

G 단백질 연결 수용체계(GPCR system)에서의
정전기적 포텐셜(Electrostatic Potential)에 따른
효과를 고려한 단백질과 리간드의 상호작용 예측
(protein-ligand interaction prediction)

최규홍,¹ 신윙희,² 이동선²

¹ 서울특별시 관악구 관악로 1 서울대학교 농업생명과학대학 응용생물학전공, 151-742.

² 서울특별시 관악구 관악로 1 서울대학교 자연과학대학 화학부, 151-742.

전화: (010) 2271-1303, 이메일: rb8989@snu.ac.kr

Chemworks (or Chemworks2) ID: ecc13_SNU_seventm

초록 (ABSTRACT)

2012 년 G 단백질 연결 수용체(G-Protein Coupled Receptors ; GPCR) 연구가 노벨 화학상을 받았다.¹ 상당히 많은 병과 관련되어 있어 잠재력이 크고, 많은 연구가 진행 중이다.² 현재 리간드와 단백질간의 정전기적 포텐셜 연구를 통한 예측 연구가 진행되고 있지만,³ GPCR 과 리간드 간의 연구에서 아직 리간드의 전하를 통한 단백질과 리간드간의 상호작용 예측 연구가 되어 있지 않다. 그렇기 때문에 이번 연구에서는 8 가지 방법으로 전하(charge)를 띠게 하여서 단백질과 리간드의 상호작용을 계산을 통하여 예측하여 보았다.

핵심어 : *GPCR, Protein Ligand interaction, Restrained electrostatic potential, GalaxyDock,*

Quantum mechanical calculation

서론

GPCR system 은 많은 척추동물의 생리에서 신호 전달과정의 기작에서 중요한 역할을 한다. 막 단백질(membrane protein) 중에서 가장 큰 family 를 가지고 있으며 호르몬과 신경전달물질(neurotransmitters)를 세포에 전달하여 시각, 후각, 미각 등의 감각의 반응을 인식하는데 사용된다.² 그렇기 때문에 다양한 병을 치료하는 타겟으로 큰 잠재력을 가지고 있다.

근래에 양적 구조-반응 관계(Quantitative structure-activity relationship ; QSAR)연구를 통하여 리간드를 기반으로 한 제약업체의 약 발견과 설계가 중요해지고 있다.⁴ 이러한 연구에서 경험적인 전하(empirical charge)와 양자계산을 통해 얻어진 전하(charge)를 사용한 리간드 연구가 있다.³ 하지만 GPCR 에 이러한 전하 계산을 통한 단백질과 리간드의 상호작용 예측에 대한 연구가 없다. 그렇기 때문에 이번 연구에서 3 가지 GPCR 계(system)과 그것의 리간드의 상호작용을 전하(charge)의 종류를 비교해 보았다.

이론 및 계산방법

사용된 GPCR 계(system)에는 총 3 가지의 adrenergic receptor 를 이용한다. $\beta 1AR(2ycw)$ 과 CAU,⁶ $\beta 2AR(2rh1)$ 과 CAU,⁷ 그리고 $\beta 2AR(3sn6)$ 과 POG⁸를 연구하였다.

리간드에 전하(charge)를 계산하는 방법에는 총 8 가지 방법을 사용하였다. 그 중 6 가지는 경험적인 전하(empirical charge)를 띠게 하였고, 전하를 띠게 하기 위해서 SYBYL8.1⁹를 이용하였다. SYBYL8.1 을 이용하여 만든 리간드의 경험적인 전하(empirical charge)에는 Del-Re charges,¹⁰ formal charges, Gasteiger-Huckel charges, Gasteiger-Marsili

charges,¹¹ Huckel charges, Pullman charges¹² 의 방법을 사용하였다.(각각 이하 DR, FC, GH, GM, H, P) 그리고 양자역학적인 계산(Quantum mechanical calculation)을 통해 2 가지 방식의 전하(charge)를 계산하였다.

양자역학적인 계산(Quantum mechanical calculation)을 통한 RESP(restrained electrostatic potential) 전하(charge)를 띠게 하기 위해서 2 가지 방식의 계산을 하였다.

