

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

개요 :

인간게놈의 서열분석 (genomics), 기능유전자의 해독과 단백질체학 (proteomics) 에서의 눈부신 진보는 새로운 생물분자의 개발과 생산에 대한 체계적 접근을 가능하도록 만들고 있다. 이러한 가능성은 생물정보학 (bioinformatics), 세포 대사/조절/신호 경로의 수학적 모델링에 기초한 체계적 생물학 (**systems biology**) 그리고 분리/정제 기술 (separation/purification technologies) 을 생물반응기와 분리공정을 포함하는 상류/하류 공정 (upstream/downstream processes) 의 설계/모사/최적화/제어에 관한 체계적 공학 (**systems engineering**) 에서의 여러 기법들과 접목시킴으로서 이룩할 수 있다.

Systems biology 는 생물세포내 대사, 유전자 조절과정 그리고 신호전달 체계를 수학적으로 모델링하여 세포내 또는 세포간 상호관계에 대한 동적거동을 체계적으로 해석하려는 학문이다. 이를 위하여, 세포내 모든 물질들의 구조, 기능에 관한 정보 (bioinformatics) 가 필수적이며, 때로는 컴퓨터를 이용한 분자모델링 (molecular simulation) 을 통하여 필요한 정보를 예측할 수 있어야 한다.

Systems engineering 은 생물분자의 생산에 있어서 공학적/기술적 문제를 체계적으로 다루는 학문으로 주로 공정개발, 설계, 최적화를 수행할 수 있는 방법론/도구 (methodology and tool) 의 개발과 이것들의 통합된 환경을 구축하는 것이다. 여기에서, 반응기 (upstream processes), 분리/정제 공정 (downstream processes) 이 주된 관심이다. **Systems biology** 가 미시적세계에 초점이 맞춰있다면, **Systems engineering** 는 거시적세계와 현실적/경제적 문제에 중심을 둔다. **Systems biology** 와 **Systems engineering** 의 공통적 특징은 체계적/수학적 모델링을 이용한 현상의 체계화이며, 이를 위해, 응용수학 (applied mathematics), 계산과학 (computational science), 그리고 시각화와 사용 편리성을 위한 software engineering 을 도구로서 이용하는 것이다.

Systems biology 와 **Systems engineering** 을 통합하여 의약적/상업적 가치있는 생물분자의 발견/생산/분리/정제에 이르는 과정/공정을 체계적/수학적으로 모델링하고 컴퓨터를 이용한 통합적 환경을 구축하려는 시도를 우리는 이 보고서에서 **Systems Biology and Engineering (SBE)** 라고 명명한다 (Lim et al., 2003)¹.

본 보고서는 앞으로 총 8회에 걸쳐, SBE 란? (Report #1), SBE 에 관한 연구동향/현황 (Report #2), 살아있는 미생물 세포모사 (Report #3), 세포조직/기관모사 (Report #4), 생물공정제어를 위한 cell population 모사 (Report #5), 생물분자 분리/정제 공정의 분류와 특성 (Report #6), 크로마토그래피 공정의 모델링/모사/최적화 (Report #7) 그리고 최종적으로 SBE 의 응용을 통한 신물질 개발과 공정개발의 구체적 예 (Report #8) 에 관하여 다루게 될 것이다.

¹ Lim et al. (2003), Integrated tools for innovative bioproduct and bioprocess development, STVF Framework Program proposal, submitted to STVF (Danish scientific research council).

I. Systems biology and engineering (SBE)?

35000 여 인간유전자 해독작업이 주도해온 유전체학 (genomics) 시대에 이어, 유전자 생산품인 단백질, 엔자임, 펩타이드등에 관한 연구 (functional genomics or proteomics) 가 post-genome 시대에 주목을 받고 있다. 단백질과 펩타이드는 20 여개의 아미노산의 조합으로 구성되어 있고, 아미노산과 아미노산 사이에는 펩타이드 결합 (-CONH-) 으로 연결되어있다. 15000 여개의 알려진 단백질이 있지만, 300000 여종의 단백질이 있을 것으로 예측되며, 결국 10^6 개의 가능성있는 단백질 및 폴리 펩타이드 분자에 관한 연구라고 단백질체학을 표면적으로 정의할 수 있다.

