

비천연 아미노산의 유전적 도입과 응용



서강대학교 이현수 교수

1. 서론

단백질은 광합성에서부터 세포 내의 신호전달과 면역 반응에 이르기까지 생명체 내의 거의 모든 과정에 관여한다. 이러한 단백질의 필수적인 역할을 고려할 때, 단백질의 구조와 기능 연구는 생명 현상의 이해와 응용에 매우 중요하다. 하지만 단백질을 구성하고 있는 아미노산들은 비교적 단순한 작용기 (functional group)들로 이루어져 있기 때문에 세포 내에 존재하는 단백질 자체만으로는 구조와 기능을 연구하는데 한계가 있다. 이러한 한계점을 극복하기 위해서 단백질 내에 존재하지 않는 작용기를 가지는 비천연 아미노산(unnatural amino acid, UAA)을 단백질 내에 도입하려는 노력이 진행되어 왔다. 이러한 UAA를 단백질에 도입하게 되면, 단백질의 구조와 기능 연구는 물론, 새로운 기능을 가지는 단백질을 합성할 수 있고, 나아가 새로운 기능을 수행하는 세포도 만드는 것이 가능할 수 있다. 여기서는 UAA를 유전적으로 도입하는 방법의 소개와 이를 이용한 최신 연구에 대해서 소개하고자 한다.

2. 비천연 아미노산의 유전적 도입법

2001년에 살아있는 세포 내에서 특정 단백질의 원하는 위치에 UAA를 도입하는 방법이 최초로 개발되었다[1]. 이 방법은 세포 내의 단백질 합성 시스템을 사용하기 때문에, 이 방법을 적용하기 위해서는 세포 내의 단백질 합성 시스템에 몇 가지 추가적인 요소를 더해 주어야 한다[2,3]. 먼저 20개의 아미노산 각각에 대하여 할당된 코돈(codon)이 존재하듯이 UAA를 위한 코돈이 필요하다. 마찬가지로 각각의 20개의 아미노산을 위한 aminoacyl-tRNA (aatRNA)와 aaRS가 있듯이 UAA를 위한 aatRNA와 aaRS도 필요하다. 그리고 중요한 것은 이 추가적인 요소들이 세포 내에서의 단백질 합성 시스템에 영향을 주지 않아야 한다는 점이다. 즉 UAA를 위한 aatRNA는 세포 내에 존재하는 다른 aaRS들에 의해서 아실화가 되어서는 안 되고, UAA를 위한 aaRS는 세포 내에 존재하는 다른 aatRNA에 아실화를 시키면 안 되며, UAA를 위한 tRNA만을 인지하여 아실화를 시켜야 한다 (orthogonal aatRNA/aaRS pair). 그리고 이 aaRS는 UAA만을 인지하여 그에 해당되는 aatRNA에 아실화를 시켜야 하고, 세포 내에

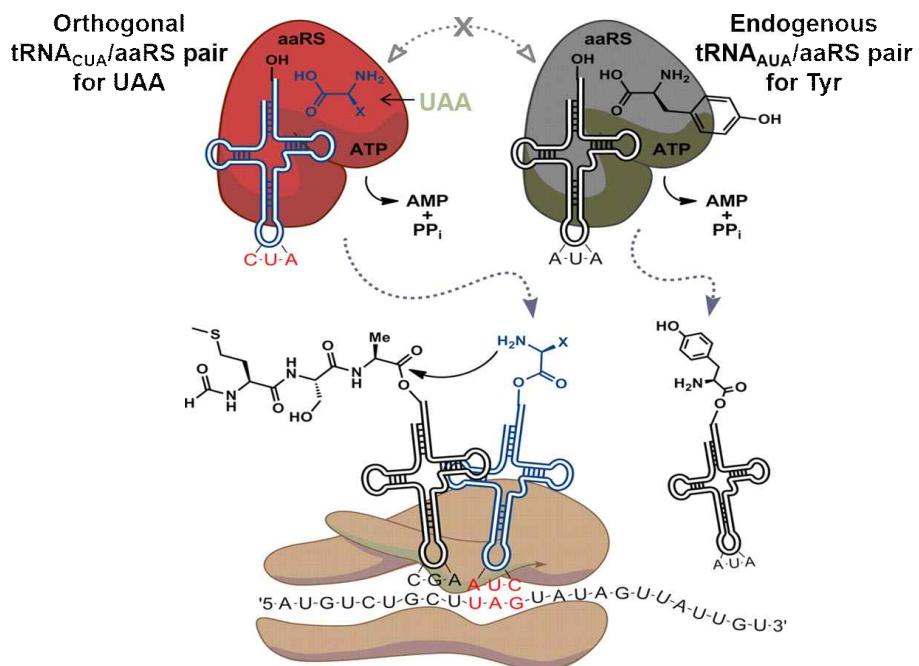


그림 1. 세포 내의 단백질 합성 시스템을 이용한 UAA의 도입.

존재하는 20개의 아미노산들은 아실화시켜서는 안 된다 (그림1). 마지막으로 UAA는 세포 내에서 단백질 합성에 사용될 수 있도록 안정해야하고, 세포 내로 잘 흡수가 되어야 한다.

