A Comparison of Methods for Pharmacophore Generation with the Catalyst Software and their Use for 3D-QSAR: Application to a Set of 4-Aminopyridine Thrombin Inhibitors

P.A. Greenidge*[a] and J. Weiser[b]

[a]*Thrombosis Research Institute, Emanuelle Kaye Building, Manresa Road, London. SW3 6LR. U.K.* [b]*Anterio Consult & Research GmbH, Augustaanlage 26, 68165 Mannheim, Germany*

Abstract: The method of structure-based pharmacophores for use in 3D-QSAR as implemented by Gillner and Greenidge [6] is further examined. Conformational models are generated using both Catalyst [3] and Macromodel [7]. K_i estimates obtained with the pharmacophore models are compared with observed values for a set of 4-aminopyridine thrombin inhibitors [8].

INTRODUCTION

We will not attempt to review the field of pharmacophores as there are many people better qualified to do so. Instead we refer interested readers to a recently published book on the subject [1] and an excellent contemporary article [2]. Here we wish to share our experiences with the popular 3D-QSAR and database searching software Catalyst [3].

In Catalyst, a conformational model is an abstract representation of the accessible conformational space of a ligand. It is assumed that the biologically active conformation of a ligand (or a close approximation there of) should be contained within this model. Given a ligand training set (ligand conformational models and experimental activity values), Catalyst is able to automatically generate a number of pharmacophore models [4]. These specify the relative alignments and active conformations of the ligands consistent with their binding to a common receptor site. It is strongly recommended that for automatic pharmacophore generation, that the *best* conformation generation method be used for the training set of inhibitors instead of the *fast* mode [3]. Although both methods emphasize conformational coverage, *best* conformer generation considers the arrangement in space of chemical features rather than simply the arrangement of atoms. However, the *fast* option generates conformers at interactive speed, whereas the *best* option requires an order of magnitude more time. Given the limitations of our computing power we sought to investigate whether there was in fact any difference in the quality of the 3D-QSAR results obtained by either method. These automatically generated pharmacophores are compared with structure-based pharmacophores [5,6]. Additionally, we consider the impact of conformer generation on database scoring. Though adequate for database searching, the suitability of *fast* conformer generation for database scoring has hitherto not been investigated in detail. We also examined how the root mean square deviation (R.M.S.D.) values and correlation coefficients between predicted and observed experimental values were influenced by the choice of mapping mode (*fast fit* (rigid) or *best fit* (flexible)) of inhibitors to the structure-based and automatically generated pharmacophore models. Conformational models (both *in vacuo* and in continuum solvent) were also generated using Macromodel [7] and the 3D-QSAR results compared with the *best* mode of Catalyst conformer generation.

*Address correspondence to this author at the Thrombosis Research Institute, Emanuelle Kaye Building, Manresa Road, London. SW3 6LR. U.K. [b]Anterio Consult & Research GmbH, Augustaanlage 26, 68165 Mannheim, Germany; e-mail: Paulette.Greenidge@denovopharma.com

METHOD

Catalyst Pharmacophore Construction and 3D-QSAR

A pharmacophore model (in Catalyst called a hypothesis) [3] consists of a collection of features necessary for the biological activity of the ligands arranged in 3D-space, the common ones being hydrogen-bond acceptor, hydrogen-bond donor and hydrophobic features. Hydrogen bond donors are defined as vectors from the donor atom of the ligand to the corresponding acceptor atom in the receptor. Hydrogen bond acceptors are analogously defined. Hydrophobic features are located at the centroids of hydrophobic atoms. Catalyst features are associated with position constraints that consist of the ideal location of a particular feature in 3D-space surrounded by a spherical tolerance [8]. In order to map to the pharmacophore it is not necessary for a ligand to possess all the appropriate functional groups capable of simultaneously residing within the respective tolerance spheres of the pharmacophoric features. However, the fewer features an inhibitor maps to, and the poorer its fit to them, then the lower its affinity will be predicted to be. Each feature is associated with a weight (a measure of its proposed importance to the pharmacophore as a whole). These models may be used for 3D-QSAR analyses or as database queries.

1389-5575/01 $20.00+.00

Training Set

The training set (inhibitor conformational models and K_i values) for generation and regression of automatically generated pharmacophores, but regression only of structure-based pharmacophores, consisted of 16 inhibitors taken from the data set of Bursi and Grootenhuis [9] (inhibitors with undefined stereochemistry were excluded), Fig. (**1**). The K_i values ranged from 4nM to 2000nM. Sixteen compounds represent the minimum recommended number of molecules to be included in a training set. In the first part of this study up to 250 conformers were generated using the *fast* conformer generation method, with a maximum conformational energy of 20 kcal/mol above the lowest energy conformation of the inhibitor found by the poling algorithm. Thereafter, the *best* conformer method was used with an energy cut off of 10 kcal/mol for all molecules with the exception of compound **8** for which a value of 15 kcal/mol was used. These latter values reflect the upper energy limit of conformations of inhibitors mapping to the structure-based pharmacophore with 30% scaling of excluded volumes according to atomic van der Waals radii and regressed with the *fast* conformer model. In the Catalyst file the variable, confAnalysis.best.max.successive.failures was set equal to 100. Catalyst default values were used throughout unless otherwise stated. All calculations were performed using an Iris Indigo Elan, R4000, memory size 128Mbytes, 100 MHZ IP20 processor.

