid and glycerol, consuming ATP and generating a phospholipid bilayer, which forms a cell membrane. A constant supply of energy (ATP) is required to maintain protein and membrane synthesis, and thus glucose is essential for the survival of the virtual cell.

E-CELL simulation system

The SSC model has 105 protein-coding genes (Table 2) and 22 RNA-coding genes, and consists of 495 reaction rules. Each reaction rule defines what to do within one single time step (one millisecond, in this case). Reactions include: (1) enzymatic reactions that increase and decrease the quantity of its substrate(s) and product(s), respectively; (2) complex formations, in which multiple substrates form a complex; (3) transportations that change the location of certain substances; and (4) stochastic processes such as a transcriptional factor binding to a specific site of the chromosome. The E-CELL system also accepts user-defined reactions, making it capable of handling many other phenomena such as diffusion and variable cell volume.

When simulation of the SSC starts, all the reaction rules are executed in parallel (in practice pseudo-parallel), and the overall behavior of the cell can be observed through various graphic interfaces (Fig. 2). The dynamic changes in the amount (molecular number) of various substances inside the cells can be observed using an interface ('Tracer Window', presented as 'Traced substances'). In Fig. 2, C00031E and C00186E show temporal patterns of change in extracellular glucose and lactic acid (waste end product), respectively. It can be seen that the cell is 'living' steadily because extracellular glucose is gradually decreasing and lactic acid is reciprocally increasing. If, at a certain time point, the extracellular glucose level is set to 0, the cell starts to starve and will eventually become incapable of glucose uptake. The activity of a specific biochemical reaction can be monitored using the 'Reactor Window' (Fig. 2, 'Reactor' panel), which shows the number of molecules being processed within a single time step. The amount of substrates can be altered by the user, even in the middle of a simulation, using the 'Substance Window' (Fig. 2, 'Substance' panel).

Finally, the expression of all genes can be monitored simultaneously using the 'Genemap Window' (Fig. 2).

