

## 3D-QSAR study of ring-substituted quinoline class of anti-tuberculosis agents

Amit Nayyar,<sup>a</sup> Alpeshkumar Malde,<sup>b</sup> Rahul Jain<sup>a,\*</sup> and Evans Coutinho<sup>b,\*</sup>

<sup>a</sup>Department of Medicinal Chemistry, National Institute of Pharmaceutical Education and Research,  
Sector 67, S.A.S. Nagar, Punjab 160 062, India

<sup>b</sup>Department of Pharmaceutical Chemistry, Bombay College of Pharmacy, Kalina, Santacruz (E), Mumbai 400 098, India

Received 9 August 2005; revised 2 September 2005; accepted 2 September 2005

Available online 7 October 2005

**Abstract**—A 3D-QSAR analysis of a new class of ring-substituted quinolines with anti-tuberculosis activity has been carried out by three methods—Comparative Molecular Field Analysis (CoMFA), CoMFA with inclusion of a hydropathy field (HINT), and Comparative Molecular Similarity Indices Analysis (CoMSIA). The conformation of the molecules was generated using a simulated annealing protocol and they were superimposed using features common to the set with *database alignment* (SYBYL) and field fit methods. Several statistically significant CoMFA, CoMFA with HINT, and CoMSIA models were generated. Prediction of the activity of a set of test molecules was the best for the CoMFA model generated with *database alignment*. Based upon the information contained in the CoMFA model, we have identified some novel features that can be incorporated into the quinoline framework to improve the activity.

© 2005 Elsevier Ltd. All rights reserved.

### 1. Introduction

Tuberculosis (TB) causes more than two million deaths each year mainly in the impoverished parts of the world.<sup>1</sup> The condition of patients suffering from tuberculosis has further complicated and worsened because of the association of the disease with human immunodeficiency virus (HIV) and the emergence of drug-resistant strains of the causative organism *Mycobacterium tuberculosis*.<sup>2</sup> TB and HIV infections complement each other and approximately 10 million adults are infected with both pathogens. New structural classes of drugs, which are affordable and more effective than the presently used drugs, are urgently required to target both, the resistant strains of *M. tuberculosis*, and the emerging pathogens, such as *M. avium*.<sup>3</sup> The past decade has seen significant advances in understanding various molecular aspects of the lifecycle of *M. tuberculosis*, culminating in publication of its complete genome.<sup>4</sup> The identification of new specific targets using the genomic annotation of *M. tuberculosis* is under development and potential target based anti-TB drugs appear to be several years away.<sup>3</sup>

More recently, our research efforts have focused on developing an entirely new structural class of anti-TB agents possibly acting on completely novel targets and with a mechanism of action different from those of the existing drugs. Our approach involves a broad structure-directed screening of new chemical entities against various pathogens including *M. tuberculosis*. This approach has proven to be highly rewarding and has resulted in the identification of a new class of ring-substituted-4-methylquinolines with potent antimycobacterial activity against both drug-sensitive and drug-resistant strains of TB.<sup>5</sup> Structural and biological optimization of this class of compounds has led to the identification of several potent anti-tuberculosis compounds exhibiting activity at 1.0 µg/mL against drug-sensitive strain of *M. tuberculosis* H37Rv.<sup>6–8</sup> A majority of these compounds were readily synthesized in 3–5 steps; each step being facile, high yielding, and involving only inexpensive reagents. The synthetic protocol also permitted us to make various substitutions on both rings of the quinoline moiety to produce a large number of analogs. In an attempt to understand the essential structural requirements for anti-tuberculosis activity of ring-substituted quinolines, we have performed a 3D quantitative structure activity relationship study (QSAR) by Comparative Molecular Field Analysis (CoMFA),<sup>9</sup> Comparative Molecular Similarity Indices Analysis (CoMSIA),<sup>10</sup> and CoMFA with inclusion of a

\* Corresponding authors. Tel.: +91 172 2214682; fax: +91 172 2214692 (R.J.); tel.: +91 22 26670871; fax: +91 22 26670816 (E.C.); e-mail addresses: [rahuljain@nipr.ac.in](mailto:rahuljain@nipr.ac.in); [evans@bcpindia.org](mailto:evans@bcpindia.org)

hydrophobic interaction field (HINT);<sup>11</sup> and the results of this study are reported herein. In the absence of any information regarding the drug target, the indirect ligand-based approaches like 3D-QSAR can assist in understanding the SARs and also serve as a guide in the design of more potent inhibitors.

## 2. Computational details

### 2.1. Dataset

A set of 112 molecules<sup>5–8</sup> was used in this study (Table 1, the molecules with prefixes 1\_, 2\_, 3\_, and 4\_ are from references 5, 6, 7, and 8, respectively). Of these, 18 molecules (1\_2I, 1\_2J, 1\_2K, 1\_3I, 1\_3J, 1\_3K, 2\_13E, 2\_13F, 2\_13G, 2\_13H, 2\_14E, 2\_14G, 2\_14H, 4\_7B, 4\_7C, 4\_7D, 4\_7E, and 4\_7F) with flexible alkyl side chains (chain length greater than 3 carbon atoms) were not included in the generation of the model. The remaining 94 molecules were divided into a training set (70 molecules) and a test set (24 molecules) by means of chemical as well as biological diversity. Daylight fingerprints and pIC<sub>50</sub> data were used to select diverse molecules with the Tanimoto similarity coefficient.<sup>12</sup>

