Comparing the Quality and Predictiveness between 3D QSAR Models Obtained from Manual and Automated Alignment

Anu J. Tervo,*[,†,‡] Tommi H. Nyrönen,[‡] Toni Rönkkö,[†] and Antti Poso[†]

Department of Pharmaceutical Chemistry, University of Kuopio, P.O. Box 1627, 70211 Kuopio, Finland, and CSC − Scientific Computing Ltd., P.O. Box 405, 02101 Espoo, Finland

Received October 19, 2003

A set of 113 flexible cyclic urea inhibitors of human immunodeficiency virus protease (HIV-1 PR) was used to compare the quality and predictive power of CoMFA and CoMSIA models for manually or automatically aligned inhibitor set. Inhibitors that were aligned automatically with molecular docking were in agreement with information obtained from existing X-ray structures. Both alignment methods produced statistically significant CoMFA and CoMSIA models, with the best q^2 value being 0.649 and the best predictive r^2 being 0.754. The manual alignment gave statistically higher values, whereas the automated alignment gave more robust models for predicting the activities of an external inhibitor set. Both models utilized similar amino acids in the HIV-1 PR active site, supporting the idea that hydrogen bonds form between an inhibitor and the backbone carbonyl oxygens of Gly48 and Gly48′ and also the backbone NH group of Asp30, Gly48, Asp29′, and Gly48′ of the enzyme. These results suggest that an automated inhibitor alignment can yield predictive 3D QSAR models that are well comparable to manual methods. Thus, an automated alignment method in creating 3D QSAR models is encouragable when a well-characterized structure of the target protein is available.

INTRODUCTION

Understanding the structural properties and features affecting the biological activity of a drug molecule is important in the ongoing process of drug design. After the initial characterization of a drug by physical means, additional quantitative information can be obtained by using three-dimensional structure−activity relationship (3D QSAR) methods.[1] However, generating 3D conformations and alignment for compounds used in 3D QSAR modeling is a difficult and time-consuming process, especially when the compounds are large in size and contain several rotatable bonds. Moreover, there is a risk of introducing user subjectivity in manual alignment. Thus, an automated generation of alignment for the 3D QSAR model could be beneficial. In general, automated molecular docking offers a more objective way to create conformations and alignment and to include the information obtained from amino acids of the active site into the process of model building. There are published reports of successful 3D QSAR models where conformations and alignment have been obtained by automated molecular docking.[2−7] However, comparisons of the quality and information value of 3D QSAR models based on manual and automated alignment methods have been given little attention.

In the present study, the information value and statistical quality of 3D QSAR models was compared between automated and manually generated alignments. The automated generation was obtained by using a molecular docking procedure. Manually generated compounds were constructed by using the bioactive conformation from the X-ray structure of a cocrystallized ligand as a template, obtained from the Protein Data Bank.[8]

3D QSAR models for the automated and manual compound alignment were created using CoMFA[9] and CoMSIA[10] methods. Both methods produce information on the favored and disfavored properties of the examined compound set. Compound alignment plays a crucial role in the creation of CoMFA models,[9] whereas CoMSIA models are less alignment-dependent.[11]

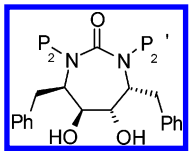
The studied compound set consisted of 113 conformationally flexible human immunodeficiency virus protease (HIV-1 PR) cyclic urea inhibitors, obtained from the literature.[12−14] HIV-1 PR is an aspartic protease encoded by the human immunodeficiency virus (HIV).[15−17] Its role has been found to be essential for viral infectivity.[18] The design of HIV-1 PR inhibitors has been identified as one of the primary targets for structure-based drug design against HIV infection.[19] HIV-1 PR is a C_2-symmetrical dimer structure, and its active site is located between two subunits. The key active site is thought to include two aspartate residues, Asp25/Asp25′, at the "floor" of the active site. The other side, "roof", the active site contains the important hydrophobic residues Ile50/Ile50′ and a structural water molecule. The studied inhibitors can also occupy four symmetrical hydrophobic pockets; S1/S1′ and S2/S2′.[19−21] Although S2 and S2′ are mainly hydrophobic, both hydrophobic and hydrophilic P2/P2′ substituents can enter these sites.[17,19]

Several 3D QSAR studies of HIV-1 PR inhibitors have been reported.[7,22−32] In one of these reports, Debnath[28] used a similar inhibitor set to create a series of CoMFA models. Debnath obtained conformations and alignment by manually

* Corresponding author phone: +358-9-4572228; fax: +358-9-4572302; e-mail: anu.tervo@csc.fi.
† University of Kuopio.
‡ CSC − Scientific Computing Ltd.

10.1021/ci0342268 CCC: $27.50
Published on Web 02/25/2004

Table 1. Structures of Inhibitors **1**−**52**[a]



compd	P_2/P_2'	compd	P_2/P_2'
1	methyl	**27**	*p*-fluorobenzyl
2	ethyl	**28**	*o*-chlorobenzyl
3	*n*-propyl	**29**	*m*-chlorobenzyl
4	*n*-butyl	**30**	*p*-chlorobenzyl
5	*n*-hexyl	**31**	*p*-bromobenzyl
6	*n*-heptyl	**32**	*m*-methylbenzyl
7	$CH_2CH_2OCH_3$	**33**	*p*-methylbenzyl
8	$CH_2CH_2OCH_2CH_2OCH_3$	**34**	*p*-(trifluoromethyl)benzyl
9	*i*-butyl	**35**	*o*-methoxybenzyl
10	*i*-pentyl	**36**	*m*-methoxybenzyl
11	*i*-hexyl	**37**	*m*-nitrobenzyl
12	*i*-heptyl	**38**	*m*-iodobenzyl
13	*i*-octyl	**39**	*p*-(hydroxymethyl)benzyl
14	allyl	**40**	*m*-(hydroxymethyl)benzyl
15	isoprenyl	**41**	*m*-hydroxybenzyl
16	$CH_2CH_2OCHCH_2$	**42**	*m*-aminobenzyl
17	3-propynyl	**43**[a]	*n*-pentyl
18	cyclopropylmethyl	**44**[a]	$CH_2CH_2OCH_2CH_3$
19	cyclobutylmethyl	**45**[a]	2-methylpropen-3-yl
20	cyclohexylmethyl	**46**[a]	cyclopentylmethyl
21	benzyl	**47**[a]	4-picolyl
22	2-picolyl	**48**[a]	*o*-fluorobenzyl
23	3-picolyl	**49**[a]	*m*-bromobenzyl
24	α-naphthylmethyl	**50**[a]	*m*-(trifluoromethyl)benzyl
25	*β*-naphthylmethyl	**51**[a]	*p*-methoxybenzyl
26	*m*-fluorobenzyl	**52**[a]	*p*-hydroxybenzyl

[a] These compounds belong to the test set.

constructing inhibitors at the active site of HIV-1 PR. The construction was followed by a systematic conformational search and minimization. That report enabled us to further evaluate our CoMFA results. Subsequently, the authenticity of our 3D QSAR models based on an automated inhibitor alignment was explored by examining if the information obtained from CoMFA and CoMSIA fields could be mapped back to the amino acids of HIV-1 PR active site.

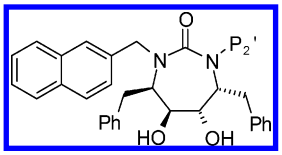
MATERIALS AND METHODS

The inhibitors were divided into a training set of 93 compounds for generating the 3D QSAR models and a test set of 20 compounds. An external test set was needed to objectively validate the quality of the final models, and therefore the test set was selected to represent different structural elements and activity intensities among the studied 113 HIV-1 PR inhibitors. The structures of the inhibitors are presented in Tables 1−4 (compounds marked with an *a* belong to the test set).

