

Comparative QSAR studies on peptide deformylase inhibitors

Ji Young Lee · Munikumar Reddy Doddareddy ·
Yong Seo Cho · Hyunah Choo · Hun Yeong Koh ·
Jae-Hoon Kang · Kyoung Tai No · Ae Nim Pae

Received: 24 April 2006 / Accepted: 18 January 2007 / Published online: 27 February 2007
© Springer-Verlag 2007

Abstract Comparative quantitative structure–activity relationship (QSAR) analyses of peptide deformylase (PDF) inhibitors were performed with a series of previously published (British Biotech Pharmaceuticals, Oxford, UK) reverse hydroxamate derivatives having antibacterial activity against *Escherichia coli* PDF, using 2D and 3D QSAR methods, comparative molecular field analysis (CoMFA), comparative molecular similarity indices analysis (CoMSIA), and hologram QSAR (HQSAR). Statistically reliable models with good predictive power were generated from all three methods (CoMFA $r^2=0.957$, $q^2=0.569$; CoMSIA $r^2=0.924$, $q^2=0.520$; HQSAR $r^2=0.860$, $q^2=0.578$). The predictive capability of these models was validated by a set of compounds that were not included in the training set. The models based on CoMFA and CoMSIA gave satisfactory predictive r^2 values of 0.687 and 0.505, respectively.

The model derived from the HQSAR method showed a low predictability of 0.178 for the test set. In this study, 3D prediction models showed better predictive power than 2D models for the test set. This might be because 3D information is more important in the case of datasets containing compounds with similar skeletons. Superimposition of CoMFA contour maps in the active site of the PDF crystal structure showed a meaningful correlation between receptor–ligand binding and biological activity. The final QSAR models, along with information gathered from 3D contour and 2D contribution maps, could be useful for the design of novel active inhibitors of PDF.

Keywords PDF · QSAR · CoMFA · CoMSIA · HQSAR

J. Y. Lee · M. R. Doddareddy · Y. S. Cho · H. Choo ·
A. N. Pae (✉)
Biochemicals Research Center,
Korea Institute of Science and Technology,
P.O. Box 131, Cheongryang,
Seoul 130-650, South Korea
e-mail: anpae@kist.re.kr

J. Y. Lee · H. Y. Koh
Department of Chemistry, Inha University,
253 Yonghyungdong Namgu,
Inchon 402-751, South Korea

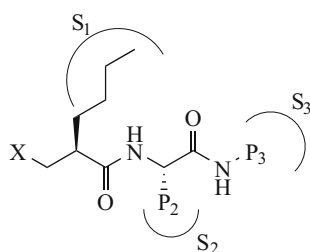
J.-H. Kang
Research Laboratories, ILDONG Pharmaceutical Co.,
260-5, Eonnam-Dong, Giheung-Gu, Yongin,
Gyeonggi-Do 449-915, South Korea

K. T. No
Department of Biotechnology, Yonsei University,
220, Sinchon-dong, Seodaemun-gu,
Seoul 120-749, South Korea

Introduction

The appearance of bacterial resistance against numerous antibiotics continues to rise rapidly, and much research effort is now focused on overcoming this phenomenon [1]. Protein synthesis is a proven rich source of targets for antibacterial drugs [2], and most known antibiotics (e.g. aminoglycosides, macrolides, tetracyclines and oxazolidinones) work by inhibiting one or more steps of this complex process. Protein synthesis in mammalian and bacterial cells always begins with the amino acid methionine. The initiating amino acid is subsequently removed by the enzyme methionine amino peptidase. Although the overall process is similar in both mammalian and bacterial cells, a significant difference is that, in bacteria, formylation first takes place and the initiating methionine is subsequently deformylated. If the formyl group is not removed by deformylase, protein synthesis is inhibited and the bacterium cannot grow [3–5]. Peptide deformylase (PDF)—a

Fig. 1 General structure of peptide deformylase (PDF) inhibitors. *X* Pharmacophore element capable of chelating metal ions. *X* attaches to a 2-substituted hexanoyl, which mimics the transition state of the hydrolysis of formyl-methionine



metalloenzyme containing Fe(II) as the catalytic metal ion in nature—catalyzes the elimination of the *N*-formyl group from the terminal methionine residue [3, 4, 6–9]. This essential role of PDF in bacterial protein synthesis provides a rational basis for selectivity, making it an attractive target for drug discovery [10, 11].

Most PDF inhibitors with sufficient potency and antibacterial activity share the common structure [2] shown in Fig. 1. In this structure, *X* represents a metal chelating group, generally thiol [12–14], hydroxamate [15–18] or reverse hydroxamate (*N*-formylhydroxylamine) [19, 20]. The *P*₁' mimicking the methionine side chain fits into a deep hydrophobic pocket, *S*₁', which lies adjacent to the metal binding site. The *P*₂' group makes hydrophobic contacts with a shallow pocket near the active site and the substituent of *P*₃' is exposed to the solvent. The *P*₂' and *P*₃' regions of the scaffold can provide additional binding energy, selectivity, and favorable pharmacokinetic properties.

The naturally occurring compound actinonin, which was first isolated from an actinomycete in 1962 [21], has a hydroxamate group in the metal binding region and has potent activity against bacteria *in vitro*, but is inactive *in vivo* [22, 23] (Fig. 2a). BB-3497, developed by British Biotech (British Biotech Pharmaceuticals, Oxford, UK), exhibits potent antibacterial activity both *in vitro* and *in vivo* [19], is a representative PDF inhibitor (Fig. 2b). BB-83698, one of the various analogues synthesized from BB-3497, is currently in phase I clinical trials [24, 25] (Fig. 2c). In this study, we have generated QSAR models utilizing derivatives of BB-3497 for the design and optimization of novel active PDF inhibitors.

Quantitative structure–activity relationship (QSAR) studies are used to quantitatively correlate biological activity to the structure of the compounds under investigation; the models thus generated can be used to predict the activities of structurally related compounds. Among the various QSAR methods, comparative molecular field analysis (CoMFA) [26] and comparative molecular similarity indices analysis (CoMSIA) [27] are three-dimensional (3D) QSAR methodologies applied widely in drug discovery processes. The CoMFA method generates models by calculating energies of steric and electrostatic interactions between the compound and the probe atoms located at the intersections of a 3D lattice according to

Lennard-Jones 6–12 and Coulomb potentials. CoMSIA is calculated using Gaussian-function similarity indices that represent steric, electrostatic, hydrophobic, hydrogen donor and acceptor interactions, and correlating these fields with experimental activity [28]. In these 3D-QSAR methodologies, alignment of compounds on an appropriate conformation is very important to get good results. Hologram quantitative structure–activity relationships (HQSAR) [29] are developed from a unity hashed fingerprint concept [30]. In HQSAR, each molecule in the dataset is divided into structural fragments, which are arranged to form a molecular hologram. HQASR encodes all possible molecular fragments (linear, branched, and overlapping). These three methods were analyzed by partial least squares (PLS) [31, 32] to derive a linear relationship between activity and the structural information available in the dataset; cross-validation [33–35] was used to check the consistency and predictive capability of the generated models.

In the present study, we have focused on the development and analysis of the various QSAR models using CoMFA, CoMSIA, and HQSAR to gain structural insight informing the design and optimization of novel and potent PDF inhibitors.

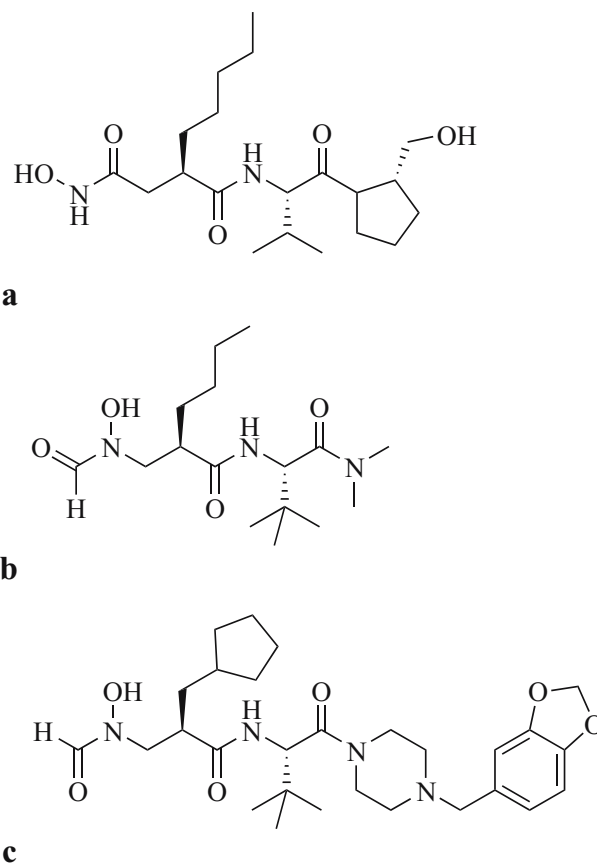


Fig. 2a–c Structures of potent PDF inhibitors. **a** Actinonin, **b** BB-3497, **c** BB-83698

Materials and methods

Dataset

A dataset reported by British Biotech [36–38] consisting of 78 reverse hydroxamate derivatives from four different scaffolds was used for the analysis. The structures of the compounds and their IC₅₀ values (nM) against *Escherichia coli* PDF are listed in Tables 1 and 2. The structure optimization of BB3497 extracted from the crystal structure (PDB ID: 1G27) was performed using standard Tripos force field [39] and Gasteiger-Hückel [40] charge with an energy gradient convergence criterion of 0.001 kcal mol⁻¹ and a distance-dependent dielectric constant. The initial structures of other compounds were constructed based on the conformation of BB3497 followed by geometrical optimization using the above method. For those molecules having rotatable bonds,

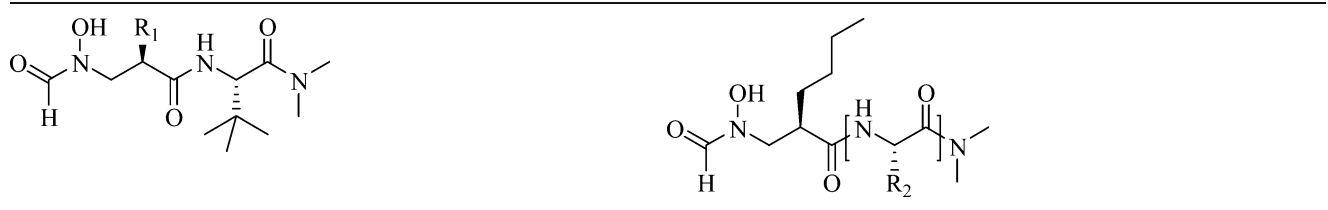
systematic conformational searches were carried out to find the lowest energy structures.

