탄성망모델을 이용한 단백질 열안정성 해석

박영걸†·원종진*·정재일**

Thermostability prediction of protein structure by using elastic network model

Young Gul Park, Chongjin Won and Jay I. Jeong

Key Words: Elastic Network Model(탄성망 모델), Thermophile(호열성 단백질), Thermostability (열안정성)

Abstract

In this study, an elastic network model is established in order to find dominant factors which reflect thermostability of protein structures. The connections in the elastic network model are selected with respect to the free energy between alpha-carbons, which is representatives of residues in the elastic network model. We carried out normal mode analysis and compared eigenvalues of the stiffness matrix from the elastic network model. In most cases, thermophilic proteins are observed to have higher values of lowest natural frequency than mesophiles and psychrophiles have. As a result, the thermophiles are calculated to be stiffer than other proteins in view of dynamic vibration.

1. 서 론

최근에 가장 많은 연구가 이루어지고 있는 분야 중 하나가 단백질이다. 단백질은 생명을 유지하는 가장 기본적인 단위이며 많은 역할을 담당하고 있으며 많은 불치병들을 치료할 수 있는 단서이기도 하다.

또한 단백질 연구에 있어서 여러 분야에서 각각의 시각으로 연구가 진행되고 있고 그 주제 또한 매우 다양하다. 그중 단백질의 열안정성(thermostability)에 관한 연구가 활발히 이뤄지고 있다. 그러나 아직 단백질의 열안정성에 관한 확실한 요인이 밝혀진 것이 없으며 앞으로 더욱 많은 연구가 이뤄질 것으로 예상된다.

이전 연구들은 단백질의 열안정성 요인을 밝히기 위해 단순한 비교들을 하였다[1,2,3,4]. 이 연구들에서는 수소결합과 이온결합의 개수, 단백질의 밀도, 크기, 아미노산의 빈도, 아미노산간의 연결종류 빈도, 금속이온의 유무처럼 단순한 비교를 하여왔다. 그러나 두드러진 특징을 찾지 못하였다.

본 연구에서는 탄성망모델(Elastic Network Model)을 이용하여 호열성 단백질(thermophile)의 특성을 밝히고자 하였다. 단백질의 동강성 특성을 파악할 수 있는 NMA(Normal Mode Analysis)를 통하여 그 특성을 밝히고자 하였다.

† 국민대학교 기계자동차공학부
E-mail : jayjeong@kookmin.ac.kr
TEL : (02)910-4419 FAX : (02)910-4839
* 국민대학교 기계설계대학원
** 국민대학교 기계자동차공학부

2. 연구방법

2.1 단백질 선정

본 연구에 사용된 단백질은 호열성 단백질(thermophile)과 중온성 단백질(mesophile)을 가지

는 family를 선정하였다. 그리고 molecule unit이 단량체(monomer)인 단백질을 선정하였다. 아미노시퀀스의 길이가 짧은 것과 긴 것으로 두 종을 선정하였다. 단백질의 PDB(Protein Data Bank) code는 1CAA/8RXN, 1THM/1BH6 이다.

2.2 Normal Mode Analysis

정규모드는 계의 운동방정식의 해를 찾아서 구할수 있다. 일반적으로 운동방정식은 계의 질량과 강성이 포함되어 있고 고유치 문제로 해결할 수 있다. 그러나 본 연구에서는 질량을 고려하지 않은 계의 강성을 표현한 강성행렬의 고유치 문제의 해를 구하여 해석을 수행하였다.

2.2.1 탄성망 모델

아미노산의 중심 원자인 alpha carbon 만을 고려하는 CGENM(coarse-grained elastic network model)을 사용하였다[5]. 우리는 CGENM의 강성행렬을 생성하기위하여 Linking Matrix를 만들었다. alpha carbon간의 연결을 결정하는 것은 펩타이드 결합, 수소결합, 이온결합, 이황결합, 그리고 반데르발스 힘(Van Der Waals Force)이 있다. 먼저 펩타이드 결합은 최소한의 안정성을 위해 (i)-(i+1), (i)-(i+2), 그리고 (i)-(i+3)를 연결하였다. 수소결합은 HB-PLUS를[6] 이용하여 계산하였고 데이터 내의 Residue와 Residue 사이의 수소결합만을 고려하였다. 이온결합은 음이온과 양이온 사이의 거리를 계산하여 연결하였고 이온결합 길이를 4Å 이하로 설정하였다. 이황결합은 황원자 사이의 거리가 3Å 이하를 가지는 것을 연결하였다. 반데르발스 힘은 alpha carbon사이의 길이를 설정하였다. 그 길이는 12Å으로 하였으며 그이하의 길이를 가지는 alpha carbon을 연결하였다.