The activity values of the inhibitors were in the literature[12,13] reported as K_i (nm). To derive the 3D QSAR models, the log $1/K_i$(M) activity values were calculated and are presented in Table 5. Unless noted otherwise, molecular modeling was performed using Sybyl 6.8.[33]

Manual Inhibitor Alignment Procedure. Manual inhibitor alignment was created by using the structure and conformation of the cocrystallized inhibitor [4*R*-(4α,5α,6*β*,7*β*)]-3,3′-[[tetrahydro-5,6-dihydroxy-2-oxo-4,7-bis(phenylmethyl)-1*H*-1,3-diazepine-1,3(2*H*)-diyl]bis(methylene)]bis[*N*-2-thiazolylbenzamide] (XV638 in Figure 1; also inhibitor **90** in our set) of the HIV-1 PR X-ray structure (PDB ID: 1qbr[21]) as a template. The 1qbr structure was selected among several existing HIV-1 PR structures in the Protein Data Bank because the cocrystallized inhibitor was structurally most similar to the compounds in the studied inhibitor set. The manual superposition was accomplished in two stages. First, structures of all inhibitors were constructed by modifying XV638, and minimizing each inhibitor for 1000 cycles with the BFGS[34−37] method, using the MMFF94s force field[38] and MMFF94 charges.[39] The energy-minimized inhibitors were superimposed to match the cyclic urea ring of XV638. Second, the superimposed inhibitors were further minimized for 100 cycles with the Steepest Descent method, using the MMFF94s force field and MMFF94 charges. The generated

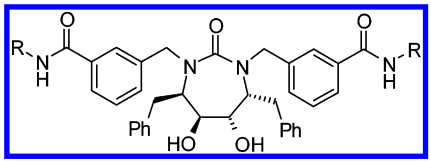
Table 2. Structures of Inhibitors **53**−**64**[a]



compd	P_2'	compd	P_2'
53	*n*-butyl	**59**	*p*-fluorobenzyl
54	allyl	**60**	*p*-(hydroxymethyl)benzyl
55	cyclopropylmethyl	**61**	*m*-aminobenzyl
56	cyclopentyl	**62**	*m*-hydroxybenzyl
57	benzyl	**63**[a]	*n*-propyl
58	4-picolyl	**64**[a]	3-picolyl

[a] These compounds belong to the test set.

Table 3. Structures of Inhibitors **65**−**98**[a]



compd	R	compd	R
65	H	**82**	2-(4,6-di-CH_3-pyridinyl)
66	NH_2	**83**	2-(5-Cl-pyridinyl)
67	OH	**84**	2-(3,5-di-Cl-pyridinyl)
68	OCH_3	**85**	2-(5-Br-pyridinyl)
69	CH_3	**86**	2-(4-CH_3-pyrimidinyl)
70	$CH(CH_3)_2$	**87**	2-(5-CF_3-pyridinyl)
71	$CH_2CH_2CH_3$	**88**	2-pyrimidinyl
72	$CH_2CH_2CH_2CH_3$	**89**	4-methyl-2-oxazolyl
73	$C(CH_3)_3$	**90**	2-thiazolyl
74	CH_2CF_3	**91**	5-methylthiazol-2-yl
75	CH_2CN	**92**	2-imidazolyl
76	4-pyridinyl	**93**	2-benzimidazolyl
77	3-pyridinyl	**94**[a]	CH_2CH_3
78	2-pyridinyl	**95**[a]	$CH_2C_3H_5$
79	2-(3-CH_3-pyridinyl)	**96**[a]	benzyl
80	2-(4-CH_3-pyridinyl)	**97**[a]	2-(6-CH_3-pyridinyl)
81	2-(5-CH_3-pyridinyl)	**98**[a]	4-methylthiazol-2-yl

[a] These compounds belong to the test set.

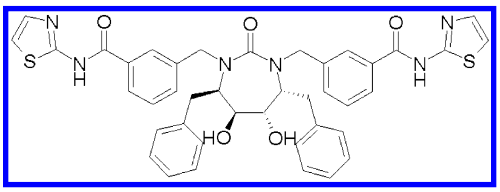
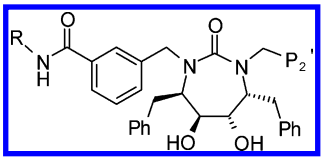


Figure 1. Structure of the cocrystallized inhibitor XV638 in the X-ray structure 1qbr.

Table 4. Structures of Inhibitors **99**−**113**[a]



compd	R	P_2'	compd	R	P_2'
99	2-pyrazinyl	3,5-dimethoxyphenyl	**107**	1,1-dimethylethyl	3-aminophenyl
100	2-(6-CH_3-pyridinyl)	3,5-dimethoxyphenyl	**108**	2-benzimidazolyl	3-aminophenyl
101	2-pyridinyl	3,5-dimethoxyphenyl	**109**	2-imidazolyl	3-aminophenyl
102	2-pyrazinyl	3-methoxyphenyl	**110**	CH_2CN	benzoylglycin-3-yl
103	2-(5-CH_3-pyridinyl)	3-methoxyphenyl	**111**[a]	2-(5-CH_3-pyridinyl)	3,5-dimethoxyphenyl
104	2-pyridinyl	3-methoxyphenyl	**112**[a]	2-(6-CH_3-pyridinyl)	3-methoxyphenyl
105	2-(5-CH_3-pyridinyl)	3-aminophenyl	**113**[a]	2-pyrazinyl	3-aminophenyl
106	2-pyridinyl	3-nitrophenyl			

[a] These compounds belong to the test set.

conformations were aligned again to match the common seven-membered cyclic urea ring and its benzyl and hydroxyl substituents. Finally, Gasteiger−Hückel[40,41] charges for building the 3D QSAR models were calculated for all inhibitors.

Automated Inhibitor Alignment Procedure. The X-ray structure of HIV-1 PR (PDB ID:1qbr) was used for docking simulations. Prior to docking, the cocrystallized inhibitor XV638 was removed from the structure, and hydrogen atoms were added by using Sybyl Biopolymer. The carbonyl oxygen of the seven-membered cyclic urea ring, which exists in all inhibitors in the studied set, is designed to displace a structural water molecule from the active site of HIV-1 PR.[20] Therefore, this water molecule was not introduced into the active site.

The inhibitor structures generated for the manual alignment offered reasonable starting conformations for docking; thus, they were docked to the active site of HIV-1 PR using the GOLD program, version 1.2.[42] The docking procedure was repeated 10 times for each inhibitor using default parameters. As we wanted to generate the alignment as automatically as possible, the representative docked conformation for each inhibitor was selected from the docking results on the basis of their rank score (shown in Table 5). Finally, the representative conformation of each inhibitor docked with GOLD was minimized inside the active site for 100 steps with the Steepest Descent method using the Tripos force field,[43] Gasteiger−Hückel charges for the inhibitors, and Kollman all-atom charges[44] for the HIV-1 PR. The enzyme was kept rigid during the minimization, allowing only the inhibitor to relax.

To check the accuracy of docking and the quality of the chosen conformations, we calculated the heavy atom root-mean-square deviation (RMSD) between the docked and cocrystallized conformations from the existing X-ray structures. HIV-1 PR structures (PDB ID: 1hwr;[20] 1dmp;[45] 1qbr) cocrystallized with inhibitors **14**, **42**, and **90**, respectively, were retrieved from the Protein Data Bank. The enzymes were superimposed with 1qbr coordinates on the basis of C_α atoms using the VERTAA program[46] in the BODIL modeling environment.[47] After superimposition, cocrystallized inhibitors were extracted from the complexes, and the heavy atom RMSD was calculated.

CoMFA and CoMSIA Analyses. The training set of 93 inhibitors was used to create 3D QSAR models with both CoMFA[9] and CoMSIA[10] methods to explore contributions to inhibitory activity. 3D QSAR models were built using the inhibitor alignment obtained from both manual and automated alignment procedures.

Steric and electrostatic CoMFA descriptor fields were calculated in a 3D cubic box extending 4 Å beyond the aligned inhibitors. Steric and electrostatic fields were calculated by using an sp^3 carbon probe atom with a +1 charge and 1.5 Å grid spacing. The Tripos Standard field with a cutoff of 30.0 kcal/mol was used for both steric and electrostatic fields. Two different 3D QSAR models were built using the CoMSIA method. The first model was generated by using steric and electrostatic fields, and the second model was generated by using hydrogen bond donor and acceptor fields. CoMSIA models were calculated by applying default settings in Sybyl.

