Comparative Homology Modeling and Ligand Docking Study of Human Catechol-O-Methyltransferase for Antiparkinson Drug Design

Jee-Young Lee and Yangmee Kim*

Department of Chemistry and Bio/Molecular Informatics Center, Konkuk University, Seoul 143-701, Korea
**E-mail: ymkim@konkuk.ac.kr*
Received July 25, 2005

Catechol-O-methyltransferase (COMT, EC 2.1.1.6) is an S-adenosylmethionine (SAM, AdoMet) dependent methyltransferase, and is related to the functions of the neurotransmitters in various mental processes, such as Parkinson's disease. COMT inhibitors represent a new class of antiparkinson drugs, when they are coadministered with levodopa. Based on x-ray structure of rat COMT (rCOMT), the three dimensional structure of human COMT (hCOMT) was constructed by comparative homology modeling using MODELLER. The catalytic site of these two proteins showed subtle differences, but these differences are important to determine the characterization of COMT inhibitor. Ligand docking study is carried out for complex of hCOMT and COMT inhibitors using AutoDock. Among fifteen inhibitors chosen from world patent, nine models were energetically favorable. The average value of heavy atomic RMSD was 1.5 Å. Analysis of ligand-protein binding model implies that Arg201 on hCOMT plays important roles in the interactions with COMT inhibitors. This study may give insight to develop new ways of antiparkinson drug.

Key Words : Methyltransferase, COMT, SAM, Parkinson's disease, Homology modeling

Introduction

Addition of methyl group to biologically active molecules such as hormones, and proteins causes a change in the physicochemical properties of the molecules and biological methylations play indispensable roles in cellular functions.[1] O-methylation of endogenous catecholamine and other catechols is catalyzed by the catechol-O-methyltransferase (COMT, EC 2.1.1.6.). Since COMT transfers a methyl group from S-adenosylmethionine (SAM, AdoMet)[2] to *meta*- or *para*- hydroxyl groups of the catechol[3,4] COMT is called as SAM-dependent methyltransferase.[5] The physiological substrates of COMT include a wide variety of catechol estrogens and endogenous and exogenous catecholamine (dopamine, nor-epinephrine and epinephrine). It plays important roles in the metabolism of these catechol substrates. Specially, in the brain, COMT is related to the functions of dopamine and epinephrine in various mental processes, such as Parkinson's disease.[6] COMT inhibitors represent a new class of antiparkinson drugs, when they are coadministered with levodopa.[8,9]

COMT is composed of two domains such as a soluble domain (S-COMT) with 221 residues and a membrane-bound domain (MB-COMT) with an additional 50 residues at the N-terminus.[10] Active site of COMT consists of a SAM binding site and an actual catalytic site located in S-COMT. The Mg^{2+}, which is bound to COMT catalytic site, converts the hydroxyl groups of the catechol substrate to be more easily ionizable and make it possible bind tightly to the catechol moiety.[3] COMT proteins are distributed in various mammalian tissues and these are encoded in various mammalian species by a single gene, such as human, rat, mouse, dog and pig. Sequence identity of these five species is over 70%.[11] The 2.0 Å resolution crystal structure of rat COMT (rCOMT) complexed with its cosubstrate SAM and a novel inhibitor shows the atomic interactions between the important residues at the active site and the inhibitor.[12] The structure of human COMT (hCOMT) is expected to be similar to that of rCOMT, because the amino acid sequence identity is about 80%.

In this study, three dimensional structure of hCOMT was determined by comparative homology modeling. The crystal structure of rCOMT was used as a template protein for homology modeling. Based on this structure, we studied ligand-protein binding model for fifteen COMT inhibitors collected from world patent, and the interactions between the hCOMT and inhibitors were investigated in this paper.

Methods

Comparative Protein Structure Modeling. The amino acid sequences of the hCOMT soluble domain was retrieved from Expasy.[13] We built structure of hCOMT using comparative homology modeling based on the structure of rCOMT. These molecular modeling procedures were carried out on an OCTANE R12000 Silicon Graphics workstation.

