

소분자 도킹에서의 평가함수의 개발 동향

정환원¹ · 조승주^{2*}

Recent Development of Scoring Functions on Small Molecular Docking

Hwan Won Chung¹ and Seung Joo Cho^{2*}

Abstract

Molecular docking is a critical event which mostly forms Van der waals complex in molecular recognition. Since the majority of developed drugs are small molecules, docking them into proteins has been a prime concern in drug discovery community. Since the binding pose space is too vast to cover completely, many search algorithms such as genetic algorithm, Monte Carlo, simulated annealing, distance geometry have been developed. Proper evaluation of the quality of binding is an essential problem. Scoring functions derived from force fields handle the ligand binding prediction with the use of potential energies and sometimes in combination with solvation and entropy contributions. Knowledge-based scoring functions are based on atom pair potentials derived from structural databases. Forces and potentials are collected from known protein-ligand complexes to get a score for their binding affinities (e.g. PME). Empirical scoring functions are derived from training sets of protein-ligand complexes with determined affinity data. Because non of any single scoring function performs generally better than others, some other approaches have been tried. Although numerous scoring functions have been developed to locate the correct binding poses, it still remains a major hurdle to derive an accurate scoring function for general targets. Recently, consensus scoring functions and target specific scoring functions have been studied to overcome the current limitations.

Key words : Small Molecule, Search Algorithm, Docking, Scoring, targeted scoring function

1. 평가함수의 중요성

소분자 도킹은 생물학적인 관점에서 분자인지에 해당한다. 즉, 공유결합이 아니라 비공유결합을 이용하여 가역적으로 결합함으로써 소분자 도킹은 생명현상에서 중요한 분자인지에서 핵심적인 사건이다. 그러므로 분자간의 결합이 될 때, 정확한 결합구조와 결합력을 예측하는 것은 대단히 중요한 일이다. 결합방식은 대단히 많은 가능성이 존재하므로 또한 일반적으로 모든 가능성을 다 탐색하는 것은 불가능하기 때문에 효과적인

탐색방법이 필요하다. 보기에는 단순히 보이지만, 사실상 소분자 도킹 문제는 엄청나게 큰 자유도를 가지고 있는 대단히 복잡한 문제이다. 즉, 도킹문제는 지구상에서 가장 높은 봉우리를 찾는 문제와 유사하다. 평가함수는 각각의 지점에서의 고도계에 해당한다. 탐색알고리즘은 고도계의 정보를 활용한 여행 전략이다. 가장 높은 봉우리를 찾는 직접적인 방법이 없다고 가정하면, 주로 높은 봉우리를 골고루 많이 찾을 수 있는 방법이 필요하다. 효과적인 탐색의 핵심은 지구상의 모든 표면을 골고루 찾을 수 있는 전역탐색과 히말라야 부근을 좀 더 찾을 수 있는 국부탐색을 적절히 활용하는 것이다. 탐색방법으로는 유전 알고리즘, 몬테 카를로, 분자동역학, 타부 탐색 등의 다양한 방법들이 있으며, 최근에는 STUN, PSO 등과 같은 새로운 알고리즘도 도킹에 도입되고 있다. 이러한 탐색알고리즘에 관한 논의는 지난번 총설에 대략적으로 기술하였다.^[1] 이러한 탐색을 가능하게 하는 것은 역시 산의 높이를 측정할 수 있

¹한국과학기술연구원 계산과학연구소 (Computational Science Center, Future Fusion Technology Division, Korea Institute of Science and Technology, P.O. Box 131, Cheongryang, Seoul 130-650, South Korea)

²조선대학교 의과대학 세포분자 연구센터 (Department of Cellular and Molecular Medicine & Center for Resistance Cells, College of Medicine, Chosun University, 375 Seosuk-dong, Dong-gu Gwangju 501-759, Korea)

*Corresponding author: chosj@chosun.ac.kr

(Received : January 15, 2010, Accepted : March 20, 2010)

는 성능이 좋은 고도계이며 도킹문제에서 이것은 평가 함수에 해당한다. 그런데 도킹문제에서 소분자가 얼마나 잘 결합하는지를 평가하는 것, 즉, 결합의 자유에너지를 구하는 일은 쉽지 않으며, 본 총설에서는 자유에너지를 구하는 방법에 대한 소개를 하고자 한다.

