

화학 정보학 및 분자모델링

HCV와 antiviral drugs against HCV의 docking 계산

화학과 201313766 오예슬

1. 서론

[1] C형 간염

<1> C형 간염이란?

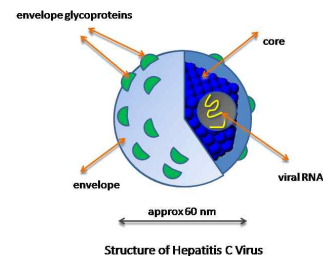
C형 간염(hepatitis C, HC)은 C형 간염 바이러스(hepatitis C virus, HCV)에 감염되어 간에 염증이 발생하고 간세포가 손상되는 질환이다. 간세포가 파괴당할 때 이에 대응하기 위한 신체의 면역반응으로 인해 간에 염증이 생긴다. C형 간염은 HCV에 감염된 사람의 체액이나 혈액이 정상인의 상처 난 피부나 점막을 통하여 전염되는 감염성 질병이다. 소독하지 않은 바늘이나 침으로 시술을 받거나 문신, 피어싱, 침술 또는 면도기, 칫솔, 손톱 깎기 등을 통해 감염이 이뤄진다. 의료기관에서도 출혈이 동반될 수 있는 시술 등을 통해 바이러스 전파가 이뤄질 수 있다.

C형 간염 바이러스(HCV)에 감염되면 70~80%가 만성 간염으로 진행하고 이 중에서 30~40% 정도가 간경변증, 간암으로 진행한다.

우리나라에서는 전 국민의 약 1%가 C형 간염 바이러스 보유자로 추정되며, 전체 만성 간 질환(간염, 간경변증, 간암) 환자의 약 10~15%가 C형 간염 바이러스에 의해 발생한다.

<2> C형 간염 바이러스란?

C형 간염 바이러스(HCV)의 크기는 55~65 nm 정도로 작고, 막이 있는 양성 단일 가닥 RNA 바이러스의 하나로, 플라비 바이러스 (Flaviviridae) 과에 속한다. C형 간염 바이러스는 C형 간염, 그리고 간암(간세포암, HCC)과 림프종의 병인이 된다.



그림(1)

<3> C형 간염의 분류와 치료

A형과 B형 간염과 달리 현재 C형 간염을 예방할 백신은 존재하지 않는다. 다만 최근 C형 간염 치료제의 눈부신 발전으로 탁월한 효과와 가벼운 부작용을 가진 경구용 항바이러스제를 사용할 수 있고 C형 간염의 완치를 언급할 수 있을 만큼 치료제가 발전하였다.

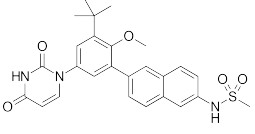
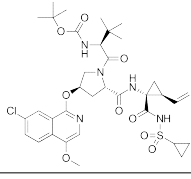
직접 작용 항바이러스 (direct acting antivirals, DAA) 치료는 경구용 항바이러스제가 C형 간염 바이러스의 생활사에 직접 작용하여 항바이러스 효과를 나타내는 것이다. 현재 사용되는 경구용 항바이러스제는 바이러스 유전자형에 따라 치료약제가 다르다. C형간염 바이러스는 유전자형이 1형에서부터 6형까지 존재하지만, 국내에는 주로 1형과 2형이 대부분이다. 유전자형 1형은 전 세계적으로 가장 널리 퍼져 있으며 아형(subtype)에 따라 1a형과 1b형으로 나뉜다.

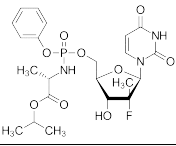
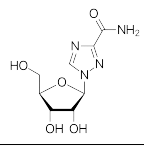
유전자형별로 구체적인 치료 약제를 살펴보면 아래 표(1)과 같다.

1형		2형	모든 유전자형
1a형	1b형	소발디 (sofosbuvir) 리바비린 (ribavirin)	마비렛 (glecavirprevir /pibrentasvir)
하보니 (ledipasvir /sofosbuvir) 제파티어 (elbasvir /grazoprevir)	하보니 (ledipasvir/sofosbuvir) 제파티어 (elbasvir/grazoprevir) 비키라 (ombitasvir/paritaprevir/ritonavir) 엑스비라 (dasabuvir) 다클린자 (daclatasvir) 순베프라 (asunaprevir)		

표(1)

위의 표(1)에서, 앞으로 다룰 1b형 치료제 중 ‘dasabuvir, asunaprevir’ 성분과 2형 치료제 중 ‘sofosbuvir, ribavirin’의 정보를 아래 표(2), (3)에 정리하였다.

