

**IgG binding domain of protein G 의 폴딩 현상에 대한  
분자 수준의 해석 연구**

이승엽<sup>(a)</sup>, Shoji Takada<sup>(b)</sup>, 김도현<sup>(a)</sup>

<sup>(a)</sup>한국과학기술원 생명화학공학과, 초미세화학공정센터

<sup>(b)</sup>고베대학교 이학부, JST PRESTO 21

| ; †

**Molecular scale simulation of folding of IgG binding  
domain of protein G**

Seung Yup Lee<sup>(a)</sup>, Shoji Takada<sup>(b)</sup>, Do Hyun Kim<sup>(a)</sup>

<sup>(a)</sup>Department of Chemical and Biomolecular Engineering and

Center for Ultramicrochemical Process Systems,

Korea Advanced Institute of Science and Technology,

Daejeon, 305-701, KOREA

<sup>(b)</sup>Graduate School of Science and Technology, Kobe University and

Japan Science and Technology, PRESTO 21,

Rokkodai, Nada, Kobe, 657-8501, JAPAN

**서론 및 연구목적**

단백질 폴딩 현상에 대한 이해는 단백질의 구조 및 작용, 그리고 넓게는 많은 생물 화학적 현상을 이해하고 적용하는데 필요하므로 많은 연구가 진행되어 오고 있다. 그러나 단백질 자체의 복잡성과 그에 작용하는 많은 물리화학적 작용들로 인해 많은 이론적, 실험적 연구에도 불구하고 밝혀진 부분은 미비하다.

최근에는 작은 크기의 단백질의 폴딩 (residues 수가 100 개 이하) 현상에 대해서 많은 연구가 이루어지고 있고, 특히 최근에 등장한 energy landscape theory [1-3] 와 minimal frustrated model 또는 Gō-like model [4-5] 이라고 불리는 방법에 기초한 연구들을 통해서 대부분의 작은 단백질이 two-state 폴딩 현상을 보이고 폴딩을 결정하는 가장 중요한 요소는 native topology 라는 것이 밝혀 졌다. 또한 이에 대한 많은 실험적 연구 결과들도 이를 뒷받침하고 있다 [6,7]. 그러나 몇몇 단백질은 이런 native topology 가 큰 역할을 하지 못한다고 알려졌는데, 그 중에서도 56개의 residues 로 구성된 IgG binding domain of protein G 는 작은 단백질이고, 구조가 완전한 대칭형임에도 minimal frustrated model 에 근거한 방법으로는 이 단백질의 폴딩 현상을 잘 설명할 수가 없었다.

따라서 본 연구에서는 IgG binding domain of protein G 의 폴딩 현상을 분자 수준의 해석 방법인 over-damped Langevin dynamics 를 이용해서 예측하였다. 단백질 구조정보 데이터 베이스와 단백질 작용의 물리화학적 지식 등을 바탕으로 한 분자간 상호 작용 [8-9] 을 topological constraints 와 함께 고려하여 protein G 의 폴딩 pathways, 이차 구조 형성 propensity, 그리고  $\alpha$  helix 와 2개의  $\beta$  hairpins 의 구조 형성과 안정성을 root mean square deviation (RMSD), local order parameter 그리고 number of helical segments 등을 통해서 조사하고 폴딩 pathways 에 대한  $\Phi$  value analysis 의 실험결과와 비교 분석 하였다.

