

미생물을 이용한 외부단백질 대량생산공정의 설계를 위한 Bioprocess Flowsheeting Software의 응용

이 종 대

생산기술연구원 화학공정연구팀

Application of a Bioprocess Flowsheeting Software to a Process Design for the Mass Production of Foreign Protein by Using Microorganism

Jongdae Lee

Chemical Process Engineering Team, Korea Institute of Industrial Technology, Seoul 153-020, Korea

ABSTRACT

An optimal process design of a foreign protein production system was carried out using a bioprocess flowsheeting software, BioPro Designer, with a capability of economic analysis. The flowsheeting program was applied to a production system of the tailspike protein of Salmonella phage P22, and helped save time and efforts in selecting an optimal process. A wild type tailspike and two types of mutant tailspikes, tsf G244→R and Su A334→V, were considered in this study to show that the folding characteristics of foreign protein produced inside host influenced the selection of the best production system. An optimal production system for mature tailspike was chosen under the criterion of capital investment per unit mass of mature protein recovered.

서 론

미생물을 이용하여 외부 단백질(foreign protein)을 대량 생산할 때 고려되어야 할 중요한 사항 중의 하나는 공정의 효율적인 설계 및 운용이다. 경제성에 입각하여 생물공장을 설계할 때 각 단위공정(unit operations)을 어떻게 배열, 연결하는가 하는 문제, 즉 공정의 최적 합성 문제는 중요한 연구대상이다. 공정합성 단계에서 각 단위 공정을 배열하고 연결할 때는 대상이 되는 미생물의 특성이 중요하게 된다. 왜냐하면 미생물내에 target 단백질이 주로 inclusion body 형태로 존재하느냐 natural protein 상태로 존재하느냐에 따라 외부 단백질의 회수공정이 달라지게 되기 때문이다.

경제성분석 기능을 가진 생물공정 flowsheeting 소프트웨어는 새로운 생물화학공정들을 개발하거나, 또는 기존의 생물화학 공정을 개량하거나 그리고 수많은 연구활동의 방향 등을 결정하는데에서 중요한 역할을 담당한다. 한 공장으로부터 최대의 이윤을 얻기 위해 자본투자를 결정할 때 좋은 숙주-벡터 시스템을 개발하는데 연구비를 투자할지 혹은 단백질을 회수하는 새로운 공정에 연구비를 투자할 지는 상당히 중요한 요소들 중의 하나이다. 물론 발효에 의하여 부가가치가 높은 단백질을 생산하기 위한 최적의 생물공정을 디자인하는 일은 학교 및 산업체 등으로부터의 모든 기초이론 및 공학이론 등 모든 분야의 이론들을 필요로 한다. 더욱이 생물 및 의약품 산업에 있어서 의학적으로 효과가 있는 물질을

발견하는데는 10년이 걸리고 하나의 치료약을 개발하려면 250백만달러가 소요된다. 생물공정에서는 속주 및 생산되는 단백질의 특성, 적절한 디자인 변수들의 선택, 그리고 경제성평가 등과 같은 요인들에 의해 process flow diagram이 달라지기 때문에 BioPro Designer같은 체계적인 디자인 도구가 필요하게 된다.

공정을 디자인할 때 컴퓨터를 이용하여 체계적으로 수행하게되면 디자인그룹간의 의사소통을 돕고, 기존의 flowsheets에 새로운 정보라든가 새로운 단위공정들을 추가하는 것이 용이하게 된다. 그리고 짧은시간 내에 여러 가지 변화요인들에 대해 반복적인 계산을 수월하게 하는 등 디자인 할 때의 문제를 해결하는데 상당한 도움을 받을 수 있고 이러한 문제들을 쉽게 해결하는 것은 공정디자인에 들어가는 비용을 절약해준다. 더구나 경제성분석 기능을 가진 디자인 소프트웨어는 수많은 선택의 생물공정중에서 가장 이윤이 많은 공정을 쉽게 결정하게 만든다.

BioPro Designer는 생물공정의 디자인 및 평가를 수행할 수 있는 소프트웨어이다. BioPro Designer는 물질 및 에너지 수지를 세워서 계산하고, 설비의 크기와 가격을 결정하고, 상세한 경제성분석을 리포트하고 공정스케줄링 문제를 다룬다. BioPro Designer는 user와 컴퓨터 사이의 커뮤니케이션을 증진시키고 배우는 시간을 최소화할 수 있도록 user에게 친근한 인터페이스를 가지고 있다. BioPro Designer는 공정디자인의 초기단계나 스케일-업 시에 사용되도록 만들어 졌다.