Table 2. A list of protein-coding genes in the virtual cell

ID	Name	ID	Name
MG005	Serine–tRNA ligase	MG215	6-Phosphofructokinase (pfkA)
MG021	Methionine–tRNA ligase	MG216	Pyruvate kinase (pyk)
MG023	Fructose–bisphosphate aldolase (tsr)	MG232	Ribosomal protein L21
MG033	Glycerol uptake facilitator (glpF)	MG234	Ribosomal protein L27
MG035	Histidine–tRNA ligase	MG249	RNA polymerase sigma S subunit
MG036	Aspartate–tRNA ligase	MG251	Glycine–tRNA ligase
MG038	Glycerol kinase (glpK)	MG253	Cysteine–tRNA ligase
MG041	Protein histidine (HPr)(ptsH)	MG257	Ribosomal protein L31
MG069	Phosphotransferase enzymeII (ptsG)	MG266	Leucine–tRNA ligase
MG070	Ribosomal protein S2	MG283	Proline–tRNA ligase
MG081	Ribosomal protein L11	MG292	Alanine–tRNA ligase
MG082	Ribosomal protein L1	MG300	Phosphoglycerate kinase (pgk)
MG087	Ribosomal protein S12	MG301	G3PD (gapA)
MG088	Ribosomal protein S7	MG311	Ribosomal protein S4
MG089	Elongation Factor G	MG325	Ribosomal protein L33
MG090	Ribosomal protein S6	MG334	Valine–tRNA ligase
MG092	Ribosomal protein S18	MG340	RNA polymerase beta subunit
MG093	Ribosomal protein L9	MG341	RNA polymerase beta subunit
MG111	Phosphoglucose isomerase B (pgiB)	MG344	Lipase
MG113	Asparagine–tRNA ligase	MG345	Isoleucine–tRNA ligase
MG114	PGP synthase (pgsA)	MG351	Inorganic pyrophosphate (ppa)
MG126	Tryptophan–tRNA ligase	MG361	Ribosomal protein L10
MG136	Lysine–tRNA ligase	MG362	Ribosomal protein L7
MG142	Translation initiation factor 2	MG363	Ribosomal protein L32
MG150	Ribosomal protein S10	MG363.01	Ribosomal protein S20
MG151	Ribosomal protein L3	MG365	Methionyl-tRNA formyltransferase
MG152	Ribosomal protein L4	MG375	Threonine–tRNA ligase
MG153	Ribosomal protein L23	MG378	Arginine–tRNA ligase
MG154	Ribosomal protein L2	MG407	Enolase (eno)
MG155	Ribosomal protein S19	MG417	Ribosomal protein S9
MG156	Ribosomal protein L22	MG418	Ribosomal protein L13
MG157	Ribosomal protein S3	MG424	Ribosomal protein S15
MG158	Ribosomal protein L16	MG426	Ribosomal protein L28
MG159	Ribosomal protein L29	MG429	Protein phosphotransferase (ptsI)
MG160	Ribosomal protein S17	MG430	Phosphoglycerate mutase (pgm)
MG161	Ribosomal protein L14	MG431	Triosephosphateisomerase (tpiA)
MG162	Ribosomal protein L24	MG433	Transcription elongation factor Ts
MG163	Ribosomal protein L5	MG437	CDP-diglyceride synthetase (cdsA)
MG164	Ribosomal protein S14	MG444	Ribosomal protein L19
MG165	Ribosomal protein S8	MG446	Ribosomal protein S16
MG166	Ribosomal protein L6	MG451	Transcription elongation factor Tu
MG167	Ribosomal protein L18	MG455	Tyrosine–tRNA ligase
MG168	Ribosomal protein S5	MG460	L-lactate dehydrogenase (ldh)
MG173	Translation initiation factor 1	MG462	Glutamate–tRNA ligase
MG174	Ribosomal protein L36	MG466	Ribosomal protein L34
MG175	Ribosomal protein S13	V1001	Nucleoside-phosphate kinase
MG176	Ribosomal protein S11	V1002	Nucleoside-diphosphate kinase
MG177	RNA polymerase alpha core subunit	V1003	Glutamine–tRNA ligase
MG178	Ribosomal protein L17	V4001	Acylglycerol lipase
MG194	Phenylalanine–tRNA ligase alpha	V4002	Glycerol-1-phosphatase
MG196	Translation initiation factor 3	EC1238	Phosphatidylglycerophosphatase
MG197	Ribosomal protein L35	EC3928	Diacylglycerolkinase (dgkA)
MG198	Ribosomal protein L20		

These 105 genes plus 22 RNA coding genes constitute the minimal cell.

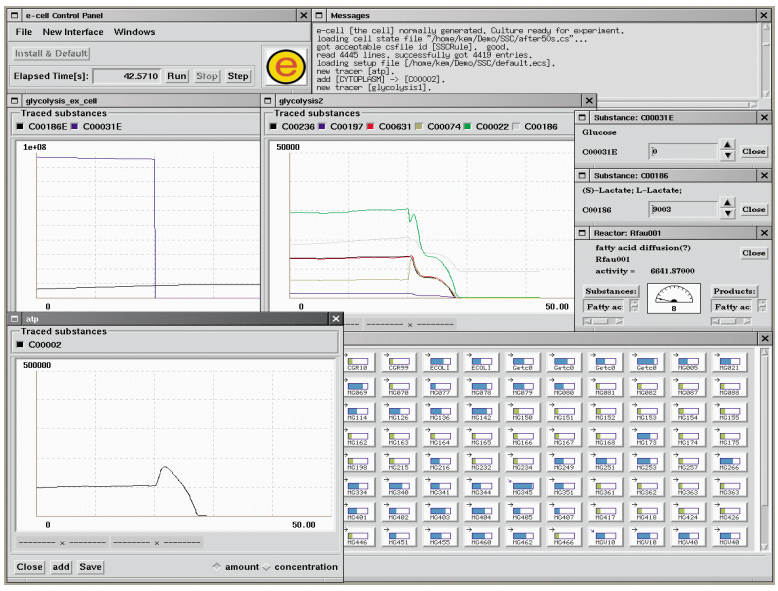

Fig. 2. Feature of the E-CELL simulation system. The user can monitor dynamic changes of substance quantities and reaction fluxes via various graphical interfaces. Quantities of any substance can also be altered during the simulation.