In this dataset, 58 analogues were selected randomly as a training set to generate QSAR models and the remaining compounds were used a test set for validation. BB022 was not included in the analysis because of having uncertain activity. The IC₅₀ values were converted to the corresponding pIC₅₀ (-logIC₅₀) values and used as dependent variables in the QSAR investigations. All calculations were performed on a Silicon Graphics, Origin 300 R14000 processor using SYBYL 7.0 (Tripos, St. Louis, MO).

Alignment

Molecular alignment plays an important role in CoMFA and CoMSIA studies, and it is preferable to choose an alignment that maintains bioactive conformation [41]. The conformation of BB3497 extracted from crystal structure

Table 1 Structures and biological activities of compounds active against *Escherichia coli* peptide deformylase (PDF). *QSAR* Quantitative structure–activity relationship

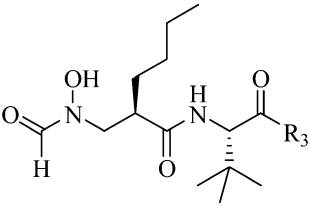
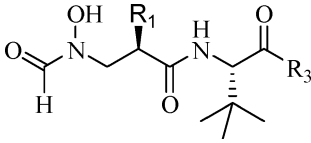


No.	R ₁	IC ₅₀ (nM)	pIC ₅₀	No.	P ₂ amino acid NH-R ₂ -CO	IC ₅₀ (nM)	pIC ₅₀
BB3497	<i>n</i> -Bu	7	8.15	BB023	Gly	100	7.00
BB001 ^a	Me	100	7.00	BB024 ^a	Ala	40	7.40
BB002	Et	70	7.15	BB025	Val	20	7.70
BB003	<i>n</i> -Pr	50	7.30	BB026	Leu	20	7.70
BB004	(S) <i>n</i> -Bu	70	7.15	BB027	Cha	30	7.52
BB005	<i>n</i> -Pentyl	10	8.00	BB028	Ile	20	7.70
BB006	<i>n</i> -Hexyl	30	7.52	BB029	(R) <i>t</i> -Leu	40	7.40
BB007	<i>n</i> -Heptyl	40	7.40	BB030	Pen(SMe)	6	8.22
BB008	<i>n</i> -Octyl	50	7.30	BB031	Cys(Bn)	70	7.15
BB009 ^a	<i>i</i> -Pr	50	7.30	BB032	Ser	50	7.30
BB010	<i>i</i> -Bu	10	8.00	BB033 ^a	Val(β -OH)	20	7.70
BB011 ^a	<i>i</i> -Pentyl	20	7.70	BB034	Val(β -OMe)	20	7.70
BB012	<i>c</i> -Pentylmethyl	8	8.10	BB035	Asp(β -Bn)	100	7.00
BB013	<i>c</i> -Pentyl	20	7.70	BB036 ^a	Glu(β -Bn)	20	7.70
BB014	<i>c</i> -Hexylmethyl	6	8.22	BB037	Lys	200	6.70
BB015	Allyl	30	7.52	BB038	Lys(ϵ -NMe ₂)	300	6.52
BB016	But-3enyl	20	7.70	BB039	Arg	20	7.70
BB017 ^a	But-2-ynyl	30	7.52	BB040 ^a	Phe	10	8.00
BB018	EtSCH ₂	20	7.70	BB041	Phe(4-Cl)	50	7.30
BB019 ^a	Bn	20	7.70	BB042 ^a	Tyr	300	6.52
BB020	Ph(4-Cl)	500	6.30	BB043	L-Tic	80	7.10
BB021	(4-MeO)PhCH ₂	200	6.70	BB044	Pro	400	6.40
BB022 ^b	1-Piperidylmethyl	>1,000					

^a Compounds used as test set for model validation

^b Compound that was not included in the construction of the 3D-QSAR models

Table 2 Structures and biological activities of compounds active against *E. coli* PDF

								
No.	R ₃	IC ₅₀	pIC ₅₀	No.	R ₁	R ₃	IC ₅₀	pIC ₅₀
BB045	OMe	30	7.52	BB062	<i>n</i> -butyl	2-Benzofuryl	4	8.40
BB046 ^a	OH	80	7.10	BB063	<i>c</i> -pentyl-Me	2-Benzofuryl	4	8.40
BB047	Me	20	7.70	BB064 ^a	<i>n</i> -butyl	4-(OH)Ph	8	8.10
BB048 ^a	Pyrrolidin-1-yl	10	8.00	BB065	<i>c</i> -pentyl-Me	4-(OH)Ph	2	8.70
BB049	Morpholin-4-yl	20	7.70	BB066	<i>n</i> -butyl	4-(MeO)Ph	1	9.00
BB050	4-Me-piperazin-1-yl	50	7.30	BB067	<i>c</i> -pentyl-Me	4-(MeO)Ph	3	8.52
BB051	4-Me-piperidin-1-yl	10	8.00	BB068	<i>n</i> -butyl	4-(NH ₂)Ph	8	8.10
BB052	4-Ac-piperidin-1-yl	10	8.00	BB069	<i>n</i> -butyl	4-(NHCOMe)Ph	8	8.10
BB053	4-EtO ₂ C-piperidin-1-yl	10	8.00	BB070 ^a	<i>n</i> -butyl	4-(NHCOCF ₃)Ph	20	7.70
BB054	4-Bn-piperidin-1-yl	20	7.70	BB071	<i>n</i> -butyl	4-(NHSO ₂ Me)Ph	4	8.40
BB055 ^a	N(Me) <i>c</i> -Hexyl	30	7.52	BB072	<i>n</i> -butyl	4-(morpholino)Ph	5	8.30
BB056	Decahydroquinolin-1-yl	40	7.40	BB073 ^a	<i>c</i> -pentyl-Me	4-(morpholino)Ph	6	8.22
BB057	Tetrahydroquinolin-1-yl	9	8.05	BB074	<i>n</i> -butyl	4-(F)Ph	3	8.52
BB058	N(Bn)CH ₂ CH ₂ Ph	90	7.05	BB075 ^a	<i>c</i> -pentyl-Me	4-(F)Ph	3	8.52
BB059 ^a	Ph	3	8.52	BB076	<i>n</i> -butyl	4-(CN)Ph	7	8.15
BB060	2-Pyridyl	3	8.52	BB077	<i>c</i> -pentyl-Me	4-(SO ₂ Me)Ph	10	8.00
BB061 ^a	2-Furyl	5	8.30					

^a Compounds used as test set for model validation

(1G27) was used as a template, and all compounds in the dataset were aligned on this conformation to obtain a consistent alignment. The molecular alignment was carried out using the "Atom Fit" method in SYBYL. The alignment of structures is shown in Fig. 3. The superimposition of all analogues showed a reasonable fit to the binding pockets.

Comparative molecular field analysis

The steric and electrostatic potential fields for CoMFA were calculated at each lattice intersection of a regularly spaced grid with a variation of 0.5 Å to 3.0 Å. The lattice was defined automatically, and is extended 4 Å units past Vanderwaals volume of all molecules in the x, y, and z directions. The Vanderwaals potential (Lennard-Jones 6–12) and columbic term, which represent steric and electrostatic fields, respectively, were calculated using Tripos force field [39]. A distance-dependent dielectric expression $\epsilon = \epsilon_0 \cdot R_{ij}$ with $\epsilon = 1.0$ was used. An sp³ carbon atom with Van der Waals radius of 1.52 Å and +1.0 charge served as the probe atom to calculate steric and electrostatic fields. The effect of probe atoms was studied using three different probes: a positively charged

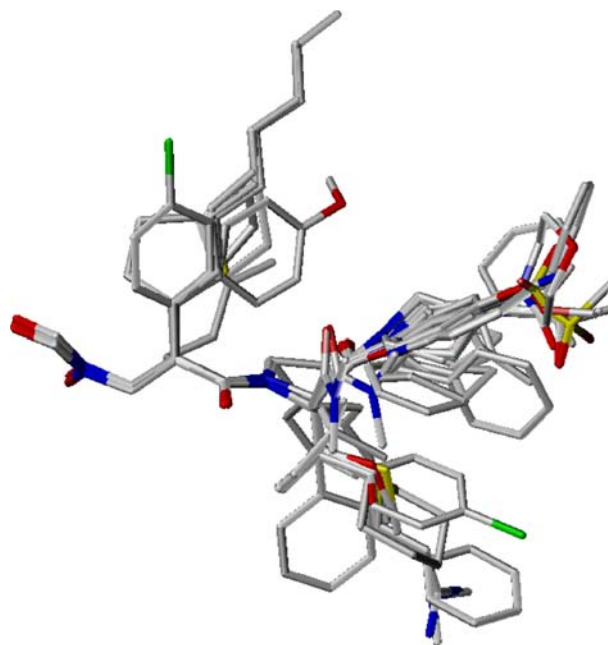


Fig. 3 Alignment of 77 PDF inhibitors used in the quantitative structure–activity relationship (QSAR) studies

sp^3 carbon atom, a negatively charged sp^3 oxygen atom, and a proton. The steric and electrostatic contributions were truncated at ± 30 kcal mol $^{-1}$, and electrostatic contributions were ignored at lattice intersections with maximum steric interactions. To investigate differences in charge, we applied partial charges using the Gasteiger-Hückel method, [42] the Gasteiger-Marsili method, [43] and the AM1 molecular electrostatic potential (AM1-ESP) fitted partial charge method using the ESP routine in MOPAC 6.0 [44].

