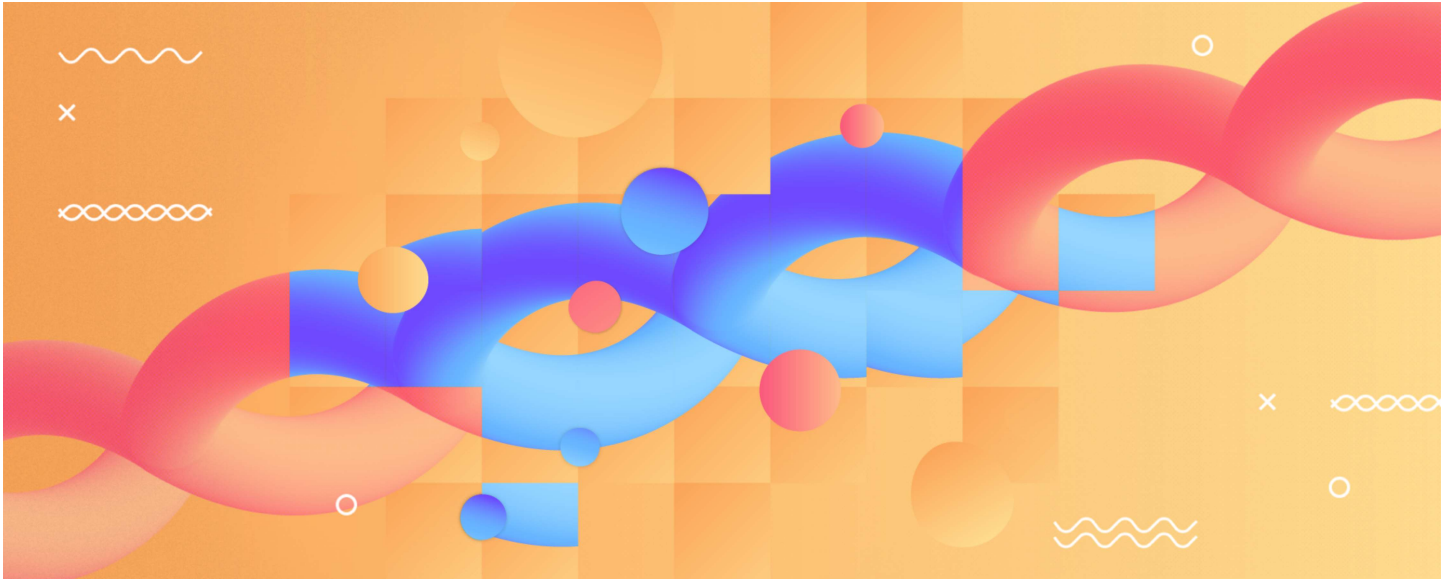


확률적 과정으로 조절되는 생명현상?

2019년 4월 12일

이종봉



액정의 위상학적 결함^{topological defect} 연구로 물리학 박사학위를 취득한 필자는 완전히 생소한 분야인 생물물리 연구실에서 박사후 연구원^{PostDoc} 과정을 시작했다. 새로운 셋업에서 몇 번의 실험을 실패한 후, 난생 처음 생명현상의 과정을 실시간으로 관찰했을 때의 신비로움을 잊을 수 없다.

당시 필자는, 이중나선 DNA가 헬리케이스^{helicase} 효소에 의해 분리되고 DNA 중합^{polymerase} 효소에 의해 다시 합성되는 DNA의 복제 과정을, 바이러스의 일종인 박테리오파지^{bacteriophage} 단백질을 통해 연구하고 있었다. 이 박테리오파지 DNA 복제는 가장 단순한 생물학적 과정이지만, 이조차도 4가지 다른 종류의 단백질들이 적재적소에서 복합체를 이뤄내고, 각각의 임무를 빈틈없이 수행해야만 가능한 일이다. 박테리오파지보다 복잡한 구조를 지닌 대장균의 경우에는, DNA복제를 담당하는 복합체만 해도 최소 16가지 다른 단백질로 구성되어 있다. 개별 생체분자의 활동성은 이들의 상호작용을 통해 수백에서 수천 배로 증가한다. 이러한 프로세스의 복잡성과 상호작용은 고등 생물로 옮겨갈수록 기하급수적으로 증폭된다.

