A Structure-Based 3D-QSAR(CoMSIA) Study on a Series of Aryl Diketoacids (ADK) as Inhibitors of HCV RNA-dependent RNA Polymerase

Jinyoung Kim, Jin Hee Han, and Youhoon Chong[*]

Division of Biosciences and Biotechnology, Konkuk University, Seoul 143-701, Korea. [*]E-mail: chongy@konkuk.ac.kr
Received August 9, 2006

The diketoacid moiety of aryl α,γ-diketoacids (ADKs, Fig. 1) proved essential for antiviral activity against hepatitis C virus (HCV) RNA-dependent RNA polymerase (RdRp), and structure-activity optimization studies have led to the identification of an ADK inhibitor as one of the most potent HCV RdRp inhibitors reported.[1] However, in spite of the extensive structure-activity relationship study, it has not been clear what controls the binding affinity of ADK analogues to the target enzyme.[2]

3D-QSAR techniques, such as comparative molecular force field analysis (CoMFA)[3] and comparative molecular similarity indices analysis (CoMSIA),[4] are based on the experimental structure-activity relationship on specific biomolecule-ligand pair. This method is based only on the ligand structure and thus the spatial arrangement (or alignment) is crucial in determining the accuracy of these approaches. In case of ADK analogues, it is obvious that the common structural unit, diketoacid moiety, could be used as a template for atom based alignment. However, the atom-based fit (using atoms of diketoacid) produced poor CoMFA and CoMSIA results (data not shown) presumably due to the highly flexible alkyl linker which connects two aromatic rings of ADK analogues (Fig. 1). This result implies that ADK analogues bind to the target enzyme in a characteristic active conformation, which cannot be identified by investigation of the ligand structures alone. Thus, ADK analogues should be docked into the binding site of the target enzyme to provide the active conformations, which can be used for the structure-based alignment. However, the lack of information about the binding site of ADK analogues at the target enzyme limits this approach. In this study, the binding site of ADK analogues in the HCV RdRp was proposed by using an unusual crystal structure of rUTP-HCV RdRp complex (PDB ID 1GX6)[5] and structural similarity between rUTP and ADK. The ADK analogues were aligned by docking into the binding site, and a structure-based 3D-QSAR study was performed to correlate the biological activities of ADKs with their three-dimensional structures.

The diketoacid moiety is famous for its metal-binding ability,[6] and ADK analogues are known to bind the divalent metal ions at the active site of HCV RdRp.[7,8] However, it remains unsolved how ADK analogues bind to the active site of a polymerase enzyme without formation of the complementary base pairing with the RNA template chain. Additionally, it has been speculated that ADK analogues might have different binding site around the active site of

HCV RdRp. Recently, Bressanelli *et al.*[5] reported the unusual crystal structure of rUTP bound to HCV RdRp and it shows that rUTP binding site is quite different from the active site with the base (uracil) hydrogen atoms bonded to the polypeptide main chain (PDB ID 1GX6).[5] This alternative binding mode of rUTP is artificial in the sense that rUTP cannot bind in such a way in the presence of template, but it is conceivable that molecules with higher binding affinity to the rUTP-binding site in this mode can inhibit the catalytic activity of the enzyme. Thus, we set out to investigate the characteristic binding mode of rUTP, which resulted in construction of a pharmacophore model composed of three key interactions between rUTP and HCV RdRp (Fig. 2): (a) electrostatic interaction with two divalent metal ions (Mn^{2+}), (b) H-bonding of triphosphate moiety to nearby amino acid residues (Phe224, Asp225), (c) H-bonding between uracil
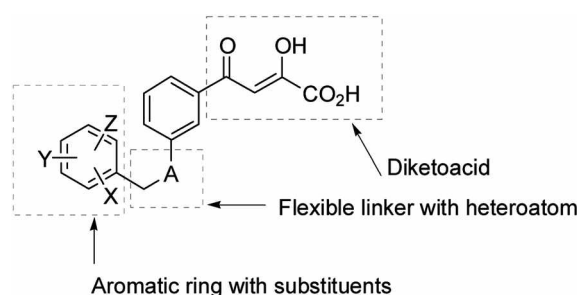

