Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters 15 (2005) 1027–1032
Bioorganic & Medicinal Chemistry Letters

Three-dimensional quantitative structure (3-D QSAR) activity relationship studies on imidazolyl and *N*-pyrrolyl heptenoates as 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR) inhibitors by comparative molecular similarity indices analysis (CoMSIA)

Ramasamy Thilagavathi, Raj Kumar, Vema Aparna, M. Elizabeth Sobhia, Bulusu Gopalakrishnan and Asit K. Chakraborti*

Department of Medicinal Chemistry, National Institute of Pharmaceutical Education and Research (NIPER), Sector 67, S. A. S. Nagar, Punjab 160 062, India

Received 2 September 2004; accepted 14 December 2004
Available online 18 January 2005

Abstract—A comparative molecular similarity indices analysis (CoMSIA) of a set of 29 imidazolyl and *N*-pyrrolyl heptenoates have been performed to find out the structural requirements for 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR) inhibitory activity. The HMG like side chain, a common moiety of statins, was used to align the molecules. The results guide to design new chemical entities with high potency.

1. Introduction

Coronary artery disease is a major health problem in developed countries and currently affects 13–14 million adults in the United States alone. Elevated cholesterol level is the primary factor in this disease.[1,2] 3-Hydroxy-3-methylglutaryl-CoA reductase (HMGR) is the primary target enzyme for chemotherapy of hypercholesterolemia. Inhibitors of HMGR, commonly referred to as statins, share a HMG moiety and bind to the active site of HMGR. Statins (Fig. 1) are effective and safe drugs and widely prescribed in cholesterol-lowering therapy. Inhibition of HMGR also induces growth arrest and cell death in several cancer cell types, presumably through the reduction of nonsterol mevalonate-derived products.[3,4]

The recent revelation of the crystal structure of the catalytic domain of human HMGR complexed with the inhibitor by Istavan and Deisenhofer[5] explains the detailed characterization of the active site and the

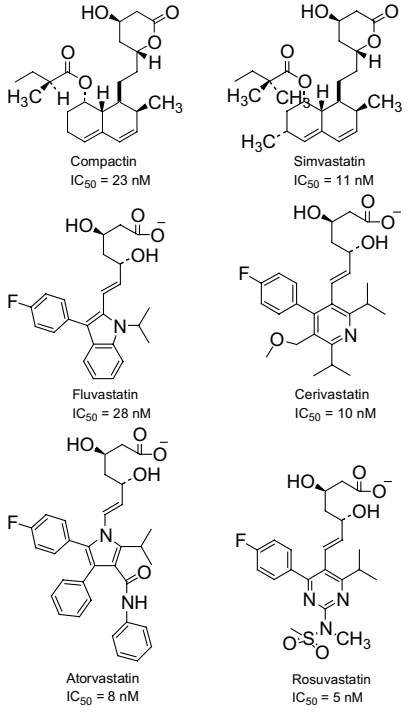


Figure 1. Structures of statins.

Keywords: 3D-QSAR; HMG-CoA; Imidazolyl heptenoates; *N*-Pyrolyl heptenoates.
* Correponding author. Tel.: +91 1722214682; fax: +91 1722214692; e-mail: akchakraborti@niper.ac.in

0960-894X/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2004.12.042

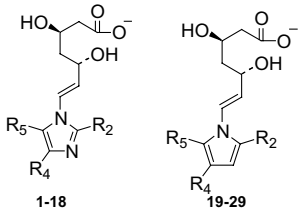
HMGR–statin interactions. Previous studies provided many compounds with IC_{50} values in nanomolar range based on the replacement of the chiral decalin moiety of the statins (Fig. 1). Further the synthesis of pravastatin analogues, methanesulfonamide pyrimidine/pyrrole substituted 3,5-dihydroxy-6-heptenoates and the thiophene-based analogues were reported to be potent inhibitors of HMGR.[6–8]