### 2.2. Biological data

The activity of all compounds used in the QSAR study was measured by the same assay<sup>5–8</sup> and is reported as % inhibition at 6.25 µg/mL. For the QSAR study, the activity values were transformed as follows:<sup>13</sup>

$$\text{Activity} = -\log c + \text{logit},$$

where  $c$  is the molar concentration = concentration (µg/mL) \* 0.001/(molecular weight)

$$\text{logit} = \log[\% \text{inhibition} / (100 - \% \text{inhibition})].$$

### 2.3. Molecular modeling

The CoMFA,<sup>9</sup> CoMSIA,<sup>10</sup> and CoMFA with HINT<sup>11</sup> studies were carried out using SYBYL 7.0<sup>14</sup> installed on a Pentium 2.8 GHz PC with the Linux OS (Red Hat Enterprise WS 3.0). The structures were built with the *Sketcher* module and energy minimized by Powell's method using the MMFF94<sup>15</sup> force field with a distance-dependent dielectric term. The minimization was terminated at a maximum value of the gradient at 0.5 kcal/mol/Å. Each structure was then subjected to molecular dynamics (MD) simulation, where it was heated to 700 K for 1 ps and annealed slowly to 200 K in steps of 100 K for 1 ps at each temperature, with a step size of 1 fs and snapshots captured every 5 fs. The lowest energy structure from the MD trajectory was sent through a final round of minimization, which was carried out with the same criteria as mentioned above. For the 18 molecules with flexible side chains (vide supra), a set of local minimum energy conformations within 5 kcal/mol of the lowest energy structure was generated using the *MULTI SEARCH* method in SYBYL. The activities of all these 18 molecules using each one of their low energy conformations were later pre-

dicted using the best 3D-QSAR model to determine the favored orientations of the flexible side chains.

### 2.4. Alignment

The most crucial input for the CoMFA is the alignment of the molecules. The molecule 2\_2H with the highest activity and the least conformational flexibility was chosen as the template and all other molecules were aligned to it using the *DATABASE ALIGNMENT* method in SYBYL (Fig. 1). The molecules were aligned with reference to the quinoline ring. A second alignment was also carried out using the *FIELD FIT* method, wherein the steric and electrostatic fields around the molecules were superimposed over the same fields of the template molecule (Fig. 2).

### 2.5. CoMFA interaction energy calculation

The steric and electrostatic fields in CoMFA were calculated at each lattice intersection of a regularly spaced grid of 2.0 Å in all three dimensions within the defined region. A sp<sup>3</sup> carbon atom with +1.0 charge was used as the probe. The van der Waals potential and Coulombic energy between the probe and the molecule were calculated using the standard Tripos force field. A distance-dependent dielectric constant of 1.0r was used in the calculation of the electrostatics. The steric field was truncated at points where the value exceeded +30.0 kcal/mol, and the electrostatic fields were ignored at those lattice points where the steric interactions were high.

### 2.6. HINT hydropathic field calculation

The hydropathic fields were calculated with the *HINT* module in SYBYL. The region defined for calculation of the CoMFA fields was also used to describe the space in which the *HINT* fields were to be generated. *HINT* calculates the hydrophobic interaction between all atom pairs in a molecule using the following equation

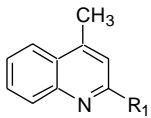
$$B = \sum \sum b_{ij},$$

where,  $b_{ij} = a_i a_j S_i S_j R_{ij} T_{ij}$ ,  $b_{ij}$  is the micro-interaction constant representing the attraction/interaction between atoms  $i$  and  $j$ ,  $a_i$  is the hydrophobic atom constant for atom  $i$ ,  $S_i$  is the solvent accessible surface area for atom  $i$ ,  $R_{ij}$  is the functional distance behavior for the interaction between atoms  $i$  and  $j$ , and  $T_{ij}$  is a discriminant function designed to keep the signs of interactions consistent with the *HINT* convention that favorable interactions are positive and unfavorable interactions are negative.

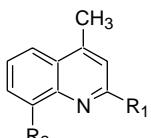
### 2.7. CoMSIA interaction energy calculations

Five CoMSIA fields—steric, electrostatic, hydrophobic, and hydrogen bond donor, and acceptor potentials were calculated at each lattice intersection of a regularly spaced grid of 2.0 Å. A probe atom with radius 1.0 Å, +1.0 charge, hydrophobicity of +1.0, and hydrogen bond donor and acceptor properties of +1.0 was used to calculate steric, electrostatic, hydrophobic, and

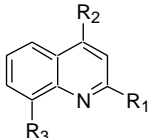
**Table 1.** Dataset<sup>5–8</sup> used for the CoMFA study

<div>  <p><b>Series 1</b></p> </div>						
Molecule	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>	Activity
1_2A	CH(CH <sub>3</sub> ) <sub>2</sub>	—	—	—	—	4.59
1_2C	<i>c</i> -C <sub>4</sub> H <sub>7</sub>	—	—	—	—	4.81
1_2D	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	—	—	—	—	4.98
1_2E	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	—	—	—	—	3.84
1_2F	<i>c</i> -C <sub>3</sub> H <sub>5</sub>	—	—	—	—	4.08
1_2G	1-Adamantyl	—	—	—	—	5.15
1_2H	(CH <sub>2</sub> ) <sub>2</sub> CH <sub>3</sub>	—	—	—	—	4.10
1_2I	(CH <sub>2</sub> ) <sub>3</sub> CH <sub>3</sub>	—	—	—	—	4.09
1_2J	(CH <sub>2</sub> ) <sub>4</sub> CH <sub>3</sub>	—	—	—	—	4.03
1_2K	(CH <sub>2</sub> ) <sub>5</sub> CH <sub>3</sub>	—	—	—	—	3.96

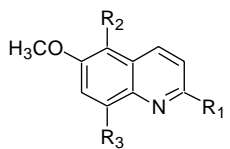
  

<div>  <p><b>Series 2</b></p> </div>						
Molecule	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>	Activity
1_3A	CH(CH <sub>3</sub> ) <sub>2</sub>	CH(CH <sub>3</sub> ) <sub>2</sub>	—	—	—	4.95
1_3C	<i>c</i> -C <sub>4</sub> H <sub>7</sub>	<i>c</i> -C <sub>4</sub> H <sub>7</sub>	—	—	—	6.30
1_3D (DCMQ)	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	—	—	—	6.65
1_3E	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	—	—	—	4.74
1_3H	(CH <sub>2</sub> ) <sub>2</sub> CH <sub>3</sub>	(CH <sub>2</sub> ) <sub>2</sub> CH <sub>3</sub>	—	—	—	4.31
1_3I	(CH <sub>2</sub> ) <sub>3</sub> CH <sub>3</sub>	(CH <sub>2</sub> ) <sub>3</sub> CH <sub>3</sub>	—	—	—	4.30
1_3J	(CH <sub>2</sub> ) <sub>4</sub> CH <sub>3</sub>	(CH <sub>2</sub> ) <sub>4</sub> CH <sub>3</sub>	—	—	—	4.27
1_3K	(CH <sub>2</sub> ) <sub>5</sub> CH <sub>3</sub>	(CH <sub>2</sub> ) <sub>5</sub> CH <sub>3</sub>	—	—	—	4.17