Flowsheeting 소프트웨어가 생물공정의 디자인과 최적화에 어떻게 효과적으로 사용될 수 있는지를 보여주기 위하여 시뮬레이션이 행해졌다. Salmonella P22의 tailspike단백질이 모델단백질로 선정이 되었는데 그 이유로는 tailspike 단백질에 대해선 실험자료와 많은 관련 논문들이 있어 정보를 구하기가 쉽기 때문이다(2, 3, 8, 1). 근래에 *E. coli*에서 tailspike 단백질을 발현시키는 시스템이 개발되었다(5). 세포 내에서 외부단백질의 과잉생산은 inclusion body를 생성하게 되는데 inclusion body는 활성이 없기 때문에 응집된 상태로부터 natural protein을 회수하는 공정이 필요하게된다.

Wild-type tailspike 단백질과 유전자 조작을 통하여 변환된 두개의 mutant 단백질(tsf G244-R과 Su A334-V)이 본 연구에서 사용되었는데 3개의 단백질들은 세포 내에서 생산될 때 natural protein과 inclusion body의 생성 비율에서 차이가 있다.

Wild-type tailspike 단백질(Fig. 1의 CASE I)의 회수공정에서는 natural protein과 inclusion body가 모두 회수되었고, tsf G244-R 변종단백질(CASE II)의 회수공정에서는 inclusion body만이 denaturation과 renaturation공정을 통하여 회수되었고, Su A334-V 단백질(CASE III)인 경우에는 오직 natural protein만이 filtering공정을 통해서 분리되었다. 세 개의 이종 단백질들은 각각 다른 종류의 생물공정 flowsheets를 형성하게되는데 본 연구에서는 BioPro Designer를 이용하여 경제성에 입각하여 공정합성과 분석이 행해졌다.

재료 및 방법

Flowsheeting 소프트웨어

BioPro Designer는 flowsheeting software로써 MIT에서 1990년에 개발되어 상용화되었다. 본래는 생물 공정의 분석및 공정의 자동 합성을 목적으로 설계되었는데 생물 공정의 분석 분야에 보다 더 치중해 확장 및 개발이 이루어지고 있다. 현재는 IBM 및 Mac 컴퓨터용 software를 개발하였으며, Object Oriented Programming(OOP) 기법을 가미하여 보다 더 효율적인 프로그램으로 update되었다. BioPro Designer는 상호작용이 가능한 그래프 인터페이스를 가지고 있고, recycle loop를 포함하는 통합된 flowsheets를 위한 물질수지, 설비크기와 구매비용의 추정, 상세한 stream과 경제분석 리포트, batch와 semicontinuous 스케줄링, 도표그리기, 최신의 on-line help 기능을 내재하고 있다(7). Flowsheeting 프로그램은 생물 혹은 제약 공정의 개발을 빠르게 한다. 그리고 새로운 생물공정제품이 다른 분야(생물학, 생화학, 효소학, 생물공학엔지니어, 공정엔지니어 등)의 전문가들의 협력을 필요로 하는 일이라면 그룹들 사이에서의 커뮤니케이션을 효과적으로 할 수 있게 도와준다. 그리고 flow-sheeting program은 공정의 수를 줄이고, 제품수율을 향상시키고, 보다 나은 스케줄링에 의해 기기사용률을 증가시키는 것에 의해 엔지니어가 통합된 공정설비의 투자 및 운영 자금을 최소화하게 해준다. BioPro Designer는 3가지의 중요한 reports를 만들어준다.

스트림 리포트(Stream Report)는 전후의 단위공정의 이름, 스트림의 성질, 각 components의 중량비와 유속 등을 보여준다. 상세한 경제성분석 리포트(Detailed Economic Analysis)는 각 단위공정에 대한 설명과 각단위공정의 구입가격 등의 정보를 제

