

Impact of input ligand conformations on protein-ligand docking performances

양진솔, 백민경, 석차옥 *

Department of Chemistry, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul, 151-742
Received Date; Accepted Date (to be inserted by the publisher after your manuscript is accepted)

대부분의 단백질-리간드 도킹 프로그램에서 리간드의 구조 유연성은 리간드의 회전 가능한 torsion angle들을 샘플링 하는데 국한된다. 따라서 도킹에 사용되는 초기 리간드 구조의 결합길이, 결합각, ring 구조 등에 따라 단백질-리간드 도킹의 성공여부가 달라질 수 있다. 실제 단백질-리간드 도킹 프로그램을 이용하여 단백질-리간드 상호작용을 연구하는 경우, 리간드의 구조를 임의의 구조 데이터베이스로부터 얻거나 다양한 리간드 3차원 구조 생성 프로그램 등을 이용하여 생성하게 된다. 따라서 리간드의 초기 구조가 단백질-리간드 도킹 프로그램의 성능에 어떤 영향을 주는지 살펴보는 것은 실제 단백질-리간드 도킹을 활용하는 경우에 단백질-리간드 도킹 프로그램의 성능이 어떤 것인지 알아볼 수 있다는 점에서 매우 중요하다. 본 연구에서는 리간드 분자의 초기 구조가 단백질-리간드 도킹에 미치는 영향을 알아보기 위해 30개의 단백질-리간드 복합체 표적에 대해 리간드의 초기 구조를 다양하게 생성하여 GalaxyDock으로 단백질-리간드 복합체 구조를 예측하였다. 도킹을 위한 리간드 분자를 만들 때 Ring library에서 여러 가능한 ring의 conformation을 가져오는 방법으로 리간드의 구조를 다양하게 만들었으며, 도킹 결과 한 가지 모델만 사용했을 때에 비해 도킹의 성공률이 70%에서 80%로 10% 증가한 것을 확인하였다. 또한 특히 구조적으로 유연한 ring이 리간드에 있는 경우 초기 구조가 도킹의 성공률에 큰 영향을 미치는 것을 확인할 수 있었다

핵심어: 단백질-리간드 도킹, 구조 유연성, 결합길이, 결합각

Introduction

단백질-리간드 도킹은 단백질과 이에 대응하는 작은 리간드 분자가 단백질의 결합자리에 결합하여 이루는 복합체의 구조를 예측하는 방법으로, 1980년대 처음 연구가 시작된 후 structure-based drug discovery 등에 유용하게 사용될 수 있는 방법으로 주목을 받으며 꾸준히 연구되고 있다.^{1,3}

단백질-리간드 복합체 구조를 예측하기 위해서는 두 가지 중요한 과정이 필요하다. 먼저 가능한

다양한 binding mode 들을 나타내는 거대한 conformational space 을 다양하게 샘플링 할 수 있어야 하고, 그 후에는 각각의 예측된 binding mode 들의 interaction energy 를 정확히 예측할 수 있어야 한다. 시중에 나와 있는 많은 단백질-리간드 도킹 프로그램들은 보통 torsion angle 에 변화를 주어 다양한 리간드 구조를 만들고 이를 scoring 하는 과정을 energy minimum 에 도달할 때까지 반복하여 리간드의 binding pose 를 예측하고 있다.²

앞서 언급한 바에서 알 수 있듯이 단백질-리간드 도킹에서는 가능한 리간드의 구조들을 잘 샘플링하는 것이 중요한데, 아무리 scoring 을 잘한다고 하더라도 실제와 들어맞는 binding pose 가

Impact of input conformation on docking performance

만들어지지 않는다면 결코 좋은 결과를 얻을 수 없기 때문이다.

