

비주얼 프로그래밍 환경하에서 가상 *Escherichia Coli* 시스템 개발에 관한 연구

이성근, 한상일, 황규석\*  
부산대학교 화학공학과  
(kshwang@pusan.ac.kr\*)

**Study on Development of *In Silico* System of *Escherichia Coli*  
Using a Visual Programming Environment**

Sung Gun Lee, Sang il Han, Kyu Suk Hwang\*

Department of Chemical Engineering, Pusan National University, Busan, Korea.  
(kshwang@pusan.ac.kr\*)

### 서론

가상 미생물 시스템 구축을 위한 모델링과 해석은 대사산물, 대사회로, 유전정보 및 단백질-단백질 반응 등의 많은 자료와 생물학과 물리학, 화학, 수학, 전산학, 화학공학적인 방법론들이 필요하게 된다. 이를 위해서 여러 가지 방법들이 제안되었으며, 그 중에서 FBA(flux balance analysis)와 최적화 기법(optimization method)을 함께 활용함으로써, 세포의 거동, 대사능력(metabolic capabilities), 그리고 원하는 성분을 생성시킬 수 있는 중요한 반응들을 정확하게 분석할 수 있게 되었다[5].

최근에 몇몇 연구자들은 조절 제약 조건(regulatory constraints)을 FBA(dynamic flux balance analysis)에 통합 시키려고 시도하고 있다. 이 조절 제약 조건은 생성규칙(production rule—If than rule)과 불린논리(boolean logic)로 표현된 논리식(logic equation)을 사용한다. 이러한 생성시스템(production system)은 어떤 조건에 따른 결과, 즉 상황을 나타내는 지식을 쉽게 규칙으로 표현할 수 있는 장점이 있다. 반면에 이러한 불린 규칙은 (I) binary system이기 때문에 복잡한 조절 단백질의 상호관계를 표현할 수 가 없는 문제가 발생될 수 있다 [2, 4, 5]. (II)어떤 규칙들이 순서대로 적용되었는지 알 수 없기 때문에 [7], 외부 환경 변화에 따른 유전자들의 조절 메커니즘을 알 수 가 없었다. 또한 (III)최적의 대사 흐름 분석, 최대 성장률, 그리고 세포의 거동을 예측하기 위해서, linear programming package(LINGO)와 spreadsheet package (EXCEL)을 사용하여왔다[1, 2]. 이러한 프로그램은 예측 결과를 비주얼하게 표현하는 데에 한계가 있고, 사용자가 모델의 입력 변화의 영향을 쉽게 알 수가 없었다.

위에서 언급한 것들을 해결하기 위해서, (I)조절 요소들(regulatory protein, genes, effectors etc.) 사이의 관계를 더욱더 유연하게 표현하기 위해서 결합의 정도를 나타내는 가중치 개념을 도입하였다. (II)외부환경 변화에 따른 유전자들의 조절 메커니즘을 표현하기 위해서, 정의된 기호로 계층적 조절 네트워크(hierarchical regulatory network)를 모델링하였다. (III)사용자가 쉽게 정량적으로 예측할 수 있고, 비주얼한 결과를 보여줄 수 있는 통합 시스템을 개발하였다. 프로그램 개발은 비주얼 프로그램 환경인 LabVIEW에서 실행되었다.

## 본론

### 1. Visual Programming

모든 프로그래밍 과정은 LabVIEW version 7.1(National Instruments Ltd.)에서 만들어 졌고, 1GByte RAM을 가진 IBM Computer에서 실행되어졌다. 최적의 플러스 분배와 최대 성장률을 구하기 위해서, LabVIEW에서 지원하는 Sub.VI인 Linear Programming Simplex Method.VI를 사용하였다.

### 2. Fundamental Modeling Symbols

계층적 조절 네트워크를 구성하고 있는 오페론은 기호(Symbol), 간선(Arc), 가중치(Weight)로 구성되어 있다. 모델링 기호는 조절(regulation)을 위한 생물학적 기호들과 가중치를 위한 비생물학적 기호들로 구분된다. 프로모터, not gate, 오퍼레이터 기호들은 하나의 클러스터를 형성하여 유전자 전사(gene transcription)의 기능을 담당한다. 간선은 각 기호들 간에 인과관계의 유무와 가중치를 전파시키는 기능을 한다. 가중치가 심볼들의 활성화유무를 결정하기 위해서 사용되지는 것 이외에, 노드사이에 결합의 정도를 나타낼 수 있다. 이는 binary system으로 나타낼 수 없는 경우에 효과적으로 사용할 수 있다.