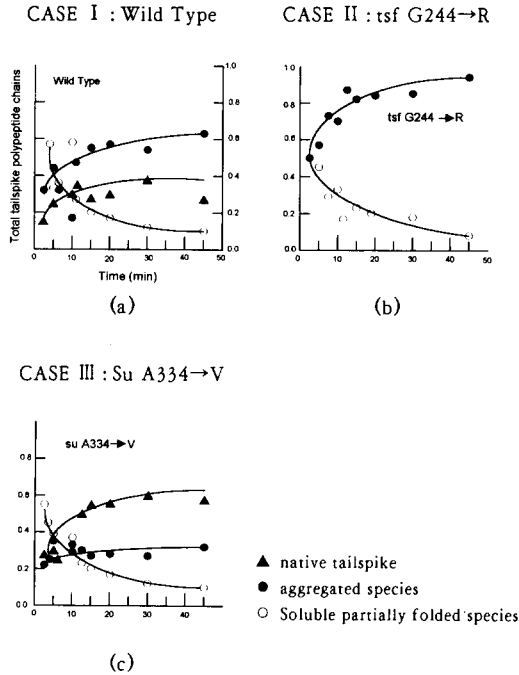


Fig. 1. Folding Characteristics of Tailspike Proteins (2).

공한다. 세분화된 가격 리포트(Itemized Cost Report)는 각 단위공정과 관련된 모든 가격정보를 제공한다.

Tailspike 단백질의 Folding 메커니즘

Tailspike는 Salmonella phage P22의 단백질로써 Salmonella의 표면에 부착한다. MIT 생물학과의 Prof. King의 group은 tailspike의 아미노산 배열이 folding과 refolding에 끼치는 영향을 연구해오고 있는데 Tailspike의 아미노산 sequence중 단 한 개를 치환하여도, 단백질의 활성이나 안정성에 영향이 없이, inclusion body 생성을 촉진 혹은 억제할 수 있다는 사실을 밝혀냈다.

Tailspike는 세 개의 똑같은 polypeptide의 trimer이다. 각 polypeptide는 666개의 아미노산으로 구성되어 있다. Tailspike는 Salmonella 표면의 O-antigen을 인식하고 그것에 달라붙는다. 성숙한 tailspike는 열안정성이 뛰어나서 녹는점이 88℃이다. Tailspike 단백질은 대부분의 proteases에 강하고 sodium dodecyl sulfate(SDS)의 용매에서 denature되지 않는다. Tailspike 단백질은 낮은 pH의 urea에

서 시작하면 세포 외부에서 성공적으로 refolding된다(8, 9). 세포 외부에서 denature된 단백질의 refolding, 그리고 세포내 folding과 maturation의 refolding에 관한 연구에서 inclusion body는 native 단백질로부터 생성되는 것이 아니고 부분적으로 fold된 중간체로부터 생성된다는 것을 알아냈다(3). 외부숙주에서 cloned genes들의 단백질 생산은 새로 합성된 polypeptide를 inclusion body 형태로 축적하는 결과를 가져온다.

Wild type 숙주인 경우에는 생성된 단백질의 33%는 native 상태로, 그리고 나머지 67%는 inclusion body 형태로 축적이 되었다(Fig. 1-a). Wild type 단백질에서 두 개의 아미노산을 치환(즉 331번째 V 대신에 R, 334번째 A 대신에 V를 치환) 하게되면 tsf G244→R이 된다(2). Tsf G244→R 인 경우(Fig. 1-b)에 생성된 단백질의 95% 이상이 inclusion body의 형태로 생성되었다. 또 하나의 변종인 Su A334→V 단백질은 native 단백질을 더 많이 생산하도록 하기 위하여 334번째에 있는 Alanine을 Valine으로 대체한 것이다. 단 한 개의 아미노산을 치환해도 활성이나 안정성에 영향을 미치지 않고 응집을 저지할 수 있다(3). Suppressor(억제제)를 포함하는 chain들이 wild type보다는 보다 효과적으로 mature됐고 거의 60%가 native한 상태로 존재하였다.

공정 디자인 목표

tailspike 단백질의 가격을 현시점에서 결정할 수 없기 때문에 회수되는 active한 단백질의 단위질량에 대한 자본 투자비를 최저 공정을 선택하는 기준으로 삼았다. Fermenter는 10,000리터의 broth를 저장할 수 있는 것으로 가정되었다. fermentation의 마지막에는 cell mass가 40g/L가 되는 것으로 가정했고, 생산되는 tailspike의 생산량은 8g/L가 되는 것으로 가정하였다. 세 개의 모델 단백질들(CASE I : wild type, CASE II : tsf G244→R, and CASE III : Su A334→V)의 dynamics와 folding 특성은 Fig. 1에 보여준다. Fig. 1에서 보여지는 데이터를 이용하여 가장 좋은 생산시스템을 경제성 분석을 통하여 결정한다.

