



# 정규모드 기법을 이용한 단백질의 동적 특성 규명

엄길호\*  
(성균관대학교)

## 1. 머리말

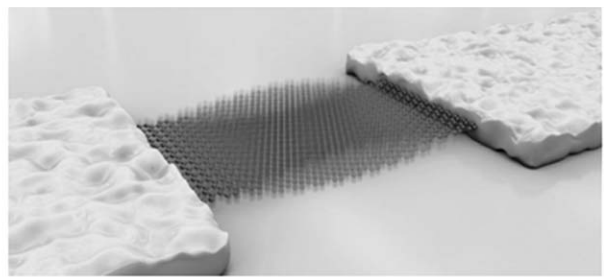
1950년대 들어서 진동특성 해석은 구조물의 거동 해석에 매우 중요한 역할을 하고 있다. 예를 들면, 1940년에 타코마 다리는 큰 하중에도 견딜 수 있도록 설계되었음에도 불구하고 바람에 의한 동적 하중에 의해 붕괴가 되었다(그림 1(a), Tacoma Narrows Bridge”, Wikipedia(en.wikipedia.org/wiki/Tacoma\_Narrows\_Bridge)). 이와 같이, 동적 하중에 의한 구조물의 진동이 구조물의 거동을 좌우하는 것을 알 수 있다. 1950년대 이후부터 구조물의 진동특성을 규명/해석하는 것이 매우 중요함이 인식되기 시작하여 지금까지 진동특성 해석이 구조물 설계에 있어서 매우 중요한 요소가 되고 있다. 진동특성 해석 기법은 자동차, 비

행기, 선박 등 다양한 분야에서 사용되는 구조물 설계에 활용되고 있다.

앞에서 언급한 바와 같이 진동특성 해석 기법은 공학적 구조물 설계에 매우 중요한 역할을 하고 있다. 1993년에 처음 탄소나노튜브가 발견됨에 따라(S. Iijima, T. Ichihashi, Nature 363, pp. 603~605 (1993)) 나노구조물을 이용한 나노소자 설계 기술이 21세기 들어서 각광을 받고 있다. 특히, 나노와이어(nanowire), 그래핀(graphene) 등 다양한 나노소재의 발견에 따라 나노소재 기반의 기계적/전기적 거동을 수행할 수 있는 소자를 개발하기에 이르렀다(K. Eom, et. al., Phys. Rep. 503, pp. 116~163 (2011)). 이와 같이 나노구조물의 기계적 거동 해석에 있어서 진동이론은 매우 중요한 역할을 하고 있다. 예를 들면, 나노와이어 및



(a)



(b)

그림 1 (a) 1940년도에 바람에 의한 동적 하중으로 인한 타코마 다리의 붕괴직전의 타코마 다리의 변형이 일어난 사진, (b) 그래핀(graphene)에 전기적 힘(electrostatic force)이 가해짐에 그래핀이 진동되는 모식도(J.S. Bunch, Nat. Nanotechnol. 6, pp. 331~332 (2011)).

\* E-mail : kilhoeom@skku.edu / Tel : (031)299-6913

그래핀 등 다양한 나노소재를 이용하여 고주파 동적 특성을 지닌 나노 레조네이터를 개발하여 주파수 필터링, 센서 및 액추에이터 등 다양한 역할을 수행할 수 있는 소자로 소개되고 있다(그림 1(b), K. Eom, et. al., Phys. Rep. 503, pp. 116~163(2011)). 이와 같이 나노 레조네이터의 고주파 동적 특성을 규명하는데 진동 이론은 매우 중요한 역할을 하고 있다. 진동 이론은 나노소재 기반의 나노소자의 기계적 거동을 해석하는데 이용되는 것 뿐만 아니라, 단백질과 같은 생체 분자의 동적 거동을 해석하는데도 유용하다.

