Hologram and Receptor-Guided 3D QSAR Analysis of Anilinobipyridine JNK3 Inhibitors

Jae Yoon Chung,[†,‡] Art E Cho,[‡,*] and Jung-Mi Hah[†,*]

[†]*Life Sciences Research Division, Korea Institute of Science and Technology, P. O. Box 131, Cheongryang, Seoul 130-650, Korea. *E-mail: jhah@kist.re.kr*

[‡]*Department of Biotechnology and Bioinformatics, Korea University, Jochiwon-Eup, Yeongi-Gun, Chungnam 339-700, Korea. *E-mail: artcho@korea.ac.kr*

Received August 14, 2009, Accepted September 15, 2009

Hologram and three dimensional quantitative structure activity relationship (3D QSAR) studies for a series of anilinobipyridine JNK3 inhibitors were performed using various alignment-based comparative molecular field analysis (COMFA) and comparative molecular similarity indices analysis (CoMSIA). The *in vitro* JNK3 inhibitory activity exhibited a strong correlation with steric and electrostatic factors of the molecules. Using four different types of alignments, the best model was selected based on the statistical significance of CoMFA ($q^2 = 0.728$, $r^2 = 0.865$), CoMSIA ($q^2 = 0.706$, $r^2 = 0.960$) and Hologram QSAR (HQSAR: $q^2 = 0.838$, $r^2 = 0.935$). The graphical analysis of produced CoMFA and CoMSIA contour maps in the active site indicated that steric and electrostatic interactions with key residues are crucial for potency and selectivity of JNK3 inhibitors. The HQSAR analysis showed a similar qualitative conclusion. We believe these findings could be utilized for further development of more potent and selective JNK3 inhibitors.

Key Words: JNK3, QSAR, CoMFA, CoMSIA, HQSAR

Introduction

c-Jun N-terminal Kinase (JNK) is a member of the mitogen-activated protein kinase (MAPK) family and functionally are a family of serine/threonine protein kinases.[1] Activation of JNKs is associated with complex receptor-mediated signaling mechanisms, which are induced by various extracellular stimuli such as cytokines, mitogens, osmotic stress and UV irradiation.[2] The activated JNKs then phosphorylate a number of substrate proteins such as c-Jun, ATF-2, APP, Elk-1, Tau, NFAT-4 for activation.[2]

At least ten different JNK isoforms derived from three genes (*JNK1, JNK2, and JNK3*) have been identified in mammalian cells.[3] While JNKs 1 and 2 are widely expressed in variety of tissues, JNK3 exhibit predominant localization in the brain and at lower levels in the heart and testis.[3,4] In the human CNS, the major JNK isoforms expressed are JNK3α1 and JNK1α1.[5] There is a growing number of evidences implicating JNK3, the neuronal specific isoform, is involved in various apoptotic neuronal cell death. It appears to play important roles in the brain to mediate neurodegeneration, such as Ab processing, Tau phosphorylation, and neuronal apoptosis in Alzheimer's Disease, as well as the mediation of neurotoxicity in a rodent model of Parkinson's Disease.[2] Therefore, identifying JNK3 selective inhibitor may contribute towards neuroprotection therapies with reduced side effect risks, and will aid in further understanding of the roles of individual JNK kinases.

In the last few years, various JNK3 inhibitors have been investigated. Highly potent and specific inhibitors, 2`-anilino-4,4`-bipyridine derivatives, of c-Jun N-terminal kinase-3 activity were reported by Swahn *et al.* in 2006.[6] The scaffold of 2`-anilino-4,4`-bipyridine derivatives and their inhibitory potency toward JNK3 are shown in Table 1. Using one of the derivatives, Swahn *et al.* solved one crystal structure of JNK3-inhibitor complex (PDB code 2EXC). This complex structure offers a lot of important structural information in the active site and this can be used to understand the essential structural features for JNK3 inhibition and more importantly, to guide the design of better inhibitors.

Figure 1 shows the inhibitor, **14** (Table 1) within the active site of the complex structure (2EXC) and the analyzed structural features are as follows: 1) the H-bond interaction between amino pyridine of **14** and the backbone of Met149 residue in the hinge region (Met146 ~ Asp150), which is common for most potent protein kinase inhibitors;[7,8] 2) the hydrophobic interaction between R_1 and Met146 residue (gatekeeper); 3)

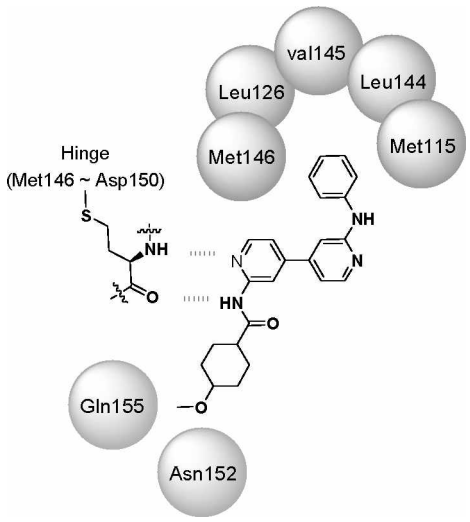


Figure 1. 2D view of interaction JNK3 active site and compound 14 (PDB: 2EXC).

the H-bond interaction between R_2 and Glu155. In particular, the Met146 and Glu155 residues are unique to JNK3 compared with other similar kinases such as p38 and ERK2 (Thr106 and Asn115 in p38 and Gln105 in ERK2 are unique to their respective proteins) and they seem to contribute to JNK3's selectivity over other kinases.[9]

The three-dimensional quantitative structure activity relationship (3D-QSAR) methods used here are comparative molecular field analysis (CoMFA), and comparative molecular similarity indices analysis (CoMSIA).[10-13] Following two 3D-QSAR analyses, hologram QSAR (HQSAR) method was also performed.[14] While HQSAR does not require molecular alignment, CoMFA and CoMSIA methods are significantly influenced by the alignment of input compounds. In order to develop reasonable models from CoMFA and CoMSIA, four different alignments were applied. They are *ligand-based alignment* using quantum mechanics (QM), *pharmacophore-based alignment*, and two *receptor-guided alignments* (docking-based and MC searching & EM based).

