

Docking and QSAR studies of PARP-1 Inhibitors

PARP-1 억제제의 Docking 및 QSAR 연구

Hye-jung Kim, Seung Joo Cho*

Biochemical Research Center, Korea Institute of Science and Technology, PO Box 131, Seoul 130-650, Korea

*To whom correspondence should be addressed. E-mail: chosj@kist.re.kr

Abstract

Poly(ADP-ribose)polymerase-1 (PARP-1) is a nuclear enzyme involved in various physical functions related to genomic repair, and PARP inhibitors have therapeutic application in a variety of neurological diseases. Docking and the QSAR (quantitative structure-activity relationships) studies for 52 PARP-1 inhibitors were conducted using FlexX algorithm, comparative molecular field analysis (CoMFA), and hologram quantitative structure-activity relationship analysis (HQSAR). The resultant FlexX model showed a reasonable correlation ($r^2 = 0.701$) between predicted activity and observed activity. Partial least squares analysis produced statistically significant models with q^2 values of 0.795 (SDEP=0.690, $r^2=0.940$, $s=0.367$) and 0.796 (SDEP=0.678, $r^2=0.919$, $s=0.427$) for CoMFA and HQSAR, respectively. The models for the entire inhibitor set were validated by prediction test and scrambling in both QSAR methods. In this work, combination of docking, CoMFA with 3D descriptors and HQSAR based on molecular fragments provided an improved understanding in the interaction between the inhibitors and the PARP. This can be utilized for virtual screening to design novel PARP-1 inhibitors.

Introduction

Poly(ADP-ribose) polymerase-1 (PARP-1) also known as poly(ADP-ribose) synthetase is an abundant nuclear enzyme implicated in cellular responses to DNA injury provoked by genotoxic stress.¹⁻³ The results obtained with PARP knockout mice have demonstrated that PARP-1 deficiency might be involved in neuronal damage derived from ischemic injury, neurotoxin-induced parkinsonism or traumatic brain injury.⁴⁻⁷ Genetic disruption of PARP-1 gene provides protection

against myocardial ischemia and reperfusion, streptozotocin-induced diabetes and endotoxic shock.⁸⁻¹⁰ Inhibition of PARP is exploited essentially with the scope of suppressing PARP-1 activity, that on one hand protects cells from certain types of DNA damage, while on the other promotes cell death in response to massive DNA breakage. PARP-1 could therefore be considered as a potential target for the development of pharmacological strategies to enhance the antitumor efficacy.¹¹ PARP-1 consists of DNA binding domain, automodification site and

catalytic domain.¹² Most of the PARP inhibitors so far developed bind to the nicotinamide-binding domain, where they act as competitive inhibitors.¹³

In this paper, we used an approach combining FlexX²³⁻²⁵, CoMFA¹⁴ and HQSAR¹⁵ methods to describe QSAR models for series of PARP inhibitors. CoMFA captures 3D information concerning the ligands and the specific conformation of the molecules, while HQSAR gives sub-structural features in sets of molecules that are relevant to biological activity. All available inhibitors were collected from the literature and reported in Table 1 along with their inhibitory IC₅₀ values.^{5,16-19} In addition to QSAR, a prediction test and randomization test were used to determine the statistical significance of the analyses.²⁰

Methods

Molecular Modeling

Both CoMFA and HQSAR studies were performed using the SYBYL 6.9.2²¹ molecular modeling software package on a Silicon Graphics workstation, Origin R3000 processors. All the compounds were fully optimized using Tripos molecular mechanics force field with a distance-dependent dielectric constant and an energy gradient convergence criterion of 0.05 kcal/mol. The partial atomic charges required for the calculation of the electrostatic potential were assigned using the PEOE method.²²

