

Systems biology and engineering (SBE)

임영일

CAPEC, Department of Chemical Engineering,
Technical University of Denmark, DK-2800 Lyngby, Denmark
Phone: +45 4525 2804, Fax: +45 4593 2906, Email: lim@kt.dtu.dk

2. SBE 에 관한 연구동향과 현황

Systems Biology and Engineering (SBE) 에 관련된 여러 학문들 (genomics, proteomics, bioinformatics, systems biology, separation science, systems engineering, applied mathematics, computational science and software engineering) 에 관한 전문적 연구 및 통합 움직임은 세계 여러 연구기관에서 엿볼 수 있다. 본 장에서는 이들 분야들과 SBE 와의 연관관계를 설명하고, 알려진 연구주제나 연구자들에 관하여 알아본다.

2.1 Bioinformatics

유전자와 생물분자의 구조와 기능의 복잡성/다양성으로 인하여 체계적 정보구축이 요구된다. Protein databank (<http://www.pdb.org/>, Berman et al., 2000) 에서는 단백질의 서열, 구조, 기능에 관한 정보를 게시하고 검색할 수 있다. Danish centre for human genome research (<http://proteomics.cancer.dk/>, Gromov et al., 2002) 에서는 인간 유전자에서 발현되는 단백질에 관한 2D-PAGE (2-dimensional polyacrylamide gel electrophoresis) 정보를 인터넷 사이트를 통하여 갱신하고 있다.

1992 년에 설립된 The Institute for Genomic Research (TIGR, <http://www.tigr.org/>) 에서는 수십여 종의 알려진 종들 (e.g., viruses, prokaryotes, eubacteria, eukaryotes) 의 유전자 서열 (genome sequence) 을 찾아볼 수 있다. TIGR 는 최근 21 개의 미생물의 전체 유전자를 해독하였다.

Princeton 대학의 전산화공 연구실 (<http://titan.princeton.edu/>)을 이끄는 C. A. Floudas (floudas@titan.princeton.edu) 교수는 자체 개발한 global optimization 알고리즘인 α BB (α -Branch and Bound) 를 이용하여 분자 자유에너지 함수의 최적화를 통해 단백질 구조/형상 예측에 관하여 연구한다 (Klepeis & Floudas, 2003). 이 연구는 초기조건으로 어느 단백질의 아미노산 서열이 주어졌을 때, 이들의 2 차구조 (e.g., α -helices and strands, β -sheet, and disulfide bridge formation) 를 예측하고, α BB 최적화 알고리즘을 이용하여 단백질의 3 차 구조 (i.e., protein folding) 를 전산모사한다.

단백질의 3 차 구조 (tertiary structure)가 단백질의 기능과 흡착에 큰 영향을 미치기 때문에, 단백질 분리 메커니즘 (i.e., separation and adsorption mechanism) 을 규명하는데 있어 단백질 구조 예측은 필수적이라 할 수 있다.

2.2 Functional genomics

Systems Biology and Engineering (SBE) 분야를 뒷받침하기 위해 빼놓을 수 없는 연구방향은 기능 유전체학 (functional genomics) 또는 단백질체학 (proteomics) 이다. 유전정보 해독 (genome sequence) 을 주도해온 genome era 이후의 주된 연구방향은 genome sequencing 보다도 훨씬 더 복잡한 functional genomics 이다 (Vukmirovic & Tilghman, 2000).

생물학은 complexity and data-rich science 로 특성지어 진다. 유전정보로부터 발현되는 단백질/효소 생성과정, 이들의 대사/조절/신호 경로 등은 DNA microarray, protein chips 등으로 직간접적으로 측정이 되며, 방대한 자료가 수집된다 (Lockhart & Winzeler, 2000). 이들 자료분석을 통하여 정보-지식화 (i.e., from information to knowledge) 하는 작업은 systems biology 분야에서의 여러 반응경로, 반응상수를 알아내는데 큰 기여를 할 것이다.