이러한 proteomics 는 여러 연구분야와 관련되어 있다:

- 단백질의 서열, 구조 그리고 기능 (sequence, structure and functions) 에 관한 database 와 이들에 관한 분자모델링과 형상/구조 (conformation/structure) 예측에 관한 생물정보공학 (bioinformatics)
- 유전정보로부터의 단백질, 엔자임등의 생성 (transcription/translation), 대사 (metabolism), 조절 (regulation), 신호전달 (signal transmission) 과정 (pathway/network) 에 대한 연구 (i.e., transcriptome and metabolome)
- 단백질 분리/정제 과학 (separation/purification science/engineering) 및 단백질 분석과 기능에 관련된 연구 (e.g., protein-chips)
- 단백질등 생물분자의 주변 환경변화에 따른 형상변화 (conformation change) 에 기초한 바이오 센서 공학 (biosensor engineering)
- 재조합 미생물로부터 단백질 생산공정에서 분리 생산공정에 관한 생물공정공학 (bioprocess engineering)

본보고서는 복잡성 (complexity) 그리고 data-rich science 로 대표되는 단백질체학 (proteomics) 과 생물공정공학 (bioprocess engineering) 의 체계적인 접목에 관한 하나의 접근방식을 다룬다. 즉, 의학용 또는 상업적 가치가 있는 생물분자의 생산에 있어서, 분자수준의 이해와 미생물세포의 동적거동으로부터, 발효/분리/정제 공정수준에 이르는 과정을 모델링과 모사를 통하여 가상공장 (*in silico* plant) 을 건설하는 것이다. 그림 1 은 현실세계 (see Fig. 1a) 와 이에 상응하는 체계적 공정공학 (systems engineering) 그리고 체계적 생물학 (systems biology) 을 보여준다.

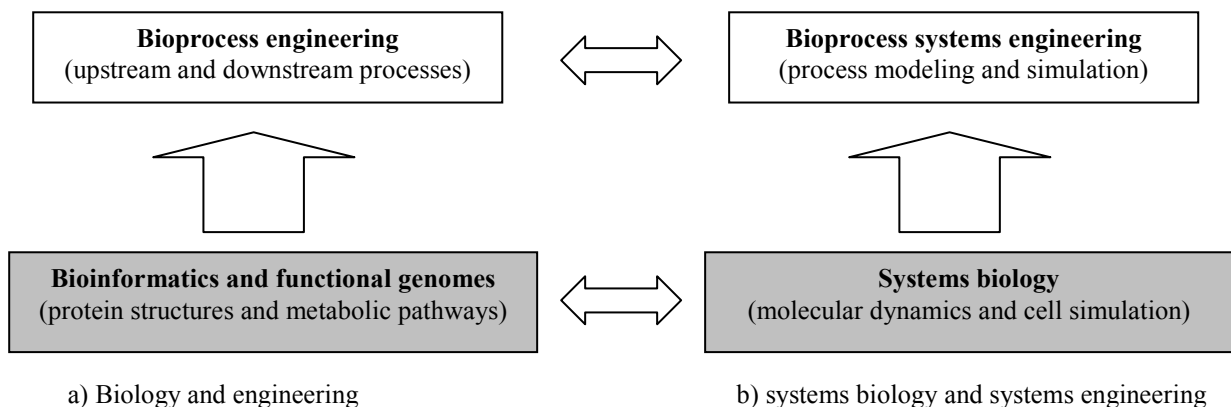


Fig. 1 Systematic approach for biomolecule production from molecular-level to process-level understanding.

systems biology and systems engineering (SBSE) 또는 좀더 줄여서 systems biology and engineering (SBE) 에 관한 본보고서의 첫번째 장에서는 이러한 체계적 접근이 왜 필요한지, 그리고 이들 학문들의 배경과 각각의 정의를 내린다. 참고로 본보고서는 최근 본인 외 13 인이 공동 제출한 연구계획서 (Lim et al., 2003) 에 기반을 두고 있음을 밝힌다.

