ELSEVIER

PARM: A practical utility for drug design

Jianfeng Pei, Jiaju Zhou, Guirong Xie, Hongming Chen, and Xianfeng He

Laboratory of Computer Chemistry (LCC), Institute of Chemical Metallurgy, Chinese Academy of Sciences, Beijing, China

To accommodate situations in which the 3D structure of the target receptor is not available, we have developed the Pseudo Atomic Receptor Model (PARM) software package. In this article we describe PARM and illustrate its use with three examples: elemenes (potential anticancer drugs), angiotensin converting enzyme inhibitors, and human HIV-1 inhibitors TTD (1,1,3-trioxo-2H, 4H-thieno[3,4-e][1,2,4] thiadiazine derivatives). The results show that PARM can build models with favorable cross-validation statistics (R_{cv}^2 values 0.7–0.9) and give helpful SAR insight. PARM has certain advantages: (a) it can be used for many systems, regardless of whether the 3D structure of the receptor is known; (b) PARM models were demonstrated to be highly statistically reliable; and (c) PARM analyses are robust and reproducible. © 2001 by Elsevier Science Inc.

Keywords: drug design, PARM, 3D QSAR, pseudoreceptor model, elemene, anti-HIV, cancer, ACE inhibitor

INTRODUCTION

In computer-aided molecular design, the molecular modeling approach based on molecular interactions between ligand and target protein plays an important role.[1–4] For systems such as enzymes and their inhibitors where the 3D structrual data of the target enzyme is known, there is reliable information on which to base a study of these interactions. This methodology has become the "gold standard" in drug design as has been shown in numerous studies.[5–10] For other systems lacking 3D structural data for the target, homology or other methods can be used to predict the 3D structure of the target.[11–14] However, this approach is more problematic, as the predicted target 3D structure is not as reliable for de novo drug design. CoMFA[15] is the most popular 3D QSAR approach in use when the 3D structure of the target protein is not known. Recently, several pseudoreceptor modelling methods, alternative 3D QSAR methods, have been developed and employed in drug design. Jain et al.[16] established a pseudoreceptor model in 1994 to predict the odor of flavorants. Jain's pseudoreceptor method used receptor shape factors and Van der Waals interactions. The receptor modelling method proposed by Walters and Hinds[17] consists of pseudo-atoms at the topological level. Hahn[18,19] used grid points in the space to describe four types of interactions—electrostatic, hydrophobic, hydrogen bond, and Van der Waals. That is, each spatial point in the receptor model 3D surface surrounding the set of ligands consists of four real numbers. Vedani et al. constructed pseudoreceptor models using fragments from amino-acids.[20] Quasi-atomistic receptor models[21] have also been developed that populate with atomistic properties (hydrogen bonds, salt bridges, aromatic and aliphatic regions, solvent) and individually fitted receptor envelopes (simulating a flexible receptor cavity) using an induced fit. All of these methods are based on a training set of ligand molecular structures together with their bioactivities. Previous work on pseudoreceptors at the LCC was based on an improvement of Walters' approach.[17] This led to a pseudoreceptor modelling method now implemented in our Pseudo Atomic Receptor Model (PARM) software. The concepts behind this method have been described to a limited extent previously.[22,23] In this article, we illustrate the usefulness of PARM using a series of case studies. These studies have shown us that PARM can produce models of good statistical quality that are robust and stable with acceptable predictivity that is as good, and in some cases, better than the popular CoMFA method.

METHODOLOGY

We have discussed the theory of the PARM method previously[23] and present only a brief description here In summary, the PARM method defines 15 kinds of pseudoreceptor atoms likely to be encountered in proteins. A set of grid points is generated around the common surface of the superimposed training-set molecules. Receptor models are

Color Plates for this article are on pages 472–473.

Corresponding author: Jiaju Zhou, Laboratory of Computer Chemistry, Institute of Chemical Metallurgy, Chinese Academy of Sciences, P.O. Box 353, 100080, Beijing, China.

