

모드 축소법을 이용한 단백질의 구조변화 연구

김재인 * · 엄길호 * · 나성수 †

Jae In Kim, Kilho Eom, and Sungsoo Na

1. 서론

단백질은 그 기능에 따라 다양한 종류의 화학물질과 결합을 하게 되고, 이런 결합으로 기능을 할 때 구조의 변화가 생기게 된다. 각 단백질의 기능이 파악되고, 5 만개 이상의 단백질의 구조가 파악된 지금에 있어서도 단백질 구조 변화 경로를 실험적으로 파악하기에는 어려움이 있다. 이는 단백질의 구조가 변화되는 시간이 피코초(ps) 단위로 짧기에 실험적인 방법으로 관찰이 어렵기 때문이다. 때문에 단백질 구조변화에 기본이 되는 안정된 두 형태, 즉 열린형태(open form)와 닫힌형태(closed form)를 기반으로 한 경로예측 연구가 진행되고 있다. 이렇게 실험적인 방법으로 관찰하기 어려운 단백질의 구조변화를 연구하기 위해, 여러가지의 전산해석 기법들이 사용되고 있다. 전통적인 분자동역학 기법은 현재에도 중요한 기법으로 사용되고 있는데, 모든 원자를 고려한 전산해석을 수행하므로 화학적 결합까지 모두 연구할 수 있다는 장점이 있다. 반면에, 모든 원자를 고려하기 때문에, 전산해석을 수행하는데 있어서 컴퓨터의 부하가 많이 걸리고, 계산시간도 오래 걸리는 단점이 있다. 더욱이 일부 단백질의 구조변화가 일어나는 시간은 마이크로초(μs) 단위인데 반해 분자동역학으로 해석 가능한 시간은 피코초(ps) 단위라는 단점이 존재한다. 이러한 단점을 극복하기 위해 정규모드 해석법이 등장하게 등장하게 되었고, 분자동역학보다 단순한 퍼텐셜 함수를 사용해서 마이크로초(μs) 단위의 해석까지도 가능하게 되었다. 하지만 연구의 대상이 거대 단백질로 확대됨에 따라 모든 원자를 사용한 정규모드 해석법도 컴퓨터 자원에 있어서 한계를 드러내게 되었다. 이에 단백질의 동역학적 특성에 가장 큰 영향력을 미치는 α 탄소만을 고려한 탄성망 모델이 제안되었다. 이 모델은 탄소원자를 질점으로 가정하고, 원자사이에 작용하는 힘을 스프링으로 나타내어 다자유도계 질량-스프링 모델로 바꾼 것이다. 이번 연구에서는 이 탄성망 모델을 이용한 단백질의 구조변화 예측을 수행하였다.

† 교신저자; 고려대학교 기계공학과
E-mail : nass@korea.ac.kr
Tel : (02) 3290-3854,
* 고려대학교 기계공학과

단백질의 구조변화 예측에는 여러가지 전산해석 기법이 사용되고 있다. 사용되는 단백질의 구조 모델 뿐 아니라 구조변화의 중간단계를 예측하는 기법 자체가 여러가지 연구되고 있는데, 열린형태와 닫힌형태의 퍼텐셜을 섞어서 구조변화를 예측하는 방법, 열린형태와 닫힌형태의 두 구조간의 차를 이용한 보간법, 열린형태에서의 정규모드를 이용하여 구조를 변화시키는 방법 등이다. 이 연구에서는 정규모드를 이용한 구조변화 예측을 수행하였다.

2. 탄성네트워크 모델

탄성네트워크 모델은 단백질 동역학을 연구하는데 사용되는 단순 질량-스프링 모델이다. 노드의 위치는 protein data bank(PDB)에서 제공 받은 좌표중 α -탄소 원자의 위치로 결정하게 된다. 이 모델에서 노드간의 거리 12\AA 보다 작으면 하모닉 스프링으로 연결된다. 이때의 포텐셜 에너지는 다음과 같이 나타낼 수 있다.

$$P = \frac{\gamma}{2} \sum_{i,j}^N (r_{ij} - r_{ij}^0)^2 H(r_c - r_{ij}^0) \quad (1)$$

$H(r_c - r_{ij}^0)$ 는 헤비사이드 함수로 $r \geq 0$ 이면 $H(r_c - r_{ij}^0) = 1$, 다른 경우에는 $H(r_c - r_{ij}^0) = 0$ 이 된다.