<div>  <p><b>Series 3</b></p> </div>						
Molecule	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>	Activity
2_2A	CH(CH <sub>3</sub> ) <sub>2</sub>	H	NO <sub>2</sub>	—	—	3.03
2_2B	C(CH <sub>3</sub> ) <sub>3</sub>	H	NO <sub>2</sub>	—	—	3.70
2_2C	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	H	NO <sub>2</sub>	—	—	3.99
2_2D	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	H	NO <sub>2</sub>	—	—	6.61
2_2E	CH(CH <sub>3</sub> ) <sub>2</sub>	CH(CH <sub>3</sub> ) <sub>2</sub>	NO <sub>2</sub>	—	—	4.07
2_2F	C(CH <sub>3</sub> ) <sub>3</sub>	C(CH <sub>3</sub> ) <sub>3</sub>	NO <sub>2</sub>	—	—	4.14
2_2G	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	NO <sub>2</sub>	—	—	6.08
2_2H	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	NO <sub>2</sub>	—	—	7.53
2_3B	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	H	NH <sub>2</sub>	—	—	3.55
2_4C	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	H	NHCOCH <sub>3</sub>	—	—	2.64
2_4D	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	H	NHCOCF <sub>3</sub>	—	—	3.71
2_4E	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	NHCOCF <sub>3</sub>	—	—	6.19

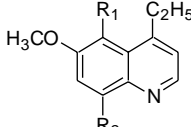
  

<div>  <p><b>Series 4</b></p> </div>						
Molecule	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>	Activity
2_6B	CH(CH <sub>3</sub> ) <sub>2</sub>	CH(CH <sub>3</sub> ) <sub>2</sub>	NO <sub>2</sub>	—	—	6.04
2_6D	C(CH <sub>3</sub> ) <sub>3</sub>	C(CH <sub>3</sub> ) <sub>3</sub>	NO <sub>2</sub>	—	—	3.95
2_6H	H	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	NO <sub>2</sub>	—	—	3.80
2_7B	H	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	NH <sub>2</sub>	—	—	4.09

(continued on next page)

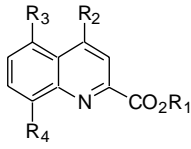
Table 1 (continued)

Molecule	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>	Activity
----------	----------------	----------------	----------------	----------------	----------------	----------



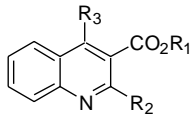
**Series 5**

2_13B	OCH(CH <sub>3</sub> ) <sub>2</sub>	NO <sub>2</sub>	—	—	—	6.18
2_13E	OC <sub>5</sub> H <sub>11</sub>	NO <sub>2</sub>	—	—	—	4.67
2_13F	OC <sub>6</sub> H <sub>13</sub>	NO <sub>2</sub>	—	—	—	5.23
2_13G	OC <sub>7</sub> H <sub>15</sub>	NO <sub>2</sub>	—	—	—	7.54
2_13H	OC <sub>8</sub> H <sub>17</sub>	NO <sub>2</sub>	—	—	—	3.07
2_14D	OC <sub>5</sub> H <sub>9</sub>	NH <sub>2</sub>	—	—	—	4.61
2_14E	OC <sub>5</sub> H <sub>11</sub>	NH <sub>2</sub>	—	—	—	6.66
2_14G	OC <sub>7</sub> H <sub>15</sub>	NH <sub>2</sub>	—	—	—	5.14
2_14H	OC <sub>8</sub> H <sub>17</sub>	NH <sub>2</sub>	—	—	—	3.77



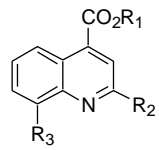
**Series 6**

3_1	H	H	H	H	—	3.49
3_2	CH <sub>3</sub>	H	H	H	—	3.47
3_3B	CH <sub>3</sub>	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	H	—	7.51
3_3C	CH <sub>3</sub>	H	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	—	6.72
3_3D	CH <sub>3</sub>	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	H	H	—	3.88
3_3E	CH <sub>3</sub>	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	H	—	6.75
3_3F	CH <sub>3</sub>	H	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	—	4.63
3_3G	CH <sub>3</sub>	1-Adamantyl	H	H	—	4.42
3_4A	H	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	H	H	—	3.61
3_4B	H	1-Adamantyl	H	H	—	3.94



**Series 7**

3_5	H	H	H	—	—	4.90
3_7A	CH <sub>3</sub>	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	H	—	—	4.20
3_7B	CH <sub>3</sub>	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	—	—	6.22
3_7C	CH <sub>3</sub>	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	H	—	—	4.37
3_7D	CH <sub>3</sub>	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	—	—	6.03
3_7E	CH <sub>3</sub>	1-Adamantyl	H	—	—	3.80

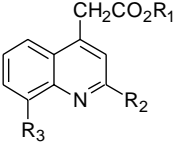


**Series 8**

3_9	H	H	H	—	—	3.25
3_11A	CH <sub>3</sub>	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	H	—	—	2.92
3_11B	CH <sub>3</sub>	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	—	—	4.50
3_11C	CH <sub>3</sub>	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	H	—	—	3.12
3_11D	CH <sub>3</sub>	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	—	—	4.65
3_11E	CH <sub>3</sub>	1-Adamantyl	H	—	—	4.69
3_11F	CH <sub>3</sub>	1-Adamantyl	1-Adamantyl	—	—	2.87
3_12A	H	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	H	—	—	3.21
3_12B	H	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	H	—	—	3.66
3_12C	H	1-Adamantyl	H	—	—	3.18

