

< 논문 >

## 최적설계 기법을 이용한 단백질 3차원 구조 예측

정민중\* · 이준성†

(2006년 2월 6일 접수, 2006년 4월 24일 심사완료)

### Prediction of Protein Tertiary Structure Based on Optimization Design

Min-Joong Jeong and Joon-Seong Lee

**Key Words:** Optimization(최적화), Structure Design(구조설계), Protein(단백질), Fragment Assembly(단편조립법), 3-Residue Fragment(3-잔기아미노산), Folding Optimization(접힘최적화)

#### Abstract

Many researchers are developing computational prediction methods for protein tertiary structures to get much more information of protein. These methods are very attractive on the aspects of breaking technologies of computer hardware and simulation software. One of the computational methods for the prediction is a fragment assembly method which shows good *ab initio* predictions at several cases. There are many barriers, however, in conventional fragment assembly methods. Argues on protein energy functions and global optimization to predict the structures are in progress for example. In this study, a new prediction method for protein structures is proposed. The proposed method mainly consists of two parts. The first one is a fragment assembly which uses very shot fragments of representative proteins and produces a prototype of a given sequence query of amino acids. The second one is a global optimization which folds the prototype and makes the only protein structure. The goodness of the proposed method is shown through numerical experiments.

#### 1. 서론

주어진 아미노산 배열에서 단백질의 3차원 구조를 해명하는 것은 현대과학의 커다란 명제중의 하나이다. 이는 단백질의 구조를 알고 이러한 구조가 형성되는 원리를 아는 것은 단백질의 기능을 이해하기 위한 가장 큰 요소이기 때문이다. 또한 이러한 구조예측은 장래 새로운 신약 제제와 생체의 신물질 창조에 직접적인 관련을 갖는다.<sup>(1-4)</sup>

이미 X-ray와 NMR(Nuclear Magnetic Resonance)과 같은 실험적인 방법을 통해 단백질의 3차원 구조를 파악하는 것은 상당한 수준의 정확도가 확보되었고 널리 사용되고 있다. 그러나 이러한 실험적 방법은 한 개의 단백질 구조를 해석하기 위한 실험시간이 짧게는 수개월에서 길게는 수년이 걸리는 점과 해석이 필요한 수만 종류의 단백질 수를 고려할 경우, 매우 제한적인 방법이다.

위와 같은 문제점을 해결하기 위하여 현재 많은 연구자들이 컴퓨터를 이용하여 신속히 단백질의 3차원 구조를 규명하고자 하고 있다.<sup>(2-8)</sup> 이러한 방법은 특히 나날이 발전하고 있는 컴퓨터 하드웨어 및 각종 시뮬레이션 기법의 발전을 고려할 경우 매우 매력적인 방법이다. 컴퓨터를 이용하여 아미노산 배열로부터 구조예측을 수행하는

† 책임저자, 회원, 경기대학교 기계시스템공학부  
E-mail : jslee1@kyonggi.ac.kr  
TEL : (031)249-9813 FAX : (031)249-9796

\* 한국과학기술정보연구원 e-Science 응용연구팀

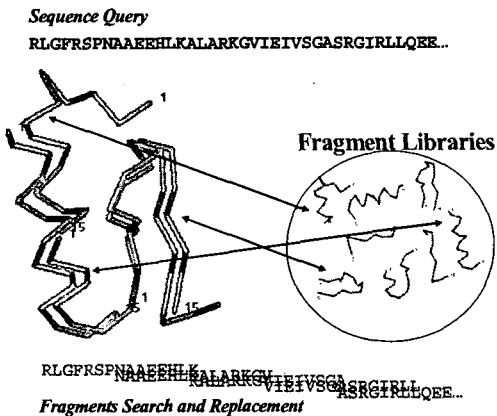


Fig. 1 A conventional fragment assembly method

대표적인 방법으로는 단편조립법(Fragment Assembly Method)<sup>(4,6,9,10)</sup>이 있다. 이 방법은 실험적으로 3차원 분자구조가 공개된 단백질을 일정한 크기로 분해하여 단편을 만들고 이를 이용하여 가능한 아미노산 배열에 대한 3차원 분자구조가 저장된 단편 라이브러리를 준비한다. 이후 예측하고 싶은 아미노산의 배열에 대하여 부분적으로 준비된 단편 라이브러리를 탐색하여 초기구조를 만든다. 만들어진 구조의 단백질 에너지 방정식을 점검하고, 에너지 방정식이 최적화 될 때까지 단편 라이브러리에 준비되어 있는 단편을 교환해 가며 최종 예측구조를 찾는다. Fig. 1에 전형적인 단편조립법의 예를 나타내었다.