Molecular Docking

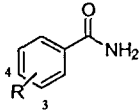
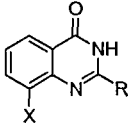
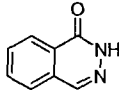
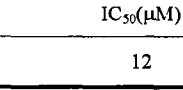
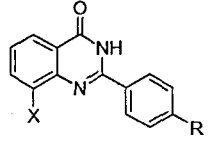
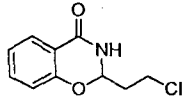
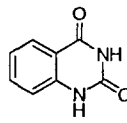
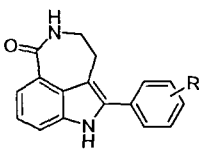
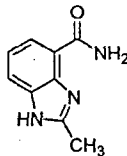
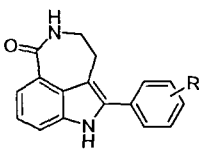
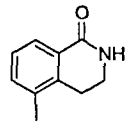
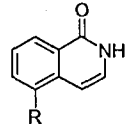
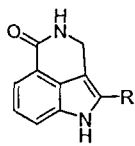
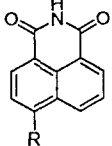
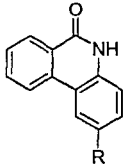
All ligand molecules that were included for 3D

QSAR studies were used in the FlexX²³⁻²⁵ docking study. FlexX is a fast automated docking program that considers ligand conformational flexibility by an incremental fragment placing technique. The crystal structure (3pax) of chicken PARP-1 was used in this study. The active site for docking was defined as all amino acid within 8.0 Å proximity of the co-crystallized compound. FlexX total energy scores were correlated with the inhibitory activities for each ligand using the least-squares fit method.

CoMFA Analysis

Superimposition of the molecules was carried out using the pose of crystallized complexes of PARP-1 with FlexX docked binding mode and some inhibitors as the template (2paw, 1efy, 1pax, 2pax, 3pax, 4pax). The overlapped molecules were surrounded by a 3D grid of points extending in three dimensions to at least 4 Å beyond the union volume occupied by the superimposed molecules. The default sp³ carbon atom with +1 charge was selected as the probe atom for the calculations of steric and electrostatic fields around the aligned molecules. The probe-ligand interaction energies were calculated using a Lennard-Jones 6-12 potential and Coulombic potential with a distance-dependent dielectric, respectively. Values of steric and electrostatic energies were truncated at 30 kcal/mol. The entire set of inhibitors was divided into two groups in the approximate ratio of 3:1, *i.e.*, 12 inhibitors were used as a test set to assess the predictive power. This experiment was repeated ten times.

Table 1. Inhibitory values (IC₅₀) of selected classes of PARP inhibitor.

									
Code	R	IC ₅₀ (μM)	Code	R	X	IC ₅₀ (μM)	Code	R	IC ₅₀ (μM)
1	H	22.0	20	H	H	9.50	41	H	0.30
2	3-F	20.0	21	CH ₃	H	5.60	42	NO ₂	0.35
3	4-F	200	22	CH ₃	OH	0.40			
4	3-Cl	22.0					Code	IC ₅₀ (μM)	
5	4-Cl	300					43	12	
6	3-NH ₂	33.0							
7	4-NH ₂	1800	Code	R	X	IC ₅₀ (μM)	Code	IC ₅₀ (μM)	
8	3-Br	55	23	H	CH ₃	0.87	44	8.5	
9	4-Br	2200	24	H	OCH ₃	4.2			
10	3-OCH ₃	17.0	25	NO ₂	CH ₃	0.13	Code	IC ₅₀ (μM)	
11	3-OH	9.10	26	NO ₂	OCH ₃	0.85	45	8.1	
12	4-OH	280	27	CF ₃	OCH ₃	39			
13	3-CH ₃	19.0	28	CN	CH ₃	0.27	Code	IC ₅₀ (μM)	
14	3-NO ₂	160	29	CN	OCH ₃	1.34	46	3-CH ₂ NMe ₂	0.008
15	3-N(CH ₃) ₂	120	30	OCH ₃	CH ₃	0.19	47	4-CH ₂ NMe ₂	0.005
			31	OCH ₃	OCH ₃	2.0	48	4-CH ₂ -(1-pyrrolidiny)	0.006
Code	IC ₅₀ (μM)		32	NH ₂	CH ₃	0.44			
16	1.00		33	CO ₂ CH ₃	CH ₃	4.80	Code	R	IC ₅₀ (μM)
			34	H	OH	1.06	49	H	0.047
Code	IC ₅₀ (μM)		35	NO ₂	OH	0.23	50	Ph	0.005
17	0.44		36	OH	OH	0.29	51	(3-CH ₂ NMe ₂)Ph	0.005
			37	OH	CH ₃	0.22	52	(4-CH ₂ NMe ₂)Ph	0.005
Code	R	IC ₅₀ (μM)	38	NH ₂	OH	0.52			
18	H	7.00					Code	R	IC ₅₀ (μM)
19	OH	0.39	Code	R	IC ₅₀ (μM)	49	H	0.047	
			39	H	1.40	50	Ph	0.005	
			40	NH ₂	0.18	51	(3-CH ₂ NMe ₂)Ph	0.005	
							52	(4-CH ₂ NMe ₂)Ph	0.005