Automatically Generated Pharmacophores

Catalyst 4.0 [3] was used to automatically generate 10 pharmacophore models with up to a total of five features. Constraints were placed on the number of features such that there could be respectively, 0-5 hydrogen bond donors, hydrogen bond acceptors, hydrophobic aliphatic or hydrophobic aromatic features. During hypothesis generation, molecules are mapped to the pharmacophore features using their pre-stored conformations (*fast fit*). The dumping score for the null hypothesis was 80 bits. The larger the difference between the cost of a generated hypothesis and that of the null hypothesis the better the expected predictive power of the hypothesis for external ligands not previously included in the training set. The total costs of the hypotheses varied over a narrow range of between

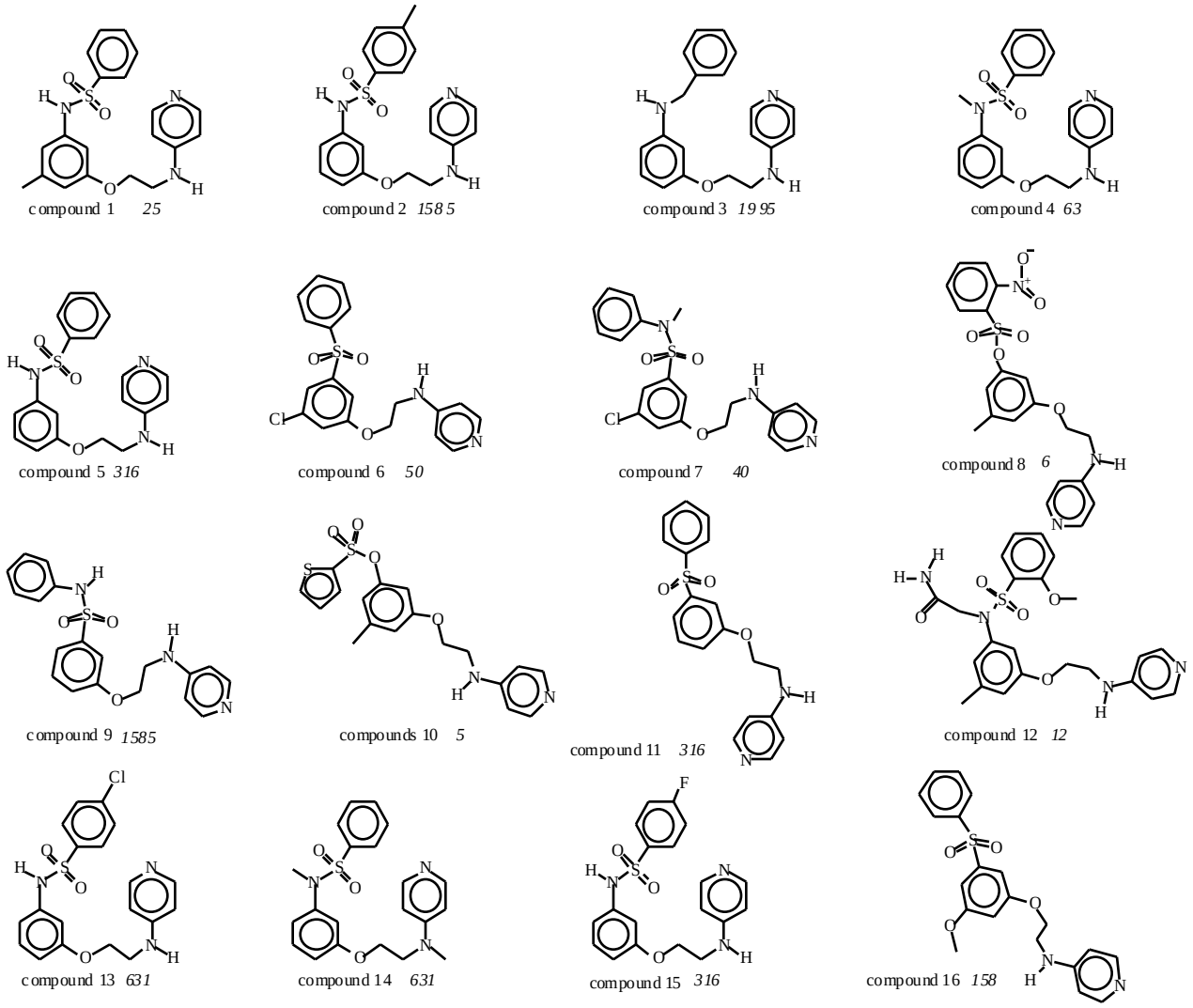

Fig. (1). Structures and experimental K_i values (nM) of 4-aminopyridine thrombin inhibitors training set members [8].