또한 섭동 행렬 Q 는 다음과 같이 나타나게 된다.

$$Q_{ij} = \langle (r_i - \langle r_i \rangle) \cdot (r_j - \langle r_j \rangle) \rangle \quad (2)$$

3. 모드 축소법

모드 축소법은 단백질 구조물에서 적은 수의 노드만 가지고 정규모드 해석을 가능하게 해주는 방법이다. 전체 노드를 N 개라고 했을 때, (N/n) 의 수로 축소해서 구하게 된다. 단백질 구조물에 대해 다음과 같은 고유치 문제가 있다고 하자.

$$Ku = \lambda Mu \quad (3)$$

이 때 K 는 탄성행렬, M 은 질량행렬, λ 는 고유치, 그리고 u 는 고유벡터를 나타낸다. 이 고유치문제는 다음과 같은 형태로 변환될 수 있다.

$$\begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix} = \lambda \begin{bmatrix} M_{11} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix} \quad (4)$$

여기서 1 은 고려하고자 하는 노드를 의미하고, 2 는 비교대상인 노드를 의미한다. 위의 식은 다음과 같이 정리할 수 있다.

$$K_1 \mathbf{u}_1 = \lambda M_1 \mathbf{u}_1 \quad (5)$$

K1 과 M1 은 축소된 탄성행렬과 질량행렬을 나타낸다. 축소하는 과정은 다음의 식으로 나타낼 수 있다.

$$K_1 = K_{11} - K_{12} K_{22}^{-1} K_{21} \quad (6-1)$$

$$M_1 = M_{11} \quad (6-2)$$

이 방법을 이용해서 축소된 모드를 구할 수 있고, 이를 구조변화에 적용하게 된다.

4. 중간 구조 생성

모드 축소법을 통해 구한 정규모드를 이용해서 단백질의 구조 변화를 일으켜 중간 단계를 생성하게 된다. 중간 구조를 구하는 식은 다음과 같다.

$$\mathbf{r}^\pm = \mathbf{r}^0 \pm \alpha \left(\frac{k_B T}{\lambda_k} \right)^{1/2} \mathbf{V}_k \quad (7)$$

\mathbf{r}^\pm 은 +, - 방향으로 생성된 중간구조의 각 노드의 위치 벡터를 의미하고, \mathbf{r}^0 는 초기상태 노드의 위치 벡터를 의미한다. α 는 변화량을 조절하는 상수이고, k_B 는 볼츠만상수, T 는 온도, λ_k 는 k 번째 고유치, \mathbf{V}_k 는 k 번째 고유 모드를 나타낸다.

5. 에너지 감소

축소된 모드를 이용해 생성된 구조는 실제구조와 비교해 비현실적인 구조일 확률이 크다. 이를 극복하기 위해 에너지 감소를 매 단계에서 생성되는 +, - 방향의 두 구조에 적용하고, 그 중 에너지가 작은 것을 선택하는 방법으로 구조를 예측해 나간다. 에너지 감소법으로는 conjugate gradient method 를 적용해서 효율성을 높였다.

6. 시뮬레이션 결과

구조변화 예측을 위해서 대장균내에 존재하는 신호전달 단백질을 모델로 이용했다. 초기구조와, 마지막 구조는 pdb 데이터를 이용하여 비교대상으

로 삼았다. Pdb 데이터의 RMSD=1.3878Å 로 측정되었고, 89 회의 구조변화를 통해 RMSD=1.3728Å 의 값을 얻을 수 있었다.

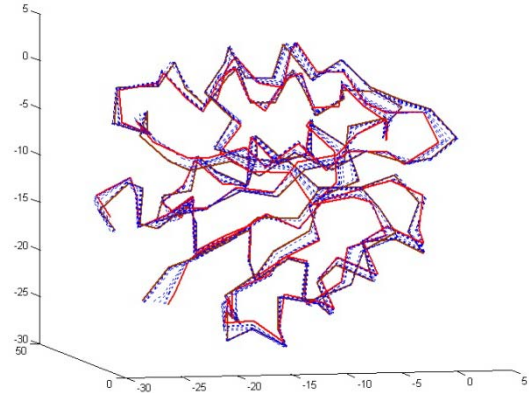


그림 1. 예측된 구조변화

7. 결론

모드 축소법을 이용해서 단백질 구조해석의 시간을 단축시킬 수 있었다. 또한 축소된 모드가 구조변화 예측에도 효과적임을 확인할 수 있었다. 이를 통해서 시작과 끝이 없는 구조에 대한 예측도 가능할 것으로 예상되며, 거대구조의 단백질에도 적용가능할 것으로 생각된다.

후 기

This work was supported by Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (No. 2009-0063170) and also acknowledges the support by Basic Research Program of the KOSEF(No. R01-2007-000-10497-0).