The structurally conserved regions (SCRs) were determined by pairwise sequence alignment with the Insight/Homology module as shown in Figure 1. The X-ray structure of rCOMT at 2.0 Å resolution (PDB entry 1H1D) is the only known structure of mammalian COMT and it was used as a structural template. For a given alignment, five comparative models of the target sequence were built by MODELLER,[14] applying the default model building routine 'model' with fast refinement. This procedure is advantageous because one can select the best model from several candidates. Further-

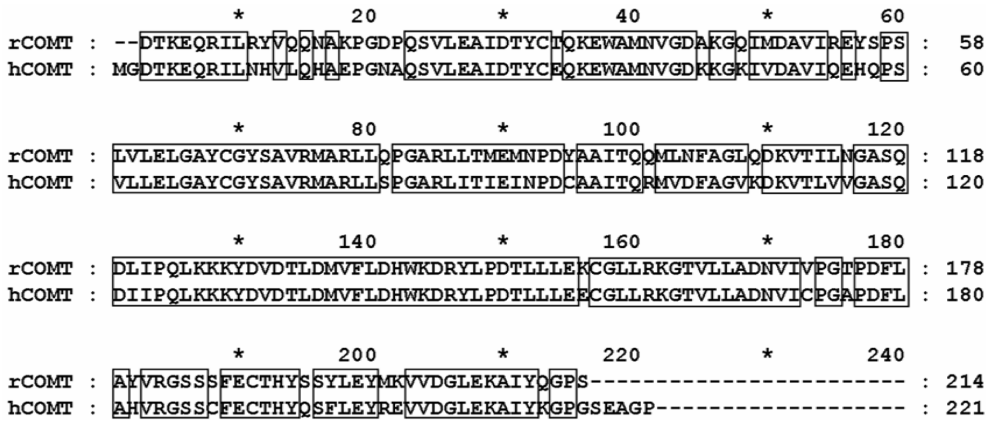

Figure 1. The sequence alignment of the template protein (rCOMT) and the target protein (hCOMT). Black Boxes denote the structurally conserved regions (SCRs).

more, the variability among the models can be used to evaluate the reliability of the modeling. Energy minimization (steepest descent and conjugated gradient algorithms; gradient on energies less than 1 kcal/mol used as convergence criteria) was performed using the consistent valence force field and the Discover program. The qualities of these models were analyzed by PROCHECK.[15]

Ligand Docking. Fifteen molecules shown in Figure 2 were collected as COMT inhibitors from world patent and these were docked using AutoDock[16] to hCOMT structure determined by comparative homology modeling. The Lamarckian Genetic Algorithm (LGA) of the Autodock 3.05 was used for docking experiments. Distance-dependent function of the dielectric constant was used for the calculation of the energetic maps and all other parameters were used by default value. We carried out 150 and 250 independent docking processes for each complex. The metal ion, Mg^{2+}, in the active site of hCOMT was modeled using amber force field provided in AutoDock 3.05.

Results and Discussion

Three Dimensional Structure of hCOMT. Five models of hCOMT were generated by MODELLER. Energy and

