J. Chosun Natural Sci.
Vol. 8, No. 1 (2015) pp. 40 – 47
http://dx.doi.org/10.13160/ricns.2015.8.1.40

A CoMFA Study of Glycogen Synthase Kinase 3 Inhibitors

Anand Balupuri[1†], Pavithra K. Balasubramanian[1†], and Seung Joo Cho[1,2†]

Abstract

Glycogen synthase kinase 3 (GSK-3) is a serine/threonine protein kinase that has recently emerged as a promising target in drug discovery. It is involved in multiple cellular processes and associated with the pathogenesis of several diseases. A three-dimensional quantitative structure-activity relationship (3D-QSAR) analysis was performed on a series of GSK-3 inhibitors to understand the structural basis for inhibitory activity. Comparative molecular field analysis (CoMFA) method was used to derive 3D-QSAR models. A reliable CoMFA model was developed using ligand-based alignment scheme. The model produced statistically acceptable results with a cross-validated correlation coefficient (q^2) of 0.594 and a non-cross-validated correlation coefficient (r^2) of 0.943. Robustness of the model was checked by bootstrapping and progressive scrambling analysis. This study could assist in the design of novel compounds with enhanced GSK-3 inhibitory activity.

Keywords: Glycogen Synthase Kinase 3, GSK-3 Inhibitors, 3D-QSAR, CoMFA

1. Introduction

Glycogen synthase kinase-3 (GSK-3) is a multifunctional serine/threonine protein kinase. It is ubiquitously expressed protein kinase that exists in two isoforms, α and β[1,2]. It is involved in diverse physiological pathways ranging from metabolism, cell cycle, gene expression, development and oncogenesis to neuroprotection. It has been implicated in several diseases such as diabetes, inflammation, cancer, Alzheimer's and bipolar disorder[3-6]. High-level expression of GSK-3 in brain is associated with a variety of neurological disorders like Alzheimer's, bipolar disorder, Huntington disease and other neurodegenerative disorders[7,8]. Due to the therapeutic potential, development of GSK-3 inhibitors is a focus of research for pharmaceutical companies and academic researchers[9].

In recent years, various efforts have been made to discover GSK-3 inhibitors. The availability of GSK-3 crystal structures allows structure-based lead discovery and optimization[10,11]. Consequently, several GSK-3 inhibitors have been reported till date. Recently, a novel class of 5-aryl-4-carboxamide-1,3-oxazoles has been reported as potent inhibitors of GSK-3[12]. However, a quantitative structure-activity relationship (QSAR) study on these inhibitors has not been carried out to establish exactly how their chemical structures relate to the inhibitory activities. Our research group has reported several molecular modeling studies[13-18]. In this study, we have performed comparative molecular field analysis (CoMFA) to identify the key structural elements that are required in the rational design of potential drug candidates of this class.

2. Methodology

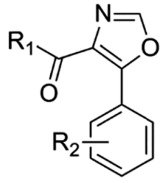
2.1. Data Set

A data set of 31 GSK-3 inhibitors was collected from the literature[12]. Biological activities of the inhibitors were reported as pIC_{50} values. These values span 4 log units, suggesting a suitable data set for a QSAR study. The reported pIC_{50} values were used as the dependent variables in the CoMFA model. The X-ray crystal structure 4AFJ with resolution of 1.98 Å[12] was obtained from the Protein Data Bank (PDB, http://www.rcsb.org/pdb). The extracted co-crystallized ligand (compound 7) was used as template to construct the 3D structures of all the compounds. The chemical structures and activities

Departments of [1]Biomedical Sciences and [2]Cellular.Molecular Medicine, College of Medicine, Chosun University, Gwangju 501-759, Korea

[†]Corresponding author : anandbalupuri.niper@gmail.com, pavithrabioinfo@gmail.com, chosj@chosun.ac.kr
(Received : March 3, 2015, Revised : March 16, 2015, Accepted : March 25, 2015)

