ELSEVIER

Pharmacophore identification and 3D-QSAR studies in N-(2-benzoyl phenyl)-L-tyrosines as PPARγ agonists[☆]

Lalit Rathi, Sushil K. Kashaw, Anshuman Dixit, Gyanendra Pandey and Anil K. Saxena*

Medicinal Chemistry Division, Central Drug Research Institute, Lucknow-226001, India

Abstract—The identification of pharmacophore and three dimensional quantitative structure–activity studies have been performed on a set of N-(2-Benzoylphenyl)-L-tyrosine for their PPARγ agonist activity by using the logico-structural based software Apex 3D—which describes the properties and distribution of primary and secondary biophore sites in the three dimensional space. Among several models, two models of comparable probability were selected on the basis of $R^2 > 0.60$, chance ≤ 0.04, size < 4, match > 0.20. These models showed a good correlation between the observed and predicted biological activity both for training and test sets.
© 2003 Elsevier Ltd. All rights reserved.

1. Introduction

Of the estimated six million individuals diagnosed with diabetes mellitus in United States, 90% are characterized as non-insulin dependent (NIDDM, type-II). Insulin resistance is characterized by impaired uptake and utilization of glucose in insulin sensitive target organs such as adipocytes and skeletal muscle and by impaired inhibition of hepatic glucose output.[1] The insulin resistant state at the peripheral level causes impaired glucose utilization leading to hyperglycemia, which may also play a role in the etiology of a wide spectrum of metabolic disorders such as obesity, hypertension, atherosclerosis, neuropathy, nephropathy, retinopathy etc.[2,3] So a tight blood glucose control is necessary with a reduce incidence of insulin resistance of peripheral tissues. As the exercise enhances tissue responsiveness to insulin, the combination of diet and exercise is the primary treatment for NIDDM patients.[4] Since, dietary adherence is difficult for most patients, medication is necessary for many NIDDM patients. Most commonly employed oral hypoglycemic agents are sulfonylureas and biguanides which have the disadvantage such as primary and secondary failure of efficacy as well as the potential for induction of severe hypoglycemia.[5] So there is a need of a new candidate molecule which may

effectively reduce insulin resistance or potentiate insulin action in genetically diabetic and or obese animals. The search for the drugs that reverse the insulin resistance without stimulating insulin release from β-cells also fulfill a major medical need in the treatment of NIDDM. Hence the search of such drugs with a potential to reduce long term complications of NIDDM, is of current interest. The pioneering discovery of ciglitazone, a thiazolidine-2,4-dione derivatives (TZDs) and several of its analogues have led to the identification of possible molecular targets peroxisome-proliferator activated receptorγ (PPARγ), selectively expressed in adipocytes.[6–10] The TZDs are found to be the promising compounds capable of ameliorating NIDDM by improving insulin resistance without inducing hypoglycemia.[11] These agents substantially increase insulin sensitivity in muscle, liver and adipose tissue, resulting in the correction of elevated plasma level of glucose without the occurrence of hypoglycemia. But undesirable effects associated with glitazones have been observed in animal and human studies, which include cardiac hypertrophy, haemodilution and severe liver toxicity.[12,13] So the scientists are trying to develop a new drug which should be equipotent to that of TZD but devoid of any toxicity. Though it still remains unclear whether the side effects are caused by the mechanism of action of these compounds or originate within the TZD chemical structure common to this class. So the emphasis is on the development of non-thiazolidinedione PPARγ agonists, which might surmount the problem, associated with the known TZDs and thus may offer an advantage as an anti-diabetic agent.

In view of above, there is a need for the development of 3D-QSAR models and identification of pharmcophore in terms of essential structural and electronic features important for the PPARγ agonistic activity. Though some 3-D-QSAR studies of TZDs have been reported however no such studies have been carried out on non-TZDs. Among various approaches available for 3D-QSAR, logico-structural based approach (Apex-3-D) has been used to identify the pharmacophore and to derive 3D-QSAR models in recent years[14,15] so the same approach has been applied in *N*-(2-benzoyl phenyl)-L-tyrosines PPARγ agonists and the results are described in this paper.