Each icon represents each corresponding gene, and exhibits the amount of mRNA of that gene (molecular number). A specific gene can be easily knocked-out by clicking its icon; it is thus possible to conduct real-time knock-out experiments *in silico*. Clicking the icon again can reactivate a knocked-out gene.

Final 'desperate efforts' before starvation

Even this simple cell model sometimes shows unpredictable behavior and has delivered biologically interesting surprises. When the extracellular glucose is drained and set to be zero, intracellular ATP momentarily increases and then decreases (Fig. 2, 'Traced substances' panel, substance ID C00002). At first, this finding was confusing. Because ATP is synthesized only by the glycolysis pathway, it was assumed that ATP would decrease when the glucose, the only source of energy, becomes zero. After months of checking the simulation program and the cell model for errors, the conclusion is that this observation is correct and a rapid deprivation of glucose supplement can lead to the same phenomenon in living cells.

The reason is, in fact, clear. In glycolysis, one molecule of glucose produces two molecules of ATP. In more detail, however, two molecules of ATP are consumed in the first part of glycolysis and then four molecules of ATP are synthesized in the second part, producing two net molecules of ATP. Thus, when glucose is rapidly deprived, first consumption of ATP stops and then, a few moments later, ATP synthesis stops. The slight difference in time causes a transient increase in ATP. This momentary increase in the level of ATP, is of great interest and is as if the cell undergoes a final 'desperate effort' when it is rapidly starved.

Virtual erythrocytes

Obviously, the SSC model described above is only a hypothetical cell; no such cells exist in nature. Thus, it was decided to model living cells so that the simulation results could be evaluated. Human erythrocytes were chosen for the model because intracellular metabolism is limited in human erythrocytes and because they do not replicate, transcribe or translate genes; also, there are already several studies on the modeling of erythrocytes[14–16]. It is possible to compare computer models with real red blood cells because a considerable amount of experimental data about red blood cells has accumulated[17,18]. Recently, the construction of a prototype of human erythrocytes using the E-CELL System has been completed (Fig. 3).

In human erythrocytes, the major metabolic pathways are glycolysis, the pentose phosphate pathway and nucleotide metabolism. In addition, there are abundant hemoglobins that carry oxygen from the lungs to peripheral tissues. Thus, the erythrocyte is a 'bag' packed with hemoglobins that maintain several metabolic pathways for ion transport to maintain homeostasis of intracellular osmosis.

An E-CELL model of the human erythrocyte has been developed, by defining reaction rules for all these metabolisms based on the previous erythrocyte model[14–17]. All the kinetic equations and parameters used in the model were obtained from previously published experimental data. After tuning the kinetic parameters, the E-CELL erythrocyte model reached a steady state in which quantities of intermediate metabolites inside the virtual cell are comparable with the experimental data of living erythrocytes. We are currently extending and improving the erythrocyte model for a more accurate simulation, by taking into account osmotic pressure, pH and variable cell volume.

Using the E-CELL for pathological analyses

It is possible to perform *in silico* experiments in which the function of an enzyme is inhibited, and to simulate the behavior of human erythrocytes from hereditary anemic patients using the E-CELL model (Fig. 4). Using the simulated erythrocytes on the E-CELL program, the activity of aldolase is blocked in our virtual erythrocytes; aldolase (fructose bisphosphate aldolase) converts fructose-1,6-bis-phosphate (X12) to glyceraldehyde-3-phosphate (X14) and dihydroxy-acetone-phosphate (X13). Aldolase-deficient humans are susceptible to hemolytic anemia. With aldolase-deficiency, the reactant fructose-1,6-bis-phosphate (X12) is markedly increased and accumulated, whereas the reaction products glyceraldehyde-3-phosphate (X14), dihydroxy-acetone-phosphate (X13) and further metabolites downstream of this reaction are significantly decreased. Although this is an expected result, such *in silico* experiments might give unexpected results, as we have seen in the SSC model.

