Improved CoMFA Modeling by Optimization of Settings

Shane D. Peterson, Wesley Schaal, and Anders Karlén*

Department of Medicinal Chemistry, Uppsala University, BMC Box 574, SE-751 23 Uppsala, Sweden

Received December 21, 2004

The possibility of improving the predictive ability of comparative molecular field analysis (CoMFA) by settings optimization has been evaluated to show that CoMFA predictive ability can be improved. Ten different CoMFA settings are evaluated, producing a total of 6120 models. This method has been applied to nine different data sets, including the widely used benchmark steroid data set, as well as eight other data sets proposed as QSAR benchmarking data sets by Sutherland et al. (*J. Med. Chem.* **2004**, *47*, 5541−5554). All data sets have been studied using training and test sets to allow for both internal (q^2) and external (r^2_{pred}) predictive ability assessment. CoMFA settings optimization was successful in developing models with improved q^2 and r^2_{pred} as compared to default CoMFA modeling. Optimized CoMFA is compared with comparative molecular similarity indices analysis (CoMSIA) and holographic quantitative structure−activity relationship (HQSAR) models and found to consistently produce models with improved or equivalent q^2 and r^2_{pred}. The ability of settings optimization to improve model predictive ability has been validated using both internal and external predictions, and the risk of chance correlation has been evaluated using response variable randomization tests.

1. INTRODUCTION

Comparative molecular field analysis (CoMFA) was introduced in 1988 by Cramer et al. to quantify the relationship between three-dimensional molecular structure and biological activity.[1] It has since become one of the most widely used 3D-QSAR methods in medicinal chemistry. CoMFA finds three-dimensional features with both positive and negative influences on activity and then uses this information to predict the activities of other untested compounds. Structure−activity correlations are detected by calculating steric and electrostatic interaction energies between each of a set of compounds and the selected probe atom at many prespecified points surrounding the molecules. Interaction energies between the probe atom and active compound are recorded to form a data matrix consisting of hundreds, or even thousands, of columns of independent data. Partial least squares regression[2] (PLS) is used to establish what, if any, correlation exists between the activity and the steric and electrostatic energy interactions.

CoMFA is implemented in the Sybyl[3] molecular modeling package and allows the modeler to choose settings for a number of parameters in order to make a model. Since the result of simultaneously varying many CoMFA settings is not well-documented in the literature, it is difficult for users to determine which settings, or combinations of settings, work best for modeling a given set of compounds. Consequently, most CoMFA users rely on the default CoMFA settings.

There have been a large number of studies performed to improve CoMFA modeling by addressing one or more of the steps leading up to the development of the model.[4] Molecular alignment is thought to be a very important factor in producing useful CoMFA models, and there have been many studies where authors have developed methods to address alignment problems.[5−10] Closely related to alignment is the orientation of the grid box with respect to the aligned compounds. Several groups have studied the effect of varying grid orientation on CoMFA modeling.[11,12] Other authors have found improved models by adjusting one or more CoMFA settings.[13−16] However, despite the fact that CoMFA has been in wide use since its introduction, we found no published study with a comprehensive analysis of the effect of varying many CoMFA settings on model usefulness.

To explore the possibility that CoMFA modeling could be improved by modifying settings, we undertook a comprehensive approach to settings selection. The use of this approach began in our lab with the study by Schaal in a CoMFA analysis of HIV-1 protease inhibitors using rate constants for association and dissociation as response data.[17] In the present study, we take a more comprehensive look at the effect of settings selection on model predictive ability by varying more settings and have applied the approach to several well-known data sets. The predictive power of each model is evaluated using q^2 and r^2_{pred}, and the risk of chance correlation is assessed using *y*-randomization. This study represents a novel approach to improved CoMFA modeling using a feasible method for testing many combinations of settings.

2. METHODS

2.1. Data Sets. The effectiveness of the settings optimization method is tested on nine data sets. To allow our results to be directly comparable with previous studies, benchmark data sets were chosen. The steroids presented in the original CoMFA article were chosen for method development.[1] Eight other data sets have recently become available electronically in the work of Sutherland et al., where a variety of QSAR methods have been compared.[18]

* Corresponding author. E-mail: Anders.Karlen@orgfarm.uu.se.

10.1021/ci049612j CCC: $33.50
Published on Web 11/03/2005

Table 1. Training and Test Set Compounds with Experimental Activity

compound number	steroid name	CBG activity (log K)
1	aldosterone[a]	6.28
2	androstanediol[a]	5.00
3	androstenediol[a]	5.00
4	androstenedione[a]	5.76
5	androsterone[a]	5.61
6	corticosterone[a]	7.88
7	cortisol[a]	7.88
8	cortisone[a]	6.89
9	dehydroepiandrosterone[a]	5.00
10	deoxycorticosterone[a]	7.65
11	deoxycortisol[a]	7.88
12	dihydrotestosterone[a]	5.92
13	estradiol[a]	5.00
14	estriol[a]	5.00
15	estrone[a]	5.00
16	etiocholanolone[a]	5.23
17	pregnenolone[a]	5.23
18	17-hydroxypregnenolone[a]	5.00
19	progesterone[a]	7.38
20	17-hydroxyprogesterone[a]	7.74
21	testosterone[a]	6.72
22	prednisolone[b]	7.51
23	cortisol-21-acetate[b]	7.55
24	4-pregnene-3,11,20-trione[b]	6.78
25	epicorticosterone[b]	7.20
26	19-nortestosterone[b]	6.14
27	16α,17-dihydroxy-4-pregnene-3,20-dione[b]	6.25
28	16α-methyl-4-pregnene-3,20-dione[b]	7.12
29	19-norprogesterone[b]	6.82
30	11β,17,21-trihydroxy-2α-methyl-4-pregnene-3,20-dione[b]	7.69
31	11β,17,21-trihydroxy-2α-methyl-9α-fluoro-4-pregnene-3,20-dione[b]	5.80

[a] Member of the training set. [b] Member of the test set.

The steroid data set has frequently been used as a benchmark for testing and validating new QSAR methods and is readily available in electronic form.[19] Training set structures were obtained as a binary molecular database included with the Sybyl molecular modeling package. Biological activity data, expressed as log(*K*) of the binding affinity of the steroids to the corticosteroid binding globulin (CBG) carrier protein for the training set was obtained from Dunn et al.[20] The test set consists of 10 steroid structures not included in the training set. Experimental activity data was taken from Coats[19] and is listed in Table 1. The structures of the test set compounds were obtained from the Gasteiger research group.[21] They were derived using CORINA[22] and downloaded from the group's web page. Several authors[19,23] have noted structural errors in the steroid data set and have suggested corrections. Steroid structures used in this study have been examined for correctness and are available in the Supporting Information in Figures 1 and 2.

The molecular structures of steroids in both the training and test sets were minimized to ensure uniformity among the data sets. Minimization was performed using the Tripos force field,[24] Gasteiger−Marsili[25,26] charges, and the conjugate gradient method. No initial optimization was performed, and structures were minimized to convergence using a gradient of 0.05 kcal/(mol Å). The Sybyl command ORIENT BEST_VIEW was used to orient deoxycortisol prior to structure alignment. Structures of the training and test sets were aligned using the rule used by Cramer in the original CoMFA article.[1] This involved least squares fitting all compounds to deoxycortisol using steroid atoms 3, 5, 6 13, 14, and 17.

