

Directed Evolution in Protein Functionality Improvement

김 학 성, 강 환 구, 남 수 완

1. 서론

생물체내에 존재하는 단백질 과 효소는 생명체의 유지에 필요한 기능을 효율적으로 수행 하도록 매우 오랜 시간 동안에 걸쳐 진화되어 왔다. 이러한 단백질 이나 효소를 실제 산업적으로 활용하기 위해서는 단백질이나 효소의 기능을 목적에 맞도록 개량하여야 한다. 예를 들어 의약품 단백질의 경우 생체내에서의 안정성, 면역 반응성, 약리 활성등을 개량할 경우 여러 가지 측면에서 의미가 있다. 또한, 생체내에서 단백질의 발현을 조절하는 regulator를 변형시킴으로써 산업적으로 중요한 단백질들의 생산성을 획기적으로 증대시킬 수 있다. 특히 효소의 경우 활성, 열 및 유기용매에 대한 안정성, enantioselectivity, 기질에 대한 특이성등을 목적에 맞도록 개량하게 되면 산업적 응용을 크게 증대시킬 수 있다.

지금까지 단백질이나 효소의 기능을 개량하기 위한 연구는 주로 효소에 국한되어 집중적으로 진행되어 왔으며, 구체적인 방법으로는 무작위로 효소의 유전자에 변이를 유도하는 방법, 그리고 효소의 구조, 또는 기능과 구조와의 관련성등 이미 알려진 정보를 기반으로 하는 단백질공학(Protein Engineering)에 의존하여 왔다. 그러나, 이러한 방법은 효소의 기능을 산업적 이용이 가능한 수준까지 효율적으로 개량하는데 한계가 있는 것으로 보고되고 있다. 최근 기존의 방법보다 매우 빠르고 획기적으로 효소의 기능을 개량할 수 있는 방향적 진화(Directed evolution) 방법이 보고되면서 이를 효소뿐만 아니라 단백질에도 광범위하게 활용하고자 하는 연구가 급속도로 증가하고 있다. 본 고에서는 단백질과 효소의 기능을 획기적으로 개량하는 방향적 진화 방법과 이의 구체적인 적용사례에 대해 기술하고자 한다.

2. 방향적 진화(Directed evolution)의 기본 개념

단백질 혹은 효소의 구조나 기능에 대한 구체적인 data가 없는 경우에 매우 효율적으로 단백질이나 효소의 기능을 개량할 수 있는 방법으로 이의 기본 개념은 그림 1과 같다. 즉, 먼저 단백질이나 효소의 유전자에 여러 가지 방법으로 변이를 유도한 다음 이를 적당한 숙주세포에 발현 시켜 mutant library를 구축한다. Library로부터 특성이 개량된 positive mutants를 적당한 방법으로 선별하고 이를 다음 단계의 template로 사용하여, 단백질이나 효소의 기능이 목적하는 만큼 개량될 때까지 이러한 과정을 반복한다. 여기서 단백질이나 효소의 유전자에 돌연변이를 가하여 다양한 mutant library를 구축하는 방법에 따라 단백질 및 효소의 진화 속도가 결정되

며 이는 곧 기능이 개량된 단백질과 효소를 얼마나 빠르게 얻을 수 있는지와 직결된다. 또한, 수 많은 변이주로부터 원하는 방향으로 기능이 향상된 변이단백질을 선별하는 방법도 단백질의 방향적 진화 효율을 결정하는 주요한 요인이 된다. Directed evolution과정에서 mutant library를 구축하기 위해 변이를 가하면 주로 negative 방향으로 진화된 변이주들이 대부분이고 원하는 방향으로 진화된 것은 매우 드물다. 따라서, 방향적 진화를 효율적으로 진행하기 위해서는 mutation rate 또는 mutation space를 적당히 조절하는 일이 중요하다. 일단 mutant library가 구축되면 이로부터 여러 가지 functional diversity를 갖도록 selection pressure를 가하여 방향적 진화를 유도하면 그림 2와 같이 목적하는 기능을 갖는 단백질이나 효소를 얻을 수 있다.

단백질 및 효소의 방향적 진화는 이를 coding 하는 유전자에 돌연변이를 가하여 mutant library를 구축하는 것으로부터 시작되는데 mutant library를 구축하는 일반적인 방법은 다음과 같다.