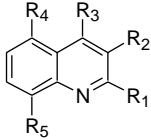
Table 1 (continued)

Molecule	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>	Activity
----------	----------------	----------------	----------------	----------------	----------------	----------



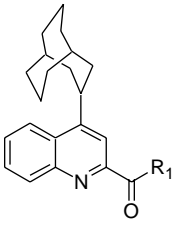
**Series 9**

3_13B	C <sub>2</sub> H <sub>5</sub>	H	H	—	—	2.85
3_14A	CH <sub>3</sub>	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	H	—	—	3.12
3_14B	CH <sub>3</sub>	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	—	—	3.22
3_14C	CH <sub>3</sub>	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	H	—	—	3.28
3_14D	CH <sub>3</sub>	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	—	—	3.39
3_14E	CH <sub>3</sub>	1-Adamantyl	H	—	—	3.78
3_14F	C <sub>2</sub> H <sub>5</sub>	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	H	—	—	4.89
3_14G	C <sub>2</sub> H <sub>5</sub>	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	—	—	6.94
3_14H	C <sub>2</sub> H <sub>5</sub>	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	H	—	—	4.63
3_14I	C <sub>2</sub> H <sub>5</sub>	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	—	—	6.78
3_14J	C <sub>2</sub> H <sub>5</sub>	1-Adamantyl	H	—	—	4.70
3_15A	C <sub>2</sub> H <sub>5</sub>	H	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	—	—	3.96
3_15B	C <sub>2</sub> H <sub>5</sub>	H	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	—	—	3.75

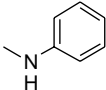


**Series 10**

4_2D	CONHNH <sub>2</sub>	H	1-Adamantyl	H	H	6.71
4_2E	CONHNH <sub>2</sub>	H	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	H	5.67
4_2F	CONHNH <sub>2</sub>	H	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	H	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	5.77
4_2G	CONHNH <sub>2</sub>	H	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	H	5.75
4_2H	CONHNH <sub>2</sub>	H	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	H	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	4.23
4_3B	1-Adamantyl	CONHNH <sub>2</sub>	H	H	H	4.38
4_3C	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	CONHNH <sub>2</sub>	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	H	H	4.82
4_3D	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	CONHNH <sub>2</sub>	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	H	H	4.65
4_4B	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	H	CONHNH <sub>2</sub>	H	H	2.92
4_4C	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	H	CONHNH <sub>2</sub>	H	H	2.94
4_4D	1-Adamantyl	H	CONHNH <sub>2</sub>	H	H	4.02
4_5A	H	H	CH <sub>2</sub> CONHNH <sub>2</sub>	H	H	3.50
4_5B	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	H	CH <sub>2</sub> CONHNH <sub>2</sub>	H	H	2.64
4_5C	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	H	CH <sub>2</sub> CONHNH <sub>2</sub>	H	H	3.79
4_5D	1-Adamantyl	H	CH <sub>2</sub> CONHNH <sub>2</sub>	H	H	4.38
4_5E	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	H	CH <sub>2</sub> CONHNH <sub>2</sub>	H	<i>c</i> -C <sub>5</sub> H <sub>9</sub>	5.04
4_5F	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	H	CH <sub>2</sub> CONHNH <sub>2</sub>	H	<i>c</i> -C <sub>6</sub> H <sub>11</sub>	4.87

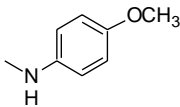
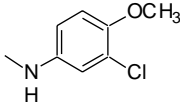
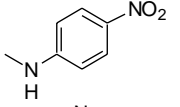
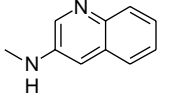
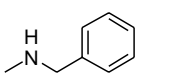


**Series 11**

4_7A	NH <sub>2</sub>	—	—	—	—	6.38
4_7B	NH(CH <sub>2</sub> ) <sub>2</sub> CH <sub>3</sub>	—	—	—	—	3.24
4_7C	NH(CH <sub>2</sub> ) <sub>3</sub> CH <sub>3</sub>	—	—	—	—	4.21
4_7D	NH(CH <sub>2</sub> ) <sub>4</sub> CH <sub>3</sub>	—	—	—	—	3.99
4_7E	NH(CH <sub>2</sub> ) <sub>5</sub> CH <sub>3</sub>	—	—	—	—	3.93
4_7F	NH(CH <sub>2</sub> ) <sub>6</sub> CH <sub>3</sub>	—	—	—	—	2.82
4_7G	N(CH <sub>2</sub> CH <sub>3</sub> ) <sub>2</sub>	—	—	—	—	3.07
4_7H		—	—	—	—	7.08

(continued on next page)

Table 1 (continued)

Molecule	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>	Activity
4_7I		—	—	—	—	4.55
4_7J		—	—	—	—	4.59
4_7K		—	—	—	—	4.71
4_7L		—	—	—	—	4.67
4_7M		—	—	—	—	7.10

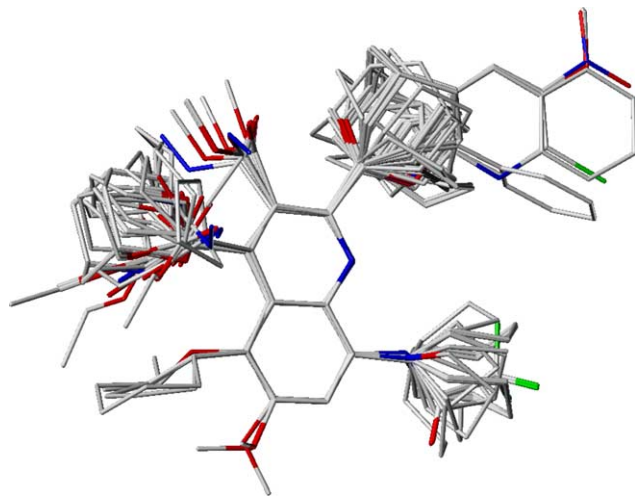


Figure 1. The training set molecules aligned over each other using DATABASE ALIGNMENT.