2. 평가함수의 양자역학적 배경

무한히 빠른 컴퓨터가 있다면, 정확한 도킹 구조를 양자역학을 응용한 계산으로 얻을 수 있을 것이다. 그런데 양자계산법은 너무 많은 시간이 걸리기 때문에 현실적으로 사용이 불가능하다. 따라서 개발된 좀 간단한 모형이 force field라고 할 수 있다. 이는 분자 자체 또는 분자간의 복잡하고 애매모호한 양자역학적인 결과를 해석하여, 공유결합(bond)과 비공유결합의 개념을 도입함으로써, 간단한 함수형태로 나타낸 것이다. 공유결합에는 bond, angle, dihedral angle 등을 묘사하는 함수를 사용하며, 비공유결합에는 수소결합, Van der waal 결합, 정전기적인 인력을 묘사하는데 역시 간단한 함수식을 사용한다. 따라서 분자동력학 시뮬레이션에서 자주 쓰이는 force field는 각 항이 간단한 함수식으로 표현되는 식 (1)과 같은 골격을 가지고 있다.

$$\begin{aligned} \text{Total binding free energy} &= \text{공유결합} + \text{비공유결합} \\ &\bullet \text{공유결합} = \text{bond} + \text{angle} + \text{dihedral angle} \\ &\bullet \text{비공유결합} = \text{van der Waals} + \text{electrostatic interaction} \end{aligned} \quad (1)$$

즉, 이것은 양자역학적인 결과를 근사적인 함수형태로 나타낸 것이다. 분자내부의 상황을 묘사하는 공유결합부분은 상호작용에 미치는 영향이 적으며, 주로 소분자와 단백질의 결합력에 중요한 요인은 분자간의 비공유결합에 기인한 상호작용이다. 식 (1)은 원론적으로는 완벽하지만 실제 계산에 사용하기 위해서는 좀 더 많은 요인들을 고려해야 한다. 우선 대부분의 생물고분자들의 작용은 수용액 상에서 이루어지기 때문에 용매효과를 고려해야 한다. 용매효과를 묘사하기 위하여 자주 사용하는 모형은 실제 물 분자 알갱이들을 고려하는 방식보다는 물 분자를 일종의 연속체로 생각하는 근사적인 방식이다. 또 식 (1)의 정확도를 높이기 위해서는 분자의 편극효과를 고려하는 경우도 있다. 식 (1)을 뼈대로 하여 더 정확한 에너지를 구하는 연구는 많이 진행되고 있으나, 계산속도가 느린 것이 단점이다. 용매효과를 보정하는 문제는 다음 단락에서 소개하는 평가 함수들에서 주로 간접적으로 보정되는 경우가 많다.

소분자와 단백질간의 결합은 평형상수를 구하는 문제이고 이 경우 분자 내부적으로는 결합에 따르는 strain이 발생하고 상호작용과 관련한 측면에서는 configurational entropy와 관련한 자유에너지의 변화가 생긴다. Strain은 직접에너지 계산을 하고, conformational entropy의 경우에는 소분자의 단일결합의 수를 계산하여 근사적으로 보정해 주게 된다. 지금까지의 대부분의 평가함수는 모두 단백질을 움직이지 않는 물체로 가정한 것이 대부분이다. 따라서, 이러한 자유에너지의 보정은 대부분의 경우 소분자(리간드)에 해당하며 단백질의 구조가 변할 때, 어떤 식으로 자유에너지를 보정하는지에 대한 연구는 별로 이루어진 바 없다. 또한 모든 평가함수는 각각의 부분의 합으로 이루어진다는 것을 가정하므로 소분자의 크기가 증가하면 결합력이 과도하게 계산되는 경향이 있다. 즉, 실제로는 소분자의 크기가 커질 때, 결합력이 작아질 수도 있지만 평가 함수에서는 이러한 효과를 묘사하기가 쉽지 않다.

3. 평가함수의 개발

어차피, 정확한 결합에너지를 얻는 것은 어렵고, 가상탐색의 경우에는 속도가 중요하므로, 결합의 자유에너지를 빨리 계산하기 위하여 근사적인 방법을 활용하게 된다. force field를 이용하는 것은 물리적인 법칙을 이용하여 원리적으로 결합에너지를 얻으려는 시도이다. 이에 반하여 empirical scoring function이나, knowledge-based scoring function들은 물리법칙에 의존하는 입장을 다소 벗어나 실험 데이터를 활용하는 보다 응용에 치중한 평가함수를 파라메타 fitting을 통하여 얻는다.

