

# 신 기능 단백질 설계를 위한 Loop Engineering

KAIST 생물과학과

김학성 · 박희성 · 남성훈  hskim76@kaist.ac.kr

자연계에는 5만 종류 이상의 다양한 기능을 수행하는 단백질이 존재한다고 알려져 있고 약 30,000 종류의 단백질 구조가 밝혀졌다. 단백질 구조의 비교 연구 결과, 특이하게도 이들 단백질들의 구조는 500~600개 정도의 한정된 단백질 골격(scaffold)에 기초해서 이루어져 있는 것으로 밝혀졌다. 현재 다양한 기능의 단백질들은 오랜 시간에 걸친 다윈식 진화 과정, 즉, 돌연변이와 선택의 과정을 통해서 분화되었다고 알려져 있다. 자연계에서 이루어지는 단백질의 진화는 진화 초기의 원시 골격으로부터 시작된다. 단백질 원시 골격은 유전자수준에서의 중복과 변형의 과정을 통해서 여러 형태의 미분화 골격으로 진화가 이루어진다. 미분화 골격은 다시 유전자에 삽입, 제거, 재조합 등과 같은 복잡한 유전자 재배열과정을 거쳐서 다양한 형태의 원시 단백질로 분화된다. 이들 원시 단백질은 구조적으로 불안정하거나 매우

약한 활성을 보인다. 원시 단백질들은 그 골격을 유지한 채로 단일 염기 돌연변이의 축적을 통한 미세 조정과정을 거쳐서 지금과 같은 다양한 형태와 기능을 갖는 단백질 거대그룹(Superfamily)이 형성되었다고 믿고 있다. (Fig. 1) 자연계는 처음부터 새로운 단백질을 만드는 de novo 합성 보다는 한정된 수의 단백질 구조 골격을 이용하여 다양한 기능과 특성의 단백질을 만들어 낸 것으로 인식되고 있다. 이러한 단백질의 진화 기작을 고려할 때, 각각 다른 기능을 수행하는 단백질이라 할지라도 3차원 구조의 기본 골격은 유사할 수 있고, 비슷한 기능을 수행하는 단백질이라도 그 구조는 상이한 경우도 있을 것이다. 기능이 다른 단백질이 구조상에서 비슷한 골격을 유지하고 있다는 사실은 단백질의 기능 및 특성을 변화시키는데 있어서 큰 힌트를 준다.

단백질이 그 특유의 기능을 발휘하기 위해서

는 3차원 구조의 형성이 중요하다. 단백질의 3차원 구조는 폴리펩타이드 사슬 내의 특정한 아미노산들이 도메인을 형성하면서 생성된다. 단백질의 1, 2차 구조에 존재하는 특정 아미노산들이 3차 구조를 형성하면서 공간적으로 서로 만나게 되고, 이렇게 접혀진 도메인에서는 서열상으로는 멀리 떨어진 아미노산들이 근접하면서 단백질들에서 발견되는 특이적인 촉매 또는 기질 결합 부위들이 만들어진다. 이를 통해 단백질에서 활성부위나 결합부위에 작용기들이 적절한 공간적 배열을 가지게 되는 것이다. 단백질의 3차원 구조는 그 기능 면에서 중요한 역할을 하는 여러 작용 부분들의 공간적 배열을 통해 형성된 활성 부위와 단백질이 안정하게 유지되도록 그 구조를 유지하는 골격 부분으로 크게 나누어 볼 수 있다. 한정된 골격을 가지는 단백질 세계에서 그보다 훨씬 많은 다양한 기능을 수행하기 위해서는 그 골격은 비슷하게 유지되지만, 특유의 기능을 수행