One major advantage of these *in silico* experiments is that they can be repeated automatically, for all enzymes at various levels of inhibition, resulting in lists of enzyme deficiencies that could kill, perturb, or preserve the cell. The results might suggest some biologically interesting phenomena, which could then be verified by *in vitro* or *in vivo* experiments. In addition, the results

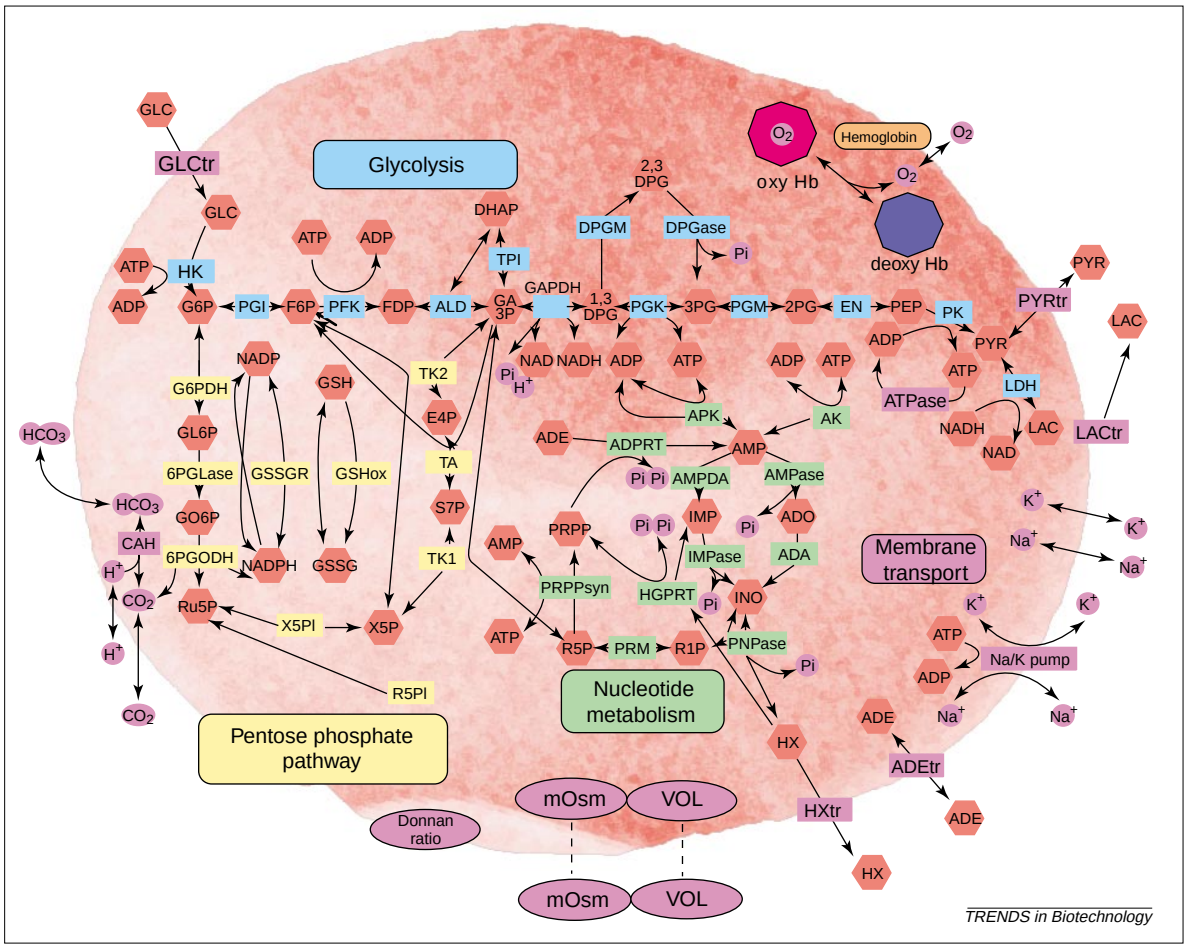

Fig. 3. The human erythrocyte model. It has three major metabolic pathways: (1) glycolysis; (2) the pentose phosphate pathway; and (3) nucleotide metabolism, as well as Na^+/K^+ pumps, transport systems, and magnesium complexation. Abbreviations: ADA; adenosine deaminase; ADE, adenine; ADPRT, adenine phosphoribosyl transferase; AK, adenosine kinase; ALD, aldolase; AMP; adenosine monophosphate; AMPDA, adenosine monophosphate deaminase; APK, adenylate kinase; CAH, carbonic anhydrase; DHAP, dihydroxy acetone phosphate; DPG, diphosphogrycerate; DPGase, diphosphogrycerate phosphatase; DPGM, diphosphogrycerate mutase; EN, enolase; E4P, erythrose 4-phosphate; FDP, fructose 1,6-diphosphate; F6P, fructose 6-phosphate; G6P glucose 6-phosphate; GA3P, glyceraldehyde 3-phosphate; GAPDH, glyceraldehyde phosphate dehydrogenase; GLC, glucose; GLCtr, glucose transport process; GO6P, gluconate 6-phosphate; GSH, glutathione; GSHox, glutathione turnover; GSSGR, glutathione reductase (NADPH); Hb, hemoglobin; HGPRT, hypoxanthanine-guanine phosphoryl transferase; HK, hexokinase; HX, hypoxanthine; HXtr, hypoxanthine transport process; IMP, inosine monophosphate; IMPase, inosine monophosphatase; INO, inosine; LAC, lactate; LACtr, lactate transport process; LDH, lactate dehydrogenase; mOsm, osmolarity; 2PG, 2-phosphogrycerate; 3PG, 3-phosphogrycerate; 6PGLase, 6-phosphogluconolactonase; 6PGODH, 6-phosphogluconate dehydrogenase; PEP, phospho-enolpyruvate; PFK, phosphofructokinase; PGI, phosphoglucoisomerase; PGM, phosphoglyceromutase; Pi, inorganic