The 3D QSAR models that were based on manually aligned inhibitors were referred to as Model 1M (CoMFA), Model 2M (CoMSIA using steric and electrostatic fields), and Model 3M (CoMSIA using hydrogen bond donor and acceptor fields). Corresponding 3D QSAR models based on automatically aligned inhibitors were referred to as Models 1A, 2A, and 3A, respectively.

Activity values (log $1/K_i$(M)) were used as dependent variables in the partial least squares (PLS)[48] statistical analysis of all models. The predictive value of the PLS models was evaluated by using the leave-one-out (LOO) cross-validation method. LOO was chosen as a cross-validation method to enable the comparison of the results with Debnath's report,[28] where the LOO method was utilized. A minimum σ value (standard deviation threshold) of 2.0 kcal/mol was utilized to improve the signal-noise ratio. The final non-cross-validated models were developed by using the optimal number of components that had both the highest q^2 value and the smallest value of standard error of predictions (Spress).

Model Validation. The predictive power of all CoMFA and CoMSIA models was further validated by using an external test set (inhibitors marked with an *a* in Tables 1−4). The inhibitors in the test set were given exactly the same pretreatment as the inhibitors in the corresponding training sets. The correlation between experimental and predicted activity for all models was calculated as a predictive r^2 value (pred-r^2). The results of the CoMFA models (Models 1M

Table 5. Experimental Activities, GOLD Rank Scores, and Predicted Activities of CoMFA and CoMSIA Models

			Model									Model					
compd	activity[a]	gold[b]	1M	2M	3M	1A	2A	3A	compd	activity[a]	gold[b]	1M	2M	3M	1A	2A	3A
1	5.24	63.18	6.67	6.61	8.01	6.04	6.91	7.80	**58**	8.16	74.77	8.85	8.10	8.22	8.54	8.86	8.21
2	7.00	65.68	6.90	6.88	8.04	6.81	7.04	7.86	**59**	8.44	56.27	8.38	8.23	8.09	8.18	8.85	8.82
3	8.10	68.44	7.76	7.63	8.10	7.88	7.21	7.95	**60**	9.03	56.15	8.57	8.73	8.83	9.17	8.94	9.41
4	8.85	77.15	8.10	8.09	8.06	7.92	7.35	7.94	**61**	9.00	55.75	8.85	8.86	8.85	8.84	9.44	9.11
5	8.34	72.68	8.20	8.28	8.08	8.65	7.70	8.02	**62**	9.48	66.91	8.59	8.67	8.58	9.35	8.74	9.20
6	6.59	78.57	6.89	6.59	8.10	7.27	7.67	8.15	**63**[c]	8.96	75.88	9.09	8.85	8.14	8.90	8.37	8.46
7	6.10	70.21	6.46	6.57	5.94	6.46	6.66	6.07	**64**[c]	8.28	69.93	8.48	8.62	8.22	9.04	9.14	8.12
8	5.11	73.02	5.35	5.38	4.65	4.99	6.73	4.93	**65**	10.41	90.14	10.07	10.57	10.09	11.03	9.97	10.70
9	7.31	66.85	7.75	7.82	8.03	7.75	7.27	7.92	**66**	10.74	99.09	10.77	11.16	10.21	10.46	9.83	10.82
10	7.92	77.11	7.94	7.89	8.10	7.99	7.45	7.93	**67**	10.70	100.89	10.65	10.88	10.58	10.70	9.99	10.53
11	8.15	79.48	8.33	8.21	8.11	8.14	7.57	7.91	**68**	10.35	92.50	10.24	10.42	9.68	10.28	10.18	10.34
12	7.52	73.45	7.99	8.05	8.11	7.36	7.53	7.84	**69**	10.18	92.04	10.03	10.18	9.82	10.46	10.03	9.51
13	6.96	80.05	7.35	7.35	8.11	6.72	7.51	8.15	**70**	9.24	93.27	9.37	9.67	9.96	9.39	9.74	9.64
14	8.28	70.19	7.82	7.77	8.12	7.81	7.62	7.91	**71**	9.44	94.86	8.61	9.43	9.64	9.43	9.66	9.49
15	8.74	75.70	8.88	8.80	8.20	8.54	7.83	8.14	**72**	9.37	101.44	9.11	9.22	9.48	9.58	9.91	9.38
16	7.22	71.51	6.76	6.70	7.95	7.19	6.50	7.89	**73**	8.62	86.88	8.15	8.46	8.74	8.75	9.65	9.32
17	7.66	69.70	7.54	7.94	8.13	7.65	7.55	8.02	**74**	9.68	94.75	9.09	9.64	9.72	9.71	10.06	9.75
18	8.68	75.04	7.85	7.80	8.10	8.36	7.51	8.08	**75**	10.20	101.59	10.29	10.34	10.24	10.41	10.87	9.85
19	8.89	54.03	7.67	7.98	8.00	9.18	7.69	7.91	**76**	9.39	103.22	9.55	9.69	9.84	9.16	9.99	9.47
20	7.43	46.66	7.76	8.22	8.00	7.85	7.68	7.77	**77**	9.54	107.48	10.06	9.97	9.70	9.36	9.79	9.76
21	8.52	83.93	8.01	8.20	8.14	7.89	8.55	7.99	**78**	10.37	110.30	10.26	10.13	10.23	10.88	10.81	10.65
22	6.84	85.81	6.61	6.90	7.01	7.13	8.52	7.23	**79**	9.59	111.20	10.04	9.66	10.12	9.54	9.70	9.58
23	8.01	86.21	8.80	8.97	8.19	8.44	8.86	8.02	**80**	10.57	113.71	10.06	9.96	10.16	10.85	10.26	10.44
24	7.07	94.22	7.79	7.66	8.00	7.29	8.50	7.96	**81**	10.96	118.23	10.53	10.23	10.21	11.32	10.65	10.67
25	9.51	86.49	9.08	9.19	8.20	9.58	8.54	7.68	**82**	10.80	118.02	10.90	10.43	10.67	11.30	10.55	10.83
26	8.52	87.53	9.00	9.21	8.14	8.76	8.51	8.10	**83**	10.92	111.18	10.46	10.15	10.40	10.65	10.56	10.38
27	8.85	86.82	8.57	8.31	8.16	8.11	8.23	8.04	**84**	9.61	107.71	10.35	10.10	10.36	9.76	9.63	9.30
28	6.62	95.26	7.24	6.60	7.94	6.86	8.06	7.91	**85**	10.46	111.80	10.49	10.15	10.33	10.30	10.49	10.46
29	9.05	85.06	9.13	9.13	8.12	8.82	8.67	8.12	**86**	9.94	115.58	10.41	10.15	10.64	9.64	10.14	10.21
30	8.28	78.02	8.34	8.26	8.09	8.43	8.65	8.01	**87**	10.07	118.16	9.99	10.15	10.18	10.08	10.11	10.72
31	7.57	72.60	8.40	8.22	8.10	8.18	8.81	8.27	**88**	9.82	111.52	10.32	10.33	10.68	10.11	10.27	10.48
32	8.15	84.98	8.78	8.56	8.13	8.37	8.86	8.20	**89**	10.19	111.48	9.64	9.82	10.00	10.13	10.22	10.22
33	8.24	63.60	8.02	8.18	8.13	8.40	8.90	8.04	**90**	10.57	105.60	10.77	10.56	10.22	10.32	10.76	10.50
34	7.29	54.32	7.13	7.22	8.07	7.17	8.52	8.09	**91**	10.85	117.03	11.43	10.93	10.46	10.72	10.91	10.84
35	5.73	90.63	6.26	5.71	8.03	5.79	8.12	7.85	**92**	10.85	105.17	10.69	11.01	10.78	10.59	10.64	10.43
36	8.80	96.59	8.66	8.45	8.16	8.82	9.25	8.08	**93**	10.62	119.72	10.82	10.63	10.54	10.46	10.15	10.43
37	8.55	94.77	8.94	8.42	8.67	8.52	8.44	8.84	**94**[c]	9.68	100.86	9.78	9.88	9.76	9.99	9.98	9.69
38	9.38	92.15	8.92	8.70	8.13	9.07	8.67	8.19	**95**[c]	9.13	103.83	9.08	9.43	8.73	9.40	9.73	9.58
39	9.47	64.48	9.19	9.13	9.58	9.26	9.08	9.49	**96**[c]	9.37	110.67	9.29	9.77	9.74	9.63	10.25	9.81
40	9.85	85.01	9.32	9.57	9.63	10.05	9.04	9.94	**97**[c]	10.70	112.95	10.34	10.14	10.26	11.32	10.72	10.88
41	9.92	85.22	9.04	9.49	9.10	9.73	8.75	9.89	**98**[c]	10.60	117.29	11.18	10.93	10.49	9.68	10.15	9.95
42	9.55	89.02	8.68	8.88	9.55	8.72	8.57	9.45	**99**	8.60	88.00	9.19	9.05	9.70	8.46	8.84	8.30
43[c]	8.80	76.01	7.37	7.06	8.10	8.02	7.65	7.81	**100**	9.15	86.68	9.26	9.15	9.66	9.31	9.09	9.51
44[c]	5.96	69.97	6.85	6.95	5.79	6.50	6.86	6.35	**101**	9.07	86.76	9.37	9.20	9.66	8.94	9.00	8.92
45[c]	8.14	75.39	7.43	7.44	8.02	7.29	7.30	7.95	**102**	10.42	106.03	9.67	10.21	9.66	10.48	9.44	9.42
46[c]	8.37	60.20	7.09	7.63	7.96	7.99	7.61	7.81	**103**	10.16	103.43	9.87	10.34	9.70	10.25	10.56	9.85
47[c]	7.05	86.98	8.64	7.88	8.34	8.55	9.03	8.28	**104**	10.33	99.19	9.68	10.17	9.61	10.37	10.29	9.90
48[c]	7.47	81.66	7.40	6.22	8.11	7.68	8.01	8.09	**105**	10.12	95.93	10.20	10.06	10.28	9.83	9.93	10.02
49[c]	8.85	87.77	9.38	9.14	8.11	8.94	8.66	8.10	**106**	10.02	89.76	9.85	9.92	10.17	9.65	9.23	10.39
50[c]	7.66	83.21	10.40	8.61	8.10	9.02	8.39	8.07	**107**	9.39	95.15	9.42	9.21	10.05	9.50	9.48	9.64
51[c]	6.80	74.27	8.13	7.48	8.17	7.60	8.45	7.88	**108**	10.64	103.12	11.22	11.21	10.79	10.15	10.43	10.81
52[c]	9.92	83.02	7.26	7.51	8.88	8.48	8.45	9.03	**109**	10.92	103.74	10.70	10.98	10.56	10.79	10.17	10.94
53	9.22	77.80	9.24	9.11	8.05	9.02	8.29	8.25	**110**	10.62	94.25	10.65	10.76	10.60	10.69	10.14	10.34
54	8.85	66.57	9.54	9.36	8.22	9.00	8.62	8.68	**111**[c]	8.72	92.56	9.42	9.20	9.66	9.14	9.27	8.91
55	8.82	77.77	9.50	9.29	8.13	8.73	8.30	8.16	**112**[c]	10.28	99.63	9.77	10.28	9.63	10.05	9.47	9.18
56	9.55	68.81	9.84	9.53	8.14	9.14	8.37	8.22	**113**[c]	10.80	96.13	9.89	9.89	10.32	9.66	9.27	10.50
57	8.64	67.93	9.13	8.88	8.15	8.87	8.99	8.20									