Furthermore, for the whole set, to detect possible chance correlations, the scrambling test was repeated ten times.

Hologram QSAR (HQSAR)

For HQSAR, the HQSAR module of SYBYL was used. All the sub-structural fragments in the size range were generated for each molecule. The SLN (SYBYL Line Notation) for each fragment generated was mapped to a unique integer. The substructure fingerprints were then hashed into hologram bins with lengths from 53 to 401. In order to get good models from HQSAR analysis, we investigated the influence of the fragment distinction and the fragment size on the key statistical parameters. Chirality flags were not utilized in the present study. The best model was chosen based on the least standard error of cross-validated predictions.

PLS Analysis

PLS analysis was carried out using the leave-one-out option to obtain the optimal number of components to be used subsequently in the final analysis. The cross-validated coefficient q^2 was calculated using formula

$$q^2 = 1 - \frac{\sum (Y_{obs} - Y_{pred})^2}{\sum (Y_{obs} - Y_{mean})^2}$$

where Y_{pred} , Y_{obs} , and Y_{mean} are predicted, observed, and mean values of the target property (pIC_{50}), respectively. The optimal number of components was designated as that which yielded the highest q^2 values and the smallest rms error values.

Results and Discussion

Molecular docking

We conducted docking analysis of all data sets on the active site of the PARP-1 enzyme (3pax) using FlexX program. The inhibitors 10 (3pax), 17 (1pax), 22 (4pax), 40 (2pax) were selected as the starting set of 52 inhibitors and were tested for ability to reproduce the crystallized binding geometry. In all results, FlexX easily found the binding geometry corresponding to the crystallized complex among the solutions with the lowest energy, being the rms deviation between the docked and the crystallized geometry in the limitation of the crystallographic resolution (rms <1.0 Å, resolution 2.4 Å). Conserved important interaction in all the docked PARP-inhibitor complexes is hydrogen bond between the amide moiety and the backbone of Gly863, and π - π interaction between the aromatic group of the inhibitors and the side chain of Tyr907. The best possible binding modes of the most active inhibitors (50, 52) at the active site are displayed in Figure 1. To evaluate the relationship between activities and binding energy scores obtained from FlexX docking, we have used least-squares fit analyses. The equations were obtained for the inhibitory activities represented as pIC_{50} values, using the binding energies as the sole descriptor variable with good correlation ($r^2 = 0.701$).

$$pIC_{50} = 0.020 - 0.255\Delta E$$

Figure 2 shows the correlation between predicted activities and observed activities for the total set.