공정설계

Upstream 공정은 동일하다고 보고 fermentation 후의 downstream 공정만을 고려하여 분석하였다. Inclusion body와 native 단백질의 생성비율에 따라

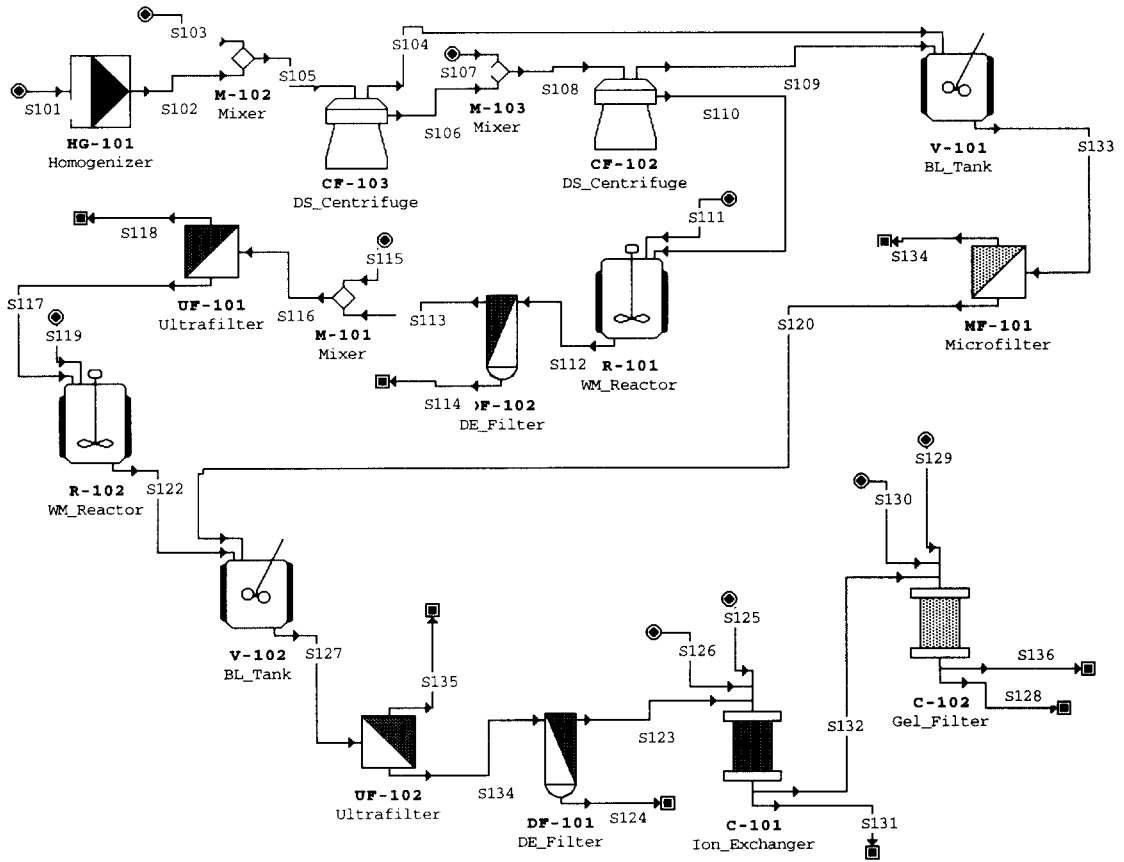


Fig. 2. Flowsheet for CASE I.

downstream 공정은 상당히 달라지게 된다. Fig. 1에 있는 데이터를 이용하여 세 개의 다른 공정 flowsheets(Fig. 2, 3과 4)가 완성되었고 분석되었다. CASE I의 경우에 natural 단백질과 inclusion body의 생성이 모두 중요하기 때문에 각각의 단백질이 각기 다른 공정을 통하여 회수된다. Inclusion body의 40% 만이 denaturation과 renaturation공정을 통하여 natural 단백질이 회수되는 것으로 가정한다. Fig. 2에서 보여지듯이 cells들은 고압의 homogenizer (HG-101)를 통하여 파쇄된다. Broth들은 homogenizer를 여러번 통과하게 된다. 밸브를 통과할 때 압력차는 크게된다. 높은 압력차로 여러번 통과시키는 것은 inclusion body와 natural 단백질을 완전하게 cells로부터 분리시켜 inclusion body의 분리를 용이하게 하기 위함이다. 원심분리기를 통과하기 전에 broth는 homogenizer 전의 부피