[a] Activity expressed as log $1/K_i$(M). [b] GOLD Rank Scores. [c] These compounds belong to the test set.

and 1A) were further compared with the best model reported by Debnath.[28]

RESULTS AND DISCUSSION

The information value and statistical quality of 3D QSAR models based on an automatically generated alignment were compared to models where the alignment was generated manually. In total, six 3D QSAR models were created from two inhibitor sets containing 113 HIV-1 PR inhibitors. Inhibitors were then subjected to 3D QSAR analysis using CoMFA and CoMSIA methods, and the statistical quality and information value of these models were also evaluated.

Automated Inhibitor Alignment. The automated molecular docking was carried out using the docking program GOLD and resulted in 10 conformations for each inhibitor. As we wanted to generate the alignment as automatically as possible, the docked conformation for each inhibitor was chosen on the basis of its GOLD rank score. GOLD scoring

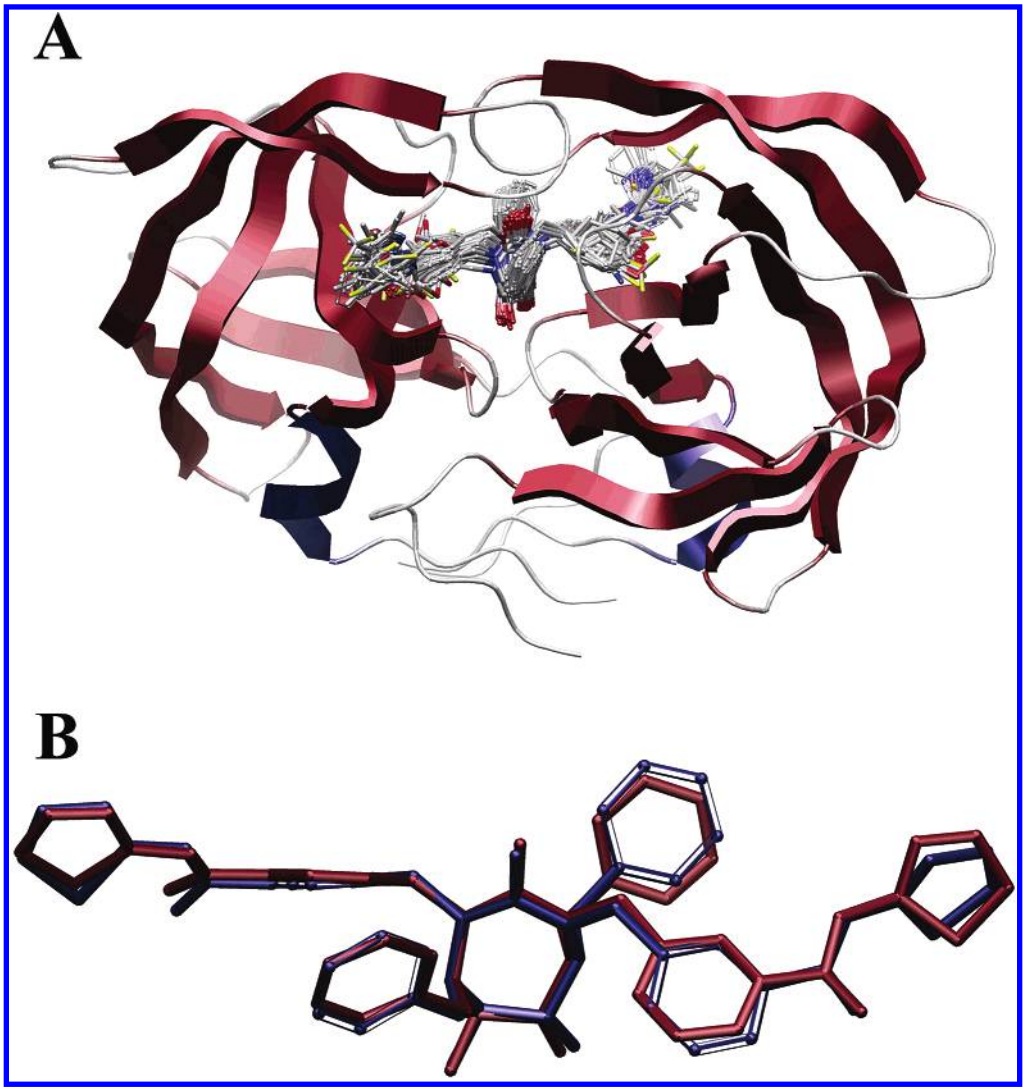


Figure 2. (A) Conformations and alignment of the docked inhibitors in the active site of HIV-1 PR. (B) Comparison between the cocrystallized inhibitor XV638 from the X-ray structure 1qbr (red) and the docked inhibitor **90** (blue). Hydrogens are omitted for clarity. The picture was rendered by the BODIL program.

is based on an estimation of hydrogen bonding stabilization energy, van der Waals internal energy for the conformation of the compound, and complex interaction energy between the compound and the target protein.[42] All selected conformations were docked in a similar orientation at the active site of HIV-1 PR (Figure 2A).

