http://dx.doi.org/10.5012/bkcs.2013.34.1.299

Modeling Partial Atomic Charge of Organic Molecule and Mutated Amino Acid in a Protein-Ligand Complex for Molecular Mechanics Simulation

Kyung Lae Park

Department of Pharmaceutics, College of Pharmacy, Chungnam National University, Daejeon 305-764, Korea
E-mail: parki@cnu.ac.kr
Received August 11, 2012, Accepted October 24, 2012

Key Words : Molecular mechanics, Force field, Partial atomic charge, Protein-ligand complex

The atoms of organic molecules and standard amino acids in protein have specific atomic attributes, such as partial charge, depending on the environmental condition and protonation status. In most of the molecular mechanics algorithm and force field these atomic properties of organic molecules are therefore categorized according to the atoms and bond types and defined as constant values in the internal database. The partial atomic charge is one of these sorts of properties which are given as constant parameters in every modeling algorithm.

In some case these limited number of atom types are not sufficient to model the system adequately. In case of amino acids there exist not only 20 standard types but also 'mutated' or covalently bound amino acids which cannot be well described with the standard parameter values in a given force field algorithm. In this case a reasonable readjustment of atomic charges is necessary to represent a balanced distribution of total charge in a molecular unit.

In recent investigation of molecular interactions in catechol oxidase with its inhibitor,[1] we found that a histidine sidechain was covalently bound to a cystine monomer and this unusual adduct acted as a stable electron donor to copper atom which again were complexed by an *N*-phenylthiourea ligand.

For the molecular mechanical simulation of this complex it was necessary to predefine the adequate atomic charge of involving nonstandard histidine-cystine adduct and *N*-phenylthiourea derivatives as ligand molecule. In this work we use the framework of AMBER program package [2] to find the consistent atomic charges of the components mentioned.

Methods

The atomic charge is one of electronic properties which can only be precisely calculated quantum mechanically. For a wide range of modest sized organic molecules J. Steward proposed AM1 (Austin Model 1) which is one of the most widely used semiempirical calculation algorithm for the atomic charges.[3] Based on the AM1 model A. Jakalian *et al.* proposed a modified AM1-BCC(Bond Charge Corrections) model to generate high quality atomic charges within the fast and efficient calculation time.[4] Actually the charges adopted in "generalized AMBER force field(GAFF)"[2] is generated using HF/6-31* RESP (Restrained Electro Static Potential) charges, but AM1-BCC was parameterized to reproduce HF/6-31* RESP charges for a variety of organic molecules. Jakalian *et al.* had further globalized the AM1-BCC method using 2755 organic molecules as test sets and reported the validation applied to virtually all types of compounds in The Merck Index and NCI Database.[5] In case of large scale molecular simulation the charges are to be calculated by AM1-BCC and fitted for the general consistency with the GAFF in the framework of AMBER molecular mechanics algorithm.

Results and Discussion

Editing Structure. The ligand *N*-phenythiourea in the catechol oxidase and the ring substituents are the first targets of the charge calculation and the atom naming was adopted from the X-ray structure 1bug.pdb [1] as is shown in Figure 1. As the ring substituents we introduced tertiary butyl group on 5 free positions; *para* substituent was on carbon C6 (Structure **2**), *ortho* on C2 (**3**) and *meta* on C1 (**5**) atom respectively.

For the purpose of the molecular dynamics simulation we calculated also the *t*-butyl substituent on counter positions to **3** and **5** and notated as **4** and **6** respectively. Because of the rotation freedom of the bond N1-C3 the structure **3** and **4** and the structure **5** and **6** are chemically identical with the same atomic charge distribution, but the structures are

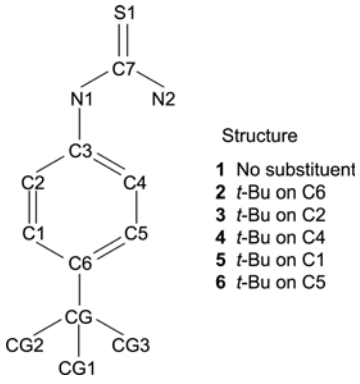

Figure 1. Atomic notations in *N*-*p*-*t*-butylphenylthiourea (**2**) and the name of other structures (**1**, **3** to **6**).