특히, 단백질과 같은 생체 분자의 거동을 규명하는데 있어서 물 안에서 생체 분자의 역학적 거동이 매우 중요하다. 특히, 인체 내에서 존재하는 생체 분자가 거동을 할 때 물 안에서 동적 거동을 한다. 생체 분자는 크기가 나노미터에서 마이크로미터 정도이다. 생체 분자가 이와 같은 크기이기 때문에 물 분자의 브라운 운동(Brownian motion)이 생체분자의 동적 특성을 유발한다. 구체적으로 브라운 운동을 하는 물 분자가 생체분자와 충돌하면서 생체 분자의 미세한 진동을 유발하여 생체 분자가 생물학적 기능을 수행할 수 있도록 동적 거동을 할 수 있다. 따라서 2000년대 이후부터 단백질과 같은 생체 분자의 동적 특성 해석을 위해 기존의 진동 이론과 물리학 분야의 통계역학 이론이 각광을 받기 시작했다. 이 글에서는 단백질과 같은 생체 분자의 동적 특성을 해석하는데 유용한 진동 이론에 대해 논하려고 한다. 이 논문에서는 생체 분자의 거동에만 국한하지만, 이 논문에서 소개된 이론 및 전산모사 기법은 나노소재 기반의 나노소자 설계 등 다양한 나노/바이오 분야에서의 동적 거동을 해석하는데 활용될 수 있다.

## 2. 이론적 배경: 탄성 네트워크 모델

생체분자의 동적 거동 해석을 위한 진동 및 통계역학 이론에 대해 소개하고자 한다. 우선 독자들의 이해를 돕고자 1차원 스프링 모델을 고려한

다. 구체적으로 생체분자가 탄성 특성을 가지기 때문에 생체분자를 단순히 1차원 스프링으로 대체한다. 이 때, 1차원 스프링의 강성도를  $\gamma$ 라고 하고, 스프링의 변위를  $u$ 라고 한다. 생체 분자가 물 분자에 의해 분자구조가 변형될 때의 변형에너지는 식 (1)과 같다.

$$V(u) = \frac{1}{2} \gamma u^2 \quad (1)$$

물 분자에 의해 생체 분자가 브라운 운동을 하기 때문에, 생체 분자의 변형에 대한 변위는 볼츠만 확률분포(Boltzmann probability distribution)를 따른다(식 (2) 참고).

$$p(u) = \frac{1}{Z} \exp\left(-\frac{V(u)}{k_B T}\right) = \frac{1}{Z} \exp\left(-\frac{\gamma u^2}{2k_B T}\right) \quad (2a)$$

여기서,  $k_B$ 와  $T$ 는 각각 볼츠만 상수, 절대온도를 의미하며,  $Z$ 는 분배함수로 식 (2b)와 같이 정의된다.

$$Z = \int_{-\infty}^{\infty} \exp\left(-\frac{V(u)}{k_B T}\right) du = \int_{-\infty}^{\infty} \exp\left(-\frac{\gamma u^2}{2k_B T}\right) du \quad (2b)$$

따라서, 물 분자에 의한 생체분자가 브라운 운동을 할 때, 생체 분자의 변위의 평균값과 변위 제곱의 평균값은 다음과 같다.

$$\langle u \rangle = \int_{-\infty}^{\infty} u \cdot p(u) du = 0 \quad (3a)$$

$$\langle u^2 \rangle = \int_{-\infty}^{\infty} u^2 \cdot p(u) du = \frac{k_B T}{\gamma} \quad (3b)$$

식 (3)에서 보여준 바와 같이, 변위의 평균값은 0임을 알 수 있다. 이는 직관적으로 생각해보면 생체분자가 브라운 운동에 의해 미세하게 진동을 해도 거시적으로 오랜 시간동안 볼 때 생체분자의 변위의 평균은 0임을 알 수 있다. 반면에, 식 (3b)에서 보여준 바와 같이 생체분자의 변위의



그림 2 Lactoferrin 단백질(왼쪽)을 탄성 네트워크(오른쪽)으로 모델링하는 모식도. 단백질 구조를 단순히 3차원 스프링 네트워크로 묘사함(J.I. Jeong, Y. Jang, M.K. Kim, J. Mol. Graph. Model. 24, pp. 296~306 (2006))

제곱에 대한 평균값은 생체분자의 강성도에 반 비례하다는 것을 알 수 있다. 따라서 실제 실험에서 생체 분자의 브라운 운동을 관찰하여 변위 제곱에 대한 평균값을 구한다면, 생체 분자의 강성도를 계산할 수 있다.