처음 만들어진 DOCK 과 같은 도킹 알고리즘은 처음에 분자들을 rigid body 로 취급하며 출발하였다.¹ 이 방법은 분자들이 복합체를 이루면서 모양이 거의 변하지 않을 경우 성공적일 수 있지만 대부분의 리간드들이 구조적으로 유연하다는 점에서 명확한 한계를 가졌다. 이러한 한계를 극복하기 위해 현재 개발된 대부분의 단백질-리간드 도킹 프로그램들은 리간드의 결합길이, 결합각은 그대로 두고 torsion angle 만을 변화시키는 방법으로 리간드의 구조 유연성을 고려하고 있다.³ 이와 같은 방법은 결합길이나 결합각은 잘 변하지 않는 값이라는 점에서 어느 정도 합리적인 선택이라고 할 수 있으나, 실제 단백질-리간드 복합체에서 리간드의 구조가 이상적인 모습에서 벗어나 있는 경우가 많이 나타난다는 점에서 이러한 방식이 리간드 분자의 구조 유연성을 충분히 고려할 수 있는 지에 대한 검증이 필요하다. 특히 분자 안에 존재하는 ring 의 경우 가능한 구조간의 편차가 크며 단백질-리간드 도킹의 성능에 크게 영향을 미친다는 것이 알려져 있다. 본 연구에서는 리간드의 초기 구조가 단백질-리간드 도킹의 성능에 어떠한 영향을 미칠 수 있는지를 알아보고, 어떻게 하면 이를 고려하여 도킹의 성능을 향상시킬 수 있는지 그 방법을 제시하고자 한다.

리간드의 초기 구조가 단백질-리간드 도킹 성능에 어떤 영향을 미치는지 알아보기 위해 이미 구조가 알려져 있는 30 개의 단백질-리간드 복합체 구조에 대하여 다양한 리간드 초기 구조를 이용하여 단백질-리간드 도킹을 수행하였다. 리간드의 초기 구조는 단백질-리간드 복합체 구조에서 가져온 실제 결합구조와, 다양한 리간드 구조를 가지고 있는 PubChem database 에서 가져온 구조, 직접 제작한 in-house program 을 이용하여 1 차원 SMILES formula 를 3 차원 리간드 구조로 변환시킨 구조를 이용하였다. 이러한 다양한 리간드 초기구조를 이용하여 각각 단백질-리간드 도킹을 수행한 결과를 서로 비교하여 에너지와 실제 binding pose 와 예측한 것 사이의 RMSD 를 비교하여 보았다.

Getting Ligand 3D structure from protein-ligand complexes

도킹에 이용할 리간드의 초기 구조 중에 하나로 타겟으로 하는 단백질-리간드 복합체 구조에서 리간드 구조만을 떼어내어 사용하였다. 단백질-리간드 복합체 구조에서 리간드를 제외한 다른 단백질 및 물 분자를 제거한 뒤, UCSF Chimera 프로그램을 이용하여 수소원자를 붙이고 Gasteiger charge 를 assign 하였다.

Getting Ligand 3D structure from PubChem database

리간드의 초기 구조를 얻는 또 다른 방법으로 PubChem database 를 이용하였다. PubChem 은 National Center for Biotechnology Information(NCBI)에 의해 운영되고 있는 여러 화학 분자들의 데이터베이스로, 수백만 개의 화합물 구조에 대한 정보를 가지고 있고, 이는 FTP 서버를 통해서 자유롭게 다운받을 수 있다. 리간드 이름을 기반으로 PubChem database 에서 타겟 리간드를 찾아 sdf format 으로 제공되는 리간드의 3 차원 구조를 얻은 뒤 이에 UCSF Chimera 프로그램을 이용하여 Gasteiger charge 를 assign 하고 mol2 format 으로 변환하였다.

Constructing Ligand 3D structure from SMILES formula

SMILES formula 는 분자의 구조를 간단한 line notation 으로 나타낼 수 있는 방법으로, 분자가 가지고 있는 원자들과 그들의 결합, 그리고 stereochemistry 에 대한 정보를 모두 가지고 있다. 한 분자의 3 차원 구조는 그 분자가 가지고 있는 원자들의 종류, 각각의 원자들 사이의 결합 길이와 결합각, 그리고 네 원자가 이루고 있는 dihedral angle 을 가지고 나타낼 수 있는데, 여러 화학적인

이론 및 계산방법

Overall Procedure

본 연구의 전체적인 방법은 Figure 1 에 나타낸 바와 같다.