2.1 사용 가능한 코돈

단백질 합성에 사용되는 유전코드는 4개의 염기 중 3개로 구성되어 64개의 코돈으로 이루어져 있고, 이 64개의 코돈 중에서 61개는 20개의 아미노산에 할당이 되어 있고, 3개는 단백질 합성을 멈추는 신호(stop codon, 정지코돈)로 사용된다. 단백질 합성을 멈추기 위해서는 이론적으로 하나의 정지코돈만 필요하기 때문에 나머지 두 개의 정지코돈은 UAA를 도입하는데 사용될 수 있다. 3개의 정지코돈 중 TAG 코돈은 대장균과 이스트에서 가장 사용 횟수가 적고, 필수적인 유전자를 멈추는 데는 거의 사용되지 않는다. 그리고 박테리아들 중에서 mRNA의 UAG 코돈(앰버 코돈, amber codon)을 인지하여 아미노산을 삽입하는 억제 tRNA (suppressor tRNA, UAG의 경우 앰버 억제 tRNA, amber suppressor tRNA)를 가지고 있는 것들이 존재한다고 알려져 있다[4,5]. 실제로 앰버 코돈은 1980년대부터 단백질 변형 실험에 사용되어 왔다[6,7]. 따라서 이 앰버 코돈은 세포의 단백질 합성 시스템에 큰 영향을 미치지 않으면서 UAA를 도입하는데 사용될 수 있다. 오펠 코돈(opal codon, TGA)도 같은 목적으로 사용될 수 있지만, 앰버 코돈에 비해 사용빈도가 높기 때문에 원하지 않은 억제가 일어날 가능성이 높다[8].

또 다른 방법은 4개로 이루어진 코돈(quadruplet codon)을 사용하는 것이다. 4개의 염기로 이루어진 코돈을 사용할 때 나타날 수 있는 문제점은 처음 3개의 염기로 구성된 코돈이 번역(translation)되는 것에 의한 코돈의 프레임쉬프트(frameshift)이다. 이러한 문제점은 처음 3개의 염기를 AGG와 같이 사용 빈도가 낮은 코돈으로 구성함으로써 어느 정도 해결될 수 있다. 최근 Schultz와 Chin 등은 4개의 염기로 구성된 AGGA 코돈을 사용하여 대장균 내에서 좋은 효율로 UAA를 도입했고, 나아가서 AGGA와 TAG를 사용하여 두 개의 다른

UAA를 한 단백질에 도입함으로써 여러 개의 다른 UAA를 동시에 한 단백질에 도입할 수 있는 가능성을 보여주었다[9,10]. 여기에서 더 나아가 Chin 등은 서로 다른 두 개의 UAA를 도입할 때의 낮은 도입 효율을 개선하여 일반적으로 사용가능한 시스템을 구축하였다[11].

2.2 사용 가능한 aatRNA/aaRS pairs

비천연 아미노산의 유전적 도입에 있어서 핵심이 되는 부분이 바로 특정 비천연 아미노산에 작용하는 aatRNA/aaRS pair라고 할 수 있다. 단백질의 생합성 과정에 관여하는 aaRS는 aatRNA의 여러 부분과 상호작용을 함으로써 다른 tRNA들로부터 그것의 짹이 되는 tRNA를 인지한다[12]. 그 상호작용이 아주 세밀하면서 복잡하고, 세포 내에 이미 각각의 아미노산에 해당하는 aatRNA/aaRS pair들이 존재하기 때문에 이들과는 독립적으로 UAA에만 작용하는 tRNA/aaRS pair를 만들어내는 것은 아주 어려운 일이다. 하지만 원핵 세포 (prokaryotic cell), 진핵 세포 (eukaryotic cell), 그리고 아케이아 (archaea)에서 사용되는 tRNA를 비교해 보면, aaRS에 의해 인지되는 부분들이 각각 다르다는 것을 알 수 있다[13,14]. 따라서 다른 종에서 사용되는 aatRNA/aaRS pair를 UAA에 대해서 사용하면 어느 정도의 orthogonality를 가질 수 있을 것이다. 실제로 몇몇 원핵 세포의 tRNA/aaRS는 진핵 세포의 tRNA/aaRS와 반응하지 않는다는 연구 결과가 알려져 있다[15,16].

현재 세포 내에서 유전적인 방법으로 UAA를 도입하는데 사용되고 있는 aatRNA/aaRS pair는 다양하게 존재하지만 주로 아래의 pair들이 많이 쓰이고 있다. 대장균에서 사용되는 것들은 *Methanococcus jannaschii* tRNA^{Tyr}/TyrRS pair[1]와 *Methanosarcina mazei* 또는 *Methanosarcina barkeri* tRNA^{Pyl}/PylRS pair[5,17-19]가 있고, 이 중 tRNA^{Pyl}/PylRS pair는 진핵 세포에서도 사용할 수 있다. 그리고 진핵 세포에서 사용되는 orthogonal tRNA/aaRS pair는 대장균 유래의 tRNA^{Tyr}/TyrRS pair²²와 tRNA^{Leu}/LeuRS pair[20]가 있다. 각 aaRS마다 선택적으로 잘 인지하는 UAA의 일반적인 구조가 다르기 때문에, UAA의 구조와 응용할 세포 등을 고려하여 적절한 tRNA/aaRS pair를 선택하는 것이 중요하다. 최근에는 tRNA^{Pyl}/PylRS pair가 많이 사용되고 있는데, 원핵/진핵 세포 모두에서 잘 작용하고, 다양한 구조의 라이신 유도체를 효과적으로 도입하는 것으로 알려져 있다[21].