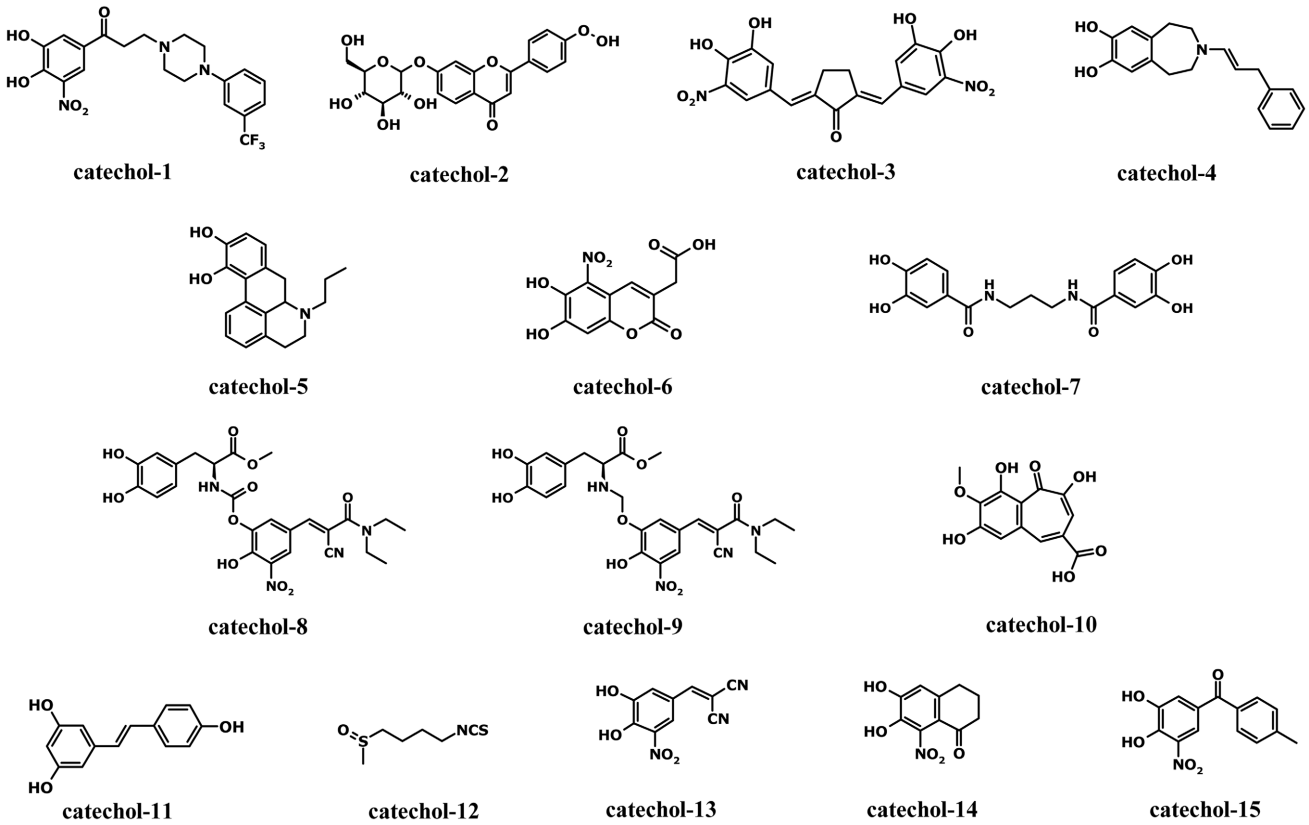

Figure 2. Two dimensional structures of fifteen COMT inhibitors from world patent.

Table 1. RMSD and energy of five hCOMTs predicted by MODELLER

	RMSD with rCOMT	Energy (kcal)
hCOMT1	0.333219	1025.70
hCOMT2	0.194803	1122.49
hCOMT3	0.158170	1067.70
hCOMT4	0.370857	1084.06
hCOMT5	0.303863	1092.19

RMSD for five hCOMT models were listed in Table 1. The five generated models of hCOMT are represented in Figure 3(a) and structural alignment of rCOMT and hCOMT is shown in Figure 3(b). Three dimensional structures of structurally conserved regions (SCRs) were very similar for each five models of hCOMT. Among these five hCOMT models, the lowest energy structure was hCOMT-1 shown in Figure 3(c). In order to select the best model, we checked the structural validity of hCOMT by PROCHECK. The torsion angles of φ and ψ in the generated models are represented in Ramachandran plot as shown in Figure 4. These torsion angles of 95.9% of the residues had values within the most favored regions and only 0.5% of the residues had values within disallowed regions and the overall G-factor is 0.19 as shown in Table 2. The overall G-factor[19] is a measure of the overall normality of the structure and low G-factors indicate that residues have unlikely conformations. The overall value is obtained from an average of G-factors for all residues in the structure. X-ray structure of rCOMT has a resolution of 2.0 Å and a G-factor of 0.34 Å. In Ramachandran plot, the stereochemical quality of a protein model can be judged by the use of φ, ψ scatter plots, with incorrect structures generally having a much larger fraction of residues lying in disallowed regions.[18] Since x-ray structure of rCOMT at 2.0 Å resolution has also 0.5% of its residues in disallowed regions, it can be said that our hCOMT structure satisfies criteria of a good model.