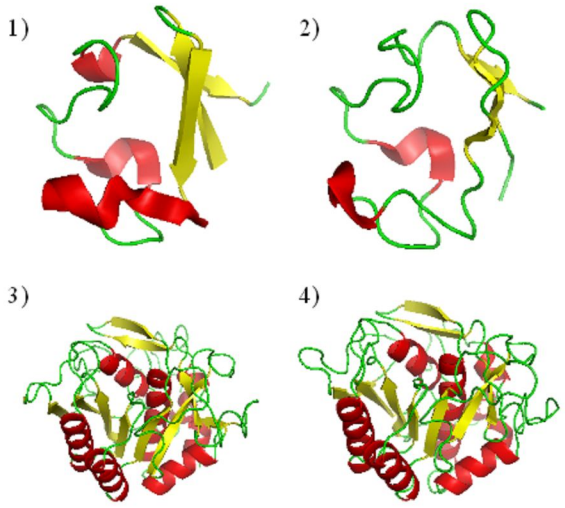

Fig. 1 shape of Rubredoxin and Subtilisin
1)1CAA, 2)8RXN, 3)1THM, 4)1BH6

2.2.2 수소결합의 에너지

수소결합은 단백질을 이루는 연결 중 가장 약한 연결이지만 많은 수로 인해 중요한 역할을 한다. 본 연구에서는 높은 온도의 영향을 반영하기 위하여 낮은 에너지를 가지는 수소결합은 끊어서 모델을 만들었다.

수소결합의 에너지는 여러가지 에너지 식이 제시되었으며 대부분 결합 길이와 각도에 대한 식으로 구성이 되어 있다[7,8]. 그래서 본 연구에서는 길이와 각도의 크기에 따라서 연결을 결정하였고 NMA를 수행하였다. 생성한 모델의 종류는 수정전 모델(반데르발스 힘 포함, 모델1), 반데르발스 힘미포함 모델(모델2), 수소결합 길이 2.0Å이하(모델3), 1.9Å(모델4) 그리고 수소결합 길이 2.0Å 이하와 각도 160°이상(모델5)을 기준으로 한 모델 다섯가지를 생성하여 NMA를 수행하였고 고유값(eigenvalue)을 비교하였다.

3. 결과

3.1 Rubredoxin에서의 고유 값의 차이

Rubredoxin은 단지 오십여개의 아미노산으로 이루어진 작은 크기의 단백질이다. 1CAA는 꼬리를 가지는 특성을 보이고 있고 8RXN은 꼬리를 가지고 있지 않다. 먼저 반데르발스 힘을 포함하는 모델1에서는 중온성 단백질이 더 높은 고유값을 가진다. 이것은 1CAA의 형상의 영향을 받은

Table 1 Number of hydrogen bond (Rubredoxin)

cutoff	origin	L>2.0Å	L>1.9Å	L>2.0Å A<160°
1CAA	37	22	12	10
8RXN	33	19	10	12

Table 2 Number of hydrogen bond (Subtilisin)

cutoff	original	L>2.0Å	L>1.9Å	L>2.0Å A<160°
1THM	274	135	66	76
1BH6	250	125	69	72