2.3 aaRS의 기질 특이성 변형 (aaRS engineering)

비천연 아미노산의 유전적 도입 방법에서는 각각의 비천연 아미노산에 대해 선택적으로 작용할 수 있는 변형 aaRS가 필수적으로 필요하다. 따라서 새로운 구조의 UAA를 단백질에 도입하기 위해서는 그에 해당되는 aaRS 변이를 찾는 과정이 수반된다. 각 aaRS가 어느 종으로부터 왔느냐에 따라 aaRS 변형에 사용할 호스트가 달라지는데, 최근에는 주로 대장균을 사용하기 때문에 여기서는 대장균에서의 스크리닝 방법에 대해서만 간단하게 설명하고자 한다. aaRS의 기질 선택성을 바꾸기 위해 X선 결정 구조로부터 해당 aaRS의 아미노산 인식 부위의 잔기들(일반적으로 5~6개)을 각각 20개의 아미노산으로 변형한 라이브러리를 제작한다 [1-3]. 이 라이브러리로부터 UAA만을 선택적으로 아실화시키는 aaRS를 찾기 위해 그림 2와 같이 양성선택과 음성선택 과정을 진행한다[2]. 양성선택은 클로람페니콜 아세틸트랜스퍼라제 (CAT, chloramphenicol acetyl transferase) 유전자의 특정 위치의 앰버 코돈 억제를 통해 생성된 클로람페니콜 저항성에 의한 선택과정이고, 음성선택에서는 독성을 가지는 바네이즈 (barnase) 유전자의 앰버 코돈 억제 작용을 통해 원하는 변이를 선택한다. 이러한 양성/음성 스크리닝 과정을 거치면 특정 UAA에 선택적인 aaRS를 얻을 수 있다. UAA의 구조에 따라

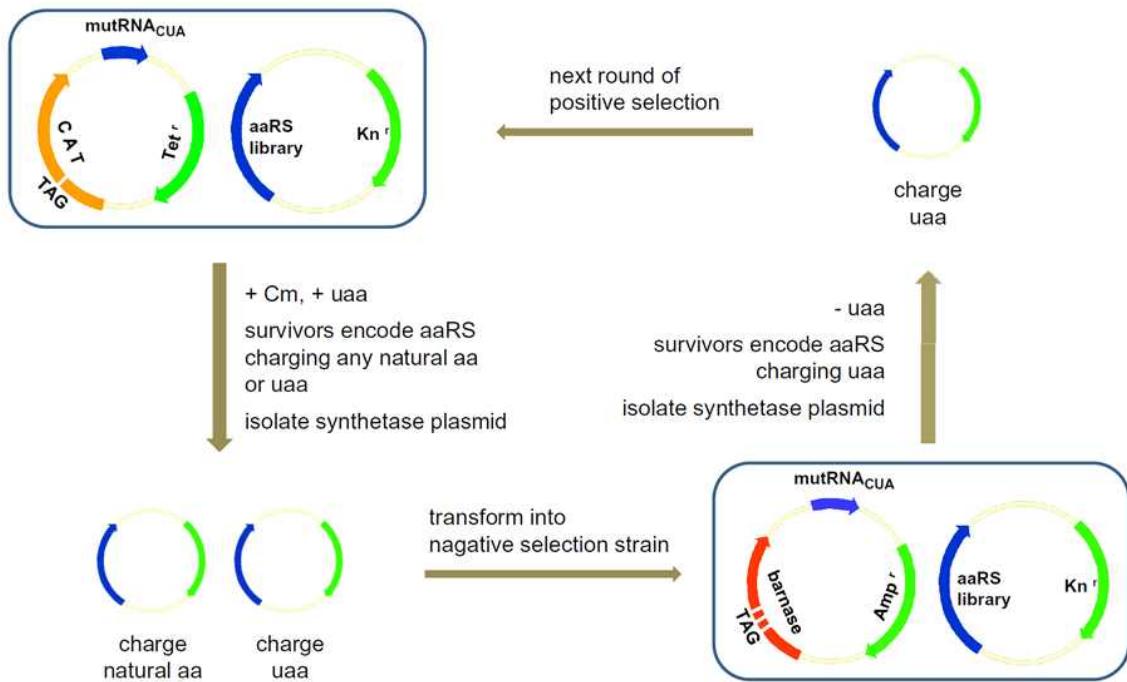


그림 2. 박테리아에서의 aaRS 기질 선택성 변형 방법. 양성 선택에서는 자연에 존재하는 20개의 아미노산 또는 UAA를 인지하는 aaRS 변이를 가지는 세포 만이 살아남고, 음성 선택에서는 자연에 존재하는 아미노산을 인지하는 aaRS 변이를 가지는 세포는 모두 죽게 된다. 따라서 양성선택과 음성선택을 거치면 UAA만을 인지하는 aaRS 변이를 선택할 수 있다.

적절한 pair를 선택하는 것이 중요하고, 구조의 특성에 따라서는 선택적인 aaRS를 얻지 못하는 경우도 많이 나타난다.

3. 비천연 아미노산의 유전적 도입법의 응용

지금까지 설명한 유전적 도입 방법으로 2017년 현재 약 150개 정도의 UAA들이 박테리아, 이스트, 그리고 포유 세포에서 성공적으로 단백질에 도입되었다. 그림 3은 대표적인 UAA들의 구조를 보여 준다. 이들 중 많은 아미노산들이 생화학적으로 유용하고 세포 내의 단백질 연구에 중요한 성질을 가지고 있다.

3-1. 단백질 표지와 변형

현재까지 단백질 내에 도입된 비천연 아미노산들 중에는 케톤 (ketone)이나 아자이드 (azide)와 같이 세포 내의 천연 아미노산들에는 없는 특별한 반응성을 가지는 (bioorthogonal) 작용기들을 포함한 것들이 있다. 이러한 UAA가 단백질에 도입이 되면, 이 UAA에 표지 물질 (biophysical probes), 변역후 변형 (posttranslational modification), 또는 단백질의 안정성이나 활성을 증가시킬 수 있는 물질 등, 다양한 작용기들을 특정 위치에 선택적으로 도입할 수 있다.