이러한 단편조립법은 현재까지의 처음으로부터 (*ab initio*) 구조예측방법 중에 가장 우수한 방법으로 알려져 있다. 그러나 이 방법은 예측의 효율성이 탐색을 해야 할 단편 라이브러리 자체의 정확도에 따라 상당한 영향을 받는다. 따라서 많은 종류의 아미노산 배열에 대응할 수 있도록, 기존에 알려진 단백질 구조에 대한 복잡한 구조 유사성 검색 및 이를 통해 만들어진 정교한 단편 라이브러리를 필요로 한다. 또한 단편을 교환하며 구조를 결정하므로 선택된 단편의 적합성을 평가하기 위한 복잡한 에너지 방정식과 이를 최적화하는 강력한 최적화 알고리즘이 필요하게 된다. 그리고 방법의 특성상 라이브러리에 유사한 형태의 분자구조가 없을 경우 진정한 의미의 *ab initio* 예측이 어려운 문제점이 존재한다.

본 논문에서는 새로운 형태의 단편조립과 이를 이용한 구조예측의 방법을 제시한다. 제안된 방

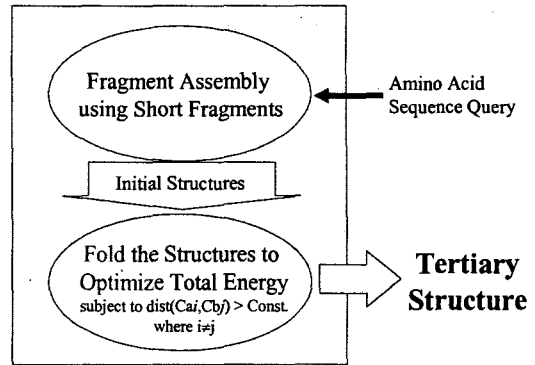


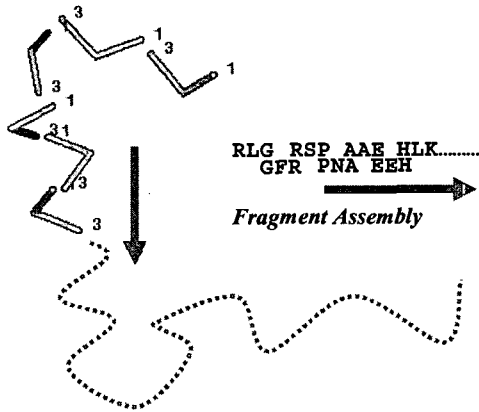
Fig. 2 Proposed protein structure prediction method

법은 기존의 단편조립법에 널리 사용되는 15개 아미노산 정도의 단편<sup>(9)</sup>이 아닌 매우 짧은 단편을 사용한다. 이러한 짧은 단편은 상대적으로 긴 단편에 비하여 단편을 만들기 위한 구조 유사성 검색 등의 복잡한 과정이 생략되게 하는 장점이 있다. 제안된 방법에서는 3-아미노산(3-residue) 단편들을 이용하여 주어진 아미노산 배열의 요구에 따라 초기구조를 만든다. 이러한 초기구조는 추가적인 단편 교환은 필요 없다. 마지막으로 구해진 초기구조에 대하여 단백질 접힘의 최적화 과정을 수행하여 최종적인 예측구조가 제안되게 된다. Fig. 2에 제안된 예측방법의 전체구성을 나타내었다.

2장에서는 본 논문에서 제안된 아미노산 단편조립의 내용을 설명하고, 3장에서는 단편조립에서 구해진 초기구조에 대한 접힘 최적화를 기술한다. 또한 접힘 최적화의 유효성을 검토하기 위한 계산실험의 결과를 나타낸다. 4장에서는 제안된 방법을 이용하여 단백질 *ab initio* 구조예측의 실제 예를 설명한다. 마지막으로 5장에서는 본 연구에서 얻어진 결론과 향후 연구방향을 제시한다.

## 2. 아미노산 단편조립

본 연구에 사용된 단편은 3개 아미노산의 고정된 아미노산 길이를 갖고 있으며 Fig. 3과 같이 아미노산 배열의 입력에 따라 3개 아미노산씩 연결해 가며 이때 전후 1개의 아미노산은 중복되게 된다.