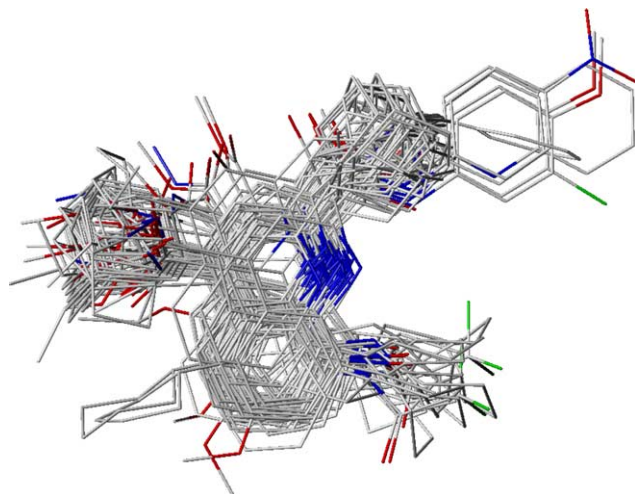


Figure 2. The training set molecules aligned over each other using FIELD FIT ALIGNMENT.

hydrogen bond donor and acceptor fields. The contribution from each one of these descriptors was truncated above 0.3 kcal/mol.

## 2.8. Partial least squares analysis

The partial least square (PLS) method was used to set up a correlation between the molecular fields and the inhibitory activity of the molecules. The optimal number of components was determined with *SAMPLS*<sup>16</sup> (Samples-distance Partial Least Square) and cross-validation was carried out by the *leave-ten-out* method. The model with the optimum number of components (highest  $q^2$ ) and with the lowest standard error of prediction (SDEP) was considered for further analysis. Equal weights were assigned to the steric and electrostatic fields by the *COMFA\_STD* scaling option. To speed up the analysis and reduce noise, columns with a value ( $\sigma$ ) below 2.0 kcal/mol were filtered off. Final analysis was performed to calculate the conventional  $r^2$  using the optimum number of components. To further assess the robustness and statistical confidence of the derived models, bootstrapping<sup>17</sup> analysis for 100 runs was performed. Bootstrapping involves the generation of many new datasets from the original dataset and is obtained by randomly choosing samples from the original dataset. The statistical calculation is performed on each of these bootstrapping samplings. The difference between the parameters calculated from the original dataset and the average of the parameters calculated from the many bootstrapping samplings is a measure of the bias of the original calculations. Models with a cross-validation ( $q^2$ ) value above 0.3 were sought, since at this value the probability of chance correlation is less than 5%.<sup>18</sup>

## 2.9. Predictive correlation coefficient

The predictive ability of each 3D-QSAR model was determined from a set of 24 test molecules not included

in the model generation. The predictive correlation coefficient ( $r^2_{\text{pred}}$ ), based on the test set molecules, is defined as

$$r^2_{\text{pred}} = (\text{SD} - \text{PRESS})/\text{SD},$$

where SD is the sum of squared deviations between the biological activity of the test set and the mean activity of the training set molecules and the PRESS is the sum of squared deviations between predicted and actual activity values for every molecule in the test set.

### 2.10. CoMFA contour maps

Contour maps were generated as a scalar product of coefficients and standard deviation ( $\text{StDev} * \text{Coeff}$ ) associated with each column. Favored and disfavored levels, fixed at 80% and 20%, respectively, were used to display the steric and the electrostatic fields. The contours for steric fields are shown in green (more bulk favored) and yellow (less bulk favored), while the electrostatic field contours are shown in red (electronegative substituents favored) and blue (electropositive substituents favored).

## 3. Results and discussion

A simulated annealing protocol was used to identify the global minimum energy conformation, which was used in the subsequent alignment and generation of the 3D-QSAR model. The biological activity of the dataset is within the range 2.64–7.54 log units (approximate five log units). The CoMFA and CoMSIA models were generated for both the alignments, namely *database* and field fit alignments. The hydrophobic field calculated using *HINT* was used in conjunction with CoMFA fields to derive additional models. The statistics of the models generated are shown in Table 2.

The CoMFA model using field fit alignment has a  $q^2$  of just 0.22 with four components, hence it was not consid-

ered further. The remaining five models have statistically significant results with regard to the training set molecules, but only Model 1 has a significant predictive correlation coefficient ( $r^2_{\text{pred}}$ ) for the external test set. Therefore, Model 1, that is, the CoMFA model derived using *database alignment* was taken up for further consideration. This model exhibits a cross-validated correlation coefficient ( $q^2$ ) of 0.49, conventional correlation coefficient ( $r^2$ ) of 0.94, and predictive correlation coefficient ( $r^2_{\text{pred}}$ ) of 0.42. A plot of the predicted versus experimental activity for the training set molecules using Model 1 is shown in Figure 3.

Table 3 provides a prediction of the activity of the test set molecules using this model. The incorporation of

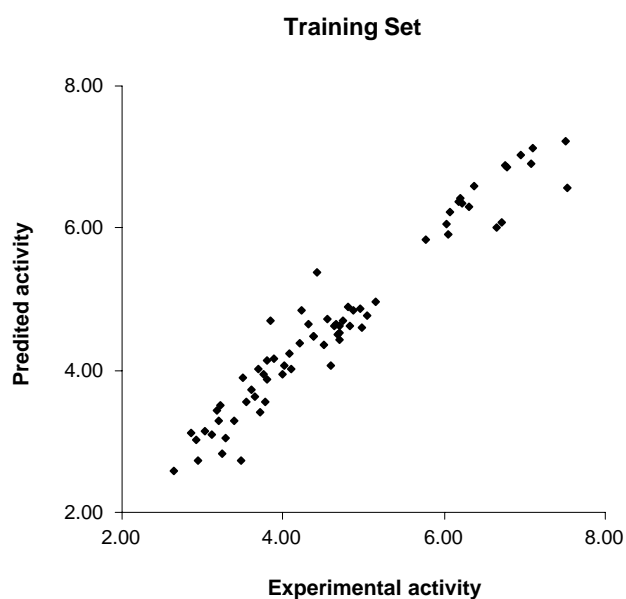


Figure 3. Predicted versus experimental activity for molecules in the training set using Model 1.