1. SBE의 목적과 필요성

의약품 단백질을 생산할 수 있는 유전정보를 보유한 어떤 재조합 미생물이 있다고 가정해 보자. 이 미생물의 전체 유전정보가 알려져 있고, 이 유전자가 만들어내는 생성물들 (e.g., enzymes, proteins)의 구조와 기능이 파악되었다고 한다면, 이 목적단백질의 대량생산을 위하여 실험실크기 (lab-scale)에서 산업용크기 (industrial-scale)에 이르는 복잡하고 길며 고비용의 생물반응기와 분리/정제 공정의 개발과 설계가 필요하다. 즉 신약개발에 성공하였다고 하여도 상업화를 위한 긴 공정개발 단계가 기다리고 있다. 사실상, 이러한 길고 위험스러운 공정개발 단계로 인하여 의약적 가치가 있는 생물분자의 극히 일부분만이 상업화에 성공하고 있다 (Cserjan-Puschmann et al., 2002).

그림 2은 유전정보 (genetic information)로부터 제품 (product)까지, 그리고 분자수준에서의 이해로부터 공정개발까지를 도식화한 것이다. 유전자 발현/조절과정 (gene expression and regulation)의 수학적 모델링과 세포모사 (living cell simulation)를 토대로 새로운 생물 분자는 개발되고, 생물반응공정 (upstream processes)은 설계된다. 반응후에 나오는 혼합물은 하류공정 (downstream processes)에서 분리/정제된다. 분자수준에서의 이해 (e.g., biomolecule structure and properties)는 분리공정의 모델링, 모사, 개발에 있어서 필요한 분리 메카니즘 규명에 요구된다. 즉, 생물제품은 유전자정보와 생물분자의 특성치로부터 개발되어진다.