일반적으로 생체 분자의 브라운 운동을 해석하는데 1차원 스프링 모델로는 부족하다. 특히, X-ray crystallography 및 NMR(nuclear magnetic resonance) 등을 이용한 실험기법은 생체 분자의 브라운 운동을 측정하여 생체 분자를 구성하는  $N$ 개의 원자의 변위에 대한 정보를 제시한다. 따라서 생체 분자를 단순히 1차원 스프링 모델로 해석하는 것보다  $N$ 개의 원자가 상호작용을 하는  $N$ 개의 노드를 스프링으로 연결한 스프링 네트워크로 해석하는 것이 적당하다(그림 2). 이와 같은 모델링 기법은 1996년 Tirion이 처음으로 제시하여 탄성 네트워크 모델이라 명명되고 있으며(M.M. Tirion, Phys. Rev. Lett. 77, pp. 1905~1908 (1996)), 현재까지 많은 생체 분자의 동적 거동을 해석하는데 유용하게 활용되고 있다(K. Eom, et al., J. Comput. Theor. Nanosci. 7, pp. 1210~1226

(2010)). 생체 단백질을  $N$ 개의 노드로 구성된 탄성 네트워크로 모델링 할 때, 생체 단백질 구조가 변형될 때 변형에너지는 식 (4)와 같다.

$$V = \frac{\gamma}{2} \sum_{i,j}^N [d_{ij} - d_{ij}^0]^2 \cdot H(r_c - d_{ij}^0) \equiv \frac{1}{2} u^T K u \quad (4)$$

여기서  $N$ 은 생체 단백질의 총 아미노산 개수를 의미하며,  $d_{ij}$ 는  $i$ 번째 아미노산과  $j$ 번째 아미노산과의 거리를 의미하며, 위첨자 0은 평형상태를 나타내며,  $H(z)$ 는 Heaviside unit step function을 의미하며,  $r_c$ 는 cut-off 거리를 나타낸다. 또한,  $u$ 는 아미노산의 변위에 대한 열벡터를 의미하며,  $K$ 는 탄성 네트워크의 강성도 행렬을 나타낸다. 식 (4)에서 제시된 바와 같이 탄성 네트워크는 다음과 같이 구성한다.

- (1) 생체 단백질을 구성하는  $N$ 개의 아미노산을  $N$ 개의 노드로 치환한다.
- (2) 임의로 2개의 노드를 선택하여, 각각  $i$ 번째 노드와  $j$ 번째 노드로 명명한다.
- (3) 위의 (2)에서 선택된 2개의 노드 사이의 거

리가 cut-off 거리보다 작으면 강성도가  $\gamma$ 인 스프링으로 연결한다. 만약 2개의 노드 사이의 거리가 cut-off 거리보다 크면 스프링으로 노드를 연결하지 않는다.

(4) 위의 (2)와 (3)의 과정을 생체 단백질을 구성하는 모든 노드에 대해 반복하여 최종적으로 스프링 네트워크를 얻는다.

(5) 위의 (4)에서 얻어진 스프링 네트워크를 통해, 탄성 네트워크의 강성도 행렬을 구한다.

위에서 언급된 바와 같이 생체 단백질의 강성도 행렬을 계산할 수 있다면, 위의 1차원 스프링 모델에서 언급된 통계역학 이론을 이용하여 생체 단백질의 mean-square fluctuation을 구하여 실제 X-ray crystallography, NMR 실험 및 분자동역학과 같은 원자모델 기반 전산모사 결과와 비교하여 탄성 네트워크 모델링 기법을 검증할 수 있다. 탄성 네트워크 모델에서 식 (4)에서 주어진 각 아미노산의 변위는 서로 커플링이 되어 있기 때문에, 곧바로 볼츠만 확률분포 이론을 변위 벡터에 적용할 수 없다. 각 아미노산의 변위들이 커플링된 것을 de-couple시키기 위해 정규모드 기법을 이용하여 정규모드 공간에서 변위 벡터를 고려한다. 정규모드 기법을 이용하여 식 (4)에서 언급된 변형 에너지는 식 (5)와 같이 표현될 수 있다.

$$V = \frac{1}{2} u^T K u \equiv \frac{1}{2} u^T (Q^T \Lambda Q) u \quad (5)$$

여기서  $\Lambda$ 는 대각 행렬로서 강성 행렬의 고유값으로 구성되어 있으며,  $Q$ 는 모달 매트릭스로 각 행은 강성 행렬의 정규모드 벡터를 나타낸다. 결국 행렬  $\Lambda$ 와  $Q$ 는 탄성 네트워크의 강성 행렬에 대한 정규모드 기법을 통해 계산할 수 있다. 정규모드 공간에서 변위 벡터를 아래 식 (6)과 같이 정의한다.