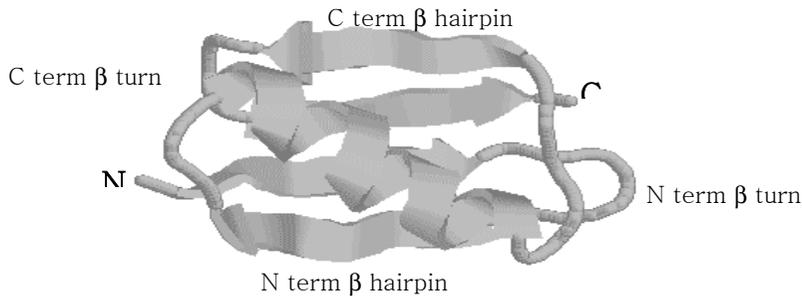
**2. 이론 및 연구내용**

*화학공학의 이론과 응용 제8권 제2호 2002년*

본 연구에서 단백질 chain 은 backbone  $C_{\alpha}$  model 과 all-atom model 의 중간 형태인 coarse grained model 의 reduced chain 으로 나타내어진다. 즉 1개의 아미노산은 3개의 통합 backbone atoms ( $C_{\alpha}$ , NH, 그리고 CO) 와 1개의 서로 다른 centroids (19 종류, Gly 은 제외) 로 구성되어진다. Protein G 는 native topology 가 폴딩에서 큰 역할을 하지 못하기 때문에 본 연구에서는 단백질 입자간의 상호 작용을 모사하기 위해서 G $\bar{o}$ -like model 에서 제안된 native topology 에 근거한 힘과 함께 단백질 입자들간의 물리화학적 작용을 나타내는 힘을 고려한 포텐셜 에너지 함수 (PEF) 를 사용 하였다. 초기 random 한 단백질 구조로부터 native 구조로까지 변화하는 폴딩 과정 동안 protein G 의 폴딩 현상 및 이차구조 형성 propensity 등을 over-damped Langevin dynamics 와 제시된 에너지 함수를 이용해서 관찰하였다. 모델에서 물 입자는 직접적으로 포함되지 않고, 용매효과를 고려한 작용들을 물리화학적 PEF 에 포함시켜서 계산을 수행하였다.

### 3. 결과 및 고찰

IgG binding domain of protein G 는 중앙에 위치한  $\alpha$  helix 를 기준으로 2개의  $\beta$  hairpins 이 위치하고 있다 (Figure 1). 각각의  $\beta$  hairpin 은 2개의 strands 와 그것을 연결하는  $\beta$  turn 으로 구성되어 있다. Kabsch and Sander's criteria [10] 에 따르면 N terminus (1 MET) 에서 helix 앞 부분 (19 GLU) 까지가 N term  $\beta$  hairpin 이고, C term  $\beta$  hairpin 은 42 GLU 부터 C terminus (56 GLU) 까지의 부분이다.



SEQUENCE: MTYKLIILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDATKTFVTVE

N term →

→ C term

Figure 1. IgG binding domain of Protein G 의 native topology

두개의 hairpins 의 folding 거동을 보기 위해서 평균 부분 RMSD 을 figure 2(a) 에 도시하였다. 본 연구에서 전체 폴딩 시간은 600 ns 이고, 모두 20개 폴딩 trajectories 에 대해서 평균한 값을 도시하였다. 회색 선은 N-term  $\beta$  hairpin 이고 검은 선은 C-term  $\beta$  hairpin 이다. 그림에서 볼 수 있듯이 C-term  $\beta$  hairpin 이 좀 더 빨리 native topology 에 접근하고 fluctuation 도 작다. 이것은 C-term  $\beta$  hairpin 이 N-term  $\beta$  hairpin 보다 폴딩 과정에서 빨리 형성되고 생성된 구조가 안정하다는 것을 의미한다. 이는 transition state 에서의  $\Phi$  value analysis 실험의 결과에서 예측한 것과 일치한다 [11]. 즉 폴딩 과정에서는 두 hairpins 의 native 구조 대칭성으로 인한 생성 과정의 대칭성은 지켜지지 않고, 그 중 하나인 C-term  $\beta$  hairpin 의 형성이 선호되어서 이 부분이 더 안정적이고 빨리 형성된다.