3.1. 기본 원리를 이용한 평가함수 (force-field based scoring functions)

물리적인 기본원리를 활용한 방법들은 결국 식 (1)에서와 같은 근사적인 평가함수를 활용하게 된다. 원론적으로는 이러한 방식이 가장 바람직하다. 더욱이 이것이 여타의 방법론과 다른 점은 실험값을 통해서 얻은 것이 아니므로 체계적인 보완이 가능하다는 점이다. 대부분의 경우에 식 (1)과 같은 기체상태의 분자역학적인 에너지를 구하고, 용매효과와 엔트로피 효과등을 보완하는 방식을 취하고 있다. 실질적으로는 엔트로피효과는 그 크기가 작기 때문에 용매효과를 적절히 보정해주는 것이 가장 중요한 과제로 남아있다.

사실은 solvation을 정확하게 계산하는 것이 어렵기 때문에, 제일원리를 이용한 평가함수는 실질적으로 정확도가 떨어지는 경향이 있다.^[24] 더욱이 짧은 거리에

서의 강한 반발력 때문에 도킹할 때 탐색과정에서 Van der Waals 힘을 작게 만들어서 사용한다. 또 하나의 단점은 이 평가함수는 정전기적인 인력을 과대평가하는 경향이 있다. 이러한 경향을 보완하기 위하여 desolvation term을 도입한 경우도 있다.^[5,6] 특히, free energy perturbation^[2]이나, linear response approximation^[3] 등은 원론적으로는 상당히 정확한 계산결과를 줄 수 있으나, 실질적으로는 force field의 불완전성, charge를 계산하는 문제, sampling을 충분히 할 수 없는 등의 문제들이 있어서 기대만큼 별로 정확하지 않다.

3.2. 실험 data를 활용한 평가함수 (empirical scoring functions)

제일원리를 이용한 평가함수의 단점을 보완하기 위하여 실험결과를 활용하여 만든 평가함수이다. 이의 기본적인 가정은 결합 자유에너지가 식 (2)와 같이 몇 가지의 기본적인 요소의 합으로서 표현될 수 있다고 본다. 실제로 평가함수를 유도하는 과정은 실험 data와 결합 pose를 가지고 각각의 항을 구하는 방법을 가정하고 파라메타를 fitting하게 된다.

$$\text{결합 자유에너지} = \text{motion} + \text{interaction} + \text{desolvation} + \text{configuration} \quad (2)$$

실험값을 얻을 수 있는 대부분의 결합들은 모두 결합을 하고 있는 화합물들이므로 최근에 나오는 평가함수들은 반발항들을 고려하고 있다. 반발력, 수소결합, 소수성 효과들을 적절하게 안배하는 것이 가장 중요한 과제이다. 수소결합을 정의하는 과정에서도 단순히 heavy atom 사이의 거리를 고려하는 방법과 중간에 끼인 수소를 포함한 각도를 고려한 방법이 있다. 이렇게 임의적이고, 엄밀하지 않은 유도과정에도 불구하고 empirical 평가함수는 경험적으로 비교적 잘 맞는 것으로 받아들여지고 있다. 약점으로는 알려진 dataset이 한정되어 있어서 일반적으로 잘 맞는 평가함수를 구하기에는 힘들다고 생각된다. 실제로 잘 맞는다고 알려지는 것은 이의 validation과정이 작은 dataset에 한정되어 있기 때문으로 여겨진다. 앞으로도 이러한 dataset과 관련된 문제는 풀기가 쉽지 않을 것으로 생각된다.

3.3. 구조 data를 활용한 평가함수 (knowledge based scoring functions)

empirical 평가함수와 유사하지만, knowledge based 평가함수는 결합에너지에 대한 정보없이 구조적인 정보만을 통계처리하여 얻는다. 기본가정은 자주 생기는

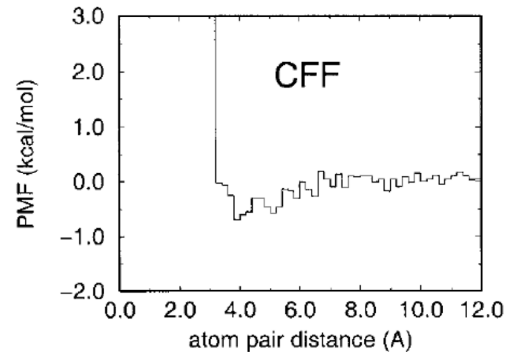


그림 1. 거리에 따른 계단식 평가함수.