로 회석하는데 그 이유는 natural 단백질이 포함되어 있는 cells 구성 성분들로부터 inclusion body의 분리를 용이하게 하기 위함이다. 원심분리기(CF-102, CF-103)를 통과한 후부터 natural 단백질과 inclusion body는 각각 다른 경로를 통하여 회수된다. Natural 단백질은 microfilter(MF101)와 depth filter(DF-103)를 통하여 filtration 된다. 다른 한편으로 disk stack centrifuges (CF-102, CF103)에 의해서 분리되는 inclusion body는 denaturation과 renaturation공정을 통하여 분리된다. Inclusion body는 잘 교반되는 연속 반응기에서 urea에 의해서 denature된다. Urea는 chaotropic agent로써 inclusion body안에 있는 denature된 단백질을 용해시킨다. 잘 교반되는 반응기(WM-101)에서 용해시킨 후에 단백질들은 depth filter(DF-102)를 통하여 정제된다. 용해된 단백질들은

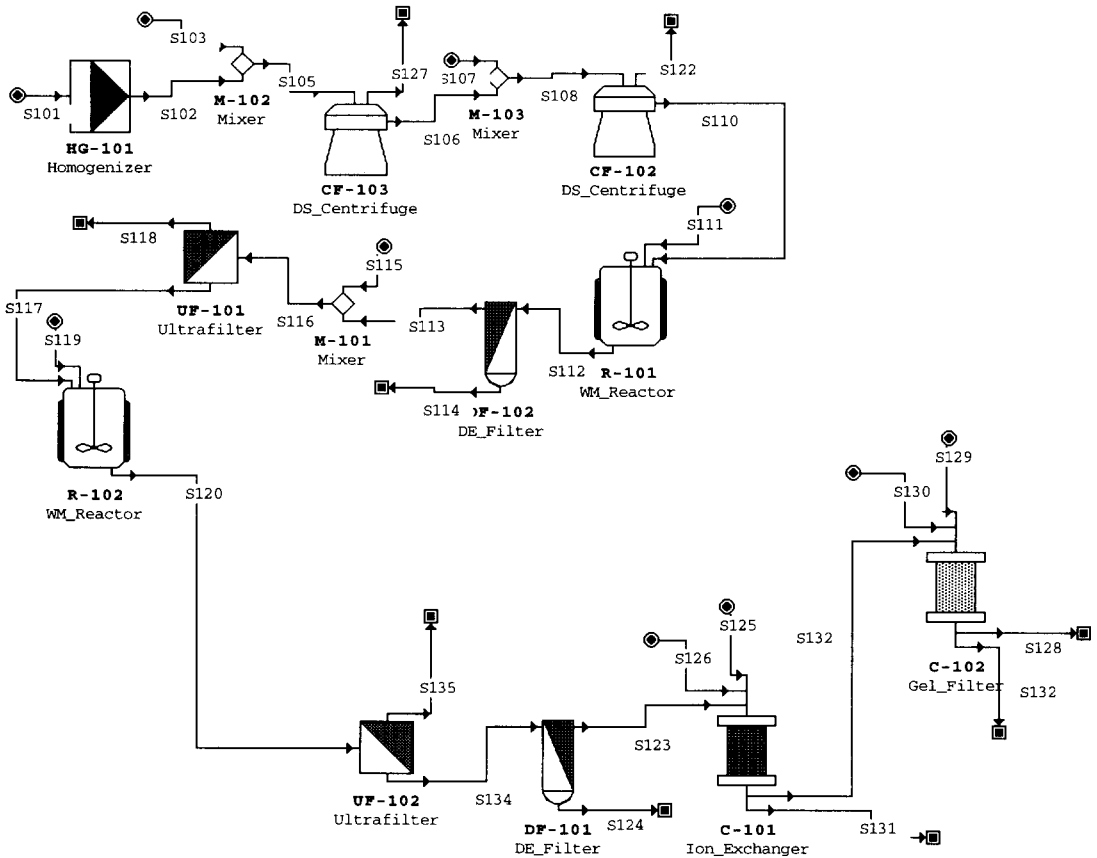


Fig. 3. Flowsheet for CASE II.