### 3. Construction of Hierarchical Regulatory Network

정의된 기호들로 구성된 계층적 조절 네트워크는 조절 단백질들이 어떻게 유전자들의 세트를 조절하는지 직관적으로 알 수 있다. 이 네트워크는 stimulus, modulon, regulon/operon, gene level, metabolic pathway level로 나눌 수 있다. Stimulus level은 단순한 환경 변화(a simple environmental change)를 나타낸다. Modulon level은 세포 외부 환경의 변화에 반응하고, 하나 이상의 효과를 나타내는 regulatory protein에 의해서 regulon/ operon을 통합적으로 조절하는 기능을 한다. Regulon/operon level은 common regulatory protein에 의해서 직접적으로 오페론들의 세트 또는 유전자들의 세트를 조절한다 [3]. Gene level은 오페론에 의해서 조절되는 유전자들의 active /inactive를 나타낸다. 이러한 유전자들은 metabolic pathway level에서 세포성장에 필요한 대사물질을 합성하거나 분해하는 기능을 하게 된다.

### 4. Composition of *In Silico* System

*In silico* system은 크게 네가지로 구성이 되어있다; Extracellular Environment Setting Mode and Constraints, Optimal Flux Distributions and Time Profile for Substrates, Dynamic Flux Distribution, Dynamic Hierarchical Regulatory Network.

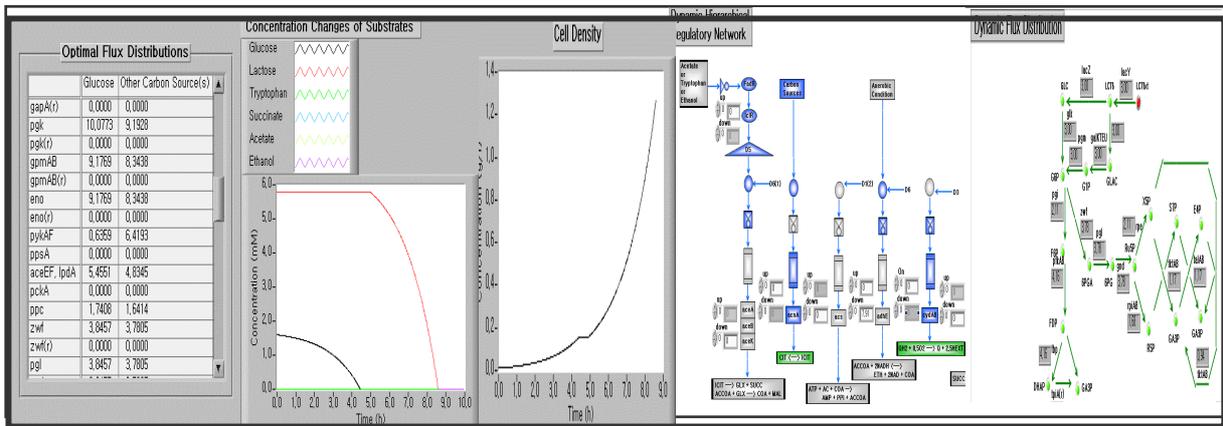


Fig. 1 *In silico* system of *Escherichia coli*.

#### Extracellular Environment Setting Mode and Constraints

세포가 어떠한 조건하에서 성장하는지를 설정하는 것으로써, 8개의 checkboxes-aerobic/anaerobic condition, glucose, lactose, succinate, acetate, tryptophan로 구성되어 있다. 최적의 흐름 분산을 구하기 위해서는 *E. Coli*의 maximal uptake rate와 cell density를 지정하여야 한다. 사용자는 단지 각각의 해당 box에 수치값을 지정하면 된다. 이러한 작업을 통해서 simulation을 위한 초기 조건 지정이 완료된다.