The remaining eight data sets were obtained in electronic form from the Supporting Information provided by Sutherland et al.[18] and were used without modification. The data sets are of various sizes and span in activity. These include 114 angiotensin converting enzyme (ACE) inhibitors, 111 acetylcholinesterase (AChE) inhibitors, 163 benzodiazepine receptor (BZR) ligands, 322 cyclooxygenase-2 (COX-2) inhibitors, 397 dihydrofolate reductase (DHFR) inhibitors, 66 glycogen phosphorylase b (GPB) inhibitors, 76 thermolysin (THER) inhibitors, and 88 thrombin (THR) inhibitors. Although nearly all of these data sets have been previously used in QSAR studies, training and test sets were redesigned by Sutherland et al. Approximately two-thirds of the compounds were assigned to the training set, and one-third were assigned to the test set. For the BZR, COX-2, and DHFR data sets, several compounds were excluded because of indeterminant activities. After removal of these compounds, the BZR, COX-2, and DHFR sets contained 147, 282, and 361 compounds, respectively. To ensure that test sets would adequately examine predictive accuracy, they were designed to maximize diversity. Care was also taken to be sure that the training and test sets contained compounds with similar ranges of activity.[18]

2.2. CoMFA Modeling. All CoMFA calculations were performed using Sybyl 7.0.[3] In the case of the steroid data set, the Sybyl default region was used while the custom regions of Sutherland et al. were used for modeling those data sets. The default CoMFA region was used for all *y*-randomization tests. All regions use a 2 Å grid spacing. The probe atom used in all CoMFA modeling was a +1 charged carbon sp^3 atom. CoMFA q^2 values were obtained

Table 2. Possible and Evaluated CoMFA Settings Available in Sybyl 7.0[a]

setting name	setting options
dielectric function	**distance**, **constant**
drop electrostatics	**yes**, **no**, **all rows union volume**
DU and LP atoms electrostatic[b]	yes, no
DU and LP atoms steric[c]	yes, no
electrostatic maximum[d]	**5**, **15**, **30**, **45**, **60**, **80**
field type	**both**, **electrostatic**, **steric**
H-bond function	**yes**, **no**
PLS scaling method	autoscale, CoMFA std, none, user specified
rep. vdW term	**8**, **10**, **12**
scale for volume av.[d]	1.0
steric maximum[d]	**5**, **15**, **30**, **45**, **60**, **80**
switch function	**yes**, **no**
transform	**none**, **indicator**, **squared**
volume averaging type	**none**, **box**, **sphere**

[a] Settings and possibilities in bold were those evaluated in this study. Underlined possibilites are those considered to be the default. [b] Setting for the inclusion of dummy atoms and lone pairs in electrostatic field calculations. [c] Setting for the inclusion of dummy atoms and lone pairs in steric field calculations. [d] Possible values not limited to those shown.

using PLS[2] and leave-one-out (LOO) cross-validation. Leave-one-out cross-validation was used instead of the more robust leave-many-out procedure to ensure the reproducibility of q^2.[27] q^2 was determined for the first six principle components, and the optimal number of components was determined on the basis of the first minimum cross-validated standard error of prediction.[28,29] Generally, this resulted in CoMFA models of lower dimensionality, thus, reducing the risk of overfitting. This may result in a different number of optimal components than Sybyl would choose as Sybyl 7.0 determines model complexity on the basis of the highest q^2 available. Calculation time was reduced by using sample-distance partial least squares (SAMPLS),[30] and standard CoMFA scaling was used for all cross-validated models. Predictive r^2 (r^2_{pred}) values were calculated using the following equation:

$$r^2_{pred} = 1 - (PRESS/SD)$$

where PRESS is the sum of the squared differences between the observed and predicted activities and SD is the sum of squared differences between the measured activities of the test set and the average measured activity of the training set.

2.3. Optimization of CoMFA Settings. The settings chosen for exploration in this study are listed in Table 2. CoMFA settings selection is complicated because of many impossible combinations of settings. For example, if one chooses to use only electrostatic fields, settings such as the repulsive vdW term are no longer applicable; the repulsive vdW term is only used in the case where steric or H-bonding fields are used. Figure 1 illustrates the procedure used for generating the possible combinations of settings. The generation of all of the possible combinations of settings resulted in the production of 6120 models.

Further complicating the process of settings optimization are the inconsistencies between the Sybyl graphical user interface (GUI) and the Sybyl global (tailor) variables. The CoMFA settings listed in Table 2 and Figure 1 are referred

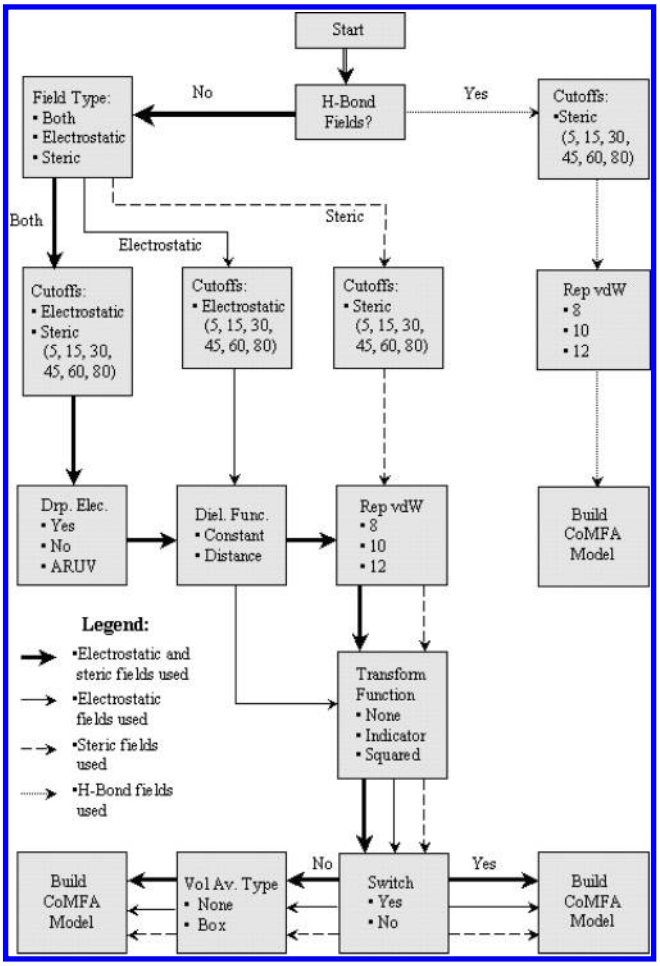

Figure 1. Decision tree for determining possible combinations of CoMFA settings.

to using their names as tailor variables. In some cases, tailor variable names as well as the names of the choices available for each variable differ from their name in the GUI. Also, some settings belonging to different tailor variables are available through the same selection panel in the GUI. Furthermore, it is not possible to set the repulsive vdW term in the GUI. Instead, it must be set as a tailor variable prior to CoMFA model derivation.

To make the optimization of settings feasible, model derivation was automated using a Sybyl programming language (SPL) script. A possible settings combination was first determined according to Figure 1, then used to derive a model. Cross-validation was performed to determine the appropriate number of principal components and q^2. Following cross-validation, a final model was derived without cross-validation and used to predict activities for test set compounds. This process was repeated for each successive model. All calculations were performed on either Intel Xeon 2.4 or 2.8 GHz Intel Pentium workstations with RedHat Linux. The generation of 6120 models required between 2.5 and 38 h of CPU time, depending on the data set and computing platform.