E-mail address: jjzhou@lcc.icm.ac.cn (J. Zhou)

Journal of Molecular Graphics and Modelling 19, 448–454, 2001

655 Avenue of the Americas, New York, NY 10010

1093-3263/01/$–see front matter
PII S1093-3263(00)00104-2

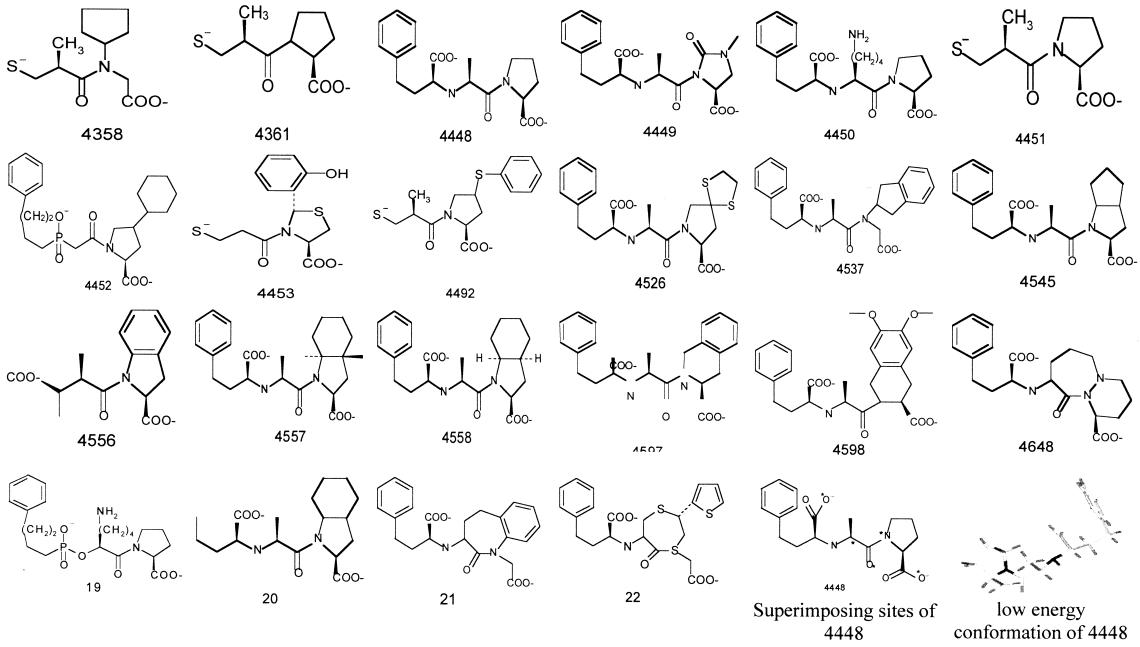


Figure 1. Structures of ACE inhibitors showing atoms used for superimposition and the low-energy. Conformation of 4,448 used as a superimposition template.

made by placing atoms at these points in 3D space to simulate a receptor site and its interactions with the ligand. A genetic algorithm is used to optimize these models and select models with high cross-validated R^2. For each model, a linear relation between log of the bioactivity and interaction energy is established, consistent with the assumed free-energy relationships in ligand binding. Using this relation, the bioactivity of unknown compounds can be predicted. Color Plate 1 shows a schematic of PARM pseudoreceptor modelling process. Using a robust evolutionary optimization process (see "Optimize using GA" in Color Plate 1), in most cases PARM can find a reliable model that can be used to predict bioactivity of unknown compounds. In our case studies, the Tripos 5.0 force field was used to calculate interaction energy between pseudoreceptor and ligand. A distance-dependent dielectric function was used in the energy calculation ($d_{ij} = r$) except where otherwise noted. Partial charges on ligand atoms were calculated by the PM3 semiempirical MO method as implemented in Sybyl 6.0.[24,25] Geometry optimizations, conformer searches, and superimpositions were all carried out using SYBYL 6.0 with default settings. While performing PARM analysis, users can adjust the following parameters: the distance between grid point and its closest atom in the ligands (the "cushion"); the initial number of models (population size); and the maximum number of generations for the genetic evolution calculation. In the CoMFA analyses, all the parameters were the defaults except that the region was extended beyond the van der Waals envelope of all molecules by at least 4.0 Å. The all-orientation searching approach[26] was used to get orientation-independent and optimum CoMFA R_{cv}^2. The training and test sets were chosen to contain maximum molecular diversity and to ensure that the test set was representative of the training set.