#### Optimal Flux Distributions and Time Profile for Substrates

원하는 생리학적 특성(objective function: growth rate)을 최적화 함으로써, 최적 흐름 분산을 테이블로 보여준다. 특히, global control mode에 의한 최적 흐름의 변화를 직접적으로 비교할 수 있고, 시간에 따른 기질의 농도 변화를 차트로 보여준다. 이것은 다른 환경조건에서 세포의 대사와 생리학적 변화를 이해하는 것을 가능하게 한다.

#### Dynamic Flux Distribution

Metabolic reaction pathway를 보여주는 dynamic graphical viewer를 제공한다. 이 viewer는 여러 환경 조건하에서 변화하는 최적 흐름의 결과값에 의해서 자동적으로 그리고 다이내믹하게 보여준다.

#### Dynamic Hierarchical Regulatory Network

조절 단백질이 환경변화에 따라서 어떻게 유전자들의 세트를 조절되는지를 보여주는 dynamic graphical viewer를 제공한다. 이 network는 stimulus level, modulon level, regulon/operon level, gene level metabolic pathway level에 의해서 계층적 구조를 이루고 있다.

## 결론

### 1. Parameters for simulation

조절 네트워크의 계층적 구조, central metabolic pathway와 parameters의 값들을 이용하여, 시간에 따른 유전자 조절을 모사함으로써, 본 연구에서 제시한 모델링이 설명되어졌다. Glucose와 lactose가 함께 존재하는 조건에서 *E. coli*를 simulation하기 위해서, 초기조건 (biomass : 0.011 g/l, glucose : 1.6 mM, lactose : 5.8 mM, delay time : 0.5h)과 흡수율제한조건 (glucose : 6.5 mM, lactose : 3.0 mM, Oxygen : 15.0mM)을 정하였다.

### 2. Glucose and Lactose in Medium

세포 외부환경에 glucose와 lactose가 있는 경우에,  $t=0(h)$ 에서  $t=4.45(h)$ 에서 glucose가 대사되었다.  $t=4.45(h)$ 에서  $t=4.95(h)$ 에서는 alternate carbon source인 lactose를 대사하기 위한 gene들이 발현되었고,  $t=4.95(h)$ 에서  $t=8.59(h)$ 구간 동안 Lactose가 대사되었다.

본 연구에서 *in silico* system은 사용자에게 예측 결과를 직관적으로 인지할 수 있는 방법을 제공할 수 있다. 정의된 기호들로 모델링한 계층적 조절 네트워크는 좋은 예이다. 이 시스템은 생물학자와 공학자에게 여러 환경 조건하에서 세포의 거동을 연구하는 할 때에 좋은 도구로서 사용될 수 있을 것이다.

## 참고문헌

- [1] W.C. Covert, C. H. Schilling, and B. O. Palsson, Regulation of gene expression in flux balance models of metabolism, *J. Theor. Biol.* **213**, 2001, 73–88.
- [2] M.W. Covert, E. M. Knight, J. L. Reed, M. J.Herrgard, and B. O. Palsson, Integrating high-throughput and computational data elucidates bacterial networks, *Nature*, **429**, 2004, 92-96.
- [3] G. N. Stephanopoulos, A. A. Aristidou, and J.Nielsen, *Metabolic Engineering: Principles and Methodologies*, 1998, pp180-193, Academicpress, London, UK.
- [4] S. G. Lee, C. M. Kim, and K. S. Hwang, "Development of a software tool for *in silico* simulation of *Escherichia coli* using a visual programming environment," *Journal of Biotechnology*, 119, 2005, 87-92.
- [5]. S. G. Lee, K. S. Hwang, and C. M. Kim, "Dynamic behavior of regulatory elements in the hierarchical regulatory network of various carbon sources-grown *Escherichia coli*," *Journal of Microbiology and Biotechnology*, 2005, 551-559.
- [6] A. Varma and B.O. Palsson, Stoichiometric flux balance models quantitatively predict growth and metabolic by-product Secretion in wild type *Escherichia coli* W3110, *Appl. Environ. Microbol.*, **60**, 1994a., 3724-3731.
- [7] P.H. Winston, *Artificial Intelligence*, 1992, pp119-137, Addison Wesley, U.S.A.