Recently, we have been involved in the computer aided design of novel PDE-IV[9–11] and COX-2[12–14] inhibitors using three-dimensional quantitative structure–activity relationship (3-D QSAR)[15,16] studies. In the present study, we have performed the 3-D QSAR studies on imidazolyl and *N*-pyrrolyl heptenoates by comparative molecular field similarity indices analysis (CoMSIA) method. CoMSIA is one of the latest techniques[17] used to produce three-dimensional models to indicate the regions that affect biological activity with the change in chemical substitution. Compared to comparative molecular field analysis (CoMFA), it provides more insight into the physico-chemical properties of molecules as it includes extra descriptors such as hydrophobicity, hydrogen bond donor and hydrogen bond acceptor properties.

2. Results and discussion

Twenty nine 3,5-dihydroxyheptenoates (Table 1) reported by the same group[18,19] were selected for the present study. Twenty three molecules were taken for construction of training set and the remaining six molecules were used as test set to validate the developed CoMSIA model. The bioactive conformation of fluvastatin was extracted from the HMGR–inhibitor complex (1HWI.pdb)[20] and used as a template to model the selected compounds. The bound conformation of

Table 1. The structures and actual and predicted inhibitory activities



Cd	R_2	R_4	R_5	Actual pIC_{50}	Pred pIC_{50}	Residual	Set[a]
1	CF_3	4-F–C_6H_4	4-F–C_6H_4	7.89	7.85	0.04	TR
2	CH_3	4-F–C_6H_4	4-F–C_6H_4	<7.00	8.17	−1.17	Outlier
3	*t*-Bu	4-F–C_6H_4	4-F–C_6H_4	8.15	8.40	−0.25	TS
4	Me_2N	4-F–C_6H_4	4-F–C_6H_4	8.04	7.97	0.07	TR
5	4-F–C_6H_4	4-F–C_6H_4	Me_2CH	9.00	8.75	0.25	TS
6	Me_2CH	4-F–C_6H_4	3-Cl–C_6H_4	8.10	7.81	0.29	TR
7	Me_2CH	4-F–C_6H_4	3,5-Di-Cl–C_6H_3	7.66	7.73	−0.07	TR
8	Me_2CH	4-F–C_6H_4	3,5-Di-Me–C_6H_3	7.25	7.04	0.21	TR
9	Me_2CH	4-F–C_6H_4	2-Me–4-F–C_6H_3	8.40	8.63	−0.23	TR
10	Me_2CH	4-F–C_6H_4	3,5-Di-Me–4-F–C_6H_2	7.52	7.52	0.00	TR
11	Me_2CH	4-F–C_6H_4	3,5-Di-Et–4-F–C_6H_2	7.00	7.26	−0.26	TS
12	Me_2CH	3,5-Di-Me–4-Cl–C_6H_2	4-F–C_6H_4	8.70	8.69	0.01	TR
13	Me_2CH	3-Pyridyl	4-F–C_6H_4	9.00	9.15	−0.15	TR
14	Me_2CH	3-SO_2Me–C_6H_4	4-F–C_6H_4	8.70	8.74	−0.04	TR
15	Me_2CH	3-NHMe–C_6H_4	4-F–C_6H_4	8.52	8.58	−0.06	TR
16*	Me_2CH	4-F–C_6H_4	4-F–C_6H_4	6.15	6.16	−0.01	TR
17*	Me_2CH	4-F–C_6H_4	4-F–C_6H_4	5.70	5.79	−0.09	TR
18	Me_2CH	4-F–C_6H_4	4-F–C_6H_4	9.00	8.75	0.25	TR
19	Me_2CH	4-F–C_6H_4	4-F–C_6H_4	8.70	8.50	0.20	TR
20	Me_2CH	C_6H_5	4-F–C_6H_4	8.70	8.50	0.20	TR
21	Me_2CH	3-Cl–C_6H_4	4-F–C_6H_4	8.70	9.08	−0.38	TS
22	Me_2CH	3-Br–C_6H_4	4-F–C_6H_4	9.30	9.26	0.04	TR
23	Me_2CH	4-Pyridyl	4-F–C_6H_4	8.30	8.33	−0.03	TR
24	Me_2CH	2-Pyridyl	4-F–C_6H_4	8.70	8.69	0.01	TR
25[b]	Me_2CH	4-F–C_6H_4	4-F–C_6H_4	9.52	9.58	−0.06	TR
26	Me_2CH	4-F–C_6H_4	C_6H_5	7.60	7.85	−0.25	TR
27	Me_2CH	4-F–C_6H_4	3-Cl–C_6H_4	7.15	7.38	−0.23	TR
28	Me_2CH	4-F–C_6H_4	3-Br–C_6H_4	7.16	7.38	−0.23	TS
29[c]	Me_2CH	4-F–C_6H_4	4-F–C_6H_4	6.98	6.97	0.01	TR