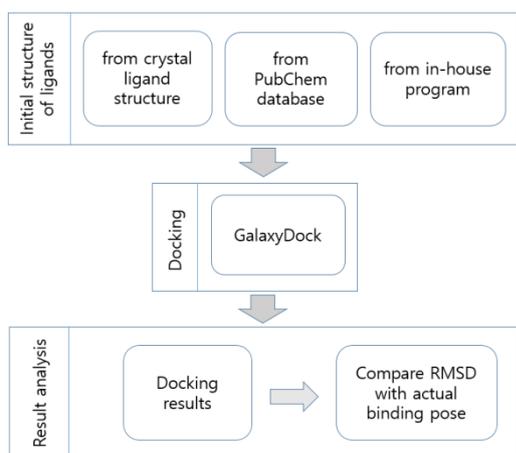


Figure 1. Overall procedure

Jinsol Yang et al.

구조들에 대해 이러한 값들은 실험을 통해 측정된 바 있다. 기존에 알려진 측정값들을 토대로 connectivity, hybridization state, 그리고 ring 에 대한 정보가 있으면 이에 대응하는 이상적인 구조를 만들어낼 수 있다.

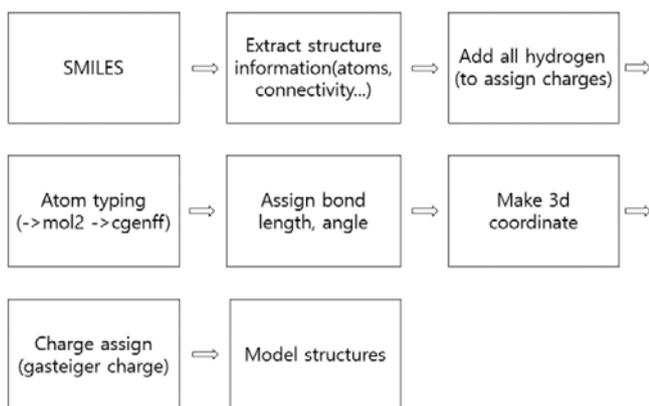


Figure 2. Flow chart of constructing ligand 3D structures from SMILES

이러한 일을 수행하기 위해서 만들어진 in-house 프로그램에서는 먼저 주어진 SMILES formula 에서 각 원자들 간의 connectivity, hybridization state, 그리고 ring 을 구성하는 원자들에 대한 정보를 얻었다, 이를 바탕으로 먼저 생략되어있는 수소 원자를 붙이고, 모든 원자들의 MOL2 type 을 지정하였다.

리간드에 ring 구조가 있는 경우에는 미리 만들어놓은 ring structure library 를 이용하여 ring 구조를 생성하였다. Ring structure library 를 구축한 방법은 다음 섹션에서 다루도록 하겠다. Ring structure library 에서 분자 안에 ring 들과 같은 구조식을 가지는 가지는 고리 구조들을 뽑았고, 이 구조를 그대로 가져와서 이용하는 식으로 분자 내에서 고리 구조를 다양하게 만들었다 Ring 이 아닌 나머지 구조에 대해서는 원자들의 MOL2 type 을 CGENFF 에서 가장 빈번하게 등장하는 atom type 으로 mapping 한 결과를 토대로, CGENFF 에서 주어지는 parameter 을 이용하였다. CGENFF 에 있는 결합길이, 결합각 정보와 주어진 connectivity 와 conformation 정보에서 가장 선호되는 몇 가지 구조에 대응하는 dihedral angle 을 설정하여 구조를 만들어냈다. (예를 들면 SN=4 인 구조의 경우 dihedral angle 을 gauche 또는 anti 에 대응하는 60°, -60°, 180° 도 중 하나로, SN=3 인 구조인 경우 0° 또는 180° 가 되도록 하였다.) 이렇게 임의로 dihedral angle 을 지정할 때 원자들 사이에 clash 가 일어나는 것을 방지하기 위하여 연결되지 않은 두 원자간의 거리가 1.3Å 이하가

되면 clash 로 간주 새로운 dihedral angle 을 지정하였다.