첫 번째 경우는 리간드의 구조를 최적화(optimization)를 하기 위해서 전기적 구조를 모델링(electronic structure modeling)을 하는 Gaussian09¹³ 를 사용하였다. 계산 방법으로는 양자역학적 모델링 방법(quantum mechanical modeling method)중에서 밀도 함수 이론(density functional theory)중에서 가장 많이 쓰이는 방식으로 에너지를 자동으로 설정하고 이온화 포텐셜(ionization potentials)과 양성자 친화성(proton affinities), 그리고 총 원자 에너지(total atomic energies)를 이용한¹⁴ B3LYP(Becke, three-parameter, Lee-Yang-Parr)¹⁵을 사용하였다. 이 계산을 하는데 사용된 basis set 은 극성을 띤(polarized)한 basis set 인 6-31G*를 사용하였다.¹⁶ 이렇게 구조가 최적화 된 리간드를 다 전자계의 파동함수의 전자배치를 구하는 방식 중 하나인 HF(Hartree-Fock theory)¹⁷ 를 사용하여 점전하(single point charge)를 계산하였다. 이 계산을 하는데 사용된 basis set 은 6-31G*를 사용하였다.¹⁴ 이렇게 구한 점전하(single point charge)는 RESP(restrained electrostatic potential) method 를 이용하여 보다 정확한 양자 역학적인 정전기적 전하(Quantum mechanical electrostatic charge)를 구하였다(이하 RESP_1).¹⁸

두 번째 양자역학적인 계산(Quantum mechanical calculation)을 통한 RESP(restrained electrostatic potential) 전하(charge) 계산 방법은 첫 번째 경우와 같은 Gaussian09¹³ 란

프로그램을 사용하였다.(이하 RESP_2) 분자 구조 최적화에는 HF(Hartree-Fork theory)¹⁷ 를 사용하고 basis set 으로는 수소와 무거운 원자(heavy atom)이 포함된 6-31G**^{16,19} 를 사용하였다. 이렇게 최적화된 구조를 갖는 리간드를 B3LYP¹⁵ 을 통하여 점전하(single point charge)를 계산하였다. 이때에 basis set 은 최외각 극성의 함수(valence polarization functions)를 정의에 따라 포함한 cc-pVTZ^{19,20} 를 이용하였다. 이것을 RESP(restrained electrostatic potential) method 를 이용하여 최종적인 양자 역학적인 정전기적 전하(Quantum mechanical electrostatic charge)를 구하였다.¹⁸

8 가지 방법으로 계산한 전하(charge)를 띠는 리간드를 총 4 번의 GalaxyDock²¹ 을 실행하여 단백질과 리간드의 상호작용에 대하여 알아보고, RMSD 와 energy 를 비교하여 보았다.

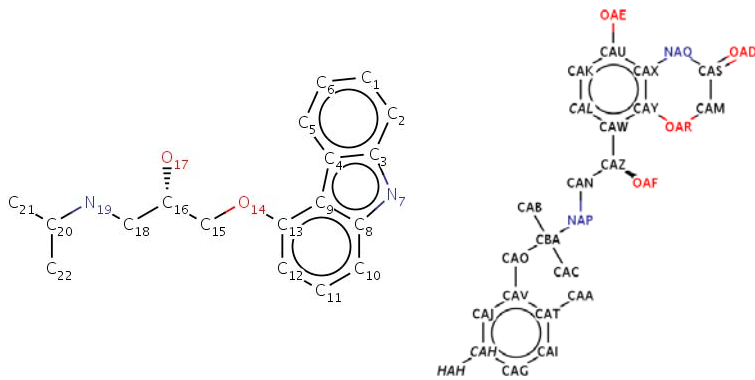


그림 1. (왼쪽) β1AR(2ycw)와 β2AR(2rh1)의 antagonist 역할을 하는 CAU, (오른쪽) β2AR(3sn6)의 agonist 역할을 하는 P0G

결과 및 토의

전하(charge)에 따른 RMSD(Root-mean square deviation; 생물정보학에서 분자간의 거리를 판단할 때 사용하는 단위)²² 와 energy 에 대한 결과는 표 1 과 표 2 에 나타나 있다. 총 4 번의 GalaxyDock²¹ 을 이용한 120 가지의 결과가 전하(charge)별로 계산되었고,

각각의 예측 결과와 실제 계(native system)의 거리를 RMSD 로 나타내어 비교하였다.

2rh1 은 β 2AR 단백질이고 이것과 CAU 사이의 상호작용을 알아보았다. CAU 는 antagonist 이다.⁷ 그렇기 때문에 방해(blocking)하는 역할을 하게 된다. 표 1 의 2rh1-CAU 의 관계를 보면 가장 낮은 RMSD 수치와 그 때의 energy rank 가 관련이 있음이 보인다. 그 중에서도 정밀한 계산이 이었던 RESP_2 의 결과가 좋았다.