$$w = Qu \quad (6)$$

여기서  $w$ 는 정규모드에 의해 변환된 변위 벡터

를 의미한다. 따라서 식 (5)에서 제시된 변형에너지는 식 (7)과 같이 나타낼 수 있다.

$$V = \frac{1}{2} w^T \Lambda w = \frac{1}{2} \sum_{i=1}^{3N-6} \lambda_i w_i^2 \quad (7)$$

여기서,  $\lambda$ 는 강성 행렬  $K$ 의 고유값이며, 아래 첨자  $i$ 는 모드 인덱스를 의미한다. 식 (7)에서 강체운동에 해당되는 zero-mode는 합산에서 제외한다. 통계역학에서 균등 분할 이론(equi-partition theorem)을 통해 정규모드 공간에서의 변위의 확률 분포는 볼츠만 확률 분포를 따른다(식 (8a)).

$$p(w_i) = \frac{1}{Z} \exp\left(-\frac{V(w_i)}{k_B T}\right) = \frac{1}{Z_i} \exp\left(-\frac{\lambda_i w_i^2}{2k_B T}\right) \quad (8a)$$

여기서,  $Z_i$ 는  $i$ 번째 모드에 해당되는 분할함수로 식 (8b)와 같다.

$$Z_i = \int_{-\infty}^{\infty} \exp\left(-\frac{\lambda_i w_i^2}{2k_B T}\right) dw_i \quad (8b)$$

식 (8a)의 확률 분포를 통해 정규모드 공간에서 표현된 변위벡터는 식 (9a)의 특성을 지닌다.

$$\langle w_i \rangle = 0 \text{ and } \langle w_i w_j \rangle = k_B T \lambda^{-1} \delta_{ij} \quad (9a)$$

여기서  $\delta_{ij}$ 는 Kronecker delta를 의미한다. 식 (9)를 벡터 표현으로 나타내면 식 (9b)와 같다.

$$\langle ww^T \rangle = k_B T \Lambda^{-1} \quad (9b)$$

식 (9)의  $\langle w_i w_j \rangle$ 를 모달 매트릭스  $Q$ 를 이용하여 변환하면, 변동 행렬은 식 (10)과 같다.

$$F \equiv \langle uu^T \rangle = Q^T \langle ww^T \rangle Q = k_B T Q^T \Lambda^{-1} Q \quad (10)$$

식 (10)을 통해 생체 단백질의 강성 행렬을 파악한다면, 고유모드 해석 기법을 이용하여 생체 단백질이 열적 에너지에 의해 브라운 운동에 의

한 요동(fluctuation)에 대한 동적 특성을 수치적으로 파악할 수 있음을 알 수 있다.

### 3. 탄성 네트워크 모델을 이용한 단백질의 동적 특성 해석

탄성 네트워크 모델은 기존의 원자모델 기반 분자동역학 전산모사 기법으로 해석이 수치적으로 어려운 분자량이 큰 단백질 구조체의 동적 특성 해석에 유용하게 활용되고 있다(K. Eom, et. al., J. Comput. Theor. Nanosci. 7, pp. 1210~1226

(2010)). 예를 들어, ATP와 같은 분자가 motor protein과 같은 거대 단백질 구조체에 화학 반응을 통해 구조의 변화를 유도하여 생물학적 기능을 하는 단백질의 동적 특성 해석에 탄성 네트워크 모델이 각광을 받고 있다. 또한 최근에 단백질이 ATP와 같은 리간드와 화학적으로 결합하여 유발되는 단백질의 구조 변화가 단백질 구조의 저주파 모드와의 상관관계가 있음이 보고되고 있다.