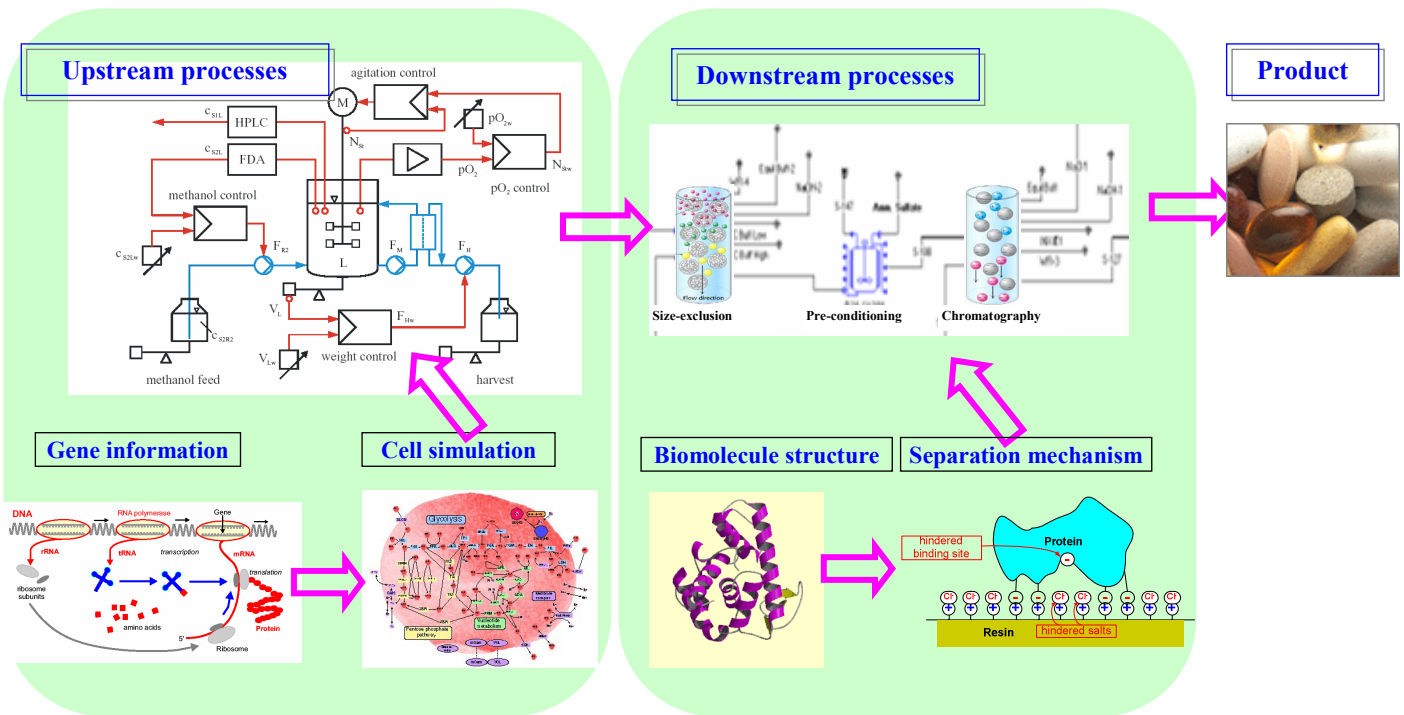


Fig. 2 From gene to product and from molecular-level to process-level (Lim et al., 2003).

만일 컴퓨터를 이용하여 이러한 다단계 공정개발을 미리 모사/설계/최적화 한다면, 요구되는 실험 횟수와 개발비를 크게 줄일 수 있고 (Kahn et al., 2001; Takahashi et al., 2002), 결국 생명공학 기업들은 시장경쟁력을 확보할 수 있게 될 것이다. 또한 생물학 분야 (genomics, proteomics, transcriptomics, metabolomics etc.)에서 빠른 속도로 축적되고 있는 정보를 체계적이고 효율적인 방법으로 저장하고, 이 정보를 바탕으로 생물분자 생산공정으로의 이용은 앞으로 다가오는 수십년 동안 생물/화학 공학자들이 해야 할 과제로 보인다.

2. 화학공학에서의 systems engineering

공정개발을 가속화시키고, 축적된 기술을 바탕으로 새로운 물질/공정 개발에 컴퓨터를 이용하는 것 (process systems engineering; PSE) 은 이미 화학공학분야에서 폭넓게 사용되고 있다 (Gani et al., 1997; Gani, 2003). Fig. 3 (Perkins, ESCAPE11, 2001) 에서 보듯이, PSE 라는 것은, 현실세계를 이해하고 관찰/실험하는 chemical engineering science 와 이를 바탕으로 한 모델링/공정모사/설계/합성 등을 통하여 엔지니어링 문제를 해결하는 학문으로 볼 수 있다.

역사적으로 볼 때, 현실세계 파악을 위한 실험위주의 chemical engineering science 는 1970년대 까지 주된 연구대상이 되었고, 이러한 축적된 정보와 지식은 1980년대로 들어오면서, 체계적으로 공정개발에 응용되었다. 즉 화학분자들의 물리적/화학적/열역학적 특성치 (e.g., DIPPR (Design Institute for Physical Properties) database for chemicals) 와 반응/분리/공정에서의 모델식과 해석기 (solver) 를 탑재한 다양한 공정모사기 (e.g., ASPEN Plus, <http://www.aspentec.com>) 는 이무렵에 개발되었다. 이를 이용하여 공정개발기간은 단축되었고 공정변수 최적화를 매우 체계적으로 수행할 수 있었다. 1990년대 그리고 현재까지 이러한 공정모사기술은 지속적으로 개발 중에 있고 더욱 정교해지고 있다.