Comparative molecular similarity indices analysis

CoMSIA calculates similarity indices at the intersections of a surrounding lattice. The charge, probe, and grid spacing used to construct the CoMFA best model were also used for the CoMSIA investigation. Five physicochemical properties of steric, electrostatic, hydrophobic, hydrogen bond donor, and hydrogen bond acceptor fields were calculated. The steric contribution was reflected by the third power of the atomic radii of the atoms. The electrostatic properties were introduced as Gasteiger-Hückel partial charges, and atom-based hydrophobicity was assigned according to the parameterization developed by Ghose and Viswanadhan [45, 46]. The lattice dimensions were selected with a sufficiently large margin (4 Å) to enclose all aligned molecules. Any singularities were avoided at atomic positions in CoMSIA fields because a Gaussian type distance dependence of the physicochemical properties was adopted, thus no arbitrary cutoffs were required. CoMSIA similarity indices (A_{F,K^q}) for a molecule j with atoms i at a grid point q are calculated by Eq. 1:

$$A_{F,K^q}(j) = -\sum \omega_{probe,k} \omega_{ik} e^{-\alpha r_{iq}^2} \quad (1)$$

where ω_{ik} is the actual value of the physicochemical property k of atom i , $\omega_{probe,k}$ is the value of the probe atom, and r_{iq} is the mutual distance between the probe atom at grid point q and atom i of the molecule. In the present study, similarity indices were computed using a probe atom with a charge of +1, a radius of 1 Å, a hydrophobicity of +1, and the attenuation factor (α) of 0.3 for the Gaussian type distance.

HQSAR analyses

HQSAR is a technique based on the concept of using the properties of a molecule expressed in holograms as descriptors in QSAR models. With the transformation of the chemical representation of a molecule into its corresponding molecular hologram, this method requires no explicit 3D information for the ligands [29]. The construction of a molecular hologram containing the

HQSAR descriptors was completed as described below. First, the molecule was hashed to a molecular fingerprint that encoded the frequency of occurrence of various molecular fragment types using a predefined set of rules. The molecular fingerprint was then cut into strings at fixed intervals as specified by a hologram length (HL) parameter. Then all of the generated strings were hashed into a fixed length array. The SYBYL line notation (SLN) for each string was mapped to a unique integer in the range of 0–2 [31] using a cyclic redundancy check (CRC) algorithm [47]. The numerical representation of molecules was exploited by a subsequent correlation analysis; typically, a PLS QSAR model was constructed. The HQSAR models were constructed by screening the 12 default HL values ranging from 53 to 401 (53, 59, 61, 71, 83, 97, 151, 199, 257, 307, 353, and 401). To improve the statistical analysis, various fragment types and lengths were also investigated.

PLS analysis

The conventional CoMFA, CoMSIA, and HQSAR descriptors derived above were used as explanatory variables, and pIC_{50} values were used as the target variable in PLS regression analyses to derive QSAR models using the standard implementation in the SYBYL package. The predictive value of the models was evaluated by leave-one-out (LOO) [33, 35] cross-validation method. All cross-validated PLS analyses were performed without using the column filtering option. The cross-validated coefficient q^2 was calculated using Eq. 2:

$$q^2 = 1 - \frac{\sum (Y_{predicted} - Y_{observed})^2}{\sum (Y_{observed} - Y_{mean})^2} \quad (2)$$

Where $Y_{predicted}$, $Y_{observed}$ and Y_{mean} are predicted, actual, and mean values of the target property (pIC_{50}), respectively. $\sum (Y_{predicted} - Y_{observed})^2$ is the predictive sum of squares (PRESS). To maintain the optimum number of PLS components and minimize the tendency to over fit the data, the number of components corresponding to the lowest PRESS value was used to derive the final PLS regression models. In addition to the q^2 and number of components, the conventional correlation coefficient r^2 and its standard errors (SEE) were also computed. The number of components giving the optimal number of components (ONC) was used to generate the final PLS regression models. The conventional correlation coefficient r^2 and its standard error were subsequently computed for the final PLS models. CoMFA and CoMSIA coefficient maps were generated by interpolation of the pairwise products between the PLS coefficients and the standard deviations of the corresponding CoMFA or CoMSIA descriptor values. A

Table 3 Statistical results of comparative molecular field analysis (CoMFA) analyses using various charges and probe atoms. *GH* Gasteiger-Hückel, *GM* Gasteiger-Marsili, *AM1* Austin Model 1, *ONC*Optimal number of components, *SDEP* Standard deviation of error of predictions, *SEE* Standard deviation of a non-cross-validated analysis

Charge	Probe ^a	ONC	q^2	<i>SDEP</i>	r^2	<i>SEE</i>	F^b -ratio
GH	C.3	6	0.569	0.406	0.957	0.128	189.730
	O.3	3	0.509	0.421	0.820	0.255	82.098
	H	7	0.517	0.434	0.954	0.134	148.171
GM	C.3	3	0.526	0.414	0.879	0.209	130.799
	O.3	3	0.512	0.420	0.827	0.250	86.153
	H	7	0.507	0.439	0.938	0.156	107.954
AM1	C.3	3	0.535	0.410	0.860	0.225	110.529
	O.3	3	0.533	0.411	0.872	0.215	122.636
	H	3	0.539	0.408	0.872	0.215	122.737

^a C.3=sp³ carbon with +1 charge; O.3=sp³ oxygen with -1 charge; H=proton

cross-validated r^2 (q^2), obtained as a result of this analysis, served as a quantitative measure of the predictive ability of the final QSAR models. The q^2 value is a statistical indication of how well a model can predict the activity of members left out of the model formation. In contrast, the conventional r^2 is simply a reflection of how well the fit equation reproduces input values [48]. The selection of the best QSAR model was chosen on the basis of a combination of q^2 , r^2 , predictive r^2 , and *SED*, *SEE*. The 10-fold cross-validation for each best model was applied. The samples were randomly split into ten groups. Nine groups of data were used to build the calibration model, and the activity of the molecules in the remaining group was predicted by the resulting model. The procedure was repeated 10 times until all groups had been held out from the data and the activity of all samples had been predicted. The cross-validation was repeated 20 times and the statistical results were averaged.

Predictive r squared

To validate the derived QSAR models, the biological activities of the external test set were predicted using models derived from the training set. The predictive ability of the models is expressed by the predictive r^2 value

(r^2_{pred}), which is analogous to cross-validated r^2 (q^2) and is calculated using Eq. 3:

$$r^2_{pred} = \frac{SD - PRESS}{SD} \quad (3)$$

Where SD is the sum of squared deviation between the biological activities of the test set molecule and the mean activity of the training set molecules, and PRESS is the sum of squared deviations between the observed and the predicted activities of the test molecules.

Results and discussion

CoMFA models

CoMFA analyses were performed using aligned structures of the entire training set with pIC₅₀ values ranging from 6.30 to 9.00. The 3D QSAR models were generated by correlation of the biological activities of the 58 selected inhibitors with variations in their CoMFA fields using the PLS method. Several CoMFA models were obtained with various different charges, probe atoms, and grid spacing values. The statistical results of these models are presented in Tables 3 and 4.

Table 4 Statistical results of CoMFA analyses for various grid spacing using Gasteiger-Hückel charge and C.3 probe

Grid (Å)	ONC	q^2	<i>SDEP</i>	r^2	<i>SEE</i>	F -ratio
0.5	3	0.391	0.469	0.735	0.309	49.940
1.0	7	0.524	0.431	0.967	0.113	212.431
1.5	6	0.503	0.436	0.955	0.131	179.014
2.0	6	0.569	0.406	0.957	0.128	189.730
2.5	1	0.382	0.464	0.507	0.414	57.702
3.0	4	0.543	0.410	0.899	0.193	117.443

Table 5 Results of comparative molecular similarity indices analysis (CoMSIA) analyses for various fields

Model	Field(s)	ONC	q^2	<i>SDEP</i>	r^2	<i>SEE</i>	<i>F</i> -ratio
1	S	4	0.448	0.451	0.766	0.293	43.483
2	E	8	0.407	0.486	0.909	0.190	61.302
3	H	8	0.504	0.445	0.943	0.150	101.918
4	D	9	0.297	0.535	0.575	0.416	7.207
5	A	3	0.119	0.564	0.454	0.444	14.949
6	S+E	9	0.470	0.464	0.946	0.148	93.490
7	S+H	6	0.520	0.429	0.924	0.170	103.800
8	E+H	7	0.455	0.461	0.950	0.140	134.928
9	S+E+H	7	0.503	0.441	0.952	0.137	141.339
10	S+E+D	1	0.404	0.456	0.540	0.400	65.701
11	S+E+A	1	0.385	0.630	0.425	0.432	55.500
12	S+E+D+A	1	0.389	0.461	0.419	0.435	54.151
13	S+E+H+D	1	0.460	0.434	0.471	0.415	66.826
14	S+E+H+A	8	0.457	0.465	0.960	0.127	146.162
15	E+H+D+A	1	0.444	0.440	0.576	0.385	75.978
16	S+E+H+D+A	1	0.453	0.437	0.581	0.382	77.769

The statistical validity of the models can be judged by high q^2 and r^2 values along with a low standard error of estimate. Among the nine generated models, the model with highest q^2 value was obtained by using Gasteiger-Hückel partial charges and a positively charged sp^3 carbon atom. This model was used for further analysis.