2.4. Selection of Optimal Models. Sorting highly predictive CoMFA models from less-predictive models depends heavily on model validation. Consideration of q^2 is the traditional way of validating PLS models, and recent findings indicate that the use of q^2 alone may be advantageous in cases where data sets are small as a result of the loss of

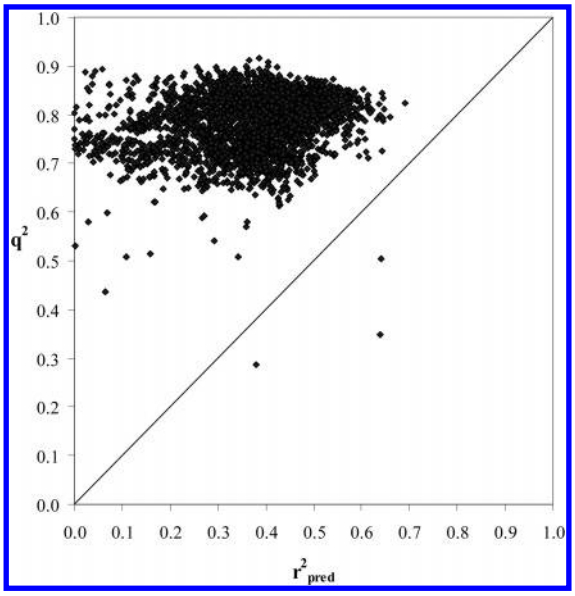

Figure 2. Scatter plot showing q^2 vs r^2_{pred}. All models with q^2 and r^2_{pred} between 0 and 1 are shown.

information associated with removing compounds to form a test set.[31] However, other authors have shown that q^2 is a useful but not sufficient criterion for model validation and have recommended the use of external test sets for the estimation of predictive ability.[32]

The cross-validated r^2 (q^2) is a useful estimate of predictive ability; it shows the degree of internal predictive ability and gives a fast estimate of the likelihood of chance correlation. A q^2 of 0.0 represents a model capable of predicting no better than random guessing, while a model with q^2 equal to 1.0 perfectly predicts the target property. It has been shown that, with a q^2 of approximately 0.3, chance correlation will occur less than 5% of the time for a 21-compound data set.[33] However, a q^2 of approximately 0.4 is often considered the minimum criterion for a useful CoMFA model.

r^2_{pred} is another widely used metric for model validation and is based only on the ability of the model to predict responses for observations never included in the model. There can be large differences between r^2_{pred} values calculated using different mean activity data. Some authors choose to use the average activity of the training set compounds,[1,34−36] while others choose to use the average activity of compounds included in both the training and test sets.[3] An $r^2_{pred} = 0$ corresponds to a model that predicts equally as well as the prediction of activity values based on the average activity of the training set compounds.[37] Although r^2_{pred} is useful in assessing external predictive ability, it is difficult to compare r^2_{pred} values calculated using different data sets. The discrepancy in r^2_{pred} values is often due to different distributions of activity in the training and test sets.[4] This possibility should be kept in mind when comparing models from these data sets.

Generally, it was difficult to make any correlation between r^2_{pred} and q^2 for any of the nine data sets. Figure 2 shows a scatter plot of all CoMFA models with q^2 and r^2_{pred} values between 0 and 1 using the steroids data set. Similar plots have been generated using the Sutherland data sets and can be found in the Supporting Information. It should be noted that, although q^2 varied only between roughly 0.6 and 0.9, there was a much larger spread in r^2_{pred}. This figure indicates that changing CoMFA settings had little effect on q^2 and, based on traditional minimum validation criterion, nearly all models are acceptable. Furthermore, many models had high q^2 values despite the fact that they predicted the activities of the test set compounds poorly.

Since there was a lack of correlation between q^2 and r^2_{pred}, models with high a q^2 did not always have high external predictive abilities. Therefore, to find highly predictive models, r^2_{pred} and q^2 were used as the primary criteria for selecting optimal models, with r^2_{pred} given preference. In all of the data sets studied, different models possessed the highest r^2_{pred} and q^2 values. Therefore, in optimal model selection, a balance between the two estimates was sought. The model thought to be optimal for each data set possessed a high (although often not the *highest*) value for r^2_{pred} and q^2. There are often several models with similar q^2 and r^2_{pred}, making the selection process difficult. In these cases, the number of principle components (c), $SDEP_{cv}$,[28,29] and model fit (r^2) were used as secondary selection criteria. A scoring function (score $= 5r^2_{pred} + 4q^2 + 2r^2$) was used to roughly screen models; however, optimal models were chosen subjectively from those with top scores. On the basis of these criteria, one model was chosen from each data set and is referred to as the "optimal model."

2.5. Validation by *y*-Randomization. Clark and Cramer[33] showed that the risk of chance correlation in CoMFA is low given a sufficiently high q^2. However, to be sure that our settings optimization process is unlikely to find models with spurious structure−activity correlation, *y*-randomization tests were used on all data sets. The risk of spurious correlation can be assessed by scrambling the activity values of the training set compounds, then rederiving the models. If equally good models can be derived from training sets described by random responses, spurious correlation is likely.[38] A thorough analysis of validation by response variable randomization can be found in the work of van der Voet.[39]

Activity values were randomized for all training set compounds. To do this, 100 iterations of randomization were performed, resulting in 100 sets with randomly assigned activities. Fundamental to the response variable randomization test is the assignment of truly random response data. If, during the randomization procedure, the correlation coefficient (comparison between actual and randomized) was outside the interval $-0.05 < r < 0.05$, randomization was repeated. To reduce the likelihood that very similar data sets would be produced by imposing the correlation coefficient restriction, activity assignment uniqueness was assured by allowing no more than 10% of the compounds to be assigned the same activity value as that in any other randomized set. *y*-randomization was performed on all nine data sets used in the study, resulting in the derivation of 612 000 models for each data set, a total of approximately 5.5 million CoMFA models.

2.6. Contour Surfaces. Contour surfaces used in the study were generated using StDev*Coeff as the field type. Favored and disfavored cutoff energies were set at the 70th and 30th percentiles for the steric contributions and at the 75th and 25th percentiles for the electrostatic contributions. This is true for all contour surfaces shown, except in the case of the electrostatic map for model 4699, where the 50th and 40th percentiles were graphed for the electrostatic contribution.

Table 3. Comparison of Optimal CoMFA Settings and Model Statistics for CoMFA Models Using the Steroid Data Set

statistic/setting	default CoMFA[a]	model 2142[a]	model 3084[a]	model 4699[a]	model 3785[a]
q^2	0.70	0.84	0.80	0.82	0.81
r^2_{pred}	0.31	0.64	0.66	0.64	0.62
r^2	0.94	0.95	0.88	0.93	0.87
c	3	2	1	3	1
SDEPcv	0.70	0.49	0.54	0.54	0.53
steric contribution	0.42	0.58	0.55	0.70	0.38
electrostatic contribution	0.58	0.42	0.45	0.30	0.62
dielectric function	distance	constant	constant	distance	constant
drop electrostatics	yes	yes	yes	ARUV	no
electrostatic maximum	30	5	5	30	30
field type	both	both	both	both	both
H-bond function	no	no	no	no	no
rep. vdW term	12	12	8	8	8
steric maximum	30	60	45	15	30
switch function	yes	no	no	no	no
transform	none	indicator	indicator	indicator	squared
volume averaging type	N/A	box	none	none	none

[a] Underlined settings are those considered to be the default.

Table 4. Test Set Experimental and Predicted CBG (log K) Activities[a]

compound number	exptl activity	predicted activity[b]	predicted activity[c]
22	7.51	7.86	7.41
23	7.55	7.84	7.67
24	6.78	6.76	6.94
25	7.20	7.63	7.48
26	6.14	5.79	6.14
27	6.25	7.32	7.06
28	7.12	7.44	7.18
29	6.82	6.67	6.63
30	7.69	7.91	7.63
31	5.80	7.93	7.38

[a] CoMFA models derived using the steroid data set. [b] Predictions made using default CoMFA settings. [c] Predictions made using optimized CoMFA settings (model 2142).