PSE 에 대한 좀더 자세한 정의는 “**응용수학, 계산과학, 여러 공학기술의 통합적 기술 또는 방법**이며, 여기에서, **응용수학**은 주로 수치해석을 위한 대수학 (Linear Algebraic, NLA), ODE/DAE/PDE solution methods, Optimization 등과 수학적 모델링을 위한 것들을 말한다. **계산과학**은 이러한 수학적 표현의 컴퓨터를 이용한 구현이라고 볼 수 있으며, 프로그래밍 (Fortran, C/C++), 시각화, 각종 계산소프트웨어 (CFD codes, simulators, Libraries, MatLab) 이용기술을 말한다. **여러 공학기술**은 화학공학을 비롯하여, 화학, 생물공학, 기상학, 유체공학, 반도체 공학 등 다양한 응용분야” 를 말한다.”

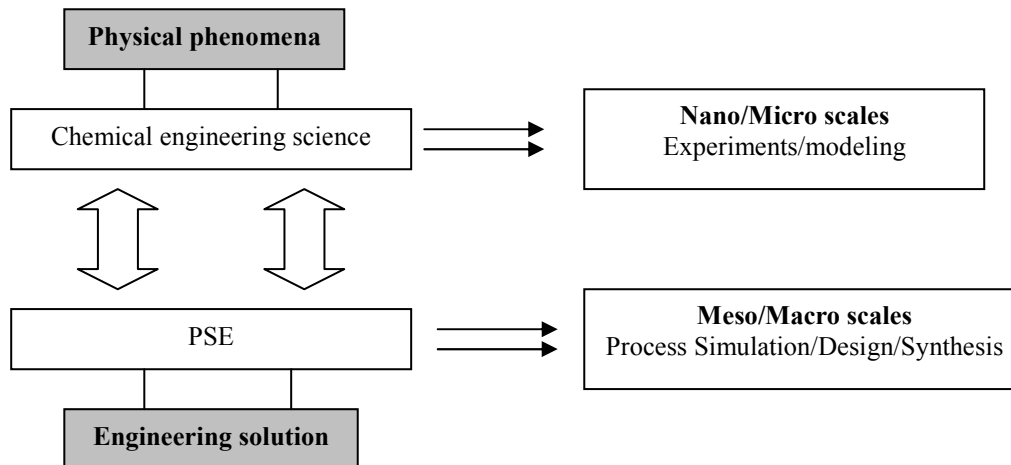


Fig. 3. The role of PSE in engineering solutions with a traditional aspect. (Perkins, 2001).

결국 process systems engineering 의 최종 결과물은 시스템 분석, 공정모사, 최적화 등 가상실험을 통한 기술개발과 database 통합에 있으며, 결국 공정의 자동화와 합리화등에 목적을 둔다. 이러한 학문적 연관관계를 바탕으로 재조명된 PSE 의 주요 연구분야는 Fig. 4 에서 보여준다. 이 그림에서는 제품의 개발단계에서 공장건설, 제품생산, 판매, 시장변화에 이르는 제품의 잉태에서 종말까지의 전반적인 단계는 독립적이지 않고, 서로 유기적으로 연결된 통합된 구조임을 보여주고 있다.