2.3 Systems biology

살아있는 세포의 동적거동을 수학적 모델링을 통하여 체계적으로 이해하려고 하는 systems biology 는 최근 떠오르는 생명공학 기술중의 하나이다. 세포의 컴퓨터 모사에 대한 중요성은 1980년대부터 부각되었고, 1990년대 후반에서야 세포 전체에 대한 모델링과 모사가 이루어졌다 (Tomita, 2001). Systems biology 의 가장 중요한 연구 목적중의 하나는 컴퓨터 모사를 통하여 새로운 유전자를 설계하고, 이 새로운 유전자를 보유한 세포의 대사, 유전자 조절, 신호전달 경로에 대한 수학적 모델링을 통해 세포의 동적거동을 분석하는 것이다. 이 분야에서 일본은 앞선 연구를 진행하고 있는 것으로 보인다: the Kitano symbiotic systems project (Japan Science and Technology Corporation, <http://www.symbio.jst.go.jp/symbio/>, Kinato, 2002), the E-Cell project (<http://www.e-cell.org>, Tomita et al., 1999), the SBML project (<http://www.sbml.org/>, Hucka et al., 2003), Institute for Advanced Bioscience (Keio university, <http://bioinfo.sfc.keio.ac.jp/IAB/>, Japan).

한국은 실험을 통한 모델개발에 충분한 경쟁력이 있어 보인다. 한국과학기술원의 이상엽교수가 이끄는 대사공학연구실 (Metabolic & Biomolecular Engineering National Research Laboratory, <http://mbel.kaist.ac.kr/>)과 박선원교수가 이끄는 Process systems lab. (<http://love.kaist.ac.kr>)에서는 대사과정에서 각 경로마다 생물분자 유량을 계산해 내는 MetaFluxNet v1.5.2 (Lee et al., 2003) 를 개발했고, 계속적으로 갱신하고 있다. 가상세포 모사를 위해서 E-Cell 의 보완프로그램으로 MetaFluxNet 는 유용해 보인다.

이 외에 국제적으로 알려진 연구진 및 연구과제는 다음과 같다: 영국 맨체스터 대학의 Control Systems Centre (University of Manchester, <http://www2.umist.ac.uk/csc/>) 소속의 Wolkenhauer (see Wolkenhauer, 2002), 독일 Max-Planck Institute Magdeburg (<http://www.mpi-magdeburg.mpg.de/>) 의 Systems Biology group 의 E. D. Gilles (see Kremling et al., 2000, 2001a, 2001b), 미국의 J. C. Schaff 가 주도하는 the V-Cell project (National Resource for Cell Analysis and Modeling, University of Connecticut Health center, <http://www.nrcam.uhc.edu>, Loew & Schaff, 2001), Microbial cell project (<http://www.microbialcellproject.org/>).

특히 1996년 부터 시작된 E-Cell project (Tomita et al., 1999) 는 세포의 동적거동을 모델링/모사할 수 있는 통합환경으로 480여개의 유전자를 보유한 *M. genitalium* 원핵세포의 127개 기본적인 유전자만을 포함한 가상 세포내 대사과정을 모사 (*in silico* experiments) 하고, 생성물들의 농도를 예측하는 작업을 성공적으로 끝냈다. 하지만, *M. genitalium* 세포 *in silico* 모델에는 유전자 조절경로나 신호전달 경로등에 관련된 반응식은 포함하고 있지 않으며, 지속적인 개발상태에 있다. E-Cell 을 이용한 적혈구세포 가상모사도 실험 결과와 비교할 만한 좋은 성과를 보이고 있다 (Tomita, 2001).

세포내 대사/조절/신호 경로 모델링과 분리공정에서 흡착 메카니즘의 복잡성으로 인하여 범용의 재활용 가능한 모델링을 위한 프로그래밍 언어 개발은 필수적이다. 이 분야는 Software engineering 과 밀접한 관계를 갖고 있으며, SBML (System biology markup language, <http://www.sbml.org>, Hucka et al., 2003) 은 이러한 목적으로 개발 중에 있으며, XML (extensible markup language) 과 UML (unified modeling language) 에 기반을 두고 있다.