Table 1. The 10 Best PARM Pseudoreceptor Models for ACE Inhibitors

Model	B_0	B_1	R	Standard Deviation	R_{cv}^2
1	4.336	−0.052	0.917	0.211	0.799
2	4.798	−0.035	0.906	0.224	0.765
3	4.373	−0.045	0.903	0.228	0.763
4	4.818	−0.043	0.909	0.221	0.760
5	4.154	−0.032	0.894	0.237	0.745
6	4.214	−0.025	0.894	0.237	0.743
7	4.706	−0.040	0.895	0.236	0.738
8	5.029	−0.042	0.895	0.236	0.733
9	4.038	−0.044	0.885	0.247	0.720
10	4.767	−0.033	0.888	0.244	0.717

Activity (log1/C) = $B_0 + B_1 \times$ Energy (kcal/mol).

Table 2. PARM Pseudoreceptor Model Results for ACE Inhibitors

Molecule	Actual log(1/C)	Predicted log(1/C)	Residual	E(kcal/mol)	IC_{50} (nM)
4361	4.5376	4.4003	0.1373	7.673	29 (23 ~ 35)
4451	5.0044	4.8175	0.1869	−2.804	9.9
4492	5.0968	4.6892	0.4076	0.418	8
4448	5.2840	5.3594	−0.0754	−16.416	5.2 (1 ~ 5.2)
4545	5.3768	5.2727	0.1041	−14.237	4.2 (1.5 ~ 4.2)
4597	5.5229	5.5263	−0.0034	−20.606	3
4648	5.7144	5.7618	−0.0474	−26.520	1.93
4537	4.3979	4.7163	−0.3184	−0.265	40
4526	6.0915	5.7511	0.3404	−26.253	0.81
4450	5.7696	5.7765	0.0069	−26.891	1.7
4598	5.8239	5.7013	0.1226	−25.001	1.85 (1.1 ~ 2.6)
4556	4.8928	5.2648	−0.3720	−14.038	12.8
4558	5.0458	5.2566	−0.2108	−13.834	9
4449	5.5686	5.6422	−0.0736	−23.519	2.7
19	4.4437	4.7651	−0.3214	−1.489	36
4452	4.9586	4.8280	0.1306	−3.070	11
		log1/C = 4.706 − 0.040*E, R = 0.895, SD = 0.236, Rcv^2 = 0.738			
*4358	4.9208	4.7647	0.1561	−1.478	12
*4453	4.8962	4.7023	0.1939	0.088	12.7
*20	5.6198	5.6800	0.0602	−24.467	2.4
*21	4.6198	4.7076	−0.0878	−0.045	24 (24 ~ 26)
*22	5.4437	5.5158	−0.0721	−20.344	3.6
*4457	5.3188	5.2736	0.0452	−14.261	4.8

RESULTS AND DISCUSSION

Case Study I: ACE Inhibitors

The angiotensin-converting enzyme inhibitors (ACEI) constitute an important class of antihypertensive drugs.[27] In this case study we used twenty-two compounds whose structures are summarized in Figure 1. We used 16 compounds in the training set and 6 compounds in the test set. All structural and bioactivity data were obtained from literature references.[27,28] We took compound 4,448 as the template compound because of its relatively simple structure and high activity, with the rest of compounds superimposed using the atoms with asterisks. These atoms were proposed as important for interaction with the receptor site.[27] We assumed planar amide bonds and used low-nergy conformations thatt matched the superimposition on compound 4,48. In several cases the S- moiety was superimposed on the second carboxyl group of 4,48. The parameters for the PARM calculation were: population size = 1,00, grid size = 49 times; 49 times; 49 (default value), maximum generations = 2,000, cushion = 0.5 Å. The top 10 models are

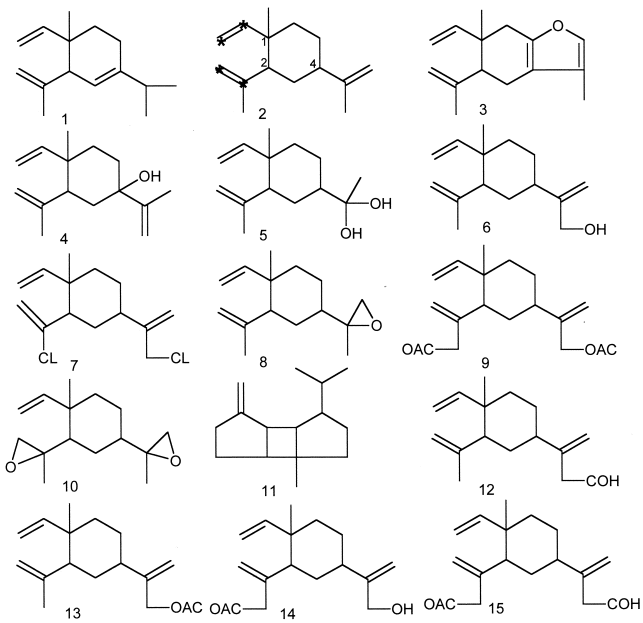


Figure 2. Structures of elemene derivatives.