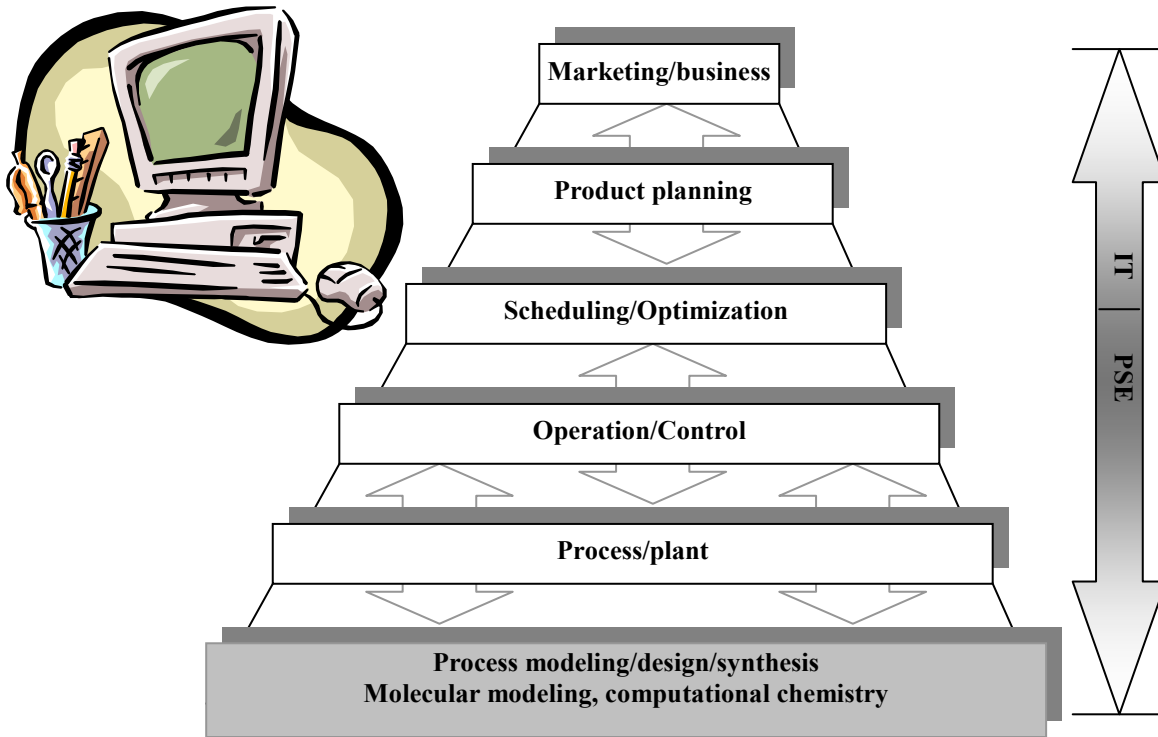


Fig. 4. Rigorous definition of PSE in the multimode/multiscale.

3. Systems biology and Systems engineering

복잡한 생물체의 거동과 확립되지 않은 여러 유전체의 기능으로 인하여, 생물공학과 체계적 공학의 접목은 화학공학에 대한 그림 3 과 4 처럼 잘 체계화되어있지 않지만, 비슷한 접근방식을 갖는다. 다른 한편으로, 생물공학에서 다루는 물질은 열에 약하고, 구조가 매우 복잡한 거대 생물분자이기 때문에 화학공학적 접근과 비슷한 면도 있지만, 새롭게 발전된 여러 학문 (multi-disciplines) 의 통합과 좀더 체계적이고 근본적인 접근이 요구된다.

예를 들어, 생물분자 생산을 위해 가장 많이 사용되는 원핵 미생물중의 하나인 4800 여개의 유전자를 보유한 *E. coli* 는, 100 여개의 대사반응, 2500 여개의 단백질 생성, 100 여개의 조절과정이 알려져 있고, 효소는 단백질 생산과 조절기능을 위해 활동한다. 여기에서, 복잡한 대사/조절/신호 경로 (metabolic/regulatory/signal pathway or network) 를 작은 기능단위로 나누고, 이 각단위를 모듈화하여 세포의 동적 거동을 체계적으로 분석할 수 있다.