3. RESULTS AND DISCUSSION

3.1. Steroid Data Set. The default CoMFA model using the 21-compound steroid training set produced a three-component model with q^2 equal to 0.70 and r^2_{pred} equal to 0.31 (Table 3). Thus, default CoMFA settings produced a model with high internal and moderate external predictive abilities. Compounds **27** and **31** were not predicted as well as the other compounds (Table 4). In the case of compound **27**, there is no obvious deviation from the experimental space covered by the training set. The compound has an experimental activity of 6.25 and was predicted to have an activity of 7.32. The worst prediction of all was that of compound **31**, which had an experimental activity of 5.80 and a predicted activity of 7.93, a difference of more than 2 logarithmic units. This compound is thought to be poorly represented by the training set as a result of its fluorine substituent and is often poorly predicted in 3D-QSAR studies.[1,19,23,40−47]

An optimal CoMFA model was selected with a $q^2 = 0.84$ and an $r^2_{pred} = 0.64$, showing an improvement in q^2 of 0.14 and an improvement in r^2_{pred} of 0.33 (model 2142 in Table 3). Figure 3 shows a comparison of the predictive abilities of both the default model and the model chosen as the best model of the 6120 models. Table 4 lists the compound numbers, their experimental activities, and their predicted activities based both on the default CoMFA model and model 2142. Figure 3 shows that test set predictions generally improved by selecting more suitable settings. With the exception of compounds **24** and **29**, all test set compounds were predicted better by the optimized CoMFA model than by the default model. Compounds **27** and **31** had respective experimental activities of 6.25 and 5.80 and were predicted worst by both models. Model 2142 predicted these compounds to have activities of 7.06 and 7.38, respectively, while the default model predicted them to have activities of 7.32 and 7.93, respectively.

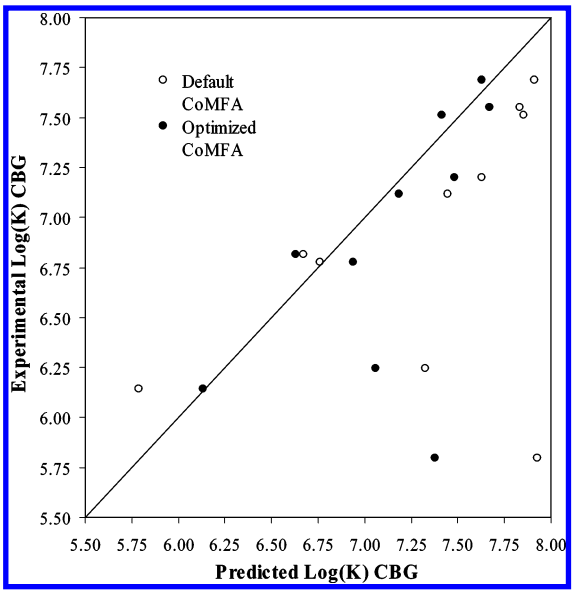

Figure 3. Experimental vs predicted activities using the default and optimized models for the steroid data set.

For the steroid data set, four highly predictive models were selected for comparison, and their settings are listed in Table 3. Interestingly, all four models were made using nonstandard fields. Three of the four models (2142, 3084, and 4699) used indicator fields, while the fourth best model used squared (parabolic) fields. Indicator fields differ from the standard fields in the way energies above and below the cutoff values are treated. When indicator fields are used, all energies above the cutoff have equal influence on the model and are assigned values equal to the corresponding steric or electrostatic cutoff; however, the sign of the interaction energy term is preserved.

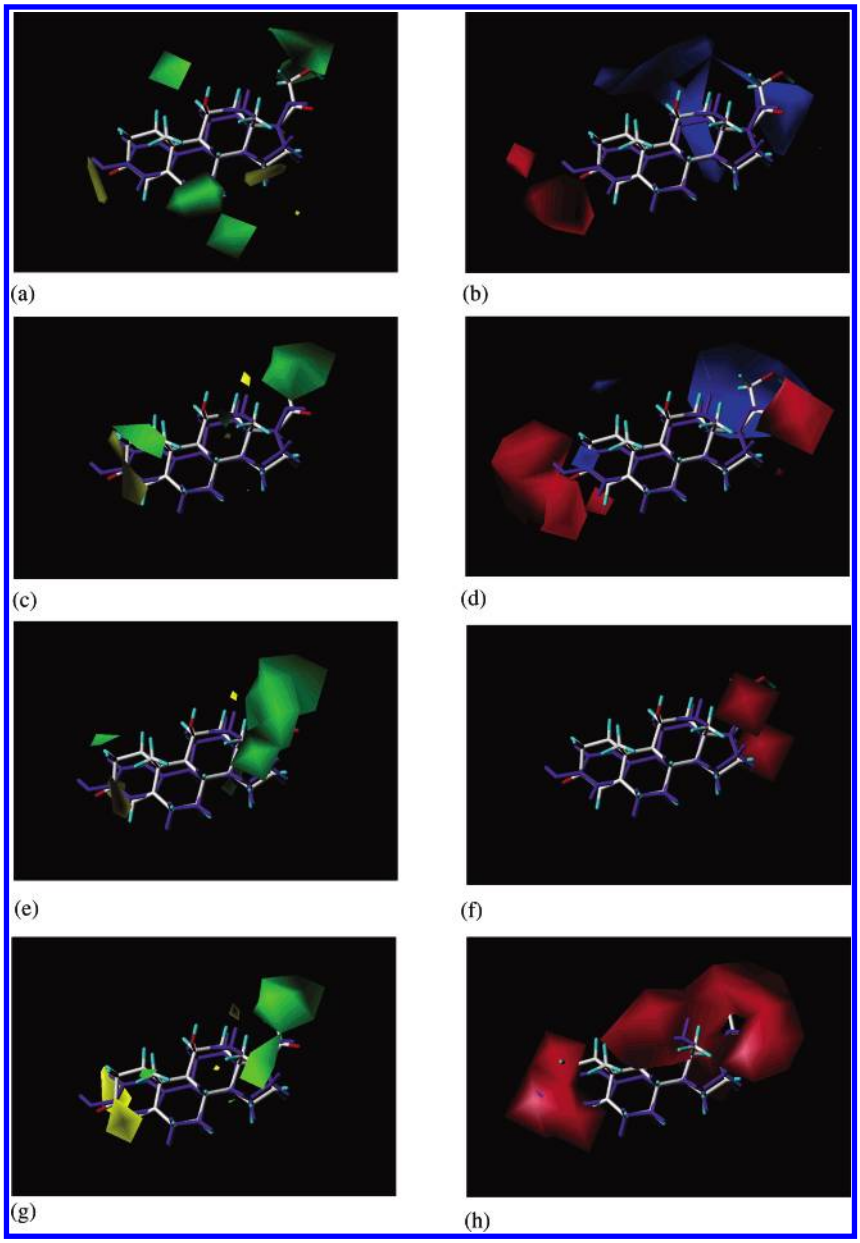

Figure 4. Comparison of contour surfaces derived using four highly predictive CoMFA models: a, c, e, and g show steric surfaces for models 2142, 3084, 4699, and 3785, respectively, and b, d, f, and h show electrostatic surfaces for models 2142, 3084, 4699, and 3785, respectively. Fields shown in green and yellow indicate regions where steric bulk is positively and negatively correlated, respectively, with an increase in activity. Fields shown in red and blue indicate regions where negative charge is positively and negatively correlated, respectively, with an increase in activity.