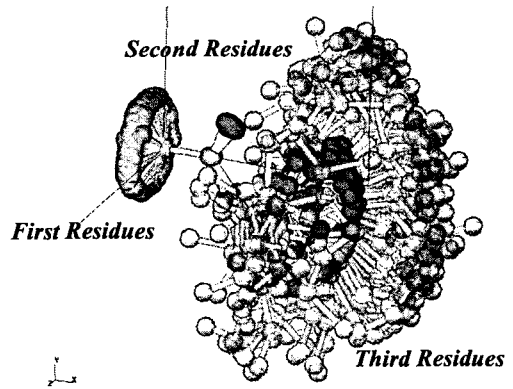
**Fig. 3** Proposed fragment assembly using 3-residue fragments

**2.1** 제안된 3-아미노산 단편

사용된 단편은 PDB-REPRDB(11)라고 하는 대표적인 단백질을 검색해 주는 데이터베이스를 이용하였다. PDB-REPRDB에 2004년 6월까지 등록된 단백질을 대상으로 4167개의 대표적인 단백질을 찾았다. 검색된 단백질들을 가능한 모든 3-아미노산의 아미노산 배열에 대하여 분해한 결과 1,267,745개의 단편이 형성되었다. 예를 들어 Alanine-Alanine-Alanine의 배열에 대하여 총 1,552개의 단편이, Tyrosine-Tyrosine-Tyrosine에 대해서는 3개의 단편이 형성되었다. 가장 많은 단편을 갖는 3-아미노산 배열은 3,237개의 Arginine-Tyrosine-Proline이고 Cystine-Methionine-Cystine의 배열은 단편이 형성되지 않았다. 구해진 모든 단편의 3차원 분자구조는 첫 번째 아미노산의 알파 탄소( $C_{\alpha}$ )를 원점으로 하여  $C_{\alpha}$ -C-O 평면이 일치하도록 모두 좌표 변환하여 저장하였다. Fig. 4에 본 연구에 사용된 일부 단편의 시각화를 나타내었다. 본 연구에서 사용된 단편에는  $N-C_{\alpha}$ -C-O(backbone) 및 베타탄소( $C_{\beta}$ )의 3차원 구조정보가 저장되었다.

**2.2** 단편을 이용한 초기구조 형성 방법

단백질의 아미노산 배열에 대한 3차원 분자구조를 형성하기 위해서는 동일한 3-아미노산 배열에 대하여 구조적으로 다양한 단편들 중에서 어떤 것을 선택하여 사용할 것인가에 대한 정의가 필요하다. 본 연구에서는 원자들의 기하학적 거리만을 이용하는 결합오차(coupling error)를 정의



**Fig. 4** Ball-and-stick visualization of 3-residue fragments

하였고, 이를 이용하여 아미노산 배열에 대한 다수의 단편으로부터 적절한 단편을 선택할 수 있었다. 단편 선택과 이를 이용한 초기 3차원 구조 형성의 자세한 내용은 다음과 같다.

구조를 형성하는 배열의 처음부분에 대한 임의의 단편을 선택한다. 이 단편에 대하여 뒤따르는 배열에 해당하는 많은 단편이 존재하게 된다. Fig. 5와 같이 이중 하나를 회전시키고 이동시켜 Fig. 5(d)와 같이 앞에 오는 단편의 3번째 아미노산과 뒤에 오는 첫 번째 아미노산의  $C_{\alpha}$ -C-O 평면을 일치시킨다. 이때 중복되는 두개의  $C_{\alpha}$ 는 같은 좌표를 갖고, 두개의 C는 동일 선상에 위치한다. 그리고 두개의 O(산소원자)는 같은 평면상에 위치한다. 이때 중복되는 아미노산의 두개의 N(질소원자)의 상호간의 거리를 측정하여 이를 결합오차로 정의하였다. 따라서 가능한 모든 단편에 대하여 모든 결합오차를 측정하고 이중 가장 작은 결합오차 값을 보이면서 원자간 충돌이 없는 단편을 선택하여 연결하였다. 이때 단편의 조립을 시작하는 처음 단편의 선택은 참고가 될 커플링오차를 구할 수 없으므로 고려 가능한 모든 단편을 모두 선택하였고, 조립의 처음 시작은 아미노산 배열의 첫 번째부터  $n$ 번째까지 모두에 걸쳐서 실시하였다. 제안된 단편조립에 따라서 100여개의 아미노산을 갖는 단백질에 대하여 수천 개의 초기구조가 형성되게 된다.