케톤기는 자연에 존재하는 아미노산에는 없는 작용기이고, 히드라자이드 (hydrazide)나 알코시아민 (alkoxyamine)과 반응하여 중성 산성도에서 안정한 히드라존 (hydrozone)이나 옥심 (oxime)을 생성한다. 따라서 *p*-acetylphenylalanine (1)과 같이 케톤기를 포함한 UAA가

단백질에 도입되면 특정 위치에 선택적으로 다양한 작용기들을 도입할 수 있기 때문에 생화학

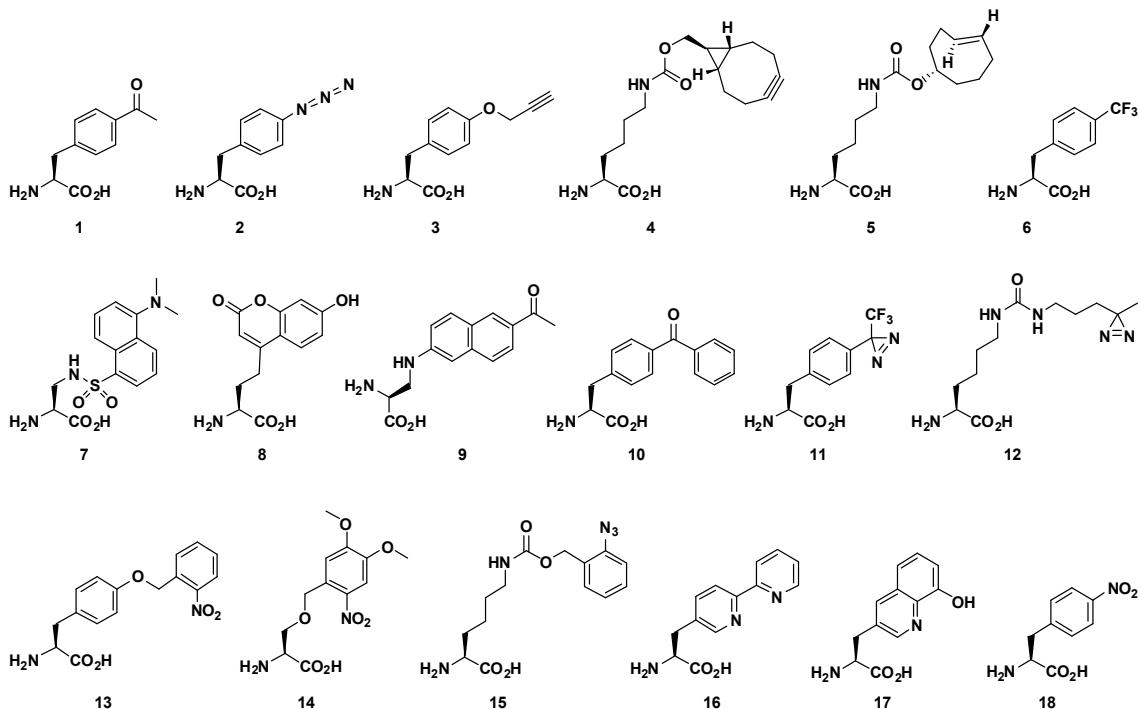


그림 3. 유전적 방법으로 도입된 대표적인 UAA의 구조.

적으로 아주 유용하게 쓰일 수 있다. 실제로 이 아미노산을 특정 단백질에 도입한 뒤, 그 단백질에 형광체 (fluorophores)[22,23], 바이오틴 (biotin)[22], 당 (sugars)[24] 등이 도입되었다.

아자이드와 아세틸렌 (acetylene)은 구리(I)를 촉매로 사용하면 [3+2] 고리첨가반응 (cycloaddition reaction)을 하는 것으로 알려져 있다 (클릭 반응, click reaction). 아자이드와 아세틸렌 작용기를 각각 포함하고 있는 *p*-azidophenylalanine (2)과 *p*-propargyloxyphenylalanine (3)은 박테리아[25]와 이스트[26]에서 성공적으로 단백질에 도입되었다. 이 아미노산들 중 하나를 포함한 단백질과 클릭반응을 할 수 있는 기질에 형광 물질[26]이나 PEG (polyethyleneglycol)[27]등을 연결하여 단백질의 특정 위치에 선택적으로 형광 물질이나 PEG을 도입하였다. 하지만 클릭반응은 구리(I)를 촉매로 사용하기 때문에 세포 내에서는 이 반응을 사용할 수가 없다. 이런 단점을 보완하기 위해서 최근에는 시클로옥타인 (cyclooctyne), 트랜스시클로옥틴(transcyclooctene)을 포함한 UAAs (4, 5)를 단백질에 도입하여 테트라진 (tetrazine)과 반응시킴으로 해서, 구리 (I)과 같이 독성이 있는 물질을 사용하지 않고 매우 효율적으로 클릭반응을 진행하였고, 이를 이용하여 세포 표면에 존재하는 단백질의 특정 위치를 표지할 수 있음이 보고되었다[28,29]. 그리고 이러한 선택적 반응을 하는 UAA를 이용하여 세포 내에서 특정 단백질의 활성을 선택적으로 조절하는 연구가 최근에 보고되었다[30].

3-2. 단백질의 구조와 기능 연구

NMR을 통한 단백질의 구조 연구를 위해 불소 (fluorine)를 포함한 아미노산 trifluoromethylphenylalanine (6)이 단백질에 도입되었다. 이 아미노산은 nitroreductase와

같은 단백질에 도입되어 ^{19}F NMR 스펙트럼 변화를 통해 기질, 효소 억제제, 코팩터 (cofactor)등의 결합과 효소반응을 연구하는데 사용되었다[31]. 또 다른 연구에서는 trifluoromethylphenylalanine과 ^{13}C 나 ^{15}N 로 표지된 α -methyltyrosine, 또는 ^{15}N 로 표지된 α -nitrobenzyltyrosine을 특정 단백질의 여러 곳에 도입하여 작은 분자와의 결합을 NMR 스펙트럼 변화로 관찰함으로써 이 분자가 단백질의 어느 부위에 결합하는지를 예측할 수 있었다 [32].