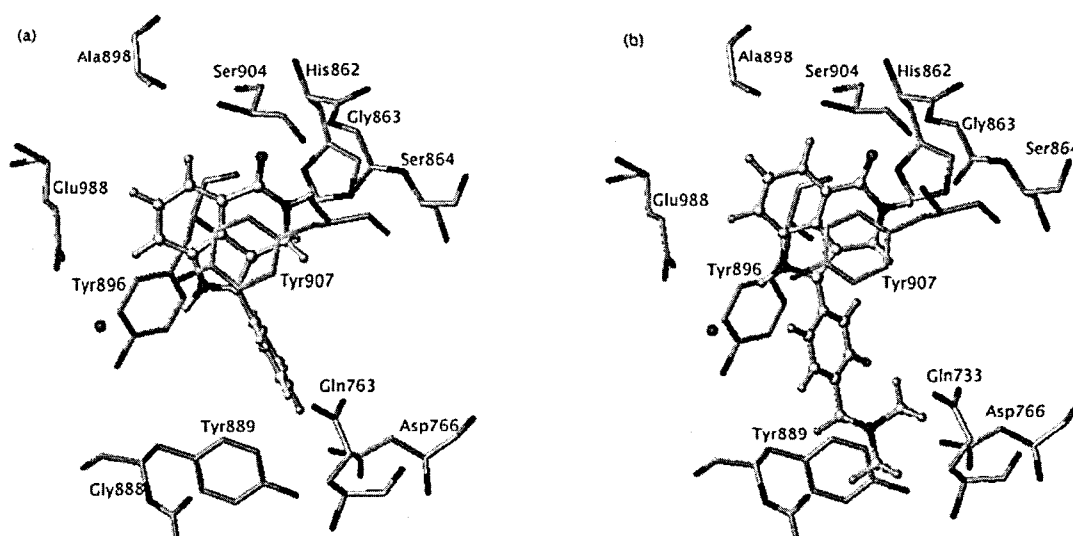


Figure 1. Docking pose of inhibitor 50 (a), 52 (b)

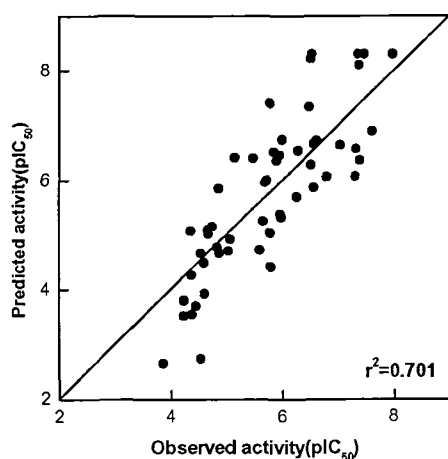


Figure 2. Observed activity (pIC_{50}) versus predicted activity (kcal/mol) calculated by FlexX docking

CoMFA Analysis

Superimposition of the molecules was carried out based on the pose of crystallized inhibitors and the docked geometry. Figure 3 shows the aligned molecules within the grid box used to generate the CoMFA column. The statistical results for CoMFA models based on steric and electrostatic in interactions are summarized in

Table 2. The optimal CoMFA model for the entire training set gave $q^2=0.795$ with 5 principal components. Figure 4 illustrates the cross-validated correlation curves for our CoMFA model.

Table 2. Statistical Results of CoMFA.

	Model
N^a	5
$q^2{}^b$	0.795
SDEP ^c	0.690
$r^2{}^d$	0.940
S ^e	0.367
F value ^f	144.965
Ste ^g : Ele ^h	0.603 : 0.397

^aNumber of PLS components in analysis ^bSquared correlation coefficient of a cross-validated analysis ^cStandard deviation of error of predictions ^dSquared correlation coefficient of a non-cross-validated analysis ^eStandard deviation of a non-cross-validated analysis ^fF-ratio ^gMolecular field used in CoMFA (Ste : steric field, Ele : electrostatic field)

More rigorous statistical tests, *i.e.*, prediction and scrambling test, were performed. Noticeable difference of the statistical parameters of the

prediction test ($r_{\text{pred}}^2 = 0.741 \pm 0.044$, $s_{\text{pred}} = 0.695 \pm 0.060$) from those of the scrambling ($q^2 = 0.008 \pm 0.098$, $\text{SDEP} = 1.192 \pm 0.054$) reflects the stability and robustness of our CoMFA model.