최종 리간드의 구조는 MOL2 file 형태로 변환하고 이후 UCSF Chimera 프로그램을 이용하여 각 원자가 가지고 있는 charge 를 Gasteiger charge 로 assign 하였다.

Ring structure library

Ring structure library 구축을 위해 Yang Zhang group 에서 만든 ligand-protein binding database 인 BioLiP 에서 해당하는 18 만개의 regular ligand 의 pdb 파일을 다운받았고, 다운받은 구조에서 ligand 가 가지고 있는 ring 구조만 추출하여 mol2 file 로 만들었다. 추출한 ring 구조는 flexible, rigid 한 구조로 분류하였고, 같은 구조식을 가지고 있는 ring 들은 상호간의 RMSD 가 $0.2*(n*6)^{1/2} \text{Å}$ 이상이 되도록 clustering 하여 non-redundant 한 ring library 를 구축하였다.

Protein-Ligand docking using GalaxyDock

본 연구에서는 단백질-리간드 복합체 구조를 예측하기 위하여 EDISON 에 탑재된 GalaxyDock 프로그램^{4,5} 을 이용하였다. GalaxyDock 프로그램은 conformational space annealing(CSA)를 사용하여 큰 자유도를 가지고 있는 리간드 구조의 conformational space 를 효과적으로 탐색하고, autodock3 에서 사용되던 에너지 함수에서 부족한 부분인 ligand 의 내부 에너지에 PLP(piecewise linear potential)의 ligand torsion energy 를 개선하여 단백질-리간드 복합체 구조를 예측하는 프로그램이다. 최근에는 단백질의 표면을 나타낼 수 있는 beta-complex 를 pre-docking 에 이용하여 CSA 를 위한 initial bank 를 얻음으로써 그 성능이 향상되었다. 모든 도킹은 10 회 반복하여 가장 좋은 결과를 가지는 것을 분석에 사용하였다.

결과 및 토의

일반적으로 docking 한 결과 구조와 실제 crystal 구조와의 RMSD 가 2Å 이하이면 docking 이 성공한 것으로 간주한다. 이러한 기준을 적용했을 때, crystal 구조에 있는 리간드 구조를 그대로 이용한 경우 30 개의 target 모두 docking 에 성공하여 100%의 성공률을 보였다. PubChem database 에서 가져온 리간드 구조를 사용했을 때는 30 개의 target 중 21 개가 성공하여 70%의 성공률을 보였다.