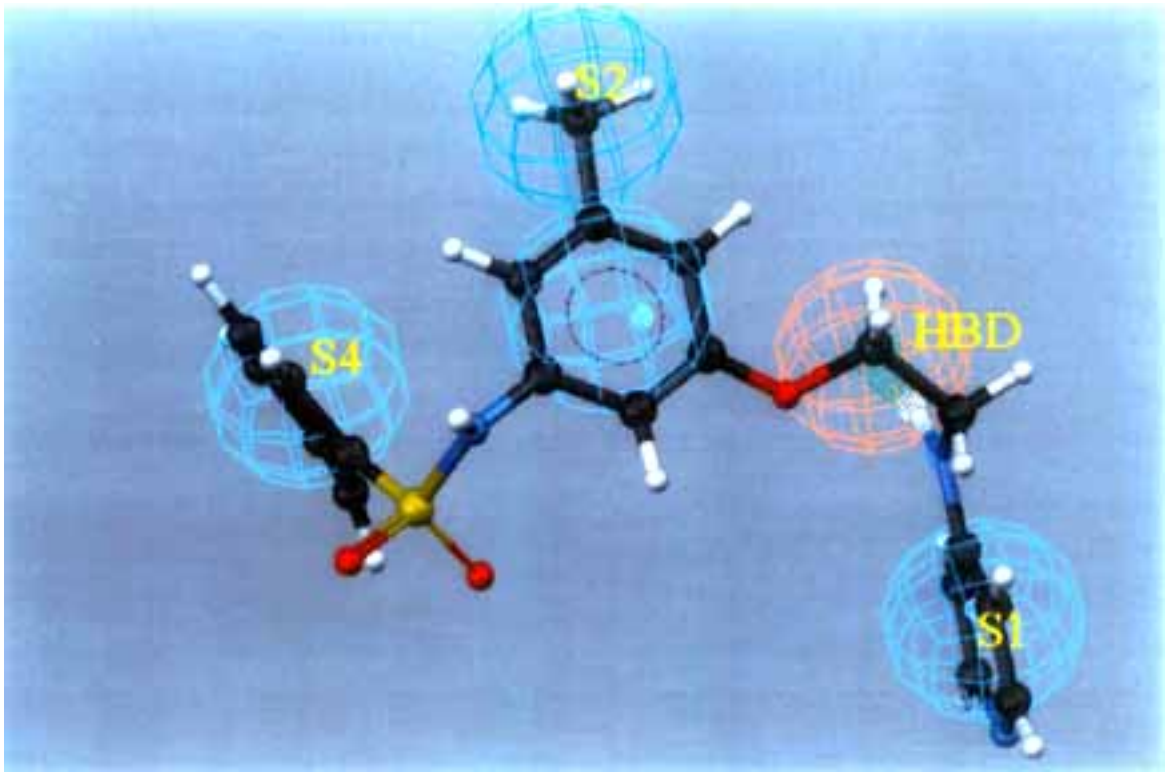

Fig. (2). Mapping of Crystal Structure Conformation of Compound 1 to the Structure-Based Thrombin Pharmacophore (no excluded volumes). The Hydrogen Bond Donor Feature is Labelled as is correspondence of hydrophobic features to thrombin active site (S1, S2 and S4).

75.94 to 78.19 bits. The cost factor *Config* is a measure of the magnitude of the hypothesis space for a given training set. If the *Config* value exceeds 17, there are more degrees of freedom in the training set than Catalyst can properly deal with and the hypothesis result may not be useful. The generated hypotheses had *Confi*g values of 18.84.

Structure-Based Pharmacophores

The location of features in the structure-based pharmacophore were defined by the crystallographic coordinates of atoms in the compound **1**/thrombin complex (PDB reference code 1UVT, resolution 2.5 Å). The inhibitor was extracted from the enzyme/inhibitor complex and this crystallographic conformation registered. The inhibitor was able to successfully map to all features of the described pharmacophore using either *fast* or *best* fit (Fig. (**2**)). Subsequently, the remaining atoms delimiting the active site were represented as excluded volumes (space which the inhibitor is not allowed to occupy) defined within acut off of 6Å from the inhibitor using Insight II [10]. The values for the van der Waals radii were taken from Pauling [11] (i.e. 1.4, 1.5, 1.7 and 1.85 Å for oxygen, nitrogen, carbon and sulphur atoms respectively).

Macromodel

All conformational searches were performed using version 7.0 of the molecular modeling program MacroModel [7]. For all compounds the MMFF94s [12] force field was used – either in vacuo or in combination with the GB/SA [13] solvation model. GB/SA treats solvent as analytical dielectric continuum that starts near the van der Waals surface of the solute and extends to infinity. The model includes both generalized Born-based (GB) [13] solvent polarization terms and surface area-based (SA) [14] solvent displacement terms. All non-bonded cutoffs were set to infinity for all calculations. Energy minimizations were performed with the Truncated Newton-Raphson Conjugate Gradient (TNCG) [15] method, which involves the use of second derivatives; the derivative convergence criterion was set to 0.05 kJ/Å-mol. Conformational search was performed by the Monte Carlo [16] for the random variation of all of the rotatable bonds method combined with the so-called Low Mode Conformational Search (LMCS) [17]. 10,000 Monte Carlo steps were carried out for each calculation and all structures up to 10 kcal/mol above the global energy minimum (GEM) were stored. All calculations were performed either with the Linux version of MacroModel on a 750 MHz Pentium III processor or the Sun version on a Sun Ultra 2 machine.

RESULTS AND DISCUSSION

We analyzed the ability of both structure-based and automatically generated pharmacophore models to explain the differing K_i values of training set members whose conformations were generated by the *fast* option.

Structure-Based Pharmacophores

Excluded volumes were incremented in steps of 10% of their respective atomic van der Waals radii, and the corresponding structure-based pharmacophore regressed against the training set (Fig. (**3**)). The structure-based pharmacophore without any excluded volumes performs worse with respect to correlation coefficent and R.M.S.D values than do the equivalent pharmacophores augmented with excluded volumes scaled from 10% to 70% of respective atomic van der Waals radii. The optimal scaling of excluded volumes as a percentage of atomic van der Waals radii is 30% (Fig. (**3**)). An initially, somewhat anomalous result occurs for excluded volumes scaled to 100% of their respective atomic radii, in that the correlation coefficient and R.M.S.D. values are comparable to that of a pharmacophore without excluded volumes. However, this is because only three training set members (compounds **1**, **8** and **10**) are able to map to the pharmacophore. The other thirteen molecules are consequently given the same affinity values by default (the intercept of the slope), which may or may not closely approximate the observed affinity. Thus, the resulting correlation coefficient and R.M.S.D. values are an artifact of the inability of most of the data set to map to the pharmacophore.