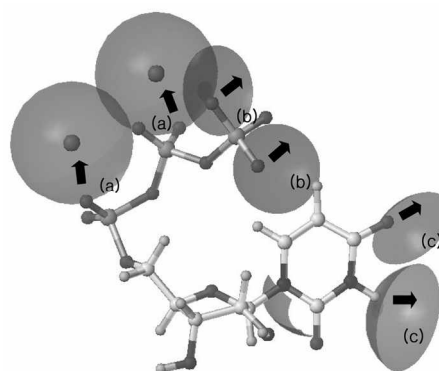Figure 1. General structure of ADK analogues.

Figure 2. Most critical interactions between rUTP and HCV RdRp used as guidelines for docking by FlexX-Pharm: (a) Electrostatic interaction with Mn^{2+}; (b) H-bonding with Phe224 and Asp225; (c) H-bonding with Leu159. Shaded dishes indicate rUTP-binding site of the enzyme characterized by these conditions.
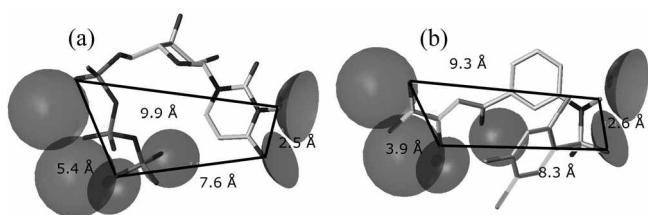
Figure 3. Comparison of the pharmacophores of (a) rUTP and (b) ADK (1). Shaded dishes indicate specific interactions around the pharmacophore with the enzyme residues: electrostatic interaction with divalent metal ions, H-bond acceptors, and H-bond donors.

base and Leu159. To our surprise, the generated pharmacophore has an interesting match with an ADK molecule of which diketoacid moiety and H-bond donor/acceptor are in good three dimensional agreements with the triphosphate and uracil base of the rUTP, respectively (Fig. 3). Based on the structural similarity of the pharmacophore model of rUTP and ADK, we assumed that ADK might bind to the same site of rUTP found in the crystal structure of 1GX6. Thus, twenty structurally distinct ADK analogues taken from the literature[2a] (Fig. 4) were docked into the rUTP binding site by the pharmacophore-guided docking[9] (FlexX-Pharm, SYBYL 7.2)[10] protocol to generate structure-based alignment of ADK analogues. The results of the superimposed images of 20 ligand structures docked at the ADK binding site of the HCV RdRp are shown in Figure 5.

With the structure-based ligand alignment in hand, we attempted 3D-QSAR study by using CoMSIA method. The

basic principle of CoMSIA is the same as that of CoMFA, but CoMSIA includes some additional descriptors such as hydrophobicity, hydrogen bond donor and hydrogen bond acceptor.[4] In this study, CoMSIA was chosen as the method of choice in preference to CoMFA because both hydrophobic and hydrogen bonding interaction fields were included in the CoMSIA model, which might play a key role in determining the binding affinity of ADK analogues to the HCV RdRp. The regression analysis of CoMSIA field energies was performed using the partial least squares (PLS) algorithm with the leave-one-out (LOO) method adopted for cross validation. The results obtained from the PLS analysis are summarized in Table 1. CoMSIA with 20 molecules in the training set produced a cross-validated r^2 of 0.966 with minimum standard error of estimation (SEE, 0.215) (Table 1). This analysis was used for the final non-cross-validated run, giving a good correlation coefficient (q^2 value of 0.525) (Table 1). Table 1 also shows high hydrophobic contribution (0.530) to the CoMSIA model compared with hydrogen bond donor or acceptor, which suggests that the contribution from the hydrophobic interaction is the key to the binding affinity of ADK analogues. Moreover, as the diketoacid moiety and hydrogen bond donor (or acceptor) are the common structural units of ADK analogues (Fig. 1), the contribution of hydrophobic field to the final CoMSIA model becomes even more important.