최근에는 형광을 가지는 아미노산들도 단백질에 도입되었다. Dansylalanine (7)은 human superoxide dismutase에 도입되어 형광의 변화를 통해 단백질의 구조 변화를 관찰하였다 [33]. 이 아미노산이 단백질의 표면에 도입되었을 때에는 단백질의 denaturation 후에 거의 형광변화가 없었지만, 단백질의 내부에 도입되었을 때에는 denaturation 후에 형광의 파장과 세기가 모두 크게 변화하였다. Coumarin을 포함하고 있는 아미노산 (8) 또한 비슷한 연구에 사용되어 우레아 (urea)의 농도에 따른 단백질의 부분적인 unfolding 현상을 형광변화를 통해 관찰하였다[34]. Prodan 유사체를 포함하고 있는 아미노산 (9)은 *E. coli* glutamine binding protein에 도입되어 glutamine의 결합에 따른 형광의 차이를 관찰하여, 형광을 가지는 비자연 아미노산을 통하여 작은 분자와 단백질과의 결합을 측정할 수 있었다[35]. 게다가 이 아미노산은 포유 세포에서 핵, 골지체, 소포체 등에 존재하는 단백질에 도입되어 각 단백질이 각 세포 소기관에 위치하는 것을 실시간으로 살아있는 세포의 이미징을 통해 관찰할 수 있었다 (그림 4)[36].

3-3. 단백질의 상호작용 분석과 활성의 조절

Photocrosslinking은 세포 내에서 단백질의 상호작용을 분석하는데 중요한 도구로 사용되어 왔다. *p*-Azidophenylalanine (1)[25], *p*-benzoylphenylalanine (10)[37], 그리고 *p*-(3-trifluoromethyl-3*H*-diazirin-3-yl)phenylalanine (11)[38]은 photocrosslinking 을 할 수 있는 작용기를 포함하고 있고, 유전적 방법에 의해 단백질에 도입이 가능하다. 이 중에서 *p*-benzoylphenylalanine은 다양한 단백질에 도입이 되어 그 단백질과 상호작용을 하는 기질이나 리간드 (ligand), 단백질, DNA와 photocrosslinking 을 일으켜서 세포 내에서 또는 세포 밖에서 단백질의 상호작용을 분석하는데 이용되었다[39-42]. 그리고 UAA 12는 세포 내에서 chaperone HdeA에 도입되어, 이 단백질과 상호작용하는 단백질을 찾아내는데 사용되었다[43].

Photocaged 아미노산은 단백질의 활성을 조절하는데 이용될 수 있다. 예를 들어 caspase 3의 활성 부위의 시스테인을 α -nitrobenzylcysteine으로 변형하면 caspase 3는 더 이상 활성을 가지지 않는다. 하지만 α -nitrobenzyl 부분은 자외선에 의해 쉽게 제거될 수 있기 때문에 이러한 광반응을 통해서 caspase 3의 활성을 쉽게 회복시킬 수 있었다[20]. 마찬가지로 α -nitrobenzyltyrosine (13)을 β -galactosidase의 Tyr503에 도입하면 활성이 없는 단백질을 만들 수 있었고, 여기에 자외선을 비춰주면 다시 단백질의 활성이 되살아났다[44]. 4,5-Dimethoxy-2-nitrobenzylserine (14)은 전사 요소 (transcription factor) Pho4에 도입되어 이스트 세포 내에서 빛을 통해 인산화 (phosphorylation)를 조절하는데 사용되었다[45]. 최근에는 ortho-azidobenzylloxycarbonyl-L-lysine (15)을 단백질에 도입하여 스타우딩거 환원 반응 (Staudinger reduction reaction)을 할 수 있는 저분자 물질을 이용하여 세포 내의 효소의 활성을 조절하는 연구가 보고되었다[46].

