

계층적 유전자 조절 네트워크와 모사 모듈을 통합한
가상 *Escherichia coli* 시스템의 모델링에 관한 연구

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

**Study on *in silico* system of *Escherichia coli* by the combination of hierarchical
regulatory network and dynamic simulation module**

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

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

서론

최근 완전히 서열화된 미생물 지놈들을 신속하고, 정확하게 전산적 분석을 실시하여, 세포 전체의 종합적인 생화학적 대사경로를 구축할 수 있게 되었다. 이로 인해서 세포의 기능을 PC상에서 모사할 수 있는 여러 가지 *in silico* models에 관한 연구가 활발히 진행되고 있다. 그 중에서 FBA(flux balance analysis)는 상세한 속도식이 부족한 상황에서 대사 능력(metabolic capabilities), 세포 성장을, 대사 부산물, 기질의 농도와 원하는 성분을 생성시킬 수 있는 중요한 반응들을 정확하게 분석할 수 있다 [6]. 분석에 필요한 정보는 대사경로의 양론계수, 대사요구량, 그리고 기본적인 물리화학적 제약조건이며, 이러한 정보는 잘 알려져 있기 때문에 쉽게 이용할 수 있다.

최근에 몇몇 연구자들은 조절 제약 조건(regulatory constraints)을 FBA(dynamic flux balance analysis)에 통합 시키려고 시도하고 있다. 그 이유는 제약조건들이 미생물의 거동에 큰 영향을 미치기 때문이다. FBA에서 조절 제약 조건은 논리적 방법 [1]으로 표현되어질 수 있다. 이 방법을 가진 FBA는 *E. Coli*의 이단계 적응성장을 정량적으로 예측하기 위해서 사용되어졌다.

논리적 표현방법은 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 표

현할 수 가 없는 문제가 발생될 수 있고[2, 4, 5], (II)어떤 규칙들이 순서대로 적용되었는지 알 수 없기 때문에 [7], 외부 환경 변화에 따른 유전자들의 조절 메커니즘을 알 수 가 없었다.

본 연구에서는 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.

본론

1. 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 활성유무를 결정하기 위해서 사용되지는 것 이외에, 노드사이에 결합의 정도를 나타낼 수 있다. 이는 binary system으로 나타낼 수 없는 경우에 효과적으로 사용할 수 있다.

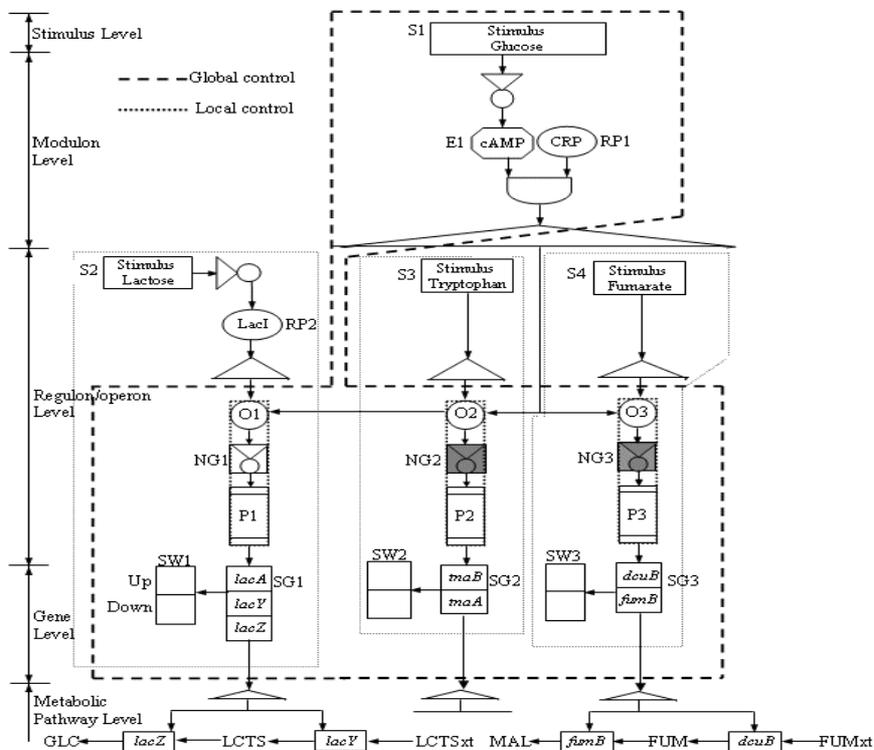


Fig. 1 Hierarchical regulatory network

2. Construction of Hierarchical Regulatory Network

정의된 기호들로 구성된 계층적 조절 네트워크(Fig. 1)는 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 세포 성장에 필요한 대사물질을 합성하거나 분해하는 기능을 하게 된다.

3. Dynamic Simulation Module

동적 모사 모듈(dynamic simulation module)은 흐름 평균 분석(flux balance analysis)과 미분식들을 이용하여 기질의 농도와 세포량을 정량적으로 모사하는 기능을 하며, 두 개의 하부모듈(submodule)인 흐름 평균 분석 모듈(flux balance analysis module)과 예측모듈(prediction module)로 이루어진다.

결론

1. Central Metabolic Network of *E. coli*

본 연구에서는 대사산물 생산의 핵심인 central metabolic pathway of *E. Coli*을 채택했다. central metabolic pathway에 관련된 21개의 조절단백질 중에서, 혐기성, 호기성, glucose, lactose, succinate, acetate, ethanol, tryptophan metabolism에 관련된 13개의 조절단백질들을 *in silico* system에 포함하였다.

2. Combination of Hierarchical Regulatory Network and Dynamic Simulation Module

가상 미생물 시스템은 계층적 유전자 조절 네트워크 모듈과 동적 모사 모듈의 연계를 통하여 이루어진다. 유전자 조절 네트워크는 환경 변화에 따른 조절 단백질(regulatory protein), 오페론 및 유전자의 활성화와 비활성, 단백질의 합성과 분해를 위한 지연시간(delay time)과 같은 정보를 제공하며, 이들 모두는 동적 모사 모듈의 제약조건으로 작용하게 된다. 유전자 및 오페론의 비활성은 대사흐름에 제약조건으로 작용하며, 지연시간은 오페론이 비활성에서 활성화로 변화할 때 발생하고, 세포 성장이 일어나지 않는 구간을 생성시킨다.

동적 모듈 모사에서는 유전자 조절 네트워크에서 생성된 제약조건들을 고려함으로써, 성장율, 세포량, 기질의 농도변화, 그리고 흐름 평균 분석 등을 실시하게 되고, 세포의 외부환경을 나타내는 기질 농도의 time profiles은 환경변화에 민감한 계층적 유전자 조절 네트워크에 환경 변화 시점을 알려준다. 변화된 환경은 다시 계층적 유전자 조절 네트워크에서 새로운 제약조건들을 생성하게 한다. 이렇게 계층적 유전자 조절 네트워크

모듈에서 생성된 정보와 동적 모사 모듈에서 생성된 정보가 서로 연계됨으로써, 가상 미생물 시스템이 실행된다.

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.

3. 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 제시한 모델링방법과 *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