(M-101)와 ultrafilter(UF-101)로 구성되어 있는 diafiltration공정으로 넘어간다. Diafiltration중에 urea는 제거되고 순수로 대체된다. 이후 단백질용액은 잘 교반된 반응기에서 물로 희석된다(WM-102). 비활성 단백질을 만드는 분자간 상호작용을 최소화하기 위해 희석하는 것이 필요하다. 두 가지의 다른 경로를 통하여 회수되는 natural 단백질들은 blend 탱크(BL-Tank)에서 같이 혼합된다. 혼합된 단백질 용액은 ultrafilter를 통하여 15배 정도로 농축된다. 농축액에 있는 미세입자들은 dead-end filter(DF-101)를 통하여 제거된다. 이 dead-end filter는 이후의 공정에서 쓰이는 크로마토그래피 설비들을 보호하는 역할을 한다. 최종적으로 ion-exchange(C-101)와 gel-filter(C-102)를 통하여 정제과정이 끝나게 된다.

CASE II(Fig. 3)인 경우에는 모든 tailspike 단백질이 inclusion body 형태로 생성된다. 따라서 CASE II에서는 CASE I에 비해서 4개의 단위조각(V-101, MF-101, DF-103, 그리고 V-102)이 생략된다. 이 경우에 생성된 natural 단백질은 회수되지 않는다. CASE III(Fig. 4)의 경우에는 inclusion body와 natural 단백질이 동시에 생산되어도 inclusion body의 회수공정은 고려하지 않았다. 왜냐하면 가정에 의하여 denaturation과 renaturation 과정을 통하여 inclusion body의 40%만이 natural 단백질로 회수되므로 CASE III의 경우에는 12% (30%의 40%)가 natural 단백질로 회수되는데 이 경우에는 inclusion body 회수공정의 투자비를 고려하여 무시하는 것으로 하였다.