그림 3에서 보여준 바와 같이 대부분의 단백질의 구조 변화는 모드 인덱스가 5 이하의 저주파

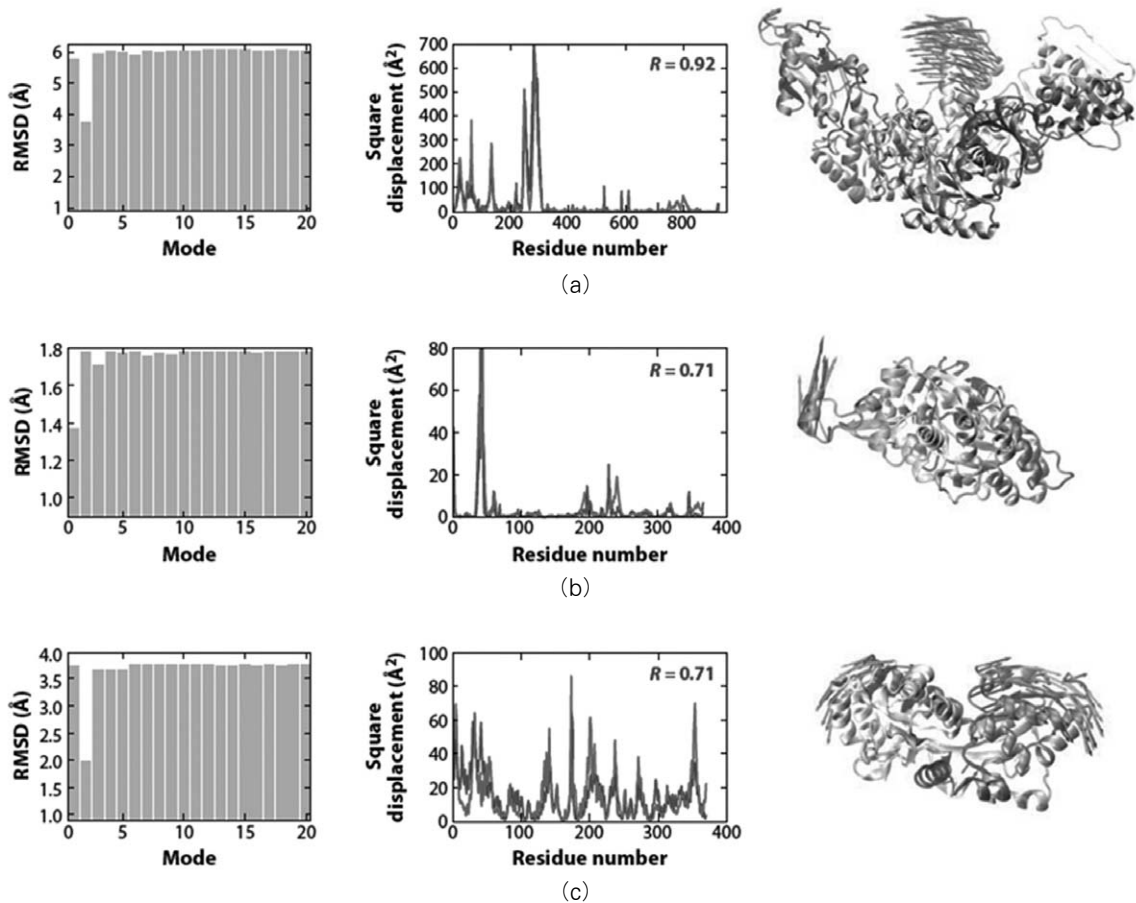


그림 3 리간드와 결합할 때 유발되는 단백질의 구조 변화와 저주파 모드와의 연관성. (왼쪽 그림) 특정 정규모드와 리간드가 결합하지 않은 구조와 리간드가 결합한 구조와의 root-mean-square distance(RMSD)와의 관계. (중간 그림) 구조 변화를 나타내는 RMSD가 감소하는 정규모드 방향으로 단백질 구조를 변화시킨 모델 결과 값과 실제 실험으로 측정된 구조 변화와의 비교. (오른쪽 그림) RMSD가 감소하는 정규모드 방향으로 단백질 구조 변화가 일어나는 모식도. (a) HIV-1 reverse transcriptase, (b) actin, (c) maltodextrin-binding protein (I. Bahar, et. al., Annu. Rev. Biophys. 39, pp. 23~42 (2010))