The inhibitor extracted from the 1UVT crystal structure is able to map to all features of the structure-based pharmacophore models using *fast* fit. A K_i estimate of 0.086 nM versus an observed value of 25nM was obtained for the model with excluded volumes scaled to 30% of atomic van der Waals radii. The Catalyst generated conformer is retrieved as a hit using either *fast* or *best flexible* search when the pharmacophore models are used as database queries. In order to be retrieved as a hit a molecule must be able to map to all features of a pharmacophore. However, using *fast* fit (in which features may be omitted) compound **1** does not map to all features of the pharmacophore models (Fig. (**4**)). This means that even though a conformer such that all features can be mapped (without minimization) is present within its conformational model it does not obtain the best overall fit score in doing so. In the *View Database* workbench no K_i estimate could be obtained for the Catalyst generated conformation of compound **1**. Instead the compound was reported to be just on the borderline of satisfying the query. In order to create a credible QSAR model it is essential that the training set ligands map to the pharmacophore features in the optimal manner. Thus, when a predefined pharmacophore model exists (i.e. one not automatically generated by the Catalyst software), it would surely be desirable to perform the regression using a flexible fit procedure even though it has previously been concluded that for automatically generated pharmacophores, this affords no advantage.

Automatically Generated Pharmacophores

By default Catalyst produced 10 alternative pharmacophore models (Table 1). The automatically generated pharmacophores perform very well for the training set of inhibitors, with correlation coefficients above 0.9 for all ten models, and R.M.S.D. values approximately ranging between 0.5 and 0.7 (Table 1). Significantly, the automatically generated pharmacophores assume that an oxygen atom of the sulphone moiety is involved in hydrogen bonding with the receptor (Fig. (**5**)) and not the 4-amino group, as in the crystal structure. Though there are nominally ten distinct models, many of them are in fact very similar (Table 2). Model 2 differs from Model 1 only in that the locations of a hydrophobic aliphatic and a hydrophobic aromatic feature are swapped (Table 2). The two models share the same feature weights and co-ordinates for all five equivalent features, hence the identical correlation coefficient and R.M.S.D. values produced by both models. Models one to seven share a common hydrophobic feature (Table 2) which is mapped by the methyl group of compound **1**. Similarly, Models 3 to 6 include a hydrogen bond accepting feature, and a hydrophobic aliphatic (aromatic Model 6) feature with identical coordinates between models (Table 2). The coordinates and type (hydrophobic aliphatic or hydrophobic aromatic) of the other two features in these models are also similar, but the weights may vary between models (Table 2), hence the limited spread of correlation

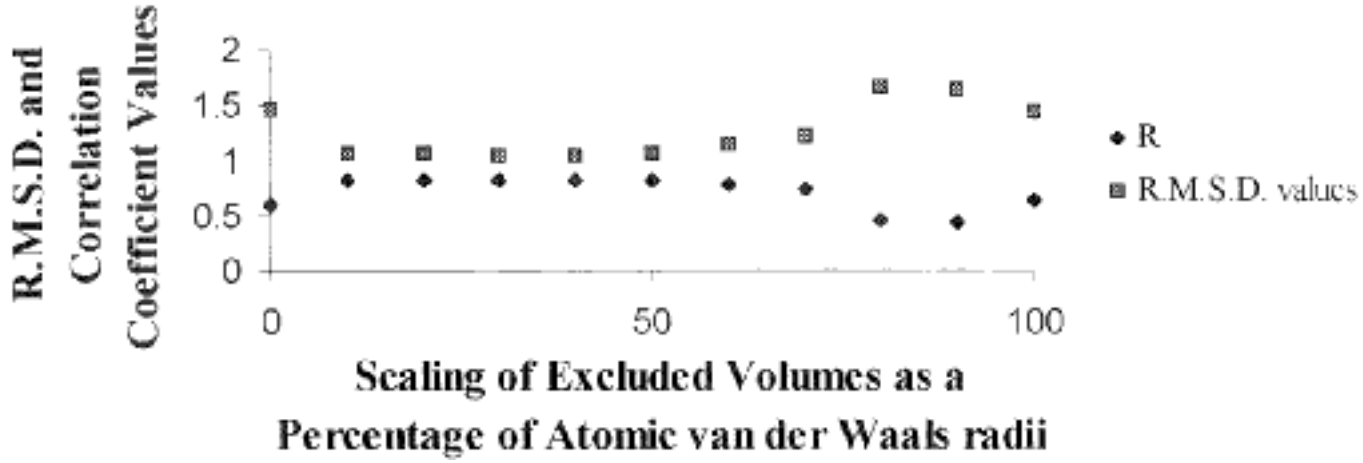

Fig. (3). Structure-Based Pharmacophores (fast conformational model): Relationship Between the Scaling of Excluded Volumes as a Percentage of Atomic van der Waals Radii and R.M.S.D. values and Correlation Coefficient R Between Observed and Calculated K_i Values.