Cd = Compound, **16*** = *3S,5R*, **17*** = *3R,5R*.

[a] TR = training set, TS = test set.

[b] Bromine substitution at C_4 of pyrrole ring.

[c] Heptanoate.

fluvastatin possessed 3*R*,5*S* configuration. This chiral form has been reported[18,19] to be active for this series of compounds and therefore used for all the molecules. The HMGR inhibitory activities reported from rat liver microsomal assay[18,19] were used for CoMSIA studies.

The X-ray crystallography studies[5] confirmed that the side chain moiety, common to the HMGR inhibitors, binds to the narrow pocket where HMG is normally bound and the structurally diverse rigid hydrophobic groups of the statins are accommodated in a shallow nonpolar groove. Thus, the rigid hydrophobic group behaves as a linker and helps to properly orient the hydrophilic side chain inside the catalytic active site. Keeping this point in mind, we aligned the molecules using the C_1–C_5 carbon atoms of HMG like side chain moiety (Fig. 2) of the most active molecule **25** (Table 1).

The summary of the statistical results obtained for CoMSIA studies are shown in Table 2. We found that the CoMSIA descriptors such as steric, electrostatic, hydrophobic and hydrogen bond donor played a significant role in the prediction of biological activity. An excellent value of 0.89 for r^2 prediction was obtained for this model with the r^2cv of 0.613. The actual and predicted value of the training and test set molecules showed a linear relationship (Fig. 3). Exclusion of both the steric and electrostatic field descriptors produced no significant change in internal predictivity. However, a little decrease in the r^2 prediction (0.89–0.84) was observed. Removal of only the steric field from the model produced marginally better r^2cv with a little drop in the r^2 prediction value. Incorporation of the hydrogen bond acceptor field resulted in marginally inferior internal and external predictions.

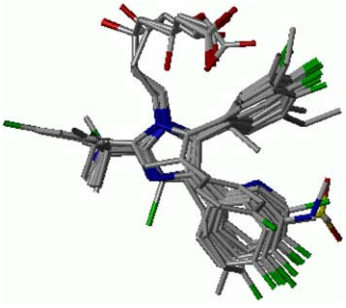


Figure 2. Alignment based on database method.

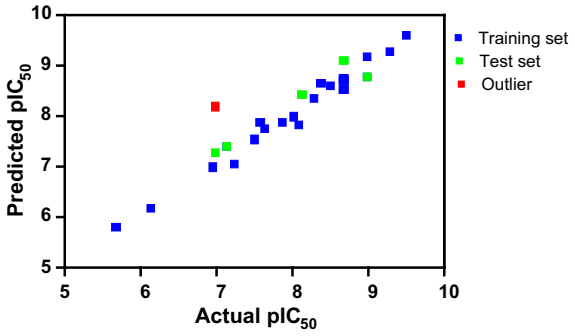


Figure 3. Actual versus predicted pIC_{50} of training and test set molecules.