Impact of input conformation on docking performance

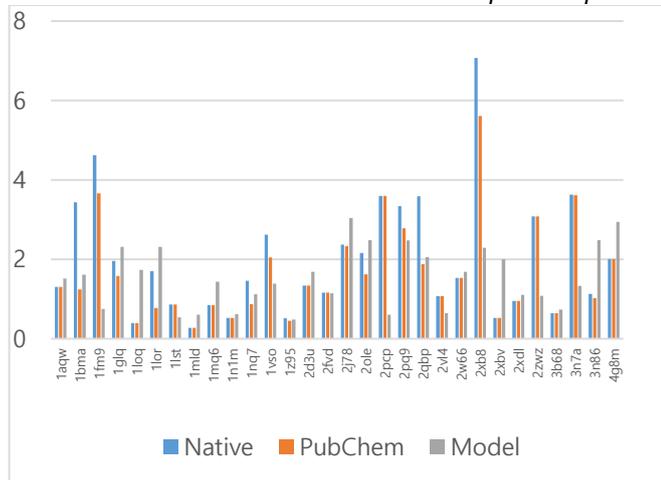


Figure 3. 각 구조를 이용해 도킹한 결과 구조의 결정 구조와의 RMSD

직접 만든 in-house 프로그램을 이용하여 만든 구조를 이용했을 때는 3 개의 model 구조 중 하나만 성공해도 docking 이 성공한 것으로 간주하였다. 30 개의 target 중 24 개가 도킹에 성공하여, 80%의 성공률을 보였다. 이는 PubChem database 에서 리간드 구조를 가지고 왔을 때에 비해 향상된 것이다. 이는 리간드에 ring 구조가 있는 경우, in-house 프로그램을 이용하는 경우 3 개의 다양한 ring 구조를 만들어 사용하기 때문으로 보인다. GalaxyDock 은 ring 의 구조 유연성을 고려하지 않으므로 초기 ring 구조가 결합구조와는 다른 경우 제대로 된 도킹 결과를 얻기 어렵다. 따라서 ring 구조를 보다 다양하게 만들어주는 in-house 프로그램을 이용하는 것이 PubChem database 를 이용하는 것보다 더 향상된 도킹 성능을 보여준 것으로 판단된다. 본 연구에서는 리간드의 구조를 3 개까지만 생성하였지만, in-house 프로그램을 이용하여 3 개보다 더 다양한 구조를 만들어 낼 수 있음을 고려해 보면 성공률은 더욱 향상 될 수 있을 것으로 보인다.

여기서는 실패한 target 들의 경우, in-house 프로그램으로 만든 리간드 구조가 crystal 구조와 어떻게 다른지, 또한 PubChem 을 이용했을 때는 실패했지만 in-house 프로그램을 이용하여 만든 구조로는 성공한 경우 어떤 차이가 있었기 때문인지에 대하여 몇몇 사례를 바탕으로 분석해보고자 한다.

먼저 crystal 리간드 구조를 제외한 다른 모든 구조에서 도킹이 실패한 2J78 을 살펴보면, 초기 리간드의 구조가 모든 모델에서 실제 crystal 구조와 확연히 차이가 있음을 볼 수 있다 (Figure 4). 자세히

보면 가운데 ring 의 구조부터 큰 차이가 있는데, ring 구조가 조금만 달라도 결과가 뻗어 나오는 방향은 크게 변할 수 있고, 이렇게 되면 구조가 크게 달라져 전체 도킹 성능에 큰 영향을 미칠 수 있다. In-house 프로그램으로 만든 모델 1, 모델 2, 모델 3 에서 가운데 ring 의 구조와 crystal ligand 에서의 ring 의 RMSD 는 각각 0.837, 0.861, 0.871Å 로 꽤 큰 차이가 있는 것을 알 수 있다.

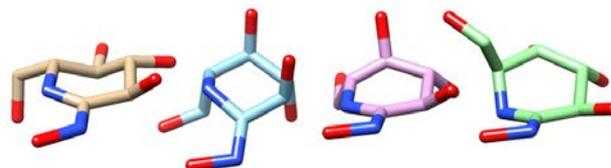


Figure 4. 2J78의 리간드 구조들. 왼쪽부터 실제 단백질-리간드 복합체에 있는 리간드 구조, in-house 프로그램으로 만든 모델 1, 모델 2, 모델 3이다.

3N7A 의 경우 여러 모델을 만들어 성공적으로 docking 이 이루어진 경우였는데, in-house 프로그램으로 만든 모델 2 를 이용했을 때 실패했던 도킹이 마찬가지로 in-house 프로그램으로 만든 모델 1 을 이용했을 때는 성공하였다. 모델 1 과 모델 2 의 구조를 살펴보면 ring 구조에서 크게 차이가 있음을 확인 할 수 있었다 (Figure 5).