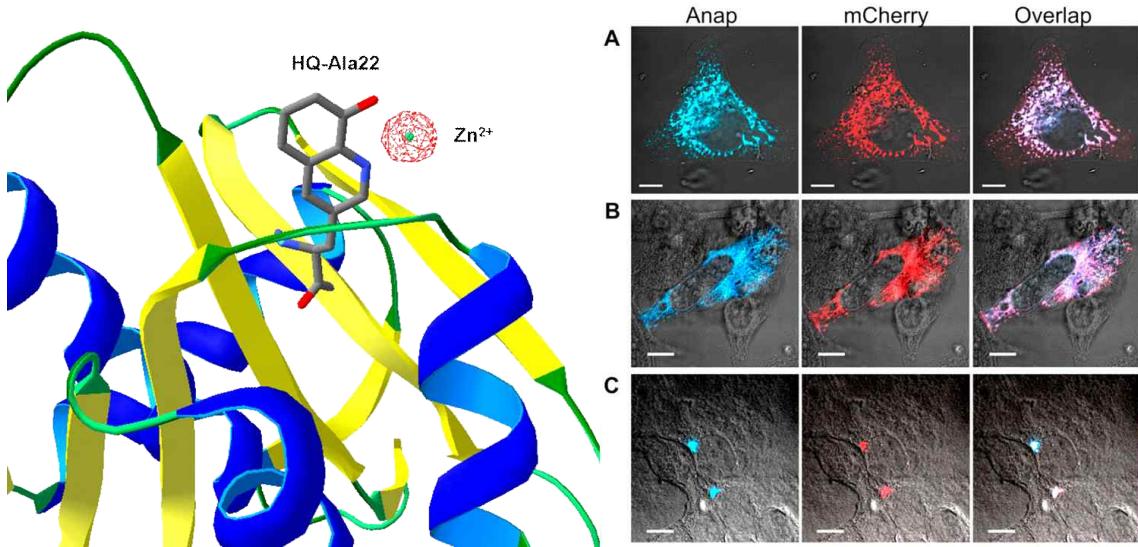


그림 4. 왼쪽, Phe22 위치에 hydroxyquinolinylalanine (HQ-Ala, 17)이 도입된 TM0665 단백질의 X-선 결정구조. TM0665 단백질은 아연(II) 이온과 함께 결정화 되었고, 아연 이온은 주변 잔기의 도움 없이 UAA 17과 안정한 결합 구조를 형성하고 있다; 오른쪽, 형광 아미노산 9를 골지체와 소포체 내에 존재하는 단백질에 도입하여 살아있는 세포 이미징을 통해서 단백질의 세포 내 위치를 형광 이미지로 분석한 결과 (참고문헌 36에서 발췌한 사진).

3-4. 새로운 기능의 단백질 생합성

단백질 내의 원하는 위치에 선택적으로 금속 결합자리를 도입할 수 있으면 금속 이온을 포함한 단백질 (metalloprotein)을 쉽게 설계/합성할 수 있을 것이다. 이러한 목적에서 금속 이온에 결합할 수 있는 작용기를 포함하고 있는 bipyridylalanine (16)[47,48]과 hydroxyquinolinylalanine (17)[49]이 대장균에서 단백질에 도입되었다. Bipyridylalanine은 *E. coli* catabolite activator protein (CAP)에 도입되었고, 이 CAP 변형체는 구리 (II)와 환원제의 존재 하에서 CAP 결합 부위를 포함한 DNA의 특정 위치만을 선택적으로 절단하였다 [48]. Hydroxyquinolinylalanine은 TM0665 단백질에 도입되어 아연 이온과 함께 결정을 만든 다음, SAD (single-wavelength anomalous dispersion) phasing을 통해서 X선 결정 구조를 밝히는데 사용되었다 (그림 7)[49]. 이 결정 구조의 금속 결합 부위에는 주위의 다른 잔기들의 도움 없이 hydroxyquinoline만으로 안정한 금속 결합 구조를 형성하고 있었다. 이 결과는 간단히 UAA의 도입만으로 쉽게 단백질 내에 금속 결합 부위를 만들어낼 수 있음을 보여준다. 최근에는 bipyridylalanine을 이용하여 단백질 내에 금속 결합 부위는 만들고, 여기에 구리 이온을 결합시켜 비대칭 Friedel-Crafts alkylation을 수행한 연구가 보고되었다[50].

p-Nitrophenylalanine (18)은 새로운 의학적 응용을 위해서 면역 반응을 촉진하는데 사용되었다. 생물체는 일반적으로 자신이 가지고 있는 단백질에는 면역 반응을 보이지 않지만 *p*-nitrophenylalanine을 항원 단백질에 도입했을 때, 면역 반응이 현저하게 높아졌다. *p*-Nitrophenylalanine은 tumor necrosis factor a (TNF-a)에 도입되었고, 이 변형 단백질을 쥐에 주사했을 때, 보조제 (adjuvant) 없이도 강한 면역반응을 보였다 (그림 8)[51].

4. 결론

2001년 처음으로 세포 내에서 유전적 방법으로 비천연 아미노산이 도입된 이후로 현재까지 150여 개의 비천연 아미노산이 다양한 세포에서 단백질에 도입되었다. 최근에는 동물 세포, 쥐의 뇌 등에서도 비천연 아미노산이 도입된 예가 보고되었다[52-54]. 이 유전적 도입 방법은 다양한 구조와 작용기를 가지는 아미노산을 다양한 단백질의 어느 위치에든지 도입할 수 있는 가장 일반적인 방법이라고 할 수 있다. 게다가 이 방법은 세포 내에서 진행되기 때문에 기술적으로 간단하면서도 높은 수율로 비천연 아미노산을 포함한 변형 단백질을 생합성할 수 있다.

유전적 방법에 의해 단백질에 도입된 새로운 아미노산의 작용기들은 이전에는 할 수 없었던 생물학적/화학적 연구들을 가능하게 할 것이다. 현재까지는 이러한 아미노산들의 응용범위가 비교적 단순한 시스템에 한정되었지만, 앞으로는 보다 도전적이고 흥미로운 실험들에 응용되어서 실제로 생물학적/화학적으로 중요한 문제들을 해결하는데 사용될 수 있을 것으로 기대된다. 이런 목적으로 형광을 가지는 아미노산, 빛에 민감한 아미노산, 생명체 내에서 선택적 반응을 수행하는 아미노산 그리고 생체 내 번역후 변형 (posttranslational modification)을 모방할 수 있는 아미노산 등이 유용하게 쓰일 것으로 예상된다.