System biology 에 관한 개괄적인 설명은 Report #3 에서 다룰 것이다.

조직/기관 모사 (tissue/organ simulation)

단일세포의 동적모사가 아직 초보적인 단계이고, 다양한 기능을 갖는 수많은 세포들의 상호관계에 대한 모사는 더더욱 복잡하고 어렵다. 즉, 유전자 정보로부터의 세포/조직/기관의 수학적 모델링을 통하여 생물체의 거동을 해석한다는 것은 아직 불가능하다. 하지만, 생물체 특히 인체내의 특정한 역할을 하는 부위 (e.g., 조직, 기관) 의 모사는 부분적으로 이루어지고 있다 (Bassingthwaight, 2002). 이러한 움직임은 The Physiome Project (<http://www.physiome.org/>) 에서 좀더 구체화되고 있다.

와싱턴대 Bassingthwaight 교수 (Dept. of bioengineering, Univ. of Washington, USA) 는 The Physiome Project 의 한 일환으로 인체 심장에 관한 물리적 현상과 동적 기능에 관한 정량적 모사에 초점을 둔 The Cardione Project (Bassingthwaight et al., 1998) 를 수행 중에 있다.

The Center for Modeling Integrated Metabolic Systems (MIMS, <http://www.csuohio.edu/mims/>) 소속의 Cabrera 교수 (Dept. Biomedical engineering, Case Western Reserve Univ., USA) 는 동맥 혈액상의 glucose/산소 등의 농도변화에 따른 조직의 대사반응에 미치는 영향을 모델링하여 정맥의 혈중 glucose/산소 등의 농도를 예측하는 연구를 진행 중에 있다. 이 모델은 혈관내 여러 화학 물질들의 대류, 확산, 반응에 관한 편미방/미방/대수식 등으로 구성되어 있다 (Cabrera, et al., 1998). 이 연구는 약물전달과정이나, 외부환경변화에 따른 인체의 반응을 모사/예측하는 것을 목적으로 하고 있다.

이 분야에 관한 개괄적인 설명은 Report #4 에서 다룰 것이다.

2.4 Bioreactor engineering

생물반응기내의 세포 배양액 속에는 수많은 미생물들이 존재한다. 이들 각각의 동적 거동은 생물반응기의 설계와 제어에 큰 영향을 준다. 앞서 설명한 Systems biology 에서 구현된 세포모사 결과를 생물반응기 모델식으로 연결함으로써 우리는 유전정보로부터 생물반응기 모사까지 체계적인 분석이 가능하다.

세포배양에 있어서 미생물수와 생성물의 자가 진동 (autonomous oscillation) 현상은 잘 알려진 사실이다 (Zhu et al., 2000). 이러한 미생물수의 동적거동은 cell population balance equation 으로 표현할 수 있다 (Jones and Kompala, 1999). 이 분야는 크게 3 가지 방법으로 접근할 수가 있다:

- i) structured/unsegregated modeling (Jones and Kompala, 1999): 세포의 외부환경변화 (e.g., feeding flowrate, substrate concentration change) 에 따른 세포내 대사반응을 화학적/생물학적 모델식 (i.e., chemically/biologically structured) 으로 표현하고 단일세포의 동적거동 (i.e., 수많은 세포들로 세분화되지 않은, unsegregated) 을 연구. 이 방법은 앞서 설명한 세포모사 (2.3 systems biology) 와 거의 같은 접근 방법이다.
- ii) segregated/unstructured modeling (Zhu et al., 2000): 다양한 크기/나이 (mass and age) 를 갖고 있는 미생물들의 입자분포식 (cell population balance equation, PBE) 을 바탕으로 세포성장속도식을 표현하기 위해 단순한 먹이 (substrate) 의 질량수지식만을 이용하는 것 (i.e., chemically/biologically unstructured).
- iii) segregated/structured modeling (Mhaskar et al., 2002): 미생물들의 입자분포식 (cell population balance equation, PBE) 과 세포외부환경변화에 관련된 세포내 대사반응 (metabolic reactions) 을 동시에 고려하는 연구.