All energies below the cutoff are assigned a value of zero. In contrast, standard fields place a higher emphasis on the energies below the cutoffs. In the case of squared fields, the magnitude of all standard electrostatic and steric energies at each lattice point are squared, but with the sign of the energy term preserved.

Figure 4 shows a comparison of contour surfaces generated using the four selected models. Model 4699 was based primarily on steric interaction energies, as shown in Table 3, and therefore, the same settings were not effective in producing surfaces for electrostatic contributions. Furthermore, it should be mentioned that contour surfaces for models derived from indicator fields tend not to be viewed well using default viewing;[48] therefore, they were altered according to the protocol described in Section 2.6.

The contour surfaces are superimposed on two compounds with respectively high and low CBG activities. The compound colored on the basis of atom type is corticosterone (compound 6) with a CBG activity of 7.88, and the compound in purple is estradiol (compound 13) with a CBG activity of 5.00. Model statistics as well as settings used for the four CoMFA models are shown in Table 3. The settings used to develop the models depicted in the figure differed in the dielectric function, to what extent electrostatic interactions should be excluded, the steric and electrostatic cutoffs, the switch function, the repulsive vdW term, and the volume averaging type used. Despite these differences in settings, the models showed similarity in steric and electrostatic contour surfaces. This indicates that the underlying SAR for the data set is captured, although different CoMFA settings are used.

3.2. Sutherland Data Sets. The optimization of CoMFA settings led to models with enhanced internal and external predictive abilities for all Sutherland data sets. Table 5 shows

Table 5. A Comparison of Optimal Models for Sutherland Data Sets

method	statistic	ACE	AChE	BZR	COX-2	DHFR	GPB	THER	THR
CoMFA def[a]	q^2	0.68	0.52	0.32	0.49	0.65	0.42	0.55	0.59
	r^2_{pred}	0.48	0.59	−0.01	0.27	0.59	0.43	0.49	0.64
	r^2	0.80	0.78	0.46	0.69	0.79	0.84	0.94	0.85
	$SDEP_{cv}$[b]	1.34	0.86	0.55	0.74	0.76	0.86	1.36	0.63
	c	3	3	2	5	5	4	6	4
CoMFA opt[c]	q^2	0.73	0.70	0.45	0.56	0.68	0.73	0.65	0.66
	r^2_{pred}	0.71	0.73	0.45	0.37	0.64	0.57	0.61	0.72
	r^2	0.96	0.96	0.78	0.75	0.86	0.98	0.95	0.87
	$SDEP_{cv}$[b]	1.27	0.70	0.50	0.69	0.73	0.61	1.19	0.59
	c	6	6	5	6	6	6	6	5
CoMSIA[d,e]	q^2	0.66	0.49	0.45	0.57	0.65	0.61	0.51	0.72
	r^2_{pred}	0.49	0.44	0.12	0.37	0.53	0.59	0.53	0.63
	r^2	0.73	0.86	0.62	0.69	0.75	0.92	0.77	0.89
	$SDEP_{cv}$[f]	1.36	0.89	0.49	0.68	0.76	0.66	1.35	0.56
	c	2/pho	4/all	3/pho	4/all	4/all	4/hyd	3/hyd	4/all
HQSAR[d,g]	q^2	0.72	0.34	0.42	0.50	0.69	0.66	0.49	0.50
	r^2_{pred}	0.30	0.37[h]	0.17	0.27	0.63	0.58	0.53	−0.25
	r^2	0.84	0.72	0.64	0.70	0.81	0.77	0.81	0.87
	$SDEP_{cv}$[f]	1.24	1.00	0.52	0.73	0.71	0.63	1.38	0.74
	c	4/HC	5/H	4/C	7/−	6/HC	2/−	4/−	6/H

[a] Model derived using default settings. [b] Calculated using LOO cross-validation [c] Model derived using optimal settings. [d] Model derived by Sutherland et al.[18] [e] CoMSIA extra fields, pho, hyd, and all indicate the use of hydrophobic, hydrogen bonding, and all types in addition to electrostatic and steric field types. [f] Calculated using L 10% O cross-validation by Sutherland et al.[18] [g] HQSAR component clarification: C and H indicate the chirality and hydrogens for defining holograms. [h] Compounds 2−36 excluded because the predicted activity exceeds the tested activity by more than 12 pIC_{50} units.

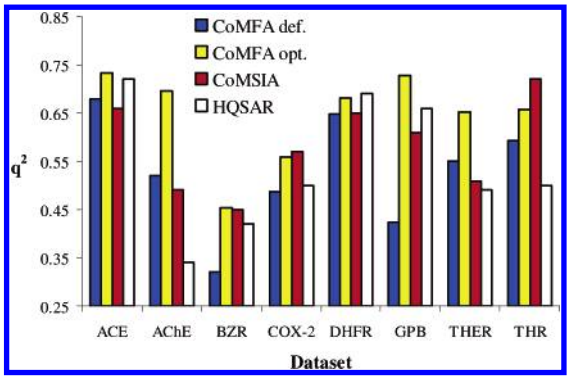

Figure 5. Comparison of q^2 for Sutherland data sets using default CoMFA, optimized CoMFA, CoMSIA, and HQSAR. CoMSIA and HQSAR data taken from the work of Sutherland et al.[18]

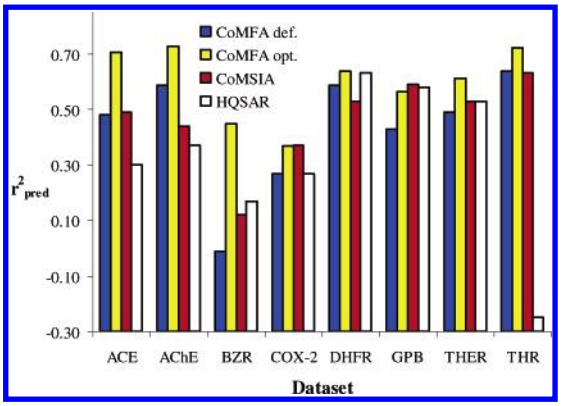

Figure 6. Comparison of r^2_{pred} for Sutherland data sets using default CoMFA, optimized CoMFA, CoMSIA, and HQSAR. CoMSIA and HQSAR data taken from the work of Sutherland et al.[18]

a comparison of internal and external predictive abilities for default and optimized CoMFA models as well as comparative molecular similarity indices analysis (CoMSIA)[47] and holographic quantitative structure−activity relationship (HQSAR)[49] models derived by Sutherland et al.[18] Sutherland et al. also test EVA, 2D-QSAR, and 2.5D-QSAR using these data sets, but we limit our comparison to CoMSIA and HQSAR since they tended to produce the most predictive models for these data sets. Furthermore, Sutherland et al. use two variants of CoMSIA: CoMSIA default and CoMSIA extra, where a combination of hydrophobic/hydrogen bonding fields are used with electrostatic and steric fields. Since CoMSIA extra tended to perform better than CoMSIA default, it was used for comparison with our optimized CoMFA models. All further mention of CoMSIA within this article refers to the CoMSIA extra results of Sutherland et al.

Figures 5 and 6 show a graphical comparison of the different QSAR methods using q^2 and r^2_{pred}, respectively. Although HQSAR or CoMSIA were able to produce models slightly higher (0.01−0.06) in either q^2 or r^2_{pred}, in the cases of the COX-2, DHFR, GPB, and THR data sets, they were not able to produce models where both q^2 *and* r^2_{pred} were higher. For the remaining four data sets (ACE, AChE, BZR, and THER), the optimization of CoMFA settings was useful in finding models equal to or better in *both* q^2 and r^2_{pred} as compared with the CoMSIA and HQSAR counterparts.