The actual and calculated inhibitory activities and the residual values for training set molecules are given in Table 2, and the plot of actual pIC_{50} values versus predicted pIC_{50}
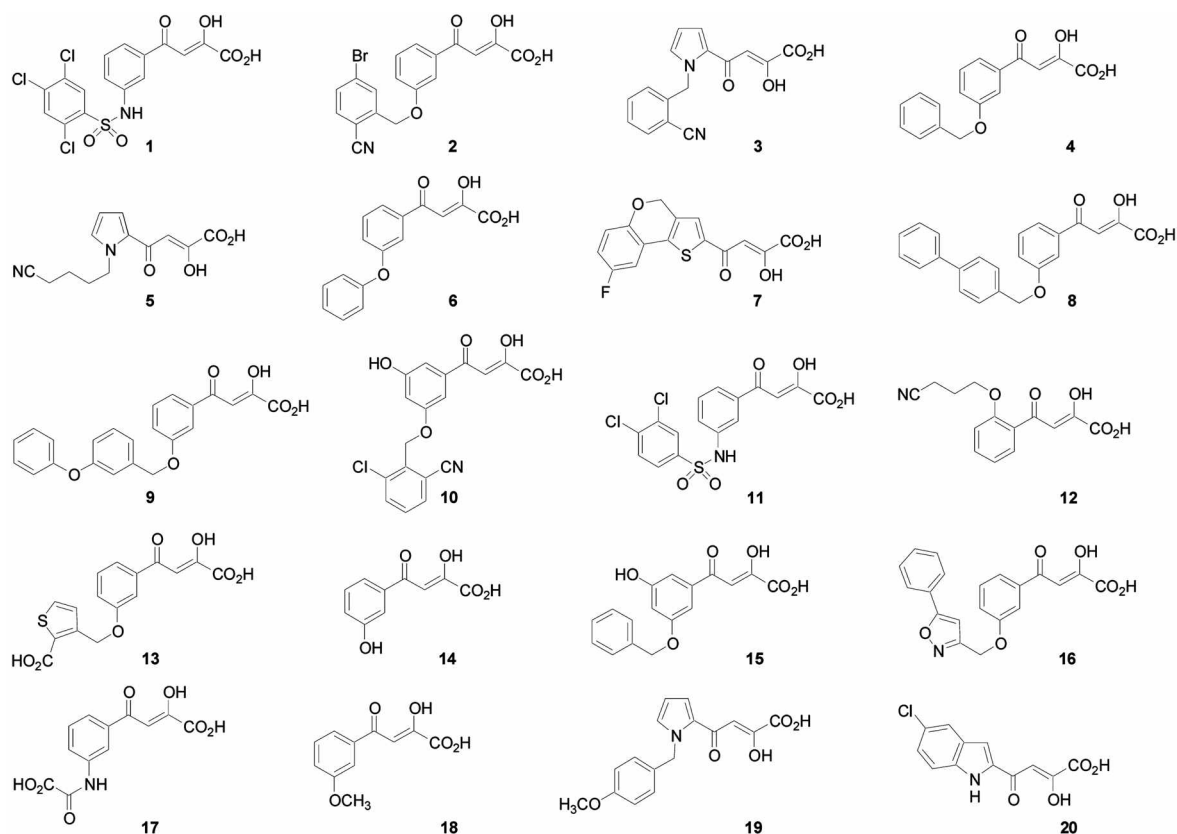