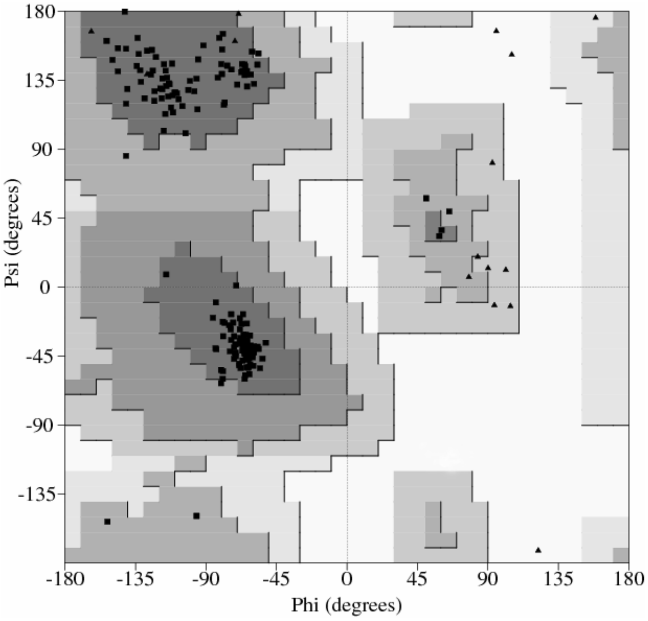

Figure 4. Ramachandran plot of hCOMT1 obtained by PROCHECK.

Ligand Docking of COMT inhibitors. Most of COMT inhibitors include catechol ring. These series of inhibitors were called as 1st generation inhibitors. The most famous drugs as COMT inhibitors are Entacapone[8] and Tolcapone,[9] which also include catechol ring. From world patent, we collected fifteen COMT inhibitors which also include catechol ring. Secondary structures of these inhibitors, catechol series, are shown in Figure 2.

Ligand docking study was carried out for complex of hCOMT-1 and fifteen COMT inhibitors using AutoDock. AutoDock showed that among fifteen inhibitors, nine inhibitor (from catechol-1 to catechol-9) binding models were energetically favorable. Since the average value of heavy atomic RMSD was 1.5 Å. We confirmed that the hCOMT model from comparative modeling is acceptable.

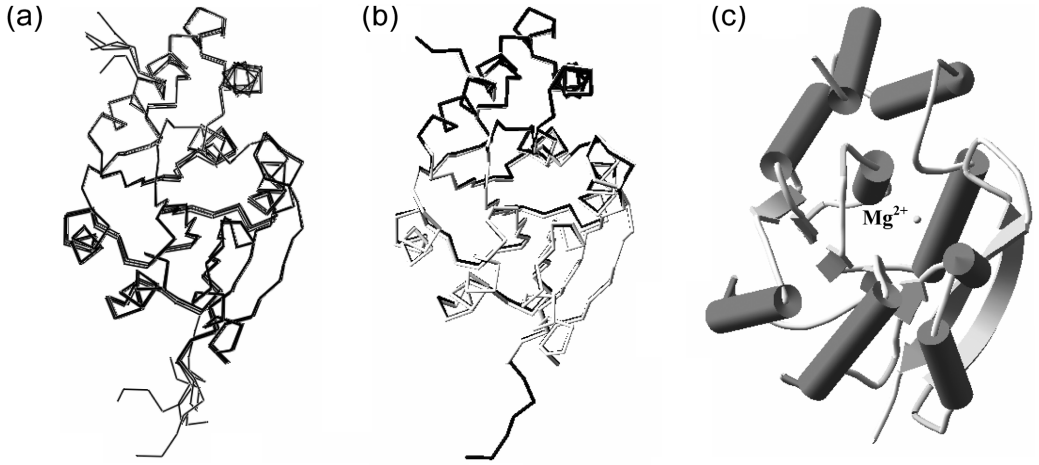

Figure 3. (a) Ribbon representation of five hCOMT structures determined by MODELLER. (b) Structure alignment of rat and human COMT. White ribbon is rCOMT and black ribbon is hCOMT. (c) Secondary structure representation of hCOMT-1. The arrow and dot circle indicate the active site of hCOMT.