1b	dasabuvir	asunaprevir
구조		
PubChem	56640146	16076883
SMILES	<chem>CC(C)(C)c1cc(cc(c1OC)c2ccc3cc(ccc3c2)NS(=O)(=O)C)n4ccc(=O)[nH]c4=O</chem>	<chem>O=C(N5[C@H](C(=O)N[C@@]2(C(=O)NS(=O)(=O)C1CC1)[C@H](\C=C)C2)C[C@@H](Oc3ncc(OC)c4c3cc(Cl)cc4C5)[C@@H](NC(=O)OC(C)(C)C)C(C)C</chem>
화학식	C ₂₆ H ₂₇ N ₃ O ₅ S	C ₃₅ H ₄₆ Cl N ₅ O ₉ S
몰 질량	493.58 g·mol ⁻¹	748.29 g·mol ⁻¹

2	sofosbuvir	ribavirin
구조		
PubChem	45375808	37542
SMILES	<chem>C[C@@H](C(=O)OC(C)C)N[P@](=O)(OC[C@H]1[C@H]([C@@]([C@H](O1)N2C=CC(=O)NC2=O)(C)F)OC3=CC=CC=C3</chem>	<chem>C1=NC(=NN1C2C(C(C(O2)CO)O)O)C(=O)N</chem>
화학식	C ₂₂ H ₂₉ F N ₃ O ₉ P	C ₈ H ₁₂ N ₄ O ₅
몰 질량	529.458 g·mol ⁻¹	244.206 g·mol ⁻¹

위-표(2), 아래-표(3)

[2] 단백질-리간드 도킹(Protein-Ligand Docking)

Protein-Ligand Docking은 단백질과 리간드의 상호작용을 모델링하여 target protein에 결합하는 ligand의 결합구조, 즉 단백질-리간드 결합구조(protein-ligand complex)를 알아내고 어떤 ligand가 가장 세게 결합할 것인지 예측하는 것을 목표로 한다. 이를 통해 target protein의 binding site 내에서 ligand가 어떻게 행동하고 있는지 규명해볼 수 있다.

Docking 계산은 ligand의 conformation을 예측하고 binding affinity를 평가한다. 일반적으로 docking을 사용하는 경우 binding site를 지정해주고 지정된 범위 안에서 ligand의 conformation을 찾는다. 이러한 계산으로 가장 많이 사용되는 응용 분야는 virtual screening이다. Virtual screening 분야는 데이터베이스 안의 다양한 분자들이 target protein과 결합하는 경향을 실험 없이 컴퓨터로 계산하여 예측하는 것이다.

특정 단백질에 붙어서 그 단백질의 역할을 선별적으로 제어하는 것이 바로 약이다. 수천수만 개의 화합종으로부터 실제 target protein에 binding 하는 ligand를 찾아내어 신약개발의 속도를 높일 수 있다. 따라서 단백질과 리간드(small molecule)가 어떤 상호작용으로 서로 붙게 되는지 연구하는 것, 즉 Protein-Ligand Docking은 신약개발에 매우 유용하다.

앞에서 정리한 C형 간염 1b형과 그 치료제 중 두 가지 성분 '다사부비르(dasabuvir), 아스나프레비르(asunaprevir)', 그리고 C형 간염 2형과 그 치료제 중 두 가지 성분 '소포스부비르(sofosbuvir), 리바비린(ribavirin)'을 각각 docking 계산시켜 총 4가지의 계산 결과를 얻어보도록 하겠다.

2. 계산 방법

4가지 계산 중 단백질: HCV 1b, 리간드: dasabuvir를 예로 들어서 설명하겠다.

[1] Step 1: 단백질 구조 열기

-RCSB PDB 사이트에서 바이러스 단백질을 찾는다.

[3HKY] HCV NS5B polymerase genotype 1b in complex with 1,5 benzodiazepine 6

-Chimera에서 다른 리간드와 연결된 단백질 복합체를 불러온다.