We were interested in how the GOLD rank scores correlate with the experimental inhibitory activity. GOLD rank scores (Table 5) for the best-ranking conformations were compared with the corresponding experimental activity using the standard linear r^2 value (Figure 3). An r^2 value of 0.416 was obtained for the entire set of 113 inhibitors. This indicates that there was a correlation between GOLD score and inhibitory activity. The correlation can also be observed from the plot in Figure 3, and the correlation was best with highly active inhibitors.

The three X-ray structures of HIV-1 PR (1hwr, 1dmp, and 1qbr) that were cocrystallized with inhibitors **14**, **42**, and **90**, respectively, were used to further validate the docking results by calculating the heavy atom RMSD between the docked and the cocrystallized inhibitors. Obtained RMSD values for inhibitors **14**, **42**, and **90** were 0.562, 0.464, and 0.608 Å, respectively. Thus, molecular docking produced

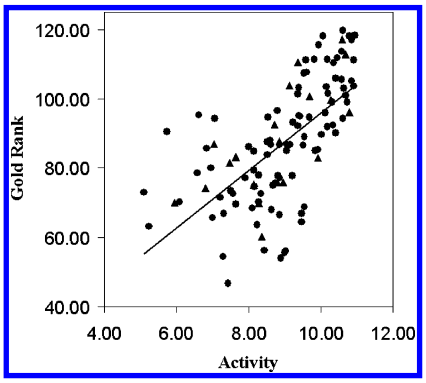


Figure 3. Correlation between experimental activities and GOLD Rank Scores (●, compounds of the training set; Δ, compounds of the test set).

good solutions when compared to existing X-ray structures and suggests that an overall inhibitor alignment could be used in 3D QSAR analyses. The docked inhibitor **90** with cocrystallized XV638 is presented in Figure 2B.

CoMFA and CoMSIA Analyses. Six 3D QSAR models were created by using both CoMFA and CoMSIA methods. Models based on manual inhibitor alignment were referred to as Model 1M (CoMFA), Model 2M (CoMSIA using steric

Table 6. Statistical Parameters of the Models[e]

	cross-validated			conventional		
	q^2	comp[a]	Spress[b]	r^2	s[c]	pred-r^{2d}
Model 1M	0.616	4	0.862	0.873	0.496	0.224
Model 2M	0.713	6	0.754	0.905	0.434	0.526
Model 3M	0.633	4	0.843	0.729	0.724	0.735
Model 1A	0.523	6	0.972	0.946	0.326	0.647
Model 2A	0.534	2	0.939	0.701	0.752	0.454
Model 3A	0.649	5	0.829	0.768	0.675	0.754

[a] Optimal number of components. [b] Standard error of predictions. [c] Standard error of estimate. [d] Predictive r^2. [e] Model 1M (CoMFA), Model 2M (CoMSIA using steric and electrostatic fields), and Model 3M (CoMSIA using hydrogen bond donor and acceptor fields) are based on manual alignment. Respective Models 1A, 2A, and 3A are based on automated alignment.

and electrostatic fields), and Model 3M (CoMSIA using hydrogen bond donor and acceptor fields), and the models based on automated inhibitor alignment were referred to as Models 1A, 2A, and 3A, respectively.

All the statistical values obtained from the CoMFA and CoMSIA models are presented in Table 6. The values obtained from Model 1M were distinctly better than those for Model 1A. This may be due to the fact that the CoMFA method is sensitive to changes in conformations and compound alignment,[9] and such accuracy might not always be achieved when compounds having several rotatable bonds are docked. However, both Model 1M and Model 1A were statistically significant CoMFA models, that can be utilized in drug design. The predicted inhibitory activities of the 92 training set inhibitors are listed in Table 5, and graphs of experimental versus predicted activities are presented in Figure 4. There were no considerable outliers in the predictions of either model, indicating both their robustness and predictive power.

The predicted inhibitory activities of the 92 CoMSIA training set inhibitors are listed in Table 5 for all models, and the graphs of experimental versus predicted activities are presented in Figure 4. There were no considerable outliers in the predictions. This also indicates the robustness and predicting power of the models, although both Models 3M and 3A had slight difficulties in predicting inhibitors with low inhibitory activity. CoMSIA statistics indicated the same phenomenon observed with CoMFA; i.e., the models generated from manual inhibitor alignment had slightly better statistics than those based on automated molecular docking. Docked inhibitors only produced statistically better 3D QSAR model in CoMSIA models with hydrogen bond donor and acceptor fields. Nevertheless, both alignment methods can yield statistically significant and predictive models.

The CoMSIA method is not sensitive to small changes in the orientation of the compounds.[11] The stronger dependence of conformation and orientation with CoMFA was seen in the statistics of Models 1M−3M and Models 1A−3A. CoMSIA models had better statistical quality than CoMFA models for the same conformations and alignment. This result can be observed also from earlier published reports.[5,49]

Model Validation. The ability to predict the activity of inhibitors that were not included in the model building is the most important test for the models' predictivity. This was tested by using an external test set of 20 inhibitors. These inhibitors were pretreated in a similar way as the inhibitors in the corresponding training set. Predicted activities of the test set inhibitors for all six models are presented in Table 5 and plotted in Figure 4. Calculated pred-r^2 values are presented in Table 6.

Figure 4 and pred-r^2 values indicate that models based on the automated inhibitor alignment procedure predicted the activities of the test set inhibitors better than models based on manual alignment. Models 1A and 3A clearly predicted the activities of the test set better than Models 1M and 3M, while Model 2M was more predictive than Model 2A (Table 6). Thus, the manual inhibitor alignment produced statistically better CoMFA and CoMSIA models, but the CoMFA and CoMSIA models based on automated inhibitor alignment better predicted the activities of the novel inhibitors.

The statistics obtained from both CoMFA models (Models 1M and 1A) were compared with the best CoMFA model reported by Debnath.[28] Debnath had used 93 inhibitors in the training set and 25 inhibitors in the test set. The best model in that study had a q^2 value of 0.727, which is better than the q^2 value of either Model 1M or Model 1A in the present study (0.616 and 0.523, respectively). The other statistical parameters were also slightly better compared to Model 1M and Model 1A. However, the number of components used by Debnath was seven in the best model, thus being higher than the number of components used in Model 1M (four) and Model 1A (six). Model 1M and Model 1A are thus statistically simpler. Interestingly, the pred-r^2 of Debnath's model with the best q^2 was 0.555, and the best pred-r^2 of all reported CoMFA models was 0.629.[28] Our Model 1A yielded a pred-r^2 value of 0.647 (Table 6). Moreover, the predictive power of Model 3M and Model 3A (CoMSIA) was even better (pred-r^2 values of 0.735 and 0.754, respectively). This demonstrates the high predictive value of models based on automated inhibitor alignment.

Comparison of CoMFA and CoMSIA Contour Maps. Statistically relevant information from both CoMFA and CoMSIA analyses was represented as 3D contour maps. The visual analysis of the contour maps is quite rough, but the general lines can be drawn. There were some differences in the size of the favored and disfavored regions, but the types of the regions were equivalent. There were no significant disagreements between the CoMFA and CoMSIA maps based on either automated or manual inhibitor alignment. Thus, the following interpretation is focused only on models based on the automated alignment.

CoMFA and CoMSIA contour maps of Models 1A−3A with the docked inhibitor **90** are displayed as stdev*coeff plots in Figure 5. The maps give suggestions on how to modify inhibitors of the examined set to improve their activity. For Model 1A (CoMFA) and Model 2A (CoMSIA of steric and electrostatic fields), sterically favored regions are shown in green, and sterically disfavored regions are in yellow. Red regions indicate that a negative potential is favored, and blue regions indicate that a positive potential is favored. For Model 3A (CoMSIA of hydrogen bond donor and acceptor fields), cyan regions indicate that hydrogen bond donors are favored in that area, and purple regions show hydrogen bond donor disfavored areas. Magenta regions indicate that hydrogen bond acceptors are favored, while red regions show disfavored areas for hydrogen bond acceptors.