2. Materials and methods

All molecular modeling and 3D-QSAR studies were performed using Apex-3-D software on a Silicon Graphics INDY R-4000 workstation employing molecular simulation Incorporation (MSI) software[16] (InsightII, builder, Discover, Search-compare, Apex-3-D). Apex-3-D expert system is based on logico-structural approach to drug design developed by Golender et al. and is used for classification and prediction of biological activity.[17,18]

A total of 23 molecules belonging to three major prototypes were divided into a training set (Compd. no **1–20**) and a test set of three compounds (Compd. no.**21–23**) representing each prototype (Table 1).[19] Some of the molecules without proper biological activity had to be removed from these studies. The structures of all compounds were constructed using the sketch programme in the builder module of InsightII software[20] and were converted to 3D for optimization of their geometry (net charge 0.0) by selecting the force field viz. potential action, partial charge action and formal charge action as fixed. The molecular structures were finally minimized using the Steepest Descent, Conjugate Gradients and Newton–Raphson's algorithms in sequence followed by Quasi-Newton–Raphson(va09a),[21] optimization techniques implemented in the Discover module (ver 2.9) by using energy tolerance value of 0.001 kcal/mol and maximum number of iteration set to 1000. In view of our earlier findings that the total energy of the molecular conformation obtained through above standard energy minimization procedure do not differ much with the near global minimum energy conformation,[22] these molecular structure was stored in MDL format and were used for the computational calculation of different physico-chemical properties like atomic charges, Pi-population, electron donor and acceptor indexes, HOMO and LUMO coefficients and hydrophobicity and molar refractivity based on atomic contributions by MOPAC 6.0 (MNDO Hamiltonian)[23] module. The data was used by Apex-3-D programme for automated identification of biophores (pharmacophores), superimposition of compounds and quantitative model building. Biophores represent a certain structural and electronic pattern in a bioactive molecule that is responsible for its activity through interaction with the receptor. These biophores can be regarded as the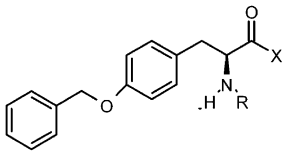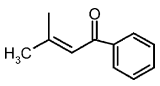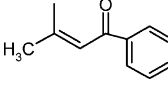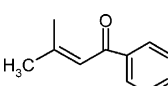 local array of descriptor centers (user-defined atoms, pseudoatoms like ring centers, hydrophobic regions or hydrogen binding sites) which are common to a series of all molecules in their bioactive confirmations, responsible for biological activity through binding with the receptors were used to derive 3D QSAR equation with the settings of radius at 0.60, the occupancy at 6.0, the sensitivity at 0.80 and random samples at 100.
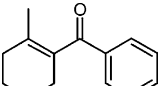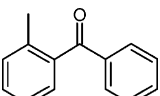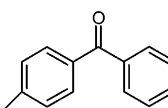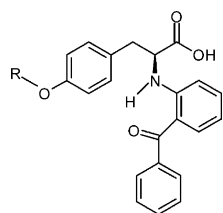
The biophoric sites were set to pi-population, charge, hydrogen donor, hydrogen acceptor, HOMO, LUMO, hydrophobicity and refractivity. The secondary sites were set to hydrogen acceptor, presence; hydrogen donor, presence; heteroatom, presence; ring, presence; hydrophobic, hydrophobicity; steri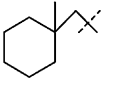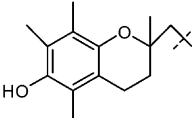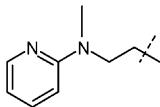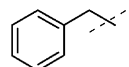c, refractivity. Quality of the each model was estimated from the observed R^2 (correlation cofficient), RMSA (calculated root mean square error based on all compounds with degrees of freedom correction), RMSP (calculated root mean square error based on 'leave-one-out' with no degree of freedom correction), chance statistics (evaluated as the ratio of the equivalent regression equations to the total number of randomized sets; a chance value of 0.01 corresponds to 1% chance of fortuitous correlation and match parameter (evaluated for the quality of superimposition for molecules having common biophores)