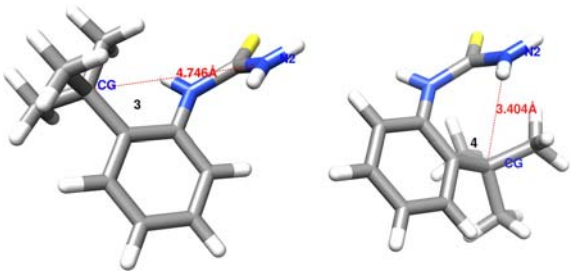

Figure 2. The two *N-o-t*-butylphenylthioureas **3** and **4** show different conformations due to the hindered rotation around the N1-C3 bond.

conformationally different firstly because of the bulky urea group on the one hand and secondly in the complex of these compounds within the cavity of the protein. The thiourea and *t*-butylphenyl group may be involved to different interaction part of the protein, which can make the N1-C3 single bond to a non-flexible bond (Figure 2). The two rotational conformers may then show different interaction patterns within the protein complex.

The source coordinates of the ligand *N*-phenylthiourea (**1**) was filtered from the pdb code and the *t*-butyl group was edited in the code and the vacant valence was filled with hydrogen using default bond geometry. The model building utility of the program CHIMERA[6] was used, where the coordinate values of heavy atoms and therefore the original conformation was conserved. The conformation of the two *N-o-t*-butylphenylthiourea was distinguishable with the distance N2-CG of 4.746 Å (**3**) and 3.404 Å (**4**), as shown in Figure 2, and the corresponding value of the *meta*-analogue were 6.438 Å (**5**) and 5.330 Å (**6**), respectively. The *N-p-t*-butylphenylthiourea has the distance of 6.734 Å (**2**) from N2 to CG atom.

Calculation of Charge. The antechamber program in the AMBER11 package was used to execute the MOPAC[7]

Table 1. The calculated and adjusted atomic charge of *N*-phenylthiourea (Structure **1**) and *t*-butyl derivatives (**2**, **3**, **4**, **5** and **6**). The naming of hydrogen is synchronized to the name of host atom and the bold-faced atoms and charges show the change of atomic charge due to the substation at different position