Table 3. Biological Activity of Elemene Derivatives

Compounds	ID_{50} (μM)	Log $1/ID_{50}$
1	31.8	1.498
2	39.1	1.418
3	108.6	0.964
4	72.6	1.139
5	82.5	1.084
6	25.0	1.602
7	61.7	1.210
8	45.4	1.343
9	18.7	1.728
10	55.0	1.260
11	63.6	1.197

* the unit of the concentration used for $Log(1/ID_{50})$ is mM.

Table 4. The 10 Best PARM Pseudoreceptor Models for Elemene Derivatives

Model	B_0	B_1	R	Standard Deviation	R_{cv}^2
1	1.211	−0.028	0.990	0.037	0.972
2	1.167	−0.025	0.982	0.052	0.948
3	1.185	−0.022	0.977	0.058	0.939
4	1.163	−0.031	0.979	0.055	0.932
5	1.164	−0.026	0.977	0.058	0.931
6	1.123	−0.030	0.974	0.061	0.930
7	1.159	−0.025	0.974	0.062	0.925
8	1.177	−0.021	0.972	0.063	0.925
9	1.101	−0.027	0.970	0.066	0.919
10	1.101	−0.0267	0.976	0.060	0.918

Activity (Log $1/ID_{50}$) = $B_0 + B_1 \times$ Energy(kcal/mol).

presented in Table 1. Model 7 was taken as the final model because its test set performance was the best. The standard deviation of the fit was 0.236 and the cross-validated R_{cv}^2 was 0.738. The calculation results and predicted bioactivities are presented in Table 2, where the first 16 compounds were in the training set and the last 6 compounds (asterisks) were the test set. The maximum error of prediction was less than 0.2 log units.

Case Study II: Elemene Derivatives

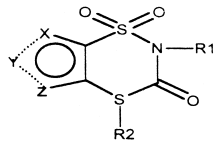
Elemene and its derivatives constitute a useful class of potential anticancer drugs showing efficacy against brain tumors. We analyzed a set of fifteen elemenes whose structures are shown in Figure 2 and biological activities summarized in Table 3. Compounds 1–11 were synthesized and tested by Professor Li Deshang's group and the remaining four compounds are hypothetical compounds selected for synthesis. The target bioactivity was the inhibition (ID_{50}) of erbB2-NIH cells transfected by erbB2 cancer genes (synthesis and biological testing will be published elsewhere.) We used compounds 1–9 as the training set while compounds 10–15 were the test set. Compound 2 (the 1S,2S,4R-stereoisomer) has a simple structure and relatively high activity, so it was used as the template. We carried out a systematic search over its rotatable bonds, followed by optimization, to get the lowest energy conformation (see Color Plate 2). The MINDO3 method in MOPAC, as implemented in Sybyl, was used to calculate partial atomic charges. Five superimposition rules were used with the ComFA method and the best result was obtained by superimposing all molecules using the asterisked atoms (Figure 2, compound 2). When running PARM, the calculation parameters of the genetic algorithm are: population size = 1,500, grid number = 49 × 49 × 49 (default value), maximum generation = 2,000, cushion = 0.5Å. The top 10 models are shown in Table 4. Within these 10 models, the largest standard deviation value was only 0.066 (model 9) and the lowest R_{cv}^2 was still 0.918. The predicted ID_{50} results for compounds 10–15 are presented in Table 5, together with the training results for all 10 models. The maximum error of prediction was 0.179 (model 1, compound 10). 11A and 11B present results for two different conformations. In each column, the highest and the lowest values are italicized. Color Plate 3 shows a pseudoreceptor model with compound 2 as the ligand. The dots in the pseudoreceptor model represent the locations of the15 PARM atom types.[23] The colors follow the normal atom coloring convetions: i.e., red = oxygen atom, blue = nitrogen atom, yellow = sulfur atom, etc.