Materials and Computational Methods

Inhibitor data set. In the set of 24 2`-anilino-4,4`-bipyridine derivatives, compounds without clear IC_{50} value were omitted to derive QSAR models. The pIC_{50} values were used as the dependent variable in QSAR models. The total of 24 compounds was divided into the training set and the test set. 20 compounds were selected as the training set to generate QSAR models. The selection of the training and test sets was done manually such that low, moderate, and high JNK inhibitory activities were all represented. The structures and their activity values are listed in Table 1.

Preparation of ligands and receptor. The complex structure of **14** and JNK3 protein was obtained from the protein data bank (PDB code 2EXC) and prepared for receptor-guided alignment using Protein Preparation Wizard of Schrödinger maestro program.[15] All water molecules were removed from the structure and **14** was selected as a template because **14** can supply reasonable bioactive conformation of the scaffold. The 3D structures of other compounds were constructed based on the template using Schrödinger LigPrep program with the force field, OPLS 2005.[16]

CoMFA, CoMSIA, and HQSAR. In deriving CoMFA descriptor, all compounds should be aligned in a certain 3D cubic lattice with grid spacing of 1.0, 1.5, and 2.0 Å. The steric (Lennard-Jones 6-12 potential) and electrostatic (Coulombic potential) field energies were calculated by CoMFA descriptors using default options. CoMSIA also works with aligned compounds in the same lattice model of CoMFA. In CoMSIA, similarity indices for compound's atoms in each grid point were calculated and physicochemical properties (steric, electrostatic, hydrophobic, hydrogen bond donor and acceptor) were evaluated. HQSAR methodology requires two-dimensional structures of input compounds and their activity values in order to derive compound's holograms and HQSAR model.[14] HQSAR model is derived from hologram arrays of input compounds and their biological activity values by partial least square (PLS). The best HQSAR model can be produced with optimal fragment size and hologram length from the stored PLS processes.

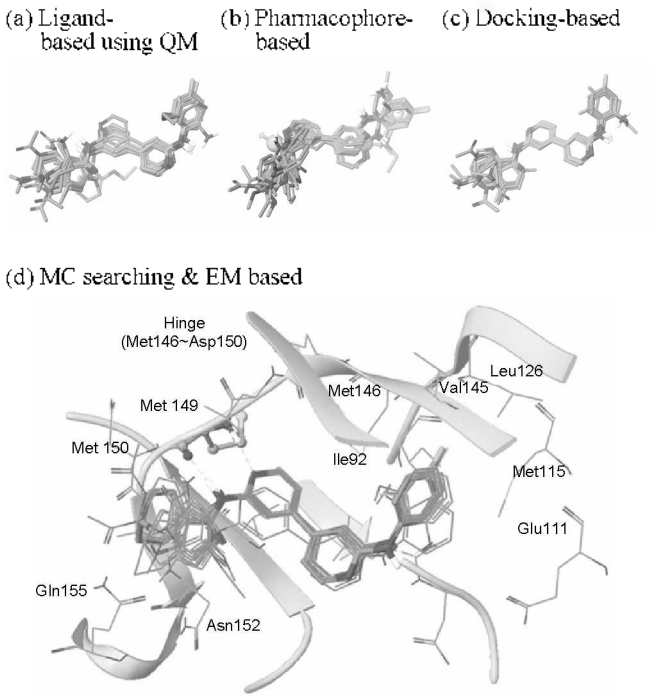


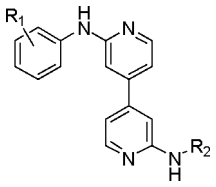
Figure 2. Different alignment results.

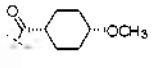
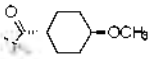
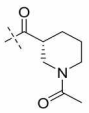
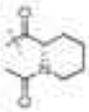
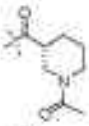
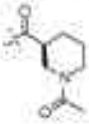
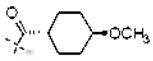
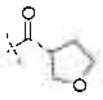
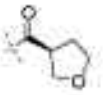
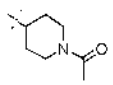
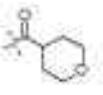
PLS. Partial least squares (PLS) methodology for cyclic cross-validation with leave-one-out (LOO) method is used to derive QSAR models.[17-18] The descriptors of CoMFA, CoMSIA, and HQSAR were used as independent variables, and pIC_{50} activity values used as dependent variables. The number of independent variables by those QSAR descriptors is usually much larger than that of dependent variables which are compounds with pIC_{50} in data set. PLS can solve those equations and generate QSAR models. Cross-validation procedure evaluates the quality of generated QSAR model. The cross-validation analysis is performed using the leave-one-out (LOO) method.

Ligand-based alignment. The structures of all compounds except **14** were fully optimized based on density functional theory (DFT B3YLP/6-31G**) calculation. The thermal correction with QM methods can be incorporated into ground state energy by vibration frequency analysis. We used Schrödinger Jaguar program to carry out QM calculations in gas phase and thermal correction at 310 K.[19] Then, the optimized compounds were aligned over the scaffold of **14**. Figure 2(a) shows the ligand-based alignment using QM.