**3. 구조 접힘(Folding) 최적화**

제안된 단편조립에 의해서 만들어지는 초기구

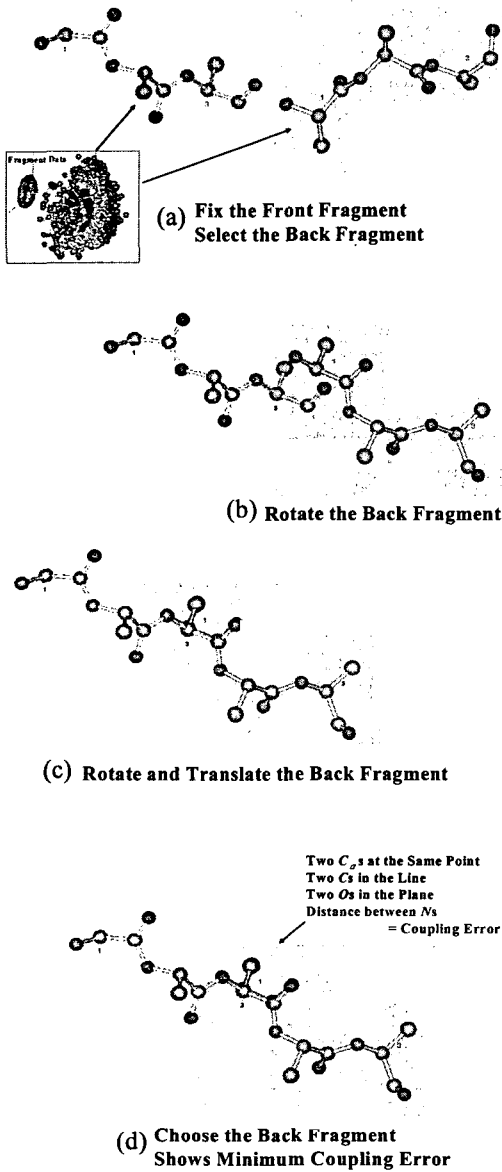


Fig. 5 New fragment assembly and coupling error

조는 각 단편 상호간의 국부적인 결합오차만을 고려하였으므로 전체적인 형태가 자연의 단백질 3차원 구조와는 달리 Fig. 6의 좌측그림과 같이 구조적으로 풀어진(unfolded) 형태를 가질 수 있다. 본 연구에서는 자연 상태의 단백질 구조처럼 예측된 단백질 3차원 구조의 내부 원자들이 매우 작은 공간에 밀집되도록 접힘 최적화를 고려하였다. 따라서 최종 예측구조를 만들기 위해서 단백질 접힘 상태에 대한 에너지 방정식을 구성하고,

이 에너지 방정식을 최소화시키는 구조를 찾는 최적알고리즘을 구성하였다.

### 3.1 접힘을 위한 에너지 방정식의 구성

단편조립에 의한 초기구조를 자연 상태의 단백질 3차원 분자구조와 유사하게 좁은 공간으로 접기 위한 전체 에너지 방정식  $E_{total}$ 은 식 (1)과 같은 4가지 에너지 항목으로 구성되었다.

$$E_{total} = E_{hydrophobic} + E_{volume} + E_{hydrobond} + E_{waals} \quad (1)$$

$E_{hydrophobic}$ 은 아미노산의  $C_{\alpha}$ 들 상호간의 거리에 Kyte와 Doolittle의 소수성 지수<sup>(12)</sup>를 가산하여 그 총합으로 구한다. 이 항목은 소수성(hydrophobic) 아미노산이 단백질의 중심으로 움직이고 친수성(hydrophilic) 아미노산은 단백질의 표면으로 움직이는 기능을 한다.

$E_{volume}$ 은 전체  $C_{\alpha}$ 간의 거리들 중 가장 먼 거리를 정육면체의 한 변의 길이로 고려하고 이때의 정육면체 체적을 계산하여 사용하였다. 따라서 이 항목은 구해진 3차원 구조가 작은 체적을 갖도록 하여 접힘이 잘 되는데 큰 영향을 준다.

$E_{hydrobond}$ 는 수소결합의 개수를 음의 값으로 나타내며 수소결합이 많아질수록 전체 에너지 값은 적어지도록 정의되었다. 이러한 수소결합의 항목에 의하여 계산되는 단백질 구조는 자연의 구조처럼 수소결합에 의한 독특한 나선구조 및 판구조<sup>(1)</sup>를 갖게 된다. 또한 이 수소결합의 작용으로 전체 에너지 방정식  $E_{total}$ 은 불연속적인 함수공간을 갖게 된다.