A. Chemical mutagenesis

DNA의 nucleotide에 변이를 주는 방법은 in-vitro DNA 합성, 또는 in-vivo DNA 복제동안에 염기들끼리의 mis-incorporation을 유발한다. 이러한 방식은 유전자안에 적은 수의 mutation을 유발시킬 때 많이 이용되어 왔다. 전통적으로 EMS (Methylsulfonate ethylester), MNNG (N-methyl-N-nitro-N-nitrosoguanidine;NTG) 등이 돌연변이를 유발하는데 주로 이용되어왔으며, 근래는 nitrous acid(HNO_2), formic acid, sodium bisulfate, hydrazine, dimethylsulfate 등의 chemical들이 DNA의 damage에 이용되고 있다.

B. Error-prone PCR (Mutagenic PCR)

PCR 반응의 주요목적은 특정 유전자를 높은 정확도로 증폭하는 것이다. DNA polymerase의 높은 fidelity ($3' \rightarrow 5'$ exonuclease activity (proof reading activity라고도 함))를 이용하여 유전자를 증폭하는 것인데 전형적인 반응조건은 $1.5 \text{ mmol l}^{-1} \text{ MgCl}_2$, $50 \text{ mmol l}^{-1} \text{ KCl}$, 10 mmol l^{-1} , $10 \text{ mmol l}^{-1} \text{ Tris-HCl}$, pH 8.3, 0.2 mmol l^{-1} 의 각각의 dNTP, 0.3 mol l^{-1} 의 각 primer, 2.5 unit thermostable DNA polymerase를 총부피가 100 l가 되게 하여 94에서 1분, 45C에서 1분, 72C에서 1분간의 thermal cycle을 반복하는 것이다. Thermostable DNA polymerase로 *Taq* polymerase를 사용할 경우 amplification과정에서 nucleotide 당 $0.1 - 2 \times 10^{-4}$ 의 빈도로 다른 nucleotide가 결합하여 돌연 변이가 발생하게 된다. Mutation frequency를 증가시키기 위해 보통 MgCl_2 의 농도를 7 mmol l^{-1} 까지 증가시키거나 0.05 mmol l^{-1} 의 MnCl_2 를 첨가하거나 dCTP와 dTTP의 농도를 1 mmol l^{-1} 로 올려주고, 또는 *Taq* polymerase를 5 unit 정도 첨가하면 된다. PCR 수행조건을 다르게 하면 mutation rate의 조절이 가능하고 이를 바탕으로 mutant library를 구축할 수

있다.

같은 맥락에서 PCR 반응에 기존에 알려진 mutagenic nucleoside analog를 보통의 deoxynucleoside triphosphates(dNTPs) pool에 첨가하여 변이를 유발시키는 방법이 있다. 8-OxodG를 포함하는 8-oxo-substituted nucleoside가 library제작에 이용되어 왔고 이 방법을 통해 mutant library를 얻을 수 있다. 또한 보통 이용되는 dNTPs의 성분비율을 한쪽으로 편향되게 하여 mutation을 유발시킬 수 있다. 이때 incorrect base는 원래의 base에 비해 incorporation되는 속도는 느리지만 충분히 library를 만들 수 있다. Deoxyinosine triphosphate(dITP)도 정상적인 dNTP 위치에 misincorporation하여 변이를 유발시킬 수 있다. 최근에는 error-prone mutant Taq polymerase가 PCR을 통한 mutagenesis에 이용되어오고 있다.

이러한 PCR을 통한 방법들은 크게 다음과 같은 단점을 가지고 있다. 하나는 짧은 길이의 유전자 부위안에서 돌연변이가 일어날 빈도가 낮아서 효소의 활성부위 안에 mutiple mutation을 유발시키기에 충분치 못하다는 것이고 또 한가지는 발생하는 mutation이 완전히 random하지 않다는 것이다. 바꿔 말하면 damage/incorporation 이 아주 잘 되는 부위가 있어서 mutation space 가 넓지 않게 된다.

C. Mutagenic PCR with Random oligonucleotides

특정위치에 4 가지의 nucleotide의 random한 조합을 가지고 있는 oligonucleotide를 이용하여 효소를 coding하는 유전자의 부분을 바꾸는 방법으로 유전자 diversity를 얻는데 유용하다. 이 방법에 의하면 원하는 위치의 base가 random하게 바뀐 재조합 plasmid의 library를 만들 수 있다. Randomization의 정도와 각각의 아미노산 치환 정도는 원래의 염기와 다른 3가지 염기간의 비율로 조절할 수 있다. 그래서 원래의 염기로 편향된 random oligonucleotides를 이용하면 치환될 아미노산의 수를 제한할 수 있다. 이 방법을 tetracycline resistance gene의 promoter region과 b-lactamase의 활성부위를 무작위로 변형시킨 연구가 보고되었다.