2ycw⁶ 는 β 1AR 단백질이고 이것과 antagonist 인 CAU 사이의 상호작용을 알아보았다. 그런데 Pullman 전하(charge)를 제외한 실험적 전하(empirical charge)에서는 energy 가 양수로 나왔다. Energy 가 양수로 나오는 것은 의미 없는 수치이기 때문에 이 계(system)에 대해서는 전하(charge)와 상호작용에 대한 예측이 효과적인 결과를 주지 못함을 알 수 있다. 특히, SYBYL8.1⁹로 계산한 실험적인 전하(DR, FC, GH, GM, H, P)의 경우 그림 1 의 N₇에 붙어있는 수소원자에 -1.0000 이라는 실제로 존재 할 수 없는 전하를 띠게 하였다. 이렇듯 실험적인 전하(empirical charge)에서는 극단적이 비현실적 전하(unrealistic charge)를 띠는 경우가 있었다.

3sn6 는 β 2AR 단백질과 그것의 agonist 인 POG 의 상호관계를 보여주는 단백질 구조이다.⁸ 이것과 단백질 리간드 상호관계 예측 결과는 Formal charge 를 제외한 경우 가장 낮은 RMSD 값을 가질 때에 상당히 낮은 energy 순위를 가짐을 알 수 있었다. 즉 예측결과와 실제가 유사하였음을 알 수 있었다.

표 1 가장 작은 RMSD를 가질 때 energy와 energy의 순위

	2rh1-CAU			2ycw-CAU			3sn6-POG		
Charge	lowest rmsd	Energy	Energy rank	lowest rmsd	Energy	Energy rank	lowest rmsd	Energy	Energy rank
DR	1.324177	-10.442	9/120	2.681947	6.549	102/120	2.627663	-2.404	11/120
FC	1.394636	-11.325	8/120	2.946054	33.724	104/120	2.918456	1.564	44/120
GH	1.40479	-10.826	19/120	3.148513	153.21	111/120	3.025177	-1.443	19/120
GM	1.496334	-11.651	3/120	3.38072	191.6	109/120	3.075522	-2.861	3/120
H	1.668616	-11.401	4/120	2.997294	1.469	86/120	2.845281	-1.837	8/120
P	1.258308	-11.286	11/120	2.618894	-3.126	84/120	2.454145	-1.15	9/120
resp_1	2.096733	-8.877	46/120	3.207059	-10.56	16/120	2.838043	-4.045	11/120
resp_2	1.501753	-11.372	3/120	3.136233	-6.776	91/120	3.587209	-1.707	16/120

각각의 정전기적 포텐셜(electrostatic potential)에 다른 종류의 전하(charge)를 띠게 하였을 때 가장 낮은 에너지와 RMSD 를 비교한 것을 표 2 에서 확인 할 수 있다.

GPCR 계에서 CAU 와 POG 는 같은 위치에서 상호작용이 일어나게 된다. 그런데 CAU 의 경우 방향족 고리(aromatic ring)을 갖기 때문에 POG 가 더욱 강한 극성(more polar)을 띠게 된다.²³ 즉, CAU 가 포함된 계(system)보다 POG 가 포함된 계(system)이 더 전하의 크기의 변화에 따라 상호관계 예측이 많이 변하게 된다.

표 2 가장 energy가 낮을 때의 RMSD 수치

표 2	2rh1-CAU		2ycw-CAU		3sn6-P0G	
	lowest energy	RMSD	lowest energy	RMSD	lowest energy	RMSD
DR	-12.092	1.612962	-11.494	4.762623	-5.331	3.140329
FC	-11.765	5.932061	-12.712	4.748197	-3.254	9.148023
GH	-12.481	5.998509	-12.276	4.325226	-3.467	4.445501
GM	-11.765	1.538015	-11.705	4.797662	-4.426	8.88573
H	-11.591	6.533489	-12.26	4.459837	-2.868	8.566912
P	-12.474	1.32932	-11.841	4.545552	-2.344	4.136153
resp_1	-11.753	5.246029	-13.024	4.028759	-6.331	3.567944
resp_2	-12.878	5.73839	-12.052	4.384795	-5.372	7.284664

단백질과 리간드 간의 상호작용을 예측 하는 데 이용한 GalaxyDock 은 단백질에 전하(charge) 뿐만 아니라 쿨롱에너지, 반데르발스, 레너드 존스 포텐셜 등과 같은 다양한 요소(term)를 사용²¹ 하였기 때문에 전하(charge)의 영향을 덜 받는 2 가지 계(system)에 대해서는 상대적으로 더 좋지 않은 결과를 나오게 되었다.