Table 2. Quality of structures checked by PROCHECK

Structure	Ramachandran plot quality (%)			Overall G-factor (Å)
	Core	Allowed	Disallowed	
hCOMT1	95.9	3.6	0.5	0.19
hCOMT2	93.8	5.7	0.5	0.16
hCOMT3	94.3	4.7	0.5	0.12
hCOMT4	93.8	5.2	0.5	0.17
hCOMT5	94.3	4.1	0.5	0.15

Table 3. *In Vitro* activities of hCOMT inhibitors, collected by world patent

Inhibitor	IC50 (μM)[a]	Ki (μM)[b]	Docking energy (Kcal/mol)	RMSD (Å)
Catechol-1	2.5		−8.2	0.8
Catechol-2	3.6		−8.0	1.0
Catechol-3	2.9		−7.7	0.6
Catechol-4			−6.4	1.7
Catechol-5	22		−6.6	1.7
Catechol-6	7.3		−6.3	1.4
Catechol-7	4.3		−6.7	1.5
Catechol-8			−6.4	1.8
Catechol-9		6.0	−7.8	1.6
Catechol-10	10		−4.0	2.3
Catechol-11	5.3		−2.1	1.9
Catechol-12	5.5		−6.3	2.0
Catechol-13	20		−4.3	2.1
Catechol-14	−	0.8	2.0	2.7
Catechol-15	−	−	2.3	2.5

[a]COMT inhibitors for the potentiation of levodopa therapy in the treatment of Parkinson's disease. [b]*In vitro* activities reported for the inhibition of binding of inhibitors at COMT.

The docking results of the fifteen catechol inhibitor complexes are listed in Table 3.

In x-ray structure of rCOMT, catechol-1 was also the inhibitor molecule. It is a typical example, where the good agreement between the docked structure and crystal structure is established. The RMSD between these two is 0.8 Å. The metal binding region, namely catechol ring moiety of catechol-1 is a little bit perturbed. Especially the piperazine moiety of the molecule is somewhat moved into hCOMT docked model while the overall active sites of hCOMT and rCOMT are very similar. Only two residues such as Cys173 and Arg 201 as listed in Table 4 are the exceptions. Based on homology, eight residues were conserved in inhibitor binding site of rCOMT and hCOMT as shown in Table 4. Three out of eight residues such as Asp141, Asn170 and Glu199, formed complicated hydrogen bonds with catechol moiety of inhibitor. One hydroxyl group of the catechol ring showed interactions with Mg^{2+} and Asp141, and the other hydroxyl group formed hydrogen bond with Glu199. Trp38, Met40, and Trp143 play an important role on hydrophobic interactions. Catechol ring can bind to these hydrophobic residues in hCOMT binding site.

There are two residues which are different between hCOMT and rCOMT inhibitor binding site. Arg201 in hCOMT is replaced for Met201 in rCOMT and Cys173 in hCOMT is replaced for Val173 in rCOMT. In case of rCOMT, the binding site of piperazine moiety of catechol-1 is exposed to the surface of rCOMT and hydrophobic residues, such as Met201 and Val173, are located at the surface of rCOMT. Therefore, it is difficult to determine interactions between rCOMT and inhibitor. However, replacements of these two residues in hCOMT could give reasonable explanation for protein-inhibitor interactions. Cys173 and Arg201 in hCOMT form polar environment near the piperazine binding site. Nitrogen atom of piperazine ring on catechol-1 can form hydrogen bonding with the side chain of Arg201 with a distance 3.2 Å. These common features were found in the results of 9 favorable docking models. The carboxyl group of catechol-7 was near by Arg201, and oxygen atom of chromen ring in catechol-2 was also close to Arg201. The interaction models between hCOMT and inhibitor, catechol-1 are represented in Figure 5. Interaction between hCOMT-1 and catechol-1 is shown in Figure 5(a) and interaction between rCOMT and catechol-1 is shown in Figure 5(b). Surface model of hCOMT structure is shown in Figure 5(c). Catechol-1 depicted by orange color shows the binding structure in x-ray structure of rCOMT and one depicted by white color shows the docking structure in hCOMT. Interactions between inhibitors and Arg201 are shown in Figure 5(d). The circle denoted by a magenta dotted line in Figure 5(d) represented the possible region for Arg201 to be a hydrogen bonding acceptor. We confirmed that in nine docking models Arg201 forms these kinds of hydrogen bondings. These interactions caused to move inhibitors in docking models of hCOMT. There are not yet known biological evidences or experimental data to support these phenomena. Therefore, it can be assumed that these charged environments can play important roles to help ligands to be fixed and to overcome the binding site exposure in hCOMT.