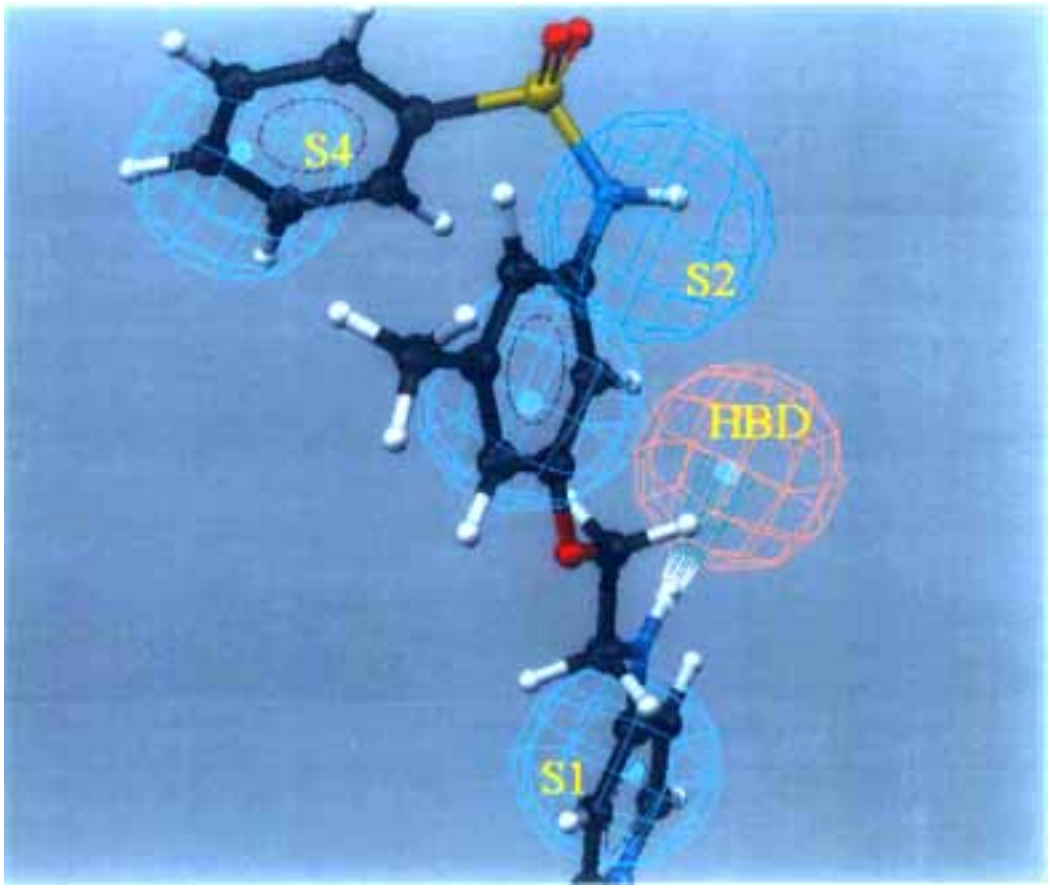

Fig. (4). Catalyst Generated Conformer and Mapping of Compound 1 to the Structure-Based Pharmacophore. The Hydrogen Bond Donor Feature is Labelled as is correspondence of hydrophobic features to thrombin active site (S1, S2 and S4).

coefficient and R.M.S.D. values for Models 3 to 6. Models 7 and 8 have some features and coordinates in common with Models 1 and 2; Models 9 and 10, unlike the other models lack a hydrogen bond accepting feature and have only four features.

Comparison of Automatically Generated and Structure-Based Pharmacophores

The automatically generated pharmacophores seem to perform better than the structure-based pharmacophores, based upon correlation coefficients and R.M.S.D. values of inhibitor conformations generated using the *fast* option (Table 1, Fig. (**3**)). However, the proposed mapping modes of the automatically generated pharmacophores gave cause for concern. The para-benzene-sulphone substituents of compounds **2**, **5**, **13** and **15** respectively, are not expected to affect the conformations that the molecules are able to adopt. It was therefore surprising that these models should differentiate between these molecules in any way as they lack excluded volumes. For instance, Model 1 ranked the compounds in the order compound **13** < compound **5** < compound **15** < compound **2**. Examination of the hypothesis log file revealed that all four of these compounds mapped to 4 out of a possible five features, hence the distinction in affinity estimates arose from their fit score. Since in the *fast fit* affinity estimation used in hypothesis generation, the pre-stored conformations of molecules are rigidly fit to a hypothesis, fit score variations are due to differences in the conformational models of the molecules. In order to remedy inconsistencies in the conformational models we performed a *best fit* of all of the molecules to the first automatically generated hypothesis. The *best fit* method minimizes the inhibitor so as to optimize the fit to the pharmacophore features. The R.M.S.D. value between experimental and estimated affinities then increased from 0.49 to 2.41. Effectively, the pharmacophore became less selective with respect to its ability to be able to distinguish between good and poor inhibitors. This increase was for the main part a consequence of the improved fit scores for many of the compounds such that their K_i values were estimated to be much better than their observed values. For over half the molecules in the training set, this overestimation was by a factor of 10 or more. The increase in R.M.S.D. values for the structure-based pharmacophore without excluded volumes was less pronounced, from 1.48 to 2.10 and for the structure-based pharmacophore with excluded volumes scaled to 30% of atomic van der Waals radii, the increase was from 1.05 to 1.29 [18]. Thus, it may be that structure based pharmacophores, especially those containing excluded volumes are less sensitive to the method used to generate conformational models for use in hypothesis regression than automatically generated pharmacophores. This is entirely consistent with the complete dependence of the automatically generated pharmacophores on the conformational model and

Table 1. Catalyst Automatically Generated Pharmacophores: Root Mean Square Deviation (R.M.S.D.) and Correlation Coefficient R, of Observed and Calculated Experimental Affinities