$E_{waals}$ 는  $C_{\alpha}$ 와 C간의 Van der Waals 에너지의 총합이며, 이 항목에 의하여 과도한 접힘에 의한 원자간의 충돌을 제거하였다.

본 연구에서 사용된 접힘을 위한 에너지 함수는 비선형적이고  $E_{hydrobond}$ 에서 보이는 것과 같이 불연속적인 함수공간을 구성한다. 이러한 복잡한 함수공간에서 최적 접힘구조를 찾기 위해 최적알고리즘을 개발하였다. 에너지 방정식에 대한 변수는 각 아미노산의  $N-C_{\alpha}$ ,  $C_{\alpha}-C$  두 결합축의 0에서 360°의 회전각으로 정의된다.

### 3.2 접힘을 위한 최적알고리즘

본 연구에서는 진화전략법(Evolutionary Strategy)<sup>(13,14)</sup>

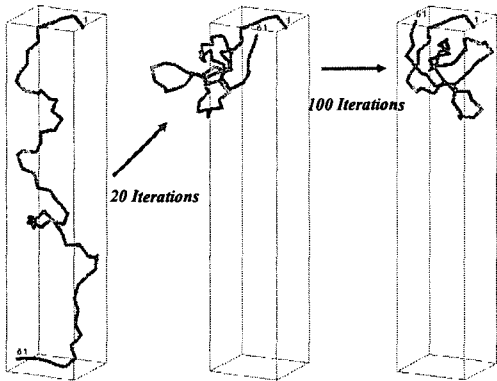


Fig. 6 Numerical experiment of proposed folding optimization

에 기초하여 교배, 돌연변이, 선택의 과정을 거치고, 반복적으로 최적화 구조를 찾게 된다. 그러나 함수공간의 복잡성으로 인하여 국부최적화(local optimum) 지점에 수렴되어 최적화 진행이 안 될 수 있다. 따라서 이러한 국부최적화를 최대한 방지하기 위해서 면역알고리즘(Immune Algorithm)<sup>(15)</sup>의 기억세포(Memory Cell)를 도입하였다. 기억세포는 최적화 과정이 일정 수 이상 반복해서 같은 함수공간을 찾을 경우, 즉 국부최적화에 빠졌을 경우 찾아진 함수공간을 저장하는 기능을 한다. 개발된 최적알고리즘은 기억세포를 사용하여 기억세포에 저장된 공간을 제외한 나머지 공간을 탐색한다. 이러한 방법으로 전체적인 최적구조를 탐색하도록 하였다.

### 3.3 접힘 최적화의 계산실험 및 검토

본 연구의 접힘 최적화의 유효성을 확인하기 위하여 Fig. 6과 7과 같은 접힘의 계산실험을 하였다. 이 실험에 사용된 접힘 최적화의 실험모델은 자연구조의 단백질 3차원 분자구조에서 일부 아미노산의  $\Phi(N-C_{\alpha}$  결합축의 회전각)와  $\Psi(C_{\alpha}-C$  결합축의 회전각)를 임의로 변형시켜서 전체 3차원 구조를 펼친 형태이다. Fig. 6 좌측의 경우는 총 5개의 아미노산을 선택하고 펼친 경우이나, 100번의 반복을 통하여 Fig. 6의 우측에서 보이는 것처럼 접혀서 자연 상태에 매우 가까운 구조로 돌아감을 알 수 있다. 이때  $C_{\alpha}$ 에 관하여 자연 상태의 3차원 구조와 접힘 최적화 구조의 Root Mean Square Deviation(RMSD)을 구해보면 2.12Å으로 거의 동일한 3차원 구조임을 알 수 있다.