Table 4. Residues in inhibitor binding site of rCOMT and hCOMT

	Residues
rCOMT	Trp38, Met40, Asp141, Trp243, Lys144, Asn170, **Val173**, Glu199, **Met201**, Val203
hCOMT	Trp38, Met40, Asp141, Trp243, Lys144, Asn170, **Cys173**, Glu199, **Arg201**, Val203

*Residues underlined are the dissimilar residues between rCOMT and hCOMT.

Conclusion

Three dimensional structure of human catechol-O-methyl-transferase (hCOMT) was determined by comparative homology modeling. The x-ray structure of rat COMT was used as a template protein. From the result of ligand docking for fifteen COMT inhibitors, nine models were energetically favorable. The average RMSD values of heavy atoms were 1.5 Å. In the complex model of hCOMT and inhibitors, the

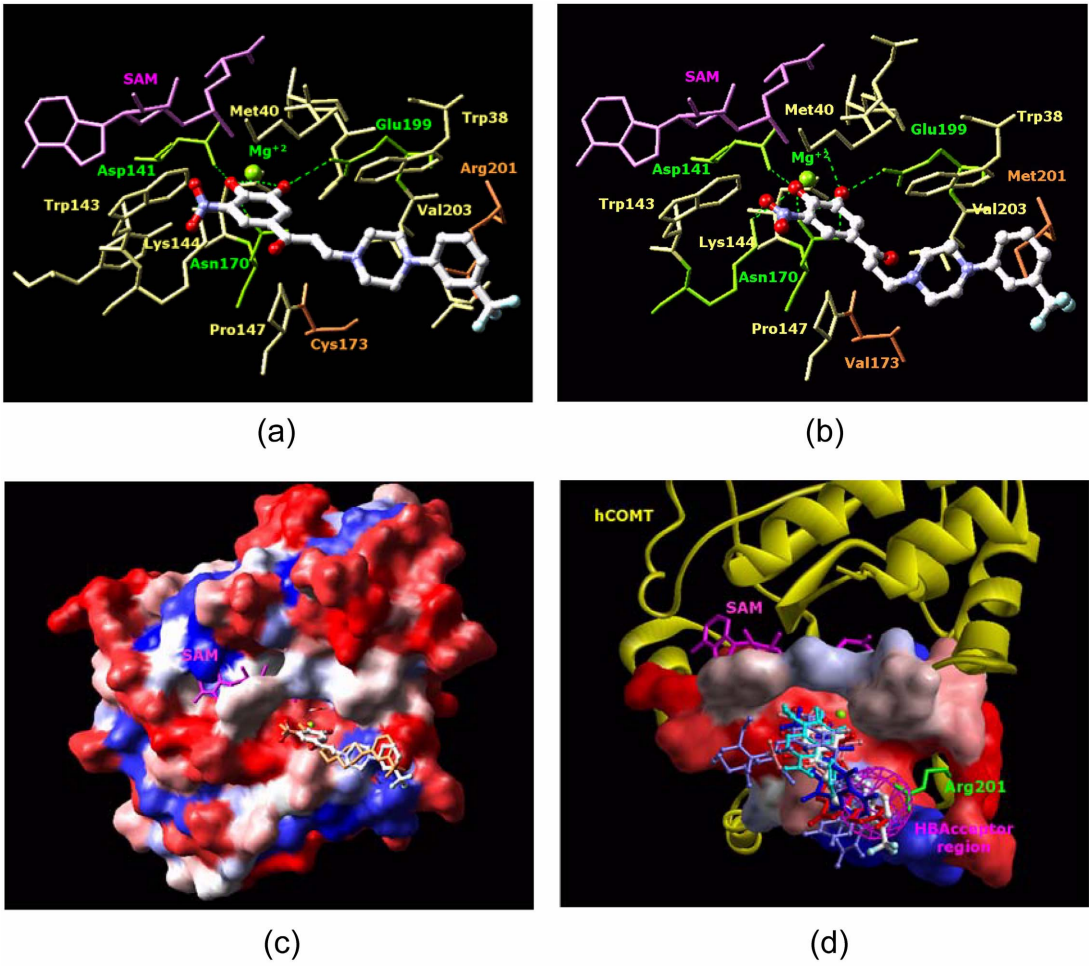

Figure 5. Representation for interaction model between hCOMT and inhibitor. (a) Interactions between hCOMT-1 and inhibitor, catechol-1. (b) Interactions between rCOMT and catechol-1. (c) Surface model of hCOMT structure. Catechol-1 depicted by orange color shows the binding structure in x-ray structure of rCOMT and one depicted by white color shows the docking structure in hCOMT. (d) Docking model of hCOMT and nine inhibitors. The binding site including the Arg 201 are shown.

metal ion, Mg^{2+}, and some charged residues, Asp141 and Glu199, formed complicated hydrogen bondings. Trp38, Met40, and Trp143 play important roles on making hydrophobic environments. Nitro group of catechol ring in most of the inhibitors was bound to Trp143 and Lys 144. From the analysis of binding model, we concluded that two residues, Arg201 and Cys173, in inhibitor binding site of hCOMT have important meaning for fixing the flexibility of inhibitor in order to overcome the binding site exposure.

In this study we have demonstrated that two residues are important for hCOMT activity and this information awaits further research for discovery of new antiparkinson drug candidate by *in silico* screening and NMR spectroscopy. This study may give insight to develop a novel antiparkinson drug.