첫번째 방법 (structured/unsegregated model) 은 복잡한 세포내 대사반응을 모사하여 세포의 동적거동을 해석하기 때문에 가장 복잡하고 현학적인 연구일 것이다. 하지만, 다양한 크기/나이를 갖는 세포들이 생물반응기내에 존재함으로써 cell population balance equation 을 이용하는 두번째와 세번째 방법이 더 현실적인 접근이다. 특히 세포들간의 상호관계와 세포외부환경에 관련된 세포내 대사반응을 함께 고려하는 세번째 방법이 가장 합리적인 접근으로 보인다.

Segregated/structured model 을 통하여 생물반응기내 세포들의 크기/나이 분포를 예측할 수 있고, 또한 목적단백질의 생성속도를 예측할 수 있으며, 생물반응기 제어를 위하여 이 모델식을 이용한다.

이 분야는 Report #5 에서 요약된다.

2.5 Separation/purification science and engineering

생물공정은 주로 미생물의 발효반응기, 세포 분쇄과정, 분리/정제 공정으로 구성된다. 비교적 작은 분자량의 단순한 물질분리에 초점이 맞추어져 있는 화학공정 분리에서는 가압/가열에 의한 증류탑이 많이 사용되지만, 폴리펩타이드 ($M_w < 10$ kDa) 나 단백질 ($M_w > 10$ kDa) 과 같은 열에 약한 거대분자를 다루는 생물공정의 분리에는 상온/상압의 전기영동법 (1D or 2D electrophoresis) 이나 크로마토그래피법 (chromatography), 분리막법 (membrane in ultra-filtrations for size exclusion) 등과 같은 정밀하고 섬세한 장치가 사용된다. 분리공정의 주요 목적은 단백질이나 펩타이드 생성반응 후 나오는 혼합용액에서 물/숙주단백질/단백질분해효소 등을 제거하는 것이다.

특히 크로마토그래픽 공정은 대용량 단백질의 분리와 정제에서 많이 사용하며, 고순도의 최종 제품을 위해서는 생물반응에 관련된 여러 불순물 (비정상 단백질, 덩어리 단백질, 불완전 단백질, 비활성 단백질 등) 의 정제과정이 필요하기 때문에, 여러 단계의 크로마토그래픽 분리공정 (e.g., ion exchange, reversed phase, affinity, hydrophobic interaction and size exclusion) 을 포함할 수 있다. 하지만, 크로마토그래픽 공정개발에 있어서 생물분자의 복잡한 흡착메카니즘으로 인하여 아직까지 실험적 공정개발에 치중하고 있다 (Iyer et al., 1999).

Rensselaer Polytechnic Institute (<http://www.rpi.edu/dept/chem-eng>, USA) 화공과 Steven M. Cramer 는 산업용 크로마토그래픽 단백질 분리공정 (preparative protein chromatography) 에 관한 모델링/모사/최적화와 실험을 통한 검증에 관한 연구를 하고 있다 (Gallant & Cramer, 1996).

크로마토그래픽 분리공정의 수학적 모델링은 일반적으로 고정상과 유동상사이의 질량보존을 위한 편미방, 흡착반응에 대한 미분식, 그리고 고체-액체상 평형농도를 예측하기 위한 대수식으로 기술된다 (Lim et al., 2003). 여기에서, 고체-액체상 단백질 평형농도 예측은 분리성능을 좌우하는 가장 중요한 모델식이기 때문에, 많은 연구자들의 관심 대상이다.