인간의 세포에서 DNA의 정보는 약 30억 개의 염기쌍에 기록되어 존재하며, A^{아데닌}-T^{티민}, G^{구아닌}-C^{시토신}의 수소결합으로 이루어진 염기쌍은 DNA의 이중 나선 구조 속에 안전하게 보관된다. 그러나 세포분열시 DNA 복제를 수행하는 DNA 중합 효소는, A-T와 G-C가 염기쌍을 이루도록 합성하는 과정에서 대략 1만 개의 DNA 중 1개의 염기쌍 오류를 발생시킨다. 다행히도 이 DNA 중합 효소는 이렇게 잘못 만들어진 오류 염기쌍(G-T, A-C 등)을 스스로 찾고 고칠 수 있는 능력을 갖고 있어서, 천만 개를 합성하면 한 개 꼴로 오류를 축소할 수 있다. 그러나 총 염기쌍 수가 30억 개인 점

을 고려하면 여전히 수백 개의 오류가 예상이 된다. 여기서 놀라운 반전은 모든 생명체가 자체적으로 DNA 복제 오류를 한 번 더 여과해서 유전정보의 오류를 현격히 줄여낼 수 있는 “DNA 염기쌍 오류 복구 시스템”을 갖추고 있다는 사실이다.

필자의 연구실은 오류 복구 시스템의 작동 원리를 밝히고자, 세포 밖에서 정제된 복구 단백질의 움직임을 DNA 상에서 실시간으로 관찰하고 있다. 이러한 단일분자-생물물리 연구로 이해하게 된 DNA 염기쌍 오류 복구의 메커니즘 모델이 [그림1]이다. 본 글에서는 이에 대한 부연설명을 하고자 한다.