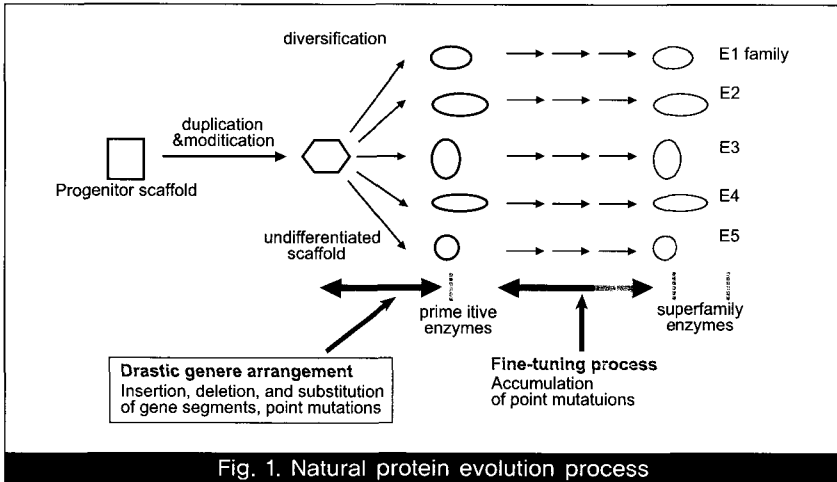


Fig. 1. Natural protein evolution process

하도록 활성 부위가 달라진 단백질들의 경우가 많을 것이며, 실제로 구조가 밝혀진 단백질들의 경우 이러한 사실을 확인할 수 있다. 따라서, 단백질의 자연 진화적 과정과 구조적 특징을 이용하게 되면 기존의 단백질 골격에 새로운 기능과 특성을 갖는 단백질을 설계하는 것이 가능하다.

최근에 발표된 논문 (Park et al., Science, 311, 535-538, 2006) 에서는 위에 언급한 단백질의 진화 과정을 단순화하여 새로운 기능을 갖는 단백질의 설계 방법에 관한 흥미있는 결과를 소개하고 있다. 즉, 단백질의 loop을 재설계하여 새로운 기능을 갖는 단백질을 창출하는 방법을 설명하고 있다. 본 논문에서 사용한 방법은 다음과 같은 여러 단계로 구성되어 있다. 즉, 기존에 존재하는 단백질 골격에 새로운 기능에 필요한 기능요소들을 설계하는 기능요소 설계 단계, 설계된 기능요소들에 해당하는 두 개 이상의 유전자 단편들을 단백질 골격 유전자에 동시에 삽입하는 기능요소 삽입단계, 새로운 변이 유전자들이 삽입된 변이체들의 라이브러리로부터 새로운 기능을 가진 단백질의 변이체를 선별하고 그 기능을 향상 및 조정하는 개량 단계이다. 이 과정을 좀더

자세히 설명하면 다음과 같다. 먼저 새로운 기능을 부여하고자 하는 단백질의 골격을 기존 단백질의 3차원 구조 데이터베이스 및 연구 문헌을 통해서 선택하게 된다. 대상 단백질의 골격이 결정되면 이 골격에 삽입되어질 새로운 기능에 필요한 기능 요소를 선택하여 설계하게 된다. 단백질에서 새로운 기능을 수행하는데 필요한 요소들인 활성 부위를 구성하는 촉매 기능 부위(catalytic element)나 기질 결합 부위(substrate binding site), 그리고 리간드 결합 부위인 단백질 루프(loop) 구조 및 아미노산 단편과 같은 중요 부위를 단백질의 아미노산 서열 유사성 비교연구, 반응 기작 및 3차원 구조에 대한 정보를 통해서 설계한다. 설계 기능 요소는 단일 아미노산 뿐 아니라 아미노산 단편 및 단백질의 2차 구조와 같이 비교적 긴 부분의 치환 및 삽입도 포함되기 때문에 구조적으로, 기능적으로 단백질 골격에 상당한 영향을 미친다. 치환되어질 부분 중 새로운 기능에 중요한 부위의 경우에는 유사 단백질들의 아미노산 서열을 비교 분석하여 보존 부위(consensus amino

acid sequence)와 임의부위(random amino acid sequence)가 포함되도록 다양한 길이와 서열의 변이 아미노산 단편들로 설계한다. 이를 통해 설계된 기능 요소들에 해당되는 합성 유전자를 해당 단백질 골격 유전자에 동시에 도입시킬 때 임의로 무작위의 아미노산이 삽입되도록 하면 새로운 기능 창출의 가능성을 높게 된다.