The contour maps (Fig. 4) produced by CoMSIA were analyzed by superimposing them onto the imidazolyl heptenoate **18** as this was the most active molecule of the imidazole series. A large green contour surrounding the isopropyl moiety indicated the importance of the presence of a bulky group in this region for biological activity. Thus, compounds **1** and **2** with less bulky group at this position exhibited decreased biological activity. Although compound **1** was predicted well by the model, compound **2** was over predicted and thus considered as outlier. The possible reason for the over prediction may be that the training set contains only one molecule (**1**) with less bulky group (i.e., CF_3) with a pIC_{50} value of 7.89, which is lower than that of the most active

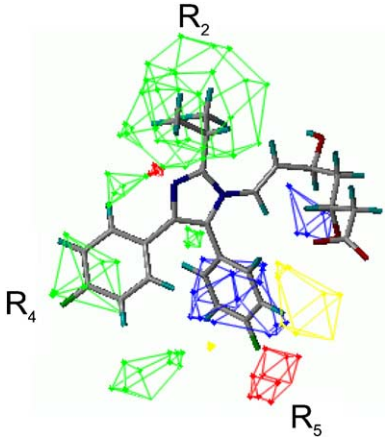


Figure 4. CoMSIA steric–electrostatic contour maps with the molecule **18**.

Table 2. The summary of the results of CoMSIA

	Steric, electrostatic, hydrophobic, donor, acceptor	Hydrophobic, donor	Steric, electrostatic, hydrophobic, donor	Electrostatic, hydrophobic, donor
r^2cv	0.595	0.605	**0.613**	0.633
NOC	6	6	**6**	6
SEP	0.720	0.710	**0.703**	0.685
r^{2-} Conv	0.963	0.953	**0.975**	0.969
SEE	0.216	0.246	**0.178**	0.20
F	70.047	53.78	**105.301**	82.1
r^2 Pred	0.8816	0.8465	**0.8940**	0.8477

NOC = number of components, SEE = standard error of estimate, SEP = standard error of prediction.

molecule of the series. Since CF_3 and CH_3 are almost of the same size they were not differentiated and the predicted pIC_{50} value of 8.17 of **2** was close to that of **1** (7.85). The importance of a bulky substituent corresponding to R_2 is manifested by the decreased pIC_{50} value of 7.97 (actual 8.04) in **4** having less bulky NMe_2 in place of Me_2CH. Similarly the predicted pIC_{50} value of 8.40 in **3** having the *t*-Bu group as R_2 is indicative of the importance of a bulky group at this position in imparting better biological activity. The green contour covering C_3–C_5 region of the R_4 aryl ring indicated that compounds with bulky substitution at these positions should possess good biological activity as observed in **12**, **14**, **15**, **21** and **22**. The increase in the actual pIC_{50} value from 8.70 to 9.30 in replacing the chlorine at C_3 of the R_4 aryl ring in **21** by the bulkier bromine as in **22** provides distinct evidence of the influence of the steric factor in this region. A yellow contour observed near the C_3 position of the R_5 aryl ring explained the intolerance to steric bulk in this region. Thus, **6**–**8**, **10**, **11**, **27** and **28** containing bulky substitutions in this area showed less biological activity (actual pIC_{50} value less than that of **18** and **25**). A red contour at C_4 of R_5 aryl ring showed the importance of electronegative atom such as F at this position in imparting better biological activity. This is reflected in the decreased biological activity of **6**–**8**, **10** and **26**–**28** that are devoid of the F substitution at this position. The overall inferior biological activity displayed by **6**–**8**, **27** and **28** may be due to the combined effect of the presence of a bulky substituent at C_3 and the lack of the presence of F at C_4 of the R_5 aryl ring. However, the decreased biological activity of **10** and **11** reflects solely the effect of bulky substituent at C_3 of the R_5 aryl ring as the replacement of the methyl group by the bulkier ethyl group induces a decrease in the actual pIC_{50} value from 7.52 to 7.00. Whereas, the poor biological activity of **26** is the result of the absence of F at C_4 of the R_5 aryl ring. The fact that **13** with the 3-pyridyl moiety as R_4 showed biological activity compared to that of **18** may be due to the reason that hydration through the pyridine electron lone pair increases the steric crowd in this region.