5. 참고문헌

1. Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. Expanding the genetic code of *Escherichia coli*. *Science* **2001**, *292*, 498-500.
2. Wang, L.; Schultz, P. G. An expanded genetic code. *Angew. Chem., Int. Ed.* **2005**, *44*, 34-66.
3. Liu, C. C.; Schultz, P. G. Adding new chemistries to the genetic code. *Ann. Rev. Biochem.* **2010**, *79*, 413-444.
4. Weigert, M. G.; Garen, A. Amino acid substitutions resulting from suppression of nonsense mutations: I. Serine insertion by the Su-1 suppressor gene. *J. Mol. Biol.* **1965**, *12*, 448-455.
5. Hao, B. et al. A New UAG-Encoded Residue in the Structure of a Methanogen Methyltransferase. *Science* **2002**, *296*, 1462-1466.
6. Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* **1989**, *244*, 182-188.
7. Dougherty, D. A. Unnatural amino acids as probes of protein structure and function. *Curr. Opin. Chem. Biol.* **2000**, *4*, 645-652.
8. Engelberg-Kulka, H. UGA suppression by normal tRNATrp in *Escherichia coli*: codon context effects. *Nucleic Acids Res.* **1981**, *9*, 983-991.
9. Anderson, J. C.; Wu, N.; Santoro, S. W.; Lakshman, V.; King, D. S.; Schultz, P. G. An expanded genetic code with a functional quadruplet codon. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 7566-7571.
10. Neumann, H.; Wang, K.; Davis, L.; Garcia-Alai, M.; Chin, J. W. Encoding

multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* **2010**, *464*, 441–444.

11. Wang, K. *et al.* Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET. *Nat. Chem.* **2014**, *6*, 393–403.
12. Ibba, M.; Soll, D. Aminoacyl-tRNA synthesis. *Ann. Rev. Biochem.* **2000**, *69*, 617–650.
13. Giege, R.; Sissler, M.; Florentz, C. Survey and summary: universal rules and idiosyncratic features in tRNA identity. *Nucleic Acids Res.* **1998**, *26*, 5017–5035.
14. Marck, C.; Grosjean, H. tRNomics: analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features. *RNA* **2002**, *8*, 1189–1232.
15. Kwok, Y.; Wong, J. T. Evolutionary relationship between Halobacterium cutirubrum and eukaryotes determined by use of aminoacyl-tRNA synthetases as phylogenetic probes. *Can. J. Biochem.* **1980**, *58*, 213–218.
16. Rould, M. A.; Perona, J. J.; Soll, D.; Steitz, T. A. Structure of *E. coli* glutamyl-tRNA synthetase complexed with tRNAGln and ATP at 2.8 Å resolution. *Science* **1989**, *246*, 1135–1142.
17. Srinivasan, G.; James, C. M.; Krzycki, J. A. Pyrrolysine encoded by UAG in archaea: charging of a UAG-decoding specialized tRNA. *Science* **2002**, *296*, 1456–1462.
18. Neumann, H.; Peak-Chew, S. Y.; Chin, J. W. Genetically encoding N-acetyllysine in recombinant proteins. *Nat. Chem. Biol.* **2008**, *4*, 232–234.
19. Chen, P. R.; Groff, D.; Guo, J.; Ou, W.; Cellitti, S.; Geierstanger, B. H.; Schultz, P. G. A facile system for encoding unnatural amino acids in mammalian cells. *Angew. Chem., Int. Ed.* **2009**, *48*, 4052–4055.
20. Wu, N.; Deiters, A.; Cropp, T. A.; King, D.; Schultz, P. G. A genetically encoded photocaged amino acid. *J. Am. Chem. Soc.* **2004**, *126*, 14306–14307.
21. Wan, W.; Tharp, J. M.; Liu, W. R. Pyrrolysyl-tRNA synthetase: An ordinary enzyme but an outstanding genetic code expansion tool. *Biochim. Biophys. Acta* **2014**, *1844*, 1059–1070.
22. Wang, L.; Zhang, Z.; Brock, A.; Schultz, P. G. Addition of the keto functional group to the genetic code of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 56–61.
23. Brustad, E. M.; Lemke, E. A.; Schultz, P. G.; Deniz, A. A. A General and efficient method for the site-specific dual-labeling of proteins for single molecule fluorescence resonance energy transfer. *J. Am. Chem. Soc.* **2008**, *130*, 17664–17665.
24. Liu, H.; Wang, L.; Brock, A.; Wong, C. H.; Schultz, P. G. A method for the

- generation of glycoprotein mimetics. *J. Am. Chem. Soc.* **2003**, *125*, 1702–1703.
- 25. Chin, J. et al. Addition of p-azido-L-phenylalanine to the genetic code of *Escherichia coli*. *J. Am. Chem. Soc.* **2002**, *124*, 9023–9027.
 - 26. Deiters, A.; Cropp, T. A.; Mukherji, M.; Chin, J. W.; Anderson, J. C.; Schultz, P. G. Adding amino acids with novel reactivity to the genetic code of *Saccharomyces cerevisiae*. *J. Am. Chem. Soc.* **2003**, *125*, 11782–11783.
 - 27. Deiters, A.; Cropp, T. A.; Summerer, D.; Mukherji, M.; Schultz, P. G. Site specific PEGylation of proteins containing unnatural amino acids. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5743–5745.
 - 28. Lang, K. et al. Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction. *Nat. Chem.* **2012**, *4*, 298–304.
 - 29. Lang, K. et al. Genetic encoding of bicyclononynes and trans-cyclooctenes for site-specific protein labeling in vitro and in live mammalian cells via rapid fluorogenic Diels–Alder reactions. *J. Am. Chem. Soc.* **2012**, *134*, 10317–10320.
 - 30. Tsai, Y-H. et al. Selective, rapid and optically switchable regulation of protein function in live mammalian cells. *Nat. Chem.* **2015**, *7*, 554–561.
 - 31. Jackson, J. C.; Hammill, J. T.; Mehl, R. A. Site-Specific incorporation of a ¹⁹F-amino acid into proteins as an NMR probe for characterizing protein structure and reactivity. *J. Am. Chem. Soc.* **2007**, *129*, 1160–1166.
 - 32. Cellitti, S. E.; Jones, D. H.; Lagpacan, L.; Hao, X.; Zhang, Q.; Hu, H.; Brittain, S. M.; Brinker, A.; Caldwell, J.; Bursulaya, B.; Spraggan, G.; Brock, A.; Ryu, Y.; Uno, T.; Schultz, P. G.; Geierstanger, B. H. In vivo incorporation of unnatural amino acids to probe structure, dynamics, and ligand binding in a large protein by nuclear magnetic resonance spectroscopy. **2008**, *J. Am. Chem. Soc.* *130*, 9268–9281.
 - 33. Summerer, D.; Chen, S.; Wu, N.; Deiters, A.; Chin, J. W.; Schultz, P. G. A genetically encoded fluorescent amino acid. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 9785–9789.
 - 34. Wang, J.; Xie, J.; Schultz, P. G. A Genetically encoded fluorescent amino acid. *J. Am. Chem. Soc.* **2006**, *128*, 8738–8739.
 - 35. Lee, H. S.; Guo, J.; Lemke, E. A.; Dimla, R. D.; Schultz, P. G. Genetic incorporation of a small, environmentally sensitive fluorescent probe into proteins in *Saccharomyces cerevisiae*. *J. Am. Chem. Soc.* **2009**, *131*, 12921–12923.
 - 36. Chatterjee, A. et al. A genetically encoded fluorescent probe in mammalian cells. *J. Am. Chem. Soc.* **2013**, *135*, 12540–12543.
 - 37. Chin, J. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11020–11024.