각각의 hairpin 은 2개의 strands 가 turn 으로 연결되어 있다. 따라서 turn 부분의 안정성과 형성과정은 hairpin 의 형성 거동에 큰 영향을 미친다. Turn 부분의 안정성과 거동을 관찰하고 계산 결과를 실험값과 비교 하기 위해서 turn 부분을 이루고 있는 2개의 residues (10 LYS, 49 THR) 에 대해서 local order parameter 인  $Q_{local}$  을 계산하였다 (Figure 2(b)).  $Q_{local}$  값이 1에 가까울수록 그 residue 가 native structure 에 근접한 것을 의미하고, 반대로 0 근처의 값을 가질수록 denatured state 상태에 있다는 것을 나타낸다.

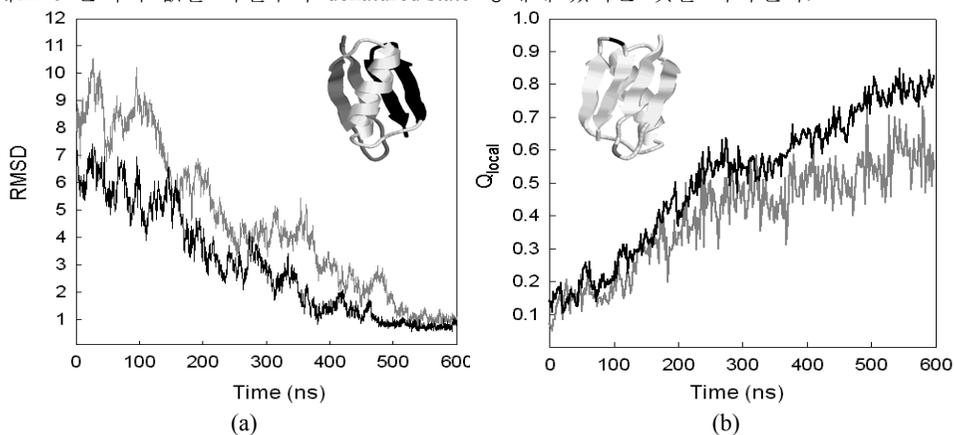


Figure 2. (a) N-and C-term  $\beta$  hairpins 의 평균 부분 RMSD  
(b) N-and C-term  $\beta$  turn residues 의 평균 local order parameter,  $Q_{local}$   
: 10 LYS and 49 THR residues

폴딩 시간에 따른 N- 와 C-term  $\beta$  turn 의 10 LYS 와 49 THR residue 의 평균 local order parameter,  $Q_{local}$  값을 figure 2(b) 에 각각 회색과 검은색 선으로 표시하였다. 생성이 느리고 불안정한 부분인 N-term  $\beta$  hairpin 부분의  $\beta$  turn 을 구성하는 10 LYS 의  $Q_{local}$  평균값도 큰 fluctuation 을 보이고, 전체 구조가 거의 native 구조에 도달하여도 ( $t=600ns$ ) 여전히 49 THR residue 보다 낮은 값을 보인다. 이처럼 N-term turn 부분이 불안정하고 생성이 느린 것은 N-term  $\beta$  hairpin 이 불안정하고 형성 과정이 C-term hairpin 보다 느린 것을 나타낸다. Local order parameter 와 부분 RMSD 의 예측 결과는 native 구조의 안정성 예측 과  $\Phi$  value analysis 실험에서 관찰한 결과와 정확히 일치한다 [11-12].

$\alpha$  helix 부분의 거동을 관찰하기 위해서 Kabsch and Sander's criteria [10] 에 의한 number of helix segments 를 계산한 결과를 Figure 3 에 나타내었다. 가는 회색 라인들은 하나의 폴딩 trajectory 에 대해서 나타낸 것이고, 굵은 검은색은 평균한 값을 나타낸다. 그림에서 볼 수 있듯이 완전한 native 구조를 형성하기 전에 helix 는 부분적인 형성과 붕괴를 반복한다. 즉 helix 는 폴딩 과정 동안 불안정하고 native 구조 형성도 늦게 되게 된다. 폴딩 과정 동안의 생성되는 구조 snapshot 을 보면  $t=240$  ns 일때 helix 는 부분 형성되었다가 다시 붕괴되고 ( $t=390$  ns), 다시 형성된다 ( $t=450$  ns). 그리고 C-term  $\beta$  hairpin 은 앞서 계산한 것과 같이 N-term  $\beta$  hairpin 보다 일찍 형성되고 안정하다.