The effect of different grid spacing values from 0.5 to 3 Å are shown in Table 4. The optimum result was obtained in 2 Å grid spacing, which showed a cross-validated correlation coefficient q^2 of 0.569 and *SDEP* of 0.406. The non-cross-validated PLS analysis gave a conventional r^2 of 0.957. The *F* value and *SEE* are 189.570 and 0.128,

Table 6 Results of hologram quantitative structure–activity relationships (HQ SAR) analyses for various fragment distinctions on the key statistical parameters using fragment size default (4–7). *HL* Hologram length

Model	Fragment distinction ^a	ONC	q^2	<i>SDEP</i>	r^2	<i>SEE</i>	HL
1	A/B	3	0.450	0.446	0.639	0.361	61
2	A/B/C	2	0.511	0.417	0.633	0.361	71
3	A/B/H	6	0.523	0.427	0.762	0.302	83
4	A/B/Ch	2	0.429	0.450	0.604	0.375	199
5	A/B/DA	3	0.437	0.451	0.632	0.365	61
6	A/C/Ch	2	0.478	0.430	0.592	0.381	53
7	A/C/DA	6	0.482	0.445	0.838	0.249	401
8	A/Ch/DA	3	0.378	0.474	0.546	0.405	353
9	A/B/C/H	6	0.560	0.414	0.821	0.262	71
10	A/B/C/Ch	2	0.475	0.432	0.598	0.378	71
11	A/B/C/DA	5	0.522	0.424	0.796	0.277	59
12	A/B/H/Ch	6	0.488	0.442	0.762	0.302	83
13	A/B/H/DA	6	0.479	0.447	0.787	0.285	151
14	A/B/Ch/DA	6	0.475	0.448	0.817	0.265	61
15	A/C/H/Ch	6	0.536	0.421	0.821	0.261	199
16	A/C/H/DA	6	0.564	0.409	0.844	0.245	257
17	A/C/Ch/DA	5	0.490	0.437	0.807	0.269	71
18	A/B/C/H/Ch	6	0.512	0.432	0.831	0.254	353
19	A/B/C/H/DA	6	0.454	0.457	0.834	0.252	151
20	A/B/C/Ch/DA	3	0.471	0.437	0.633	0.364	61
21	A/B/H/Ch/DA	6	0.490	0.442	0.843	0.245	257
22	A/C/H/Ch/DA	6	0.514	0.431	0.807	0.272	53
23	A/B/C/H/Ch/DA	6	0.454	0.457	0.836	0.251	307

^a A Atoms, B bonds, C connections, H hydrogen atoms, C chirality, DA donor and acceptor

respectively. These values indicate a good statistical correlation and reasonable predictability of the CoMFA model.

CoMSIA models

CoMSIA analyses were performed using combinations of five steric, electrostatic, hydrophobic, hydrogen bond donor and acceptor fields. The same conditions used for the best CoMFA model were also used for CoMSIA analyses. The CoMSIA results are summarized in Table 5. The results show that the factor contributing most to the model is the hydrophobic field, indicating the importance of the lipophilicity for the present series of molecules. By contrast, hydrogen donor and acceptor fields showed least contribution due to the fact that all the compounds of the dataset have almost same number of hydrogen bond donors and acceptors. The best CoMSIA model was generated using steric and hydrophobic fields. A cross-validated q^2 value of 0.520 with six as optimum number of components and a

Table 8 Partial least squares (PLS) statistics of CoMFA, CoMSIA, and HQSAR models

	CoMFA	CoMSIA	HQSAR
q^2	0.569	0.520	0.578
<i>SDEP</i>	0.406	0.429	0.402
Optimum components	6	6	6
$r^2_{10\text{-fold}}^a$	0.560	0.512	0.564
$SEE_{10\text{-fold}}^b$	0.412	0.432	0.408
r^2	0.957	0.924	0.860
SEE^c	0.128	0.170	0.232
<i>F</i> -ratio	189.730	103.800	199(HL)
Contribution			
Steric	0.608	0.397	
Electrostatic	0.392		
Hydrophobic		0.603	
r^2_{pred}	0.687	0.505	0.178

^a Average of 10-fold cross-validation results

^b Standard deviation of error of 10-fold cross-validation

^c Standard deviation of a non-cross-validated analysis

Table 7 HQSAR analysis of the influence of various fragment sizes on the key statistical parameters using the same fragment distinction as in Table 6

Fragment distinction	Fragment size	ONC	q^2	<i>SDEP</i>	r^2	<i>SEE</i>	HL
A/B/C/H	2–5	4	0.496	0.431	0.656	0.356	71
	3–6	6	0.563	0.409	0.771	0.296	53
	4–7	6	0.560	0.414	0.821	0.262	71
	5–8	6	0.501	0.437	0.805	0.273	307
	6–9	6	0.527	0.425	0.840	0.248	199
A/C/H/Ch	7–10	6	0.493	0.440	0.845	0.244	401
	2–5	4	0.448	0.451	0.632	0.368	257
	3–6	5	0.464	0.449	0.731	0.318	97
	4–7	6	0.536	0.421	0.821	0.261	199
	5–8	6	0.543	0.418	0.819	0.263	71
A/C/H/DA	6–9	6	0.517	0.430	0.823	0.260	71
	7–10	6	0.563	0.409	0.862	0.230	353
	2–5	4	0.467	0.443	0.674	0.347	59
	3–6	6	0.465	0.453	0.801	0.276	307
	4–7	6	0.564	0.409	0.84	0.245	257
A/B/C/H/Ch	5–8	6	0.483	0.445	0.821	0.262	401
	6–9	3	0.451	0.446	0.644	0.359	59
	7–10	6	0.501	0.437	0.807	0.272	97
	2–5	4	0.483	0.437	0.647	0.361	71
	3–6	6	0.558	0.411	0.777	0.292	53
A/C/H/Ch/DA	4–7	6	0.512	0.432	0.831	0.254	353
	5–8	6	0.527	0.426	0.831	0.255	307
	6–9	6	0.555	0.413	0.842	0.246	257
	7–10	6	0.578	0.402	0.860	0.232	199
	2–5	4	0.477	0.439	0.693	0.336	59
A/C/H/Ch/DA	3–6	2	0.425	0.452	0.549	0.400	53
	4–7	6	0.514	0.431	0.807	0.272	53
	5–8	6	0.463	0.454	0.835	0.251	401
	6–9	6	0.485	0.444	0.830	0.255	307
	7–10	3	0.472	0.437	0.617	0.381	71

Table 9 Experimental activities, predicted activities and residual values of molecules used in the training set for the CoMFA, CoMSIA, and HQSAR models