단백질의 구조 (structure) 와 형상 (conformation) 의 변화에 따른 흡착 메커니즘에 관한 연구 (K. K. Unger, Mainz university, <http://www.uni-mainz.de>, Germany) 와 온도, pH, 염농도 등의 변화에 따른 평형농도 예측 모델 개발 (G. Carta, Virginia University, <http://www.che.virginia.edu/>, USA) 등은 현재 활발히 진행 중에 있다. 연속적 크로마토그래픽 공정으로 생물분자의 대량생산에 적합한 SMB (simulated moving bed) protein chromatography 는 최근 각광을 받고 있는 분리 기술이다 (Juza et al., 2000; Xie, Moon & Wang, 2003). 미국 Purdue university 의 생물분자 분리실험실 (<http://atom.ecn.purdue.edu/~biosep/>, USA) 을 이끄는 N.-H. L. Wang 교수는 SMB 공정에 많은 연구결과를 발표하고 있다. M. Morbidelli (<http://www.morbidelli-group.ethz.ch/>, ETH Zurich, Switzerland) 는 SMB 공정 설계와 최적화등에 관하여 이론개발과 실험결과에 관하여 연구를 하고 있다.

이 분야는 Report #6-#7 에서 좀더 자세하게 설명할 것이다.

2.7 결론

SBE 의 목적은 유전정보로부터 발견되는 단백질과 같은 생물분자를 예측하고, 생물공정설계 단계에서 세포모사 결과를 응용하는 것이며, 생성물의 분리와 정제에 이르는 분리공정을 모델링/모사하는 것이다. 또한 인위적 공정이라기 보다는 생체내 생물학적 공정이라고 볼 수 있는 조직/기관들에 관한 수학적 모델링을 통하여, 물리적/생물학적 현상을 예측하는 것이다.

생물공학에 대한 관심이 높아 가고 있고, 이에 관련된 과학과 기술이 팽창하고 있지만, 생물분자의 다양성과 복잡성으로 인하여 좀더 근본적이고 기초적인 연구가 계속 진행되어야 한다. 예를 들어, 비교적 많은 유전자수를 보유한 세포에 관한 모델링과 가상모사에 있어서 가장 절실한 문제중의 하나는 정량적 정보의 부족이다. 어떤 단백질이 무슨 분자와 서로 반응한다는 정성적 자료와 경로 등은 알려져 있지만, 필요한 효소의 농도, 반응속도 상수, 해리상수 등은 잘 알려져 있지 않은 경우가 많다 (Tomita, 2001). 따라서, 이러한 기초연구가 선행되어야 하는 컴퓨터를 이용한 생물세포 모델링과 공정 개발은 많은 부족한 점을 안게 될 것이고, 지속적으로 보완되어야 한다.

Systems Biology and Engineering (SBE) 분야에 있어서 부분적이고, 전문적인 연구는, 앞서 기술했듯이 많이 찾아볼 수 있지만, 전체적이고 통합된 환경을 구축하려고 하는 움직임은 아직 시기상조인 듯하다. 하지만, Systems Biology and Engineering 분야로의 빠른 진출은 생물공학기술 (biotechnology) 의 체계화 및 상업화에 큰 기여를 할 것이고, 고급기술을 선점할 수 있는 기회를 갖게 될 것이다.

References

- Bassingthwaite et al. (1998), Blood flows and metabolic components of the cardiome, Prog. Biophy. & Mol. Bio., 69, 445-461.
- Bassingthwaite, J. B. (2002), Proteins, cells, organs: hierarchical structuring in modeling biological systems, Preceding of the 2nd joint EMBS/BMES conferences, p2176-2177, Houston, USA.
- Berman, H. M. et al.(2000), [The Protein Data Bank](http://www.rcsb.org/), Nucleic Acids Research, 28, 235-242.
- Cabrera, M. E. et al. (1998), Modeling metabolic dynamics from cellular processes to organ and whole body responses, Prog. Biophy. & Mol. Bio., 69, 539-557.
- Gallant, S. R., S. Vunnum, S., S. M. Cramer, (1996), Optimization of preparative ion-exchange chromatography of proteins: linear gradient separations, J. Chromatogr. A, 725(2), 295-314.
- Gromov, P. S., M. Østergaard, I. Gromova and J. E. Celis (2002), Human proteomic databases: a powerful resource for functional genomics in health and disease, Prog. Biophys. & Mol. Biol., 80, 3-22.
- Hucka, M. et al. (2003), The systems biology markup language (SBML); a medium for representation and exchange of biochemical network models, Bioinformatics, 19(4), 524-531.
- Iyer H, Tapper S, Lester P, Wolk B, Van Reis R. 1999. Use of the steric mass action model in ion-exchange chromatographic process development. J Chromatogr A 832: 1-9.
- Jones, K. D. and D. S. Kompala (1999), Cybernetic model of the growth dynamics of *S. cerevisiae* in batch and continuous cultures, J. Biotechnol., 71, 105-131.
- Juza et al. (2000), Simulated moving bed chromatography and its application to chirotechnology, Trends in Biotech., 18, p108.
- Kitano, H. (2002), Systems biology: a brief overview, Science, 295, 1662-1664.