File→Fetch by ID→PDB: 3HKY

-같은 단백질 두 개가 대칭으로 불러와 졌으므로 한 개를 삭제한다.

Select→Chain→A

Action→Atoms/Bonds→Delete

-리간드를 포함하여 단백질을 제외한 모든 것을 제거한다.

Select→Residue→All nonstandard

Action→Atoms/Bonds→Delete

[2] Step 2: Use DockPrep tool-docking 계산을 하기 위한 준비과정

-수소를 붙여주고, 부분 전하를 계산한다.

Tools→Structure editing→Dock Prep

Add Hydrogen: Unspecified (determined by method) = '알아서 붙이기' 체크

Add charge: Gasteiger = '경험적(실험적)으로 추측하여 부분 전하 계산' 체크

단백질 구조를 'rec_charged.mol2' 로 저장한다. (rec: receptor 수용체)

[3] Step 3: 수소 없는 단백질 구조를 PDB 포맷으로 저장

- 나중에 docking 계산할 때 단백질 표면 계산을 위해 수소가 없는 단백질 구조도 저장한다.
Select→Chemistry→Element→H
Action→Atoms/Bonds→Delete
File→Save PDB→‘rec_noH.pdb’ 로 저장 (use untransformed coordinates 체크)

[4] Step 4: 리간드 파일 준비

- 찾아냈던 정보로 Chimera에서 리간드를 연다. (Pubchem, Marvin sketch, PDB 이용)
File→Fetch by ID→PubChem:56640146
- 수소를 붙여주고, 부분 전하를 계산한다.
Tools→Structure editing→AddH
Tools→Structure editing→Add Charge
(단백질 부분 전하 계산할 때와 같은 방식으로 설정)
리간드 구조를 lig_charged.mol2 로 저장한다. (lig: ligand)

[5] Step 5: 단백질 표면 정보 저장

- Chimera의 Write DMS 도구를 사용해서 단백질의 표면 정보를 저장한다.
File→open→‘rec_noH.pdb’
Action→Surface→Show
Tools→Structure Editing→Write DMS
‘rec.dms’ 파일로 저장한다.

[6] Step 6: Generating spheres around the receptor

- winSCP로 지금까지 생성한 파일을 PuTTY에 복사해 준다.
- nano로 ‘INSPH’라는 이름의 input file 생성한다.
nano INSPH→해당 내용 입력→^o(저장)→^x(나감)
- sphgen 명령을 사용해서 sphere를 생성한다.
제대로 실행되었다면 ‘OUTSPH’와 ‘rec.sph’ 파일이 생성된다.
- nano로 ‘sphgen_cluster.in’ 이라는 파일을 생성한다.
nano sphgen_cluster.in→해당 내용 입력→^o(저장)→^x(나감)
- showsphere < sphgen_cluster.in 명령을 실행한다. (꺾쇠의 오른쪽 내용을 왼쪽으로)
- PuTTY에서 윈도우로 ‘sphgen_cluster.pdb’파일을 복사 후 Chimera로 연다.
- 점들을 구로 표시한다. (리간드가 결합할 수 있는 자리들이 노란 구로 표시된다.)
Action→Atoms/Bonds→sphere

[7] Step 7: Grid generation

- nano로 ‘box.in’ 파일을 생성한다.
nano box.in→해당 내용 입력→^o(저장)→^x(나감)
- showbox 명령 후 ‘box.in’에 입력했던 내용을 질문이 나올 때마다 순서대로 입력한다.
- ‘rec_box.pdb’ 파일이 생성되었는지 확인한다.

nano grid.in→receptor_file: 'rec_charged.mol2', box_file: 'rec_box.pdb' 입력→^o(저장)→^x(나감)

```
grid -i grid.in
```

-dock.in 파일을 열어 ligand_atom_file과 receptor_site_file의 파일명을 수정한다.

-dock6 명령어를 입력하여 docking 계산을 실행한다.

제대로 실행되었다면 'anchor_and_grow_ranked.mol2' 파일과 'anchor_and_grow_conformers' 파일이 생성된다.

-PuTTY에서 윈도우로 'anchor_and_grow_conformers.mol2' 파일을 복사한다.