As mentioned previously, our settings selection algorithm derives 6120 models using various combinations of settings. Included in the 6120 models are those with very high, moderate, and very low predictive abilities. With that in mind, it is interesting to note how well the default CoMFA model performs relative to all others in the design matrix. To make models comparable, they were given a score based on a weighted sum of q^2, r^2_{pred}, and r^2 (see Section 2.4). This score gives a rough indication of model usefulness relative to the other models derived using the same data set. To allow for a comparison of relative performance, models have been ranked using this score. The data set for which the default model performed best is the DHFR data set, where it is ranked 802 and still far from being an optimal model. Default models for COX-2 (1507), THER (1682), and GPB (1849)

Table 6. Comparison of Optimal Settings for Sutherland Data Sets

statistic/setting	ACE [a]	AChE[a]	BZR[a]	COX-2[a]	DHFR[a]	GPB[a]	THER[a]	THR[a]
model #	4556	5846	5095	2354	6079	4622	4078	5922
q^2	0.73	0.70	0.45	0.56	0.68	0.73	0.65	0.66
r^2_{pred}	0.71	0.73	0.45	0.37	0.64	0.57	0.61	0.72
dielectric function	distance	distance	distance	distance	distance	distance	constant	constant
drop electrostatics	no	no	ARUV	no	ARUV	no	ARUV	yes
electrostatic maximum	45	15	5	80	5	45	5	15
field type	both	both	both	both	both	both	both	both
H-bond function	no	no	no	no	no	no	no	no
rep. vdW term	8	12	8	12	12	12	10	12
steric maximum	15	5	15	45	5	15	30	5
switch function	no	no	no	no	no	no	no	no
transform	indicator	indicator	squared	squared	squared	squared	indicator	squared
volume averaging type	box	none	box	box	box	box	box	box

[a] Underlined settings are those considered to be the default.

Table 7. Comparison of Optimized Model Performance[a] Using Subsequent Data Sets[b]

model	steroids	ACE	AChE	BZR	COX-2	DHFR	GPB	THER	THR
2142	**0.84/0.64**	0.66/0.49	0.47/0.61	0.39/0.29	0.32/0.25[c]	0.66/0.55	0.37/0.38[c]	0.58/0.51	0.65/0.61
4556	0.86/0.29	**0.73/0.71**	0.46/0.55[c]	0.30/0.23	0.42/0.15[c]	0.64/0.68	0.48/0.51	0.60/0.49	0.64/0.50
5846	0.82/0.53	0.68/0.49	**0.70/0.73**	0.39/0.13	0.43/0.29	0.64/0.59[c]	0.26/0.15[c]	0.61/0.49	0.64/0.57
5095	0.84/0.44	0.68/0.53	0.52/0.61	**0.45/0.45**	0.51/0.30	0.66/0.62	0.37/0.10[c]	0.60/0.46	0.63/0.60
2354	0.74/0.47	0.68/0.71	0.44/0.60	0.38/0.26	**0.56/0.37**	0.63/0.59[c]	0.48/0.48	0.49/0.33[c]	0.62/0.63
6079	0.83/0.47	0.65/0.46[c]	0.43/0.63	0.39/0.29	0.49/0.26[c]	**0.68/0.64**	0.40/0.20[c]	0.38/0.27[c]	0.66/0.62
4622	0.83/0.40	0.68/0.70	0.48/0.63	0.37/0.19	0.48/0.37	0.62/0.57[c]	**0.73/0.56**	0.50/0.35[c]	0.61/0.62
4078	0.81/0.53	0.65/0.44[c]	0.50/0.64	0.42/0.21	0.36/0.27[c]	0.64/0.54[c]	0.21/0.10[c]	**0.65/0.61**	0.66/0.63
5922	0.79/0.35	0.64/0.56	0.35/0.50[c]	0.35/0.23	0.41/0.27[c]	0.64/0.51[c]	0.47/0.36	0.41/0.38[c]	**0.66/0.72**
default	0.70/0.31	0.68/0.48	0.52/0.58	0.32/−0.01	0.49/0.27	0.65/0.59	0.42/0.43	0.55/0.49	0.59/0.64

[a] Performance is estimated by q^2 and r^2_{pred}, shown here as q^2/r^2_{pred}. [b] Values shown in bold represent the optimal model for the given data set. [c] Model with q^2 and r^2_{pred} less than or equal to the default model.

are in the top third of models and further from being optimal models. The default model is slightly lower for AChE; with a ranking of 2198, it falls outside of the top third of highly ranked models. In the cases of ACE and THR, the default models are in the bottom 50% with ranks of 4131 and 4005, respectively. Finally, with a rank of 5500, the default model for the BZR data set performed the worst of all data sets. These results show that there is great potential for improving the predictive ability of models generated with default settings.

3.3. Comparison of Optimal Models. With evidence that CoMFA models can be so dramatically improved by finding better parameters, one may wonder if some combinations of CoMFA settings consistently perform well for all data sets. By considering the optimal models derived from each of the nine data sets, some similarity in settings becomes apparent, as can be seen in Tables 3 and 6. For example, all data sets use both steric and electrostatic fields and do not use H-bonding fields. Another noticeable feature among the optimal models is that they all use either indicator or squared transform types. Figure 3 of the Supporting Information shows scatter plots of models and their respective r^2_{pred} and q^2 values for all data sets. Models are color-coded to show differences between indicator, squared, standard transform, and H-bonding fields. The figure shows that the models derived using either indictor or squared transforms are more frequently found among the most predictive models for all nine data sets. However, many of the worst models for each data set were also derived using these transforms, which shows that indicator and squared transforms do not always perform well. H-bonding fields produce models with a varying degree of predictive ability, depending on the data set. However, there is only one case (GPB) where H-bonding fields produce models with a predictive ability comparable with models derived using either indicator or squared transforms. This suggests that H-bonding fields are less likely to produce highly predictive CoMFA models (for the data sets studied), making them less interesting than the combination of electrostatic and steric fields used with a squared or indicator transform.

The predictive statistics for the optimal models from each data set were also compared to all other data sets (Table 7). For example, the parameters found to be optimal for the ACE data set (model 4556) are used to calculate CoMFA models using the remaining eight data sets. The last row in Table 7 shows the predictive ability of the default model. Since our algorithm for model derivation evaluates parameter combinations in the same order each time it is run, one can easily find a given settings combination based on its model number in the data matrix. Table 7 shows that optimal settings are unlikely to be consistent among different data sets when the optimal model is considered, though good models are found in many cases. Indeed, there are many instances (shown in the table with a footnote) where optimal models for one data set perform worse than the default for a different data set. This was particularly noticeable for data sets DHFR, COX-2, THER, and GPB where the default model performed best (see Section 3.2). However, there were also several optimal models which performed worse than default for the ACE and AChE data sets.

Although no particular combination of parameters performed optimally for all nine data sets, it may be interesting to know if any models performed consistently better than the default for all data sets. To understand this, the q^2 and

r^2_{pred} values for all models from each of the nine data sets were compared. There were no models with higher values in both q^2 and r^2_{pred} among all data sets; however, five models were found to be superior in seven of nine data sets. In all five cases, both electrostatic and steric fields were used, no hydrogen bonding fields were used, the switch function was set to "no", and a distance-dependent dielectric function was used. Also of note is the fact that either indicator or squared transform types were used in all five consistently good models. With the best average rank of the five selected models, model 4125 may be thought to perform best. Model 4125 was different from the default model in the settings electrostatic cutoff, transform, switch function, volume averaging type, and repulsive vdW term. Those parameters are set to 5, squared, no, box, and 8, respectively. All other settings were set to their default values.