D. Saturation mutagenesis

이 방법은 유전자의 특정 아미노산을 가능한 모든 아미노산으로 치환하여 변이를 유발시키는 것이다. 가장 직접적인 방법은 site-directed mutagenesis를 통하여 모든 아미노산으로 치환되도록 필요한 nucleotide exchange를 도입하는 것인데, 이는 많은 mutagenic oligonucleotide primer를 요구하며, 생성된 변이주의 변이 여부를 DNA sequencing으로 확인하여야 한다. 일반적으로 이 방법은 특정 gene fragment를 제한 효소로 절단하여, bacteriophage M13 phage에 삽입한 후, C to A transversion, T to G transversion, A to C transversion이 가능한 primer를 제작하여, T4 DNA polymerase, T4 DNA ligase를 이용하여, circular DNA를 만든 후, 이를 *E.coli* 에 transformation 시킨다. 이렇게 생성된 circular DNA에 ATP를

labeling한 wild-type DNA를 probe로 이용하면 한 개의 mismatch codon을 갖고 있는 phage DNA는 wild-type DNA와 hybridization하지 않으므로 따로 분리할 수 있다.

E. Cassette mutagenesis

일반적으로 Cassette는 한 개에서 수 백 개의 아미노산을 coding하는 3 개에서 수 백개의 nucleotide로 정의된다. 구체적인 방법은 먼저 combinatorial cassette mutagenesis(CCM)를 통하여 이미 알려진 3차원적 구조의 정보를 이용하여 유전자의 특정 부위를 치환한다. CCM은 randomized codon을 함유한 oligonucleotide를 이용하여 이를 유전자의 특정 부위에 mutagenic cassette로 도입하여 돌연변이를 유발시킨다

F. Incremental truncation

대상이 하나의 유전자(ORF1)를 phagemid vector에 cloning하고 동일한 방법으로 다른 유전자(ORF2)를 같은 replication origin을 갖는 vector에 cloning한다. 이때 용이한 선별을 위해 각각의 vector는 서로 다른 항생제 내성을 갖도록 구축한다. 제한 효소의 절단부위를 이용하여 3' 말단으로부터 internal truncation을 수행한 ORF1과 5' 말단으로부터 internal truncation을 실행한 ORF2를 확보한다. 이로부터 하나의 library의 유전자 단편을 분리하여 다른 library의 유전자에 fusion시키게 되면 mutant library가 구축된다.

G. Homologous recombination

먼저 대상이 되는 하나의 유전자(ORF1)를 특정한 vector에 cloning한다. Homologous recombination을 유도하기 위해 다른 하나의 유전자(ORF2)는 자체를 이용한다. 두 개의 유전자를 electro-transformation을 통하여 동시에 하나의 숙주 세포에서 발현시킴으로써 mutant library를 구축하는 방법이다. 이때 숙주세포는 homologous recombination에 필요한 효소가 발현되는 시스템을 이용하여야 한다.

H. Bacterial mutator strain

생물체는 chemical agent나 UV irradiation, 효소자체의 error rate등 여러 가지 요인에 의해 자연적으로 발생할 수 있는 mutation을 극복하기 위해서 다양한 mechanism을 이용한다. 만약 이러한 repair mechanism을 억제한다면 DNA 복제시 mutation frequency가 높아질 것이고 유전자의 library 구축에 사용할 수 있다. 이를 기초로 고안된 방법이 *E.coli* mutator strain을 이용하는 것이다. DNA repair mechanism인 mismatch repair(mutS), oxo-GTP repair(mutT) 그리고 DNA pol III

의 3' → 5' exonuclease activity(mutD)가 결여된 *E.coli* 균주는 wild type에 비해서 mutation rate가 5000 배나 높은 것으로 확인되었고, 이러한 균주를 이용하면 특정 유전자의 mutant library를 구축하는데 유용하게 활용될 수 있다. Mutator strain을 이용할 경우 다른 방법과는 달리 in-vitro에서 mutation을 일으켜 숙주에 transformation하는 과정을 생략하고 바로 selection 이나 screening할 수 있는 장점을 가지고 있다.

I. Staggered Extension Process

하나의 특정 primer를 template DNA에 부착하여 DNA polymerase에 의한 primer extension reaction을 진행하다가 이를 다시 annealing 하여 다른 DNA template에 부착시켜 extension reaction을 진행한다. 이러한 반응을 full-length gene이 만들어질 때까지 반복하여 mutant library를 구축하는 방법이다. 비슷한 방법으로 random-priming recombination으로 random primer를 이용하여 target DNA의 각각 다른 segment에 상보적인 50-500 bp의 short DNA fragment를 만든다. 다음 단계로는 mispriming과 base misincorporation에 의해 생성된 short DNA fragment를 서로 priming하게 하여 full-length recombination gene을 만드는 것이다. 이 과정을 그림으로 도식화하면 아래 그림과 같다. 이러한 방법이 DNA shuffling에 비하여 갖는 이점은 mRNA를 포함하여 single strand template를 이용하므로 template DNA의 모든 nucleotide의 mutation rate를 유사하게 조절할 수 있다는 것이다.