Pharmacophore Model	R	R.M.S.D.	Feature Weights	Total Cost
1	0.965	0.488	1.78707	75.9360
2	0.965	0.488	1.78707	75.9359
3	0.951	0.568	1.92524	76.3927
4	0.950	0.576	1.95088	76.4479
5	0.948	0.585	1.94759	76.5332
6	0.948	0.585	1.94760	76.5331
7	0.945	0.607	1.79145	76.9602
8	0.938	0.638	1.86299	77.1424
9	0.918	0.731	2.14544	78.1692
10	0.915	0.743	2.01069	78.1921

Table 2. Features and Coordinates of Automatically Generated Pharmacophores

Pharmacophore Model	Features and coordinates				
1	HBA 7.58, 6.75. -3.00 9.67, 5.11, -4.42	Hphobic 3.48, 7.44, -4.35	Aromatic 6.92, 3.51, 0.08	Hphobic 1.59, 0.44, 4.67	Hphobic 9.26, 1.54, -1.26
2	HBA ibid	Aromatic ibid	Hphobic ibid	Hphobic ibid	Hphobic ibid
3	HBA 7.44, 6.55, -3.04 9.82, 5.31, -4.38	Hphobic 3.08, 6.23, -4.73	Hphobic 7.02, 3.29, 0.05	Hphobic 2.23, -0.35, 4.66	Hphobic ibid
4	HBA ibid	Hphobic ibid	Hphobic 7.05, 3.33, 0.07	Aromatic ibid	Hphobic ibid
5	HBA ibid	Hphobic ibid	Aromatic ibid	Hphobic 2.30, -0.25, 4.78	Hphobic ibid
6	HBA ibid	Aromatic ibid	Hphobic ibid	Aromatic ibid	Hphobic ibid
7	HBA 7.66, 6.82, -2.99 9.59, 5.04, -4.44	Aromatic 3.48, 7.44, -4.35	Hphobic 2.03, 0.72, 5.05	Hphobic 2.30, -0.25, 4.78	Hphobic ibid
8	HBA 7.58, 6.75. -3.00 9.67, 5.11, -4.42	Aromatic 3.48, 7.44, -4.35	Hphobic 6.78, 3.32, -0.06	Hphobic 1.59, 0.44, 4.67	Aromatic 6.78, 3.32, -0.06
9		Hphobic 3.08, 5.43, -5.13	Hphobic 7.11, 3.23, -0.03	Hphobic 2.23, -0.35, 4.66	Hphobic 9.26, 1.54, -1.26
10		Hphobic 4.22, 7.83, -4.21	Aromatic 6.93, 2.71, 0.72	Aromatic 6.92, -4.11, 2.93	Aromatic 9.39, 3.80,2.25

KEY: HBA = Hydrogen Bond Acceptor, Hphobic = Hydrophobic, Aromatic = Hydrophobic Aromatic.

inhibitor structures of the training set. It follows that structure-based pharmacophores are influenced to a lesser extent by the conformational model, as pharmacophoric features are defined using 3D experimental information from both the inhibitor and enzyme. Though of course the inhibitors must nevertheless possess a conformation which allows them to map to the pharmacophore model.

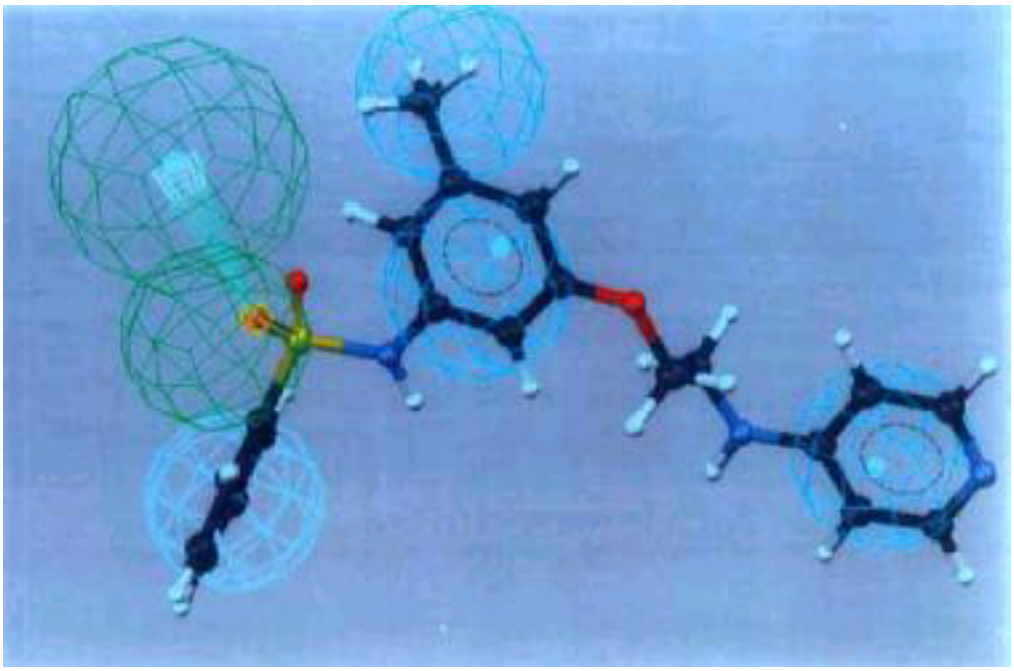

Fig. (5). Mapping of Compound 1 to the Automatically Generated Pharmacophore (Model 1). All features are hydrophobic with the exception of the hydrogen bond accepting feature to which the sulfone group maps.