No.	pIC ₅₀	CoMFA		CoMSIA		HQSAR	
		Pred. ^a	Res. ^b	Pred.	Res.	Pred.	Res.
BB3497	8.15	7.96	0.19	7.94	0.21	7.75	0.40
BB002	7.15	7.39	-0.24	7.36	-0.21	7.47	-0.32
BB003	7.30	7.55	-0.25	7.62	-0.32	7.62	-0.32
BB004	7.15	7.20	-0.05	7.15	0.01	7.27	-0.12
BB005	8.00	7.87	0.13	7.93	0.07	7.62	0.38
BB006	7.52	7.57	-0.05	7.71	-0.19	7.61	-0.09
BB007	7.40	7.35	0.05	7.47	-0.07	7.55	-0.15
BB008	7.30	7.21	0.09	7.32	-0.02	7.49	-0.19
BB010	8.00	7.84	0.16	7.83	0.17	8.09	-0.09
BB012	8.10	8.06	0.04	8.11	-0.01	7.84	0.26
BB013	7.70	7.71	-0.01	7.74	-0.04	7.83	-0.13
BB014	8.22	8.14	0.08	8.05	0.17	8.08	0.14
BB015	7.52	7.21	0.31	7.34	0.18	7.25	0.27
BB016	7.70	7.64	0.06	7.68	0.03	7.39	0.31
BB018	7.70	7.89	-0.19	7.81	-0.11	7.23	0.47
BB020	6.30	6.31	-0.01	6.27	0.04	6.49	-0.19
BB021	6.70	6.67	0.03	6.70	0.00	6.93	-0.23
BB023	7.00	6.97	0.03	6.92	0.08	7.17	-0.17
BB025	7.70	7.66	0.04	7.69	0.01	7.67	0.03
BB026	7.70	7.75	-0.05	7.75	-0.04	7.61	0.09
BB027	7.52	7.56	-0.04	7.52	0.00	7.58	-0.06
BB028	7.70	7.78	-0.08	7.92	-0.22	7.71	-0.01
BB029	7.40	7.38	0.02	7.38	0.02	7.58	-0.18
BB030	8.22	8.05	0.17	8.05	0.17	8.41	-0.19
BB031	7.15	7.13	0.03	7.03	0.12	7.23	-0.07
BB032	7.30	7.32	-0.02	7.23	0.07	7.26	0.04
BB034	7.70	7.94	-0.24	7.92	-0.22	7.88	-0.18
BB035	7.00	6.93	0.07	6.98	0.02	7.10	-0.10
BB037	6.70	6.72	-0.02	6.68	0.03	6.99	-0.29
BB038	6.52	6.54	-0.02	6.55	-0.03	6.41	0.11
BB039	7.70	7.80	-0.10	7.48	0.22	7.18	0.53
BB041	7.30	7.43	-0.13	7.22	0.08	7.11	0.19
BB043	7.10	7.11	-0.01	7.19	-0.09	6.87	0.23
BB044	6.40	6.34	0.06	6.41	-0.01	6.49	-0.09
BB045	7.52	7.62	-0.10	7.79	-0.27	7.71	-0.19
BB047	7.70	7.75	-0.05	7.74	-0.04	7.80	-0.10
BB049	7.70	7.58	0.12	7.56	0.14	7.80	-0.10
BB050	7.30	7.45	-0.15	7.46	-0.16	7.21	0.09
BB051	8.00	7.97	0.03	7.88	0.12	7.99	0.01
BB052	8.00	7.97	0.03	8.09	-0.09	8.11	-0.11
BB053	8.00	8.08	-0.08	8.02	-0.02	7.78	0.22
BB054	7.70	7.68	0.02	7.70	0.00	7.67	0.03
BB056	7.40	7.47	-0.06	7.48	-0.08	7.40	0.00
BB057	8.05	7.94	0.11	7.95	0.10	8.07	-0.02
BB058	7.05	7.06	-0.01	7.07	-0.02	7.28	-0.23
BB060	8.52	8.46	0.06	8.36	0.16	8.38	0.14
BB062	8.40	8.25	0.15	8.19	0.21	8.40	0.00
BB063	8.40	8.34	0.06	8.35	0.05	8.48	-0.08
BB065	8.70	8.68	0.02	8.57	0.13	8.45	0.25
BB066	9.00	8.69	0.31	8.39	0.61	8.38	0.62
BB067	8.52	8.84	-0.32	8.53	-0.01	8.47	0.05
BB068	8.10	8.20	-0.10	8.27	-0.17	8.36	-0.26
BB069	8.10	8.12	-0.02	8.21	-0.11	8.30	-0.20

Table 9 (continued)

No.	pIC ₅₀	CoMFA		CoMSIA		HQSAR	
		Pred. ^a	Res. ^b	Pred.	Res.	Pred.	Res.
BB071	8.40	8.53	-0.13	8.33	0.07	8.34	0.06
BB072	8.30	8.35	-0.05	8.23	0.07	8.39	-0.09
BB074	8.52	8.45	0.07	8.63	-0.11	8.33	0.19
BB076	8.15	8.08	0.07	8.55	-0.40	8.37	-0.22
BB077	8.00	8.01	-0.01	8.27	-0.27	8.28	-0.28

^a Predicted activity^b Residual (Experimental activity-Predicted activity)

conventional r^2 of 0.924 were obtained. The F value, $SDEP$ and SEE are 103.8, 0.429 and 0.17, respectively. The statistical data shows that a reliable CoMSIA model was successfully constructed.

HQSAR models

Using the training set, analyses were first performed using the default fragment size of 4–7 for different combinations of fragment distinction parameters: atoms (A), bonds (B), connections (C), hydrogen atoms (H), chirality (Ch), and donor and acceptor (DA). HQSAR analyses were performed by screening the 12 default series of HL values ranging from 53 to 401 bins. The patterns of fragment counts from the training set compounds were then related to the measured biological activity. The statistical results from

the PLS analyses for the training set compounds using several fragment distinction combinations are given in Table 6. Only models with $q^2 > 0.5$ and $r^2 > 8.0$ were selected for further analysis.

The influence of different fragment sizes was further investigated for the selected HQSAR models (Table 6). Fragment size parameters control the minimum and maximum length of fragments to be included in the hologram fingerprint. The HQSAR results for the different fragment sizes used are seen in Table 7. The results show that variations in fragment size did not lead to any improvement in the HQSAR models.

The best HQSAR model was derived using atoms, bonds, connections, hydrogen atoms and chirality information with a fragment size of 7–10. An $SDEP$ of 0.402 occurs at a cross-validated r^2 (q^2) of 0.578 with six optimal

Table 10 Experimental activities, predicted activities and residual values of molecules used in the test set for each best model of CoMFA, CoMSIA, and HQSAR

No.	pIC ₅₀	CoMFA		CoMSIA		HQSAR	
		Pred.	Res.	Pred.	Res.	Pred.	Res.
BB001	7.00	7.12	-0.12	7.18	-0.18	7.31	-0.31
BB009	7.30	7.57	-0.27	7.46	-0.16	7.67	-0.37
BB011	7.70	7.73	-0.03	7.97	-0.27	8.11	-0.41
BB017	7.52	7.44	0.08	7.42	0.10	7.28	0.24
BB019	7.70	7.48	0.22	7.21	0.49	6.93	0.77
BB024	7.40	7.12	0.28	7.15	0.25	7.19	0.21
BB033	7.70	7.82	-0.12	7.74	-0.04	7.73	-0.02
BB036	7.70	7.48	0.22	7.24	0.46	7.37	0.33
BB040	8.00	7.42	0.58	7.33	0.68	7.13	0.87
BB042	6.52	7.23	-0.71	7.29	-0.77	7.14	-0.62
BB046	7.10	7.46	-0.36	7.63	-0.53	7.83	-0.73
BB048	8.00	7.97	0.04	8.04	-0.04	7.85	0.15
BB055	7.52	7.78	-0.26	7.66	-0.14	8.24	-0.72
BB059	8.52	8.30	0.22	8.59	-0.07	8.30	0.22
BB061	8.30	8.35	-0.05	8.10	0.20	8.42	-0.12
BB064	8.10	7.96	0.14	8.47	-0.37	8.36	-0.26
BB070	7.70	7.98	-0.28	8.08	-0.38	8.33	-0.63
BB073	8.22	8.48	-0.26	8.40	-0.18	8.48	-0.26
BB075	8.52	8.56	-0.04	8.78	-0.26	8.42	0.10

components. The PLS analysis yields a conventional r^2 value of 0.860 and a standard error of 0.232 for all the studied compounds. The hologram that gives the lowest standard error has a length of 199.

Validation of the QSAR models

The statistical data of the best models selected by the three QSAR methods are summarized in Table 8. In order to

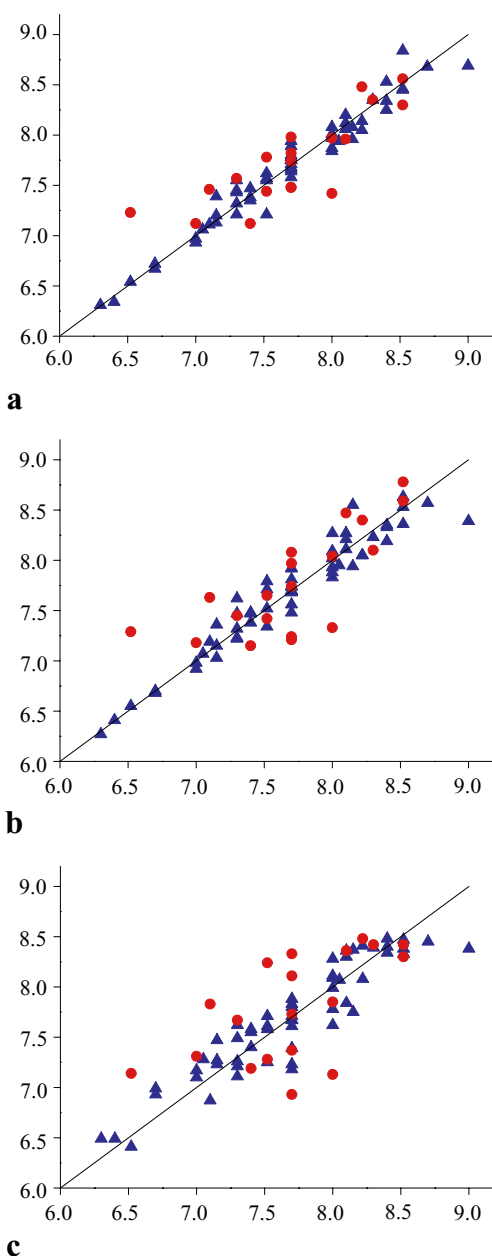


Fig. 4a–c Experimental activity (pIC_{50}) vs predicted activity values of QSAR models. Results from **a** comparative molecular field analysis (CoMFA), **b** comparative molecular similarity indices analysis (CoMSIA), and **c** hologram quantitative structure–activity relationships (HQSAR). *Blue triangles* Training set compounds, *red circles* test set compounds

study the robustness of our models, the 10-fold cross-validation option was also applied. Experimental vs predicted activities with residues of the training set are shown in Table 9.

To verify the predictive ability of the constructed 3D-QSAR models based on the training set, an external test set containing 19 compounds was used for validation. The activities of the test were predicted using the three models. The CoMFA and CoMSIA models gave satisfactory predictive r^2 values of 0.687 and 0.505, respectively, whereas the HQSAR model gave a poor predictive r^2 value of 0.178 (Table 8). The experimental versus predicted activities of the test set compounds based on the three models are listed in Table 10, and the correlation between experimental data and predicted values is presented graphically in Fig. 4.