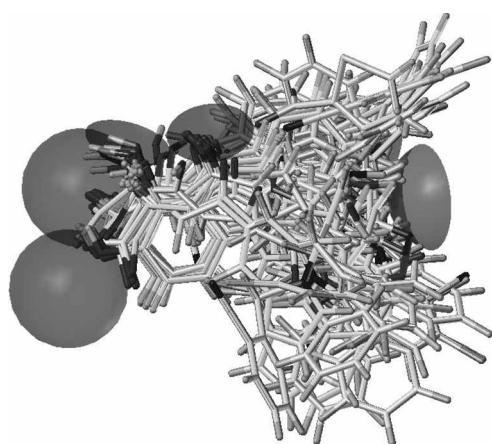**Figure 4.** Structures of ADK analogues docked at HCV RdRp by FlexX-Pharm.

Figure 5. Superimposed images of 20 ADK analogues after FlexX-Pharm docking.

Table 1. The results of PLS analysis in the Training Set using CoMSIA

	Field	HDA[a]
	r^2	0.966
PLS	Components	4
Analysis	SEE	0.215
	q^2	0.525
	Hydrophobic	0.530
Contribution	H Bond Donor	0.228
	H Bond Acceptor	0.242

[a]H = hydrophobic, D = hydrogen bond donor, A = hydrogen bond acceptor

Table 2. Actual and predicted activities (pIC$_{50}$) of the training set molecules

	pIC$_{50}$				pIC$_{50}$		
Compd.	Actual	Pred.	Residual	Compd.	Actual	Pred.	Residual
1	7.83	8.04	−0.21	11	6.85	6.69	0.16
2	7	6.67	0.24	12	6.42	6.61	0.26
3	5.2	5.35	−0.15	13	6.02	5.86	0.16
4	5.1	5.04	0.06	14	6	5.9	0.1
5	4.89	5.02	−0.13	15	5.85	5.94	−0.09
6	4.72	4.85	−0.13	16	5.22	5.21	0.01
7	4.57	4.5	0.07	17	4.72	4.52	0.2
8	4.3	4.25	0.05	18	4.4	4.95	−0.55
9	4.3	4.31	−0.01	19	5.23	5.13	0.1
10	7	7.15	−0.15	20	5.12	5.12	0
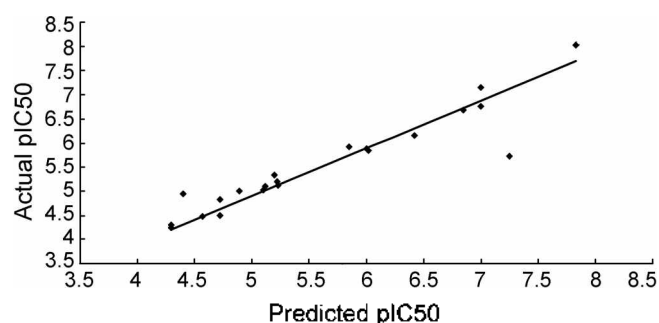
Figure 6. Comparison of actual vs predicted pIC$_{50}$.

is shown in Figure 6. As a result, both of them indicate good predictivity of the CoMSIA model.

Graphical representations of CoMSIA maps obtained by the field type "stDev*coeff" are displayed in Figure 7. The contour maps were superimposed on the most active compound **1** shown as a capped stick. The regions where hydrogen bond donor or acceptor is associated with enhanced affinity are found near the diketoacid moiety and heteroatom linker part (amino or sulfonamino group) (Figs. 7a and 7b). Thus, hydrogen bonding interaction seems to be essential for ADK analogues to bind to the target enzyme. However, the hydrogen bonding interaction is not likely to be involved in fine-tuning of binding affinities of ADK analogues as the diketoacid and heteroatom linker moieties are the common structural units of ADK analogues (Fig. 1). On the other hand, dark grey contour region where hydrophobicity is associated with enhanced affinity is located near the aromatic ring, which suggests that substitution of the aromatic ring with bulky substituents would enhance the binding affinity of ADK analogues (Fig. 7c). It is also worth while to note that the light grey contour region where hydrophobicity is associated with diminished affinity is found around the diketoacid as well as the heteroatom linker part, which suggests that substitutions at these positions would not be beneficial (Fig. 7c).