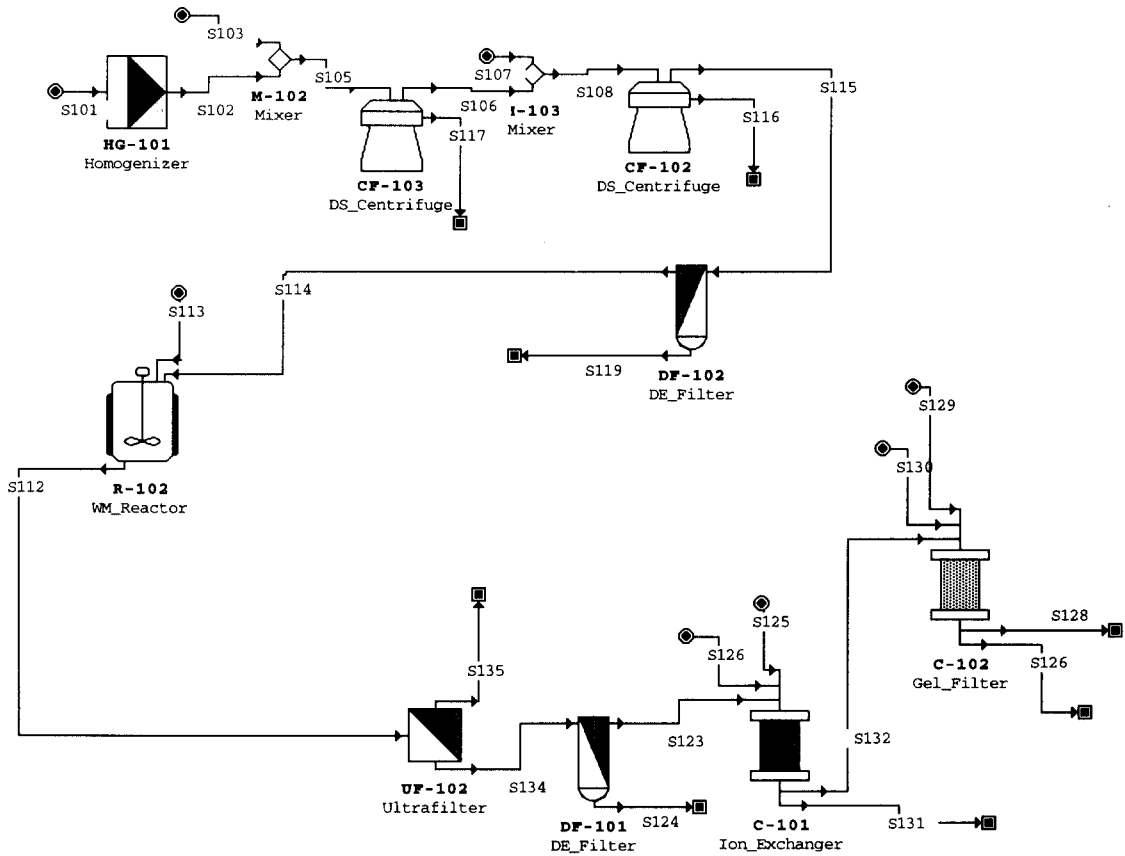


Fig. 4. Flowsheet for CASE III.

결과 및 고찰

경제성 분석에 의한 공정 최적화

위에서 언급한 것과 같이 가장 좋은 단백질 회수 공정의 합성과 선택은 inclusion body와 natural 단백질의 생성 비율에 따라 달려있다. BioPro Designer를 통하여 얻어진 정보를 바탕으로 분석하여 tailspike 단백질을 생산하는 가장 좋은 시스템을 선택했다. Table 1은 경제적 분석의 결과를 보여준다. 예상했던바와같이 CASE I을 위한 총 설비구입비는 \$1,636,999로 가장 높았으며, CASE III의 경우가 \$690,000으로 최소였다. BioPro Designer에 의해 annual operating cost가 리포트되었는데 이의 계산은 total equipment costs에 기준한 방법에 따른다(4). CASE I, II, III의 각각의 경우에 natural 단백질의 생산비율은 0.58, 0.38, 0.60이었

다. 회수된 tailspike 단백질 단위중량비 투자비는 CASE I, II와 III를 위하여 각각 \$130/g, \$180/g, \$45/g이었다. CASE III와 같이 inclusion body를 위한 회수 설비가 없는 공정이 tailspike 단백질을 생산하기위한 가장 좋은 것으로 판명되었다 (Fig. 4). 외부단백질의 과대 생산은 일반적으로 inclusion body를 주로 형성하기 때문에 일반적으로는 당연히 CASE II의 공정을 선택해야한다. 그러나 연구활동에의해 natural 단백질의 생산비율을 늘릴 수 있다면 CASE I의 공정을 당연히 선택해야한다. 그러면 inclusion body에 비교하여 natural 단백질이 어느 정도 일 때 CASE I 보다 CASE III의 경우가 유리하게 될 것인가 하는 문제를 풀기 위하여 다음의 시뮬레이션을 수행하였다. 인위적으로 inclusion body와 natural 단백질의 비를 가정하여 CASE I 과 CASE III의 경우를 비교하였다(Fig. 5).

Table 1. Economic Analysis.

	CASE I	CASE II	CASE III
Total Equipment Purchase Cost	\$ 1,636,000	\$ 1,464,000	\$ 690,000
Annual Operating Cost	\$ 6,321,000	\$ 5,476,000	\$ 2,151,000
Relative Production Yield	58%	38%	60%
Protein Production	46,400g	30,400	48,000g
Cost per protein Recovered	\$ 130/g	\$ 180/g	\$ 45/g

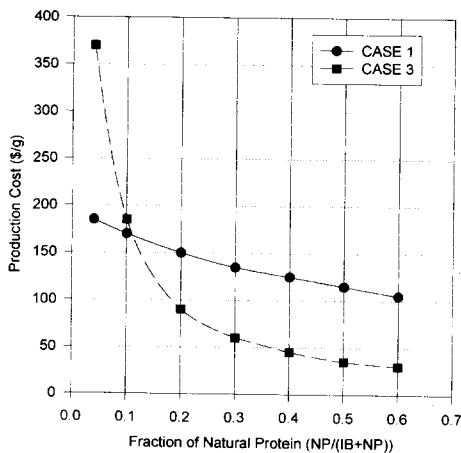


Fig. 5. Influence of Production Methods on Production Cost as a Function of Natural Protein Fraction.

Fig. 5에서 CASE III를 위한 점선은 CASE I을 위한 선보다 급하게 감소하는 것을 볼 수 있다. 그 이유로는 natural 단백질의 양이 0에 가까울수록 투자비의 생산되는 단백질의 양에 대한 비는 무한대로 증가한다. Fig. 5에서 natural 단백질의 비가 0.12일 때 두개의 선이 교차하는데 natural 단백질이 12% 보다 많을 경우에는 CASE III 타입의 공정 flowsheet가 선호된다.