The hydrophobic field (majenta) contours (Fig. 5) observed near C_3 of R_5 aryl ring and C_2, C_4 of the R_4 aryl ring suggest that these positions are not suitable for substitution with hydrophobic group. Thus, **6**–**8**,**10**, **11**, **27** and **28** having a methyl, chloro or bromo substitution at C_3 of the R_5 aryl ring exhibited decreased biological activity (compared to that of **18** and **25**). The presence of the hydrophilic nitrogen lone electron pair of the 4- and 2-pyridyl groups in **23** and **24**, respectively, may be the reason for the very good biological activity of these compounds. The yellow contours near C_3 and C_6 positions of the R_4 aryl ring indicates that introduction of hydrophobic moieties at these positions should improve the biological activity. This accounts for the better biological activity of **12**, **15**, **21** and **22** in having a chloro, methylamino or bromo substituent at C_3 of the R_4 aryl ring. The yellow contour near the C_6 of R_4 extends over the region of the C_4 carbon of the central five-membered ring. Thus, **25** with bromo substitution in this region produced better biological activity. The

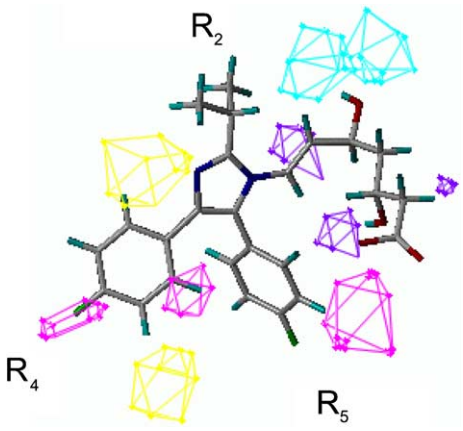


Figure 5. CoMSIA hydrophobic and donor contour maps with the molecule **18**.

CoMSIA steric contour maps produced at C_3 of R_4 aryl ring and C_3 of R_5 aryl ring were comparable with the hydrophobic contours. The purple contour represented the position where hydrogen bond donor fields disfavour the biological activity and the cyan contours showed that the presence of a donor group in this region should produce better biological activity (Fig. 5). The importance of the donor field is best represented by the orientation of the hydroxyl group in the region of the cyan contour (Fig. 5) and indicates the importance of 5*S* configuration of the hydroxyl group in the side chain. In the 5*R* configuration of the hydroxyl group in **16** and **17**, the orientation of hydroxyl group is away from the region occupied by the cyan contour and explains the poor biological activity in the inactive chiral form. Thus, the 3D-QSAR CoMSIA model developed under this study predicts the stereochemical requirement of the ligand to exhibit good biological activity. Our results are in conformity of the recent crystallography studies[5] that the bioactive chiral form of heptenoate side chain of statins having the 3*R*,5*S* configuration undergoes various hydrogen bonding and salt bridge interactions with polar residues in the active site. The poor biological activity of **29** emphasizes the requirement of the ethene linker between the heterocyclic ring and the 3-HMG moiety.

The 3-D QSAR study of 9,9-bis-(4-fluorophenyl)-3,5-dihydroxy-8-subtituted-6,8-nonadienoic acids as HMG-CoA reductase inhibitors was earlier reported by Sit et al.[21] The significant difference of the present work with this report is that the work of Sit et al. was to delineate the topographical and physio-chemical features of the binding site from the IC_{50} values in the absence of any molecular information concerning the active site of HMG-CoA reductase and the approach consists of exploring the sterically allowed conformations relative to the butadiene unit.[21] The present study is based on the structure of the active site derived from the X-ray crystal structure[5] and the bioactive conformation of fluvastatin, a known HMG-CoA inhibitor, extracted from the HMGR–inhibitor complex.[15] However, the results of the present study corroborate the assumptions made in the work of Sit et al.[21] based on topological approach