Systems biology 는 이러한 생물세포내 대사, 유전자 조절과정 그리고 신호전달 체계를 수학적으로 모델링하여 세포내 또는 세포간 상호관계에 대한 동적거동을 체계적으로 해석하려는 학문이다 (Wolkenhauer, 2002; Tomita et al., 1999; Kitano, 2002). 이를 위하여, 세포내 모든 물질들의 구조, 기능에 관한 정보 (bioinformatics) 가 필수적이며, 때로는 컴퓨터를 이용한 분자모델링 (molecular simulation) 을 통하여 필요한 정보를 예측할 수 있어야 한다.

Systems engineering 은 생물분자의 생산에 있어서 공학적/기술적 문제를 체계적으로 다루는 학문으로 주로 공정개발, 설계, 최적화를 수행할 수 있는 컴퓨터내 통합된 환경을 구축하는 것이다. 여기에서, 반응기 (upstream processes), 분리/정제 공정 (downstream processes) 이 주된 관심이다. Systems biology 가 미시적세계에 초점이 맞춰있다면, Systems engineering 는 거시적세계와 현실적/경제적 문제에 중심을 둔다. Systems biology 와 Systems engineering 의 공통적 특징은 수학적 모델링을 이용한 현상의 체계화이며, 이를

위해, 응용수학 (applied mathematics), 계산과학 (computational science), 그리고 시각화와 사용 편리성을 위한 software engineering 을 도구로서 이용하는 것이다.

그림 5 는 지금까지 설명한 SBE 에 관련된 학문들과 연구분야의 관련성을 보여준다. 6 개 분야로 나뉘어진 이들 연구분야는 다음과 같다.

- **Bioinformatics:** property and structure databases of genomes and proteins, modeling for thermodynamic property estimation of biomolecules (e.g., by molecular simulation).
- **Systems biology:** modeling and simulation of cell and cell population dynamic behaviors through gene expression and regulation pathway identification.
- **Separation/purification science and technology:** modeling and identification of adsorption isotherms and mechanisms.
- **Systems engineering:** simulation and scheduling of upstream (fermentation reactor) and downstream (membrane, crystallizer and chromatography) processes.
- **Software engineering:** tools integration, modeling language development (e.g., SBML; Hucka et al., 2003) and a computer-aided environment development
- **Applications:** industrial applications for innovative product/process development

이들 연구분야의 통합은 결국 유전자정보로부터 생물분자 개발까지 그리고 분자수준에서의 이해로부터 생물공정개발까지의 체계적 접근을 가능하도록 할 것이다.