Among the three models, the CoMFA model had the best predictive r^2 value and showed good predictions for the external test set. In contrast, HQSAR has much lower predictive power than the two models for the prediction of external molecules. It could be that including 3D information is more important in the case of datasets consisting of similar skeletons. Thus, 3D QSAR methods like CoMFA and CoMSIA are better suited to this kind of dataset than 2D QSAR methods like HQSAR. Figure 5 shows the residual values of the three models. The CoMFA and CoMSIA models yielded relatively smaller residuals than the HQSAR model. The CoMFA model has smaller residuals than the CoMSIA model, indicating that the CoMFA model is better compared to the CoMSIA model.

Of the three models tested, the CoMFA model showed the best predictability and thus was applied to predict the activities of actinonin, BB022 and BB-8369, which were not included in the model. Actinonin and BB-83698 are very active compounds, with IC_{50} of 10 nM (pIC_{50} , 8.00) against *E. coli* PDF. The activities predicted by the model were 7.50 and 7.91, respectively. The prediction for BB-83698 closely matched the experimental value. In the case of actinonin, a small difference in residual of 0.5 was observed. This may

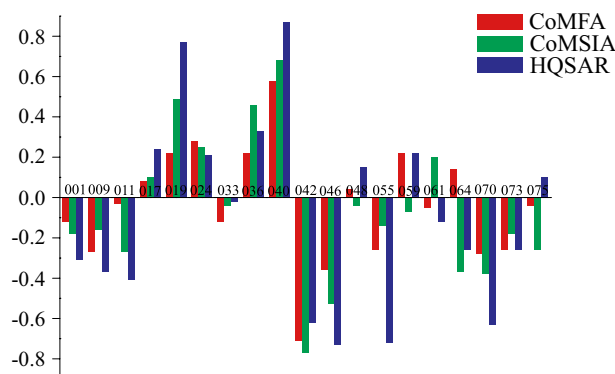


Fig. 5 The residuals between experimental activities and predicted activities of the test set in the QSAR models

be due to the fact that actinonin has a hydroxamate group as a metal binding group instead of a reverse hydroxamate moiety. In the case of BB022, which has low activity ($IC_{50} > 1,000$), there was a large difference between predicted (pIC_{50} , 7.61) and actual activities ($pIC_{50} < 6.0$); a piperidine moiety is unfavorable in the S_1' hydrophobic pocket. No compound in the training set has positive ionizable or electronegative substituents. BB022 was an outlier of the PLS analysis, and thus was poorly predicted.

Graphical interpretation of QSAR models

The steric and electrostatic contour plots of CoMFA analysis are shown in Fig. 6. The steric field (60.8 %) contributed more than the electrostatic field (39.2%). The

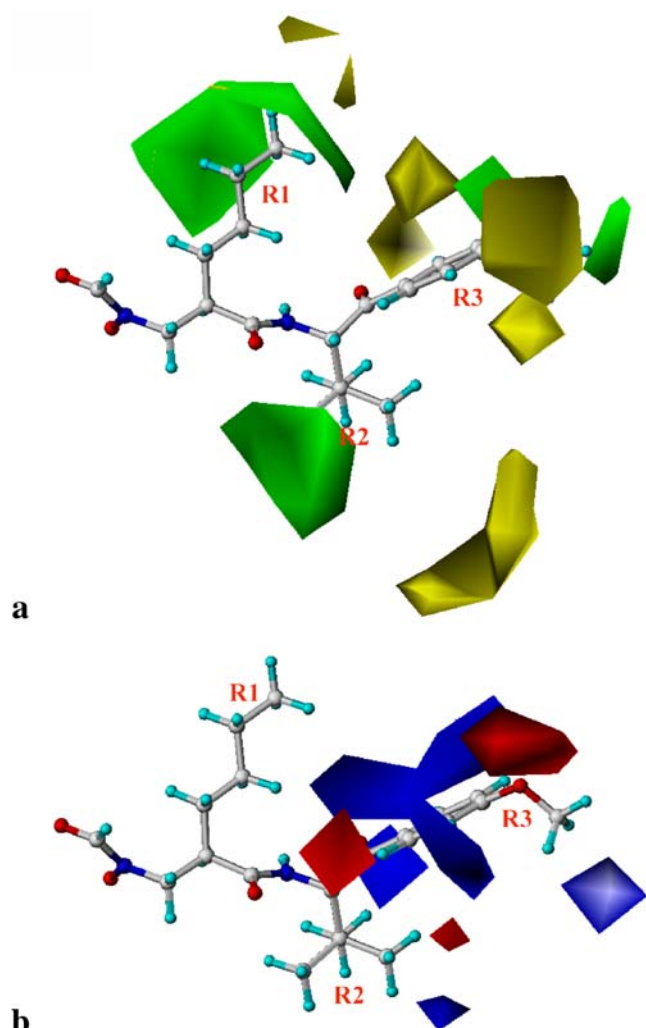


Fig. 6a,b CoMFA stdev*coeff contour plots in combination with BB066. **a** Steric field: *green contours* regions where bulky groups increase activity, *yellow contours* regions where bulky groups decrease activity. **b** Electrostatic field: *blue contours* regions where positive groups increase activity, *red contours* regions where negative charge increases activity

most active compound (BB066) is displayed in the maps. Sterically favorable and unfavorable regions are displayed as green and yellow contours, respectively. Two green and yellow isopleths around the R_1 portion of the ligand represent the fact that this region has restrictively bulky groups. This is in agreement with the fact that the S_1' region

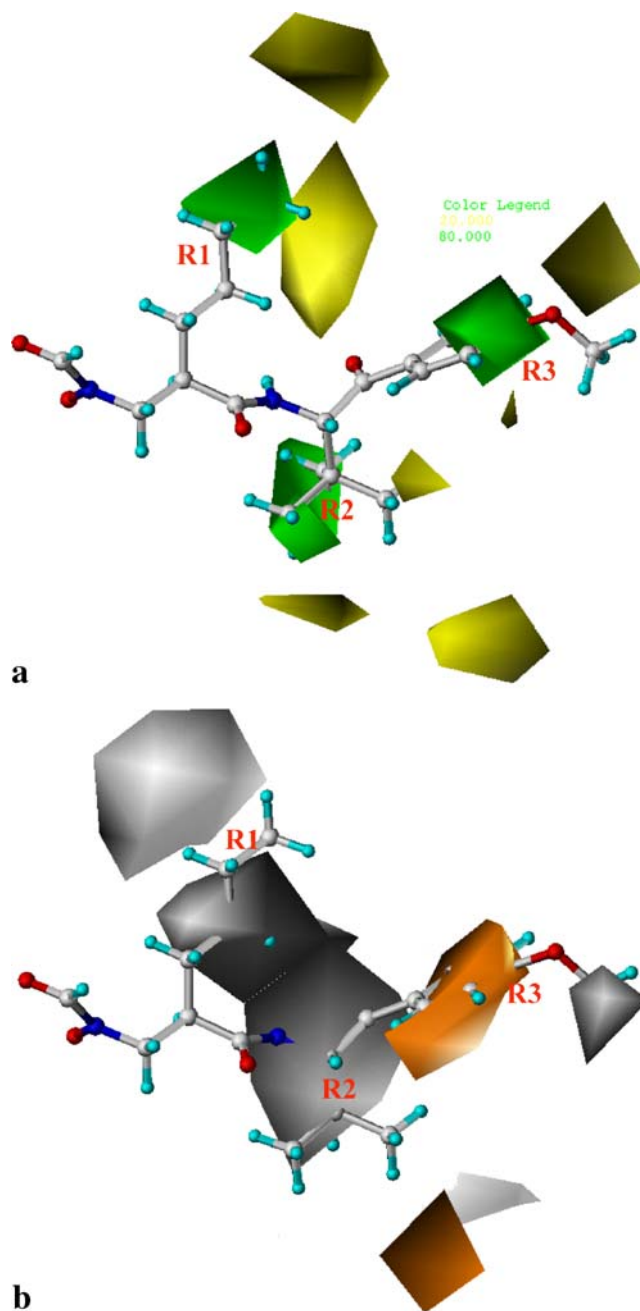


Fig. 7a,b CoMSIA stdev*coeff contour plots in combination with BB066. **a** Steric field: *green contours* regions where bulky groups increase activity, *yellow contours* regions where bulky groups decrease activity. **b** Hydrophobic field: *orange contours* regions where hydrophobic (lipophilic) groups increase activity, *white contours* regions where hydrophobic (lipophilic) groups decrease activity

is a pocket of limited size in the crystal structure. Substituents like *n*-butyl or cyclopentyl-methyl are preferred for the R_1 region. The green polyhedron on the R_2 portion indicates that *t*-butyl is preferred over other substituents with one or two carbons in that position. The yellow below the R_2 area indicates that any substituents bulkier than *t*-butyl are not favored. For example, BB037 and BB038 have lower activity than BB-3497 because they have longer R_2 side chains than BB-3497. Several green and yellow isopleths near the solvent-exposed region show that compounds with substituents like that in BB058 that extend straight out toward the solvent increase activity.