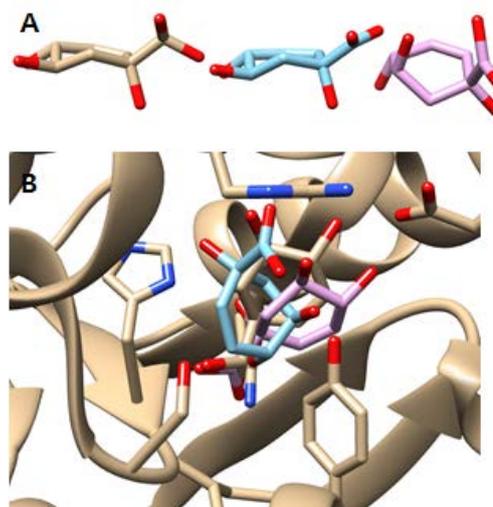


Figure 5. (A) 3N7A의 리간드 구조들. 왼쪽부터 실제 단백질-리간드 복합체에 있는 리간드 구조, in-house 프로그램으로 만든 모델 1, 모델 2이다. (B) 모델1과 모델2를 이용한 도킹의 결과 구조

Jinsol Yang et al.

2PQ9 의 경우도 이와 비슷하게 in-house 프로그램으로 만든 여러 모델들 중에 crystal structure 과 비슷한 구조를 가지는 모델 1 만이 docking 에 성공하였다.

2PCP 의 구조를 보면 한 가지 주의하여야 할 점을 알 수 있는데, 2PCP 의 리간드는 3 개의 고리로 이루어져 있다. 그런데 PubChem 을 통해 얻어진 구조는 각각의 ring 의 구조는 비슷하였지만, 그들이 결합하고 있는 stereochemistry 가 실제 리간드 구조와 일치하지 않았고 이에 따라 도킹에 실패하게 되었다. 반대로 ring 의 구조도 비슷하고 stereochemistry 도 일치하는 모델 1 의 경우 성공적인 도킹 결과를 얻을 수 있었다 (Figure 6). 이로부터 stereochemistry 가 정확히 정의되지 않은 구조를 도킹할 때는 이를 고려하여 샘플링을 할 수 있도록 구조를 생성하여야 한다는 점을 확인 할 수 있었다.

났다. 모델 1 의 경우 crystal 구조와의 RMSD 가 0.745Å 로 성공한 반면, PubChem 구조의 경우 3.663Å 로 실패하였다. 그 원인으로 생각할 수 있는 것은 conformation space 가 매우 크기 때문에 충분한 샘플링이 일어나지 못했거나, 아니면 두 구조가 결합각과 결합길이 측면에서 조금씩 차이가 있고, 분자의 길이가 긴 편이기 때문에 이러한 차이들이 누적되어 큰 차이를 가져오게 되었을 수 있다고 생각할 수 있다.

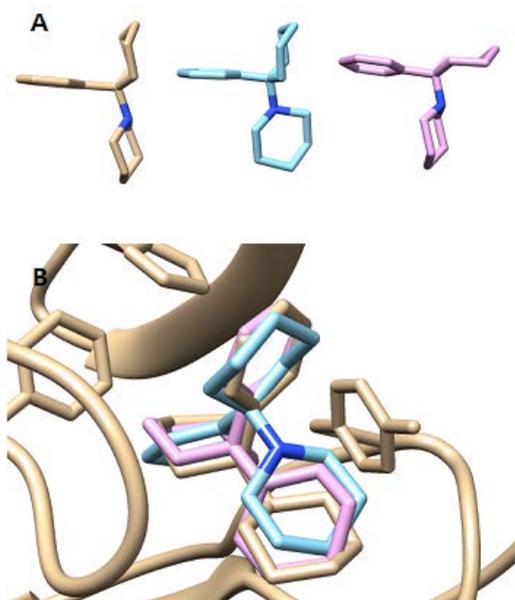


Figure 6. (A) 2PCP의 리간드 구조들. 왼쪽부터 실제 단백질-리간드 복합체에 있는 리간드 구조, in-house 프로그램으로 만든 모델 1, PubChem으로부터 얻은 구조이다. (B) 모델 1, PubChem 구조를 이용한 도킹 결과

한 가지 흥미로웠던 점은 상대적으로 rigid 하여 초기 리간드 구조에 따른 도킹 성능 차이가 별로 없을 것으로 예상되었던 target 도 큰 차이를 보이고는 한다는 것인데, 한 가지 예가 target 1FM9 인다. In-house 프로그램으로 만든 모델 1 과 PubChem 에서 가져온 리간드 구조를 살펴보면 거의 비슷해 보이고, aromatic ring 만 존재하기 때문에 리간드가 rigid 해 보이지만, 도킹 결과는 크게 차이가