To further validate our results, four compounds[2a] with pIC$_{50}$ range between 4.77 and 7.25 which were not included in the training set were assigned as test set molecules and their biological activities were predicted from the PLS equation derived from CoMSIA model. Predicted and actual activities of test set molecules are summarized in Table 3. Predicted pIC$_{50}$ values agree well with the experimental ones ($r_{pred}^2 = 0.797$) with average deviation of 0.44, which suggests that our model is good for prediction of the pIC$_{50}$
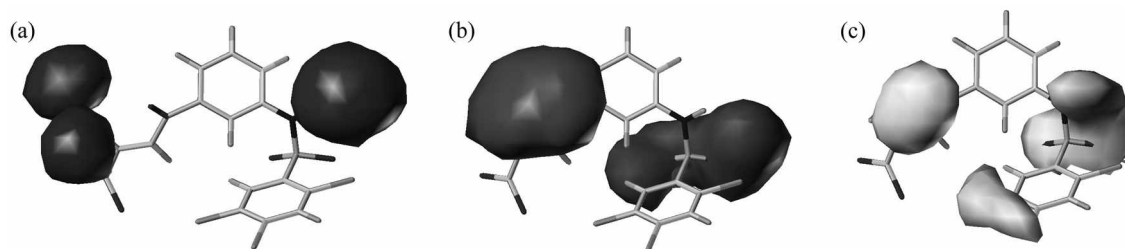

Figure 7. (a) Superposition of the CoMSIA hydrogen bond donor contour plots; (b) Superposition of the CoMSIA hydrogen bond acceptor contour plots; (c) Superposition of the CoMSIA hydrophobic contour plots.
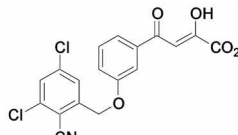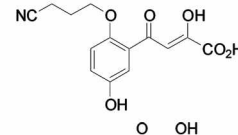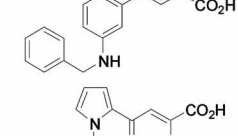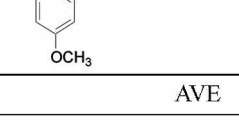
Table 3. Actual and predicted activities (pIC_{50}) of test set molecules

Compd.	Structure	Actual pIC_{50}	CoMSIA	
			Predicted pIC^{50}	Residual
T1		7.25	6.74	0.51
T2		6.96	6.27	0.69
T3		4.77	5.01	−0.24
T4		5.85	5.54	0.31
	AVE			0.44

values of ADK analogues against HCV RdRp.

In summary, a CoMSIA model constructed by structure-based 3D-QSAR study could be successfully applied to predict the biological activity of ADK analogues. The binding affinity of ADK analogues are found to be highly dependent upon the hydrogen bonding interaction as well as hydrophobic interaction around the aromatic ring of ADK analogues. In particular, the CoMSIA model proposes that the hydrophobic aromatic ring play a key role in determining the antiviral activity of ADK analogues. Thus, hydrophobic substituents around the aromatic ring reinforce hydrophobic interaction with the target enzyme, whereas the lack of aromatic substitution and thereby insufficient size of the inhibitor molecule can be primarily ascribed to their inability to bind to the hydrophobic binding site.

Experimental Section

All calculations were done on a Linux Enterprise OS using SYBYL 7.2 software packages.[10]

Data set. ADK molecules were prepared using the sketch module in the SYBYL package and conformational searches were performed by a grid search, which calculates energies by systematically changing the dihedral angles of each ligand using standard TRIPOS force field.[11] The lowest energy structures were selected as conformers for the pharma-cophore-guided docking (FlexX-Pharm) studies. Finally, all ligands were fully optimized using the standard TRIPOS force field with Gasteiger-Hückel charges[12] until the energy gradient converged below 0.05 Kcal/mol.