to receptor mapping to quantify potential intermolecular steric effects. The literature report considered overlapping volume as a molecular shape descriptor and it was concluded that the inactive molecules may experience detrimental interactions at the active site and thus the bioactivity of the compounds is conditioned by intermolecular interactions between the substituent attached to the tetrazole moiety at C-8 and its corresponding binding site.[21] In the present work the molecular design was based on the various electrostatic, steric and hydrophobic effects of the substituents on the aryl rings attached to the central heterocyclic moiety and the substituent directly attached to the central ring.

3. Conclusion

CoMSIA studies on 29 3,5-dihydroxy heptenoates were carried out to develop a 3-D QSAR model that provided good internal and external predictivity. The resulted model can be extrapolated to predict novel and more potent molecules. The contour maps obtained from the CoMSIA analysis guide to design new chemical entities with high HMGR inhibitory activity.

4. Molecular modelling

The molecular modelling and 3-D QSAR studies were performed on a Silicon Graphics Octane 2 workstation using SYBYL 6.8.[22] The bioactive conformation of fluvastatin extracted form 1HW1.pdb[20] was used to model the selected molecules. Geometry optimization was performed using Tripos force field, Powell method[23] including Gasteiger–Huckel[24] charges till the gradient convergence 0.05 kcal/mol was reached.

4.1. Alignment rules

One of the important steps in the 3-D QSAR methods is the active conformation and alignment of molecules. The success of these methods strongly depended on the relative positioning of the ligands in the fixed lattice, before the generation of 3-D descriptors. We have performed database alignment method. The fragment used for the alignment is shown in Figure 2.

4.2. CoMSIA interaction energy fields

The recently reported CoMSIA method is based on molecular indices. Using a common probe atom, similarity indices are calculated for a data set of prealigned molecules at regularly spaced grid points. When the atoms of the molecules approach the probe very nearer then there will be sudden rise in energy. Therefore, the cut off value of 30 kcal/mol is included in CoMFA. This restriction may give some false values, which sometimes lead to error in the predictions. The GAUSSIAN type distance dependant function forms used by CoMSIA method to calculate such properties overcome this problem. Similarity indices can be calculated at all grid points inside as well as outside the molecules and are subsequently evaluated in a PLS analysis.

4.3. PLS analysis

The regression analysis of CoMSIA field energies was performed using the partial least squares (PLS) algorithm with the leave-one-out (LOO) method adopted for cross-validation. The cross-validation was performed to obtain the optimum number of components which were then used in deriving final CoMSIA model without cross-validation. The column filtering value (σ) was set to 2.0 for cross-validated final analysis was carried out to calculate the conventional r^2 value using the optimum number of components.

4.4. Predicted r^2 value

To validate the derived models, biological activities of the test set molecules were predicted using models derived form training set. Predictive r^2 value was calculated using formula. r^2 Predictive = (SD − PRESS)/SD. Where SD is the sum of squared deviation between the biological activities of the set molecule and the mean activity of the training set molecules and PRESS is the sum of squared deviations between the actual and the predicted activities of the test set molecules.

Acknowledgements

R.T. thanks to CSIR, New Delhi, India for award of Senior Research Fellowship.