Table 5. Prediction of Bioactivity for Elemenes 10-15 by the Top 10 PARM models*

	10	11A	11B	12	13	14	15
Model 1	*1.081*	*1.340*	1.243	1.405	1.188	2.010	1.893
Model 2	1.397	1.286	1.235	1.409	1.115	*1.964*	1.870
Model 3	1.288	1.246	*1.178*	1.280	*0.869*	2.083	1.853
Model 4	1.176	1.288	1.200	*1.243*	1.030	2.113	1.810
Model 5	1.354	1.331	*1.279*	1.367	*1.221*	1.994	1.870
Model 6	1.312	1.298	1.256	1.376	1.161	2.087	1.961
Model 7	1.308	1.314	1.231	1.284	0.897	*2.288*	1.986
Model 8	1.350	1.230	1.191	*1.468*	1.200	2.035	*1.730*
Model 9	1.319	*1.215*	1.187	1.357	1.100	2.127	*1.960*
Model 10	*1.429*	1.232	1.253	1.397	1.029	2.061	1.938
Actual	1.260	1.197	1.197	—	—	—	—

* The Unit of activity data is Log($1/ID_{50}$).

Table 6. Anti-HIV-1 Activity and Cytotoxicity of 2,4-Disubstituted 1,1,3-Trioxo-2H,4H-hetero[1,2,4]thiadiazines 16-20



						HIV-1(III^B) molar		
Compound	X	Y	Z	R_1	R_2	EC_{50}	CC_{50}	SI
16a	CH	S	CH	Bn	2–Cl–benzyl	0.10	>119.0	>1190
16b	CH	S	CH	Bn	3–Cl–benzyl	0.8	>119.0	>149
16c	CH	S	CH	Bn	4–Cl–benzyl	1.4	18.6	13
16d	CH	S	CH	Bn	CH_2CO_2Et	7.4	129.6	18
16e	CH	S	CH	2–Cl–benzyl	Me	>446.3	446.3	<1
16f	CH	S	CH	2–Cl–benzyl	Et	7.3	605.3	83
16g	CH	S	CH	3–Cl–benzyl	Me	7.6	>729	>96
16h	CH	S	CH	3–Cl–benzyl	Et	2.1	>700.6	>334
16i	CH	S	CH	3–Cl–benzyl	CH_2CN	0.09	>340	>3778
16j	CH	S	CH	3–Cl–benzyl	$CH_2C{\equiv}CH$	0.1	65.7	657
16k	CH	S	CH	4–Cl–benzyl	Me	>417	417	<1
16l	CH	S	CH	4–Cl–benzyl	Et	>273	273	<1
16m	CH	S	CH	2,6–di–Cl–benzyl	Me	>527.5	527.5	<1
16n	CH	S	CH	2,6–di–Cl–benzyl	Et	>74.1	74.1	<1
16o	CH	S	CH	3-Br-benzyl	CH_2CN	0.09	68.6	762
16p	CH	S	CH	3-F-benzyl	CH_2CN	0.05	93.6	1872
16q	CH	S	CH	3,5–di–Cl–benzyl	CN_2CN	0.3	102.0	340
16r	CH	S	CH	Phenethyl	Et	8.6	330.0	38
16s	CH	S	CH	Phenethyl	CH_2CN	3.6	40.3	11
16t	CH	S	CH	Phenethyl	Bn	10.9	226.0	21
16u	CH	S	CH	2-picolyl	Et	41.4	>387.0	>9
16v	CH	S	CH	2-picolyl	CH_2CN	1.1	>374	>340
16w	CH	S	CH	2-picolyl	Bn	0.4	166.0	415
16x	CH	S	CH	3-picolyl	CH_2CN	0.2	>150.0	>750
17a	CH	CH	S	Bn	Me	>811	>811.0	1
17b	CH	CH	S	Bn	Et	20.5	374	18
17c	CH	CH	S	Bn	n-Pr	>403	403.0	>1
17d	CH	CH	S	Bn	CH_2CN	3.0	219.0	73
17e	CH	CH	S	Bn	Bn	8.3	>650	>78
18a	N	NCH_3	CH	Bn	Me	>498.5	498.5	<1
18b	N	NCH_3	CH	Bn	Et	>320.6	320.6	<1
19a	S	CCl	CH	Bn	CH_2CN	24.5	152.0	6
19b	S	CCl	CH	Bn	Bn	4.5	>119.0	>26
20a	CH	S	CH	Bn	CH_2CN	0.9	502.7	559
20b	CH	S	CH	Bn	Me	2.7	527.9	196
20c	CH	S	CH	Bn	Et	2.5	>775.0	>310
20d	CH	S	CH	Bn	$CH_2C{=}CH$	1.0	>376	>376
20e	CH	S	CH	Bn	Bn	2.2	>650	>296

EC_{50} indicates dose of compound required to achieve 50% protection of MT-4 cells from HIV-1 induced cytopathogenicity, as determined by the MIT method. CC_{50}, dosage required reducing the viability of mock-infected cells by 50%, as determined by the MTT method. All data represent mean value for at least two separate experiments.