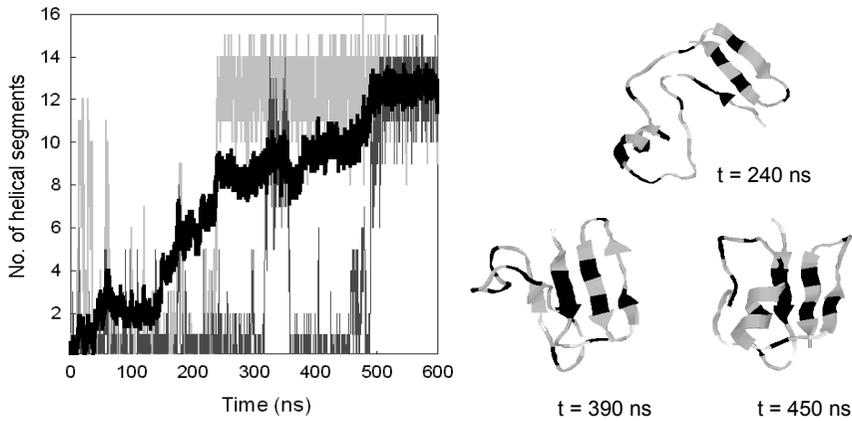


Figure 3. 폴딩 시간에 따른 number of helical segments 와 폴딩 과정에서 구조 snapshots : 240 ns, 390 ns 와 450 ns

### 결론

IgG binding domain of protein G 폴딩 현상을 topological constraints 와 물리화학적 작용을 고려해서 over-damped Langevin dynamics 을 이용해 예측하였다. 폴딩 과정동안 C-term  $\beta$  hairpin 이 N-term  $\beta$  hairpin 보다 일찍 형성되고 안정한 구조를 유지하고,  $\beta$  turn 부분도 (10 LYS, 49 THR) 같은 경향을 나타낸다.  $\alpha$  helix 는 부분적 형성과 붕괴를 반복하면서 구조의 불안정성을 나타내고, 형성도 늦게 된다.

### 사사

본 연구는 BK 21 프로그램과 CUPS 에 의해 지원되었기에 이에 감사한다.

### 참고문헌

1. Dill, K. A. and Chan, H. S., *Nature Struct. Biol.*, **4**, 10(1997).
2. Onuchic, J. N., Wolynes, P. G., Luthey-Schulten, Z. and Socci, N. D., *Proc. Natl. Acad. Sci. USA*, **92**, 3626(1995).
3. Wolynes, P. G., Onuchic, J. N. and Thirumalai, D., *Science*, **267**, 1619(1995).
4. Clementi, C., Nymeyer, H. and Onuchic, J. N., *J. Mol. Biol.*, **298**, 937(2000).
5. Klimov, D. K., Newfield, D. and Thirumalai, D., *Proc. Natl. Acad. Sci. USA*, **99**, 8019(2002).
6. Baker, D., *Nature*, **405**, 39(2000).
7. Guerois, R. and Serrano, L., *Curr. Opin. Struc. Biol.*, **11**, 101(2001).
8. Takada, S., *Proteins*, **42**, 85(2001).
9. Takada, S., Luthey-Schulten, Z. and Wolynes, P. G., *J. Chem. Phys.*, **110**, 11616(1999).
10. Kabsch, W. and Sander, C., *Biopolymers*, **22**, 2577(1983).
11. McCallister, E. L., Alm, E. and Baker, D., *Nature Struct. Biol.*, **7**, 669(2000).
12. Scheinerman, F. B. and Brooks III, C. L., *Proteins*, **29**, 193(1997).