38. Tippmann, E. M.; Liu, W.; Summerer, D.; Mack, A. V.; Schultz, P. G. A genetically encoded diazirine photocrosslinker in *Escherichia coli*. *ChemBioChem* 2007, 8, 2210–2214.
39. Schlieker, C.; Weibeahn, J.; Patzelt, H.; Tessarz, P.; Strub, C.; Zeth, K.; Erbse, A.; Schneider-Mergener, J.; Chin, J. W.; Schultz, P. G.; Bukau, B.; Mogk, A. Substrate recognition by the AAA+ chaperone ClpB. *Nat. Struct. Mol. Biol.* 2004, 11, 607–615.
40. Hino, N.; Okazaki, Y.; Kobayashi, T.; Hayashi, A.; Sakamoto, S.; Yokoyama, S. Protein photo-cross-linking in mammalian cells by site-specific incorporation of a photoreactive amino acid. *Nat. Methods* 2005, 2, 201–206.
41. Huang, L.-Y.; Umanah, G.; Hauser, M.; Son, C.; Arshava, B.; Naider, F.; Becker, J. M. Unnatural amino acid replacement in a yeast G protein-coupled receptor in its native environment. *Biochemistry* 2008, 47, 5638–5648.
42. Lee, H. S.; Dimla, R. D.; Schultz, P. G. Protein-DNA photo-crosslinking with a genetically encoded benzophenone-containing amino acid. *Bioorg. Med. Chem. Lett.* 2009, 19, 5222–5224.
43. Zhang, M. et al. A genetically incorporated crosslinker reveals chaperone cooperation in acid resistance. *Nat. Chem. Biol.* 2011, 7, 671–677.
44. Deiters, A.; Groff, D.; Ryu, Y.; Xie, J.; Schultz, P. G. A genetically encoded photocaged tyrosine. *Angew. Chem., Int. Ed.* 2006, 45, 2728–2731.
45. Lemke, E. A.; Summerer, D.; Geierstanger, B. H.; Brittain, S. M.; Schultz, P. G. Control of protein phosphorylation with a genetically encoded photocaged amino acid. *Nat. Chem. Biol.* 2007, 3, 769–772.
46. Luo, J. et al. Small-molecule control of protein function through Staudinger reduction. *Nat. Chem.* 2016, 8, 1027–1034.
47. Xie, J.; Liu, W.; Schulz, P. G. A genetically encoded bidentate, metal-binding amino acid. *Angew. Chem., Int. Ed.* 2007, 46, 9239–9242.
48. Lee, H. S.; Schultz, P. G. Biosynthesis of a site-specific DNA cleaving protein. *J. Am. Chem. Soc.* 2008, 130, 13194–13195.
49. Lee, H. S.; Spraggon, G.; Schultz, P. G.; Wang, F. Genetic incorporation of a metal-ion chelating amino acid into proteins as a biophysical probe. *J. Am. Chem. Soc.* 2009, 131, 2481–2483.
50. Drienovská, I. et al. Novel artificial incorporation of metalloenzymes by in vivo metal-binding unnatural amino acids. *Chem. Sci.* 2015, 6, 770–776.
51. Grunewald, J.; Tsao, M. L.; Perera, R.; Dong, L.; Niessen, F.; Wen, B. G.; Kubitz, D. M.; Smider, V. V.; Ruf, W.; Nasoff, M.; Lerner, R. A.; Schultz, P. G. Immunochemical termination of self-tolerance. *Proc. Natl. Acad. Sci. U.S.A.* 2008, 105, 11276–11280.
52. Greiss, S.; Chin, J. W. Expanding the genetic code of an animal. *J. Am. Chem. Soc.* 2011, 133, 14196–14199

53. Bianco, A.; Townsley, F. M.; Greiss, S.; Lang, K.; Chin, J. W. Expanding the genetic code of Drosophila Melanogaster. *Nat. Chem. Biol.* **2016**, *12*, 776–778.
54. Ernst, R. J. et al. Genetic code expansion in the mouse brain. *Nat. Chem. Biol.* **2016**, *12*, 776–778.