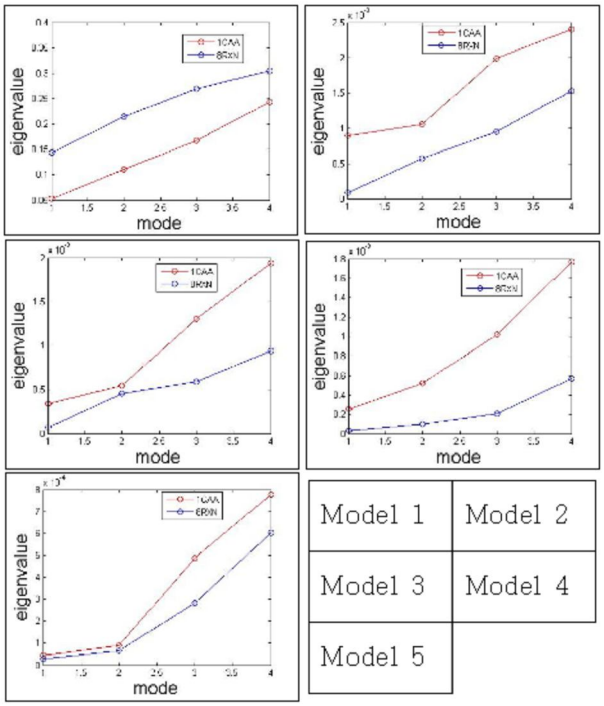

Fig. 2 eigenvalue of Rubredoxin

것으로 반데르발스 힘이 나쁜 영향을 주고 있다고 판단하여 반데르발스 힘을 제거하는 것이 좀 더 높은 차별성을 보이는 것을 알수있었다..

수소결합의 에너지가 강한 것만을 연결한 모델에서 호열성 단백질이 높은 고유값을 가지게 된다. 모든 화학결합의 수를 비교해보면 그 차이가 크지 않음을 알 수 있다. 그럼에도 불구하고 수소결합의 길이가 더 짧아진 모델에서 고유값의 차이가 두드러지는 것을 볼 수 있다. 모델3에서는 호열성의 고유값이 크지만 그 차이가 다른 모델에 비해 적은 것을 볼 수 있다. 모델 5에서도 비슷한 경향을 볼 수 있다.

3.2 SUBTILISIN에서의 고유값 차이

270여개의 아미노산으로 구성된 Subtilisin은 대부분의 모델에서 호열성 단백질의 고유값이 크게 나타난다. Rubredoxin에서와는 달리 반데르 발스 힘이 영향을 끼치지 못한다. 그러나 모델4에서는 경향이 뚜렷이 보이지 않고 두 번째 모드에서 중온성 단백질의 고유값이 크게 나타나는 것을 볼 수 있다. 이단백질에서는 모델5의 결과에서 가장 뚜렷한 차이를 보이고 있다.

4. 토 의

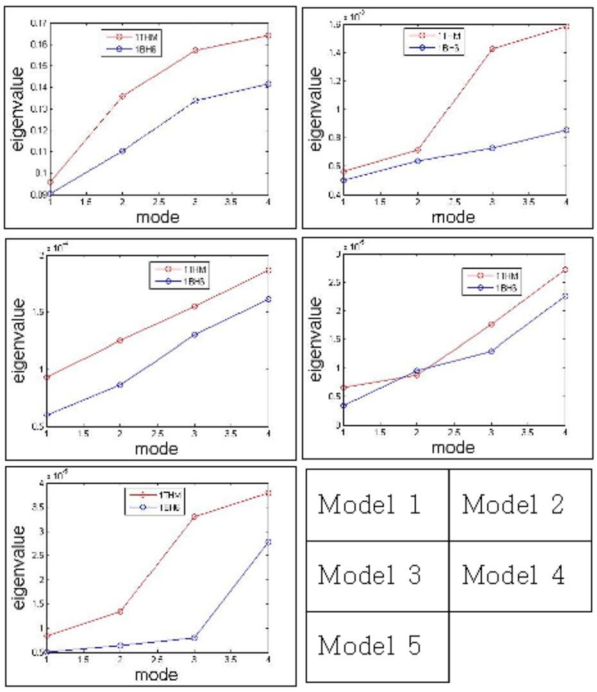

Fig. 3 eigenvalue of Subtilisin

반데르발스 힘을 적용시킨 결과에서 중온성 단백질이 호열성 단백질보다 큰 고유치를 가지게 된다. 그 이유는 1CAA가 가지는 꼬리 때문이며 반데르발스 힘이 1CAA의 몸통의 강성을 높여서 꼬리와 분리되게 만들었다. 결과적으로 꼬리만 따로 움직이는 모드를 만들었기 때문이다. 약한 화학결합의 영향을 확실히 보기 위하여 반데르발스 힘을 제거하고 해석 하여야 한다는 것을 알 수 있다.