Table 1. Chemical structures and biological activities of GSK-3 inhibitors



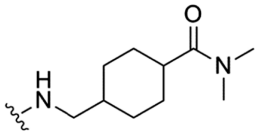
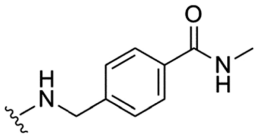
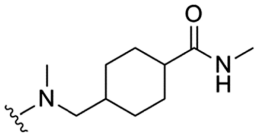
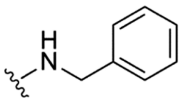
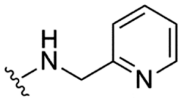
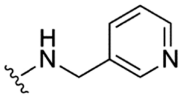
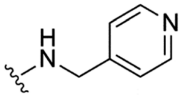
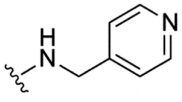
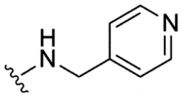
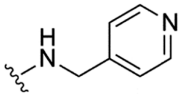
Compound	R_1	R_2	pIC_{50}
1		4-OMe	6.5
2*		4-OMe	6.6
3		4-OMe	4.6
4		4-OMe	6.0
5		4-OMe	5.7
6		4-OMe	6.6
7		4-OMe	6.4
8		3-OH	6.2
9		4-CF_3	5.2
10		4-CN	5.5

Table 1. Continued

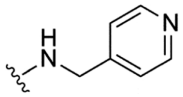
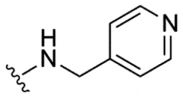
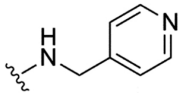
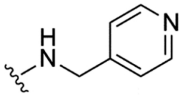
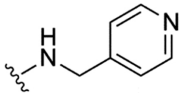
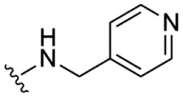
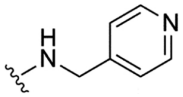
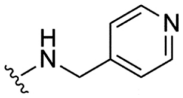
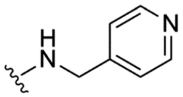
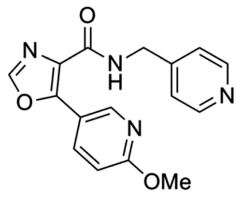
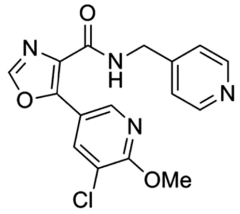
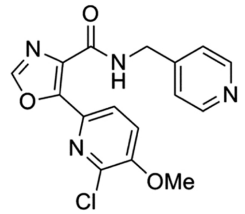
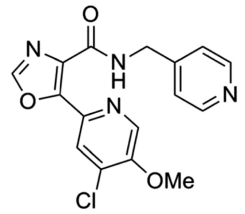
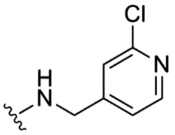
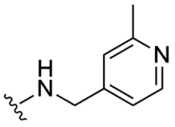
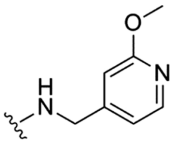
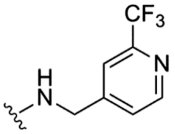
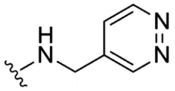
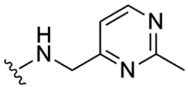
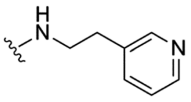
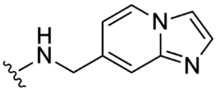
Compound	R_1	R_2	pIC_{50}
11		3-OMe	5.6
12		3-$CONH_2$	5.1
13		3-Cl, 4-OMe	7.9
14		3-F, 4-OMe	7.5
15		2-F, 4-OMe	6.4
16*		2-Me, 4-OMe	5.5
17		3-(4-Pyridyl), 4-OMe	5.9
18		3-(4-Morpholinyl), 4-OMe	5.7
19		3-Cl, 4-OPh	6.1
20			7.1
21			7.2

Table 1. Continued

Compound	R_1	R_2	pIC_{50}
22			6.4
23			5.0
24		3-Cl, 4-OMe	7.7
25		3-Cl, 4-OMe	7.5
26		3-Cl, 4-OMe	7.5
27		3-Cl, 4-OMe	7.0
28		3-Cl, 4-OMe	7.7
29		3-Cl, 4-OMe	7.4
30		3-Cl, 4-OMe	7.7
31		3-Cl, 4-OMe	8.3

*Compounds are considered as outliers.

of the inhibitors are given in Table 1. Structures of the compounds were sketched using Sybylx2.0[19]. Gasteiger-Hückel partial atomic charges were applied. The common moiety was constraint for each compound and only the varying parts were energy minimized using Tripos force field. The minimized structures were aligned to the template compound using common substructure-based alignment method.