Best Conformational Model

We then went on to generate a conformational model for the inhibitors by the recommended method using the *best* option. This took approximately one week on an Iris Indigo Elan (see Methods). We regressed the structure-based pharmacophores with the new conformational models, but did not attempt to automatically generate new hypotheses. The training set contains some inhibitors (compounds **2** and **13**) whose affinity differences can only be expected to be reproduced if some kind of steric restriction is included in the model, e.g. excluded volumes. It might be argued that including such inhibitors in the training set for automatically generated pharmacophores is then inappropriate. However, it is likely that in the early stages of model development that such potential steric problems would not be known.

The findings for the *fast* and *best* method of conformer generation are similar. There is very little variation in the R.M.S.D. values and correlation coefficients for the structure-based pharmacophore without excluded volumes regressed against the *best* generated conformational model of inhibitors and for the *fast* model (R.M.S.D. values of 1.38 and 1.47 respectively and correlation coefficients of 0.66 and 0.60 respectively). These values are even more similar for pharmacophores with excluded volumes (e.g. scaling of excluded volumes of 30% with respect to atomic van der Waals radii, R.M.S.D values of 1.04 and 1.15 and correlation coefficients of 0.82 and 0.79 respectively for *fast* and *best* conformation generation). These results support our assertion of the likely lack of sensitivity of structure-based pharmacophores (as opposed to automatically generated pharmacophores) to the mode of generation of their conformational models, *fast* or *best*.

Scoring- Fast Versus Best Conformational Model

A marked difference in scores is obtained for the Catalyst generated conformational models of compound **1**, *fast* and *best* mode in the View Database workbench. The structure-based pharmacophore with excluded volumes scaled at 30% with respect to atomic radii and regressed with the training set whose conformational models had been generated with the *best* mode was used. It is not possible to obtain a Ki estimate for compound **1** generated via *fast* mode but it is for the conformation generated by the *best* mode. Additionally, it was found that there is a 1000 fold difference in the K_i estimate obtained for the stored crystal structure conformation of compound **1** (0.28nM) versus the Catalyst generated conformation (280nM). Thus, given that 'chemically meaningful tolerances are on the same scale as the resolutions that are possible using point conformers for small- to medium-sized drug molecules' [19], the above

Table 3. Structure-based Pharmacophore Model with 30% Scaling of Excluded Volumes as a Percentage of Atomic van der Waals Radii Regressed with Different Conformational Models of the Training Set (Figure 1, 16 Molecules). A Comparison of R.M.S.D. and Correlation Coefficients Between Observed and Calculated K_i Values

Conformational Model	R.M.S.D.	R
Catalyst-*best*	1.15	0.79
Macromodel-*in vacuo*	1.52	0.57
Macromodel-continuum solvent	1.21	0.75

result suggests that the scoring function needs to be more tolerant of the generated conformers. To further investigate this point we generated conformations both in continuum solvent and *in vacuo* using Macromodel [7]. A precedent for this exists in the work of Langgård et al.[20].

Macromodel Conformation Models Versus Catalyst Best Conformational Model

Though they did not lead to any overall improvement in K_i predictions (Table 3), it can be argued that the Macromodel generated conformers result in a better 3D-QSAR model. They map to the pharmacophore in a manner which is more consistent with the experimentally observed inhibitor binding modes (Table 4) than the Catalyst generated conformers (using fast fit). Another study [21] suggests that solvation effects may play a role in influencing K_i values for this series of inhibitors. Indeed the two outliers of the Macromodel *in vacuo* model correspond to two of the molecules highlighted in that study (compounds **3** and **9**). If these molecules are removed from the data set R.M.S.D and correlation coefficient values improve to 1.11 and 0.8 respectively. Hence, using an alternative means to generate the conformational models represents only one of the steps required to improve the 3D-QSAR model whilst retaining the current scoring function. When both Table 3 and Table 4 are considered, the macromodel search with GB/SA turned on produced the best results. Further evidence that with Macromodel generated conformational models, Catalyst is able to more closely reproduce experimentally observed inhibitor binding modes is given by the pharmacophore with 100% scaling of excluded volumes as a percentage of atomic van der Waals radii. For the Catalyst *fast* conformer model only three inhibitors could map to this pharmacophore, and with the *best* model it was not possible to obtain a QSAR model. However, with the Macromodel *in vacuo* conformation model, only five molecules are unable to map to the pharmacophore, and for at least three of these molecules, this lack of mapping is readily attributable to the presence of bulky substituents [6](compounds **2**, **13** and **14**).

The work presented here has aspects in common with a number of recently published articles in that it addresses the question of how to best utilize existing structural information for docking, correct ranking of binding modes, and scoring of molecules in a receptor binding site [22].

CONCLUSIONS

Structure-based pharmacophores appear to be less sensitive to the method of conformational model generation of the training set than those automatically generated by the

Table 4. A Comparison of the Structure-based Pharmacophore (30% scaling of Excluded Volumes, Figure 2) Features Mapped by the Six Inhibitors of the Training Set With the Best K_i Values and Whose Conformational Models are Generated By Different Methods

compound	Method*	S1	HBD	S2	'Other'	S4
12	Catalyst	+	+	+	+	-
	In vacuo	+	-	+	+	+
	solvent	+	+	+	+	-
10	Catalyst	+	+	+	+	-
	In vacuo	+	+	+	+	+
	solvent	+	+	+	+	+
8	Catalyst	+	+	+	+	-
	In vacuo	+	-	+	+	+
	solvent	+	-	+	+	+
1	Catalyst	+	+	+	+	-
	In vacuo	+	+	+	+	+
	solvent	+	+	+	+	+
7	Catalyst	+	+	+	+	-
	In vacuo	+	+	+	+	+
	solvent	+	-	+	+	+
6	Catalyst	+	+	+	+	-
	In vacuo	+	+	+	+	-
	solvent	+	+	+	+	-

* Catalyst = Catalyst [3] *best* method of conformer generation, and *in vacuo* and solvent refer to Macromodel [7] generated conformers(see Methods Section)
# Benzene-sulfone moiety occupies the SI pocket instead of pyridine ring as observed in the crystal structure. S1, S2, S4 refer to enzyme active site. 'Other' refers to aromatic feature mapped by central ring of compound **1**. HBD = Hydrogen Bond Donor. If the S4 site is *not* mapped an aromatic ring of the compounds tends to part into the solvent.