To visualize the information content of the model, the CoMFA contour maps on the one (50) of the most active inhibitors is shown in Figure 5. In steric field, the green contour around tricyclic ring

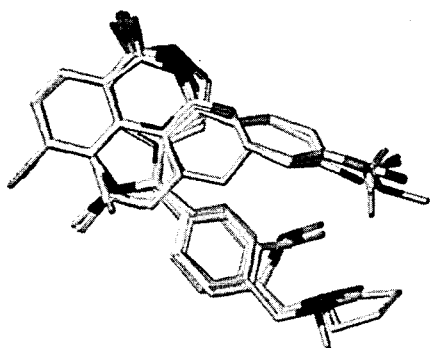


Figure 3. Alignment of PARP inhibitors used in the CoMFA model.

suggests that more bulky groups are required to increase activity. In our data set, cyclic derivatives have better activity than that of benzamide derivatives and inhibitors with C-2 aryl substituted indole core display relatively good activity by forming hydrophobic interaction with Tyr896 and Tyr889 residue.

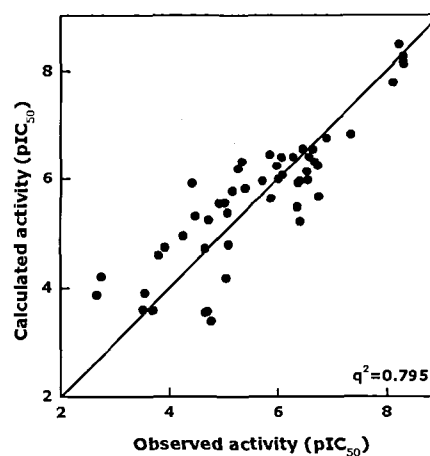


Figure 4. Observed versus calculated activity (pIC₅₀) of CoMFA model

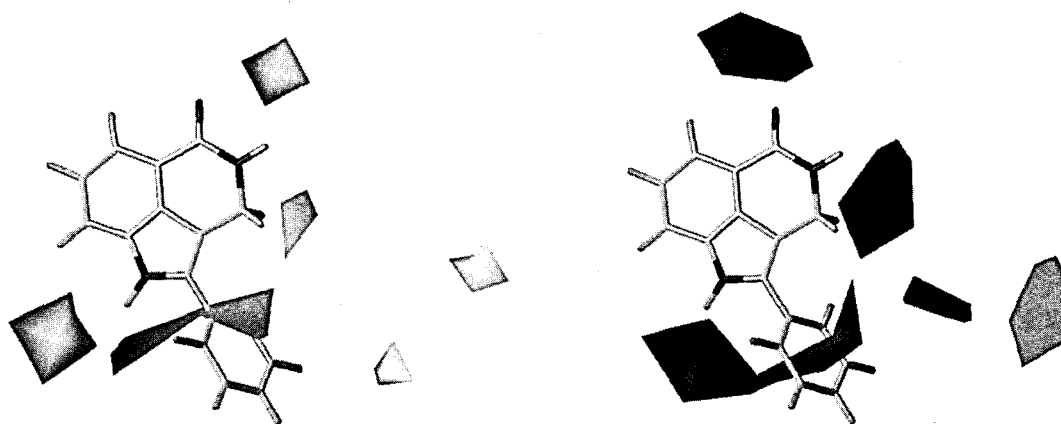


Figure 5. CoMFA STDEV*COEFF contour plots. Green contour (contribution level of 80%) indicates region where bulky group increases activity, whereas yellow contour (contribution level of 20%) indicates region where bulky group decreases activity. Blue contour (contribution level of 80%) indicates region where positive group increases activity, whereas red contour (contribution level of 20%) indicates region where negative charge increases activity.

The yellow contours suggest that bulky groups in these regions are not beneficial to activity because of bad contact with Glu988, Gln763 and Asn767 residues. In electrostatic interaction field, the blue contour around the sp^2 oxygen of tricyclic ring suggests that the cyclic derivatives of benzamide are more potent inhibitors. The blue contour on right of tricyclic ring shows that groups with positive charge in this area increase activity. The blue contour near the lower part of tricyclic ring indicates that hydrophobic groups in these regions are beneficial to activity and a hydrogen bond donor at this position enhances activity by forming hydrogen bonds with the carboxyl group of Glu 988 residues. The red contours display that groups with negative charge in these regions are favorable to activity by forming hydrogen bond with Asn767 residue.