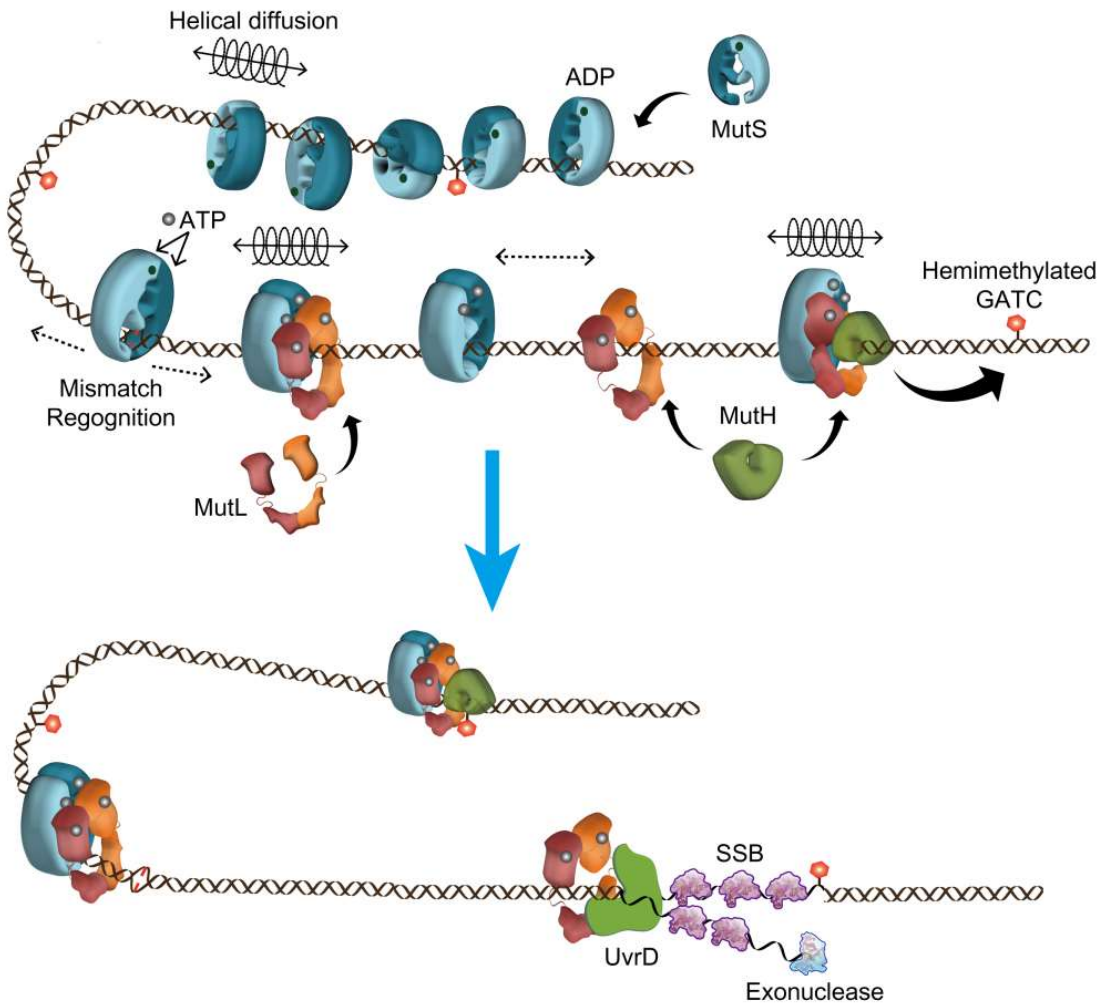


그림1. 대장균 *E. coli*의 유전정보 오류 복구 메커니즘: (위) MutS 단백질의 회전 확산 운동으로 오류 염기쌍 탐색 및 인식 → ATP-MutS의 형성 및 확산 운동 → ATP-MutS에 의한 ATP-MutL 형성 → ATP-MutS/ATP-MutL/MutH의 회전 확산 운동 → MutH에 의한 hemimethylated GATC 인식 및 DNA 가닥 절단 → (아래) ATP-MutL/UvrD 복합체 형성 → 이중나선 DNA 풀림 → DNA 단편 제거 또는 핵산 분해효소 exonuclease에 의한 DNA 제거 (→ 오류 염기쌍 제거 → DNA 중합 효소에 의한 DNA 재합성)

우선 DNA 염기쌍 오류 복구의 시작 단백질인 MutS는, DNA의 이중나선 구조를 따라 확산 운동을 하며 오류 염기쌍을 찾아내 신호를 보내는 센서 역할을 한다. 오류 염기쌍에 결합한 MutS는 ATP와의 결합을 통해 극적인 구조의 변화를 일으킨다. 이렇게 결합된 ATP-MutS는 오류 염기쌍의 존재를 알리기 위해 신호 전달의 매개체 역할을 하는 MutL 단백질을 불러들여 복합체를 형성한다. MutS-MutL 복합체는 DNA상에서 결합과 분리를 반복하며 확산 운동을 통해

이동한다. 이때 특이할 사항은, MutS 단백질이 MutL과 복합체를 형성하는 경우에는 DNA 나선을 따라 느리게 움직이고, 분리된 상태에서는 매우 빠르게 무작위로 운동한다는 것이다. 또한 이 단백질들의 동력은 결합된 ATP의 가수분해 에너지를 사용하지 않고, 오직 DNA상에서 생체분자의 확산 운동만으로 이루어진다.

박테리아나 인간 세포에서 오류 염기쌍의 제거를 시작하기 위한 단계까지의 분자메카니즘은 동일하나, 막상 오류 염기쌍의 제거 단계에서는 양자의 메커니즘이 매우 상이하다. 박테리아의 일종인 대장균^{E. coli}에서는, 위에서 설명한 확산 운동을 하며 이동하는 MutL에 또 다른 효소인 MutH가 결합하면서 오류 염기쌍 신호 전달과 오류 염기쌍 제거의 시작 메커니즘이 완성된다. MutH는 DNA 두 가닥 중 한쪽 가닥에만 메틸기(CH₃)가 붙어있는 GATC 염기서열을 인식하는 특성을 가지고 있다. 메틸기가 붙어있지 않은 DNA 가닥이 새로 합성된 DNA로 오류 염기쌍을 제거하기 위해서는, 이 가닥에 있는 유전정보를 수정해야 한다. MutL에 결합한 MutH는, 오류 염기쌍으로부터 평균 230여 개의 염기쌍마다 존재하는 GATC 염기서열 인식과 DNA 가닥의 무작위적인 절단을 통해 오류 염기쌍 제거를 촉발하는 역할을 한다. MutS-MutL-MutH 복합체가 DNA 나선 구조를 따른 확산 운동에 의해 GATC 염기서열을 찾아내는 것이다.

MutL 단백질이 UvrD 헬리케이스와 함께 두 가닥의 DNA를 풀어서, MutH에 의해 잘려진 DNA 조각을 없애면 마침내 오류 염기쌍이 제거된다. 흥미로운 사실은 오류 염기쌍에 의해 형성된 ATP-MutS 단백질이 DNA 상에 존재하는 한, MutH에 의해 잘려진 DNA 조각들이 지속적으로 제거된다는 것이다. 그러므로 오류 염기쌍이 사라진 후, DNA 상에 있는 ATP-MutS가 모두 소진되어야 더 이상 오류 염기쌍 신호를 보낼 수 없게 되어 비로소 오류 복구 시스템의 작동이 멈추게 된다. 기본적으로 인간 세포를 포함한 진핵세포에서 오류 염기쌍을 제거하는 메커니즘은 박테리아와 다르지만, MutS에 의해 오류 염기쌍을 인식하고 복구하는 과정이 개개의 단백질 또는 복합체의 1차원 확산 운동과 확률적인 단백질의 상호작용으로 이루어진다는 점은, 모든 종류의 세포에서 공통적으로 일어나는 현상이다.