새로운 기능 요소를 삽입하기 전에 우선 새로운 기질 및 리간드에 대한 결합을 방해하거나 혹은 공간적으로 제약을 가하는 불필요한 부분을 유전자 재조합을 통해서 해당 단백질의 골격으로부터 제거한다. 다음으로, 새로운 기능에 필수적인 아미노산들은 유전자의 점 돌연변이를 통해 해당 아미노산으로 치환한다. 이렇게 얻어진 단백질 골격에 길이가 다양한 아미노산 단편 및 단백질 2차 구조 형성에 필요한 기능 요소들을 유전자 재조합을 통해서 동시에 삽입, 치환한다. 이 과정 중에 전체 유전자상에서 임의적으로 돌연변이가 유발되도록 반응 조건을 조정함으로써 부가적인 돌연변이를 통한 새로운 단백질 기능의 변화를 효율적으로 유도할 수 있다. 이와 같은 일련의

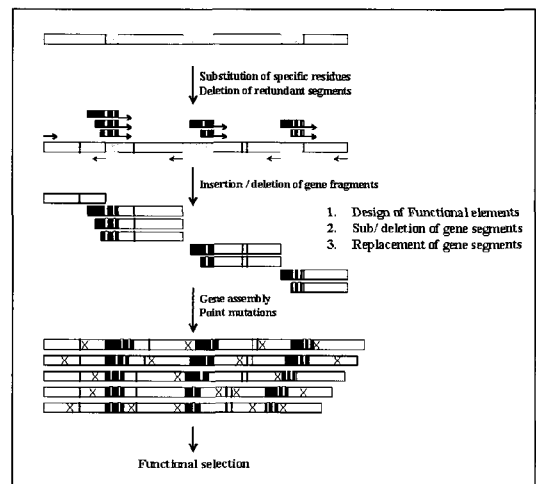


Fig. 2. SIAFE (Simultaneous Insertion and Adjustment of Functional Elements)

유전자 재조합 과정을 통해서 새로운 기능의 활성 부위를 위해 필요한 기능 요소들을 임의적으로 또한 복합적으로 동시에 해당 단백질 골격에 조합적으로 삽입하여 다양한 변이체의 library를 만들게 된다.(Fig. 2)

특히, 단백질의 일종인 효소는 최근 선진국을 중심으로 대대적인 연구개발 및 산업화를 추진하고 있는 White Biotechnology의 핵심으로 부각되어 세계적 화학 기업, 제약기업, 생명공학 기업들이 맞춤형 효소의 개발에 연구개발 투자를 집중하고 있다. 이런 관점에서 본 기술은 산업용 효소의 개발에도 널리 활용될 것으로 기대되고 있다.

다양한 서열의 변이 단백질 집단 중에서 새로운 기능을 획득한 단백질은 활성도, 리간드와의 결합 정도, 형광 물질을 이용한 측정 등의 여러 가지 방법을 통해서 선별하게 된다. 새로운 기능을 가진 것으로 선별된 변이 단백질의 경우 대부분 아주 낮은 활성을 가지거나 구조적으로 불안정한 경우가 많다. 따라서, 선별된 변이 단백질들은 새로운 기능의 향상 및 안정화를 위해 임의의 유전자 부위에 돌연변이를 유발함으로써 단백질의 특성 향상에 효과적인 방향적 진화 방법을 이용하여 그 특성을 점진적으로 개량한다. 상기와 같은 일련의 과정을 통해서 기존의 단백질 골격에 새로운 기능을 지닌 새로운 단백질을 효과적으로 제조할 수 있다.

본 논문에서는 실제로 위의 언급한 일련의 과정을 통해 사람의 몸에서 독성물질을 해독하는 기능을 하는 'Glyoxylase II (GlyII)', 골격에 베타 락탐 항생제를 분해하는 기능을 하는 'Metallo  $\beta$ -Lactamase (IMP-1)' 활성을 창출하였다. 이 두 단백질은 각각 다른 생명체 내에서 다른 기능을 수행하지만, 그 골격 구조는 유사하며 촉매작용에 결정적 역할을 하는 금속 결합부위

(Metal binding site)를 공유하고 있다는 점에서 원형(Progenitor) scaffold의 분화적 진화의 전형적인 사례로 인식되어 왔다. 본 연구는 자연계에서 오랜 시간에 걸쳐 진화되어 온 두 단백질을 새로운 단백질 설계 기술을