These five models were generally more predictive than default but were still far from optimal. For example, model 4125 performed best for the BZR data set, where it was ranked 134 of all 6120 models. Despite this high rank, this parameter combination did not perform as well on the ACE data set, where it received a ranking of 3866 and was only a marginal improvement over the default model, which was ranked 4131. Furthermore, model 4125 did not have a better q^2 and r^2_{pred} than the default model for data sets GPB and THER.

On the basis of these findings, optimal models for one data set do not perform optimally for different data sets. Considering the potential performance gain in optimal models, users may consider investing the time required to optimize parameters. A global search of parameter combinations has been done here, but the time required to find better models might be reduced by focusing on settings frequently present in optimal models. As previously discussed, several modifications may be made to the default model and could improve the predictive performance of CoMFA.

3.4. Validation by *y*-Randomization. To validate this approach, it is necessary to show that, by simply selecting various combinations of settings, it is not possible to make a seemingly good CoMFA model from random responses. If it is possible to develop predictive models from incorrect biological data by simply varying CoMFA settings, one could conclude that such a settings selections process is dangerous in terms of model validity.

To test the possibility of chance correlation in CoMFA settings optimization, training set activity values were randomly assigned 100 times for each data set. The CoMFA settings optimization algorithm was applied to all 100 random activity sets using all nine data sets, resulting in the derivation of more than 5.5 million CoMFA models. Tables 8 and 9 show a comparison of the number of models with both q^2 and r^2_{pred} values above the specified cutoff. Table 8 was made using correct activity data, while Table 9 was made from the randomly assigned sets.

Of the 5.5 million models included in Table 9, there are only 534 models where both q^2 and r^2_{pred} are greater than 0.2, only 75 models are above 0.3, and only 5 are above 0.4. Furthermore, where randomly assigned responses were used, the optimization of CoMFA parameters was not successful in finding any models with both q^2 and r^2_{pred} above 0.5 using any of the nine tested data sets. The steroid and GPB data sets were the only data sets for which models with both internal and external predictivities above 0.2 could be derived. Due to the fact that Table 8 was compiled using correct responses, only 55 080 models are included there. Despite the fact that this table contains only 1% as much data as Table 9, 18 931 models have both internal and external predictive abilities above 0.5. Upon examination of Tables 8 and 9, it is clear that models derived using randomly assigned activity stand in stark contrast to models derived using correctly assigned responses. It is, therefore, highly unlikely that sampling many combinations of parameters will result in spurious CoMFA models.

Table 8. Total Number[a] of Models with Both q^2 and r^2_{pred} above the Specified Cutoff

data set	>0.00	>0.10	>0.20	>0.30	>0.40	>0.50
steroids	6020	5926	5603	4727	2519	507
ACE	6116	6116	6116	6107	5524	4051
AChE	6102	6088	6070	5913	5003	1897
BZR	5598	4221	2347	509	43	0
COX2	6099	5784	4584	961	4	0
DHFR	6120	6120	6119	6113	6056	5508
GPB	5979	5562	4950	3603	1579	99
THER	6114	6075	5990	5794	4823	1210
THR	6115	6104	6086	6049	5995	5659

[a] Represents total number of models from 6120 models derived using real responses.

Table 9. Total Number[a] of Models with Both q^2 and r^2_{pred} above the Specified Cutoff

data set	>0.00	>0.10	>0.20	>0.30	>0.40	>0.50
steroids	8940	2556	491	75	5	0
ACE	6738	3	0	0	0	0
AChE	7738	294	0	0	0	0
BZR	15 531	0	0	0	0	0
COX2	6089	0	0	0	0	0
DHFR	22 284	0	0	0	0	0
GPB	17 235	240	43	0	0	0
THER	5386	32	0	0	0	0
THR	11 524	152	0	0	0	0

[a] Represents total number of models from 612 000 models derived for *y*-randomization tests.

4. CONCLUSION

Ten CoMFA parameters have been explored for the benchmark steroid data set and the eight additional data sets from Sutherland et al. A total of 6120 distinct CoMFA models were calculated for each data set evaluated. This approach of testing many combinations of CoMFA settings was shown to be successful in finding models with enhanced internal and external predictive abilities. Importantly, extensive *y*-randomization tests have shown that spurious correlations were highly unlikely. In the cases of the eight Sutherland data sets, optimized CoMFA models are essentially as good as or better than both HQSAR and CoMSIA models. Given the significant improvements in predictive ability, an evaluation of many CoMFA parameters may be a valuable tool for CoMFA users, especially in cases where low q^2 and/or r^2_{pred} values are obtained.

Supporting Information Available: Steroid training set structures, steroid test set structures, and scatter plots of r^2_{pred} and q^2 values of all nine data sets used in this study. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Cramer, R. D., III.; Patterson, D. E.; Bunce, J. D. Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959−5967.

(2) Wold, S. Validation of QSARs. *Quant. Struct.-Act. Relat.* **1991**, *10*, 191−193.

(3) *Sybyl*, version 7.0; Tripos Inc.: St. Louis, MO.

(4) Kim, K. H.; Greco, G.; Novellino, E. A Critical Review of Recent CoMFA Applications. *Perspect. Drug Discovery Des.* **1998**, *12/13/14*, 257−315.

(5) Jewell, N. E.; Turner, D. B.; Willett, P.; Sexton, G. J. Automatic Generation of Alignments for 3D QSAR Analyses. *J. Mol. Graph. Model.* **2001**, *20*, 111−121.

(6) Talele, T. T.; Kulkarni, S. S.; Kulkarni, V. M. Development of Pharmacophore Alignment Models as Input for Comparative Molecular Field Analysis of a Diverse Set of Azole Antifungal Agents. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 958−966.

(7) Lemmen, C.; Lengauer, T. Computational Methods for the Structural Alignment of Molecules. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 215−232.

(8) Jalaie, M.; Erickson, J. A. Homology Model Directed Alignment Selection for Comparative Molecular Field Analysis: Application to Photosystem II Inhibitors. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 181−197.

(9) Cho, S. J.; Serrano Garsia, M. L.; Bier, J.; Tropsha, A. Structure-Based Alignment and Comparative Molecular Field Analysis of Acetylcholinesterase Inhibitors. *J. Med. Chem.* **1996**, *39*, 5064−5071.

(10) Tropsha, A.; Cho, S. J. Cross-Validated R^2 Guided Region Selection for CoMFA Studies. *Perspect. Drug Discovery Des.* **1998**, *12/13/14*, 57−69.

(11) Wang, R.; Gao, Y.; Liu, L.; Lai, L. All-Orientation Search and All-Placement Search in Comparative Molecular Field Analysis. *J. Mol. Model.* **1998**, *4*, 276−283.

(12) Cho, S. J.; Tropsha, A. Cross-Validated R^2-Guided Region Selection for Comparative Molecular Field Analysis: A Simple Method to Achieve Consistent Results. *J. Med. Chem.* **1995**, *38*, 1060−1066.

(13) Kellogg, G. E. Finding Optimum Field Models for 3D QSAR. *Med. Chem. Res.* **1997**, *7*, 417−427.

(14) Melville, J. L.; Hirst, J. D. On the Stability of CoMFA Models. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1294−1300.

(15) Bursi, R.; Grootenhuis, P. D. J. Comparative Molecular Field Analysis and Energy Interaction Studies of Thrombin-Inhibitor Complexes. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 221−232.

(16) Dinan, L.; Hormann, R. E.; Fujimoto, T. An Extensive Ecdysteroid CoMFA. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 185−207.

(17) Schaal, W. Computational Studies of HIV-1 Protease Inhibitors. In *Comprehensive Summaries of Uppsala Dissertations From the Faculty of Pharmacy*; Uppsala University: Uppsala, Sweden, 2002; p 88.