결 론

수많은 선택 가능성중에서 최적의 생물화학공정을 합성하기 위해서는 수율뿐만아니라 생산되는 단백질

당 투자비에 기초한 경제성분석을 고려하여야 한다. 생물공정 합성용 flowsheeting 프로그램은 최적공정을 선택하는데 시간과 노력을 절감하게 해준다. Wild-type tailspike 단백질과 유전자 조작을 통하여 변환된 두개의 mutant 단백질(tsf G244-R과 Su A334-V)들이 본 연구에서 고려되었는데 세 가지 타입의 단백질들은 숙주 안에서 생성될 때 inclusion body와 natural 단백질의 생성비율이 서로 다르게 된다. Wild-type tailspike 단백질과 두개의 mutant 단백질(tsf G244-R과 Su A334-V)들은 각각 다른 공정을 통하여 단백질을 회수 정제하게 되는데, 예측한대로 Su A334-V 생산 시스템이 가장 경제성이 있는 것으로 추정되었다(Su A334-V 단백질의 경우 natural protein이 inclusion body보다 많이 생산되고 tsf G244-R 단백질의 경우 inclusion body만 생산하게 됨). 본 실험에서 사용된 tailspike 단백질의 경우 natural protein이 12% 이상일 때는 inclusion body 회수공정은 경제적으로 불필요하게 된다. 결론적으로 외부단백질을 생산하는 숙주-벡터 시스템에서는 natural 단백질만을 회수하는 CASE III 형태의 공정이 CASE I 형태의 공정보다 경제적으로 유리하게되는 특이점(natural 단백질과 inclusion body의 비)이 있다. 따라서 natural 단백질을 더 많이 생산하는 시스템의 개발은 아주 중요하게 된다. 유용한 단백질을 생산하기 위한 생물공정을 설계하고 합성하고자 할 때에는 경제성분석 기능을 가진 flowsheeting 소프트웨어가 필수적이다.

요 약

경제성분석 기능을 가진 생물공정 flowsheeting 소프트웨어인 BioPro Designer를 이용하여 외부단백질 생산 시스템의 최적 공정설계가 수행되었다. BioPro Designer를 이용하여 Salmonella phage P22의 tailspike protein의 생산시스템에 적용되었는데 최적공정을 선택하는 과정에서 투입되는 시간과 노력에 획기적인 개선이 있었다. 숙주의 내부에서 생산되는 외부 단백질의 folding 특성이 가장 좋은 생산시스템을 선택하는데 영향을 미친다는 것을 보여주기 위하여 wild type의 tailspike와 두개의 다른 변종 tailspike들, 즉 tsf G244→R과 Su A334→V를 비교 검토되었다. 회수된 tailspike 단백질의 단위질량 대비 자본투자비를 비교 척도로 하여 최적의 tailspike 생산 시스템이 선택되었다.

참고문헌

1. C. A. Haase-Pettingell and J. King(1988), *J. Biol. Chem.*, **263**, 4977.
2. A. Mitraki, B. Fane, C. Haase-Pettingell, J. Sturtevant, and J. King(1991), *Science*, **253**, 54.
3. A. Mitraki, C. Haase-Pettingell, and J. King (1991), *ACS Symposium Series*, **470**, 35.
4. M. S. Peters and K. D. Timmerhaus(1991), *Plant Design and Economics for Chemical Engineers*, 4th ed., McGraw-Hill, Inc.
5. J. Kim, H. Koh, D. Y. Kim, and M. Yu(1990), *Korean Biochem. J.*, **23**, 75.
6. D. Petrides, C. L. Cooney, L. B. Evans, R. P. Field, and M. Snoswell(1989), *Computers & Chemical Engineering*, **13**, 553.
7. D. Petrides(1990), *Computer-Aided-Design of Integrated Biochemical Processes : Development of Biodesigner*, Ph. D. Thesis, Dept. of Chemical Engineering, MIT.
8. J. King, B. Fane, C. Haase-Pettingell, A. Mitraki, R. Villafane, and M. Yu (1990), *Protein Folding*, p. 2205, AAAS Press, Washington.
9. R. Seckler, A. Fuchs, J. King, and R. Jaenicke (1989), *J. Biol. Chem.*, **264**, 11750.