2.2. CoMFA

CoMFA is based on the hypothesis that changes in the biological activity are related to the changes in the steric and electrostatic fields of compounds[20]. CoMFA descriptors (steric and electrostatic field energies) were calculated using default settings. A 3D cubic lattice with a grid spacing of 2.0 Å was created and fields were generated using a sp^3 carbon probe atom carrying +1 charge and van der Waals radius of 1.50 Å. Steric and electrostatic fields were calculated from Lennard-Jones and Coulomb potentials, respectively. An energy cut-off of 30 kcal mol^{-1} was used.

Partial least squares (PLS) regression algorithm was used to derive relationship between the structural parameters and the biological activities[21]. In PLS analysis, pIC_{50} values were used as dependent variables and CoMFA descriptors were used as independent variables. The leave-one-out (LOO) cross-validation was performed to obtain cross-validated correlation coefficient (q^2), optimal number of components (NOC) and standard error of prediction (SEP). In LOO cross-validation, one compound was removed from the data set and its activity was predicted using the model derived from the remaining compounds. The non-cross-validated analysis was carried out to determine non-cross-validated correlation coefficient (r^2), standard error of estimate (SEE) and F-test value (F). Developed model was further validated using bootstrapping analysis and progressive scrambling. Bootstrapping of 100 runs and a total of 100 independent scramblings with a minimum of 2 bins and a maximum of 10 bins were carried out.

3. Results and Discussion

3.1. CoMFA Model

CoMFA models were developed using a series of GSK-3 inhibitors. Ligand-based alignment scheme was used for generation of CoMFA models from the whole data set. All compounds were aligned over the template using common substructure alignment method. The aligned compounds are shown in Fig. 1. Data set was not divided into training and test sets because only 31 compounds were available for model generation. On the basis of predictions from the CoMFA models, two compounds (**2** and **16**) with high residual values were removed from data set as outliers.

A reliable CoMFA model was obtained with Gasteiger-Hückel partial charge in terms of several statistical parameters. The detailed statistical values for the ligand-based CoMFA model are listed in Table 2. PLS analysis showed a high q^2 value of 0.594 with 4 components. The non-cross-validated analysis produced r^2,

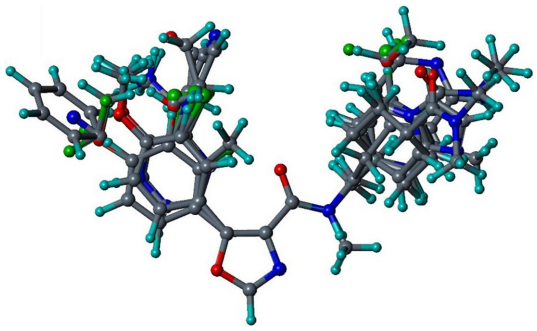


Fig. 1. Alignment of data set compounds based on common substructure using compound **7** as a template.

Table 2. Statistical parameters of the CoMFA model

Parameters	CoMFA
q^2	0.594
NOC	4
SEP	0.687
r^2	0.943
SEE	0.257
F	99.574
BS-r^2	0.956
BS-sd	0.019
Q^2	0.485
Steric contribution	47.9
Electrostatic contribution	52.1

Note: q^2 is cross-validated correlation coefficient, NOC is number of components, SEP is standard error of prediction, r^2 is non-cross-validated correlation coefficient, SEE is standard error of estimation; F is F-test value, BS-r^2 is bootstrapping r^2 mean, BS-SD is bootstrapping standard deviation, Q^2 is corrected q^2 dependency.