Pharmacophore-based alignment. Schrödinger Phase program was used to develop common pharmacophore hypothesis and to align all compounds based on it.[20] Using ConfGen in Schrödinger package, 619 of various conformers for all compounds except **14** were generated with distance dependant dielectric solvation treatment and OPLS-2005 force field.[21] Those pharmacophore features are hydrogen bond acceptor (**A**), hydrogen bond donor (**D**), hydrophobic group (**H**), negatively charged group (**N**), positively charged group (**P**), and aromatic ring (**R**). They were mapped for each compound using a set of the pharmacophore features. Compounds such that

Table 1. The streutres and observed JNK3 inhibitory activities



No	R1	R2	IC_{50}(M)	pIC50	No	R1	R2	IC_{50}(M)	pIC50
1	-	Bn	0.030×10^{-6}	7.523	13	-		0.018×10^{-6}	7.745
2	4-F	4-F	0.017×10^{-6}	7.770	14	4-F		0.015×10^{-6}	7.824
3	2-F	2-F	0.032×10^{-6}	7.495	15 *	4-F		0.235×10^{-6}	6.629
4	$2\text{-}NH_3$	$2\text{-}NH_3$	0.108×10^{-6}	6.967	16	-		0.009×10^{-6}	8.046
5 *	$2\text{-}CH_3$	$2\text{-}CH_3$	0.652×10^{-6}	6.186	17	-		0.207×10^{-6}	6.684
6	$2\text{-}OCH_3$	$2\text{-}OCH_3$	0.528×10^{-6}	6.277	18	4-F		0.005×10^{-6}	8.301
7	$2\text{-}OCH_2CH_3$	$2\text{-}OCH_2CH_3$	0.693×10^{-6}	6.159	19	4-F		0.010×10^{-6}	8.000
8	-		0.032×10^{-6}	7.495	20	-		0.007×10^{-6}	8.155
9 *	-		0.020×10^{-6}	7.699	21 *	4-F		0.003×10^{-6}	8.523
10	-		0.033×10^{-6}	7.481	22	4-F		0.009×10^{-6}	8.046
11 *	-		0.044×10^{-6}	7.357	23	-		0.008×10^{-6}	8.097
12	-		0.175×10^{-6}	6.757	24	4-F		0.006×10^{-6}	8.222

*Test set compounds

pIC_{50} values are more than 8 were selected as active compounds except test set for producing the pharmacophore hypotheses. **ADHRR** pharmacophore features were selected as the common pharmacophore hypothesis (CPH). The PLS analyses carried out using Phase produced pharmacophore-based models using grid spacing of 1 Å with training set. The best statistics pharmacophore model was selected for the alignment of all compounds except **14**. Figure 2(b) shows the pharmacophore-based alignment.

Molecular docking and docking-based alignment. Glide docking program in Schrödinger package was used to predict reasonable binding poses of input compounds within the active site.[22] The default Glide parameters were used except the constraint option of H-bond interaction with Met149 (hinge). Docking performance was tested through redocking of the **14** separated from the complex structure (2EXC) into the prepared JNK3 protein. The redocked binding poses of **14** were quite similar to its original geometry. RMSD values (heavy atoms) were between 0.376 ~ 1.694 Å (average RMSD: 1.046 Å, median RMSD: 1.087 Å). The grid spacing of CoMFA was assigned between 1.0 ~ 2.0 Å, therefore an average RMSD value close to 1.0 Å should be small enough for CoMFA analysis. Because the redocking test was successful, docking poses of other compounds could be accepted as plausible bioactive conformations in the docking-based alignment. Top-ranked Glide docking pose for each compound was chosen as the result from docking-based alignment. Figure 2 (c) shows the docking-based alignment.

Monte Carlo (MC) searching & Energy Minimization (EM) based alignment. Receptor-guided alignment using docking method is very effective for a large data set which has more than 100 compounds because docking programs can automatically select reasonable bioactive conformations among a huge number of possible binding poses. In a small data set, it is possible to generate more bioactive conformations manually for each compound using modeling tools than using automated docking programs. Especially when the structural similarity between the co-crystallized compound and other compounds is quite high, the scaffold which input compounds share can be replaced by that of co-crystallized compound. MC searching & EM based alignment modifies different substituents for each compound within the active site. First, the input compounds are superimposed based on the conformations of the common scaffold within the active site. After positioning input compounds in the active site, MC searching can determine possible conformations of substituents for each compound within the active site. MC searching method can find possible conformations of compounds which avoid steric hindrance in the active site. The conformations of input compounds found by MC searching method are slightly modified through energy minimization to find input compound's conformation with more stable binding energy. MacroModel program in Schrödinger package has MC searching and energy minimization methods.[23] Top-ranked binding pose by MacroModel was chosen for each compound.

According to the MC searching & EM based alignment procedure, several plausible conformations of input compounds were found using Macromodel application based on the conformation of common scaffold in the **14**. For each compound, conformation having most stable binding energy within the active site was selected. Figure 2 (d) shows the MC searching & EM based alignment within the active site. As one would

Table 2. Statistics of generated CoMFA models using different alignment methods

Model	Grid spacing (Å)	Leave-one-out cross-validation			Non-cross-validation			Bootstrap		Predictive R^2	Field contribution	
		q^2	n	SDEP	r^2	SEE	F-value	r^2_{boot}	StdDev	r^2_{pred}	Steric	Electro
Alignment 1 (ligand-based alignment)												
1	1.0	0.410	2	0.579	0.706	0.409	22.788	0.777	0.099	-	0.639	0.361
2	1.5	0.458	4	0.587	0.930	0.211	56.628	0.965	0.029	0.840	0.650	0.350
3	2.0	0.358	2	0.604	0.717	0.402	24.025	0.814	0.077	-	0.722	0.278
Alignment 2 (pharmacophore-based alignment)												
4	1.0	0.385	2	0.587	0.785	0.347	31.015	0.847	0.034		0.572	0.428
5	1.5	0.387	3	0.604	0.836	0.312	27.176	0.899	0.052		0.555	0.445
6	2.0	0.323	2	0.616	0.798	0.336	33.630	0.807	0.090		0.623	0.377
Alignment 2 (docking-based alignment)												
7	1.0	0.342	2	0.598	0.726	0.386	27.804	0.783	0.071	-	0.631	0.369
8	1.5	0.364	2	0.588	0.723	0.388	27.368	0.821	0.062	0.475	0.650	0.350
9	2.0	0.231	2	0.646	0.740	0.376	29.839	0.823	0.079	-	0.663	0.337
Alignment 3 (MC & EM based alignment)												
10	1.0	0.592	2	0.458	0.861	0.268	52.644	0.898	0.028	-	0.614	0.386
11	1.5	0.631	2	0.436	0.862	0.267	52.994	0.870	0.054	0.771	0.618	0.382
12	2.0	0.600	2	0.454	0.878	0.251	60.982	0.880	0.051	-	0.689	0.311