(18) Sutherland, J. J.; O'Brien, L. A.; Weaver, D. F. A Comparison of Methods for Modeling Quantitative Structure−Activity Relationships. *J. Med. Chem.* **2004**, *47*, 5541−5554.

(19) Coats, E. A. The CoMFA Steroids as a Benchmark Dataset for Development of 3D QSAR Methods. *Perspect. Drug Discovery Des.* **1998**, *12/13/14*, 199−213.

(20) Dunn, J. F.; Nisula, B. C.; Rodbard, D. Transport of Steroid Hormones: Binding of 21 Endogenous Steroids to Both Testosterone-Binding Globulin and Corticosteroid-Binding Globulin in Human Plasma. *J. Clin. Endocrinol. Metab.* **1981**, *53*, 58−68.

(21) Gasteiger, J. *Dataset of 31 Steroids Binding to the Corticosteroid Binding Globulin (CBG) Receptor*; Gasteiger Research Group: Erlangen, Germany. http://www2.chemie.uni-erlangen.de/services/steroids/.

(22) Gasteiger, J.; Rudolph, C.; Sadowski, J. Automatic Generation of 3D Atomic Coordinates for Organic Molecules. *Tetrahedron Comput. Methodol.* **1990**, *3*, 537−547.

(23) Wagener, M.; Sadowski, J.; Gasteiger, J. Autocorrelation of Molecular Surface Properties for Modeling Corticosteroid Binding Globulin and Cytosolic Ah Receptor Activity by Neural Networks. *J. Am. Chem. Soc.* **1995**, *117*, 7769−7775.

(24) Clark, M.; Cramer, R. D., III.; Van Opdenbosch, N. Validation of the General Purpose Tripos 5.2 Force Field. *J. Comput. Chem.* **1989**, *10*, 982−1012.

(25) Gasteiger, J.; Marsili, M. Iterative Partial Equalization of Orbital Electronegativity: A Rapid Access to Atomic Charges. *Tetrahedron* **1980**, *36*, 3219−3222.

(26) Hinze, J.; Jaffe, H. H. Electronegativity. I. Orbital Electronegativity of Neutral Atoms. *J. Am. Chem. Soc.* **1962**, *84*, 540−546.

(27) Shao, J. Linear Model Selection by Cross-Validation. *J. Am. Stat. Assoc.* **1993**, *88*, 486-494.

(28) Baroni, M.; Clementi, S.; Cruciani, G.; Costantino, G.; Riganelli, D.; Oberrauch, E. Predictive Ability of Regression Models. Part II. Selection of the Best Predictive PLS Model. *J. Chemom.* **1992**, *6*, 347−356.

(29) Cruciani, G.; Baroni, M.; Clementi, S.; Costantino, G.; Riganelli, D.; Skagerberg, B. Predictive Ability of Regression Models. Part I. Standard Deviation of Prediction Errors (SDEP). *J. Chemom.* **1992**, *6*, 335−346.

(30) Bush, B. L.; Nachbar, R. B., Jr. Sample-Distance Partial Least Squares: PLS Optimized for Many Variables, with Application to CoMFA. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 587−619.

(31) Hawkins, D. M.; Basak, S. C.; Mills, D. Assessing Model Fit by Cross-Validation. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 579−586.

(32) Golbraikh, A.; Tropsha, A. Beware of q^2. *J. Mol. Graph. Model.* **2002**, *20*, 269−276.

(33) Clark, M.; Cramer, R. D., III. The Probability of Chance Correlation Using Partial Least Squares (PLS). *Quant. Struct.-Act. Relat.* **1993**, *12*, 137−145.

(34) Gokhale, V. M.; Kulkarni, V. M. Comparative Molecular Field Analysis of Fungal Squalene Epoxidase Inhibitors. *J. Med. Chem.* **1999**, *42*, 5348−5358.

(35) DePriest, S. A.; Mayer, D.; Naylor, C. B.; Marshall, G. R. 3D-QSAR of Angiotensin-Converting Enzyme and Thermolysin Inhibitors: A Comparison of CoMFA Models Based on Deduced and Experimentally Determined Active Site Geometries. *J. Am. Chem. Soc.* **1993**, *115*, 5372−5384.

(36) Sulea, T.; Purisima, E. O. Comparative Molecular Field Analysis of Protein Tyrosine Phosphatase Low-Molecular-Weight Substrates. *THEOCHEM* **1998**, *434*, 139−153.

(37) Waller, C. L.; Marshall, G. R. Three-Dimensional Quantitative Structure−Activity Relationship of Angiotesin-Converting Enzyme and Thermolysin Inhibitors. II. A Comparison of CoMFA Models Incorporating Molecular Orbital Fields and Desolvation Free Energies Based on Active-Analog and Complementary-Receptor-Field Alignment Rules. *J. Med. Chem.* **1993**, *36*, 2390−2403.

(38) Wold, S.; Eriksson, L. Statistical Validation of QSAR Results. Validation Tools. *Methods Princ. Med. Chem.* **1995**, *2*, 309−318.

(39) van der Voet, H. Comparing the Predictive Accuracy of Models Using a Simple Randomization Test. *Chemom. Intell. Lab. Syst.* **1994**, *25*, 313−323.

(40) Norinder, U. 3D-QSAR Investigation of the Tripos Benchmark Steroids and Some Protein-Tyrosine Kinase Inhibitors of Styrene Type Using the TDQ Approach. *J. Chemom.* **1996**, *10*, 533−545.

(41) Oprea, T. I.; Ciubotariu, D.; Sulea, T. I.; Simon, Z. Comparison of the Minimal Steric Difference (MTD) and Comparative Molecular Field Analysis (CoMFA) Methods for Analysis of Binding of Steroids to Carrier Proteins. *Quant. Struct.-Act. Relat.* **1993**, *12*, 21−26.

(42) Silverman, B. D. The Thirty-One Benchmark Steroids Revisited: Comparative Molecular Moment Analysis (CoMMA) with Principal Component Regression. *Quant. Struct.-Act. Relat.* **2000**, *19*, 237−246.

(43) Tuppurainen, K.; Viisas, M.; Laatikainen, R.; Peraekylae, M. Evaluation of a Novel Electronic Eigenvalue (EEVA) Molecular Descriptor for QSAR/QSPR Studies: Validation Using a Benchmark Steroid Data Set. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 607−613.

(44) Silverman, B. D.; Platt, D. E. Comparative Molecular Moment Analysis (CoMMA): 3D-QSAR without Molecular Superposition. *J. Med. Chem.* **1996**, *39*, 2129−2140.

(45) Maw, H. H.; Hall, L. H. E-State Modeling of Corticosteroids Binding Affinity: Validation of Model for Small Data Set. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1248−1254.

(46) Jain, A. N.; Koile, K.; Chapman, D. Compass: Predicting Biological Activities from Molecular Surface Properties. Performance Comparisons on a Steroid Benchmark. *J. Med. Chem.* **1994**, *37*, 2315−2327.

(47) Klebe, G.; Abraham, U.; Mietzner, T. Molecular Similarity Indices in a Comparative Analysis (CoMSIA) of Drug Molecules to Correlate and Predict Their Biological Activity. *J. Med. Chem.* **1994**, *37*, 4130−4146.

(48) *QSAR and Advanced CoMFA*, Sybyl 7.0 Manual (for Linux); Tripos Inc.: St. Louis, MO, 2004.

(49) *HQSAR*, Sybyl 7.0 Manual (for Linux); Tripos Inc.: St. Louis, MO, 2004.

CI049612J