Structure **1**			Structure **2**			Structure **3**			Structure **4**			Structure **5**			Structure **6**		
Atom	AM1-BCC	Ad-justed	Atom	AM1-BCC	Ad-justed	Atom	AM1-BCC	Ad-justed	Atom	AM1-BCC	Ad-justed	Atom	AM1-BCC	Ad-justed	Atom	AM1-BCC	Ad-justed
N1	-0.4451	-0.450	N1	-0.4431	-0.450	N1	-0.4501	-0.450	N1	-0.4461	-0.450	N1	-0.4441	-0.450	N1	-0.4441	-0.450
H11	0.3565	0.350	H11	0.3565	0.350	H11	0.3605	0.350	H11	0.3595	0.350	H11	0.3555	0.350	H11	0.3555	0.350
N2	-0.6260	-0.626	N2	-0.6260	-0.626	N2	-0.6270	-0.626	N2	-0.6260	-0.626	N2	-0.6260	-0.626	N2	-0.6260	-0.626
H21	0.3135	0.314	H21	0.3135	0.314	H21	0.3135	0.314	H21	0.3145	0.314	H21	0.3145	0.314	H21	0.3145	0.314
H22	0.3455	0.345	H22	0.3445	0.345	H22	0.3445	0.345	H22	0.3445	0.345	H22	0.3455	0.345	H22	0.3445	0.345
C1	-0.1090	-0.105	**C1**	**-0.1100**	**-0.110**	**C1**	**-0.1240**	**-0.110**	**C1**	**-0.1200**	**-0.110**	**C1**	**-0.0413**	**-0.039**	**C1**	**-0.1100**	**-0.110**
H1	0.1390	0.139	**H1**	**0.1390**	**0.138**	**H1**	**0.1380**	**0.138**	**H1**	**0.1380**	**0.138**	**CG**	**-0.0127**	**-0.008**	**H1**	**0.1380**	**0.138**
C2	-0.1510	-0.150	C2	-0.1490	-0.150	C2	-0.0623	-0.069	C2	-0.1550	-0.150	C2	-0.1520	-0.140	C2	-0.1540	-0.150
H2	0.1420	0.142	**H2**	**0.1410**	**0.141**	**CG**	**-0.0087**	**-0.008**	**H2**	**0.1440**	**0.141**	**H2**	**0.1420**	**0.138**	**H2**	**0.1420**	**0.138**
C3	0.0306	0.026	C3	0.0236	0.026	C3	0.0276	0.026	C3	0.0266	0.026	C3	0.0296	0.026	C3	0.0296	0.026
C4	-0.1580	-0.150	C4	-0.1570	-0.150	C4	-0.1540	-0.150	**C4**	**-0.0693**	**-0.069**	C4	-0.1610	-0.150	**C4**	**-0.1580**	**-0.140**
H4	0.1420	0.142	**H4**	**0.1430**	**0.141**	**H4**	**0.1440**	**0.141**	**CG**	**-0.0067**	**-0.008**	**H4**	**0.1430**	**0.141**	**H4**	**0.1440**	**0.141**
C5	-0.1120	-0.105	**C5**	**-0.1120**	**-0.110**	**C5**	**-0.1220**	**-0.110**	**C5**	**-0.1180**	**-0.110**	**C5**	**-0.1130**	**-0.110**	**C5**	**-0.0453**	**-0.039**
H5	0.1390	0.139	**H5**	**0.1400**	**0.138**	**H5**	**0.1390**	**0.138**	**H5**	**0.1410**	**0.138**	**H5**	**0.1370**	**0.138**	**CG**	**-0.0127**	**-0.008**
C6	-0.1370	-0.131	**C6**	**-0.0693**	**-0.057**	**C6**	**-0.1310**	**-0.135**	**C6**	**-0.1340**	**-0.135**	**C6**	**-0.1350**	**-0.135**	**C6**	**-0.1350**	**-0.135**
H6	0.1390	0.139	**CG**	**-0.0097**	**-0.008**	**H6**	**0.1380**	**0.138**	**H6**	**0.1370**	**0.138**	**H6**	**0.1410**	**0.138**	**H6**	**0.1410**	**0.138**
C7	0.5934	0.595	C7	0.5944	0.595	C7	0.5974	0.595	C7	0.5954	0.595	C7	0.5944	0.595	C7	0.5944	0.595
S1	-0.6024	-0.614	S1	-0.6064	-0.614	S1	-0.6184	-0.614	S1	-0.6164	-0.614	S1	-0.6074	-0.614	S1	-0.6064	-0.614
			CG1	-0.0861	-0.091	CG1	-0.0881	-0.091	CG1	-0.0901	-0.091	CG1	-0.0881	-0.091	CG1	-0.0881	-0.091
			HG11	0.0417	0.040	HG11	0.0407	0.040	HG11	0.0387	0.040	HG11	0.0397	0.040	HG11	0.0397	0.040
			HG12	0.0357	0.040	HG12	0.0357	0.040	HG12	0.0447	0.040	HG12	0.0387	0.040	HG12	0.0377	0.040
			HG13	0.0387	0.040	HG13	0.0397	0.040	HG13	0.0417	0.040	HG13	0.0377	0.040	HG13	0.0397	0.040
			CG2	-0.0861	-0.091	CG2	-0.0901	-0.091	CG2	-0.1001	-0.091	CG2	-0.0861	-0.091	CG2	-0.0861	-0.091
			HG21	0.0357	0.040	HG21	0.0357	0.040	HG21	0.0437	0.040	HG21	0.0397	0.040	HG21	0.0407	0.040
			HG22	0.0397	0.040	HG22	0.0417	0.040	HG22	0.0417	0.040	HG22	0.0407	0.040	HG22	0.0317	0.040
			HG23	0.0387	0.040	HG23	0.0397	0.040	HG23	0.0447	0.040	HG23	0.0367	0.040	HG23	0.0387	0.040
			CG3	-0.0891	-0.091	CG3	-0.0911	-0.091	CG3	-0.0901	-0.091	CG3	-0.0861	-0.091	CG3	-0.0871	-0.091
			HG31	0.0367	0.040	HG31	0.0357	0.040	HG31	0.0387	0.040	HG31	0.0387	0.040	HG31	0.0387	0.040
			HG32	0.0387	0.040	HG32	0.0507	0.040	HG32	0.0377	0.040	HG32	0.0367	0.040	HG32	0.0427	0.040
			HG33	0.0397	0.040	HG33	0.0407	0.040	HG33	0.0407	0.040	HG33	0.0417	0.040	HG33	0.0387	0.040
Sum	0.000	0.000	Sum	-0.0030	0.000	Sum	-0.0040	0.000	Sum	0.0010	0.000	Sum	0.0000	0.000	Sum	-0.0010	0.000

program. The input pdb file was converted into Sybyl mol2 format which contains the bond and atom type information necessary for the calculation. We performed the semi-empirical calculation for the original *N*-phenylthiourea (**1**) and the phenyl derivatives, **2**, **3**, **4**, **5**, and **6** and the calculated charge values of the derivatives are listed in column AM1-BCC of Table 1.