HQSAR analysis

The data in Table 3 shows that predictive HQSAR models are readily derived using atoms, bond and connectivity-distinction information. Adding hydrogen distinction into molecular hologram does not appear to improve the model as measured by statistical parameters. We employed different fragment sizes on the best model for fragment distinction and the results are summarized in Table 4. The best model was found when default fragment size option (4-7) was used with 6 optimum components. A plot of observed and calculated activities for cross-validated model is shown in Figure 6.

As an additional test of the predictability of the models, prediction and scrambling test were performed as for CoMFA model.

Table 3. HQSAR analysis for various fragments distinction on statistical parameters using fragment size default (4-7).

Fragment distinction	Statistical parameters				
	q^2	SDEP	r^2	s	PCA
At-Bon ^a	0.795	0.681	0.931	0.396	5
Con ^b	0.796	0.678	0.919	0.427	5
Con-H ^c	0.771	0.704	0.901	0.518	4

^aAtoms and Bond ^bAtom, bond and connection ^cAtom, bond, connection and Hydrogen Bond

Table 4. HQSAR analysis for the influence of various fragments sizes on the key statistical parameters using the best fragment distinction (atom, bonds and connectivity).

Fragment size	Statistical parameters				
	q^2	SDEP	r^2	s	PCA
3-6	0.769	0.731	0.930	0.403	6
5-8	0.777	0.711	0.924	0.416	6
6-9	0.758	0.747	0.914	0.445	6
7-10	0.752	0.749	0.915	0.438	6

The result of prediction test ($r^2_{pred} = 0.748 \pm 0.075$, $s_{pred} = 0.628 \pm 0.069$) showed obvious difference from those of the scrambling test ($q^2 = 0.075 \pm 0.065$, SDEP = 1.132 ± 0.038).

To visualize information for activity from its fingerprint, the most important fragments of most active inhibitor (50) were color-coded by results of the final PLS analysis. Figure 7 shows tricyclic core contribute positively to activity.

In summary, the hologram length that leads to the best PLS analysis was used in development of the pharmacophore. The CoMFA model reveals regions can be correlated with the appropriate enzyme residues in 3D space around inhibitors. Combination of CoMFA and HQSAR with considerable statistical quality provided pharma-

cophore and 3D information surrounding it to design novel PARP inhibitors.

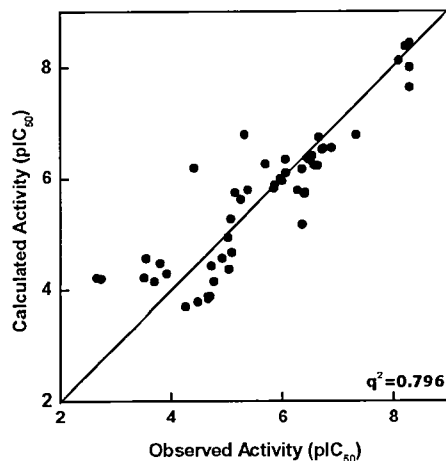


Figure 6. Observed versus calculated activity (pIC₅₀) of HQSAR model

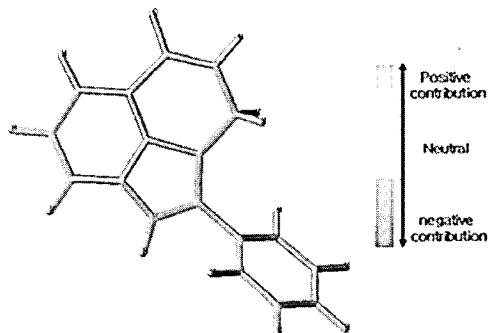


Figure 7. The HQSAR contribution map for PARP inhibitors (51).

Conclusion

In the present study, we tried to derive reasonable QSAR model through combining of CoMFA and HQSAR method. The prediction power for the entire inhibitor set was validated by prediction and scrambling tests for both QSAR methods. Although it is still a widely used tool for the study of QSAR at the 3D level, CoMFA tool is limited by the need to align the database molecules correctly within 3D space. On the other hand,

HQSAR requires only information about 2D molecular structure. The similar statistical results of them make it possible to overcome and complement the limitations for CoMFA and HQSAR method. This successful combination of CoMFA model with 3D descriptors and HQSAR model based on representation of the molecular fragments would be useful to address the design of new potent PARP inhibitor.