The CoMFA and CoMSIA contour maps of Models 1A−3A are in agreement with the chemical functional groups of

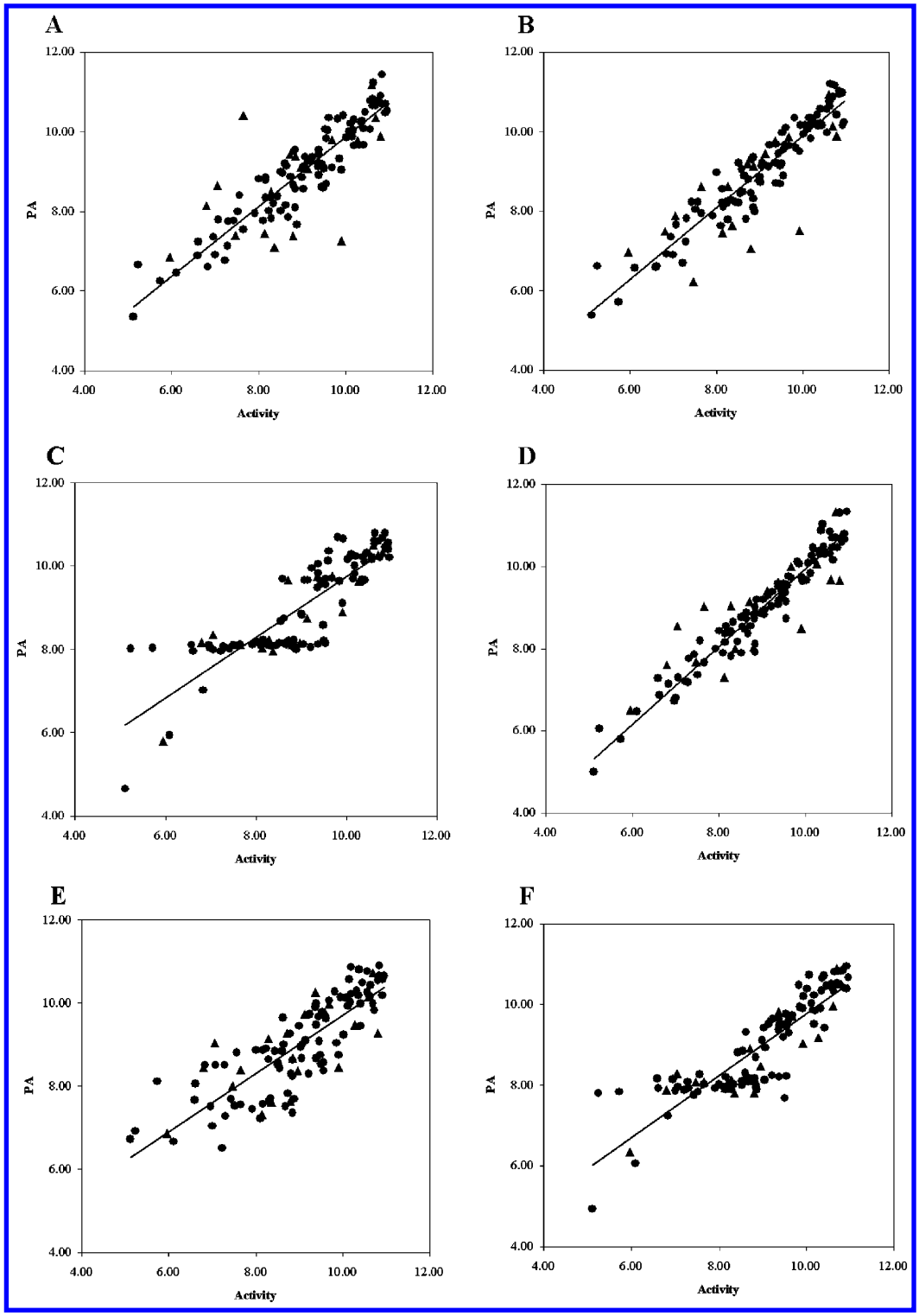


Figure 4. Correlations between the experimental and predicted activities (PA) of Models 1M−3M and 1A−3A (correlations are plotted in A−F, respectively; ●, compounds of the training set; Δ, compounds of the test set).

amino acids in the active site of HIV-1 PR. The blue CoMFA and CoMSIA regions were located in front of the backbone carbonyl oxygen of Gly48 and Gly48′ and the carboxyl group of Asp30 side chain, suggesting that a positively charged group on an inhibitor is favorable in these areas. CoMSIA indicated that a hydrogen bond between an inhibitor and the backbone carbonyl oxygens of Gly48 and Gly48′ increases activity. The red CoMFA and CoMSIA regions were located in front of the backbone NH groups of Asp29, Asp29′, and Asp30′. According to these models, an increase in the electronegativity of an inhibitor near these areas would be favorable for activity. CoMSIA proposed that hydrogen bond acceptors in an inhibitor close to the backbone NH groups of Asp30, Gly48, Asp29′, and Gly48′ could increase the binding activity. Oprea et al.[32] had already suggested the possibility of beneficial hydrogen bonds between an inhibitor and the backbone NH groups of Asp30 and Gly48, which is supported by our findings. The amino acids involved in hydrogen bonding are shown in Figure 5, together with the contour maps and the docked inhibitor **90**.