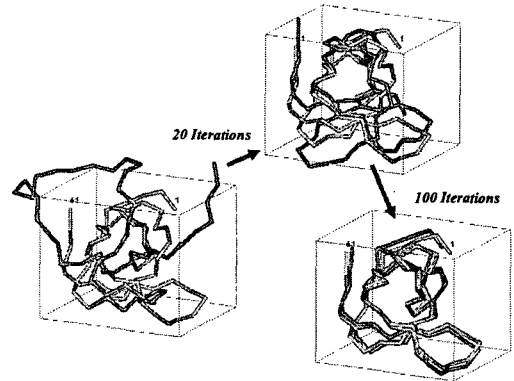


Fig. 7 Superposition of folded and native structures in optimization iteration

Fig. 7은 접힘 계산실험의 다른 예로서 자연 상태의 구조와 함께 접힘이 되는 구조를 중첩하여 나타내었다. 여기서도 Fig. 6과 같이 100번의 반복을 통하여 자연 상태에 매우 유사한 구조를 찾아감을 알 수 있다. 이러한 계산실험을 통하여 정의된 접힘을 위한 에너지 방정식과 이를 최적화하는 최적알고리즘의 타당성이 검증되었다.

## 4. 단백질 구조예측

개발된 접힘 최적화를 이용하여 *ab initio* 예측 실험을 실시하였다. *Ab initio* 예측의 계산실험을 위하여 3개의 나선구조와 2개의 평판구조를 갖는 72개의 아미노산으로 구성된 단백질의 아미노산 배열을 사용하였다. 이 단백질의 Protein Data Bank(PDB)<sup>(16)</sup> ID는 1LEB이다.

아미노산 배열만을 입력하여 제안된 단편조립에 의해 구해진 초기구조에서 자연 상태의 3차원 구조와 거의 동일한 구조가 생성되었다. 이는 사용된 단편 라이브러리에 1LEB와 3차원 분자구조가 유사한 PDB ID 1JHF의 단편이 존재하였고 제안된 단편조립법이 이를 사용하여 구조를 생성한 것이 원인으로 확인되었다. 따라서 제안된 단편조립만으로도, 만일 구조적으로 유사한 단백질 구조가 단편 라이브러리에 있을 경우, 상당히 정확한 구조를 생성할 수 있음을 알 수 있다.

제안된 구조예측법이 *ab initio* 예측을 할 수 있는지 검증하기 위해서 1LEB 자신을 포함한 모든 관련된 단백질을 공개 단백질 검색엔진인 SCOP<sup>(17)</sup>에서 검사하였다. 본 연구에 사용된 단편

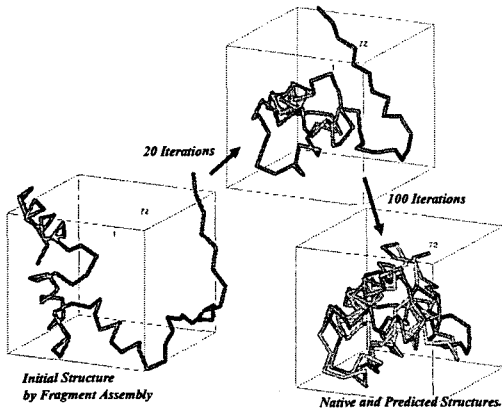


Fig. 8 Folding optimization to initial structure and structural comparison of native and predicted structures

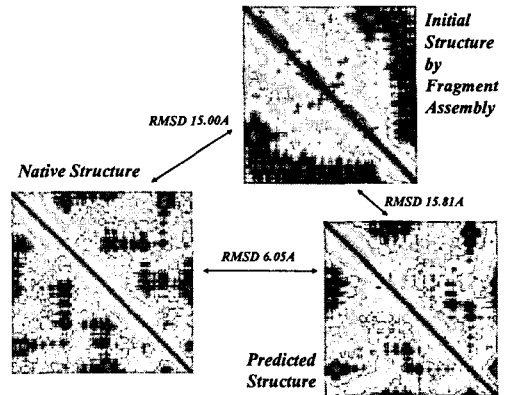


Fig. 10  $C_{\alpha}$  contact maps of native, initial, and finally predicted structure with RMSD

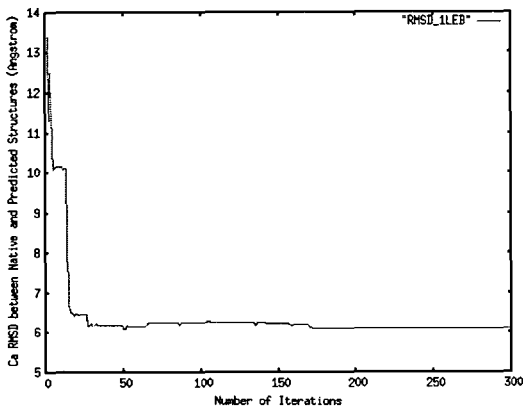


Fig. 9 Convergence property of protein folding optimization

라이브러리로부터 SCOP에서 검사된 단백질과 관련이 있는 모든 단편들을 제거 하였다. 이렇게 정리된 단편 라이브러리를 사용하여 다시 단편조합을 이용한 초기구조 형성을 실시하였다.