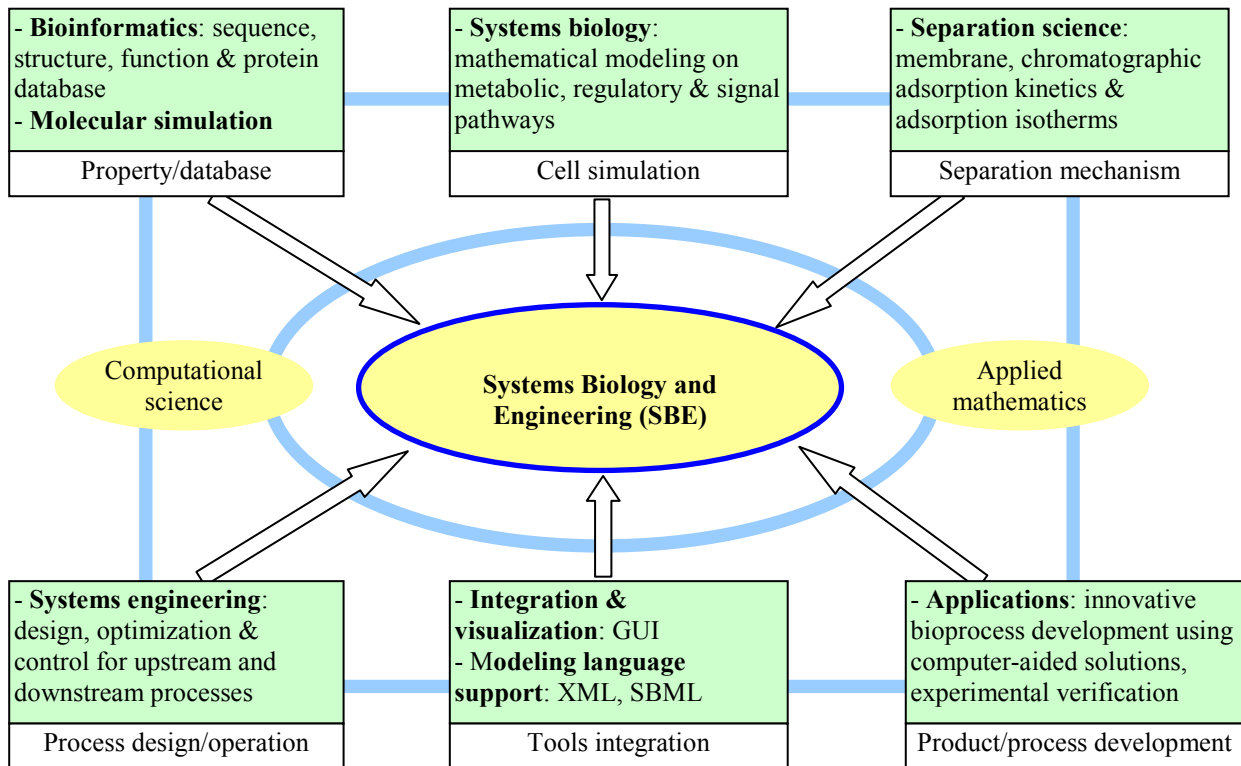


Fig. 5. Schematic approach for systems biology and engineering (SBE).

Systems biology 와 Systems engineering 을 통합하여 의약적/상업적 가치있는 생물분자의 발견/생산/분리/정제에 이르는 과정/공정을 수학적으로 모델링하고 컴퓨터를 이용한 통합적 환경을 구축하려는 시도를 Systems Biology and Engineering (SBE) 라고 본보고서에서는 명명한다.

본연구의 특징은 고가의 실험장치 비용이 들지 않는 대신에 고급인력을 투입하는 지식집약적 연구이고, 때로는 대용량 계산기기가 요구된다. 즉, information technology 와 biotechnology 의 접목이라고 간략하게 말할 수 있다.

References

- Cserjan-Puschmann, M., R. Grabherr, G. Striedner, F. Clementschitsch and K. Bayer (2002), Optimizing recombinant microbial fermentation processes: an integrated approach, *BioPharm*, July 2002, 26-34.
- Gani, R. (2003), Product design perspectives, *Comp. Chem. Eng.*, (in press).
- Gani, R.; Hyltoft, G.; Jaksland, C. and Jensen, A. K. (1997), An Integrated Computer Aided System for Integrated Design of Chemical Processes, *Comp. Chem. Eng.*, 21, 1135-1146.
- 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.
- Kahn, D., R. Plapp and A. Modi (2001), Modeling a multistep protein synthesis and purification process: a case study of a CAPE application in the pharmaceutical industry, ESCAPE11, Denmark.
- Kitano, H. (2002), Systems biology: a brief overview, *Science*, 295, 1662-1664.
- Lim, Y. I. et al. (2003), Integrated tools for innovative bioproduct and bioprocess development, STVF Framework Program proposal, submitted to STVF (Danish scientific research council), DTU, Denmark.
- Perkins, J. D. (2001), Summary of ESCAPE11, conference presentation on ESCAPE11, Denmark.
- Takahashi et al. (2002), Computational challenges in cell simulation: a software engineering approach, [IEEE Intelligent Systems](#), 17(5), 64-71.
- Tomita, M. et al. (1999), E-CELL: software environment for whole-cell simulation, *Bioinformatics*, 15, 72-84.
- Wolkenhauer, O. (2002), Mathematical modeling in the post-genome era: understanding genome expression and regulation-a system theoretic approach, *BioSystems*, 65, 1-18.