모드와 연관이 깊음을 알 수 있다. 예를 들면, HIV-1 reverse transcriptase 구조는 2번째 저주파 모드로 변화를 일으키면 결국 리간드가 결합한 구조에 근접하는(RMSD 감소) 반면에, 다른 저주파 모드로 구조 변화를 일으켜도 리간드와 결합한 구조에 근접하지 않는 것을(RMSD 변화가 없음) 알 수 있다. 이는 HIV-1 reverse transcriptase 가 리간드와 화학 결합할 때 일어나는 구조 변화는 2번째 저주파 모드와 연관이 있음을 보여준다(그림 3(a) 참고). 결국 energetics 관점에서 보면 자유에너지 표면에서 barrier가 낮은 저주파 모드 방향으로 단백질의 구조 변화가 유발됨을 알 수 있다. 이와 같이 탄성 네트워크 모델을 통해 리간드가 단백질에 화학적으로 결합할 때 단백질 구조 변화 메커니즘을 규명할 수 있기 때문

에, 1950년대부터 제시된 단백질과 리간드와의 결합에 대한 이론에 대해서도 탄성 네트워크 모델을 통해 검증이 가능함이 보고되고 있다. 단백질 구조와 리간드가 결합하는 화학결합 이론은 다음과 같이 3가지 이론적 모델로 제시되고 있다(그림 4).

- (1) **lock-and-key** 모델: 단백질 구조와 리간드가 마치 자물쇠와 열쇠처럼 서로 구조가 맞물려서 결합이 된다고 가정한다. 이 모델은 1950년대가 지나면서부터 단백질 구조가 정적인 구조가 아니라 동적 구조임을 규명하면서, 이 모델로는 열에너지에 의한 브라운 운동을 하는 단백질과 리간드와의 결합을 설명하기에는 불충분하다.
- (2) **Induced fit (Koshland-Nemethy-Filmer;**

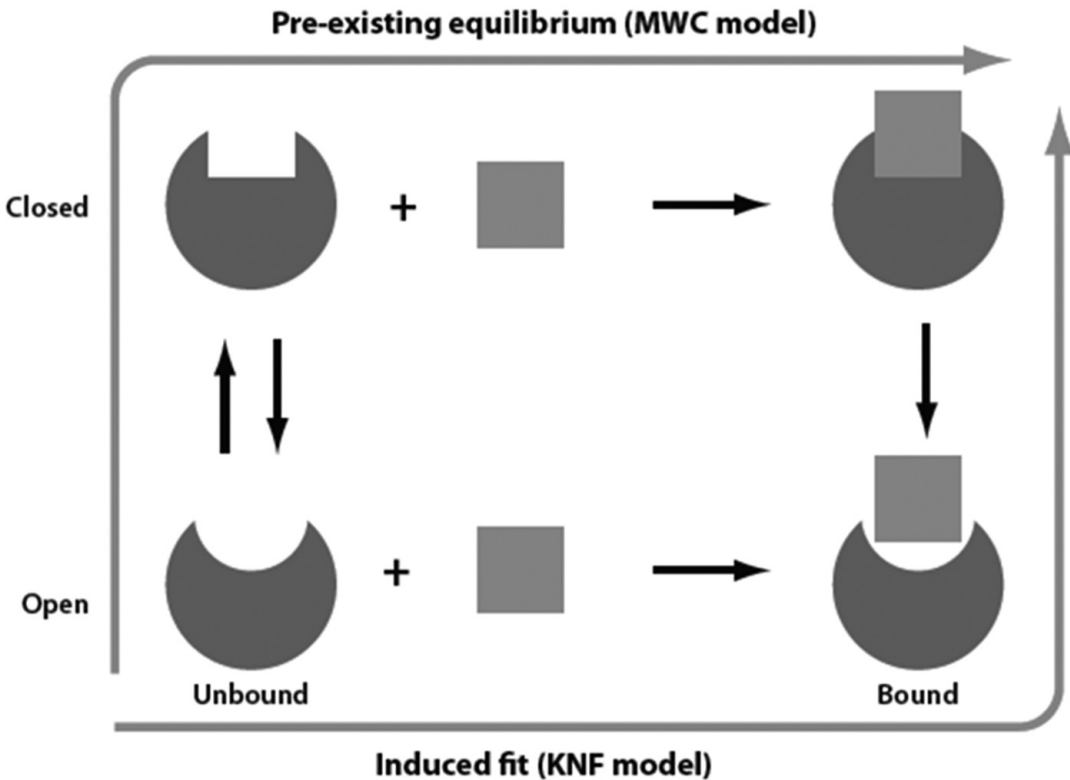


그림 4 리간드와 결합할 때 유도되는 단백질 구조 변화 메커니즘 모델: KNF 모델은 리간드와 단백질이 근접할 때 상호작용에 의해 단백질 구조 변화가 유도되어 화학 결합이 일어난다고 가정함. MWC 모델은 단백질 구조가 리간드와 결합하기 이전에 브라운 운동에 의해 구조 변화가 일어난 이후에 리간드와 결합한다고 가정함. (I. Bahar, et. al., Annu. Rev. Biophys. 39, pp. 23~42 (2010))