-
- Klepeis, J. L. and C. A. Floudas (2003), *Ab initio* Tertiary Structure Prediction of Proteins, *J. Glob. Opti.*, 25(1), 113-140.
- Kremling A. and E. D. Gilles (2001a), The organization of metabolic reaction networks: II. Signal processing in hierarchical structured functional units, *Metab. Eng.*, 3, 138-150.
- Kremling A., K. Bettenbrock, B. Laube, K. Jahreis, J. W. Lengeler and E. D. Gilles (2001b), The organization of metabolic reaction networks: III. Application for diauxic growth on glucose and lactose, *Metab. Eng.*, 3, 362-379.
- Kremling A., K. Jahreis, J. W. Lengeler and E. D. Gilles (2000), The organization of metabolic reaction networks: I. a signal-oriented approach to cellular models, *Metab. Eng.*, 2, 190-200.
- Lee D.Y, Yun H.S, Lee S.Y. and Park S.W. (2003) MetaFluxNet: the management of metabolic reaction information and quantitative metabolic flux analysis. *Bioinformatics* (In press).
- Lim, Y. I., S. C. Chang and S. B. Jørgensen (2003), A novel PDAE solver: iterative CE/SE method, *Comp. Chem. Eng.* (in press).
- Lockhart D. J. & E. A. Winzeler (2000), Functional genomics: genomics, gene expression and DAN arrays, *Nature (insight review)*, 405, 827-836.
- Loew, L. M. and J. C. Schaff (2001), The Virtual Cell: a software environment for computational cell biology, *Trends in Biotechnology*, 19(10), 401-406.
- Mhaskar, P., M. A. Hjortso and M. A. Henson (2002), Cell population modeling and parameter estimation for continuous cultures of *S. cerevisiae*, *Biotechnol. Prog.*, 18, 1010-1026.
- Tomita, M. (2001), Whole-cell simulation: a grand challenge of the 21st century, *TRENDS in Biotechnology*, 19(6), 205-210.
- Tomita, M. et al. (1999), E-CELL: software environment for whole-cell simulation, *Bioinformatics*, 15, 72-84.
- Vukmirovic O. G. & S. M. Tilghman (2000), Functional genomics: Exploring genome space, *Nature (insight review)*, 405, 820-822.
- Wolkenhauer O. (2002), Mathematical modeling in the post-genome era: understanding genome expression and regulation-a system theoretic approach, *BioSystems*, 65, 1-18.
- Wolkenhauer, O., H. Kitano and K.-H. Cho (2003), Systems biology: looking at opportunities and challenges in applying systems theory to molecular and cell biology, *IEE Control systems Magazine*, Aug. 2003, 38-48.
- Xie, Y., S.-Y. Mun, and N.-H. L. Wang (2003), Startup and Shutdown Strategies of Simulated Moving Bed for Insulin Purification, *Ind. Eng. Chem. Res.* 42, 1414-1425.
- Zhu, G.-Y., A. Zamamiri, M. A. Henson and M. A. Hjortso (2000), Model predictive control of continuous yeast bioreactors using cell population balance models, *Chem. Eng. Sci.*, 55, 6155-6167.