Tools→Surface/Binding Analysis→ViewDock→Dock4/5/6, File name: 'anchor_and_grow_conformers.mol2'

File→open→‘rec_noH’

[1] 단백질: HCV 1b - 리간드: dasabuvir

*맨 위: 예측된 리간드 중 가장 안정한 구조
(가장 위부터 순서대로 안정한 구조)

*리간드가 조금씩 다른 구조를 가짐
(\therefore 리간드가 flexible 하므로 다양한 구조를
가질 수 있다.)
(\therefore single bond rotation)

*Grid Score ≡ 에너지의 총합 : -52
(값이 낮을수록 단백질과 강한 binding)

*van der Waals E: -43

```
*electrostatic E: -9
```

*Internal energy repulsive: 39
($\hat{=}$ steric hindrance)

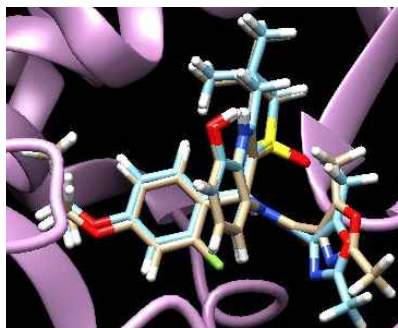
[2] 단백질: HCV 1b - 리간드: asunaprevir	[3] 단백질: HCV 2 - 리간드: sofosbuvir	[4] 단백질: HCV 2 - 리간드: asunaprevir
*Grid Score: -70	*Grid Score : -57	*Grid Score: -33
*van der Waals E: -57	*van der Waals E: -55	*van der Waals E: -32
*electrostatic E: -13	*electrostatic E: -1	*electrostatic E: -0.8
*Internal energy repulsive : 53	*Internal energy repulsive : 20	*Internal energy repulsive : 3

4. 결론 및 고찰

계산 결과, 가장 안정한 리간드 결합구조의 Grid Score는 case① [HCV 1b - dasabuvir]=-52, case② [HCV 1b - asunaprevir]=-70, case③ [HCV 2 - sofosbuvir]=-57, case④ [HCV 2 - asunaprevir]=-33으로 case②의 Grid Score가 가장 낮았다.

Docking 계산 결과의 값이 좋을수록 약효가 좋은지 비교해 보고 싶었지만, 조사 결과 C형 간염 치료를 위해 환자의 상태, 이전 치료 실패 여부, 간경변증 유무와 중증도 등 다양한 변수에 따라 약의 종류를 달리하고 일반적으로 두 개의 약을 병합 투여하는 경우가 많아서 생각했던 것과 같은 단순한 비교는 불가능했다.

계산의 정확도를 알아보기 위해, 추가로 RCSB PDB에서 찾았던 단백질-리간드 복합체 '[3H KY] HCV NS5B polymerase genotype 1b in complex with 1,5 benzodiazepine 6'를 이용하여 기존에 결합 되어 있던 리간드를 docking 계산시켜 보았다.



그림(2)

그 결과 가장 안정한 리간드 결합구조의 Grid Score는 -68로 나왔다.

그림(2)는 기존의 단백질-리간드 복합체에 결합 되어 있던 리간드(상아색)와 docking 계산을 실행하여 가장 낮은 Grid Score 값이 나온 리간드(하늘색)를 동시에 불러온 것이다. 그 결과 두 리간드가 거의 포개어져 두 리간드의 conformation, 결합구조 등이 거의 같은 것을 확인할 수 있었다.

프로그램으로 결합자리를 지정한 타겟 단백질 파일 하나를 생성하면 docking 시켜보고 싶은 리간드들을 쉽게 docking 시켜볼 수 있고, docking 시킨 각 리간드의 구조를 가장 안정한 것부터 차례로 도출해 주며, 그 정확도 또한 높아서 protein-ligand docking 계산이 상당히 유의미한 것이고 실제 약물 개발에 유용한 프로그램임을 간접적으로 확인할 수 있었다.

5. 참고자료

-대한간학회

-WIKIPEDIA

-[BRIC Bio 통신원] [컴퓨터를 이용한 신약개발 (CADD)] Molecular Docking (Sampling 알고리즘)

-The Author 2015. Published by Oxford University Press on behalf of the Infectious Diseases Society of America.