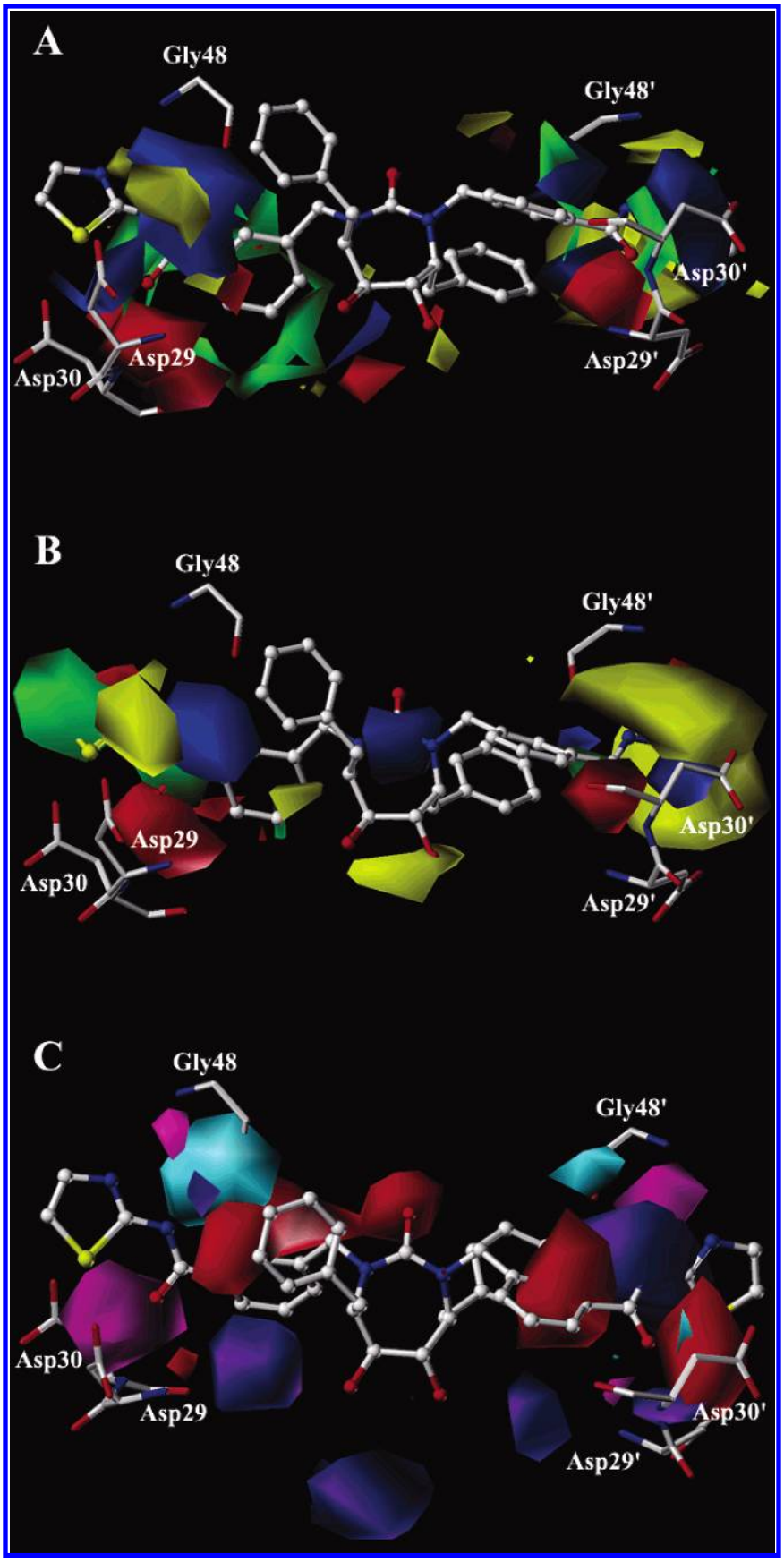


Figure 5. The contour maps of Models 1A−3A with the docked inhibitor **90** (ball-and-stick) and the amino acids involved in hydrogen bonding. All hydrogens are omitted for clarity. (A) The CoMFA steric and electrostatic fields of Model 1A. Sterically favored regions are in green; sterically disfavored regions are in yellow. Red regions indicate that negative potential is favored; blue regions indicate that positive potential is favored. (B) The CoMSIA steric and electrostatic fields of Model 2A. The color scheme is the same as previously stated. (C) The CoMSIA hydrogen bond donor and acceptor fields of Model 3A. Cyan regions indicate that hydrogen bond donors are favored in that area; purple regions show hydrogen bond donor disfavored areas. Magenta regions indicate that hydrogen bond acceptors are favored in that area; red regions show hydrogen bond acceptor disfavored areas.

Steric factors also affect the activity. Two yellow CoMFA regions are close to the hydrophobic part of Ile47, and in the volume outside the binding cavity, pointing out the areas where steric bulk in an inhibitor is not favored. Green CoMFA regions are located between amino acids Asp29 and Asp30 and in front of amino acids Val32′, Ile47′, Ile50′, and Ile84′, suggesting that steric bulk in those areas is favorable for inhibitory activity. This suggestion is consistent with the results of Debnath[27,28] and Huang et al.[7] for their inhibitor sets. The compatibility of CoMFA and CoMSIA maps with both HIV-1 PR active site amino acids and previously published results indicates the reliability of the obtained structure−activity relationship information.

CONCLUSIONS

We have compared the information value and statistical quality of CoMFA and CoMSIA models based on automatically generated alignment with identical models where the inhibitor alignment was generated manually. Results show that automated and unrestrained molecular docking can produce an inhibitor alignment that yields 3D QSAR models of comparable quality as conventional manual alignment. In addition to the statistical quality, a model based on automated alignment was able to give information about the interactions between an inhibitor and the active site of an enzyme, and thereby to improve the design of new inhibitors. The predicting power of 3D QSAR models based on automated alignment can be even better than models based on manual alignment. However, the results are only for the alignment based on the conformations automatically selected on the basis of the ranking score. Thus, better models might be obtained when examining alignments based on different combinations of the conformations.

Our study indicates that 3D QSAR models based on the automated alignment can give reliable activity predictions of new inhibitors, thus making these models potentially useful in future drug design. These results encourage the usage of an automated alignment method for creating 3D QSAR models when a reliable structure of the target protein is available.

ACKNOWLEDGMENT

We thank the National Technology Agency of Finland and the Finnish Cultural Foundation for financial support. CSC − Scientific Computing Ltd. is gratefully acknowledged for computational resources.

REFERENCES AND NOTES

(1) Kubinyi, H. *3D QSAR In Drug Design: Theory, Methods and Applications*; ESCOM: Leiden, 1993.

(2) Tervo, A. J.; Nyrönen, T. H.; Rönkkö, T.; Poso, A. A Structure−Activity Relationship Study of Catechol-O-Methyltransferase Inhibitors Combining Molecular Docking and 3D QSAR Methods. *J. Comput.-Aided. Mol. Des.* 2003, *accepted for publication*.

(3) Sippl, W. Receptor-based 3D QSAR analysis of estrogen receptor ligands − merging the accuracy of receptor-based alignments with the computational efficiency of ligand-based methods. *J. Comput.-Aided. Mol. Des.* **2000**, *14*, 559−572.

(4) Liu, H.; Huang, X.; Shen, J.; Luo, X.; Li, M.; Xiong, B.; Chen, G.; Yang, Y.; Jiang, H.; Chen, K. Inhibitory Mode of 1,5-Diarylpyrazole Derivatives against Cyclooxygenase-2 and Cyclooxygenase-1: Molecular Docking and 3D QSAR Analyses. *J. Med. Chem.* **2002**, *45*, 4816−4827.

(5) Buolamwini, J. K.; Assefa, H. CoMFA and CoMSIA 3D QSAR and docking studies on conformationally restrained cinnamoyl HIV-1 integrase inhibitors: exploration of a binding mode at the active site. *J. Med. Chem.* **2002**, *45*, 841−852.

(6) Cui, M.; Huang, X.; Luo, X.; Briggs, J. M.; Ji, R.; Chen, K.; Shen, J.; Jiang, H. Molecular docking and 3D-QSAR studies on gag peptide analogue inhibitors interacting with human cyclophilin A. *J. Med. Chem.* **2002**, *45*, 5249−5259.

(7) Huang, X.; Xu, L.; Luo, X.; Fan, K.; Ji, R.; Pei, G.; Chen, K.; Jiang, H. Elucidating the inhibiting mode of AHPBA derivatives against HIV-1 protease and building predictive 3D-QSAR models. *J. Med. Chem.* **2002**, *45*, 333−343.

(8) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, I. N.; Shindyalov, P. E.; Bourne, P. E. The Protein Data Bank. *Nucleic Acid Res.* **2000**, *28*, 235−242.

(9) 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.

(10) 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.

(11) Böhm, M.; Stürzebecher, J.; Klebe, G. Three-dimensional quantitative structure−activity relationship analyses using comparative molecular field analysis and comparative molecular similarity indices analysis to elucidate selectivity differences of inhibitors binding to trypsin, thrombin, and factor Xa. *J. Med. Chem.* **1999**, *42*, 458−477.

(12) Lam, P. Y.; Ru, Y.; Jadhav, P. K.; Aldrich, P. E.; DeLucca, G. V.; Eyermann, C. J.; Chang, C. H.; Emmett, G.; Holler, E. R.; Daneker, W. F.; Li, L.; Confalone, P. N.; McHugh, R. J.; Han, Q.; Li, R.; Markwalder, J. A.; Seitz, S. P.; Sharpe, T. R.; Bacheler, L. T.; Rayner, M. M.; Klabe, R. M.; Shum, L.; Winslow, D. L.; Kornhauser, D. M.; Hodge, C. N.; et al. Cyclic HIV protease inhibitors: synthesis, conformational analysis, P2/P2′ structure−activity relationship, and molecular recognition of cyclic ureas. *J. Med. Chem.* **1996**, *39*, 3514−3525.

(13) Wilkerson, W. W.; Dax, S.; Cheatham, W. W. Nonsymmetrically substituted cyclic urea HIV protease inhibitors. *J. Med. Chem.* **1997**, *40*, 4079−4088.

(14) Debnath, A. K. *J. Med. Chem.* **2000**, *43*, 764.

(15) Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K. J.; Starcich, B.; Josephs, S. F.; Doran, E. R.; Rafalski, J. A.; Whitehorn, E. A.; Baumeister, K.; et al. Complete nucleotide sequence of the AIDS virus, HTLV−III. *Nature* **1985**, *313*, 277−284.

(16) Seelmeier, S.; Schmidt, H.; Turk, V.; von der Helm, K. Human immunodeficiency virus has an aspartic-type protease that can be inhibited by pepstatin A. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 6612−6616.