J. DNA shuffling

현재 유전자의 mutant library를 구축하는데 가장 효율적인 방법으로 이용되는 것이 1994년 Stemmer에 의해 개발된 DNA shuffling 또는 sexual PCR이다. DNA shuffling이라고 불리는 이유는 마치 카드를 섞듯이 DNA 유전자 조각들을 무작위로 순서없이 섞어서 매우 다양한 mutant library를 구축하기 때문이다. 또한 sexual PCR이라고도 하는 이유는 sequence space를 확대하기 위해 같은 family에 속하는 homologous 유전자를 template로 사용하거나 하나의 template로부터 유래한 다양한 유전자를 template로 사용하므로 이것이 sexual reproduction과 유사한 방법으로 매우 다양한 변이 유전자를 생성시키기 때문이다. 전자의 경우를 family shuffling이라고 부른다.

DNA shuffling은 염기 서열이 비슷한 유전자들 또는 mutation에 의해 발생한 유전자들로부터 다양한 조합을 만들어 library를 구축한다. 염기서열이 비슷한 유전자들은 DNase I에 의하여 자르면 random한 DNA 조각들이 만들어진다. 이러한 조각들을 이용하여 PCR을 하게 되는데, 이때 각각의 조각들은 비슷한 염기서열을 가지고 있는 다른 조각에 대해 primer로 작용하게 된다. 예를 들어 A라는 유전자의 조각은 비슷한 서열을 가지고 있는 B라는 유전자 조각에 primer로 작용하게 되어 A 유전자 조각과 B 유전자 조각이 동시에 PCR에 의해 증폭이 된다. 이렇게

reassembled 된 조각들은 다시 PCR을 통해 full-length를 갖는 전혀 새로운 유전자가 된다. 이 방법은 유전자 조각들의 순서를 바꿀 뿐 아니라 DNase I 자체의 randomness, 그리고 PCR을 거칠 때의 error가 겹쳐지면서 훨씬 다양한 유전자 library가 만들어지게 한다. 반면 만약 이미 밝혀진 유용한 mutation을 가지고 있는 유전자를 이용하여 shuffling할 때는 염기의 치환은 바람직하지 않기 때문에 DNase I fragmentation 과정에서 Mg 대신 Mn을 이용하거나, 가능한 한 PCR회수를 줄이고, high fidelity를 갖는 DNA polymerase를 이용하여 PCR을 하여야 한다.

3. Gene Expression System

효소의 방향적 진화를 위해서는 위에 열거한 방법으로 변이 유전자를 만든 다음 이를 적절한 숙주세포에서 발현시켜야 한다. 최근까지 대장균이 숙주세포로 많이 이용되었는데 목적하는 유전자를 적당한 promoter와 결합하여 발현 vector에 삽입한다. 많이 사용되는 promoter로는 lac, tac, T7-lac 등이 있다. 한편, lipase 나 protease 등과 같은 분비 단백질은 세포내에서 발현될 경우 folding 등의 문제로 활성이 없거나 매우 낮은 경우가 발생하므로 이들을 효율적으로 분비시키는 숙주세포의 구축이 요구된다. 효소의 방향적 진화에서 대장균과 더불어 사용될 수 있는 숙주세포로는 gram-positive에 속하는 *Bacillus*, *Clostridium*, *Lactococcus*, *Staphylococcus* 등이 있다. 최근에는 yeast가 eukaryotic expression host로 중요성이 점차 증가하고 있다.

4. 변이주의 screening

위에 설명된 방법들을 이용하면 대강 10^{10} 까지의 거대한 mutant library가 구축된다. 이러한 library 에서 기능이 향상된 변이 효소를 선별하는 것이 방향적 진화의 성패를 좌우 할 수도 있다. 많이 사용되는 방법으로는 효소의 활성을 비교할 수 있는 activity staining method, colorimetric method, 숙주세포의 성장과 관련된 complementation method 등이 있는데 각 각의 효소에 따라 민감도가 충분한 선별 방법을 구축하는 것이 요구된다. 예를 들어 최근에 개발된 thermographic method는 chiral alcohol의 R, S-configured enantiomer의 반응에 의해 생성되는 black-body radiation을 IR-camera로 시간의 흐름에 따라 선별하는 방법이다.