Table 3. Statistics of the resulting CoMFA models

Model	Resion-Focusing	Leave-one-out cross-validation			Non-cross-validation			Bootstrap		Predictive R^2	Field contribution	
		q^2	n	SDEP	r^2	SEE	F-value	r^2_{boot}	StdDev	r^2_{pred}	Steric	Electro.
Receptor-based CoMFA model												
11	1.5	0.631	2	0.436	0.862	0.267	52.994	0.870	0.054	0.771	0.618	0.382
CoMFA model, region focusing method												
13	PLS	0.728	2	0.374	0.865	0.263	54.627	0.89	0.066	0.647	0.631	0.369

Table 4. The regression summary of CoMSIA models

No.	Field					LOO cross-validation		Non-cross-validation			r^2_{pred}
	S	E	H	D	A	q^2	n	r^2	SEE	F	
1	1.000	-	-	-	-	0.604	2	0.803	0.318	37.749	-
2	-	1.000	-	-	-	0.467	2	0.742	0.364	24.480	-
3	-	-	1.000	-	-	0.565	4	0.960	0.152	90.741	-
4	-	-	-	1.000	-	0.495	4	0.639	0.459	6.648	-
5	-	-	-	-	1.000	0.452	1	0.621	0.429	29.471	-
6	0.278	0.722	-	-	-	0.548	2	0.827	0.298	40.631	-
7	0.254	-	0.746	-	-	0.630	9	0.996	0.058	291.298	-
8	0.516	-	-	0.484	-	0.602	4	0.856	0.290	22.308	-
9	0.366	-	-	-	0.634	0.688	5	0.947	0.182	49.807	-
10	-	0.537	0.463	-	-	0.490	3	0.913	0.218	56.032	-
11	-	0.619	-	0.381	-	0.674	4	0.912	0.226	39.103	-
12	-	0.550	-	0.450		0.660	5	0.975	0.125	108.391	-
13	-	-	0.601	0.399	-	0.654	4	0.924	0.210	45.754	-
14	-	-	0.480	-	0.520	0.780	13	0.999	0.045	330.595	-
15	-	-	-	0.667	0.333	0.479	1	0.616	0.432	28.828	-
16	0.165	0.438	0.397	-	-	0.529	4	0.953	0.165	76.231	-
17	0.169	0.474	-	0.358	-	0.691	4	0.926	0.208	46.981	-
18	0.135	0.445	-	-	0.420	0.695	6	0.987	0.094	163.490	-
19	0.169	-	0.446	0.385	-	0.681	4	0.929	0.203	49.435	-
20	0.121	-	0.397	-	0.482	0.771	12	0.998	0.049	297.599	-
21	0.320	-	-	0.239	0.441	0.673	6	0.959	0.166	50.661	-
22	-	-	0.437	0.207	0.356	0.748	12	0.998	0.049	298.242	-
23	-	0.481	-	0.221	0.298	0.640	4	0.945	0.180	63.931	-
24	-	0.314	0.274	-	0.412	0.656	8	0.995	0.060	301.972	-
25	-	0.376	0.316	0.308	-	0.644	4	0.949	0.172	70.144	-
26	-	0.310	0.249	0.199	0.242	0.653	4	0.961	0.150	92.922	-
27	0.111	-	0.363	0.214	0.313	0.754	12	0.998	0.049	301.651	-
28	0.129	0.391	-	0.214	0.266	0.706	4	0.960	0.152	90.897	0.777
29	0.092	0.295	0.240	-	0.373	0.685	9	0.997	0.055	326.175	-
30	0.102	0.327	0.268	0.304	-	0.680	4	0.953	0.165	76.471	-
31	0.085	0.249	0.221	0.173	0.272	0.685	9	0.997	0.055	317.603	-

S, steric; E, electrostatic; H, hydrophobic; D, hydrogen bond donor; A, hydrogen bond acceptor; n, number of statistical components; q^2, the LOO cross-validated correlation coefficient; r^2, the correlation coefficient; SEE, standard estimated error; F, the Fisher value; $r^2_{predictive}$, the correlation coeffient for test set.

Table 5. The Observed *vs.* predicted activities (pIC50) with deviations by

No.		pIC50	CoMFA		RF-CoMFA		CoMSIA		HQSAR	
			prediction	Δ2	prediction	Δ2	prediction	Δ2	prediction	Δ2
1		7.523	7.256	0.071	7.326	0.039	7.277	0.061	7.353	0.029
2		7.770	7.447	0.104	7.481	0.084	7.615	0.024	7.848	0.006
3		7.495	7.373	0.015	7.430	0.004	7.463	0.001	7.53	0.001
4		6.967	7.127	0.026	7.113	0.021	7.169	0.041	6.792	0.031
5	*	6.186	7.078	0.796	7.118	0.869	7.087	0.812	7.059	0.762
6		6.277	6.298	0.000	6.206	0.005	6.302	0.001	6.307	0.001
7		6.159	5.881	0.077	5.912	0.061	6.066	0.009	6.128	0.001
8		7.495	7.980	0.235	7.996	0.251	7.518	0.001	7.654	0.025
9	*	7.699	7.981	0.080	8.038	0.115	8.098	0.159	8.201	0.252
10		7.481	7.610	0.017	7.597	0.013	7.405	0.006	7.342	0.019
11	*	7.357	7.920	0.317	7.964	0.368	7.710	0.125	7.663	0.094
12		6.757	7.317	0.314	7.274	0.267	7.087	0.109	7.251	0.244
13		7.745	7.898	0.023	7.920	0.031	7.716	0.001	7.744	0.000
14		7.824	7.883	0.003	7.870	0.002	7.906	0.007	7.902	0.006
15	*	6.629	7.872	1.545	7.853	1.498	7.798	1.367	7.902	1.621
16		8.046	7.943	0.011	7.872	0.030	8.043	0.000	7.885	0.026
17		6.684	7.030	0.120	7.064	0.144	6.701	0.000	6.669	0.000
18		8.301	8.159	0.020	8.073	0.052	8.365	0.004	8.132	0.029
19		8.000	7.931	0.005	7.918	0.007	8.036	0.001	8.132	0.017
20		8.155	8.108	0.002	8.142	0.000	8.042	0.013	8.03	0.016
21		8.523	8.215	0.095	8.225	0.089	8.521	0.000	8.278	0.060
22		8.046	8.139	0.009	8.168	0.015	8.249	0.041	8.278	0.054
23		8.097	7.962	0.018	7.976	0.015	7.952	0.021	8.033	0.004
24		8.222	8.011	0.045	8.003	0.048	8.135	0.008	8.28	0.003
SD			0.180		0.184		0.130		0.151	

[a]SD: standard deviation.