원자간의 거리는 이것이 선호되는 물리적인 이유가 있다는 것이다. 이것을 유도하기 위해서는 일련의 일반적인 소분자와 단백질의 원자형(atom type)을 정의하고 모든 가능한 원자쌍의 거리의 분포를 구한다. 정규화된 분포 $[g(r)]$ 에서 평가함수를 유도할 수 있다.

$$w(r) = -k_B T \ln[g(r)] \quad (3)$$

구조 database에서 $g(r)$ 을 얻기 때문에 보통 하나의 score는 거리에 따라 달라지는 수백개의 작은 조각으로 이루어진다(그림 1). 따라서 이것을 체계적으로 개선하는 것은 거의 불가능하다. 그럼에도 단백질 database (PDB, <http://www.rcsb.org>)에서 보고되는 구조의 개수가 해마다 늘고 있기 때문에 이 평가함수는 앞으로도 개선될 여지가 크다고 할 수 있다.^[7]

4. 평가함수의 평가

평가함수의 평가문제는 대단히 중요한 문제이기 때문에 많은 연구가 되어왔다. 평가함수를 평가하는 것은 일견 쉬워보이지만 사실상 많은 어려움이 존재한다. 우선 그 목표가 다양하다. 평가함수는 주로 신약개발등에서 활용하기 위하여 개발하므로 실용적인 관점에서 보면 크게 다음의 3가지 목표가 존재한다.

- 구조예측 - 소분자와 단백질의 결합 pose
- 가상검색 - 활성이 있는 소분자 선별
- 친화도예측 - 활성의 정도를 서열화

구조예측의 문제는 가장 본질적인 문제이고 가상검색은 선도물질 도출과정에서, 친화도예측은 선도물질의 최적화과정에서 중요하다. 다양한 목적이외에도 평

가함수의 성능자체가 도킹 프로그램과 독립적이지 않은 점도 평가함수 평가의 어려움에 기여하고 있다.

4.1. 평가함수를 평가하는데 관련된 문제들

위에서 제시한 3가지 목표 중에서 가장 본질적이라고 생각할 수 있는 구조예측에 관한 평가를 하려면 일단 도킹 프로그램을 사용하여 결과를 얻어야 한다. 이 과정에서 몇 가지 고려사항이 있다. 우선, dataset을 어떻게 설정하는가 하는 문제이다. 구조예측의 영역에서는 몇 가지의 test set 이 보고되어 있다. Astex사에서 305개의 test set을 제안하였고,^[8] Wang 등은 100개의 구조를 제안하였다.^[9] 이러한 test set을 사용하여 구조예측을 한 다음, 실질적으로 정확히 일치하는 구조는 얻을 수 없으므로 “어느정도까지를 정확한 예측으로 볼 수 있는가” 하는 문제이다. 보통 RMSD 2Å 이내의 구조를 정확한 것으로 본다. x-ray의 해상도와 관련하여 이러한 기준은 타당한 것으로 생각되지만, 좁은 주관적인 사항을 중요시하는 사람들도 있다. 실제 도킹과 관련해서 x-ray 구조에는 수소원자가 보이지 않으므로 이 수소가 실질적으로 어떤 배향을 하고 있는지, 또는 acid의 경우에는 실제 수소원자의 존재여부, C, N, O 등의 heavy atom 등은 서로 구별이 되지 않으므로 imidazole 등의 실제 배향문제 등을 결정하는 과정이다소 임의적인 성격을 갖게 된다. 즉, 평가함수와 직접 관련이 없는 단백질의 활성자리를 도킹 프로그램에 맞게 손질하는 과정이 도킹의 성능에 영향을 준다. 이러한 문제는 논리적으로 엄밀하게 해결하기는 매우 어렵다. 더구나, 최근까지의 많은 연구의 결론은 특정한 상황에 맞지만, 일반적인 상황으로 받아들이기 어렵다. 왜냐하면, 많은 경우, 작고 임의적인 dataset을 사용하거나, 각각 다른 프로그램을 사용하고, 또는 active site의 설정이 임의적이기 때문이다. 이렇게 평가함수 외적인 요소를 배제하고 평가함수를 평가한 연구가 최근에 Wang등에 의해서 수행되었다. 11개의 평가함수를 평가한 객관적인 방법으로 평가하였다. 먼저, binding pose를 최대한 다양하게 얻은 다음, 각각의 평가함수가 실험적으로 알려진 구조에 가장 높은 점수를 주는가? 하는 방식으로 평가하였다.^[10]