phosphate; PNPase, purine nucleotide phosphorylase; PRM, phosphoribomutase; PRPP, 5-phosphoribosyl 1-phosphate; PRPPsyn, phosphoribosyl pyrophosphate synthetase; PYRtr, pyruvate transport process; PYR, pyruvate; R5P, ribose 5-phosphate; RIP, ribose 1-phosphate; Ru5P, ribulose 5-phosphate; S7P, sedoheptulose 7-phosphate; TA, transaldolase; TK1, transketolase I; TPI, triose phosphate isomerase; VOL, volume; X5P, xylulose 5-phosphate; X5PI, xylulose 5-phosphate isomerase.

would provide kinetic sensitivities of different enzymes. If a small change in enzyme activity results in a dramatic change inside the cell, the enzymatic reaction is 'sensitive' and therefore would need accurate kinetic parameter values. However, other enzymatic reactions that are not kinetically sensitive would need only rough values for their kinetic parameters.

In this way, it will be possible to reproduce the behavior of human erythrocytes not only in normal physiological conditions but also in abnormal pathological conditions. Thus, one could investigate how an abnormal anemic condition occurs in human erythrocytes, for example, in patients with hereditary oxygen deficiency or patients with drug-induced inhibition of specific enzymes. Because there are many known cases of anemia with abnormal erythrocytes, computer simulation will have an important role in the development of remedies of many human diseases.

SNPs and other types of genetic polymorphism could result in different enzymes with different kinetic parameters, producing different fluxes and therefore different metabolisms inside the cell. By assigning kinetic parameters based on SNP information, it might become possible to conduct 'personalized simulation' of an individual's cellular metabolism. Such simulation could, in the future, play an important role in 'customized medicine', by predicting appropriate drugs and dosage for individual patients.