3. Results and discussion

In vitro activity data of *N*-(2-benzoylphenyl)-L-tyrosine for PPARγ agonists activity (converted into logarithmic value) is given in Table 1. Several 3D-pharmacophore models with different size and arrangements (hydrophobicity, refractivity, hydrogen bonding sites, hydrogen bond donor, hydrogen bond acceptor) were generated for training set. After the generation of 3D-pharmacophore models, the compounds left for the test set were predicted (Table 1) so as to check the validity of the 3D-QSAR models. Among the several 3D-pharamacophore models with different size and spatial arrangements (center of aromatic rings, hydrophobicity, refractivity, hydrogen bond donor, hydrogen bond acceptor, hydrogen binding site etc.), two models: model no.1 (Fig. 1) and model No.2, (Fig. 2) of comparable probability were selected, based on RMSA (calculated root mean square error based on all compounds with degrees of freedom correction) = 0.04, RMSP (calculated root mean square error based on 'leave-one-out' with no degree of freedom correction) = 0.04, $R^2 > 0.65$, chance ≤ 0.05, match value > 0.20, size = 3 number of compounds **19** and **18**, respectively (Table 2). All the primary and secondary biophoric sites responsible for pharmacological activity were same in both the models (Fig. 3) except compound No.15 in model No. 1 (Fig. 4) where primary biophoric site A has been changed from the oxygen atom attached to the oxazole ring, to the nitrogen atom present in the side chain of the benzoxazole ring. In figure, site A, B and C represent the three biophoric sites corresponding to the oxygen atom attached to the phenyl ring, oxygen atom of carboxylic group and electron density clouds over oxygen atom attached to the phenyl ring respectively. The spatial disposition of the biophoric sites A

Table 1. In vitro activity data (converted into logarithmic value) of *N*-(2-benzoylphenyl)-L-tyrosine as PPARγ agonists; Compound no. **1–20** training set ; Compound no. **21–23** test set

Compd	Substituents		Obs. Act. log Pki	Model no.1		Model no.2	
	X	R		Cal. via model 1	Pred. via model 1	Cal. via model 2	Pred. via model 2
Training set							
1	OH		0.899	0.82	0.79	0.82	0.79
2	NH₂		0.769	0.76	0.75	(—)[a]	—
3	OCH3		0.786	0.83	0.84	0.84	0.86
4	OH		0.785	0.79	0.79	0.80	0.80
5	OH		0.831	0.85	0.85	0.82	0.82
6	OH		0.770	0.85	0.86	0.82	0.84
7			0.862	0.87	0.87	0.88	0.89
8			0.918	0.92	0.92	0.94	0.95
9			0.946	0.89	0.89	0.91	0.90
10			0.847	—	—	—	—

Table 1 (*continued*)

Compd	Substituents		Obs. Act. log Pki	Model no.1		Model no.2	
	X	R		Cal. via model 1	Pred. via model 1	Cal. via model 2	Pred. via model 2
11			0.95	0.92	0.92	0.94	0.94
12			0.951	0.94	0.94	0.93	0.93
13			0.877	0.86	0.86	0.87	0.87
14			0.879	0.87	0.87	0.88	0.89
15			0.874	0.85	0.84	0.83	0.82
16			0.849	0.85	0.85	0.87	0.87
17			0.918	0.93	0.93	0.92	0.92
18			0.928	0.93	0.93	0.92	0.92
19			0.839	0.87	0.87	0.85	0.86
20			0.831	0.85	0.86	0.83	0.83
Test Set							
21			0.756		0.68		0.67

Table 1 (*continued*)

Compd	Substituents		Obs. Act. log Pki	Model no.1		Model no.2	
	X	R		Cal. via model 1	Pred. via model 1	Cal. via model 2	Pred. via model 2
22			0.913	0.89		0.88	
23			0.925	0.84		0.82	

[a] (−)activity not present in model.

Figure 1. Photograph showing the superimposition of all the 19 compounds with biophoric sites (solid sphere) and secondary sites (red circle) via model no.1.
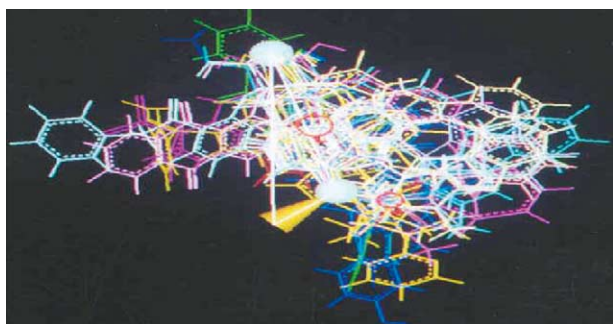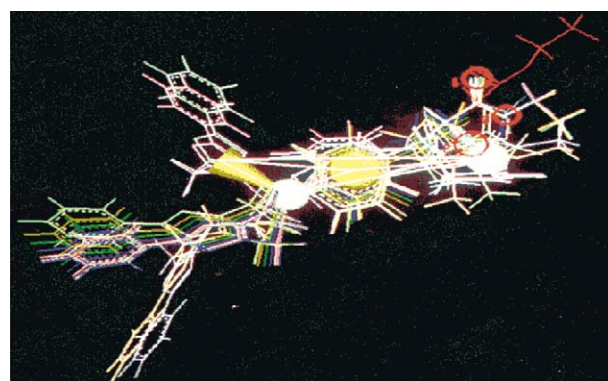