Table 6. HQSAR analyses for various fragment distinctions using default fragment size (4-7)

No.	Distinction	q^2	r^2	Ave	SEE	LV	Length
1	A	0.807	0.912	0.803	0.315	2	199
2	A/B	0.718	0.936	0.701	0.393	3	353
3	A/B/Co	0.730	0.962	0.659	0.397	4	97
4	A/B/Co/H	0.648	0.882	0.620	0.439	3	353
5	A/B/Co/Cr	0.753	0.969	0.699	0.380	4	97
6	A/B/Cr	0.732	0.928	0.698	0.383	3	151
7	A/Co	0.667	0.867	0.659	0.414	2	307
8	Co	0.687	0.837	0.656	0.402	2	257
9	A/Co/DA	0.775	0.965	0.705	0.362	4	199
10	A/B/Co/DA	0.741	0.877	0.719	0.365	2	151

[a]AVE: Ensemble q^2.

expect, most compounds have H-bond interaction with M149 (hinge). R_1 and R_2 substitutions interact with M146 and Glu155 which are important residues for JNK3 selectivity as mentioned before.

Result and Discussion

CoMFA and CoMSIA analyses. Based on four different alignments, CoMFA models were derived with different CoMFA grid spacings by PLS analyses. The statistics summaries for different CoMFA models are listed in Table 2. The best CoMFA model 11 was obtained from MC Searching & EM based alignment and grid spacing 1.5 Å (q^2 = 0.631, r^2 = 0.862, F = 52.994, SEE = 0.267). However, other alignment methods could give lower statistics. With the CoMFA model **11**, region focusing (RF) strategy was applied by reducing the number of independent variables in CoMFA calculation in order to get more predictable CoMFA model. RF-CoMFA model **13** with better statics was produced (q^2 = 0.728, r^2 = 0.865, F = 54.627, SEE = 0.263). Table 3 shows the statistics summary for RF-CoMFA model **13**. To verify the statistical confidence and robustness of CoMFA models, 100 bootstrapping samplings were performed. A bootstrapped r^2_{boot} of 0.890 and a standard deviation (StdDev) of 0.666 from RF-CoMFA model **13** were obtained (Table 3). These statistical values (RF-CoMFA **13**) indicate a good internal consistency in the training data set.

The aligned compounds used for this CoMFA model **13**

Table 7. HQSAR analysis for the influence of various fragment sizes using the best fragment distinction (A)

No.	Size	q^2	r^2	Ave	SEE	LV	Length	r^2_{pred}
1	1-2	0.269	0.582	0.269	0.653	4	199	-
2	1-3	0.424	0.731	0.444	0.561	3	97	-
3	1-7	0.814	0.910	0.792	0.310	2	199	-
4	2-7	0.810	0.910	0.794	0.313	2	199	-
5	3-7	0.809	0.911	0.795	0.313	2	199	-
6	3-8	0.836	0.948	0.812	0.299	3	307	-
7	3-9	0.838	0.950	0.805	0.298	3	151	0.700
8	4-7	0.807	0.912	0.803	0.315	2	199	-
9	4-8	0.833	0.933	0.812	0.293	2	151	-
10	5-7	0.805	0.923	0.803	0.317	2	353	-
11	5-8	0.833	0.930	0.814	0.294	2	151	-
12	5-10	0.830	0.947	0.788	0.305	3	151	-
13	6-8	0.838	0.935	0.824	0.289	2	151	-
14	7-9	0.825	0.944	0.807	0.300	2	307	-
15	7-10	0.834	0.951	0.783	0.301	3	151	-

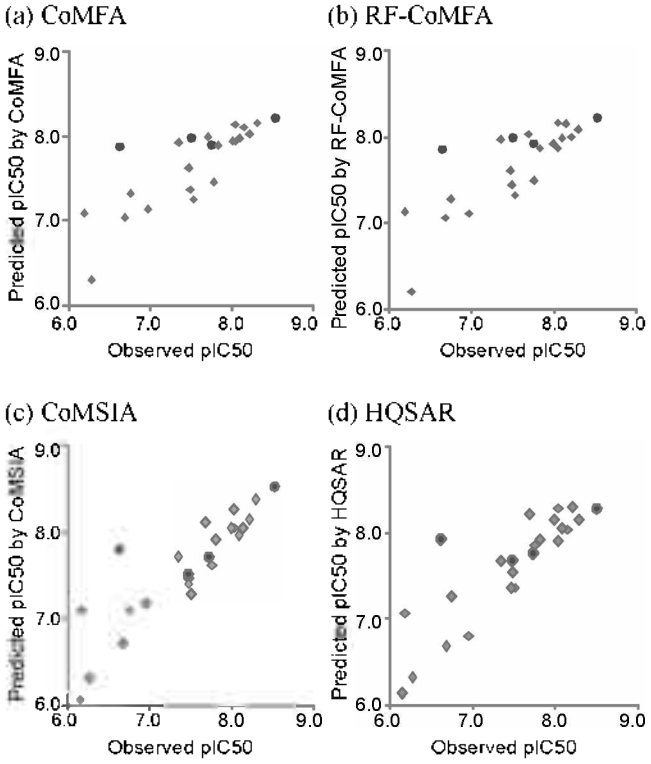


Figure 3. Plots of predicted *vs.* observed values of (a) CoMFA, (b) RF-CoMFA, (c) CoMSIA, and (d) HQSAR.

were applied to derive CoMSIA model. In CoMSIA studies, allpossible cases using five fields were calculated, and table 4 shows the statistics summaries of CoMSIA. The important values to determine the quality of QSAR models are q^2 and statistical component. When considering the number of input compounds, the number of statistical components below six is preferred. Among CoMSIA models which satisfy these criteria, the CoMSIA model **28** that has the highest q^2 value was chosen. The CoMSIA model **28** was produced by using four field descriptors (steric, electrostatic, H-bond donor, and acceptor). The combination of four fields (steric, electrostatic, H-bond acceptor/donor) gave the best statistics result (q^2 = 0.706, r^2 = 0.960, F = 90.897, SEE = 0.152).