Fitting the Total Charge. The resulting charges shows the expected electro-static potential of general organic molecules, which can otherwise be found in the standard force field parameter database. The substitution of *t*-butyl group on phenyl ring shows the change of the charge distribution beyond the ring system. The aromatic carbon atoms are normally more polar than the aliphatic carbon with average charge between −0.10 and −0.15 for C1, C2, C4, C5 and C6 of *N*-phenylthiourea (**1**). The substitution on carbon C1 (**5**), C2 (**3**), C4 (**4**), C5 (**6**) and C6 (**2**) lowers the aromatic polarity to −0.04 ~ −0.07 and the amount of excessive charge will be distributed over the butyl group.

The individual atoms of different isomers or conformers, however, show slightly different values. This discrepancy cannot be explained by the substituent positions alone. The sulfur atoms of thiourea group, for examples, show different values in all 6 conformers, what might be a computational artifact in the semiempirical calculation. The existence of artifacts can be also seen in the total charge of nonzero residual values. From these results of calculation the charges are to be generalized within the 6 conformers depending on the bonding environments and neighboring atoms to represent a consistent physical property.

The start point of fitting procedure is the hydrophobic *t*-butyl group. Firstly, from 5 AM1-BCC calculations the most probable charge values of *t*-butyl group was deduced and thus the total charge of *t*-butyl group was set to +0.079, which is the sum of three CG's of −0.091 and 9 HG's of charge 0.040 plus −0.008 from one CG, introduced direct to the phenyl ring. In the second step, the calculation result of unsubstituted phenylthiourea was adopted beside of small correction for the chemically identical atoms. The atomic charges thiourea group were set identical for all 6 molecules. The charge of substituted atoms, C1, C2, C4, C5 and C6 was set consistently according to the positional character. The *ortho* carbon C2 and C4 have the value of −0.069, the *meta* carbon C1 and C5 have −0.039 and *para* carbon C6 was set to −0.057. Also starting from the AM1-BCC calculation the charges of the rest atoms of phenyl ring were corrected partly to give a total charge of zero. The final fitted charge values are listed in Table 1.

Nonstandard Amino Acid. The catechol oxidase receptor protein contains nine different types of histidine residue depending on the neighboring groups. The crystallization medium was reported neutral[1] and therefore all histidine residues should be protonated either ND1 or NE2 of imidazole ring. The structural analysis revealed that NE2 of three imidazole rings served as electron donors to one copper atom and the two copper atoms are complexed to NE2 atoms of 6 histidines. The protonation position at these 6 histidines is therefore ND1 atom of imidazole rings. Besides of NE2 atoms the sulfur atom of the phenythiourea also donates the two electron pairs to two copper atoms and thus two copper atoms are complexed tetrahedrally with the distance of about 2.3 Å around the copper atoms.

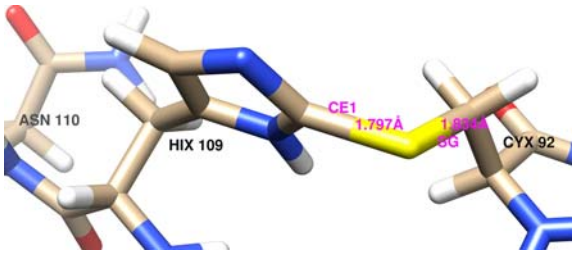

Figure 3. New type of CE1-SG bond between histidine 109 (HIX) and cystine 92(CYX) with bond length of 1.802 Å, which is shorter than of the length CB-SG (1.817 Å) of CYX.

One of the 6 histidine had an additional special feature in the imidazole ring. The CE1 atom of normal histidine has HE1 hydrogen as 4th valence, however the CE1 atom of histidine 109 is bonded to sulfur atom of monomeric cystine residue from the neighborhood. The former 5 of 6 histidines can be designated as HID as in standard AMBER convention, but the later histidine has a new CE1-SG bond with bond length of 1.802 Å instead of CE1-HE1 bond of normal histidine. The last type of histidine was named here as HIX and the atomic charges were to be redefined for the new type of histidine and cystine bridge bond as in Figure 3.