Acknowledgment. This work was supported financially by Bio/Molecular Informatics Center of Konkuk University (KRF2004-F00019). This work was supported, in part, by a Molecular and Cellular BioDiscovery Research Program grant from the Ministry of Science and Technology of Korea.

References

1. Zhou, Z. S.; Zhao, G.; Wan, W. *Frontiers of Biotechnology & Pharmaceuticals*; Science press Ltd.; 2003; vol. 2, p 256.
2. Sistla, S.; Rao, D. N. *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 1.
3. Männisto, P. T.; Kaakkola, S. *Pharmacol. Rev.* **1999**, *51*, 593.
4. Zhu, B. T. *Curr. Drug Metabol.* **2002**, *3*, 321.
5. Martin, J. L.; McMillan, F. M. *Curr. Opin. Struc. Biol.* **2002**, *12*, 783.
6. Kaakkola, S. *Drugs* **2000**, *59*, 1233.
7. Sangwon, L.; Yangmee, K. *Bull. Korean Chem. Soc.* **2004**, *25*, 838.
8. Nissinen, E.; Kaheinen, P.; Penttila, K. E.; Kaivola, J.; Linden, I. B. *Europ. J of Pharmacol.* **1997**, *340*, 287.
9. Offen, D.; Pane, H.; Galili-Mosberg, R.; Melamed, E. *Clinic. Neuropharmacology* **2001**, *24*, 27.
10. Männisto, P. T.; Reenilä, I. *Medical Hypotheses* **2001**, *57*, 628.
11. Tilgmann, C.; Ulmanen, I. *J. of Chromatography B* **1996**, *684*, 147.
12. Bonifácio, M. J.; Archer, M.; Rodrigues, M. L.; Matias, P. M.; Learmonth, D. A.; Carrondo, M. A.; Soares-Da-Silva, P. *Mole. Pharmacol.* **2002**, *62*, 795.
13. Wilkins, M. R.; Gasteiger, E.; Bairoch, A.; Sanchez, J. C.; Williams, K. L.; Appel, R. D.; Hochstrasser, D. F. *Methods Mol.*

Biol. **1999**, *112*, 531.
14. Marti-Renom, M. A.; Stuart, A.; Fiser, A.; Sánchez, R.; Melo, F.; Sali, A. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291.
15. Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Cryst.* **1993**, *26*, 283.
16. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. W.; Olson, A. J. *J. Computational Chemistry* **1998**, *19*, 1639.
17. Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577.
18. Pal, D.; Chakrabarti, P. *Biopolymers* **2002**, *63*, 195.
19. Morris, A. L.; MacArthur, M. W.; Hutchinson, E. G.; Thornton, J. M. *Proteins* **1992**, *12*, 345-364.