Table 3. Actual and predicted pIC_{50} values of data set compounds with their residuals

Compound	Actual pIC_{50}	CoMFA	
		Predicted pIC_{50}	Residual
1	6.5	6.635	-0.135
3	4.6	4.485	0.115
4	6.0	5.931	0.06935
5	5.7	5.672	0.028
6	6.6	6.485	0.115
7	6.4	6.787	-0.387
8	6.2	5.854	0.346
9	5.2	5.222	-0.022
10	5.5	5.538	-0.038
11	5.6	5.915	-0.315
12	5.1	5.107	-0.007
13	7.9	7.358	0.542
14	7.5	6.949	0.551
15	6.4	6.956	-0.556
17	5.9	5.935	-0.036
18	5.7	5.901	-0.202
19	6.1	6.083	0.017
20	7.1	7.190	-0.091
21	7.2	7.123	0.077
22	6.4	6.268	0.132
23	5.0	5.103	-0.103
24	7.7	7.429	0.271
25	7.5	7.613	-0.113
26	7.5	7.592	-0.092
27	7.0	7.119	-0.119
28	7.7	7.769	-0.069
29	7.4	7.343	0.057
30	7.7	7.553	0.147
31	8.3	8.485	-0.185

SEE and F values of 0.943, 0.257 and 99.574, respectively. The steric and electrostatic contributions were found to be 47.9% and 52.1%, respectively. The predicted and actual activities for the inhibitors along with the residual values are given in Table 3. The scatter plot for actual versus predicted pIC_{50} values is presented in Fig. 2. Activities predicted by the model are in agreement with the experimental values suggesting that a reliable CoMFA model was developed.

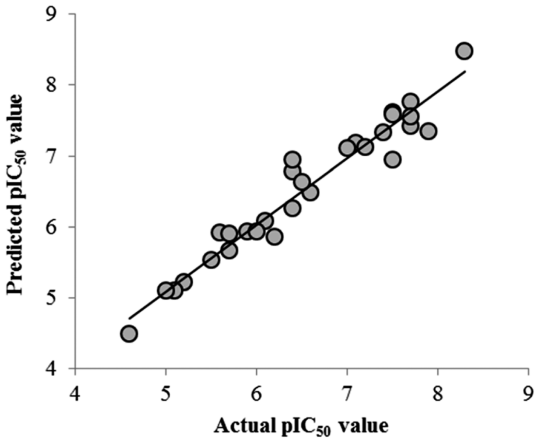


Fig. 2. Scatter plot of the actual versus predicted activities based on the CoMFA model.

3.2. CoMFA Contour Maps

The CoMFA results were graphically interpreted by the field contribution maps using the STDEV*COEFF field type. The contour maps of different field contribution of the model are presented with template (compound **7**). Maps describe default 80% and 20% level contributions for favorable and unfavorable regions, respectively. Maps show spatial requirement of steric and electrostatic fields for improving the inhibitory activity.

Steric contour map is shown in Fig. 3. Green contours represent favorable regions for bulky group substitution while yellow contours represent unfavorable regions for

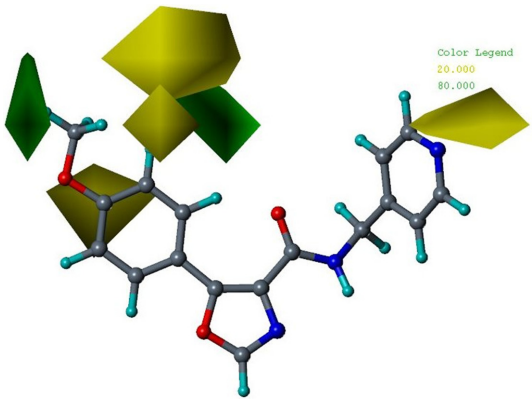


Fig. 3. CoMFA steric contour map with template (compound **7**) as a reference. Green contours represent sterically favored regions while yellow contours represent sterically unfavorable regions.