//

필자는 현재까지 축적된 선행 연구와 앞서 소개한 유전정보 오류의 복구 과정을 근거로,

복구 단백질의 작용 원리를 '단계적 cascading'이고 '확률적 stochastic'인 상호작용의 결과로 정의하고자 한다.

//

필자의 연구실에서 연구하는 PCNA^{Proliferating cell nuclear antigen} 단백질은, 생체 분자 간 확률적 상호작용을 통해 생명 현상을 조절하는 또 다른 예이다. PCNA는 인간 세포를 포함한 진핵 세포에서 DNA 중합 효소와 결합하여, DNA 합성을 증가시키는 DNA 복제의 핵심 단백질이다. 이것은 복제 외에도 DNA 복구, 전사, 세포주기조절 등 20여 가지의 생명현상에 직접적으로 관여한다고 알려져 있다. 이러한 생명현상을 수행하기 위해 30여 가지의 단백질들이 PIP^{PCNA}

interacting protein-box 시퀀스를 통해 PCNA 세 곳의 특정 영역과 직접 접촉하게 된다. 그러나 PCNA에 결합할 수 있는 영역이 극히 제한적이기 때문에, 이들 단백질들은 필연적으로 서로 경쟁 관계에 놓이게 되어, PCNA에 먼저 결합한 단백질이 다른 단백질의 상호작용을 방해하는 결과가 나타날 수 있다. 그러나 놀랍게도 세포 내에서는 아무 문제 없이 PCNA와 다른 단백질 간의 상호작용을 통해 다양한 생명현상이 진행되고 있다. 세포에서의 생명현상은 생체분자 간의 친화성^{affinity} 정도와, 농도에 따른 확률적 상호작용에 의해 모든 것이 순조롭게 진행될 것으로 기대되지만, 그 조절은 매우 정교하고 섬세하게 이루어져야만 한다.

물리학이 현상을 단순화하여 원리를 찾는 것과 달리, 생명현상은 생체분자 간의 상호작용이 매우 다양하고 복잡하여 모든 현상을 명료하게 정리하기가 어렵다. 그리하여 필자는 실시간 관측을 통해서 현재까지 축적된 선행 연구와 앞서 소개한 유전정보 오류의 복구 과정을 근거로, 복구 단백질의 작용 원리를 “단계적^{cascading}”이고 “확률적^{stochastic}”인 상호작용의 결과로 간명하게 정의하고자 한다. “단계적”이란 표현은 오류의 인식에서부터 제거까지 각 단계마다 복구 단백질들의 물리적 접촉으로 신호가 전달된다는 것을 의미하고, “확률적”이란 표현은 복구 단백질이 열적 요동^{thermal fluctuation}에 의한 확률적인 상호작용으로 이루어진다는 것을 의미한다. 역설적이지만 PCNA에서 살펴본 바와 같이 확률적 상호작용은 매우 정교한 조절이 필요하다. DNA의 물질대사는 10^{-13} L의 부피를 갖는 세포핵에서 이루어지고, 세포핵에는 다양한 종류의 단백질, RNA 등과 함께 30억개의 DNA 게놈^{genome}이 23쌍의 염색체로 나뉘어져 채워져 있다. 이토록 복잡한 구조와 다양한 생체분자들이 존재하는 세포 내에서 일어나는 생명현상이, 단지 확률적인 상호작용을 통해 정확한 시점과 장소에서 이루어지고 있다는 것은 참으로 경이로운 일이다.

후기 생명현상은 흔히 잘 짜여진 오케스트라의 연주에 비유되곤 한다. 이 오케스트라에 모든 악기를 아우르며 조율하는 지휘자가 존재하지 않는다면, 그 조화로운 소리는 기대할 수 없다. 하물며 이토록 효율적이고 정교한 생명현상을 단지 확률적인 상호작용에만 기대어 설명하고 설득할 수 있을까? 마치 오케스트라의 모든 권한과 지휘는 막상 단 하나의 악기도 들지 않은 지휘자에게 속한 것처럼 혹시 우리의 생명현상에는 아직 우리가 볼 수도 없고 발견되지도 않은 미지의 물질이나 에너지가 관여되어 있는 것은 아닐까?