In electrostatic contour maps, the red-colored contours indicate negatively charged favored regions, whereas blue contours represent positively charged favored regions. The electrostatic contour map of the CoMFA model is shown in Fig. 6b. The red (negative charge favorable) and blue (negative charge unfavorable) contours in the CoMFA electrostatic field represent 17% and 83% level contributions, respectively. The small red and blue contours around the R_2 region might come from the sulfur of BB030 and the guanidine moiety of BB039, respectively. At the R_3 position, the large blue polyhedron around the phenyl ring indicates that the electropositive groups are favorable to the activity. The red contour appearing at the phenyl group substituent indicates that electronegative groups are favored. This can be explained by the fact that when electronegative oxygen is present (BB064–067) it increases the activity. Additionally, the blue isopleths just below the R_3 region indicate that the electropositive group near the methyl group is expected to enhance activity (BB066, BB067 and BB077).

In the CoMSIA model, the hydrophobic field descriptor explains 60.7% of the variance, while the proportion of steric descriptor accounts for 39.7%. Accordingly, in this CoMSIA model, the hydrophobic field has more influence than the steric field. The CoMSIA coefficient contour maps are depicted with BB066 in Fig. 7.

The steric contour map of the CoMSIA model is similar to the CoMFA steric map. In the case of the R_1 side chain, this model proves a more effective representation than CoMFA. Two green and yellow contours in similar positions as in the CoMFA model indicate that groups either shorter or longer than *n*-butyl are not good for activity. This is in agreement with the fact that the inhibitory activity of compounds BB001, BB002 and BB003, which have shorter substituents, and compounds like BB005, BB006, BB007, BB008, which have longer substituents, showed lower activities than BB3497. The contour maps around the R_2 and R_3 regions are also similar for both CoMFA and CoMSIA models.

Figure 7b shows a hydrophobic contour map of the CoMSIA model. The orange and white contours in the

CoMSIA hydrophobic field represent 80% and 20% level contributions, respectively. The orange contours indicate that hydrophobic groups increases activity, whereas white contours indicate regions in which hydrophobic groups decrease activity. The white contour of R_1 comes from the influence

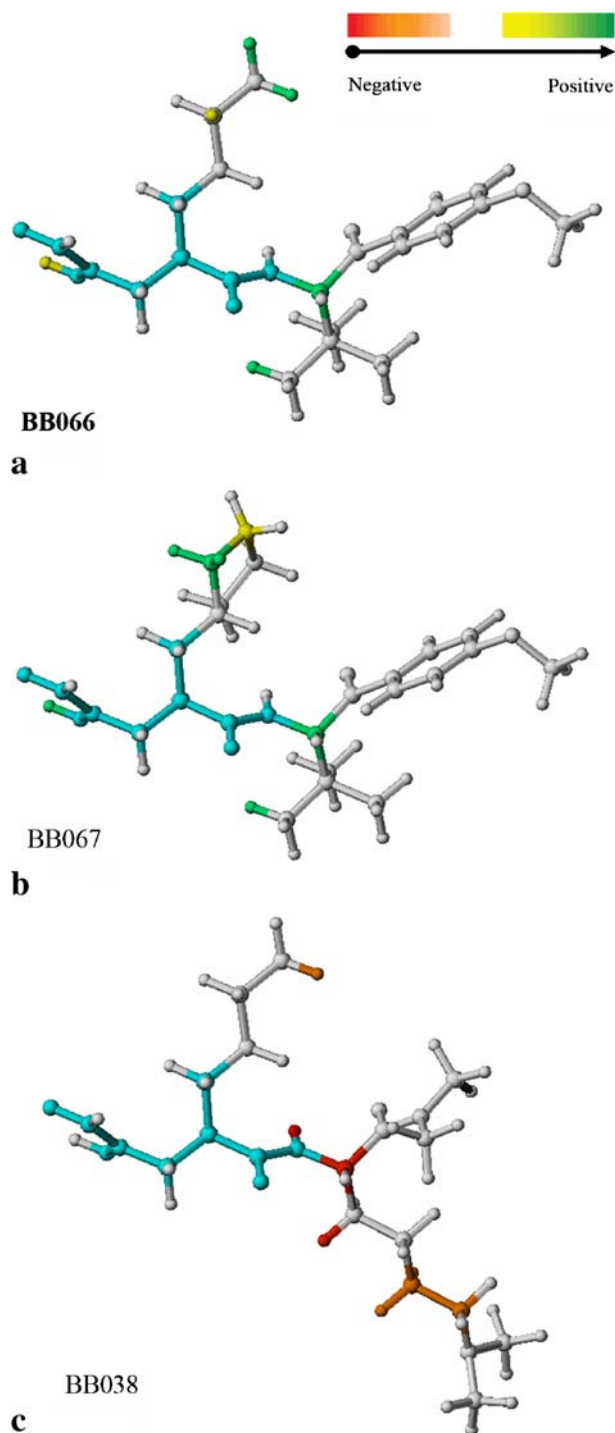


Fig. 8 The individual atomic contributions to activities of **a** BB066, **b** BB067 and **c** BB038. Colors at the red end of the spectrum reflect poor contributions, while colors at the green end indicate favorable contributions

of the chlorine substituent of BB020. The other big white polyhedron seen alongside R_1 and R_2 might come from the methoxybenzyl group at R_1 in BB021 and the hydrophobic cyclic substituents proline and tetrahydroisoquinoline carboxylic acid (Tic) group at R_2 in BB043 and BB044, respectively. In BB039, the guanidine moiety of arginine, which has hydrophobic properties, leads to an orange contour of R_2 . A favorable orange contour of R_3 appeared on the hydrophobic ring system, suggesting that hydrophobic groups like phenyl improve the inhibitory activity.

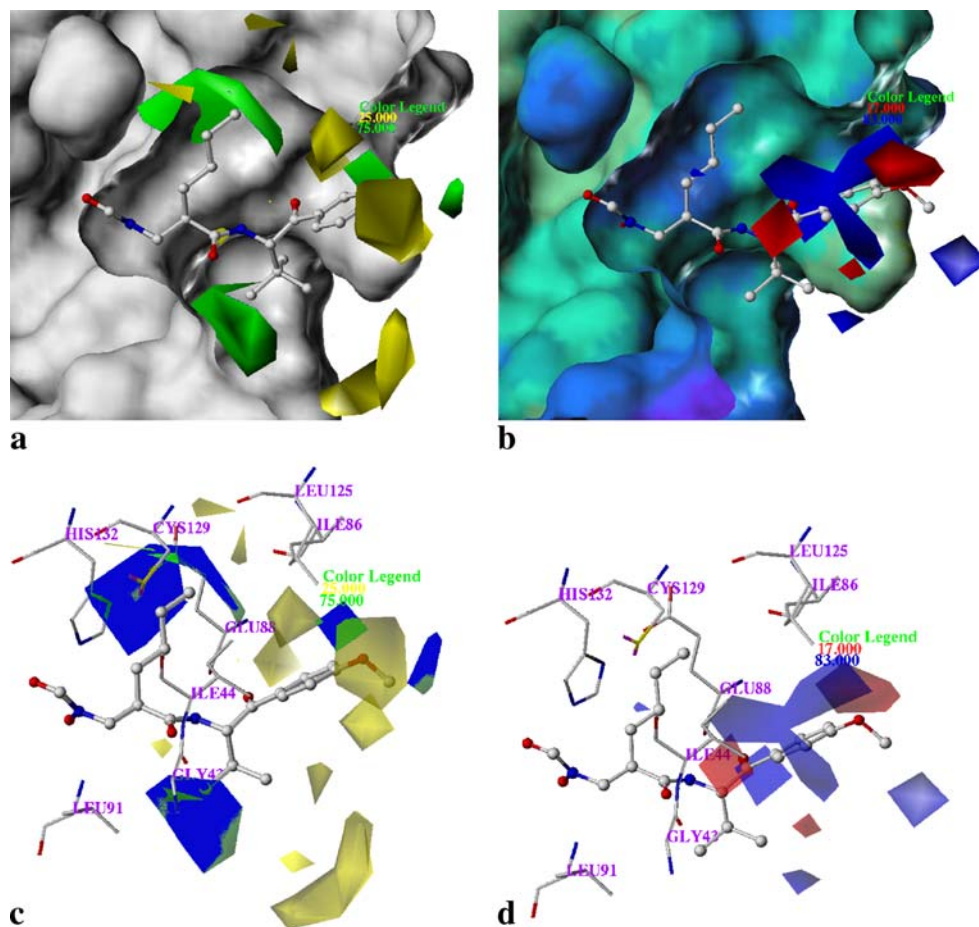
HQSAR reveals the individual atomic contributions to the activity for each molecule by coloring the atoms. The colors at the red end of the spectrum (red, red-orange and orange) reflect poor (or negative) contributions, while colors at the green end (yellow, green-blue and green) indicate favorable (positive) contributions. Atoms with intermediate contributions are colored in white. Atomic contribution maps of BB066 and BB067 as potent inhibitors, and BB038 as an inactive compound, are shown in Fig. 8. The maximal common structure is highlighted by a cyan color since it is common to all compounds and contributes in the same manner to all molecules in the training set. In the representations of BB066 and BB067, the terminal atoms

of *n*-butyl or atoms forming the ring of cyclopentyl-methyl in the R_1 substituent and several atoms of *t*-butyl group in R_2 are colored in green or yellow, indicating that they make a positive contribution to activity. Green-colored atoms in the *t*-butyl group of R_2 also indicate their favorable influence on activity. However, in the case of BB038, some atoms of the *N,N*-dimethylpentyl-1-amine group were colored orange-red, indicating a negative effect on activity.