결론

3 가지 계(system)에 대해서만 연구를 하였기 때문에 일반적인(general) 결론을 내기는 어렵다. 3 가지 계에서 양자 계산(Quantum mechanical calculation)이 보다 좋아 보이는 경향성을 갖지만, 어떤 양자계산 방식이 더 나은 가를 입증 할 수 는 없었다. 그래도

RESP(restrained electrostatic potential) 방식을 사용하면 단백질과 리간드 간의 상호관계 예측(Protein Ligand Interaction prediction)에 도움이 된다는 사실을 알 수 있었다.

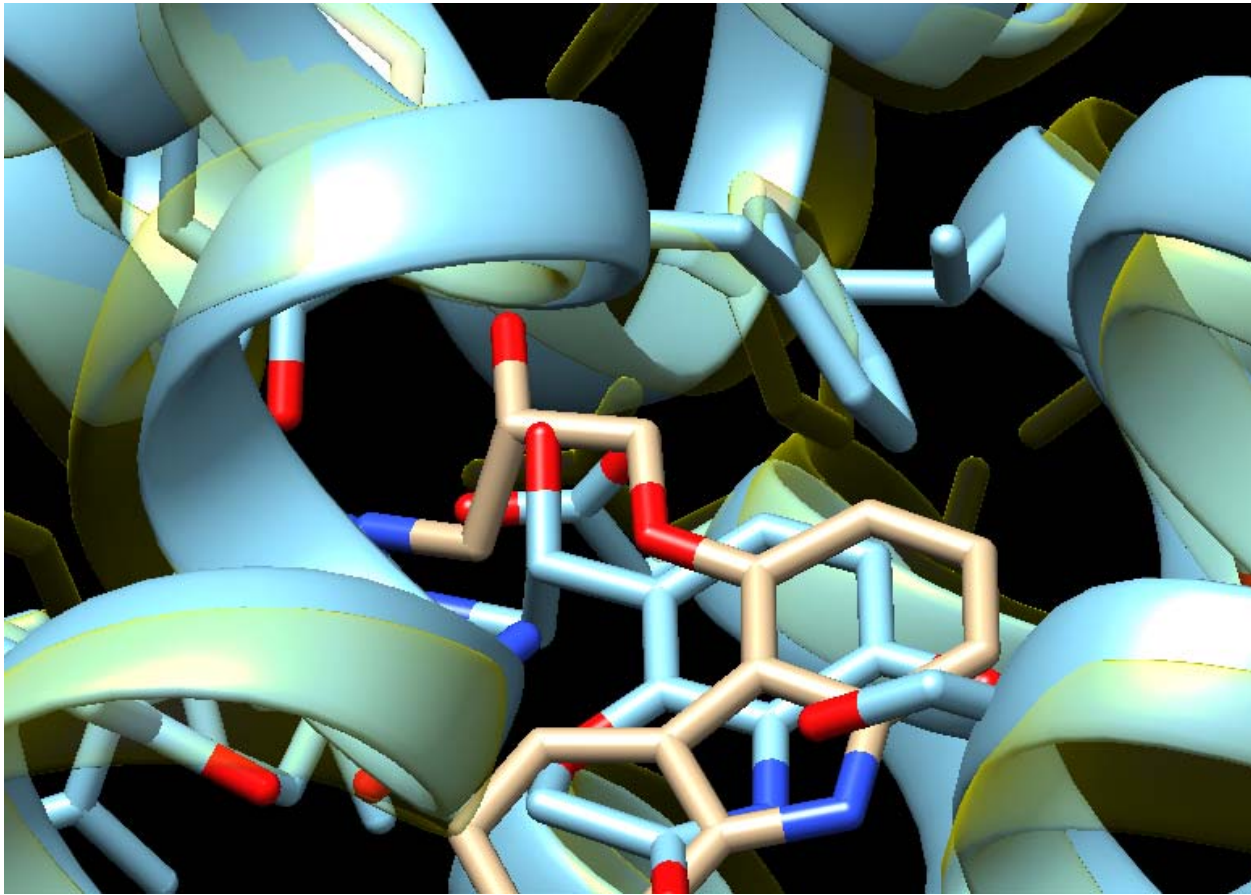


그림 2. 하늘색은 단백질과 POG, 베이지색은 CAU. Chimera 1.7rc²⁴를 통하여 확인한 구조.

특히 그림 2 을 보면 β2AR 단백질에서 antagonist 인 리간드 CAU 와 agonist 인 POG 가 각각 같은 단백질의 잔기(residue)의 부사슬(sidechain)과 반응하고 있음을 볼 수 있다. 이러한 상황으로 보아 상대적으로 리간드가 더 극성(polar)을 갖는 경우 더 정확한 전하(charge)를 띠게 한 후 상호작용 예측(interaction prediction)을 하는 데 유리하였다.