Table 2. A summary of statistics of the 3D-QSAR models

Parameter	Database_Alignment			Field_Fit Alignment	
	CoMFA Model 1	CoMFA + HINT Model 2	CoMSIA Model 3	CoMFA + HINT Model 4	CoMSIA Model 5
$a_r^2$	0.49	0.49	0.46	0.42	0.46
$b_r^2$	0.50	0.49	0.49	0.40	0.42
$N$	8	5	9	4	6
$r^2$	0.94	0.87	0.90	0.83	0.86
SEE	0.23	0.47	0.314	0.54	0.43
$F$ value	123.64	88.75	63.87	80.5	62.86
$r^2_{\text{pred}}$	0.42	−0.51	−0.37	−0.54	−0.41
$r^2_{\text{bs}}$	0.97	0.92	0.95	0.89	0.89
SD	0.012	0.016	0.015	0.034	0.021
<i>Contributions (%)</i>					
Steric	52	36	16	28	14
Electrostatic	48	37	22	40	26
Hydrophobic	—	27	26	32	21
Donor	—	—	21	—	19
Acceptor	—	—	15	—	20

$a_r^2$ , cross-validated correlation coefficient using *SAMPLS*;  $b_r^2$ , cross-validated correlation coefficient using leave-ten-out;  $N$ , optimum number of components;  $r^2$ , conventional (non-cross-validated) correlation coefficient; SEE, standard error of estimate;  $r^2_{\text{pred}}$ , predictive (test molecules) correlation coefficient;  $r^2_{\text{bs}}$ , correlation coefficient after 100 runs of bootstrapping analysis; SD, standard deviation from 100 bootstrapping runs.



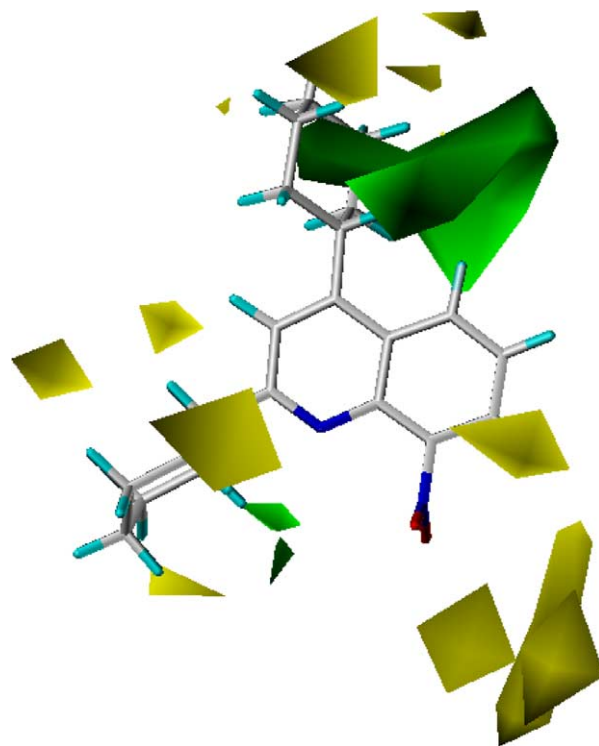
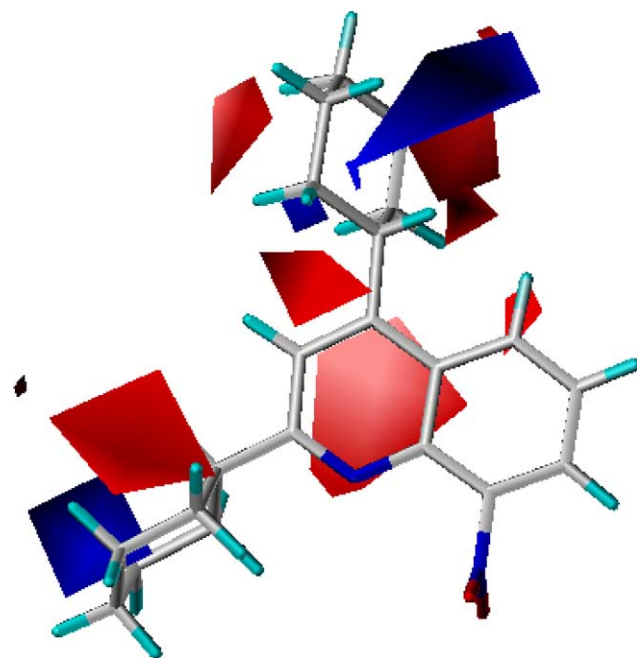
**Table 3.** Experimental, predicted activity, and residuals for the test set molecules using Model 1

Molecule	Experimental activity	Predicted activity	Residual
<b>1_2A</b>	4.59	4.01	0.58
<b>1_2H</b>	4.10	4.49	−0.39
<b>2_14D</b>	4.61	5.64	−1.03
<b>2_2E</b>	4.07	5.23	−1.16
<b>2_2F</b>	3.95	5.10	−1.15
<b>2_6D</b>	3.95	5.16	−1.21
<b>2_6H</b>	3.80	4.69	−0.89
<b>3_1</b>	3.49	2.40	1.09
<b>3_11F</b>	2.87	4.19	−1.32
<b>3_14A</b>	3.12	3.70	−0.58
<b>3_14F</b>	4.89	3.92	0.97
<b>3_15A</b>	3.96	3.61	0.35
<b>3_3C</b>	6.72	7.63	−0.91
<b>3_3F</b>	4.63	4.35	0.28
<b>3_4B</b>	3.94	3.51	0.43
<b>3_5</b>	4.90	4.07	0.83
<b>4_2E</b>	5.67	5.01	0.66
<b>4_2G</b>	5.75	5.16	0.59
<b>4_3D</b>	4.65	5.20	−0.55
<b>4_4B</b>	2.92	3.41	−0.49
<b>4_5B</b>	2.64	3.19	−0.55
<b>4_5D</b>	4.38	4.35	0.03

the hydrophathic field to the CoMFA model generated using *database alignment* does not improve the statistical quality of the model. However, the inclusion of the hydrophathic field to the CoMFA model generated using field fit alignment improves the  $q^2$  significantly from 0.22 to 0.44, but decreases the  $r^2_{pred}$ , for the external test set. The CoMSIA models are statistically comparable to the CoMFA models and no CoMSIA model gives a good predictive correlation coefficient. The statistical data in Table 2 indicate that the steric and electrostatic fields are sufficient to generate a 3D-QSAR model with good correlation and predictive powers.

