

# 다중복합탄성망 구조를 이용한 단백질의 동특성 연구

## Protein dynamics using Multi-scale Elastic Network Model

장효선\*·엄길호\*\*·나성수† \*

Hyoseon Jang, Kilho Eom and Sungsoo Na

### 1. 서 론

단백질의 생물학적 역할을 규명하기 위한 중요한 목적은 생체 내에서 이루어지고 있는 조절작용이 어떻게 일어나는지 이해하기 위함이다. 이러한 생물학적 역할을 이해하기 위해 단백질의 동특성 파악이 필수적으로 수반된다. 단백질의 동특성을 파악하기 위해 지금까지 분자동역학시물레이션(Molecular dynamics, MD), 고유모드해석기법(Normal Mode Analysis, NMA) 등 다양한 방법이 제시되었다. 하지만 이러한 기법들은 계산 시 소요되는 시간의 비효율성과 거대 단백질 해석에 따른 제약이 받는다는 단점이 있다. 이러한 단점을 보완하기 위해 고안된 방법이 탄성망구조(Elastic Network Model, ENM)를 이용한 연구기법이다. 우리는 단백질의 구조적 변화와 열 섭동에 의한 단백질의 움직임 등의 역할을 이해하고 좀 더 효율적으로 계산하기 위해 기존의 탄성망모델과 단순화된 탄성모델이 결합된 다중복합탄성망 구조를 이용하여 단백질의 동특성 연구를 진행하였다.

이 기법은 모델을 축소할 때에 우리가 관심을 가지고 있는 부분의 정보누락을 최소화하는 동시에, 계산의 효율성을 높인다는 점에서 기존의 모델축소 기법을 보완한 모델링기법이라 할 수 있다.

### 2. 본 론

#### 2.1 탄성망 모델기법과 모델축소기법

단백질의 구조해석 및 동적특성을 파악을 위한 시물레이션을 수행하기 위해 모델링을 할 때에, 단백질을 구성하고 있는 모든 원자를 사용하지 않고 움직임에 지배적인 역할을 하고 있는 원자인 Alpha

carbon(C<sub>α</sub>)만을 고려하여 모델을 구성한다.

#### (1) 탄성망 모델기법

단백질의 움직임에 지배적인 역할을 하는 C<sub>α</sub>에 대해 기계적인 질량, 스프링 모델을 접목하여 탄성망 모델의 1 차원 모델인 GNM(Gaussian network model)의 포텐셜 에너지를 구할 수 있으며 다음과 같이 정의된다.

$$E = \frac{\gamma}{2} \sum_i \sum_j \Gamma_{ij} u_i u_j \quad (1)$$

여기서, i, j 는 각각 C<sub>α</sub>의 순서, γ는 스프링 상수, u<sub>i</sub>은 i 번째 C<sub>α</sub>의 거동(fluctuation), Γ<sub>ij</sub>는 각 스프링의 연결성을 나타내는 행렬(Kirchhoff or connectivity matrix)를 의미한다. 기준 원자를 중심으로 일정범위(cutoff radius, r<sub>c</sub>)안에 있는 주위 원자들은 서로 운동에 영향을 주고 받는 연결성이 있다고 간주하고 -1 값을 사용하고, r<sub>c</sub>를 벗어난 원자들은 연결성이 없다고 여기고 zero 값을 사용한다. r<sub>c</sub> 값은 경험적으로 7~13Å의 값을 사용한다. 열의 섭동으로 인해 생기는 단백질의 거동은(fluctuation)은 다음과 같이 정의되며,

$$\langle \Delta R_i \cdot \Delta R_j \rangle = \frac{3k_B T}{\gamma} [\Gamma^{-1}]_{ij} = \frac{3k_B T}{\gamma} \sum_{k=2}^N [u_k \lambda_k^{-1} u_k^T]_{ij} \quad (2)$$

여기서, k는 k 번째 모드를 의미하는데, 한 개의 강체 모드에 의하여 zero 모드가 존재한다. 평형상태에서 열에너지에 의한 i 번째 C<sub>α</sub>의 거동은 다음과 같이 정의된다.

$$B_i = \frac{8\pi^2}{3} \langle \Delta R_i^2 \rangle = \frac{8\pi^2 k_B T}{\gamma} [\Gamma^{-1}]_{ii} \quad (3)$$

시물레이션 결과로 얻어진 값과 실험을 통하여 얻어진 값을 비교하여 스프링상수 값을 구할 수 있다.