The statistics summary by CoMSIA models is listed in Table 4. The observed activities and the predicted activities by the CoMFA, RF-CoMFA, and CoMSIA models are listed in Table 5. Table 5 shows that activities predicted by those produced models are in good agreement with the observed data, suggesting that those three models should have satisfactory predictive abilities. Figures 3 (a) ~ (c) show plots of the observed data and the predicted data by the CoMFA, RF-CoMFA and CoMSIA models.

HQSAR analysis. HQSAR model is influenced by various fragment distinction factors (atom types (A), bond types (B), atom hybridization or connectivity (Co), chirality (Cr), hydrogen bond and donor(H)) and fragment sizes. These variations were examined to produce highly predictable HQSAR model. To find the best fragment distinction factors in the initial HQSAR stage various combinations of the above fragment distinction factors were employed with the default fragment size 4 ~ 7 (Table 6). The HQSAR model which was produced from fragment distinctions of A was selected based on PLS analysis that gives the best q^2. Next, several analyses with various fragment sizeswere investigated to find a better HQSAR model using the fragment distinction factor (A). Table 7 shows the evaluated fragment sizes and their statistical results. The HQSAR model using fragment size 3 ~ 9 shows the best PLS analysis (q^2 = 0.838, r^2 = 0.950, SEE = 0.298, LV = 3, Length = 151). The predictive activities of all compounds were shown in Table 5. Figure 3 (d) shows the plot of the observed versus the predicted activities by the HQSAR model.

Validation of 3D QSAR and HQSAR models. Table 5 shows predicted pIC_{50} values from different QSAR models (CoMFA, RF-CoMFA, CoMSIA, and HQSAR). To validate the predictability and accuracy of the models, predictive correlation coefficient r^2_{pred} was investigated for all compounds including training

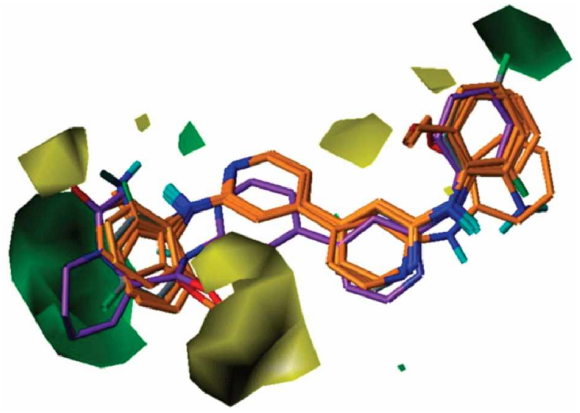


Figure 4. CoMFA steric contour maps with compound **2** (gray), **3** ~ **7** (orange), and **17** (purple). Green contours indicate regions where bulky groups increase activity, whereas yellow contours indicate regions where bulky groups decrease activity.

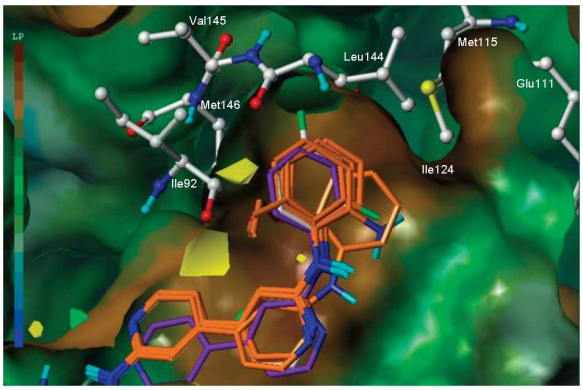


Figure 5. The hydrophobic surface map of JNK3's active site with CoMFA steric contour map. Dark brown color indicates the highest hydrophobicity, while blue color signifies the least. hydrophobic maps of receptor surface were produced by MOLCAD program in SYBYL.

and test set. The predicted r^2_{pred} values for QSAR models are shown in Table 2 - 4, and 7. The final r^2_{pred} values are 0.771 for CoMFA, 0.647 for RF-CoMFA, 0.777 for CoMSIA, and 0.700 for HQSAR. These high scores of r^2_{pred} imply that our 3D QSAR and HQSAR models are highly predictable. Figure 3 shows correlation between predicted and observed activity values by these models.

RF-CoMFA contour map. Figure 4 shows 3D contour map of RF-CoMFA steric field interaction with compounds **2** ~ 7 and **14**. This RF-CoMFA steric contour map can explain key factors affecting compound's activity. Firstly, the 4-fluorine atom as R_1 is better in increasing biological activity than other atoms on the 2-substitutions in aniline moiety. RF-CoMFA steric contour map has sterically favored region (green) around 4-position in aniline and unfavored regions (yellow) around 2-position. Compounds **3** ~ 7 have various substituents on 2-position, and they are located close to yellow contour maps. On the other hand, compound **2** has fluorine atom on 4-position in aniline, and this is located close to green contour map. This can be the reason why compound **2** has higher biological activity value than compounds **3** ~ 7. In the same way, compounds **18**, **21**, and **24**

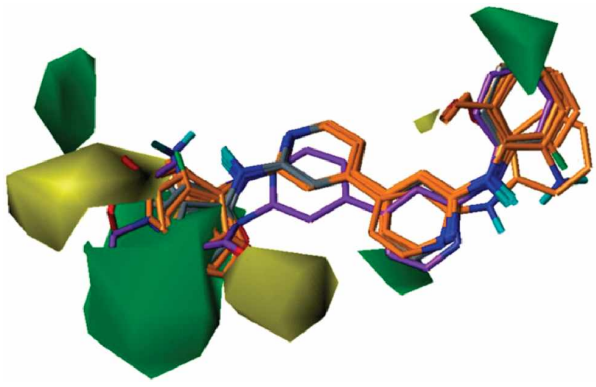


Figure 6. CoMSIA steric contour map with compounds **2**, **20** ~ **22** (gray) and compound **3** ~ 7 (orange), and **17** (purple). Green contours indicate regions where bulky groups increase activity; yellow contours indicate regions where bulky groups decrease activity.