References and notes

1. Grundy, S. M. *J. Am. Med. Assoc.* **1986**, *256*, 2849.
2. Thelle, D. S. *Drug Invest.* **1990**, *2*(Suppl. 2), 1.
3. Hawk, M. A.; Cesen, K. T.; Siglin, J. C.; Stoner, G. D.; Ruch, R. J. *Cancer Lett.* **1996**, *109*, 217–222.
4. Caruso, M. G.; Notarnicola, M.; Santillo, M.; Cavallini, A.; Di Leo, A. *Anticancer Res.* **1999**, *19*, 451.
5. Istavan, E. S.; Deisenhofer, J. *Science* **2001**, *292*, 1160–1164.
6. Turabi, N.; DiPietro, R. A.; Mantha, S.; Ciosek, C.; Rich, L.; Tu, J.-I. *Bioorg. Med. Chem.* **1995**, *3*, 1479–1485.
7. Wantnabe, M.; Koike, H.; Ishiba, T.; Okada, T.; Seo, S.; Hirai, K. *Bioorg. Med. Chem.* **1997**, *5*, 437–444.
8. Coppola, G. M.; Damon, R. E.; Yu, H.; Engstrom, R. G. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 549–554.
9. Chakraborti, A. K.; Gopalakrishnan, B.; Sobhia, M. E.; Malde, A. *Eur. J. Med. Chem.* **2003**, *38*, 975–982.
10. Chakraborti, A. K.; Gopalakrishnan, B.; Sobhia, M. E.; Malde, A. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2473–2479.
11. Chakraborti, A. K.; Gopalakrishnan, B.; Sobhia, M. E.; Malde, A. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1403–1408.
12. Chakraborti, A. K.; Thilagavathi, R. *Bioorg. Med. Chem.* **2003**, *11*(18), 3989–3996.
13. Desiraju, G. R.; Gopalakrishnan, B.; Jetti, R. K.; Nagaraju, A.; Raveendra, D.; Sarma, J. A.; Sobhia, M. E.; Thilagavathi, R. *J. Med. Chem.* **2002**, *45*, 4847–4857.
14. Desiraju, G. R.; Sarma, J. A. R. P.; Raveendra, D.; Gopalakrishnan, B.; Thilagavathi, R.; Sobhia, M. E.; Subramanya, H. S. *J. Phys. Org. Chem.* **2001**, *14*, 481–487.

15. Samuel, P. M.; Vos, D.; Raveendra, D.; Sarma, J. A. R. P.; Roy, S. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 61–64.
16. Liu, G.; Zhang, Z.; Luo, X.; Shen, J.; Liu, H.; Shen, X.; Chen, K.; Jiang, H. *Bioorg. Med. Chem. Lett.* **2004**, *12*, 4147–4157.
17. Klebe, G.; Abraham, U.; Meitzner, T. *J. Med. Chem.* **1994**, *37*, 4130–4146.
18. Chan, C.; Bailey, E. J.; Hartley, C. D.; Hayman, D. F.; Hutson, J. L.; Ingalis, G. G. A.; Jones, P. S.; Keling, S. E.; Kirk, B. E.; Lamont, R. B.; Lester, M. G.; Prithard, J. M.; Barry, C. P.; Scicinski, J. J.; Spooner, S. J.; Smith, G.; Steeples, I. P.; Watson, N. S. *J. Med. Chem.* **1993**, *36*, 3646–3657.
19. Procopiou, P. A.; Draper, C. D.; Hutson, J. L.; Ingalis, G. G. A.; Ross, B. C.; Watson, N. S. *J. Med. Chem.* **1993**, *36*, 3658–3662.
20. Abola, E. E.; Berstein, F. C.; Bryant, S. H.; Koetzle, T. F.; Weng, J. Protein Data Bank. In *Crystallographic Database–Information Content, Software Systems, Scientific Applications*; Allen, F. H., Berjerhoff, G., Sievers, R., Eds.; Data Commission of the International Union of Crystallography: Bonn, 1987; p 171.
21. Sit, S. Y.; Parker, R. A.; Motoc, I.; Han, W.; Balasubramanian, N.; Catt, J. D.; Catt, J. D.; Brown, P. J.; Harte, W. E.; Thompson, M. D.; Wright, J. J. *J. Med. Chem.* **1990**, *33*, 2982.
22. Hanley, S. Tripos Associates, Inc., St. Louis, MI 63144, USA, 1699.
23. Powell, M. J. D. *Math. Programming* **1977**, *12*, 241.
24. Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219–3228.