(17) Brik, A.; Wong, C.-H. HIV-1 protease: mechanism and drug discovery. *Org. Biomol. Chem.* **2003**, *1*, 5−14.

(18) Kohl, N. E.; Emini, E. A.; Schleif, W. A.; Davis, L. J.; Heimbach, J. C.; Dixon, R. A.; Scolnick, E. M.; Sigal, I. S. Active human immunodeficiency virus protease is required for viral infectivity. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4686−4690.

(19) Wlodawer, A.; Vondrasek, J. Inhibitors of HIV-1 Protease: A Major Success of Structure-Assisted Drug Design. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 249−284.

(20) Lam, P. Y.; Jadhav, P. K.; Eyermann, C. J.; Hodge, C. N.; Ru, Y.; Bacheler, L. T.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Wong, Y. N.; et al. Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. *Science* **1994**, *263*, 380−384.

(21) Jadhav, P. K.; Ala, P.; Woerner, F. J.; Chang, C. H.; Garber, S. S.; Anton, E. D.; Bacheler, L. T. Cyclic urea amides: HIV-1 protease inhibitors with low nanomolar potency against both wild type and protease inhibitor resistant mutants of HIV. *J. Med. Chem.* **1997**, *40*, 181−191.

(22) Senese, C. L.; Hopfinger, A. J. Receptor-independent 4D-QSAR analysis of a set of norstatine derived inhibitors of HIV-1 protease. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1297−1307.

(23) Di Santo, R.; Costi, R.; Artico, M.; Massa, S.; Ragno, R.; Marshall, G. R.; La Colla, P. Design, synthesis and QSAR studies on N-aryl heteroaryl2-propanolamines, a new class of nonpeptidic HIV-1 protease inhibitors. *Bioorg. Med. Chem.* **2002**, *10*, 2511−2526.

(24) Nair, A. C.; Jayatilleke, P.; Wang, X.; Miertus, S.; Welsh, W. J. Computational studies on tetrahydropyrimidine-2-one HIV-1 protease inhibitors: improving three-dimensional quantitative structure−activity relationship comparative molecular field analysis models by inclusion of calculated inhibitor- and receptor-based properties. *J. Med. Chem.* **2002**, *45*, 973−983.

(25) Schaal, W.; Karlsson, A.; Ahlsen, G.; Lindberg, J.; Andersson, H. O.; Danielson, U. H.; Classon, B.; Unge, T.; Samuelsson, B.; Hulten, J.; Hallberg, A.; Karlen, A. Synthesis and comparative molecular field

analysis (CoMFA) of symmetric and nonsymmetric cyclic sulfamide HIV-1 protease inhibitors. *J. Med. Chem.* **2001**, *44*, 155−169.
(26) Jayatilleke, P. R.; Nair, A. C.; Zauhar, R.; Welsh, W. J. Computational studies on HIV-1 protease inhibitors: influence of calculated inhibitor-enzyme binding affinities on the statistical quality of 3D-QSAR CoMFA models. *J. Med. Chem.* **2000**, *43*, 4446−4451.
(27) Debnath, A. K. Comparative molecular field analysis (CoMFA) of a series of symmetrical bis-benzamide cyclic urea derivatives as HIV-1 protease inhibitors. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 761−767.
(28) Debnath, A. K. Three-dimensional quantitative structure−activity relationship study on cyclic urea derivatives as HIV-1 protease inhibitors: application of comparative molecular field analysis. *J. Med. Chem.* **1999**, *42*, 249−259.
(29) Krömer, R. T.; Ettmayer, P.; Hecht, P. 3D-quantitative structure−activity relationships of human immunodeficiency virus type-1 proteinase inhibitors: comparative molecular field analysis of 2-hetero-substituted statine derivatives-implications for the design of novel inhibitors. *J. Med. Chem.* **1995**, *38*, 4917−4928.
(30) Waller, C. L.; Oprea, T. I.; Giolitti, A.; Marshall, G. R. Three-Dimensional QSAR of Human Immunodeficiency Virus (I) Protease Inhibitors. 1. A CoMFA Study Employing Experimentally-Determined Alignment Rules. *J. Med. Chem.* **1993**, *36*, 4152−4160.
(31) Oprea, T. I.; Waller, C. L.; Marshall, G. R. Three-Dimensional Quantitative Structure−Activity Relationship of Human Immunodeficiency Virus (I) Protease Inhibitors. 2. Predictive Power Using Limited Exploration of Alternate Binding Modes. *J. Med. Chem.* **1994**, *37*, 2206−2215.
(32) Oprea, T. I.; Waller, C. L.; Marshall, G. R. 3D-QSAR of Human Immunodeficiency Virus (I) Protease Inhibitors. III. Interpretation of CoMFA Results. *Drug. Des. Discov.* **1994**, *12*, 29−51.
(33) *Sybyl 6.8*; Tripos Inc.: St. Louis, MO.
(34) Broyden, C. The Convergence of a Class of Double-Rank Minimization Algorithms. *J. Inst. Math. Appl.* **1970**, *6*, 76−90, 222−231.
(35) Flethcer, R. A New Approach to Variable Metric Algorithms. *Comput. J.* **1970**, *13*, 317−322.
(36) Goldfarb, D. A Family of Variable-Metric Methods Derived by Variational Means. *Math. Comput.* **1970**, *24*, 23−26.
(37) Shanno, D. Conditioning of Quasi-Newton Methods for Function Minimization. *Math. Comput.* **1970**, *24*, 647−656.
(38) Halgren, T. A. MMFF VI. MMFF94s Option for Energy Minimization Studies. *J. Comput. Chem.* **1999**, *20*, 720−729.
(39) Halgren, T. A. Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. *J. Comput. Chem.* **1996**, *17*, 520−552.
(40) Gasteiger, J.; Marsili, M. Iterative partial equalization of orbital electronegativity: a rapid access to atomic charges. *Tetrahedron* **1980**, *36*, 3219−3222.
(41) Purcell, W. P.; Singer, J. A. A Brief Review and Table of Semiempirical Parameters Used in the Hückel Molecular Orbital Method. *J. Chem. Eng. Data* **1967**, *12*, 235−246.
(42) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727−748.
(43) 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.
(44) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. An all atom force field for simulations of proteins and nucleic acids. *J. Comput. Chem.* **1986**, *7*, 230−252.
(45) Hodge, C. N.; Aldrich, P. E.; Bacheler, L. T.; Chang, C. H.; Eyermann, C. J.; Garber, S.; Grubb, M.; Jackson, D. A.; Jadhav, P. K.; Korant, B.; Lam, P. Y.; Maurin, M. B.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Reid, C.; Sharpe, T. R.; Shum, L.; Winslow, D. L.; Erickson-Viitanen, S. Improved cyclic urea inhibitors of the HIV-1 protease: synthesis, potency, resistance profile, human pharmacokinetics and X-ray crystal structure of DMP 450. *Chem. Biol.* **1996**, *3*, 301−314.
(46) Johnson, M. S.; Lehtonen, J. V. Comparison of protein three-dimensional structures. In *Bioinformatics: Sequence, Structure and Databanks*; Higgins, D., Taylor, W., Eds.; Oxford University Press: 2000; pp 15−50.
(47) Lehtonen, J. V.; Rantanen, V.-V.; Still, D.-J.; Ekholm, J.; Björklund, D.; Iftikhar, Z.; Huhtala, M.; Jussila, A.; Jaakkola, J.; Pentikäinen, O. T.; Nyrönen, T.; A., S. T.; Gyllenberg, M.; Johnson, M. S. *BODIL: a molecular modeling environment for structure−function analysis and drug discovery, unpublished*: http://www.abo.fi/fak/mnf/bkf/research/johnson/bodil.html.
(48) Cramer, R. D., III.; Bunce, J. D.; Patterson, D. E. Crossvalidation, Bootstrapping, and Partial Least Squares Compared with Multiple Regression in Conventional QSAR Studies. *Quant. Struct.-Act. Relat.* **1988**, *7*, 18−25.
(49) Makhija, M. T.; Kulkarni, V. M. 3D-QSAR and molecular modeling of HIV-1 integrase inhibitors. *J. Comput.-Aided. Mol. Des.* **2002**, *16*, 181−200.

CI0342268