Case Study III: TTD

1,1,3-trioxo-2*h*,4*h*-thieno[3,4-e] [1,2,4] thiadiazine (TTD) derivatives[29,30] belong to a new family of nonnucleoside HIV-1 reverse transcriptase (RT) inhibitors. We used PARM to perform a pseudoreceptor 3D-QSAR study of existing TTDs to obtain more insight into the features of the molecules that correlate with anti-HIV-1 activity, and to predict the activities of new TTDs. The thirty-eight TTDs and their activities[29,30] used in this study are listed in Table 6. The lowest-energy

Table 7. Summary of PARM Pseudoreceptor Modelling of TTDs

	PARM			CoMFA		
Different training set	R_{cv}^2	SD	SD_{pred}	R_{cv}^2	SD	SD_{pred}
A	0.627	0.331	0.623	0.614	1.011	0.669
B	0.731	0.647	0.471	0.612	0.886	0.503
C	0.719	0.660	0.411	0.649	0.814	0.434

SD_{pred} indicates Standard Deviation for predicting set.

conformation of **20a**, obtained by systematic search of all rotatable bonds, agreed well with its crystal structure.[29] The other compounds were built by modification of **20a**, with the modified parts subject to systematic search. After the conformational search, all compounds were reoptimized using the Tripos force field molecular mechanics package and partial charges were calculated by AM1 method in MOPAC. Compounds **16u–16w** adopted a conformation whose energies were a little higher (less than 0.25kcal/mol) than their respective lowest-energy ones, but the N atom on picolyl group was at the same side of the ring compared with the chlorine atom on the benzyl groups of **16e** and **16f**. Compound **20a** was chosen as the template molecule for alignment due to its relative simplicity and rigidity. All other compounds were least-squares-fitted to the thiadiazine ring of **20a**. This alignment may best describe the binding sites of TTDs-RT complex, according to the superposition study of some known nonnucleoside inhibitors.[29,31]. In the PARM analyses, the calculation parameters of the genetic algorithm are: population size = 1,000, maximum generations = 4,000, grid number = 60 × 60 × 60 with a cushion of 0.5 Å. Several training and test data sets were created and evaluated to verify the predictive ability of PARM:

A. training on 16 compounds (**16a, 16e, 16h, 16j, 16m–16n, 16o–16p, 16u, 17b–17c, 17e, 18a, 19a, 20a–20b**, and **20c**); B. training on 28 compounds (**16a–16f, 16h, 16j–16p, 16s, 16u, 16w, 17b–17e, 18a, 19a–19b, and 20a–20d**); C. training on 28 compounds (**16a–16b, 16d–16e, 16g–16l, 16n–16o, 16q–16r, 16t–16u, 16w–16x, 17a–17d, 18a–18b, 19b, 20a–20b, 20d**).

All the remaining compounds were relegated to their respective test sets. Results are summarized in Table 7, in which the CoMFA results are also listed for comparison. Table 8 summarizes the results of PARM pseudoreceptor modelling of the three case studies presented here and three additional studies that we have reported earlier.[6] The additional studies cover steroids, potassium channel openers, and pyrethroids.

CONCLUSION

We have illustrated the effectiveness of PARM for pseudoreceptor modelling in comparison with CoMFA. In Table 8, most PARM statistical parameters for training sets are better than the corresponding statistics for CoMFA. In addition, all PARM models gave high R_{cv}^2 values. We observed that the PARM R_{cv}^2 value increased when the number of molecules in the training set decreased, although this may be due to decreased diversity or overfitting. PARM can produce models with high R_{cv}^2 if the evolution generation number is large enough because there are more than 15^{55} possible receptor points in the PARM pseudoreceptor pool. The genetic evolution method can effectively search these large solution spaces to find models with high R_{cv}^2. In addition, a relatively large number of pseudoreceptors can fit the smaller number of training examples. A larger training set is more difficult to fit than a smaller one and gives rise to a much smaller number of pseudoreceptors with high R_{cv}^2. Consequently, for the same number of evolution generations, larger training sets produced lower R_{cv}^2 models. As in Walters' approach,[17] PARM was able to build models with high R_{cv}^2 using a small number of training molecules. PARM exhibited a higher robustness than CoMFA in that it can almost always produce a statistically valid model. CoMFA also appears to be more sensitive to the conformation chosen and occasionally finds a model with low R_{cv}^2 values (even negative values) when the training set contains an outlier. We find PARM is not as sensitive to conformations or outliers. As