Reference

- [1]. D'Amours, D.; Desnoyers, S.; D'Silva, I.; Poirier, G. G. *Biochem. J.* 342: 249-268 (1999).
- [2]. Ziegler, M.; Oei, S. L. *Bioessays.* 23: 543-548 (2001),.
- [3]. Pieper, A. A.; Verma, A.; Zhang, J.; Snyder, S. H. *Trends. Pharmacol. Sci.* 20: 171-181 (1999).
- [4]. Shall, S.; de Murcia, G. *Mutat. Res.* 460: 1-15 (2000).
- [5]. Eliasson, M. J.; Sampei, K.; Mandir, A. S.; Hurn, P. D.; Traystman, R. J. et al. *Nat. Med.* 3: 1089-1095 (1997).
- [6]. Mandir, A. S.; Przedborski, S.; Jackson-Lewis, V.; Wang, Z. Q.; Simbulan-Rosenthal, C. M. et al. *Proc. Natl. Acad. Sci. U S A* 96: 5774-5779 (1999).
- [7]. Whalen, M. J.; Clark, R. S.; Dixon, C. E.; Robichaud, P.; Marion, D. W. et al. *J. Cereb. Blood. Flow. Metab.* 19: 835-842 (1999).
- [8]. Zingarelli, B.; Salzman, A. L.; Szabo, C. *Circ. Res.* 83: 85-94 (1998).
- [9]. Pieper, A. A.; Brat, D. J.; Krug, D. K.; Watkins, C. C.; Gupta, A. et al. *Proc. Natl. Acad. Sci. U S A* 96: 3059-3064 (1999).
- [10]. Oliver, F. J.; Menissier-de Murcia, J.; Nacci, C.; Decker, P.; Andriantsitohaina, R. et al. *Embo.*

- J.* 18: 4446-4454 (1999).
- [11]. Tentori, L.; Portarena, I.; Graziani, G. *Pharmacol. Res.* 45: 73-85 (2002).
- [12]. Rolli, V.; O'Farrell, M.; Menissier-de Murcia, J.; de Murcia, G. *Biochemistry* 36: 12147-12154 (1997).
- [13]. Costantino, G.; Macchiarulo, A.; Camaioni, E.; Pellicciari, R. *J. Med. Chem.* 44: 3786-3794 (2001).
- [14]. Cramer, R. D.; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* 110: 5959-5967 (1988).
- [15]. HQSAR Tripos Associates, I., St. Louis, MO, USA.
- [16]. Banasik, M.; Komura, H.; Shimoyama, M.; Ueda, K. *J. Biol. Chem.* 267: 1569-1575 (1992).
- [17]. Griffin, R. J.; Srinivasan, S.; Bowman, K.; Calvert, A. H.; Curtin, N. J. et al. *J. Med. Chem.* 41: 5247-5256 (1998).
- [18]. Bowman, K. J.; White, A.; Golding, B. T.; Griffin, R. J.; Curtin, N. J. *Br. J. Cancer.* 78: 1269-1277 (1998).
- [19]. Canan Koch, S. S.; Thoresen, L. H.; Tikhe, J. G.; Maegley, K. A.; Almassy, R. J. et al. *J. Med. Chem.* 45: 4961-4974 (2002).
- [20]. van der Voet, H. *Chemometrics and Intelligent Laboratory Systems* 25: 313 (1994).
- [21]. Tripos Associates, I. MO 63144 (1699).
- [22]. Gasteiger, H.; Marsili, M. *Tetrahedron* 36: 3219-3288 (1980).
- [23]. Meng, E. C. S., B. K.; Kuntz, I. D. *J. comput. Chem.* 13: 505-524 (1992).
- [24]. Knegt, R. M.; Bayada, D. M.; Engh, R. A.; von der Saal, W.; van Geerestein, V. J. et al. *J. Comput. Aided. Mol. Des.* 13: 167-183 (1999).
- [25]. Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. *J. Comput. Aided. Mol. Des.* 11: 425-445 (1997).