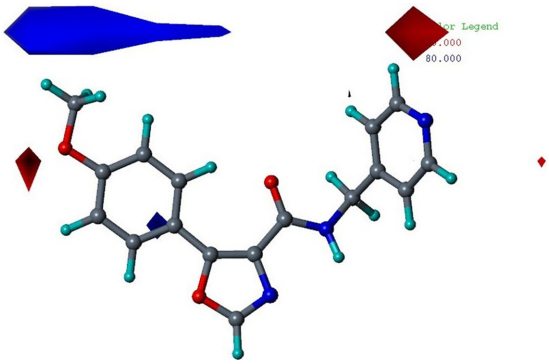


Fig. 4. CoMFA electrostatic contour map with template (compound **7**) as a reference. Blue contours indicate favorable regions for electropositive substituents while red indicate favorable regions for electronegative substituents.

bulky group substitution. A green contour seen near 4-position of the phenyl ring (R_2 substitution) suggested that bulky group in this region is favorable to enhance the activity. This might the possible reason for better inhibitory activities of compounds **6** and **7** as compared to compounds **9** and **10**. Yellow contours observed around 3-position of the phenyl ring of indicate that bulky substituents in this region could reduce the activity. This could be the possible reason for the lower activities of compounds **11** and **12** as compared to compound **8**. Similarly, compounds **17** and **18** with bulky substitutions exhibit lower activities as compared to compounds **13** and **14**. One more yellow contour near pyridine ring of the R_1 substitution indicates that small groups in the region could be favorable to enhance the activity. Thus, compound **4** with smaller R_1 substituent than compound **3** display higher activity.

Electrostatic contour map is shown in Fig. 4. Blue contours indicate regions where electropositive substitutions enhance the activity whereas red contours indicate regions where electronegative substitutions increase the potency. A big blue contour observed near methoxy group of R_2 substitution suggests that electropositive substitution in that position could increase the activity of compounds. This could be the reason why compounds **7** exhibit higher activity than compounds **9** and **10**. The small red contour seen near the pyridine ring of the R_1 position suggests that the substitution that have electronegative group in that particular position could improve the activity. This could be the reason why compound **24** containing chlorine atom in that position exhibit higher activity than compound **25**.

4. Conclusions

In this study, 3D-QSAR models were developed for a series of GSK-3 inhibitors. CoMFA provided some insights into the key structural factors affecting the bioactivity of these inhibitors. A reliable ligand-based CoMFA model was obtained with Gasteiger-Hückel partial charges. The model exhibited statistically significant results in terms of q^2 and r^2 values. Model was validated using bootstrapping and progressive sampling. Validation results indicated that the model is predictable and robust. Analysis of contour maps provided information to understand the SAR and highlighted regions to improve the activity of the compounds. Smaller groups with electronegative properties at R_1 substitution whereas electropositive groups at R_2 substitution are desirable to improve the inhibitory potency. These results may provide some useful and rational suggestions for further design of novel GSK-3 inhibitors.

Acknowledgements

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 (NRF-2012R1A1A4A01001465).

References

[1] S. Frame, P. Cohen, and R. M. Biondi, “A common phosphate binding site explains the unique substrate specificity of GSK3 and its inactivation by phosphorylation”, Mol. Cell, Vol. 7, pp. 1321-1327, 2001.

[2] N. Embi, D. B. Rylatt, and P. Cohen, “Glycogen synthase kinase-3 from rabbit skeletal muscle”, European journal of biochemistry, Vol. 107, pp. 519-527, 1980.

[3] S. E. Nikoulina, T. P. Ciaraldi, S. Mudaliar, P. Mohideen, L. Carter, and R. R. Henry, “Potential role of glycogen synthase kinase-3 in skeletal muscle insulin resistance of type 2 diabetes”, Diabetes, Vol. 49, pp. 263-271, 2000.