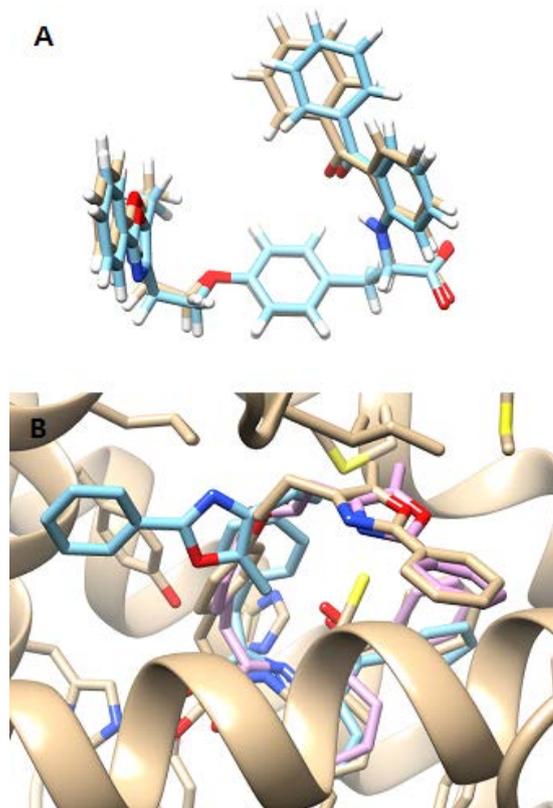


Figure 7. (A) 1FM9의 실제 단백질-리간드 복합체에 있는 구조와 PubChem 구조를 align 한 것. (B) PubChem 구조(하늘색), 모델구조(보라색)를 이용하여 도킹한 결과

1FM9 타겟에 대한 도킹 결과가 샘플링에 의한 문제인지, 결합길이 및 결합각의 차이에 의한 문제인지를 확인해보기 위하여, PubChem 으로부터 얻은 리간드 구조와 실제 결정구조에서 얻은 리간드 구조를 align 해 보았다. 이를 위해 PubChem 에서 얻은 리간드 구조에서 결합 길이, 결합각을 제외한 나머지 torsion angle 들을 실제 결정구조와 같도록 조정해 주었다. 그 결과, 크게 두 부분의 결합각 차이가 도킹 성능에 영향을 준 것으로 판단된다. 두 구조를 align 한 결과를 보면, Figure 7A 에서 오른쪽 두 벤젠 고리와 그 가운데 탄소가 이루고 있는 결합각이 실제 리간드에서는 130.4° 이지만,

Impact of input conformation on docking performance

PubChem 에서 얻은 리간드에서는, 그 각도가 122.3° 이다. 또한 오른쪽 아래 질소 원자와 그 인접하는 두 탄소 사이의 결합각이 실제 구조에서는 133.2° 이지만, PubChem 구조에서는 122.0° 이고, 그 결과 align 된 구조를 보면 끝의 벤젠 고리 부분이 서로 어긋나 있는 것을 볼 수 있다. 이는 작은 차이처럼 보이지만 도킹에서 큰 차이를 불러오게 되는데, PubChem 으로부터 얻은 구조의 도킹 결과에서 보면 한 쪽 부분은 알맞게 위치하였지만, 이 결합각의 차이로 인해 나머지 부분이 실제 결합 구조와 가까워지기 위해서는 단백질의 chain 과 clash 가 일어나야 하고, 그러므로 이는 에너지적으로 불리해진다(Figure 7B). 이러한 차이 때문에 PubChem 으로부터 얻은 구조에서 도킹이 실패한 것으로 보인다. 이러한 차이가 적은 모델 구조의 경우 도킹에 성공하였는데, 이는 리간드 결합각의 차이가 도킹에 큰 영향을 줄 수 있는 예시로 리간드의 구조 유연성을 고려해 주어야 하는 근거가 된다.