Docking with FlexX and FlexX-Pharm. Three dimensional crystal structure of HCV RdRp was obtained from the

protein data bank (PDB code 1GX6). Active site of the enzyme was defined as all the residues within 15' of the bound ligand and two metal ions. Standard parameters were used as implemented in SYBYL 7.2 package. Formal charges and the particle concept options were always checked. In each case, a maximum of 100 poses were saved for each docked compound, although typically many fewer poses (30) were saved because biased sampling was used in the docking process.

PLS analysis. The regression analysis of CoMSIA field energies was performed using the partial least squares (PLS) algorithm with the leave-one-out (LOO) method adopted for cross validation. The optimum number of components to be used in conventional analyses was chosen from (i) the analysis with the highest cross validated r^2 value, and (ii) the model with the smallest standard error of prediction for component models with identical r^2 values. Final analysis was carried out to calculate the conventional r^2 (q^2) value using the optimum number of components. This analysis was used for the final non-cross-validated run.

References

1. Summa, V.; Petrocchi, A.; Pace, P.; Matassa, V. G.; De Francesco, R.; Altamura, S.; Tomei, L.; Koch, W.; Neuner, P. *J. Med. Chem.* **2004**, *47*, 14.
2. (a) Di Santo, R.; Fermeglia, M.; Ferrone, M.; Silvia Paneni, M.; Costi, R.; Artico, M.; Roux, A.; Gabriele, M.; Tardif, K. D.; Siddiqui, A.; Pricl, S. *J. Med. Chem.* **2005**, *48*, 6304. (b) Altamura, S.; Tomei, L.; Koch, U.; Neuner, P. J.; Summa, V. PCT Int. Appl. WO 200006529, 2000; CAN 132: 132323, 2000.
3. (a) Clark, M.; Cramer III, R. D.; Jones, D. M.; Patterson, D. E.; Simeroth, P. E. *Tetrahedron Comput. Methodol.* **1990**, *3*, 47. (b) San Juan, A. A.; Cho, S. J. *Bull. Korean Chem. Soc.* **2005**, *26*, 952.
4. (a) Klebe, G.; Abraham, U.; Mietzner, T. *J. Med. Chem.* **1994**, *37*, 4130. (b) Kim, J.; Lee, M.; Kang, S.-Y.; Park, J.; Lim, Y.; Koh, D.; Park, K. H.; Chong, Y. *Bull. Kor. Chem. Soc.* **2006**, *27*, 1025.
5. Bressanelli, S.; Tomei, L.; Rey, F. A.; De Francesco, R. *J. Virol.* **2002**, *76*, 3482.
6. Sechi, M.; Bacchi, A.; Carcelli, M.; Compari, C.; Duce, E.; Fisicaro, E.; Rogolino, D.; Gates, P.; Derudas, M.; Al-Mawsawi, L. Q.; Neamati, N. *J. Med. Chem.* **2006**, *49*, 4248.
7. Grobler, J. A.; Stillmock, K.; Hu, B.; Witmer, M.; Felock, P.; Espeseth, A. S.; Wolfe, A.; Egbertson, M.; Bourgeois, M.; Melamed, J. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 6661.
8. Pais, G. C.; Zhang, X.; Marchand, C.; Neamati, N.; Cowansage, K.; Svarovskaia, E. S.; Pathak, V. K.; Tang, Y.; Nicklaus, M.; Pommier, Y. *J. Med. Chem.* **2002**, *45*, 3184.
9. Hindle, S. A.; Rarey, M.; Buning, C.; Lengaue, T. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 129.
10. *SYBYL Molecular Modeling Software*; Tripos Inc.: St. Louis, MO 2002.
11. Mohammadi, M.; McMahon, G.; Sun, L.; Tang, C.; Hirth, P.; Yeh, B. K.; Hubbard, S. R.; Schlessinger, J. *Science* **1997**, *276*, 955.
12. Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219.