OBIYagns: a grid-based biochemical simulator with a parameter estimator

Shuhei Kimura1,*, Takuji Kawasaki2, Mariko Hatakeyama1, Takashi Naka3, Fumikazu Konishi1 and Akihiko Konagaya1

1Bioinformatics Group, RIKEN Genomic Sciences Center, 1-7-22 Suehirocho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan, 2Fuji Research Institute Corporation, 2-3 Kanda Nishikicho, Chiyoda-ku, Tokyo 101-8443, Japan and 3Department of Intelligent Informatics, Faculty of Information Science, Kyushu Sangyo University, 2-3-1 Matsukadai, Higashi-ku, Fukuoka 813-8503, Japan

Received on August 4, 2003; revised on November 21, 2003; accepted on January 19, 2004
Advance Access publication February 12, 2004

ABSTRACT

Summary: OBIYagns (yet another gene network simulator) is a biochemical system simulator that comprises a multiple-user Web-based graphical interface, an ordinary differential equation solver and a parameter estimators distributed over an open bioinformatics grid (OBIGrid). This grid-based biochemical simulation system can achieve high performance and provide a secure simulation environment for estimating kinetic parameters in an acceptable time period. OBIYagns can be applied to larger system biology-oriented simulation projects.

Availability: OBIYagns example models, methods and user guide are available at https://access.obigrid.org/yagns/

Contact: konagaya@gsc.riken.jp

Supplementary information: Please refer to Bioinformatics online.

INTRODUCTION

Computer simulations of mathematical models of cellular signaling networks have been recognized as useful tools for biological researchers seeking to understand the system dynamics and regulation mechanism of cell behavior and to set up subsequent working hypotheses in their studies (Hatakeyama et al., 2003). Recently, several programs such as Gepasi (Mendes, 1993), E-Cell (Tomita et al., 1999), Virtual Cell (Loew and Schaff, 2001) and GENESIS (Sivakumaran et al., 2003) have been made available for such purposes. Mathematical modeling requires the appropriate values for a set of parameters (kinetic constants and physiological concentration of the cellular components); however, at present, there are no experimental methods that measure all such parameters with accuracy. Therefore, in order to construct a model, it is essential to estimate biological parameters based on available observable wet-lab data (Swameye et al., 2003).

For this purpose, biochemical simulator programs have been developed with a parameter estimation function (Mendes and Kell, 1998). However, these parameter estimators do not work properly when many unknown parameters are being solved at the same time because the parameter estimation process requires many computational resources beyond those a single personal computer can provide. Therefore, high-performance computing systems are becoming necessary for large-scale biochemical network simulation.

OBIYagns (yet another gene network simulator) incorporates parameter estimators distributed over an open bioinformatics grid (OBIGrid; Konagaya, 2002). OBIYagn’s Web-based graphical interface lets biological researchers use the system without any knowledge of parallel processing or distributed processing, which is necessary for high-performance computing based on the grid computing system (Foster et al., 2001; Russel et al., 2002).

The core of OBIYagns is composed of an ordinary differential equation (ODE) solver and a parameter estimator. It accepts as user inputs lists of reactant concentration, kinetic constants such as mass action and Michaelis–Menten type reactions, and formulas for biochemical reactions written in chemical notation and graphically displays simulation results (Fig. 1). Biochemical reactions are automatically converted into ODEs and then solved by the ODE solver based on the numerical differentiation formula (NDF), which is suitable for solving the stiff problems that often appear in equations for biological systems (Michalski et al., 1998; Stevenson et al., 1984). The NDF is an implicit formula that has better stability than the Gear method (Shampine and Reichelt, 1997). As for the unknown parameter estimation function, the problem of estimating unknown parameters is formulated as a function optimization problem in OBIYagns. When observed wet-data are provided, the system automatically generates an objective function of the optimization problem. The parameter estimator employs an in-house genetic algorithm called distance...
Fig. 1. OBIYagns operation screens: (1) selection of reaction types, such as mass action and Michaelis–Menten action; (2) input of kinetic parameters; (3) input of initial concentrations of reactants; (4) input of fitting data for parameter estimation (if needed); (5) input of fitting expressions (if needed); (6) input of GA parameters; and (7) output of results as a graphical presentation.

independent diversity control (DIDC) as a function optimizer. Although DIDC can solve a large number of ODEs by estimating many unknown parameters, the computational cost is high. OBIYagns is able to resolve this problem by executing DIDC in parallel in the grid (Kimura and Konagaya, 2003a,b).

OBIGrid is a practical grid system being developed by the Japan Committee on Very Large Biocomputing (VLSB; http://www.obigrid.org/vlsbio/) and the Initiative for Parallel Bioinformatics Processing (IPAB; http://www.ipab.org). The grid comprises computers connected over the Internet that are voluntarily provided by universities, national research institutes in HPB and enterprises in IPAB. At present, OBIGrid comprises 492 CPUs at 27 locations. The grid is equipped with functions that are indispensable in the life sciences fields, such as database systems, computation systems and access control systems, as well as various bioinformatics application suites protected by a virtual private network (VPN). OBIYagns is expected to become a large-scale simulation tool for biological research. Thus OBIYagns can be used for large computer simulation projects in the area of system biology.

REFERENCES