수소결합에서 공여체(donor)와 수용체(accepter) 사의의 길이를 기준으로 연결을 결정하게 되면 호열성의 고유값이 커지는 경향이 뚜렷이 보이지 않는다. 수소결합의 에너지 식을 너무 간단화했기 때문인데 수소결합의 길이와 각도가 모두 포함되어야만 한다. 길이와 각도를 포함한 모델에서는 더욱 두드러진 차이를 보이게 된다.

화학결합의 수는 많은 차이를 보이고 있지 않음을 알 수 있다. 하지만 고유주파수에서는 많은 차를 나타냄을 볼 수 있는데 화학결합의 수의 차이보다는 단백질의 topology의 영향이 큼을 알 수 있다. 각 단백질들의 active cite 사이에서의 연결들이 많은 영향을 끼치고 있고 호열성 단백질이 많은 수가 존재함을 확인할 수 있다.

대부분의 결과에서 보듯이 호열성이 중온성보다 높은 고유값을 가진다. 즉 동강성이 높으며

이로 인해 높은 온도에서의 안정성을 높을 수 있다. 그러므로 이결과를 이용하여 같은 family내에서 활성온도의 높고 낮음을 평가할 수 있을 것으로 예상한다.

4. 결 론

본 논문에서는 호열성 단백질의 열안정성 요인을 파악하고자 단백질의 topology를 반영하는 탄성망 모델을 만들었다. 이 모델은 단백질내의 주요한 화학결합인 수소결합, 이온결합 그리고 이황결합을 포함하고 있다. 특히 수소결합은 단백질의 열환경을 고려하기 위하여 연결에너지에 따른 변화를 반영하였다.

위의 모델들은 질량을 고려하지 않은 NMA를 수행하였다. 그 결과로 각 모델들의 고유값 즉 단백질의 고유진동수로 표현될수 있는 값을 얻을 수 있었다. 그리고 호열성 단백질의 고유값과 중온성이 고유값을 비교하여 호열성 단백질의 특성을 찾을수 있었다.

호열성 단백질은 일반적으로 중온성 단백질보다 높은 고유진동수를 가진다. 즉 높은 동강성을 가지게 된다. 또한 높은 동강성은 화학결합의 수보다는 topology의 영향이 크다.

이 연구의 결과로 우리는 단백질의 NMA 결과로 단백질의 열안정성으로 평가할 수 있다. 또한 고유진동수를 높이는 설계로 호열성 단백질을 설계할 수 있을 것으로 예상된다.

참고문헌

(1) Jason K Yano, Thomas L Poulos, "New understandings of thermostable and peizostable enzymes", *Curr. Opin. Biotechnol.*, Vol. 14, pp.360-365.

(2) Gerhard Vogt, Stefanie Woell, Patrick Argos, "Protein Thermal Stability, Hydrogen Bonds, and Ion Pairs", *JMB*, Vol. 269 pp.631-643.

(3) K.Saraboji, M. Michael Gromiha, M. N. Ponnuswamy, 2005, "Importance of main-chain hydrophobic free energy to the stability of thermophilic proteins", *Int. J. Biol. Macromol*. Vol. 35 pp.211-220.

(4) B. Syed Ibrahim, Vasantha Pattabhi, 2004, "Role of weak interactions in thermal stability of proteins", *BBRC*, Vol. 325, pp.1082-1089.

(5) Jay I. Jeong, Yunho Jang, and Moon K. Kim, 2006, "A connection rule for α-carbon coarse-grained elastic network models using chemical bond information," *JMGM*, Vol. 24, pp. 296~306.

(6) I.K. McDonald, J.M. Thornton, 1994, "Satisfying hydrogen-bonding potential in proteins", J. Mol. Biol, Vol. 238, pp.777－ 3.

(7) Donald F. Facobs, A.J.Rader, Leslie A. Kuhn, M.F. Thorpe, 2001, "Protein Flexibility predictions Using Graph Theory", *Proteins: Struct. Funct. Genet*. Vol. 44, pp. 150－ 5.

(8) M.F. Thorpe, Ming Lei, A.J. Rader, Donald J. Jacobs, Leslie A. Kuhn, 2001, "Protein flexibility and dynamics using constraint theory", *JMGM*, Vol. 19, pp. 60-69.