Figure 2. Photograph showing the superimposition of all the 18 compounds with biophoric sites (solid sphere) and secondary sites (red circle) via model no.2.

and B in the models involved in the interaction with the specific receptor site in addition to site C most likely involved in hydrogen bonding. The pharmacological activity depends not only on the physico-chemical properties of biophoric sites A (π population 0.262 ± 0.018 charge_heteroatom: -0.284 ± 0.064); DON_0.1: 8.235 ± 0.05),

Table 2. 3D-QSAR models describing correlation and statistical reliability for PPARγ agonist activity

M[a]	RMSA	RMSP	R^2	Chance	Size	Match	V[b]	C[c]
1	0.04	0.04	0.66	0.01	3	0.21	2	19
2	0.04	0.04	0.69	0.05	3	0.21	3	18

[a] Model no.
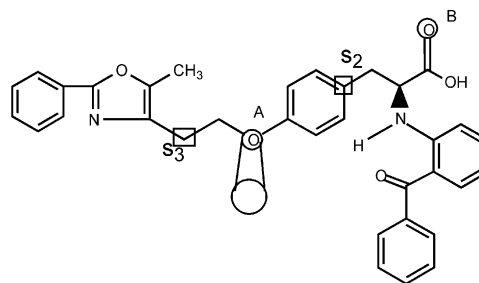[b] Number of variables.
[c] Number of compounds.

Figure 3. Pictorial representation of biophoric sites O (A,B,C) and secondary □ sites (S2 and S3) represented on the most active compound no. **11**.

B (π-population 1.014 ± 0.02; charge_heteroatom: -0.306 ± 0.033 DON_01: 8.421 ± 0.07) in model no.1 and A (π-population: 0.263 ± 0.02; charge_heteroatom: -0.294 ± 0.02; DON: 8.242 ± 0.06), B (π-population: 1.020 ± 0.012; charge heteroatom: -0.304 ± 0.01; DON_01: 8.430 ± 0.07) in model no.2 but also on their spatial arrangement in terms of mean biophoric distances (in Å unit) between site A–B (8.584), B–C (10.300) and C–A (2.999) in model no.1 and A–B (8.591), B–C (10.298) and C–A (2.999) in model no.2. Sites A and B are electron-rich sites capable of donating electrons and may involve in electrostatic, ionic and π–π interactions. Site C, which is an electronic cloud on the oxygen atom (Fig. 1) may form hydrogen bonding

between ligand and receptor. All these physico-chemical properties, distances and the spatial arrangement of the biophoric sites are important to determine the binding affinity of the ligand to the active site of PPARγ receptors, the potency and extend of pharmacological response.

In addition to the identification of three essential structural features described above as biophoric sites for all the molecules, 3D-QSAR equations (Eqs. 1 and 2) was derived using the above pharmacophore as a template for superimposition.

$$
\begin{aligned}
BA = {}& 0.004(\pm0.001)\ \text{Total refractivity} + 0.017 \\
& \times (\pm0.005)\ \text{Steric refractivity at}\ S_2 \\
& + 0.308
\end{aligned}
\tag{1}
$$

$$n = 19,\ r = 0.814,\ F_{(2,16)} = 15.694$$

$$
\begin{aligned}
BA = {}& 0.004(+0.001)\ \text{Total refractivity} + 019 \\
& \times (+0.005)\ \text{Steric refractivity at}\ S_2 \\
& - 0.010(+0.005)\ \text{Steric refractivity at}\ S_3 \\
& + 0.252
\end{aligned}
\tag{2}
$$

$$n = 18,\ r = 0.83,\ F_{(3,16)} = 10.36$$

These equations describe physicochemical properties of the total molecule (global) and of secondary sites S_2 and S_3 as independent and the PPARγ receptor agonistic activity (converted into logarithmic value) as dependent parameter. Secondary sites are also important for a compound to show PPARγ agonistic activity, parameter values of which are given in Table 3, No. 5. Both the models have same secondary site S_2 on the carbon atom of the phenyl ring attached to the propionic acid group. In addition, models No 2 also possess one more secondary site indicated by S_3 (Fig. 4).