총 6,111개의 초기구조가 형성되었고 이들의 아미노산 배열에 대한 나선구조 위치를 각각 파악하였다. 그 결과 가장 빈도가 높은 나선구조의 위치를 갖는 3차원 구조를 접힘 최적화의 초기구조로 선택하였다. 구해진 초기구조는 32개의 서로 다른 단백질로부터 만들어진 단편으로 구성되었다.

접힘이 된 최적구조를 찾기 위해서 구해진 초기구조에서 결합오차가 상대적으로 크고 분자구

조상 회전이 용이한 Glycine 아미노산을 회전의 대상으로 지정하였다. 4개의 Glycine 아미노산이 회전의 대상으로 지정되었고 이 아미노산들의  $\Phi$ 와  $\Psi$ 를 사용하여 3차원 구조를 회전시켜 탐색을 위한 개체군을 만들었다. 개체군의 크기는 반복과정에서 100으로 일정하게 유지되었고 최대 1,000번의 반복을 탐색조건으로 하였다. 100번 이상의 반복에서도 더 이상 최적구조의 변화가 없을 경우 최적알고리즘이 수렴하였다고 고려하였다.

Fig. 8에 초기구조와 자연구조를 중첩하여 표시하였다. 그림에서와 같이 반복수가 증가하면서 자연구조에 유사한 구조가 형성됨을 알 수 있었다. 접힘 최적화 과정의 수렴특성은 Fig. 9에 표시하였다. 51번의 반복에서 최적 접힘구조가 탐색되었으며 200번 이상의 반복에서 최적화 과정이 수렴됨을 알 수 있다. 최종적으로 얻어진 구조는 자연구조와 6.08Å의 RMSD를 보였다. 100 아미노산 이하의 구형 단백질의 경우 예측구조와 자연 상태의 3차원 구조의 RMSD가 약 6.5Å 이하의 경우, 예측이 잘 된 것으로 판단된다.<sup>(10,18)</sup> 따라서 Fig. 8의 3차원 예측구조는 우수한 결과로 고려된다. Fig. 10에는 자연구조와 초기구조 그리고 접힘 최적화에 의해 구해진 최종 예측구조의 아미노산  $C_{\alpha}$  상호간의 거리에 의한 Contact Map을 나타내었다. 그림과 같이 최종 예측구조는 아미노산  $C_{\alpha}$  상호간의 거리에 대한 패턴도 상당히 일치함을 알 수 있다.

5. 결 론

본 연구에서는 아미노산의 배열에 대한 단백질 3차원 분자구조 예측의 방법으로 새로운 단편조립과 접힘 최적화를 제안하였다. 제안된 단편조립은 단편간의 기하학적인 편차인 결합오차를 사용하여 구조를 형성하였다. 이러한 단편조립은 기존의 단편조립의 구조 유사성 탐색이 필요하지 않는다. 단편조립에 의하여 찾아진 초기구조는 접힘 최적화에 사용되며, 접힘 최적화를 통하여 최종 예측구조를 구할 수 있다. 접힘 최적화를 위하여 새로운 에너지 방정식과 최적알고리즘이 제안되었다. 제안된 예측방법을 사용하여 *ab initio* 예측을 실시한 결과, 예측된 3차원 구조와 자연 상태의 3차원 구조가 RMSD 6.08Å의 근사성을 나타내었다.

제안된 예측방법은 아직 개발단계에 있으며, 특히 단편조립으로부터 형성된 다수의 초기 구조에서 어떤 구조를 접힘의 후보로 선택할 것인지에 관한 고려가 필요하다. 현재 이러한 다수의 초기 구조를 한국과학기술정보연구원의 그리드컴퓨팅 자원<sup>(19)</sup>을 이용하여 모두 해석하려는 방법이 시도되고 있다. 또한 이러한 다수의 초기구조들을 클러스터 분석하여 복수의 후보를 선택하는 방법 역시 동시에 고려되고 있다. 접힘 최적화를 위한 회전대상이 되는 아미노산의 선택에서도 보다 다양한 조건이 검토되고 있다. 첫 번째 저자는 이러한 점들을 보완하기 위하여 다양한 종류의 예측문제를 대상으로 계산실험을 거듭하고 있다.