단백질에 대한 양자역학적인 계산(Quantum mechanical calculation)이 없이 오로지 리간드에 대한 계산만 하였기 때문에 상대적으로 정확성이 떨어지는 예측 결과를

보여주었다. 그리고 3 가지 계(system)을 통한 연구를 통해서도 일반적인 경향성을 알기에는 부족함이 많았다. 하지만 이 연구에서 극성이 강한 리간드의 경우 실험적인 전하(empirical charge)보다 양자역학적인 계산(Quantum mechanical calculation)을 통해 얻은 RESP(restrained electrostatic potential)이 정확한 단백질과 리간드간의 상호작용 예측(Protein Ligand interaction prediction)에 도움이 된다는 것을 알 수 있었다.

감사의 말. 이 연구는 EDISON 의 Chemworks2 프로그램의 Gaussian09 와 GalaxyDock 을 사용하여 진행되었고, 서울대학교 자연과학대학 화학부 계산생물학 연구실의 기자재를 지원받아 수행되었습니다.

참고문헌

1. “The Nobel Prize in Chemistry 2012 – Press Release”. Nobelprize.org 6 Apr(2012)
2. D. M. Rosenbaum, S. G. F. Rasmussen, B. K. Kobilka, The structure and function of G-protein-coupled receptors., *Nature* **459**, 356–63 (2009).
3. K. Tsai *et al.*, A comparison of different electrostatic potentials on prediction accuracy in CoMFA and CoMSIA studies, *European Journal of Medicinal Chemistry*, (2010)
4. M.A. Lill *et al.*, Multi-dimensional QSAR in drug discovery, *Drug Discovery Today* Volume 12, Issues 23–24, 1013–1017(2007)
6. R. Moukhametzianov *et al.*, Two distinct conformations of helix 6 observed in antagonist-bound structures of a β 1 -adrenergic receptor, *Proc. Natl. Acad.Sci(USA)*, 2–6 (2011).

7. V. Cherezov *et al.*, High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor., *Science (New York, N.Y.)* **318**, 1258–65 (2007).
8. S. G. F. Rasmussen *et al.*, Crystal structure of the β 2 adrenergic receptor-Gs protein complex., *Nature* **477**, 549–55 (2011).
9. Tripos Inc., St Louis, MO 2008 USA
10. G. Delre, B. Pullman, T. Yonezawa, Electronic Structure of the Alpha-Amino Acids of Proteins. I. Charge Distributions and Proton Chemical Shifts., *Biochimica et biophysica acta* **75**, 153–82 (1963).
11. J. Gasteiger, M. Marsili, Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges, *Tetrahedron* **36**, 3219–3228 (1980).
12. B. Pullman, N. Gresh, H. Berthod, A. Pullman, Cation Binding to Biomolecules Conformational Effects on the Phosphodiester Linkage and the Polar Head of Phospholipids, **163**, 151–163 (1977).
13. Gaussian Inc., Wallingford, CT 2009 USA
14. A. D. Becke, Density-functional thermochemistry. III. The role of exact exchange, *The Journal of Chemical Physics* **98**, 5648 (1993).
20. K. Kim, K. D. Jordan, Comparison of Density Functional and MP2 Calculations on the Water Monomer and Dimer, doi:10.1021/j100091a024.

16. James B. Foresman, Exploring Chemistry with Electronic Structure Methods, 2nd ed (1996)
17. D. R. Hartree, The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part I. Theory and Methods, *Mathematical Proceedings of the Cambridge Philosophical Society* **24**, 89 (2008).
18. W. D. Cornell *et al.*, A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules, *Journal of the American Chemical Society* **117**, 5179–5197 (1995).
19. Y. Duan *et al.*, A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations., *Journal of computational chemistry* **24**, 1999–2012 (2003).
20. T. H. Dunning Jr., Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen, *J. Chem. Phys.*, **90** (1989) 1007–23.
21. W.-H. Shin, C. Seok, GalaxyDock: protein-ligand docking with flexible protein side-chains., *Journal of chemical information and modeling* **52**, 3225–32 (2012).
22. Coutsias EA, Seok C, Dill KA. Using quaternions to calculate RMSD. *J Comput Chem* **25** (15): 1849–1857.(2004)
23. DL Nelson, MM Cox, Lehninger Principles of Biochemistry, *WHFreeman*, 5th ed,

(2008)

24. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. *J Comput Chem.* Oct;25(13):1605-12.(2004)