이러한 사례분석을 통해 대부분의 경우 리간드의 ring 구조의 차이에 따라 도킹의 성공여부가 달라진 것을 알 수 있었다. 이에 더하여 aromatic ring 과 같이 상대적으로 크기가 큰 요소들이 리간드에 존재하는 경우에 이러한 작용기들을 연결하는 결합길이나 결합각에 따라서 도킹 성능의 차이가 발생할 수 있음을 확인하였다. 단백질-리간드 도킹 프로그램이 실제 응용되는 분야에서 더욱 보탬이 되기 위해서는 ring 의 구조 유연성이나 결합길이, 결합각의 변화 등을 다양한 리간드 conformation 을 이용하여 도킹을 함으로써 고려하거나, ring 의 구조 유연성 및 결합길이, 결합각의 변화까지 고려할 수 있는 단백질-리간드 도킹 방법 개발이 필요할 것으로 보인다.

결론

본 연구에서는 리간드의 초기 구조가 단백질-리간드 도킹의 결과에 미칠 수 있는 영향을 알아보기 위해 30 개의 target 에 대해 다양한 리간드 구조를 이용하여 도킹을 시도하였다. PubChem 을 통해 얻은 각각 하나씩의 구조로 도킹했을 때 보다 여러 모델을 생성해서 도킹에 사용했을 때 성공률이 올라가는 것을 확인 할 수 있었다. 다양한 사례분석을 통하여 구조적으로 유연한 ring 이 리간드에 존재하는 경우 가능한 여러 ring 구조를 고려해 주는 것이 도킹의 성공률에 큰 영향을 미칠 수 있다는 사실을 확인할 수 있었다. 또한 aromatic ring 과 같이 크기가 큰 rigid 한 group 이 리간드에

존재하는 경우 결합각의 작은 차이가 도킹의 성능에 큰 차이를 줄 수 있음 역시 확인되었다. 이에 미루어 볼 때 앞으로 docking simulation 에서 이러한 구조를 더 다양하게 고려하는 것이 그 정확도를 향상시키는데 도움이 될 수 있을 것이다.

Ring 구조의 유연성이나 결합각의 차이를 다양한 초기구조를 만듦으로써 고려할 수도 있겠지만, 이 방법은 초기 리간드 구조를 어떤 식으로 만드는지, 얼마나 만드는지에 따라 성능이 좌우된다는 점에서 한계가 존재한다. 단백질-리간드 도킹 프로그램 자체에서 리간드의 구조 유연성을 torsion angle 들을 샘플링하는 것뿐만 아니라 추가적인 ring 의 구조 유연성, 결합각의 변화 등을 고려할 수 있도록 추가해 준다면 리간드의 초기구조에 크게 영향을 받지 않는 단백질-리간드 도킹을 할 수 있을 것으로 기대된다.

감사의 말. 이 연구는 EDISION 과제를 통해 지원받아 수행되었습니다.

참고문헌

1. KUNTZ, Irwin D., et al. A geometric approach to macromolecule-ligand interactions. *Journal of molecular biology*, **1982**, 161.2: 269-288.
2. DHANIK, Ankur; KAVRAKI, Lydia E. Protein-Ligand Interactions: Computational Docking. *eLS*, **2001**.
3. SOUSA, Sérgio Filipe; FERNANDES, Pedro Alexandrino; RAMOS, Maria João. Protein-ligand docking: current status and future challenges. *Proteins: Structure, Function, and Bioinformatics*, **2006**, 65.1: 15-26.
4. SHIN, Woong-Hee; SEOK, Chaok. GalaxyDock: protein-ligand docking with flexible protein side-chains. *Journal of chemical information and modeling*, **2012**, 52.12: 3225-3232.
5. SHIN, Woong-Hee, et al. GalaxyDock2: Protein-ligand docking using beta-complex and global optimization. *Journal of computational chemistry*, **2013**, 34.30: 2647-2