본 논문에서 제시된 방법은 아직 몇 가지 개선점을 가지고 있으나 본문에서 보여준 예측결과와 같이 처음으로부터의 예측의 새로운 방법으로 제시될 수 있을 것이다.

후 기

본 연구를 위하여 일본 산업기술총합연구소(AIST)의 AIST 슈퍼컴퓨터 사용에 도움을 주신 Yutaka Akiyama 박사께 감사드립니다. 또한 본 연구는 한국과학재단 목적기초연구(R01-2004-000-10469-0)지원으로 수행되었기에 관계자 여러분께 감사드립니다.

참고문헌

- (1) Branden, C. and Tooze, J., 1999, *Introduction to Protein Structure*, Garland Publishing, New York.
- (2) Moulton, J., Hubbard, T., Fidelis, K. and Pedersen, J.T., 1999, "Critical Assessment of Methods of Protein Structure Prediction (CASP): Round III," *Proteins*, Suppl. 3, pp. 2~6.
- (3) Moulton, J., Fidelis, K., Zemla, A. and Hubbard, T., 2001, "Critical Assessment of Methods of Protein Structure Prediction (CASP): Round IV," *Proteins*, Suppl.5, pp. 2~7.
- (4) Baker, D. and Sali, A., 2001, "Protein Structure Prediction and Structural Genomics," *Science*, Vol. 294, pp. 93~96.
- (5) Betancourt, M. and Skolnick, J., 2001, "Finding the Needle in a Haystack: Educing Native Folds from Ambiguous *ab initio* Protein Structure Predictions," *Journal of Computational Chemistry*, Vol. 22, pp. 339~353.
- (6) Bonneau, R., Tsai, J., Ruczinski, I., Chivian, D., Rohl, C., Strauss, C.E.M., and Baker, D., 2001, "Rosetta in CASP4: Progress in *ab initio* Protein Structure Prediction," *Proteins*, Suppl. 5, pp. 119~126.
- (7) Henikoff, S. and Henikoff, J., 1992, "Amino Acid Substitution Matrices from Protein Blocks," *Proceedings of the National Academy of Sciences*, Vol. 89, pp. 10915~10919.
- (8) Jones, D.T., 1999, "Protein Secondary Structure Prediction Based on Position-specific Scoring Matrices," *Journal of Molecular Biology*, Vol. 292, pp. 195~202.
- (9) Julian Lee, J., Kim, S.Y., Joo, K., Kim, I. and Lee, J., 2004, "Prediction of Protein Tertiary Structure Using PROFESY, a Novel Method Based on Fragment Assembly and Conformational Space Annealing," *Proteins*, Vol. 56, pp. 704~714.
- (10) Ishida, T., Nishimura, T., Nozaki, M., Inoue, T., Terada, T., Nakamura, S. and Shimizu, K., 2003, "Development of an *ab initio* Protein Structure Prediction System ABLE," *Genome*

- Informatics*, Vol. 14, pp. 228~237.
- (11) Noguchi, T. and Akiyama, Y., 2004, "PDB-REPRDB," *Nucleic Acid Research*, Vol. 32, Online Summary Paper, <http://www3.oup.co.uk/nar/database/summary/277>.
- (12) Kyte, J. and Doolittle, R.F., 1982, "A Simple Method for Displaying the Hydrophobic Character of a Protein," *Journal of Molecular Biology*, Vol. 157, pp. 105~132.
- (13) Jeong M.J. and Lee J.S., 2005, "Shape Design of Passages for Turbine Blade Using Design Optimization System," *Trans. of the Korean Society of Mechanical Engineers*, Vol. 15, No. 3, pp. 1013~1021.
- (14) Jeong, M.J., 2003, *Integrated Support System for Decision-Making in Design Optimization*, Ph.D. Thesis, The University of Tokyo, December.
- (15) Castro L.N. and Timmis, J., 2002, *Artificial Immune Systems: A New Computational Intelligence Approach*, Springer.
- (16) PDB: Protein Data Bank, <http://www.rcsb.org/pdb/>
- (17) SCOP: Structural Classification of Proteins, <http://scop.mrc-lmb.cam.ac.uk/scop/>
- (18) Kihara, D., Kolniski, A., and Skolnick, J., 2001, "Touchstone: An *ab initio* Protein Structure Prediction Method that uses Threading-based Tertiary Restraints," *Proceedings of the National Academy of Sciences*, Vol. 98, pp. 10125~10130.
- (19) KISTI Supercomputing Center, <http://www.ksc.re.kr>