Table 8. Summary of PARM Pseudoreceptor Models for Six Studies

	Number of compounds		PARM			CoMFA	
Testing system	Training	Predicting set	R_{cv}^2	SD	SD_{pred}	R_{cv}^2	SD_{pred}
Steroids	21	10	0.806	0.491	0.504	—	0.637
KCO	12	5	0.965	0.345	0.470	—	—
Pyrethoid	21	7	0.810	0.272	0.310	—	—
ACEI	16	6	0.738	0.236	0.013	no result	—
Elemene	9	2	0.972	0.037	0.163	0.769	0.366
TTD	28	10	0.719	0.660	0.411	0.649	0.434

SD_{pred} = (Σ (Yexp − Ycalc)2/N)**1/2; KCO, K+ channel openers; ACEI, Angiotensin, converting enzyme inhibitors; elemene; potential anti-cancer drug; TTD, 1,1,3-trioxo-2H,4H-thieno[3,4-e][1,2,4]thiadiazine derivatives.

expected, results for the test set were less impressive than for the training set but the PARM prediction results were comparable to those of CoMFA. PARM lacks the ability to evaluate the quality of molecular alignment and conformation selection, which are important for 3D-QSAR. However, this can be performed by other methods. The CoMFA calculations generate a 3D contour map that can guide design of new molecules with high bioactivity. Similarly, PARM generates pseudoatom models around the ligands. We can obtain useful SAR information from these models. For example, a negatively charged receptor atom means that a positively charged ligand group nearby would be favorable for high activity; a void receptor atom means that bulky substituents may be desirable. However, we have not systematically defined a set of rules for this kind of SAR information. This will be done in future work. In general, The PARM is a practical utility for drug design because of its robustness, stability, and minimal requirements for experimental data.

REFERENCES

1 Böhm, H.J. The computer program LUDI: a new method for the de novo design of enzyme inhibitors. *J. Comput. Aided Mol. Des.* 1992, **6**, 61–78

2 Böhm, H.J. LUDI: rule-based automatic design of new substituents for enzyme inhibitor leads. *J. Comput. Aided Mol. Des.* 1992, **6**, 593–606

3 Moon, J.B., and Howe, W.J. Computer design of bioactive molecules: a method for receptor-based de novo ligand design. *Proteins: Struct. Funct. Genet.* 1991, **11**, 314–328

4 Jiang, H., Hu, Z., Chen, J., Gu, J., Zhu, W., Chen, K., and Ji, R. Computer modeling for interaction of ligands and receptors and their application in drug design.*Prog. Chem.* (in Chinese), 1998, **10**, 427–441

5 Sikorski, A., Kolinski, A., and Skolnick, J. Computer simulations of de novo designed helical proteins. *Biophys. J.* 1998 Jul, **75:1**, 92–105

6 Miranker, A., and Karplus, M. An automated method for dynamic ligand design *Proteins* 1995, **23:4**, 472–490

7 Murray, C.W., Clark, D.E., and Byrne, D.G. PRO_LIGAND: an approach to de novo molecular design. 6. Flexible fitting in the design of peptides. *J. Comput. Aided Mol. Des.* 1995, **9:5**, 381–395

8 Pearlman, D.A., and Murcko, M.A. CONCERTS: Dynamic connection of fragments as an approach to de novo ligand design. *J. Med. Chem.*, 1996, **39:8**, 1651–1663

9 Gehlhaar, D.K., Moerder, K.E., Zichi, D., Sherman, C.J., Ogden, R.C., and Freer, S.T. CONCERTS: Dynamic connection of fragments as an approach to de novo ligand design. *J Med Chem*, 1995, **38:3**, 466–472

10 Todorov, N.P., and Dean, P.M. A branch-and-bound method for optimal atom-type assignment in de novo ligand design. *J. Comput. Aided Mol. Des.* 1998, **12:4**, 335–349

11 Johnson, M.S., Srinivasan, N., Sowdhamini, R., and Blundell, T.L. Knowledge-based protein modeling. *Cric. Rev. Biochem. Mol. Biol.* 1994, **29**, 1–68