There are also another type of histidine with the NE2

Table 2. The comparison of atomic charge of histidine residue in different protonation and bonding states. In HIE the NE2 atom is protonated and in HID and HIX the ND1 atom is protonated. The CE1 atom of imidazole ring of HIX residue is bonded to sulfur atom of neighboring CYX residue

HIE		HID		HIX	
Atom	Charge	Atom	Charge	Atom	Charge
N	-0.4157	N	-0.4157	N	-0.4157
H	0.2719	H	0.2719	H	0.2719
CA	-0.0581	CA	0.0188	CA	0.0188
HA	0.1360	HA	0.0881	HA	0.0881
CB	-0.0074	CB	-0.0462	CB	-0.0462
HB2	0.0367	HB2	0.0402	HB2	0.0402
HB3	0.0367	HB3	0.0402	HB3	0.0402
CG	0.1868	CG	-0.0266	CG	-0.0266
ND1	**-0.5432**	**ND1**	**-0.3811**	**ND1**	**-0.3463**
CE1	**0.1635**	**HD1**	**0.3649**	**HD1**	**0.3649**
HE1	**0.1435**	**CE1**	**0.2057**	**CE1**	**0.2753**
NE2	**-0.2795**	**HE1**	**0.1392**	-	-
HE2	**0.3339**	**NE2**	**-0.5727**	**NE2**	**-0.5379**
CD2	-0.2207	CD2	0.1292	CD2	0.1292
HD2	0.1862	HD2	0.1147	HD2	0.1147
C	0.5973	C	0.5973	C	0.5973
O	-0.5679	O	-0.5679	O	-0.5679
Sum	0.0000	Sum	0.0000	Sum	0.0000

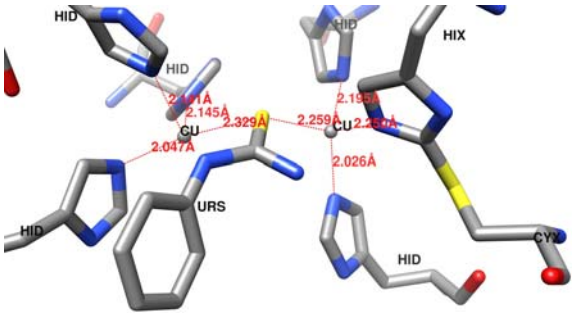

Figure 4. The tetrahedral orientations electron donors around copper atoms. The copper on the left is surrounded by three HIDs and sulfer and the right copper is complexed by two HIDs, one HIX and sulfer atom of *N*-phenylthiourea. The distance of copper revealed 2.250 Å to less polar NE2 of HIX, whereas the average distance to NE2 of five HIDs was 2.111 Å.

atoms protonated, but these three HIE are not involved in the copper and phenylthiourea complex.

The atomic charge of the HID and HIE were taken from the AMBER force field as is. The charges of HIXwere based on the values in HID. The HE1 atom was removed from the coordinates and charge was distrituted to the neighboring atoms. 50% to the CE1 atom and 25% to the ND1 and NE2 atoms of the imidazole ring. The final values are listed in Table 2.

As a result of the redistribution of charge 0.1392 of HE1 the CE1 atom became more polar and ND1 and NE2 atom of HIX less polar than in the HID residue. The more detailed analysis of the *N*-phenylthiourea and protein complex in 1bug.pdb showed the difference of this polarity in the complex bond. The bond distance of positive Copper to less polar NE2 (–0.5379) of HIX (2.250 Å) is slightly longer than with other NE2 (–0.5727) of 5 HID residue (average 2.111 Å). The tetrahedral arrangement of ligand is shown in Figure 4, where sulfur atom of *N*-phenylthiourea (URS) in the middle serves two electron pairs to two copper atoms.

Conclusions

The calculated and fitted atomic partial charges of the *N*-phenylthiourea and 5 derivatives represent reasonable charge distributions which show mostly consistent values with the GAFF of standard AMBER. Also the intuitive fitting by redistribution of atomic charge of 'lost' atom HE1 of histidine is consistent with structural status in X-ray structure of catechol oxidase.

The proposed atomic charge can be used as standard force field parameters in AMBER for the diverse molecular mechanics simulation.

Based on this study we performed various molecular dynamics simulations extensively with the reliable and confirming results which will be reported in the ongoing publications.

References

1. Klabunde, T.; Eichen, C.; Sacchettini, J.C.; Krebs, B. *Nat. Struct. Biol.* **1998**, *5*, 1084.
2. Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Wang, B.; Hayik, S.; Roitberg, A; Seabra, G.; Kollosvary, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Goehlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER11*, 2010; University of California, San Francisco.
3. Stewart, J. J. P. *J. Comp. Aided Mol. Design* **1990**, *4*, 1.
4. Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. *J. Computat. Chem.* **2000**, *21*, 132.
5. Jakalian, A.; Jack, D. B.; Bayly, C. I. *J. Computat. Chem.* **2002**, *23*, 1623.
6. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Computat. Chem.* **2004**, *25*, 1605.
7. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.