Catalyst software, in terms of R.M.S.D. values and correlation coefficients between predicted and observed values. However, neither conformational models generated using *fast* or *best* mode fared well in producing mapping modes that were consistent with that crystallographically observed for compound **1**. In this study, Macromodel generated conformers were much more successful in this regard. Thus, at least for structure-base pharmacophores, and catalyst generated conformers, it may be necessary to re-examine the current scoring algorithm which seemingly rewards molecules mapping to fewer features 'well' as compared to more features 'poorly'. We also showed that it is possible to optimize pharmacophore models for 3D-QSAR by scaling the excluded volumes.

ACKNOWLEDGEMENTS

Mikael Gillner for previous fruitful collaborations. John Deadman and Sandrine Merette for encouragement with this work and helpful interactions.

ABBREVIATION

R.M.S.D. = Root mean square deviation

REFERENCES

[1] Pharmacophore Perception, Development and Use in Drug Design, International University Line: La Jolla, **1999**.

[2] Van Drie J.H. Nugent R.A. SAR and QSAR in Environmental Research 1**998**, 9,1.

[3] Catalyst 4.0 Tutorials, August **1998**. San Diego: Molecular Simulations Inc.

[4] Gund P. In: *Progress in Molecular and Subcellular Biology*, Hahn F.E., Ed., Springer Verlag: New York, **1977**; Vol. 5, pp. 117-143.

[5] Greenidge P.A., Carlsson B., Bladh L-G., Gillner M. *J Med Chem* **1998**, *41*, 2503.

[6] Gillner M., Greenidge P. In**.** *Pharmacophore Perception, Development and Use in Drug Design*; Guner O.F., Ed., International University Line: La Jolla, **1999**; pp. 371-384.

[7] Mohamadi, F., Richards, N. G. J., Guida, W. C., Liskamp, R., Lipton, M., Caufield, C., Chang, G., Hendrickson, T., Still, W. C. *J Comput Chem* **1990**, *11*, 440-467. We used MacroModel Version 7.0 from Schrödinger Inc., Portland, OR, USA.

[8] Greene J., Kahn S., Savoj H., Sprague P., Teig S. J. *Chem. Inf. Comput. Sci.*, **1994**, *34*, 1297.

[9] Bursi R., Grootenhuis P.D. *J. Comput.-Aided Molec Des*, **1999**, *13*, 221.

[10] Insight II User Guide, September **1997**. San Diego: MSI, **1997**.

[11] Pauling L. *The Nature of the Chemical Bond*, Cornell University Press: New York, **1960**.

[12] Halgren, T.A. *J. Comput Chem.*, **1999**, *20*, 720-729.

[13] Still, W. C., Tempczyk, A., Hawley, R. C., Hendrickson, T. *J Am Chem Soc* **1990**, *112*, 6127-6129.

[14] a) Weiser, J., Shenkin, P.S., Still, W.C. *J. Comput. Chem.* **1999**, *20*, 217-230.b) Weiser, J., Weiser, A. A., Shenkin, P. S., Still, W.C. *J. Comput. Chem.* **1998**, *19*, 797-808.c) Weiser, J., Shenkin, P.S., Still, W.C. *J. Comput. Chem.* **1999**, *20*, 586-596.

[15] J.W. Ponder and F.M. Richards, *J. Comput. Chem.* **1987**, *8*, 1016-1024 .

[16] Chang, G., Guida, W. C., Still, W. C. *J. Am. Chem. Soc.* **1989**, *111*, 4379.

[17] Kolossvary, I., Guida, W. C. *J. Am. Chem. Soc.* **1996**, *118*, 5011.

[18] Counter intuitively *best* fit gave higher fit scores than did *fast fit* (i.e. higher K_i estimates). A possible explanation for the anomalous result is that in *fast* fit a lot of features are fitted within a large error margin, but *best* fit "improves" the fitting by mapping on to these features with a lesser R.M.S.D. error. However in doing so it may leave out one of the features all together, thereby reducing the fit value.

[19] Smellie A., Kahn S.B., Teig S.C. *J. Chem. Inf. Comput. Sci.*, **1995**, *35*, 295.

[20] Langgård M; Bjørnholm B., Gundertofte K.. In**.** *Pharmacophore Perception, Development and Use in Drug Design*; Guner O.F., Ed., International University Line: La Jolla, **1999**; pp. 237-250.

[21] Greenidge P., manuscript in preparation.

[22] a) Fradera X., Knegtel R.M., Mestres J., Proteins **2000**, *40*, 623. b) Knegtel R.M., Bayada D.M., Engh R.A., von der Saal W., van Geerstein V.J., Grootenhuis P.D. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 167. c) Vieth M., Cummins D.J., *J. Med. Chem.* **2000**, *43*, 3020. d) Glide (Grid Based Ligand Docking with Energetics), Manual, February **2001**, OR, USA; Schroedinger Inc.

N.B. Fradera et al. liken their approach to a pharmacophore in a binding site.