† 교신저자; 고려대학교 기계공학과

E-mail : nass@korea.ac.kr

Tel : (02) 3290-3854, Fax : (02) 6008-3855

\* 고려대학교 기계공학과

\*\* 고려대학교 기계공학과

### (2) 모델 축소 기법

거대 단백질의 해석 시 계산의 효율성을 증대시키기 위하여 고안된 방법으로 전체 N 개의 원자로 이루어진 단백질의 구조를  $N/n(n=1,2,3,\dots,n)$ 개의  $C_\alpha$  만을 사용하여 모델링을 하고 시뮬레이션을 한 기법이다. 모델 구성은 탄성망 구조와 같은 방법으로 이루어지지만 적은 개수의  $C_\alpha$  만을 사용하기 때문에 매 개변수를 새롭게 계산하여 적용하였다. 적은 개수의  $C_\alpha$  만으로도 단백질 구조의 동적특성이 잘 묘사되고 원래의 분자모델과 상당히 유사한 결과를 얻을 수 있음이 입증되었다.

## 3. 다중복합탄성망 구조를 이용한 모델 축소 기법

### 3.1 다중복합탄성망 구조 기법

단백질의 구조변화와 동적특성을 파악할 때에, 우리가 관심을 가지고 있는 부분의 정보손실을 최소화하고 계산의 효율성 증대를 위해서 고안된 기법이다. 실제로 결합이 이루어지는 부분(Binding sites)은 모든  $C_\alpha$  를 사용하여 모델링을 하고, 그 외 나머지 부분은 적은 개수의  $C_\alpha$  만을 사용하여 모델링을 하여 시뮬레이션을 수행하였다.

#### (1) Cutoff radius( $r_c$ )

다중복합탄성망 구조는 결합이 이루어지는 부분과 나머지 부분이 서로 다른 해상도로 모델이 구성된다. 따라서, 각 부분에서 원자들 사이에 전자기적 힘이 나 반데르발스 같은 힘이 미치는 범위가 달라지기 때문에 값이 다르게 적용된다. 결합이 이루어지는 부분의  $r_c$  값은 위에서 언급한 7~13Å 값을 사용하지 않, 적은 개수의  $C_\alpha$  로 구성된 부분의  $r_c$  값은 재규격화(renormalization)시킨 값에 본래의  $r_c$  을 합산한 값을 사용하며, 식은 다음과 같이 정의된다.

$$r_c = 2R_G + R_0 \quad (4)$$

여기서,  $R_G$  값은 재규격화된 값이고,  $R_0$  는 7~13Å 값이다. 그리고 서로 다른 해상도로 이루어진 부분이 맞닿아 있는 부분(interface parts)의  $r_c$  값은 다음과 같이 정의된다.

$$r_{c1,2} = \left( \frac{r_{c1}^3 + r_{c2}^3}{2} \right)^{1/3} \quad (5)$$

여기서, 각 지역의  $r_c$  값 안에는 동일한 개수의 원자가 포함되어 있고, 구 형태로 이루어진 범위의 부피에 비례하여  $r_c$  값이 변하게 된다고 간주된다.

### (2) 스프링 상수 (Spring constants)

새롭게 구성된 다중복합탄성망 구조 단백질은 각 부분별로 해상도가 다르기 때문에 스프링상수 값 역시 각기 다르게 적용하여야 한다. 결합이 이루어지는 부분은 위의 (3)식과 같이 실험치와 비교하여 스프링상수를 결정하고, 그 외 축소된 형태로 구성된 부분의 스프링상수는 Kullback-Libler Divergence 를 이용하여 그 값이 최소가 되는 스프링상수 값을 사용한다. 그리고 서로 다른 지역이 맞닿아 있는 부분(interface parts)의 스프링 상수는 각 부분에서 얻어진 값을 선형보간법(linear interpolation)을 이용하여 도출한다.

## 4. 결 론

이 논문에서는 다중복합탄성망 구조를 이용하여 단백질의 동적특성 연구를 진행하였다. 거대단백질의 해석 시 효율적인 계산시간을 확보하는 동시에 단백질의 구조변화가 일어날 때, 큰 폭으로 변화하는 부분의 정보누락을 최소화시킬 수 있는 모델을 구성하여 시뮬레이션 한 결과 원래의 탄성망 구조나 실험을 통하여 얻어진 결과와 비교해 보았을 때, 특정 단백질의 동적특성이 잘 보존되는 것을 확인할 수 있었다. 따라서, 향후 이 모델을 기반으로 하여 단백질의 구조변화 해석을 하는데 적용하여 연구를 진행할 예정이다.

## 후 기

This work was supported by the Korea Science and Engineering Foundation (No.R11-2007-028-03002) and also acknowledges the support by Basic Research Program of the KOSEF under grant No. R01-2007-000-10497-0.