Future prospects

In addition to the 'virtual self-surviving cell' and the 'human erythrocyte model' described other E-CELL models are currently under construction; a

Acknowledgements
We thank Koichi Takahashi, Kenta Hashimoto and Yoichi Nakayama for their leaderships in developing the basic E-CELL software, the Virtual Self-Surviving Cell Model, and the Human Erythrocyte Model, respectively. Special thanks are extended to many other colleagues in the Laboratory for Bioinformatics of Keio University, including Yuri Matsuzaki, Katsuyuki Yugi, Fumihiko Miyoshi, Yusuke Saito and Naota Ishikawa, for their dedication to the E-CELL project. The author acknowledges the support of Japan Science and Technology Corporation (JST) and the Ministry of Education, Japan.

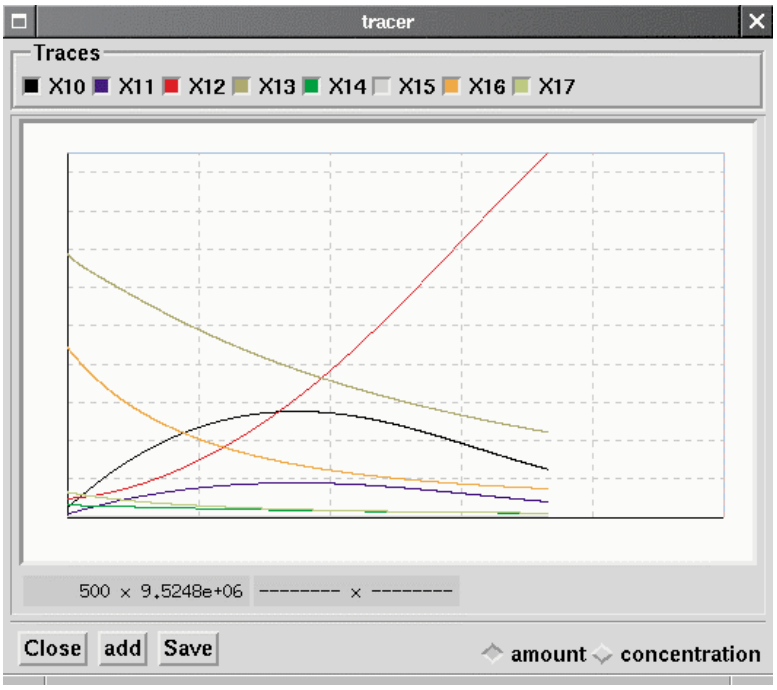

Fig. 4. Human erythrocyte with aldolase deficiencies. The E-CELL system can be used to conduct virtual experiments for pathological analyses by substituting enzymatic parameters.

'mitochondria model' and a 'signal transduction model' for the chemotaxis of *E. coli*. Examples of other successful systems for integrative simulation of the cell include DBSolve by Goryanin and colleagues[19] and the V-Cell by Schaff and co-workers[20].

One of the major problems in constructing large-scale cell models is lack of quantitative data. Most of the biological knowledge available is qualitative (such as functions of genes, pathway maps, which proteins interact with what) but for simulation quantitative data (such as concentrations of metabolites and enzymes, flux rates, kinetic parameters and dissociation constants) are needed. A major challenge is to develop high-throughput technologies for measurement of inner-cellular metabolites. A large amount of data for a variety of cell states can then be collected with the technologies to construct quantitative models, and the models can be refined iteratively until the simulation results match the data.

For this new type of simulation-orientated biology, we set up the Institute for Advanced Biosciences of Keio University (http://bioinfo.sfc.keio.ac.jp/IAB/). The institute consists of three centers for metabolome research, bioinformatics, and genome engineering, respectively. The ultimate goal of this international research institute is to construct a whole-cell model *in silico* based on a large amount of data generated by high-throughput metabolome analyses, and then to design a novel genome based on the computer simulation and create real cells with the novel genome by means of genome engineering.

In the USA, the National Institute of General Medicine Sciences (NIGMS) has announced the Alliance for Cellular Signaling (AFCS) for analyzing signal transduction of cardiomyocytes and B cells from the mouse. The ultimate goal of the project is 'to create virtual cells' of these particular cell types (http://www.nih.gov/news/pr/sep2000/nigms-05.htm). In addition, the Department of Energy (DOE) has set up a Microbial Cell Project (MCP), of which the ultimate goal is to construct a 'virtual microbe' *in silico* (http://www.microbialcellproject.org), and the Special Interest Group of Biological Simulation (SIGSIM) has been formed under the International Society for Computational Biology (ISCB) for the purpose of effective communication among scientists working on large scale cell modeling (http://www.iscb.org).