12 Rost, B. TOPITS: Threading one-dimensional predictions into three-dimensional structures. *Ismb* 1995, **3**, 314–321

13 Greer, J. Comparative modeling of homologous proteins. *Method. Enzym.* 1991, **202**, 239–252

14 Thornton, J.M., and Swindells, M.B. Molecular Structures in Biology. Diamond, R., Koetzle, T.F., Drout, K., and Richardson, J.S., eds., Oxford, New York, 1993

15 Cramer III, R.D., Patterson, D.E., and Bunce, J.D. Comparative molecular field analysis (CoMFA). Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* 1988, **110**, 5959–5967

16 Jain, A.N., Dietterich, T.G., Lathrop, R.H., Chapman, D., Critchlow, R.E., Bauer, B.E., Webster, T.A., and Lozano-Perez, T. A shape-based machine learning tool for drug design. *J. Comput. Aided Mol. Des.* 1994, **8**, 635–652

17 Walters, D.E., and Hinds, R.M. Genetically evolved receptor models: A computational approach to construction of receptor models. *J. Med. Chem.* 1994, **37**, 2527–2536

18 Hahn, M. Receptor surface models. 1. Definition and construction. *J. Med. Chem.* 1995, **38**, 2080–2090

19 Hahn, M. Receptor surface models. 2. Application to quantitative structure–activity relationships studies. *J. Med. Chem.* **1995**, **38**, 2091–2102

20 Vedani, A., Zbinden, P., Snyder, J.P., and Greenidge, P.A. Pseudoreceptor modeling: The construction of three-dimensional receptor surrogates. *J. Am. Chem. Soc.* 1995, **117**, 4987–4994

21 Vedani, A., and Zbinden, P. Quasi-atomistic receptor modeling. A bridge between 3D QSAR and receptor fitting. *Pharm. Acta. Helv.* 1998, **73:1**, 11–18

22 Chen, H. *Receptor Model Approach in Drug Design*, Doctoral Thesis, Institute of Chemical Metallurgy, Chinese Academy of Sciences, Beijing, 1998

23 Chen, H., Zhou, J., and Xie, G. PARM: A genetic evolved algorithm to predict bioactivity. *J. Chem. Inf. Comput. Sci.* 1998, **38**, 243–250

24 *Sybyl Theory Manual*. SYBYL Molecular Modeling Software, Version 6.0, 1992

25 *SYBYL Molecular Modeling System*, Tripos Associate. St Louis, USA

26 Wang, R., Gao, Y., Liu, L., and Lai, L. All-orientation search and all-placement search in comparative molecular field analysis. *J. Mol. Model.* 1998, **4**, 276–283

27 Unger, T., and Gohlke, P. Converting enzyme inhibitors in cardiovascular therapy: current status and future potential. *Cardiovasc. Res.* 1994, **28:2**, 146–158

28 Wang, Z. *Drug Structure Universal Set of The Time*, Beijing Science and Technology Press, Beijing, 1993

29 Witvrouw, M., Arranz, M.E., Pannecouque, C., Declercq, R., Jonckheere, H., Schmit, J.C., Vandamme, A.M., Diaz, J.A., Ingate, S.T., Desmyter, J., Esnouf, R., Van Meervelt, L., Vega, S., Balzarini, J., and De Clercq, E. 1,1,3-Trioxo-2H,4H-thieno[3,4-e][1,2,4]thiadiazine (TTD) derivatives: A new class of nonnucleoside human immunodeficiency virus type 1 (HIV-1) reverse transcriptase inhibitors with anti-HIV-1 activity. *Antimicrob. Agents Chemother.* Mar. 1998, 618–623

30 Arranz, E., Diaz, J.A., Ingate, S.T., Witvrouw, M., Pannecouque, C., Balzarini, J., De Clercq, E., and Vega, S. 1,1,3-trioxo-2H,4H-thieno[3,4-e][1,2,4]thiadiazine derivatives as non-nucleoside reverse transcriptase inhibitors that inhibit human immunodeficiency virus type 1 replication. *J. Med. Chem.* 1998, **41**, 4109–4117

31 Ren, J., Esnouf, R., Garman, E., Somer, D., Ross, C., Kirby, I., Keeling, J., Darby, G., Jones, Y., Stuart, D., and Stammers, D. High resolution structures of HIV-1 RT from four TR-inhibitor complexes. *Struct. Biol.* 1995, **2**, 293–302