[4] E. Planel, X. Sun, and A. Takashima, “Role of GSK-3β in Alzheimer's disease pathology”, Drug

Develop. Res., Vol. 56, pp. 491-510, 2002.
[5] C. J. Phiel and P. S. Klein, "Molecular targets of lithium action", Annu. Rev. Pharmacol., Vol. 41, pp. 789-813, 2001.
[6] M. K. Rowe, C. Wiest, and D.-M. Chuang, "GSK-3 is a viable potential target for therapeutic intervention in bipolar disorder", Neurosci. Biobehav. R., Vol. 31, pp. 920-931, 2007.
[7] R. S. Jope and M.-S. Roh, "Glycogen synthase kinase-3 (GSK3) in psychiatric diseases and therapeutic interventions", Curr. Drug Targets, Vol. 7, pp. 1421-1434, 2006.
[8] M. P. Mazanetz and P. M. Fischer, "Untangling tau hyperphosphorylation in drug design for neurodegenerative diseases", Nat. Rev. Drug Discov., Vol. 6, pp. 464-479, 2007.
[9] R. G. Gentles, S. Hu, and G. M. Dubowchik, "Recent advances in the discovery of GSK-3 inhibitors and a perspective on their utility for the treatment of Alzheimer's disease", Annu. Rep. Med. Chem., Vol. 44, pp. 3-26, 2009.
[10] M. Aoki, M. Iwamoto-Sugai, I. Sugiura, C. Sasaki, T. Hasegawa, C. Okumura, S. Sugio, T. Kohno, and T. Matsuzaki, "Expression, purification and crystallization of human tau-protein kinase I/glycogen synthase kinase-3β", Acta Crystallogr. D, Vol. 56, pp. 1464-1465, 2000.
[11] B. Bax, P. S. Carter, C. Lewis, A. R. Guy, A. Bridges, R. Tanner, G. Pettman, C. Mannix, A. A. Culbert, M. J. Brown, D. G. Smith, and A. D. Reith, "The structure of phosphorylated GSK-3β complexed with a peptide, FRATtide, that inhibits β-catenin phosphorylation", Structure, Vol. 9, pp. 1143-1152, 2001.
[12] G. Gentile, G. Merlo, A. Pozzan, G. Bernasconi, B. Bax, P. Bamborough, A. Bridges, P. Carter, M. Neu, G. Yao, C. Brough, G. Cutler, A. Coffin, and S. Belyanskaya, "5-Aryl-4-carboxamide-1,3-oxazoles: potent and selective GSK-3 inhibitors", Bioorg. Med. Chem. Lett., Vol. 22, pp. 1989-1994, 2012.
[13] A. Balupuri and S. J. Cho, "Exploration of the binding mode of indole derivatives as potent HIV-1 inhibitors using molecular docking simulations", J. Chosun Natural Sci., Vol. 6, pp. 138-142, 2013.
[14] P. K. Balasubramanian, A. Balupuri, and S. J. Cho, "A CoMFA study of phenoxypyridine-based JNK3 inhibitors using various partial charge schemes", J. Chosun Natural Sci., Vol. 7, pp. 45-49, 2014.
[15] P. K. Balasubramanian, A. Balupuri, and S. J. Cho, "Ligand-based CoMFA study on pyridylpyrazolopyridine derivatives as PKCè kinase inhibitors", J. Chosun Natural Sci., Vol. 7, pp. 253-259, 2014.
[16] P. K. Balasubramanian and S. J. Cho, "HQSAR analysis on novel series of 1-(4-phenylpiperazin-1-yl-2-(1H-Pyrazol-1-yl) ethanone derivatives targeting CCR1", J. Chosun Natural Sci., Vol. 6, pp. 163-169, 2013.
[17] S. J. Cho, "The importance of halogen bonding: a tutorial", J. Chosun Natural Sci., Vol. 5, pp. 195-197, 2012.
[18] C. G. Gadhe, G. Kothandan, and Seung Joo Cho, "Large variation in electrostatic contours upon addition of steric parameters and the effect of charge calculation schemes in CoMFA on mutagenicity of MX analogues", Mol. Simulat., Vol. 38, pp. 861-871, 2012.
[19] SYBYLx2.1, Tripos International, 1699 South Hanley Road, St. Louis, Missouri, 63144, USA.
[20] R. D. Cramer, D. E. Patterson, and J. D. Bunce, "Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins", J. Am. Chem. Soc., Vol. 110, pp. 5959-5967, 1988.
[21] S. Wold, A. Ruhe, H. Wold, and W. J. Dunn III, "The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses", SIAM Journal on Scientific and Statistical Computing, Vol. 5, pp 735-743, 1984.