



# Modelling topoisomerase I inhibition by minor groove binders

David Winkler\*

CSIRO Materials Science and Engineering, Bag 10, Clayton South MDC 3169, Australia  
Monash Institute of Pharmaceutical Sciences, 381 Royal Parade, Parkville 3052, Australia

## ARTICLE INFO

### Article history:

Received 13 November 2010

Revised 3 January 2011

Accepted 4 January 2011

Available online 7 January 2011

### Keywords:

Topoisomerase I inhibition

QSAR

Minor groove binder

Molecular field analysis

## ABSTRACT

Topoisomerase inhibition is an extremely useful target for anticancer and antimicrobial drugs, and an undesirable side effect of some drugs targeting other proteins. Published modelling studies are sparse, and have used small data sets with relatively low molecular diversity. Given the important role of minor groove binding in the mechanism of topoisomerase I inhibition, we have conducted the first 3D QSAR study of topoisomerase I inhibition of a large, diverse set of minor groove binders using the minor groove binding conformation as the alignment template. The highly significant QSAR models resulting from this alignment identify the roles played by molecular features, most importantly the hydrogen bond donor properties.

Crown Copyright © 2011 Published by Elsevier Ltd. All rights reserved.

## 1. Introduction

DNA topoisomerases play an essential function in DNA replication, transcription and chromatin condensation.<sup>1</sup> Topoisomerases have a very complex mechanism of action, involving DNA cleavage to form a covalent enzyme–DNA intermediate, DNA relaxation and finally, religation of the phosphate backbone to restore the continuity of the DNA, thus changing its topology. Given their intimate relationship with DNA replication and repair, they have been targets of important anticancer and antibacterial drugs.<sup>2</sup> As with most drugs, unwanted topoisomerase activity in compounds targeting other enzymes or receptors can contribute to substantial off-target toxicity and is usually ‘designed out’ of drug leads if possible. If the drug target is the minor groove of DNA, as is the case with agents that protect DNA from radiological damage, or antitumour or antibacterial agents, this is often quite difficult.

Structure/function studies have shown that topoisomerases are flexible enzymes that relax DNA through coordinated, controlled movements of distinct enzyme domains.<sup>1</sup> Surprisingly, given this complexity, it is possible to model topoisomerase I inhibition using relatively simple QSAR methods, at least for small sets of closely related inhibitors. Mekapati and Hansch<sup>3</sup> reviewed topoisomerase I inhibition by bis- and terbenzimidazoles, and generated individual QSAR models for each class of inhibitors. The data sets for each of these models were small (usually around 10 compounds) and the molecular diversity relatively low, but the statistical quality of the models was high.

As no 3D QSAR models for topoisomerase I inhibition of minor groove binders have been reported in the literature, and several studies have reported that DNA binding is important for topoisomerase inhibition,<sup>4,5</sup> we undertook to develop the first quantitative, 3D QSAR model to test this reported relationship. Several literature data sets of topoisomerase I inhibition, measured by essentially identical protocols, were combined into a much larger, more chemically diverse data set of minor groove binding compounds. Our aims were: to develop a more general, 3D QSAR model of topoisomerase inhibition by minor groove binders; to assess whether molecular alignments based on the assumption of minor groove binding were consistent with topoisomerase I inhibition; and to understand the molecular requirements for topoisomerase inhibition in a larger, more chemically diverse set of molecules. This will provide additional understanding of the relationship between minor groove structure and topoisomerase I effects, and allow topoisomerase I inhibition to be estimated for a larger range of minor groove chemotypes than the previous QSAR models.

## 2. Materials and methods

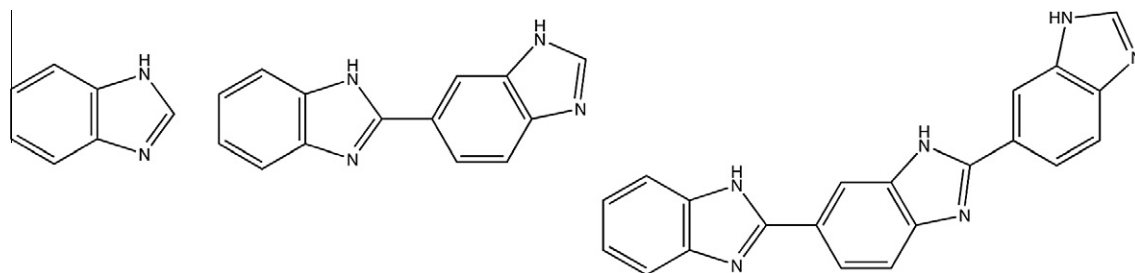
2D and 3D QSAR models were generated for a larger, more diverse combined set of topoisomerase I inhibitors reported in the literature.<sup>6–11</sup> The chemical classes in the data set were benzimidazoles, bisbenzimidazoles, and terbenzimidazoles (Fig. 1).

### 2.1. Biological activity

The minor groove binders in the combined literature data set (107 compounds) had topoisomerase I inhibition measured by essentially identical protocols. The biological activities were for

\* Corresponding author. Tel.: +61 3 9545 2477.

E-mail address: [dave.winkler@csiro.au](mailto:dave.winkler@csiro.au)



**Figure 1.** Main minor groove structure classes: benzimidazoles; bisbenzimidazoles; terbenzimidazoles.

calf thymus topoisomerase I inhibition, expressed as pREC, the  $-\log$  of the relative effective concentration (REC) of the drug that gave the same degree of DNA cleavage as Hoechst 33342 ( $\approx 1.0$ ). Smaller REC values (larger pREC values) mean more potent topoisomerase I inhibition. The biological activities for the one hundred and seven compounds in the modelling study are listed in Tables 1–9.

All studies used this relative measure (REC) to report the topoisomerase activities except in two papers by Kim et al.<sup>8,9</sup> who used another minor groove binder as a reference, and its activity relative to Hoechst 33342 was also quoted. For these data, the REC values reported in the paper were corrected back to the Hoechst 33342 standard values. Several of the minor groove binders in papers by Kim et al. had low and indeterminate ( $>$ ) activity. For indeterminate values we used REC values double the highest concentration ratio tested, and a log REC value of  $-6$  for inactive compounds in one of Kim et al.'s data sets.<sup>8</sup>

## 2.2. Multiple linear regression (MLR) QSAR modelling

The MLR models used relatively simple, interpretable molecular descriptors related to the size of the molecules (molecular volume, surface area, and number of rings), lipophilicity ( $\log P$  octanol/water calculated using the Pomona College  $C \log P$  algorithm), molecular flexibility (number of rotatable bonds), and number of hydrogen bond donor ( $N_{\text{donor}}$ ) and acceptor ( $N_{\text{acceptor}}$ ) groups in the molecule. These were calculated using Sybyl8.1 and the properties calculated in the molecular spreadsheet. Tables 1–9 also

**Table 1**  
Bis- and terbenzimidazole minor groove data set from Table 1 of Alper et al.<sup>6</sup>

The figure shows two chemical structures of bisbenzimidazole derivatives. The first structure has a substituent R on the imidazole ring and a substituent X on the benzene ring. The second structure has a substituent R on the imidazole ring and a cyano group (CN) on the benzene ring.

R	X	REC	pREC	$C \log P$	$N_{\text{accept}}$	$N_{\text{donor}}$
H	H	1.1	−0.04	4.61	6	6
H	CN	1	−0.00	4.22	7	6
H	<i>n</i> -Pr	100	−2.00	6.16	6	6
H	Ph	2	−0.30	6.50	6	6
H	2-pyridyl	3.3	−0.52	5.21	7	7
H	3-pyridyl	2	−0.30	5.00	7	7
H	4-pyridyl	2	−0.30	5.00	7	7
4-MeOPh	CN	1000	−3.00	6.34	8	6
H	—	1000	−3.00	2.70	5	4
4-MeOPh	—	3.3	−0.52	4.82	6	4

**Table 2**  
Terbenzimidazole minor groove data set from Table 2 of Alper et al.<sup>6</sup>

The figure shows a chemical structure of a terbenzimidazole derivative with a substituent R on the benzene ring.

R	REC	pREC	$C \log P$	$N_{\text{accept}}$	$N_{\text{donor}}$
1-Naphthyl	10	−1.00	7.67	6	6
2-Naphthyl	10	−1.00	7.67	6	6
Ph	1	0.00	6.50	6	6
Pr	0.5	0.30	6.16	6	6
Br	1	0.00	5.55	6	6
Piperidine	0.5	0.30	5.86	6	7
Cl	1	0.00	5.40	6	6
F	0.05	1.30	4.83	6	6
H	1	0.00	4.61	6	6
OMe	0.05	1.30	4.91	7	6
NO <sub>2</sub>	0.5	0.30	4.52	8	6
CN	0.1	1.00	4.22	7	6
OH	0.5	0.30	4.66	7	7
NH <sub>2</sub>	1	0.00	4.06	6	7

**Table 3**  
Terbenzimidazole minor groove data set from Table 3 of Alper et al.<sup>6</sup>

The figure shows a chemical structure of a terbenzimidazole derivative with substituents R<sub>1</sub>, R<sub>2</sub>, and R<sub>3</sub> on the benzene rings.

R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	REC	pREC	$C \log P$	$N_{\text{accept}}$	$N_{\text{donor}}$
Br	H	H	1	0.00	5.55	6	6
Br	H	OH	1	0.00	5.83	7	7
Br	H	Pr	1	0.00	6.87	6	6
4-ClPh	H	CF <sub>3</sub>	1	0.00	8.27	6	6
Br	Br	H	1	0.00	6.24	6	6
Br	Br	Cl	1	0.00	7.05	6	6
Br	Br	CF <sub>3</sub>	1	0.00	7.30	6	6
Ph	Ph	H	1	0.00	7.78	6	6
Br	OMe	H	2	−0.30	5.59	7	6
Ph	OMe	H	0.5	0.30	6.24	7	6

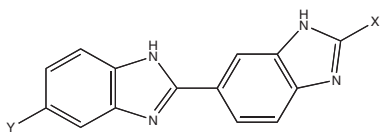
summarise these molecular properties of minor groove binders in the data set.

### 2.3. 3D QSAR modelling

Several studies have reported that DNA binding is important for topoisomerase inhibition.<sup>4,5</sup> Bailly synthesised two close analogues of Hoechst 33258. One analogue was a pure DNA minor groove

**Table 4**

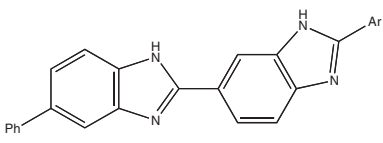
Bisbenzimidazole minor groove data set from Table 4 of Alper et al.<sup>6</sup>

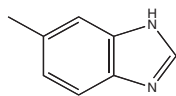
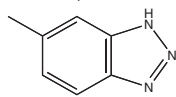
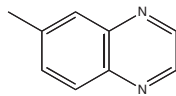
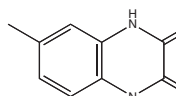
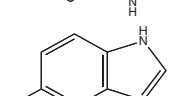
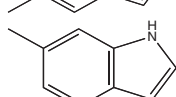


X	Y	REC	pREC	C log P	N <sub>accept</sub>	N <sub>donor</sub>
Ph	CN	>100	−2.30	4.80	5	4
2-Tolyl	CN	>500	−3.00	5.00	5	4
3-Tolyl	CN	100	−2.00	5.30	5	4
4-Tolyl	CN	1	0.00	5.30	5	4
1-Naphthyl	CN	10	−1.00	5.97	5	4
2-Naphthyl	CN	10	−1.00	5.97	5	4
Ph	CONH <sub>2</sub>	10	−1.00	4.19	5	5
2-Tolyl	CONH <sub>2</sub>	100	−2.00	4.39	5	5
3-Tolyl	CONH <sub>2</sub>	100	−2.00	4.69	5	5
4-Tolyl	CONH <sub>2</sub>	10	−1.00	4.69	5	5
1-Naphthyl	CONH <sub>2</sub>	20	−1.30	5.36	5	5
2-Naphthyl	CONH <sub>2</sub>	100	−2.00	5.36	5	5
Ph	4-Me piperazinyl	1	0.00	5.61	5	6
2-Tolyl	4-Me piperazinyl	10	−1.00	5.81	5	6
3-Tolyl	4-Me piperazinyl	1	0.00	6.11	5	6
4-Tolyl	4-Me piperazinyl	1	0.00	6.11	5	6
1-Naphthyl	4-Me piperazinyl	20	−1.30	6.79	5	6
2-Naphthyl	4-Me piperazinyl	10	−1.00	6.79	5	6
4-EtOPh	4-Me piperazinyl	1	0.00	6.04	6	6

**Table 5**

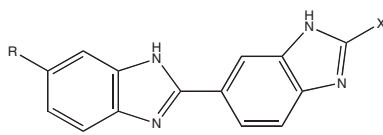
Terbenzimidazole minor groove data set from Jin et al.<sup>7</sup>

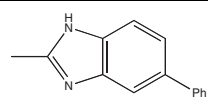
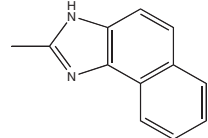
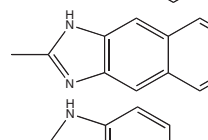
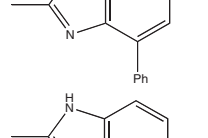
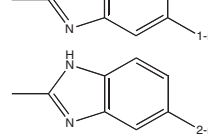
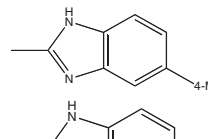
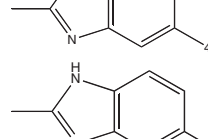
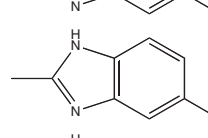
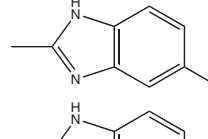
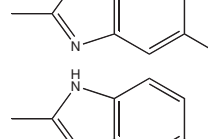
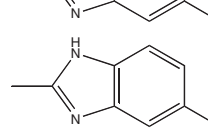
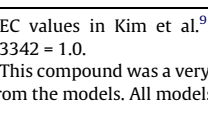



Ar	REC	pREC	C log P	N <sub>accept</sub>	N <sub>donor</sub>
	1	0.00	6.50	6	6
	1	0.00	6.34	7	6
	>100	−2.30	6.21	6	4
	1	0.00	4.84	6	6
	50	−1.70	7.06	4	5
	5	−0.70	7.06	4	5

**Table 6**

Terbenzimidazole minor groove data set from Kim et al.<sup>9</sup>

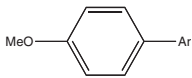
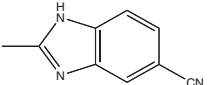
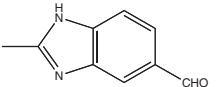
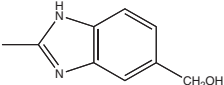
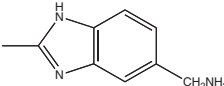
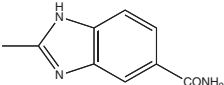
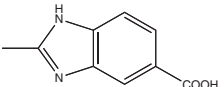
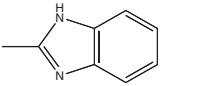
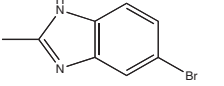
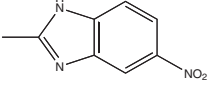
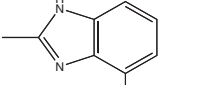
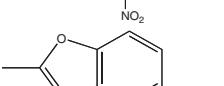
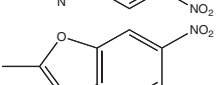
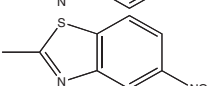
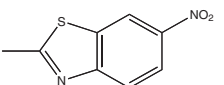
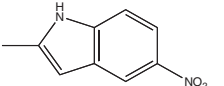
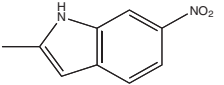


R	X	REC	pREC	C log P	N <sub>accept</sub>	N <sub>donor</sub>
	H	2	−0.30	6.50	6	6
	H	2	−0.30	5.78	6	6
	H	>2000 <sup>a</sup>	−3.60	5.78	6	6
	H	10	−1.00	6.50	6	6
	H	4	−0.6	7.67	6	6
	H	2	−0.3	7.67	6	6
	H	1	0.00	6.41	7	6
	H	4	−0.60	7.21	6	6
	Me	4	−0.60	6.76	6	6
	Et	2	−0.3	7.29	6	6
	<i>n</i> -Pr	1	0.00	7.82	6	6
	<i>i</i> Pr	1	0.00	7.69	6	6
	Ph	40	−1.60	8.59	6	6
	4-MeOPh	40	−1.60	8.61	7	6

REC values in Kim et al.<sup>9</sup> have been multiplied by 2 to normalise to Hoechst 33342 = 1.0.

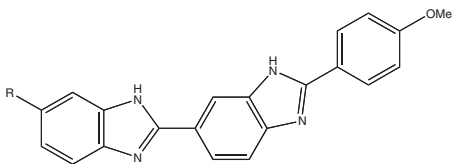


<sup>a</sup> This compound was a very significant outlier in all QSAR analyses, and was omitted from the models. All models predicted the REC to be 2–6, not >2000 as reported.

**Table 7**Benzimidazole minor groove data set from Kim et al.<sup>8</sup>

					
Ar	REC	pREC	C log P	N <sub>accept</sub>	N <sub>donor</sub>
	>300,000	−6.00	3.29	4	2
	300	−2.48	3.49	4	2
	30,000	−4.48	2.65	4	3
	>300,000	−6.00	2.64	3	3
	3000	−3.48	2.68	4	3
	>300,000	−6.00	3.74	5	2
	>300,000	−6.00	3.68	3	2
	>300,000	−6.00	4.62	3	2
	300	−2.48	3.59	5	2
	>300,000	−6.00	3.59	5	2
	Inactive	−6.00	3.40	5	0
	Inactive	−6.00	3.40	5	0
	Inactive	−6.00	4.04	4	0
	Inactive	−6.00	4.04	4	0
	Inactive	−6.00	4.26	3	1
	Inactive	−6.00	4.26	3	1

binder whereas the other behaved as a typical DNA intercalating agent. The purely intercalating analogue had no effect on the relaxation of supercoiled plasmid by topoisomerase I whereas the pure

**Table 8**Bisbenzimidazole minor groove data set from Sun et al.<sup>11</sup>

					
R	REC	pREC	C log P	N <sub>accept</sub>	N <sub>donor</sub>
(CH <sub>2</sub> ) <sub>3</sub> NMe <sub>2</sub>	2	−0.30	5.56	5	5
(CH <sub>2</sub> ) <sub>2</sub> NMe <sub>2</sub>	5	−0.70	5.18	5	5
CH <sub>2</sub> NMe <sub>2</sub>	2.5	−0.40	5.04	5	5
NMe <sub>2</sub>	1000	−3.00	5.78	5	5
NO <sub>2</sub>	1000	−3.00	5.12	7	4
NH <sub>2</sub>	100	−2.00	4.65	5	5
	1	0.00	5.52	6	6
	2	−0.30	5.40	6	5

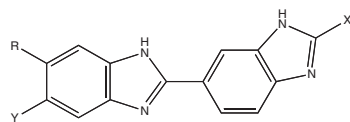
minor groove binder strongly inhibited the enzyme. The results suggest that the inhibition of topoisomerase I is affected by the ability to bind to the minor groove of DNA.<sup>4</sup>

Consequently, we used the minor groove binder conformation from the X-ray structure of a closely-related benzimidazole, methylproamine, bound to the minor groove of the DNA model dodecamer d(CGCGAATTCGCG)(2) (RCSB refcodes 1qv4 and 1qv8)<sup>12</sup> as an alignment template. This is also very similar to the conformation of the minor groove binder Hoechst 33258 bound to DNA.<sup>13</sup> The binding conformation requires that the bis- or terbenzimidazoles adopt a slightly twisted planar structure, with the benzimidazole NH moieties on the same side of the molecule. This is required for hydrogen bond interactions of these moieties in the minor groove, according to the X-ray studies.

All structures were minimised using the Tripos force field, Gasteiger-Huckel charges, and a dielectric constant of 4.0, approximating the environment of a receptor. The force field was able to reproduce the binding conformation observed in the X-ray structure quite accurately. The 3D QSAR models were derived using two molecular field methods, CoMFA,<sup>14</sup> and COMSIA,<sup>15</sup> both of which require consistent alignment of structures to a template. The terbenzimidazoles could be aligned unambiguously using the benzimidazole nitrogen atoms to carry out a least squares fit. The RMS errors for the fit were very low in most cases being below 0.1 Å.

Once aligned using the benzimidazole nitrogen atoms, the other heavy atoms occupying analogous positions in molecules also superimposed very well. The bis-benzimidazoles were superimposed on the X-ray structure of methylproamine, and also the terbenzimidazoles using the benzimidazole nitrogen atoms, with essentially no ambiguity. Several smaller minor groove binders could not be aligned in only one way, and several possible alignments were trialed. In all cases, one of these trial alignments was well predicted by the model, and the others were relatively poorly predicted. The alignment that was most consistent with the model was used in the analysis. In several molecules containing relatively flexible substituents, two or more low energy conformations of the substituents were possible. In this case, alternative conformations were trialed and only one was consistent with the model.

One compound (the naphthyl substituted compound **10** from Kim et al.<sup>9</sup>) was a very large outlier in all models. It was reported as exhibiting low or negligible topoisomerase I activity, while all models predicted the activity to be within an order of magnitude of Hoechst 33342, like its isomeric analogues. It was omitted from

**Table 9**Terbenzimidazole minor groove data set from Sun et al.<sup>10</sup>

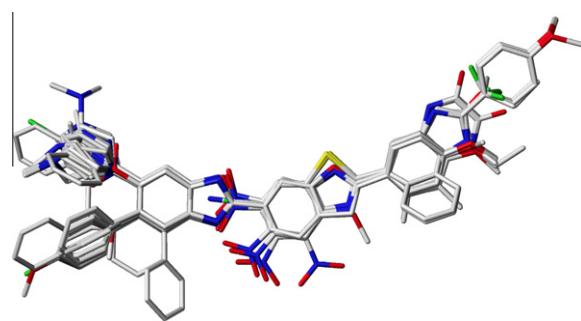
R	X	Y	REC	pREC	C log P	N <sub>accept</sub>	N <sub>donor</sub>
	H	H	1.1	−0.04	4.61	6	6
	H	H	1	0.00	4.22	7	6
	H	H	100	−2.00	6.16	6	6
	H	H	2	−0.30	6.50	6	6
	H	H	3.3	−0.52	5.21	7	7
	H	H	2	−0.30	5.00	7	7
	H	H	2	−0.30	5.00	7	7
H	H	CN	1000	−3.00	2.70	5	4
H	4-MeOPh	CN	3.3	−0.52	4.82	6	4
	4-MeOPh	H	1000	−3.00	6.34	8	6

**Table 10**

Summary of statistics for 3D QSAR models for topoisomerase inhibition

	COMFA	COMSIA
SEP	0.77	0.79
LOO $q^2$	0.85	0.85
Number of PCs	4	4
SEE	0.59	0.60
$r^2$	0.91	0.91
F	270	259
Number of compounds	106	106
Contributions of steric, electrostatic, lipophilic, donor, acceptor fields (%)	44, 56, 0, 0, 0	13, 24, 25, 20, 19
Test set SEP	0.53	0.49
Test set $r^2$	0.94	0.95

the models reported in this work. It would be of interest to measure the topoisomerase I activity of this compound again to determine whether the reported value is accurate. It is also possible that planar extended polycyclic structure may undergo unfavourable steric interactions with the minor groove, limiting its efficacy as a ligand. If so this would not be captured by the 3D QSAR models. The 3D alignment used in the molecular field QSAR studies is illustrated in Figure 2. However, this suggests that steric bulk contributed by the outlier at the left-hand end of the molecule is unlikely to be unfavourable.

**Figure 2.** Alignment of 107 minor groove binders used in the 3D QSAR studies. The structures are hydrogen suppressed for clarity.

The models employed the two CoMFA fields (steric and electrostatic) and five COMSIA molecular fields (steric, electrostatic, lipophilic, donor, and acceptor) in Sybyl8.1 on Dell Precision workstation running Red Hat Linux. Both standard leave-one-out (LOO) cross validation and an independent test set were used to validate models and assess their predictivity. When using the test set, a random selection of 20% of the compounds was withheld, and the remaining 80% used to build the model used to predict the test set.

### 3. Results and discussion

The topoisomerase I models had high statistical significance, explaining up to 75–91% of the variance in the data, and the training and cross-validation results suggested the models were robust. The very high predictivity of the external test set also suggested that the model captured the most relevant molecular properties of the topoisomerase I inhibitors.

#### 3.1. 2D QSAR models

A range of molecular descriptors were used to model the relationship between the topoisomerase I activity and molecular properties. Interestingly, the most relevant descriptor was the number of hydrogen bond donors in the molecules,  $N_{\text{donors}}$ . A relatively good QSAR model could be obtained using this descriptor alone:

$$\text{pREC} = -6.39 + 0.99N_{\text{donors}}$$

$$n = 106, \text{SEE} = 0.98, r^2 = 0.75, F = 317, P < 0.005, \text{LOO SEP} = 1.01, q^2 = 0.74$$

where SEE and SEP and the standard errors of estimation and prediction for the training and leave-one-out (LOO) cross-validation respectively,  $r^2$  and  $q^2$  are the squared correlation coefficients for the training set and LOO cross-validation.

#### 3.2. 3D topoisomerase inhibition models

Both 3D molecular field models were of similar statistical quality, and had substantially higher statistical power than the QSAR model employing only the numbers of hydrogen bond donors. The results are summarised in Table 10.

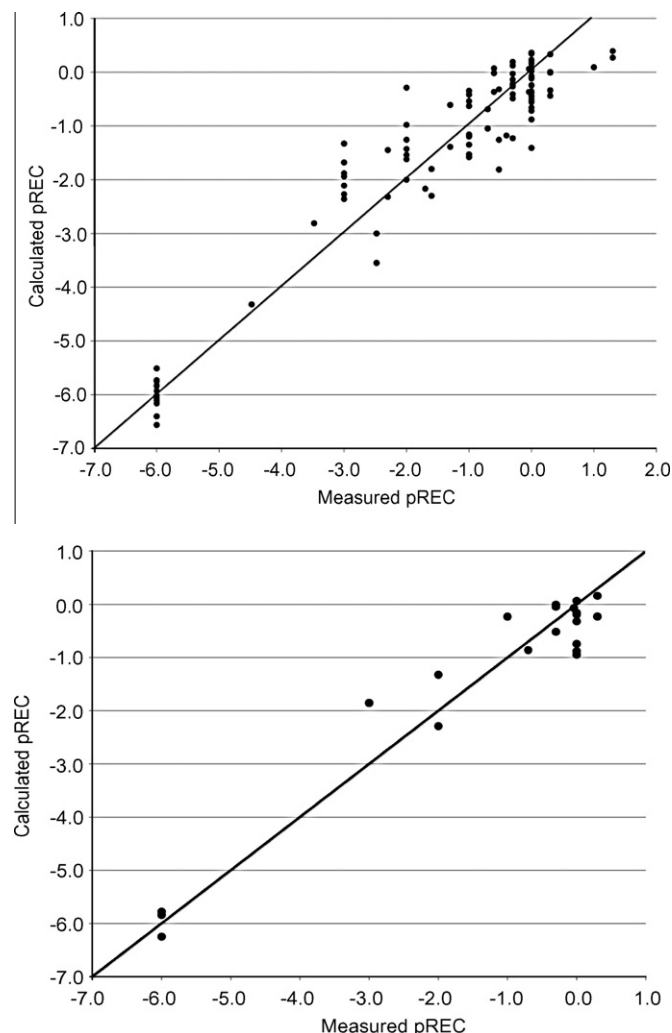
The PLS analysis showed that four principal components were optimum for both types of molecular field QSAR models. The CoMFA model showed that both steric and electrostatic fields were of similar importance to the model. The COMSIA model suggested that all five molecular fields made substantial contributions to the QSAR model.

The predictivity of the two 3D QSAR models was similar as assessed by the quality of the test set prediction, and by cross-validation. The quality of prediction of the training set for the 3D QSAR model is illustrated in Figure 3.

Although the CoMFA model was similar to the COMSIA model, the latter models are much easier to interpret. Figures 4–6 illustrate the steric, lipophilic and hydrogen bond donor properties of the model. The hydrogen bond acceptor properties are not shown as they simply reflect the positions of the heterocyclic N: atoms in the benzimidazoles. The electrostatic field maps are also not shown, as they are often hardest to interpret, and essentially reflect the positive charge on the benzimidazole NH groups in the donor map.

Figure 4 shows the COMSIA steric field map indicating where steric bulk can increase or decrease the topoisomerase activity of the minor groove binders. The green, sterically favoured region reflects the fact that terbenzimidazoles tend to have higher topoisomerase activities, unless they have large substituents attached to 2-position of the right hand benzimidazole ring. Clearly, the potent topoisomerase I inhibitor shown in 4A interacts with the green (sterically favourable) field region, but is not large enough to interact with the yellow (sterically unfavourable) field regions, as it has no substituent at the 2-position of the right-hand benzimidazole ring. The low topoisomerase I activity of the minor groove binder in 4B is due to its inability to interact with the sterically favoured green field region as the compound is too small.

Figure 5 shows regions around a potent topoisomerase I inhibitor where lipophilic and hydrogen bond donor moieties contribute

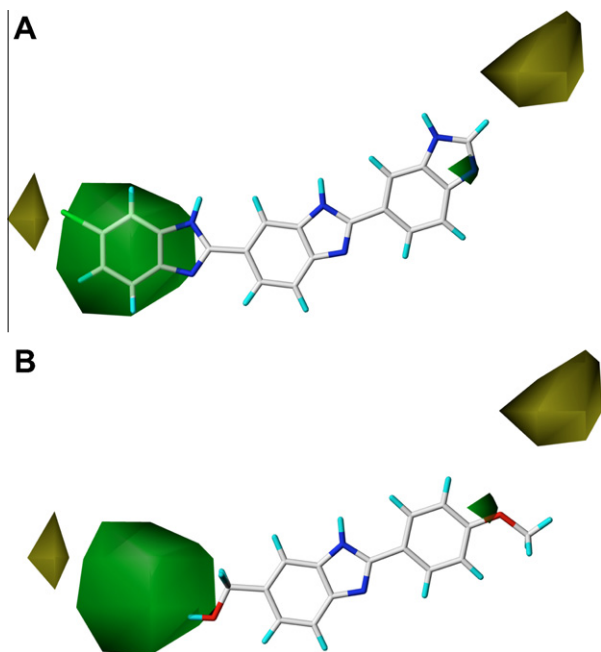


**Figure 3.** Calculated (3D CoMFA QSAR model) versus observed log relative effective concentration for calf thymus topoisomerase I inhibition. Upper graph is training set prediction, lower graph is test set prediction.

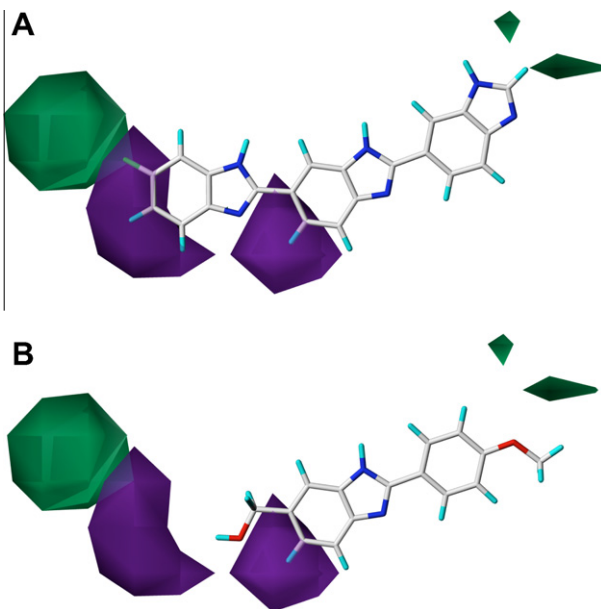
strongly to activity. This molecule clearly has lipophilic halogen atoms on the left hand phenyl ring that interact with the purple (lipophilic favoured) region. The map also shows that other compounds containing polar substituents at the 6-position of the left-hand benzimidazole would interact strongly with the green region where the presence of polar substituents enhances topoisomerase activity. Apart from being too short to interact with the sterically favoured field regions, the weak inhibitor shown in 5B also does not interact with the left-hand purple field region where lipophilicity enhances activity, and also had a polar hydroxyl substituent close to the second purple field region that favours lipophilic, rather than polar substituents.

Figure 6 illustrates the hydrogen bond donor field regions around both potent and weak topoisomerase I inhibitors. The field maps show that the central imidazole NH moiety appears to contribute less to the activity than do the equivalent groups on the other two benzimidazoles. The potent inhibitor shown in 6A is a terbenzimidazole whose three NH donor groups interact favourably with the hydrogen bond donor favoured region shown in cyan. The weak topoisomerase I inhibitor shown in 6B, however, does not have hydrogen bond donor groups that can interact with the main lobes of the donor field, and has only a single NH donor near the central, less important donor field region.



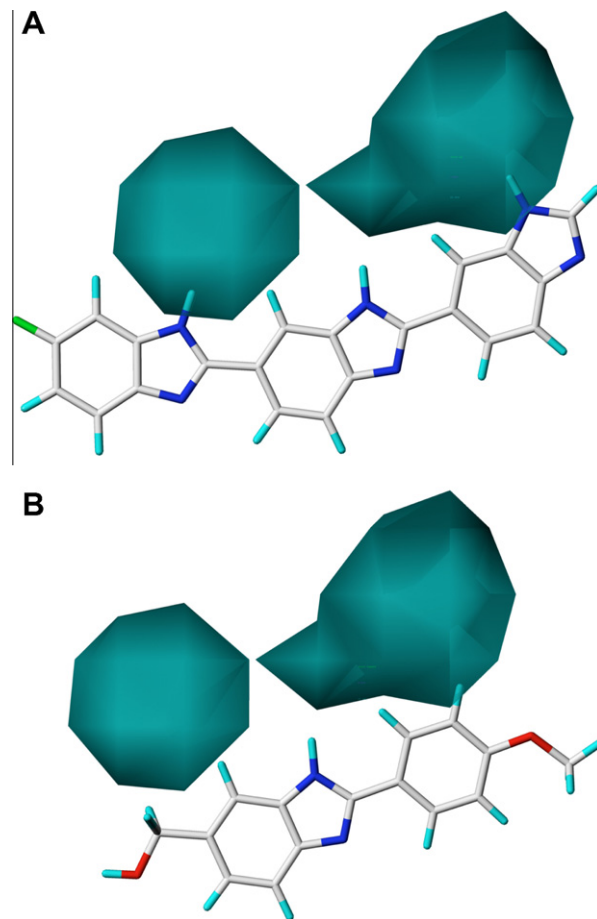


**Figure 4.** COMSIA steric molecular field mapped onto (A) potent topoisomerase inhibitor and (B) weak topoisomerase inhibitor. The green regions denote where steric bulk increases the topoisomerase I activity, and the yellow regions denote where steric bulk decreases topoisomerase I activity.



**Figure 5.** COMSIA lipophilic molecular field mapped onto (A) potent topoisomerase inhibitor and (B) weak topoisomerase inhibitor. The purple regions denote where lipophilic regions of the molecule increases topoisomerase I activity, and the green regions where lipophilic moieties in the molecule decreases topoisomerase activity.

The 3D QSAR models of topoisomerase I inhibition based on an alignment from the X-ray structure of oligobenzimidazoles bound to the minor groove of DNA were quite robust, statistically significant, and predictive. This is entirely consistent with the observations that topoisomerase I activity is related to DNA minor groove binding. Capranico et al.<sup>16</sup> recently reported that minor-groove DNA binding bibenzimidazoles and terbenzimidazoles, such as Hoechst 33258 and Hoechst 33342, stimulate topoisomerase I cleavage complex formation via an undefined mechanism.



**Figure 6.** COMSIA donor molecular field mapped onto (A) potent topoisomerase inhibitor and (B) weak topoisomerase inhibitor. The cyan regions denote where the presence of a hydrogen bond donor increases the topoisomerase I activity.

Published data showed that bibenzimidazoles can bind from position +4 to +8 downstream to topoisomerase I cleavage site. Capranico et al. suggested that they may interfere with the secondary distal DNA–enzyme contact regions, or could generate specific distortions of the DNA duplex that stabilize the topoisomerase cleavage complex. There are similarities between the MLR and 3D models, as in both cases the hydrogen bond donor efficacy of the minor groove binders is very important. As benzimidazole-like minor groove binders show a preference for binding into AT-rich regions,<sup>12</sup> and the binding is modulated by hydrogen interactions between the benzimidazole NH and the DNA bases, this result is mechanistically consistent. The MLR QSAR models for topoisomerase inhibition reported in the literature are not immediately consistent with these findings. The comprehensive review of topoisomerase I literature MLR QSAR models reported by Mekapati and Hansch<sup>3</sup> concluded that for all models involving minor groove binders, log *P* was an important factor. Lipophilicity, represented by the  $\pi$  parameters for phenyl substituents was also reported as important in a QSAR study by Kim et al.<sup>17</sup> and more generally for topoisomerase I and II by Verma.<sup>18</sup> However, these models had a much smaller number of molecules in the data set and substantially lower molecular diversity than the 3D QSAR models reported here. As the number of hydrogen bond donors and the lipophilicity of compounds are likely to be correlated in these studies, the smaller studies may have given a biased estimate of the importance of lipophilicity. In our data set  $N_{\text{donor}}$  and C log *P* were significantly correlated with a correlation coefficient of 0.58. In our MLR QSAR study, log *P* was not as significant a descriptor as  $N_{\text{donor}}$  for the

training set. The  $r^2$  value for a QSAR relationship between pREC and  $C \log P$  was only 0.26 compared to 0.75 when  $N_{\text{donor}}$  was used as the descriptor.

In summary, we report the first 3D QSAR model for topoisomerase I inhibition by a large set of minor groove binders. The data set size and diversity was substantially higher than were those used in previous classical QSAR models of this important biological response. The excellent 3D QSAR models obtained using alignments based on X-ray structures of DNA minor groove binding is consistent with such binding being important for topoisomerase I inhibition. The results from this study provide additional information on molecular properties important for topoisomerase inhibition than earlier studies had identified.

### Acknowledgments

The author acknowledges financial support from the Peter MacCallum Cancer Institute and Sirtex Medical Limited.

### References and notes

1. Leppard, J. B.; Champoux, J. J. *Chromosoma* **2005**, *114*, 75.
2. Pommier, Y.; Leo, E.; Zhang, H. L.; Marchand, C. *Chem. Biol.* **2010**, *17*, 421.
3. Mekapati, S. B.; Hansch, C. *Bioorg. Med. Chem.* **2001**, *9*, 2885.
4. Bailly, C. *Curr. Med. Chem.* **2000**, *7*, 39.
5. Nabiev, I.; Chourpa, I.; Riou, J. F.; Nguyen, C. H.; Lavelle, F.; Manfait, M. *Biochem.* **1994**, *33*, 9013.
6. Alper, S.; Arpacı, O. T.; Aki, E.; Yalcin, I. *Il Farm.* **2003**, *58*, 497.
7. Jin, S.; Kim, J. S.; Sim, S. P.; Liu, A.; Pilch, D. S.; Liu, L. F.; LaVoie, E. J. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 719.
8. Kim, J. S.; Sun, Q.; Gatto, B.; Yu, C.; Liu, A.; Liu, L. F.; LaVoie, E. J. *Bioorg. Med. Chem.* **1996**, *4*, 621.
9. Kim, J. S.; Yu, C.; Liu, A.; Liu, L. F.; LaVoie, E. J. *J. Med. Chem.* **1997**, *40*, 2818.
10. Sun, Q.; Gatto, B.; Yu, C.; Liu, A.; Liu, L. F.; Lavoie, E. J. *J. Med. Chem.* **1995**, *38*, 3638.
11. Sun, Q.; Gatto, B.; Yu, C. A.; Liu, A.; Liu, L. F.; Lavoie, E. J. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2871.
12. Martin, R. F.; Broadhurst, S.; Reum, M. E.; Squire, C. J.; Clark, G. R.; Lobachevsky, P. N.; White, J. M.; Clark, C.; Sy, D.; Spothem-Maurizot, M.; Kelly, D. P. *Cancer Res.* **2004**, *64*, 1067.
13. Squire, C. J.; Baker, S. L. J.; Clark, G. R.; Martin, R. F.; White, J. *Nucleic Acids Res.* **2000**, *28*, 1252.
14. Cramer, R. D.; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
15. Klebe, G. *Perspect. Drug Disc. Des.* **1998**, *12*, 87.
16. Capranico, G.; Marinello, J.; Baranello, L. *Biochimica et Biophysica Acta* **2010**, *1806*, 240.
17. Kim, J. S.; Sun, Q.; Yu, C.; Liu, A.; Liu, L. F.; LaVoie, E. J. *Bioorg. Med. Chem.* **1998**, *6*, 163.
18. Verma, R. P. *Bioorg. Med. Chem.* **2005**, *13*, 1059.