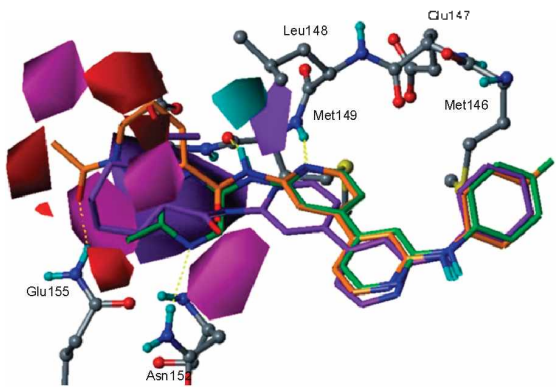


Figure 7. CoMSIA H-bond donor contour map with compound **10** (green), **17** (purple), and **19** (orange) within the active site. Yellow dot lines show H-bond interaction.

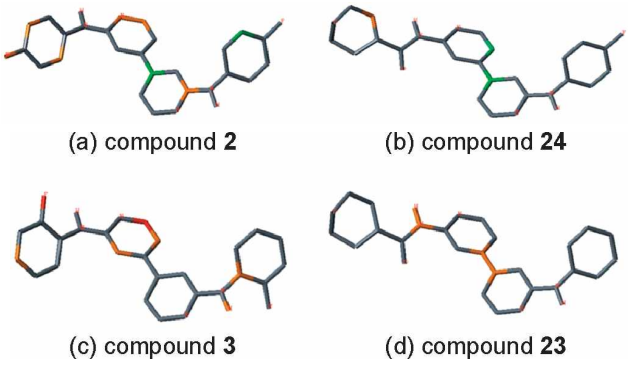


Figure 8. HQSAR atomic contribution maps. Green and yellow colors imply highest to higher atomic contributions to the activity while red and orange colors denote lower to lowest atomic contributions to the activity. The atoms with intermediate contributions are colored white.

with fluorine atom on 4-position have higher biological activities than compounds **16**, **20**, and **23** with no R_1 substitution. The hydrophobic map of the active site surface was produced by MOLCAD program in SYBYL and this hydrophobic map also supports this fact.[24] In Figure 5 it is suggested that the two residues, Ile92 and Met146, whichare located close to R_1, could

supply a steric hindrance with 2-substitution for R_1. However, there is an empty hydrophobic region surrounded by Leu144, Val145, and Met146 residues (Figure 5) over the fluorine atom of R_1. The compounds with 2-fluorine atom of R_1 can have hydrophobic interaction with those residues (Leu144 ~ Met146). Especially, the gatekeeper Met146 which is important for selectivity, is located close to 2-substitution of R_1. The hydrophobic 4-substitutions on aniline could have a substantial interaction with hydrophobic residues (Leu144 ~ Met146) without steric clash with Met146 at the same time. Consequently, the RF-CoMFA steric contour map and the hydrophobic surface map suggest that hydrophobic substitutions on 4-position as R_1 would be best for the activity.

Secondly, RF-CoMFA steric contour map shows also sterically favored and unfavored regions around R_2 (Figure 4). This could explain the low potency of compound **17** (purple in Figure 4) compared to compounds **16** ~ **24** although they have very similar molecular structures. Most of R_2 groups of **16** ~ **24** are located in the major green region, while R_2 of **17** is located in the yellow regions. This fact was also supported in the structure of JNK3's active site (Figure 2 (d)). In Figure 2 (d), the Asp150 residue is located at similar positions of yellow region (Figure 4) and the Asp150 belongs to hinge region which is quite important for H-bond interaction. Due to the steric hindrance between R_2 substitution and Asp150, compound **17** moved away from the hinge region and lost H-bond interaction with Met149 (hinge). Poor biological activity of compound **17** can be explained by the fact that there is neither R_1 substitution nor H-bond interaction with Met149.

CoMSIA contour map. Overall, the steric contour map from the CoMSIA analysis in Figure 6 is quite similar to the RF-CoMFA steric map (Figure 4). The CoMSIA steric contour map also has sterically favored (green) and unfavored (yellow) regions similar to those of the RF-CoMFA steric contour map. In Figure 6, R_1 substitutions of compound **3** ~ **7** are located close to a yellow region, whereas the R_1 of **2** is positioned in green region. This could explain the better activity of **2** over compound **3** ~ **7**. Also, the R_2 substituent of **17** is positioned in another yellow region, while the substitutions of compound **20** ~ **22** are not positioned there. The CoMSIA steric contour map is well in agreement with the RF-CoMFA steric contour map. Both CoMFA and CoMSIA steric contour maps explain that 4-R_1 substitution is preferred to 2-R_1 substitution with a hydrophobic group for higher biological activity.

Figure 7 shows the CoMSIA H-bond donor and acceptor contour maps with compounds **10**, **17**, and **19** within the active site. The magenta contours regions are favored for H-bond acceptor, while the red contours regions are not. This CoMSIA H-bond acceptor contour map can explain why compound **17** has much lower biological activity than compounds **18** and **19**. The terminal carbonyl groups of compounds **18** and **19** are oriented toward magenta regions. However, the carbonyl group of **17** is oriented toward red volume. In Figure 7, the cyan color indicates regions where H-bond donor substituents enhance activity, whereas the purple color indicates regions where H-bond donor substituents reduce activity. Compounds **18** ~ **24** have NH group close to R_2 substitution which can work as H-bond donor, while NH group of **17** is not oriented toward cyan region. This can be added to the explanation of compound **17**'s poor biological activity with other contour maps of CoMFA and CoMSIA.