The CoMFA model, with its hundreds or thousands of terms, is generally represented as a 3D ‘coefficient contour.’ Colored contours in the map represent those areas in 3D space where changes in the steric and electrostatic field values of a compound correlate strongly with concomitant change in its biological activity. The CoMFA steric and electrostatic contour plots of Model 1 are shown in Figures 4 and 5, respectively.

Analysis of the steric contours (Fig. 4) reveals one green colored contour and few small yellow colored contours. The favorable green contour is seen near the C4 and C5 substituents on the quinoline ring. There are three small contours and one medium size yellow contour in proximity of the C7 and C8 positions on the quinoline ring. There is also one small and one medium size contour near the N1 and C2 sites on the quinoline ring. Molecules with bulky substituents at the C4 and/or C5 positions exhibit good activity. These include **2\_2H** with a C4 cyclohexyl, **3\_3B** with C4 and C5 dicyclopentyl, **3\_3E** with C4 and C5 dicyclohexyl, and **4\_7M**, **4\_2D**, and **4\_7H** each with an adamantyl group at the C2 position. The sterically unfavorable yellow contour away from the C8 position of the quinoline ring indicates that too bulky substituents will not be favored at this posi-

**Figure 4.** CoMFA contour maps (Model 1) for steric fields drawn around molecule **2\_2H**. The green colored contour favors steric bulk while sites where steric bulk is disfavored are shown in yellow.**Figure 5.** CoMFA contour maps (Model 1) for electrostatic fields drawn around molecule **2\_2H**. The red contour shows regions where electronegative substituents are favored, while the blue contour is associated with positions where electropositive substituents improve activity.

tion. Molecules with less bulky substituents at C8 like cyclobutyl (**1\_3C**), cyclopentyl (**1\_3D**, **3\_3B**, **4\_2F**, **4\_5E**), exhibit good activity, but more bulky substitu-



ents at the C8 position like cyclohexyl (**1\_3E**, **3\_3F**, **4\_2H**, **4\_5F**), *t*-butyl (**2\_2B**, **2\_2F**, **2\_6D**), and adamantyl in **3\_11F** are responsible for the low activity of these molecules. For the latter group of molecules their low activity stems from the fact that the bulky group is partially buried within the yellow contour.

Analysis of the electrostatic contours (Fig. 5) shows one big and one small red colored contour (electronegative substituents favored) near the C2 and C3 positions of the quinoline ring. There is a big red contour away from the C4 and C5 positions, and a small red contour near the C4 point of the quinoline ring. There is a blue contour (electropositive substituents favored) away from the C2 position and one smaller contour near the C4 and C5 positions of the quinoline ring. The molecules like **3\_3B**, **3\_3E**, **4\_2d**, **4\_7H**, and **4\_7M**, which possess a substituted ester or amide function at the C2 position of the quinoline ring, exhibit good activity. These molecules have the amide and ester groups oriented toward the red contour, and the substituted alkyl groups of the amide and ester are directed toward the blue contour, making favorable interactions. Molecules like **3\_7A**, **3\_7B**, **3\_7C**, and **3\_7D** with a polar group at the C3 position and a non-polar group at the C2 position like methyl carboxylate at the C3 position and cyclopropyl/cyclohexyl at the C2 position also exhibit good activity as the polar group at C3 is near the favorable red contour, and the non-polar group at C2 points away from this red contour. In the same series, the molecule **3\_7E** with a methyl carboxylate at the C3 position and the bulky non-polar adamantyl group at the C2 position exhibits low activity as the large adamantyl group is partially buried within the unfavorable red contour. This indicates that both polar groups as well as less bulky non-polar groups can be accommodated at the C2 position of the quinoline ring. Polar groups at the C4 position of the quinoline ring are not favored. The CoMFA analysis thus provides deep structural insights into possible modifications of the quinoline ring that can improve the anti-tuberculosis activity.

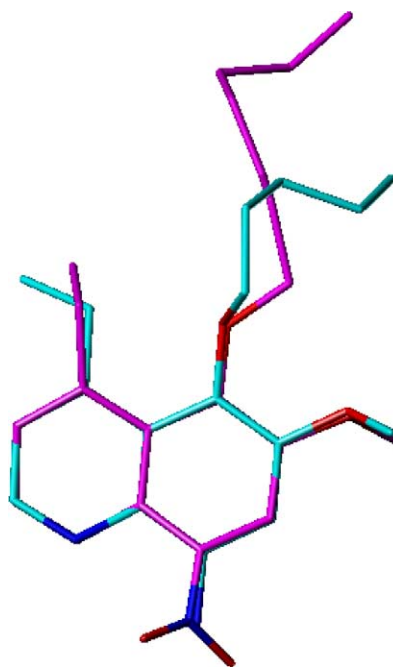
Prediction of the activities of the test set molecules based on Model 1 is given in Table 3. Out of 24 molecules in the test set, the activities of 22 are predicted well by the model. The activities of two molecules **2\_2D** and **4\_7G** in the test set are poorly predicted ( $\pm 2.0$  U) by all five models. Inclusion of these two molecules in Model 1 lowers the statistical quality of the model drastically. Hence, these two molecules were considered as outliers and not included in the calculation of the predictive correlation coefficient ( $r^2_{\text{pred}}$ ).

For the 18 molecules with the flexible alkyl side chains, an attempt was made to predict their activity based on Model 1 using the whole set of conformations generated by Systematic Search (*MULTI SEARCH*, vide supra). The activity predicted for each of these eighteen molecules is presented in Table 4.