Summary

The cell is never 'conquered' until its total behavior is understood and the total behavior of the cell is never understood until it is modeled and simulated. Whole-cell modeling, which was thought intractable until recently, has suddenly become realistic. There is no doubt that *in silico* construction of complex living cells is an exciting scientific challenge and we are just opening the door to this new area of biological research in the 21st century.

References

1 Meyers, S. and Friedland, P. (1984) Knowledge-based simulation of genetic regulation in bacteriophage lambda. *Nucl. Acids Res.* 12, 1–9
2 Koile, K. and Overton, G.C. (1989) A qualitative model for gene expression. *Proc. of the 1989 Summer Comp. Simu. Conf.* 415–421
3 Karp, P.D. (1993) A qualitative biochemistry and its application to the regulation of the tryptophan operon. In Hunter, L. (ed) *Artificial Intelligence and Molecular Biology* pp. 289–324 AAAI Press/ The MIT Press, MA, USA
4 McAdams, H.H and Arkin, A. (1997) Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci. U. S. A.* 94, 814–819
5 Leloup, J.C. and Goldbeter, A. (2000) Modeling the molecular regulatory mechanism of circadian rhythms in *Drosophila*. *Bioessays* 22, 84–93
6 Novak, B. and Tyson, J.J. (1995) Quantitative analysis of a molecular model of mitotic control in fission yeast. *J. Theor. Biol.* 173, 283–305
7 Chen, K. *et al.* (2000) Kinetic analysis of a molecular model of the budding yeast cell cycle. *Mol. Biol. Cell* 11, 369–391
8 Morton-Firth, C.J. and Bray, D. (1998) Predicting temporal fluctuations in an intracellular signalling pathway. *J. Theor. Biol.* 192, 117–128
9 Cornish-Bowden, A. and Hofmeyr, J.H. (1991) MetaModel: a program for modeling and control analysis of metabolic pathways on the IBM PC and compatibles. *Comput. Appl. Biosci.* 7, 89–93
10 Sauro, H.M. (1993) SCAMP: a general-purpose simulator and metabolic control analysis program. *Comput. Appl. Biosci.* 9, 441–450
11 Mendes, P. (1993) GEPASI: a software package for modeling the dynamics, steady states and control of biochemical and other systems. *Comput. Appl. Biosci.* 9, 563–571
12 Mendes, P. (1997) Biochemistry by numbers: simulation of biochemical pathways with Gepasi 3. *Trends Biochem. Sci.* 22, 361–363
13 Tomita, M. *et al.* (1999) E-CELL: Software environment for whole cell simulation. *Bioinformatics* 15, 72–84
14 Palsson, B.O. *et al.* (1989) Computer model of human erythrocyte metabolism. *Prog. Clin. Biol. Res.* 319, 133–150
15 Ni, T.C. *et al.* (1996) Model assessment and refinement using strategies from biochemical systems theory: application to metabolism in human red blood cells. *J. Theor. Biol.* 179, 329–368
16 Lee, I.D. *et al.* (1992) A Macintosh software package for simulation of human red blood cellmetabolism. *Comput. Meth. Prog. Biomed.* 38, 195–226
17 Joshi, A. *et al.* (1989) Metabolic dynamics in the human red cell. *J. Theor. Biol.* 141, 515–545
18 Joshi, A. *et al.* (1990) Metabolic dynamics in the human red cell. *J. Theor. Biol.* 142, 41–85
19 Schaff, J. *et al.* (1997) A general computational framework for modeling cellular structure and function. *Biophys. J.* 73, 1135–1146
20 Goryanin, I. *et al.* (1999) Mathematical simulation and analysis of cellular metabolism and regulation. *Bioinformatics* 15, 749–758