**KNF) 모델:** 단백질이 열에너지에 의해 브라운 운동을 하는 동적 구조이기 때문에, 단백질 구조와 리간드가 화학 결합이 일어날 때 리간드와 단백질과의 상호작용에 의해 단백질의 구조가 변하면서 화학 결합이 형성된다고 가정하는 모델이다.

### (3) Population Shift (Monod-Wyman-Changeux;

**MWC) 모델:** 단백질이 열에너지에 의해 브라운 운동을 하면서 다양한 구조로 전이가 일어난다. 이 때, 열에너지로 인한 브라운 운동에 의한 단백질의 구조 변화가 일어날 때의 에너지 변화는 열에너지 수준이다. 특정 리간드가 없을 때에는 단백질의 구조가 브라운 운동에 의해 여러 구조로 변동이 되지만 특정 구조에 있을 확률이 높다. 반면에, 특정 리간드가 단백질 구조에 근접하게 되면 단백질이 여러 구조로 변동할 때 결합이 형성되는 특정 구조에 있을 확률이 높아지도록 conformation 분포 확률이 변하게 된다.

위와 같이 제시된 이론적 모델 중에서 그동안 KNF 모델이 단백질과 리간드가 결합하는 메커니즘 설명에 적합하다고 여겨졌다. 하지만, 최근에 탄성 네트워크 모델을 통해 리간드가 없이 단백질 구조를 저주파 모드로 임의로 변화를 시키면, 그 변화된 구조가 리간드가 결합할 때의 구조와 매우 유사함이 보고되었다. 이를 통해 리간드가 없이 단백질이 열에너지에 의해 브라운 운동을 할 때 저주파 모드로 구조가 변하기 때문에, 단백질이 리간드와 결합하는 메커니즘을 MWC 모델로 설명이 가능함이 제시되었다. 따라서 최근에는 분자동역학 및 탄성 네트워크 모델 등 다양한 전산모사 기법을 통해 단백질마다 KNF 모델 또는 MWC 모델로 단백질과 리간드와의 화학

결합 메커니즘이 설명될 수 있음을 제시하였다. 이와 같이 단백질이 화학결합이 일어날 때 유발되는 단백질의 구조 변화를 이해하는데 진동 이론에서 많이 사용되는 정규모드 해석 기법이 매우 중요한 역할을 할 수 있음을 보여준다.

## 4. 맺음말

이 글에서는 진동 이론 중에서 정규모드 해석 기법이 생체 단백질의 생물학적 기능과 연관이 있는 동적 특성 해석에 활용 가능하다는 것을 보여주었다. 이는 단백질과 같은 생체 분자도 미시적 관점에서 보면 기계공학에서 다루는 구조 모델링 기법으로 해석이 가능하다는 것을 암시한다. 향후 기계공학에서 개발된 진동 및 고체역학 이론이 거시적인 구조물(예: 건물, 비행기, 자동차 등)의 동적 특성 해석뿐만 아니라 미시 스케일에서의 구조물(예: 단백질, 나노구조물 등)의 동적 특성 규명에 매우 유용하게 응용될 수 있음을 알 수 있다. 향후 21세기에는 기계적 성능이 우수한 소재 및 구조물 개발에 있어서, 거시적인 관점에서 최적 설계 기법을 개발하는 것뿐만 아니라 미시적인 관점에서의 최적 설계 기법 개발이 요구되고 있다. 특히, 미시적인 관점에서 생체 모방 구조물 설계를 위해서 생체 분자의 구조와 동적 특성 연관성 해석이 필수적이다. 이와 같이 단백질과 같은 생체분자의 구조와 동적 특성 연관성 해석을 위해서 기존의 진동 이론 및 고체역학 이론 등이 매우 유용하다. 따라서 향후 다양한 소재 및 구조물 개발에 있어서 미시적인 관점에서 설계 기법으로 진동 이론이 매우 중요한 역할을 할 것이 예상된다. **KSNVE**