Most of these conformationally flexible molecules exhibit low activity, with the exception of two molecules **2\_13G** and **2\_14E**, which show good activity. From

**Table 4.** Predictions of the set of conformers for 18 molecules with flexible side chains using Model 1

Molecule	Activity	Total no. of conformers within 5 kcal/mol	No. of conformers predicted $\pm 0.5$ of activity
<b>1_2I</b>	4.09	4	4
<b>1_2J</b>	4.03	9	9
<b>1_2K</b>	3.96	16	12
<b>1_3I</b>	4.30	20	8
<b>1_3J</b>	4.27	52	28
<b>1_3K</b>	4.17	111	35
<b>2_13E</b>	4.67	31	13
<b>2_13F</b>	5.23	52	26
<b>2_13G</b>	7.54	87	1 (activity = 7.20)
<b>2_13H</b>	3.07	106	2
<b>2_14E</b>	6.66	22	1 (activity = 6.40)
<b>2_14G</b>	5.14	62	19
<b>2_14H</b>	3.77	85	5
<b>4_7B</b>	3.24	42	10
<b>4_7C</b>	4.21	54	12
<b>4_7D</b>	3.99	63	9
<b>4_7E</b>	3.93	71	6
<b>4_7F</b>	2.82	92	11



**Figure 6.** Conformations of molecules **2\_13G** (magenta) and **2\_14E** (cyan) predicted well using Model 1.

the set of conformations generated for these two molecules, there was only one conformation of each molecule for which the activity could be predicted accurately, and this particular conformation for the two molecules **2\_13G** and **2\_14E** is depicted in Figure 6.

#### 4. Conclusions

A 3D-QSAR analysis of a novel class of anti-tuberculosis agents was carried out using CoMFA alone, CoMFA in conjunction with a hydrophobic field evaluated using

HINT and CoMSIA, to map the structural features contributing to the inhibitory activity of these molecules. Inclusion of the *HINT* hydropathic field to the CoMFA models does not improve the quality of the models. The CoMSIA models are comparable to the CoMFA model but lack good predictive power. The *database alignment* of molecules produced models with better statistics than those with field fit alignment. Out of the various models evaluated, the CoMFA model based on database alignment produced a statistically sound model with a good correlation and predictive power. Analysis of the CoMFA contours provides details on the fine relationship linking structure and activity, and provides clues for structural modifications that can improve the activity. This study also discloses several new derivatives of quinolines (data not included) with activity higher than that of the molecules in this study. Attempts are currently underway in our laboratory to synthesize and evaluate the anti-tuberculosis activities of the newly proposed structures.

### Acknowledgments

The computational facilities at the BCP were provided by the Department of Science and Technology (DST) through a grant (SR/FST/LS1—083/2003) under the FIST program. Amit Nayyar and Alpeshkumar Malde thank the Council of Scientific and Industrial Research (CSIR), India, for financial support.

### References and notes

- Global tuberculosis control—surveillance, planning, financing, WHO Report **2004**, please see: [http://www.who.int/tb/publications/global\\_report/en/](http://www.who.int/tb/publications/global_report/en/).
- (a) Chan, E. D.; Iseman, M. D. *BMJ* **2002**, 325, 1282; (b) Bastian, I.; Colebunders, R. *Drugs* **1999**, 58, 633; (c) Basso, L. A.; Blanchard, J. S. *Adv. Exp. Med. Biol.* **1998**, 456, 115.
- Global Alliance for TB Drug Development. In *Scientific Blueprint for Tuberculosis Drug Development*, **2001**, please see: <http://www.tballiance.org/>.
- Cole, S. T.; Brosch, R.; Parkhill, J., et al. *Nature* **1998**, 393, 537.
- Jain, R.; Vaitilingam, B.; Nayyar, A.; Palde, P. *Bioorg. Med. Chem. Lett.* **2003**, 13, 1051.
- Vangapandu, S.; Jain, M.; Jain, R.; Kaur, S.; Singh, P. *Bioorg. Med. Chem.* **2004**, 12, 2501.
- Vaitilingam, B.; Nayyar, A.; Palde, P.; Monga, V.; Jain, R.; Kaur, S.; Singh, P. *Bioorg. Med. Chem.* **2004**, 12, 4179.
- Monga, V.; Nayyar, A.; Vaitilingam, B.; Palde, P.; Jhamb, S.; Kaur, S.; Singh, P.; Jain, R. *Bioorg. Med. Chem.* **2004**, 12, 6465.
- Cramer, R. D.; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, 110, 5959.
- Klebe, G.; Abraham, U.; Mietzner, T. *J. Med. Chem.* **1994**, 37, 4130.
- (a) Kellogg, G. E.; Semus, S. F.; Abraham, D. J. *J. Comput. Aided Mol. Des.* **1991**, 5, 545; (b) Kellogg, G. E.; Joshi, G. S.; Abraham, D. J. *J. Med. Chem.* **1992**, 36, 444.
- Martin, E. J.; Blaney, J. M.; Siani, M. A.; Spellmeyer, D. C.; Wong, A. K.; Moos, W. H. *J. Med. Chem.* **1995**, 39, 1431.
- CoMFA and QSAR Manual, Sybyl 7.0, Associates Inc., 1699 S Hanley Rd., St. Louis, MO 63144, USA.
- Sybyl 7.0, Tripos Associates Inc., 1699 S Hanley Rd., St. Louis, MO 63144, USA.
- Halgren, T. *J. Am. Chem. Soc.* **1990**, 112, 4710.
- Bush, B. L.; Nachbar, R. B. *J. Comput. Aided Mol. Des.* **1993**, 7, 587.
- Cramer, R. D., III; Bunce, J. D.; Patterson, D. E.; Frank, I. E. *Quant. Struct.-Act. Relat.* **1988**, 7, 18.
- Clark, M.; Cramer, R. D.; Jones, D. M.; Patterson, D. E.; Simeroth, P. E. *Tetrahedron Comput. Methodol.* **1990**, 3, 47.