The steric refractivity index at the refractivity site S_2 in the vicinity of site A and B in both the models and S_3 site in the vicinity of substituent R in model no.2 best described the observed variation in PPARγ receptor agonistic activity in these molecules with a good correlation coefficient r (0.81 and 0.83 for model 1 and 2 respectively) and of high statistical significance ($F_{(2,16\alpha0.001)} = 12.7$; $F_{(2,16)} = 15.6$ (>99.9%) and $F_{(3,14\alpha0.002)} = 9.73$; $F_{(3,14)} = 10.36$ (>99.8%) for model 1 and 2 respectively) (Eqs. 1 and 2).
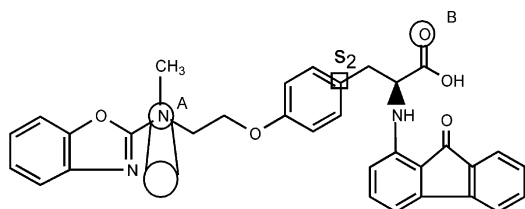

Figure 4. Pictorial representation of biophoric sites O (A,B,C) and secondary □ sites (S2 and S3) represented on the compound no. **15**.

These 3D-QSAR models showed zero difference between RMSA and RMSP variables which is also reflected in the comparison of observed, calculated and predicted (Leave-One-Out) values of biological activity (Table 1). Total refractivity, which is a global property, contributes positively for the PPARγ agonistic activity in both the models. Steric refractivity on the carbon atom of the phenyl ring indicated by S_2 positively contributes towards the biological activity that is, large bulky groups at this site is favorable for the PPARγ agonist activity in both the models. The site S_3 (model no.2) on the carbon atom attached to the oxazole ring contributes negatively for the steric interactions that is, small chemical moieties at this site will increase the potency of the molecules. Parameter values for secondary sites S2 and S3 in model 1 and 2 are shown in Table 3. These models were used to predict the biological activity of test set (Comp. no. **21**–**23**) to check validity of the models where a high correlation (0.92 for model no.1 and 0.89 model no.2) between observed and predicted activities for the test set compounds suggests that the models can be useful in designing the PPARγ agonists molecules for the treatment of hyperglycemia (Table 1).

4. Conclusion

The 3D-QSAR studies on *N*-(2-benzoylphenyl)-L-tyrosines for PPARγ agonistic activity resulted in the identification of pharmacophore in terms of physico-chemical properties, distances and the spatial arrangement of the three important biophoric sites A,B and C for binding affinity of the ligand to the active site of PPARγ receptors. The two 3D-QSAR models describing the properties and distributions of biophoric and secondary biophoric sites, showed a good correlation

Table 3. Parameter values for secondary sites S2 and S3 in model 1 and 2

| Compd | Model 1 | | Model 2 | | |
	T. R.[a]	S.R.[b] at S_2	T. R	S.R. at S_2	S.R. at S_3
1.	0.44	0.07	0.50	0.07	–
2.	0.45	(−)[c]	—	—	–
3.	0.46	0.07	0.52	0.07	–
4.	0.48	−0.07	0.54	0.07	−0.04
5.	0.48	0.07	0.54	0.07	−0.04
6.	0.48	0.07	0.54	0.07	–
7.	0.50	0.07	0.56	0.07	–
8.	0.55	0.07	0.62	0.07	–
9.	0.52	0.07	0.49	0.07	–
10.	0.55	0.07	0.62	0.07	−0.03
11.	0.57	0.07	0.64	0.07	–
12.	0.55	—	0.62	—	–
13.	0.56	—	0.63	—	−0.03
14.	0.54	—	0.61	—	–
15.	0.55	—	0.61	—	−0.03
16.	0.57	0.06	0.61	0.06	−0.03
17.	0.57	0.06	0.64	0.06	–
18.	0.56	—	0.63	—	−0.03
19.	0.54	—	0.61	—	−0.03

[a] Total refractivity.
[b] Steric refractivity.
[c] (—) indicates absence of property.

between the observed and predicted activity both in training and test set and thus may be useful in designing new chemical entities as PPARγ agonists.