통해 실험실에서 매우 짧은 시간에 진화를 유도하는 것이다. 이를 구현하기 위해 두 가지 단백질의 3차 구조, 아미노산 서열, 반응기작의 정보를 바탕으로 Gly II 단백질에 IMP-1의 기능을 하는 활성부위를 만들기 위해 필요한 기능 요소를 동시에 incorporation하고 조정하는 방법을 사용하였다. 먼저 Gly II의 C-terminal region은 구조적으로 새로운 기질의 접근을 방해할 가능성이 높기 때문에 제거하였고 다음으로 metal binding ligand를 IMP-1 반응에 적합하게 변형시켰다. 그리고 새로운 기질, 즉 Cefotaxime과 결합할 수 있는 활성부위를 만들기 위해 GlyIII 골격에서 4개의 단백질 Loop 구조를 베타 락탐 항생제를 분해할 수 있도록 기능요소를 설계해서 이를 동시에 삽입·치환시킴으로써 새로운 기

능인 베타락타메이즈 활성을 갖도록 하였다. 전체 단백질 구조에서 기본적인 골격은 유지한 채 기능부위에 해당되는 4개의 아미노산 단편의 loop 삽입, 치환시킴으로써 원래 Gly II의 기능은 사라지고 새로운 기능인 IMP-1의 활성을 갖는 새로운 단백질이 창출되도록 하였다. (Fig. 3)

가장 널리 알려진 단백질 진화 이론에 의하면, 단백질의 원형 구조를(Progenitor scaffold)은 유지되면서 다른 기능 요소들이 첨가되거나 변화되면서 현재와 같은 다양한 기능을 갖는 단백질의 superfamily가 형성되었을 것으로 추정되고 있다. 따라서, 이번 연구 결과는 현존하는 다양한 단백질이 어떻게 진화되어 왔는지에 대한 중요한 해답을 줄수 있어서 기초 생명과학 분야에서 큰 관심을 불러일으키고 있다. 또한, 본 연구에서 개발된 기술은 자연계에서 단백질이 진화한 과정을 단순화하여 새로운 기능을 갖는 단백질을 효율적으로 설계하고 창출하는 방법으로써, 앞으로 새로운 단백질 의약품 개발, 산업용 효소 개발, 합성생물학, White Biotechnology, 생유기 합성 및 단백질 공학 분야에서 크게 활용될 것으로 기대되어 생명공학의 산업화에 크게 기여할 것이다. 특히, 단백질의 일종인 효소는 최근 선진국을 중심으로 대대적인 연구개발 및 산업화를 추진하고 있는 White Biotechnology의 핵심으로 부각되어 세계적 화학 기업, 제약기업, 생명공학 기업들이 맞춤형 효소의 개발에 연구개발 투자를 집중하고 있다. 이런 관점에서 본 기술은 산업용 효소의 개발에도 널리 활용될 것으로 기대되고 있다.

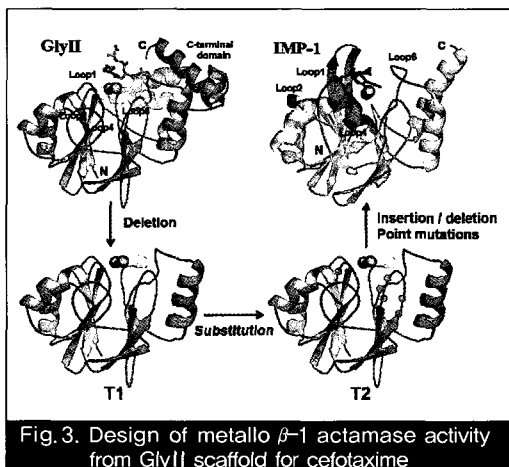


Fig. 3. Design of metallo  $\beta$ -1 actamase activity from GlyII scaffold for cefotaxime

:: Reference :: Hee-Sung Park, Sung-Hun Nam, Jin-Kak Lee, Chang No Yoon, Bengt Mannervik, Stephen J. Benkovic, Hak-Sung Kim, Science, 311, 535-538, 2006