4.2. 평가함수들을 평가한 결과

Factor Xa의 응용한 결과를 보면 Warren등의 결과는 GOLD가 GLIDE 보다 좋은 가상검색결과를 보인 반면에 Chen등의 결과를 보면 반대의 결과가 나타난다. Wang 등의 결과를 보면 thymidine 카이네이즈의 경우에 가장 잘 맞는 2개의 평가함수 (PMF와 FlexX)가

estrogen receptor의 검색에서는 가장 나쁜 것으로 나타났다. 이것은 평가함수가 작용점(target)에 얼마나 많이 의존하는지를 잘 나타내고 있다. 이것은 여러 가지 복잡한 결과의 산물로서 평가함수를 평가하는 것이 쉽지 않음을 단적으로 보여준다. 평가함수의 개발 논리와 이러한 경향의 상관성을 아는 것은 상당히 중요한 일이다. 그런데 많은 결과에서 이러한 상관관계가 별로 뚜렷해 보이지 않는 점이 보다 바람직한 평가함수를 개발하는데 어려운 점이다. 최근의 한 결과를 보면 pairwise potential에 의거한 평가함수는 binding pose를 가장 잘 설명하는 것 같고, 가상검색이나 친화도예측에는 regression을 응용한 평가함수가 잘 맞는 것으로 보이는 경향이 있었다.

5. 미래의 평가함수

일반적으로 잘 맞는 평가함수는 아직 나오지 않았다. 신약개발의 관점에서 보면 아직도 평가함수의 성능은 그다지 만족스럽지 못하다. 문제는 평가함수의 성능을 증대시키려고 많은 노력을 해 왔으나, 그 구체적인 방법론이 마땅하지 않은 점이다. 본 총설에서 제시한 3가지 방법론, 즉, force field, empirical scoring function, knowledge-based scoring function 등의 표준적인 과정이나 이를 변형한 방법으로는 신약개발에서의 실질적인 필요성을 충족시킬 만큼의 좋은 평가함수가 나오기 어려울 것이다. 이것은 단백질의 성질이 다양하기 때문이기도 하다. consensus 평가함수가 조금이나마 단일 평가함수에 비하여 성능이 좋은 것도 이러한 단백질의 다양성에 기인한다. 일반적으로 잘 맞는 평가함수를 개발하기 위해서는, 지금까지의 일반적인 방법과는 다른 혁명적인 방법이 도입되어야 할 것으로 보인다. 일반적으로 좋은 평가함수를 개발하는 것은 어려우나, 비교적 좁은 범위에서 충분히 좋은 성능을 갖는 평가함수(targeted scoring function)의 개발은 어느 정도 가능해 보인다. 따라서 지금도 많은 연구가 이 방향으로 진행되고 있으며, 또한 많은 진전이 있어왔다. 소분자 도킹 분야에서 다른 하나의 커다란 난제는 단백질의 유연성을 적절히 고려하는 문제이다. 이렇게 단백질 구조의 변화가 있을 때, 평가함수를 어떻게 처리하느냐 하는 부분도 평가함수분야에서 개척되어야 할 영역으로 남아있다.

감사의 글

이 연구는 교육과학기술부 · 한국연구재단 지원 산하

내성세포연구센터지원(R13-2003-009)으로 수행되었음.

참고문헌

- [1] S. J. Cho and H. W. Chung "Recent Development of Search Algorithm on Small Molecule Docking" *Journal of the Chosun Natural Science*, 2, p1, 2009.
- [2] H. Chen, P. D. Lyne, G. Giordanetto, T. Lowell and J. Li "On Evaluating Molecular-Docking Methods for Pose Prediction and Enrichment Factors" *J. Chem. Inf. Model.*, 46, p401, 2006.
- [3] M. H. J. Seifert "Targeted scoring functions for virtual screening", *Drug Dis. Today*, 14, p562, 2009.
- [4] W. Rexiao, L. Luhua and W. Shaomeng, "Further development and validation of empirical scoring functions for structure-based binding affinity prediction" *J. Comput. Design*, 16, p11, 2002.
- [5] B. K. Shoichet, A. R. Leach and I. D. Kuntz "Ligand solvation in molecular docking" *proteins*, 1, p4, 1999.
- [6] B. Q. Wei, W. A. Baase, L. H. Weaver, B. W. Matthews and B. K. Shoichet "A modeling binding site for testing scoring functions in molecular docking" *J. Mol. Biol.*, 13, p322, 2002.
- [7] N. Moitessier, P. Englebienne, D. Lee, J. Lawandi and C. R. Corbell "Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go" *British Journal of Pharmacology*, 153, p57, 2008.
- [8] I. Muegge and Y. C. Martin "A general and fast scoring function for protein-ligand interactions: A simplified potential approach" *J. Med. Chem.* 42, p791, 1999.
- [9] J. W. M. Nissink, C. Murray, M. Chris, H. Mike, V. K. Marcel, J. C. Cole and T. Robin, "A new test set for validating predictions of protein-ligand interaction" *proteins*, 49, p457, 2005.
- [10] R. Wang, Y. Lu and S. Wang "Comparative evaluation of 11 scoring functions for molecular docking" *J. Med. Chem.*, 46, p2287, 2003