Acknowledgements

References and notes

1. Reaven, G. M. *Diabetes* **1988**, *37*, 1595.
2. DeFronzo, R. A.; Bonadonna, R. C.; Ferrannini, E. *Diabetes Care* **1992**, *15*, 318.
3. Reaven, G. M. *Metabolism* **1980**, *29*, 445.
4. Isley, W. L. *Drugs Today* **1990**, *26*, 59.
5. Dow, T. L.; Beehle, B. M.; Chow, T. T.; Clard, D. A.; Hulin, B.; Stevenson, R. W. *J. Med. Chem.* **1991**, *34*, 1538.
6. (a) Fujita, T.; Sugiyama, Y.; Taketomi, S.; Sohda, T.; Kawamatsu, Y.; Iwatsuka, H.; Suzuoki, Z. *Diabetes* **1983**, *32*, 804. (b) Chang, A. Y.; Wyse, B. M.; Gilchrist, B. J.; Peterson, T.; Diani, A. R. *Diabetes* **1983**, *32*, 830.
7. Rieusset, J.; Andreelli, F.; Auboeuf, D.; Roques, M.; Vallier, P.; Riou, J. P.; Auwerx, J.; Laville, M.; Vidal, H. *Diabetes* **1999**, *48*, 699.
8. Willson, T. M.; Cobb, J. E.; Cowan, D. J.; Wiethe, R. W.; Correa, I. D.; Prakash, S. R.; Beckd, K. D.; Moore, L. B.; Kliener, S. A.; Lehmann, J. M. *J. Med. Chem.* **1996**, *39*, 665.
9. Lehmann, J. M.; Moore, L. B.; Smith-Oliver, T. A.; Wilkison, W. O.; Wilson, T. M.; Kliener, S. A. *J. Biol. Chem.* **1995**, *270* (22), 15953.
10. Willson, T. M.; Brown, P. J.; Sterbach, D. D.; Henke, B. R. *J. Med. Chem.* **2000**, *43*, 527.
11. Oguchi, M.; Wada, K.; Honma, H.; Tanaka, S.; Kaneko, T.; Sakakibara, S.; Ohsumi, J.; Serizawa, N.; Fujiwara, T.; Horikoshi, H.; Fujita, T. *J. Med. Chem.* **2000**, *44*, 3052.
12. Ishii, S.; Wasaki, M.; Ohe, T.; Veno, H.; Tanaka, H. *MCC 555 Diabetes* **1996**, *45* (2), 141A (abstract).
13. Watkins, P. B.; Whitecomb, R. *N. Engl. J. Med.* **1998**, *338*, 916.
14. Kashaw, S. K.; Rathi, L. G.; Mishra, P.; Saxena, A. K. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2481.
15. Arockia Babu, M.; Shakya, N.; Prathipati, P.; Kaskhedikar, S. G.; Saxena, A. K. *Bioorg. Med. Chem.* **2002**, *10*, 4035.
16. Insight II, Version 2.3.0, San Diego: Biosym Technologies: 1993.
17. Golender, V. E.; Vorpogel, E. R. Computer Assisted-Pharmacophore Identification 3D-QSAR in Drug Design. *Theory, Methods and Applications*, 137.
18. Golender, V. E., Rosenblit, A. B. Logical and Combinatorial Algorithms in Drug Design. 1983.
19. Henke, B. R.; Blanchard, S. G.; Brackeen, M. F.; Brown, K. K.; Cobb, J. E.; Collins, J. L.; Harrington, W. W.; Hashim, M. A., Jr.; Hull-Ryde, E. A.; Kaldor, I.; Klieuer, S. A.; Lake, D. H.; Leesnitzer, L. M.; Lenmann, J. M.; Lenhard, J. M.; Noble, S. A., Jr.; Oliver, W., Jr.; Parks, D. J.; Plunket, K. D.; Szewezyk, J. R.; Wilson, T. M. *J. Med. Chem.* **1998**, *41*, 5020.
20. Dinur, U.; Hagler, A. T. *Review of Computational Chemistry* **1991**, *2*, 4.
21. Powell, M. J. D. Math. Program. 197, 241
22. Pandey, T.; Pandey, S. K.; Tiwari, M.; Chaturvedi, S. C.; Saxena, A. K. *Bioorg. Med. Chem.* **2000**, 1.
23. Stewart, J. J. P. *QCPE Bull.* **1990**, 455.