Comparison of the active site in 3D QSAR models

Among the models constructed by the three QSAR methods, the CoMFA model showed the best predictability. We therefore tried to compare CoMFA contour maps with the PDF active site to check the validity of the model. The CoMFA coefficient contour maps with BB066 were superimposed on the active site of the crystal structure of *E. coli* PDF (1G27) as shown in Fig. 9. These studies established a meaningful correlation between receptor binding site and the ligand-based CoMFA contour map. The steric CoMFA contour maps were projected onto the MOLCAD-generated solvent-accessible (Connolly) surface of the active site (Fig. 9a). The superimposition of the steric fields on the

Fig. 9a–d Superimposition of the CoMFA contour plot in the active site of PDF. **a** CoMFA steric contours projected over the solvent-accessible (Connolly) topological surface of the active site. **b** CoMFA electrostatic contours projected over the electrostatic potential surface (blue negative potential, red/brown positive potential) of the active site. **c** Steric and **d** electrostatic CoMFA contour plot in the active site of PDF. The most active compound (BB066) is shown in the background



binding site of the PDF enzyme is shown in Fig. 9c. The *n*-butyl group of R₁ in BB3497 near the green contour could be mapped to the hydrophobic region of the active site, which is composed of amino acids Ile44, Ile86, Glu88, Leu125, Cys129, and His132. The green contour of the R₂ substituent indicated a bulky group showing favorable interactions with Gly43 and Leu91. However, yellow isopleths are located in the solvent-exposed region, showing relative steric influences without reflecting receptor interactions. The projection of the CoMFA electrostatic contour map onto the electrostatic potential surface map of the binding site generated by MOLCAD shows a general complementarity as depicted in Fig. 9b. Superimposition of the electrostatic fields on the binding site is shown in Fig. 9d. The big blue contour seen at the R₃ group indicates the electropositive groups may cause an increase in hydrophobic interactions with residues Ile44, Ile86, and Leu125. The phenyl ring of BB059-077 has an especially marked influence in increasing activity.

Conclusion

In this study, we have used both 3D-QSAR and 2D-QSAR methods to investigate the relationship between 58 PDF inhibitor analogues and their activities. We successfully obtained three rational and predictive QSAR models. The best model each obtained by the CoMFA, CoMSIA and HQSAR methods has values of $r^2=0.957$, $q^2=0.569$, $ONC=6$; $r^2=0.924$, $q^2=0.520$, $ONC=6$; and $r^2=0.860$, $q^2=0.578$, $ONC=6$, respectively. The predictabilities were validated using a test set comprising 19 external compounds. The predictive r^2 values obtained with the three models were 0.687, 0.505, and 0.178, respectively. Based on the predictive r^2 value of the test set, the CoMFA model appears to be better than the other models. The CoMFA contour map yielded some important structural information that can be used for the design of novel PDF inhibitors. Sterically bulky substitutions like *n*-butyl or cyclopentyl-methyl group are favorable in the R₁ position. For the R₂ region, sterically moderately bulky groups increase the biological activity. The para-substituted phenyl group in the R₃ position is beneficial for activity. CoMSIA contour plots were consistent with CoMFA maps and they additionally offered important information about hydrophobicity. The atomic contribution maps generated by HQSAR provided some information on how individual atoms contribute to the activity for each molecule. The models could be useful in designing and predicting the activity of new PDF inhibitors as antibacterial agents.

Acknowledgment This work was supported by the Korean Ministry of Commerce, Industry and Energy.

References

- Barrett CT, Barrett JF (2003) *Curr Opin Biotechnol* 14:621–626
- Yuan Z, Trias J, White RJ (2001) *Drug Discov Today* 6:954–961
- Adams JM (1968) *J Mol Biol* 5:571–589
- Adams JM, Capecchi MR (1966) *Proc Natl Acad Sci USA* 55:147–155
- Meinzel T, Mechulam Y, Blanquet S (1993) *Biochimie* 75:1061–1075
- Groche D, Becker A, Schlichting I, Kabsch W, Schultz S, Wagner AF (1998) *Biochem Biophys Res Commun* 246:342–346
- Rajagopalan PTR, Yu XC, Pei D (1997) *J Am Chem Soc* 119:12418
- Livingston DM, Leder P (1969) *Biochemistry* 8:435–443
- Lucchini G, Bianchetti R (1980) *Biochim Biophys Acta* 608:54–61
- Pei D (2001) *Expert Opin Ther Targets* 5:23–40
- Giglione C, Meinzel T (2001) *Expert Opin Ther Targets* 5:41–57
- Meinzel T, Patiny L, Ragusa S, Blanquet S (1999) *Biochemistry* 38:4287–4295
- Huntington KM, Yi T, Wei Y, Pei D (2000) *Biochemistry* 39:4543–4551
- Wei Y, Yi T, Huntington KM, Chaudhury C, Pei D (2000) *J Comb Chem* 2:650–657
- Apfel C, Banner DW, Bur D, Dietz M, Hirata T, Hubschwerlen C, Locher H, Page MG, Pirson W, Rosse G, Specklin JL (2000) *J Med Chem* 43:2324–2331
- Jayasekera MM, Kendall A, Shammas R, Dermeyer M, Tomala M, Shapiro MA, Holler TP (2000) *Arch Biochem Biophys* 381:313–316
- Thorarensen A, Deibel MR Jr, Rohrer DC, Vosters AF, Yem AW, Marshall VD, Lynn JC, Bohanon MJ, Tomich PK, Zurenko GE, Sweeney MT, Jensen RM, Nielsen JW, Seest EP, Dolak LA (2001) *Bioorg Med Chem Lett* 11:1355–1358
- Roblin PM, Hammerschlag MR (2003) *Antimicrob Agents Chemother* 47:1447–1448
- Clements JM, Beckett RP, Brown A, Catlin G, Lobell M, Palan S, Thomas W, Whittaker M, Wood S, Salama S, Baker PJ, Rodgers HF, Barynin V, Rice DW, Hunter MG (2001) *Antimicrob Agents Chemother* 45:563–570
- Smith HK, Beckett RP, Clements JM, Doel S, East SP, Launchbury SB, Pratt LM, Spavold ZM, Thomas W, Todd RS, Whittaker M (2002) *Bioorg Med Chem Lett* 12:3595–3599
- Gordon JJ, Kelly BK, Miller GA (1962) *Nature* 195:701–702
- Chen D, Patel D, Hackbarth C, Wang W, Dreyer G, Young D, Margolis P, Wu C, Ni Z, Trias J, White R, Yuan Z (2000) *Biochemistry* 39:1256–1262
- Broughton BJ, Chaplen P, Freeman WA, Warren PJ, Wooldridge KR, Wright DE (1975) *J Chem Soc Perkin Trans* 1:857–860
- Lofland D, Difuntorum S, Waller A, Clements JM, Weaver MK, Karlowsky JA, Johnson K (2004) *J Antimicrob Chemother* 53:664–668
- Azoulay-Dupuis E, Mohler J, Bedos JP (2004) *Antimicrob Agents Chemother* 48:80–85
- Cramer R, Patterson D, Bunce J (1998) *J Am Chem Soc* 110:5959–5967
- Klebe G, Abraham U, Mietzner T (1994) *J Med Chem* 37:4130–4146
- Klebe G (1998) *Perspect Drug Discovery Des* 12:87–104
- Tong W, Lowis DR, Perkins R, Chen Y, Welsh WJ, Goddette DW, Heritage TW, Sheehan DM (1998) *J Chem Inf Comput Sci* 38:669–677
- Glennon RA (2003) *J Med Chem* 46:2795–2812
- Dunn WJ, Wold S, Edlund U, Hellberg S, Gasteiger J (1984) *Quant Struct Act Relat* 3:31–137
- Geladi P (1988) *J Chemom* 2:231–246

33. Wold S (1978) *Technometrics* 4:397–405
34. Diaconis P, Efron B (1984) *Sci Am* 116:96–117
35. Cramer RDI, Bunce JD, Patterson DE (1988) *Quant Struct Act Relat* 7:18–25
36. Davies SJ, Ayscough AP, Beckett RP, Bragg RA, Clements JM, Doel S, Grew C, Launchbury SB, Perkins GM, Pratt LM, Smith HK, Spavold ZM, Thomas SW, Todd RS, Whittaker M (2003) *Bioorg Med Chem Lett* 13:2709–2713
37. Davies SJ, Ayscough AP, Beckett RP, Clements JM, Doel S, Pratt LM, Spavold ZM, Thomas SW, Whittaker M (2003) *Bioorg Med Chem Lett* 13:2715–2718
38. East SP, Beckett RP, Brookings DC, Clements JM, Doel S, Keavey K, Pain G, Smith HK, Thomas W, Thompson AJ, Todd RS, Whittaker M (2004) *Bioorg Med Chem Lett* 14:59–62
39. Clark M, Cramer RD III, van Opdenbosch N (1989) *J Comput Chem* 10:982–1012
40. Purcel WP, Singer JA (1967) *J Chem Eng Data* 12:235–246
41. Burke Jr TR, Fesen MR, Mazumder A, Wang J, Carothers AM, Grunberger D, Driscoll J, Kohn K, Pommier Y (1995) *J Med Chem* 38:4171–4178
42. SYBYL6.9 Force Field manual 57. Tripos, St Louis, MO
43. Gasteiger J, Marsili M (1980) *Tetrahedron* 36:3219–3222
44. Stewart JJP (1990) *J Comput Aided Mol Des* 4:1–103
45. Ghose A, Crippen G (1986) *J Comput Chem* 7:565–577
46. Viswanadhan VN, Ghose AK, Revankar GR, Robins RK (1989) *J Chem Inf Comput Sci* 29:163–172
47. Solis FJ, Wets RJB (1981) *Maths Opera Res* 6:19–30
48. Kubinyi H (1993) *ESCOM Leiden* 93:1281–1306