The two steric contour maps of RF-CoMFA and CoMSIA closely reflect the structural shape around R_1 region in the active site (Figure 5). The CoMSIA H-bond acceptor and donor contour maps also reflect H-bond properties around R_2 region in the active site (Figure 7). The CoMSIA H-bond donor and acceptor contour maps accord closely with H-bond interactions with Met149, Asn152, and Gln155. The cyan region favored for H-bond donor group indicates that the corresponding H-bond acceptor group should be near. In the same way, magenta region indicates that H-bond donor group of the active site should be around. In Figure 7, the cyan region matches the carbonyl group of Met149's backbone, and magenta regions match well with Gln155 and Asn152's NH groups. These CoMSIA H-bond contour maps describe H- bond properties of the active site quite satisfactorily.

As mentioned earlier, interactions with Met146 (gatekeeper), Met149 (hinge) and Glu155 are quite important for biological activity and selectivity. The unique gatekeeper residue, Met146 is involved in selectivity, and the inhibitors without interaction with Met146 could lose selectivity for JNK3. An H-bond interaction with Met149 should be necessary for JNK3 inhibitors like most other kinase inhibitors' H-bond interaction with hinge region. The Glu155 residue in JNK3 corresponds to Asn115 residue in p38. Since they are different in size and electrostatic property, interactions with Glu155 would help alot in gaining selectivity for JNK3. In these points of view, these CoMSIA steric and H-bond contour maps can give significant insight to highly active and selective inhibitors of JNK3.

HQSAR atomic contribution map. PLS analysis of HQSAR produces color-coded structure diagram representing the degree of each atom's contribution according to their biological activity. In general, the green and yellow colors in HQSAR diagrams mean the highest to higher atomic contributions to the activity, while orange and red colors denote lower to lowest atomic contributions to the activity. The atoms with intermediate contribution are colored white. Figure 8 shows atomic contributions of model compounds **2**, **3**, **23**, and **24** to their biological activities. In Figure 8, the models explain that the 4-fluorine atom of R_1 is important to biological activity: (a) and (b) show compounds **2** and **24** with 4-fluorine substitution has green atoms while (c) and (d) show compounds **3** and **23** do not. This HQSAR atomic contribution models are in good agreement with the steric contour maps of RF-CoMFA and CoMSIA.

Conclusion

In the present study, three QSAR analyses (CoMFA, CoMSIA, and HQSAR) were built with a seriesof JNK3 inhibitors. Various alignment methods (ligand-based, pharmacophore-based, and receptor-guided) were tried to derive reasonable CoMFA and CoMSIA models. Among those alignments, receptor-guided alignment using MC searching & EM produced the best alignment of input compounds. Based on this alignment, the RF-CoMFA and CoMSIA models were developed. The RF-CoMFA and CoMSIA models show critical steric and

H-bond interactions necessary for designing new JNK3 inhibitors. The CoMFA and CoMSIA steric contour maps all suggest that a hydrophobic substitution at 4-R_1 in aniline moiety would increase biological activity. This suggestion was also supported by HQSAR. The H-bond contour maps by CoMSIA can explain the differences of compounds activities. These steric and H-bond contour maps were investigated through comparison with the properties of important residues in the active site. These interactions found by CoMFA, CoMSIA, and HQSAR will be extremely useful for design of a new generation of JNK3 inhibitors.

Acknowledgments. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2009-0087992; J. M. H., J. Y. C.) and (2009-0072095; A. E. C., J. Y. C.).

Reference

1. Davis, R. J. *Cell* **2000**, *103*, 239.
2. Manning, A.; Davis, R. *Nature Reviews Drug Discovery* **2003**, *2*, 554.
3. Gupta, S.; Barrett, T.; Whitmarsh, A.; Cavanagh, J.; Sluss, H.; Derijard, B.; Davis, R. *The EMBO Journal* **1996**, *15*, 2760.
4. Martin, J.; Mohit, A.; Miller, C. *Molecular Brain Research* **1996**, *35*, 47.
5. Kuan, C.; Yang, D.; Roy, D.; Davis, R.; Rakic, P.; Flavell, R. *NEURON-CAMBRIDGE MA-* **1999**, *22*, 667.
6. Swahn, B.; Huerta, F.; Kallin, E.; Malmstroem, J.; Weigelt, T.; Viklund, J.; Womack, P.; Xue, Y.; Oehberg, L. *Bioorganic & Medicinal Chemistry Letters* **2005**, *15*, 5095.
7. Noble, M. E. M.; Endicott, J. A.; Johnson, L. N. *Science* **2004**, *303*, 1800.
8. Nolen, B.; Taylor, S.; Ghosh, G. *Molecular Cell* **2004**, *15*, 661.
9. Scapin, G.; Patel, S.; Lisnock, J.; Becker, J.; LoGrasso, P. *Chemistry & Biology* **2003**, *10*, 705.
10. Cramer III, R.; Patterson, D.; Bunce, J. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
11. Klebe, G.; Abraham, U.; Mietzner, T. *J. Med. Chem.* **1994**, *37*, 4130.
12. Klebe, G.; Abraham, U. *Journal of Computer-Aided Molecular Design* **1999**, *13*, 1.
13. Shemetulskis, N.; Weininger, D.; Blankley, C.; Yang, J.; Humblet, C. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 862.
14. Lowis, D. R. *Tripos Technical Notes* **1997**, *1*.
15. Maestro version 8.5, S., LLC, New York, NY **2008**.
16. LigPrep version 2.2, S., LLC, New York, NY **2005**.
17. Wold, S.; Sjöström, M.; Eriksson, L. *Chemometrics and Intelligent Laboratory Systems* **2001**, *58*, 109.
18. Kettaneh, N.; Berglund, A.; Wold, S. *Computational Statistics and Data Analysis* **2005**, *48*, 69.
19. Jaguar version 7.5, S., LLC, New York, NY **2008**.
20. Phase version 3.0, S., LLC, New York, NY **2008**.
21. ConfGen version 2.0, S., LLC, New York, NY **2008**.
22. Glide version 5.0, S., LLC, New York, NY **2008**.
23. MacroModel version 9.6, S., LLC, New York, NY **2008**.
24. SYBYL 7.3, T. I., 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA.