Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry Letters 17 (2007) 1025–1032
Bioorganic & Medicinal Chemistry Letters

Quantitative structure–activity relationship (QSAR) of indoloacetamides as inhibitors of human isoprenylcysteine carboxyl methyltransferase

Jo-Lene Leow,[a] Rudi Baron,[b] Patrick J. Casey[b] and Mei-Lin Go[a,*]

[a]*Department of Pharmacy, National University of Singapore, 18 Science Drive 4, Singapore 117543, Singapore*
[b]*Department of Pharmacology and Cancer Biology, Duke University Medical Centre, Durham, NC 27710, USA*

Received 18 July 2006; revised 26 October 2006; accepted 13 November 2006
Available online 15 November 2006

Abstract—A QSAR is developed for the isoprenylcysteine carboxyl methyltransferase (ICMT) inhibitory activities of a series of indoloacetamides ($n = 72$) that are structurally related to cysmethynil, a selective ICMT inhibitor. Multivariate analytical tools (principal component analysis (PCA) and projection to latent structures (PLS)), multi-linear regression (MLR) and comparative molecular field analysis (CoMFA) are used to develop a suitably predictive model for the purpose of optimizing and identifying members with more potent inhibitory activity. The resulting model shows that good activity is determined largely by the characteristics of the substituent attached to the indole nitrogen, which should be a lipophilic residue with fairly wide dimensions. In contrast, the substituted phenyl ring attached to the indole ring must be of limited dimensions and lipophilicity.

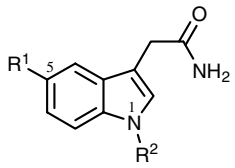
A large number of mammalian proteins contain the CaaX box motif, where C is cysteine, a is generally an aliphatic residue and X is one of several different amino acids. These proteins undergo a sequential three-step post-translational modification involving (i) prenylation of the cysteine residue mediated by one of two soluble isoprenyltransferases (protein farnesyltransferase FTase or protein geranylgeranyltransferase type I GGTase-1),[1] (ii) proteolytic removal of the aaX residues by the endoprotease Ras-converting enzyme 1 (Rce1)[2] and (iii) methylation of the newly exposed prenylcysteine by the enzyme isoprenylcysteine carboxyl methyltransferase (ICMT).[2–4] These changes result in a more hydrophobic protein with a unique structure at the C-terminus that serves as a specific recognition motif in protein–protein interactions.[5–7] Several CaaX proteins including members of the Ras family of GTPases have been implicated in oncogenesis and tumour progression. Processing by the prenylation pathway is widely thought to contribute to these roles.[6,8] For this reason, the enzymes involved in this pathway have received considerable attention as targets in drug discovery programmes.[9,10] A number of FTase inhibitors (FTIs) with high efficacy and low toxicity have been evaluated in clinical trials, but their efficacies in patients have turn out to be less than initially expected.[11] As a consequence, attention has shifted to the post-prenylation enzymes Rce1 and ICMT as potential alternative targets to FTase, in particular developing inhibitors to ICMT which catalyzes the final step of methylation of the prenylcysteine.[12–18] Recently, an indole-based selective inhibitor of ICMT was identified through the screening of a diverse chemical library comprising of over 70 subfamilies derived from unique scaffolds.[15] This compound called cysmethynil [compound **1D** in Table 1] had an in vitro IC_{50} of 2.4 μM in the initial screen for enzyme inhibitory activity.

The indole-based library from which cysmethynil was identified consists of some 70 analogues which differ in the substitution at the phenyl ring and the indole nitrogen (Table 1). These compounds were screened for ICMT inhibitory activities[19] and had IC_{50} values ranging from 2 μM to more than 50 μM.[20] Thus, they constitute a valuable database from which quantitative structure–activity relationships (QSAR) can be deduced for the purpose of optimizing and identifying indoloacetamides with more potent inhibitory activity. The

Keywords: Quantitative structure–activity relationship; ICMT inhibitory activity; Indoloacetamides; Cysmethynil; PCA and PLS; CoMFA; Multiple linear regression.
* Corresponding author. E-mail: phagoml@nus.edu.sg

0960-894X/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2006.11.030

Table 1. Structures and experimental pIC_{50}[a] values of compounds in database



No.	R^1	R^2	pIC_{50}
1A1	3-Methylphenyl	Isobutyl	4.80
1B1	3-Methylphenyl	Cyclopropylmethyl	5.04
1C1	3-Methylphenyl	*n*-Hexyl	5.13
1D[b]	3-Methylphenyl	*n*-Octyl	5.68
1E	3-Methylphenyl	Benzyl	5.19
1G1	3-Methylphenyl	*3*-Trifluoromethylbenzyl	5.10
1H1	3-Methylphenyl	2-Naphthylmethyl	4.76
1I1	3-Methylphenyl	3-Phenoxypropyl	4.00
1J1	3-Methylphenyl	H	4.00
2A1	3-Fluorophenyl	Isobutyl	5.43
2B1	3-Fluorophenyl	Cyclopropylmethyl	4.83
2C1	3-Fluorophenyl	*n*-Hexyl	4.89
2D	3-Fluorophenyl	*n*-Octyl	5.62
2E	3-Fluorophenyl	Benzyl	4.97
2F1	3-Fluorophenyl	4-*tert*-Butylbenzyl	5.30
2G1	3-Fluorophenyl	3-Trifluoromethylbenzyl	5.24
2I1	3-Fluorophenyl	3-Phenoxypropyl	4.00
2J1	3-Fluorophenyl	H	4.40
3A1	4-Methylphenyl	Isobutyl	5.01
3B1	4-Methylphenyl	Cyclopropylmethyl	4.95
3C1	4-Methylphenyl	*n*-Hexyl	5.13
3D	4-Methylphenyl	*n*-Octyl	4.97
3G1	4-Methylphenyl	3-Trifluoromethylbenzyl	5.19
3H1	4-Methylphenyl	2-Naphthylmethyl	4.71
3J1	4-Methylphenyl	H	4.00
4A1	3-Ethoxyphenyl	Isobutyl	5.21
4B1	3-Ethoxyphenyl	Cyclopropylmethyl	5.19
4C1	3-Ethoxyphenyl	*n*-Hexyl	5.18
4D	3-Ethoxyphenyl	*n*-Octyl	5.66
4E	3-Ethoxyphenyl	Benzyl	5.11
4F1	3-Ethoxyphenyl	4-*tert*-Butylbenzyl	4.90
4G1	3-Ethoxyphenyl	3-Trifluoromethylbenzyl	5.31
4H1	3-Ethoxyphenyl	2-Naphthylmethyl	5.08
4I1	3-Ethoxyphenyl	3-Phenoxypropyl	4.00
4J1	3-Ethoxyphenyl	H	4.00
5A1	2-Methoxyphenyl	Isobutyl	4.66
5B1	2-Methoxyphenyl	Cyclopropylmethyl	4.79
5C1	2-Methoxyphenyl	*n*-Hexyl	5.09
5D	2-Methoxyphenyl	*n*-Octyl	5.22
5E	2-Methoxyphenyl	Benzyl	4.88
5G1	2-Methoxyphenyl	3-Trifluoromethylbenzyl	5.21
5H1	2-Methoxyphenyl	2-Naphthylmethyl	4.53
5I1	2-Methoxyphenyl	3-Phenoxypropyl	4.00
5J1	2-Methoxyphenyl	H	4.00
6A1	3-Chloro-4-fluorophenyl	Isobutyl	5.04
6B1	3-Chloro-4-fluorophenyl	Cyclopropylmethyl	5.24
6C1	3-Chloro-4-fluorophenyl	*n*-Hexyl	5.03
6D	3-Chloro-4-fluorophenyl	*n*-Octyl	5.15
6E	3-Chloro-4-fluorophenyl	Benzyl	4.98
6F1	3-Chloro-4-fluorophenyl	4-*tert*-Butylbenzyl	5.12
6G1	3-Chloro-4-fluorophenyl	3-Trifluoromethylbenzyl	5.21
6I1	3-Chloro-4-fluorophenyl	3-Phenoxypropyl	4.00
6J1	3-Chloro-4-fluorophenyl	H	4.00
7A1	3,5-Bis(trifluoromethyl)phenyl	Isobutyl	4.59
7B1	3,5-Bis(trifluoromethyl)phenyl	Cyclopropylmethyl	4.00
7D	3,5-Bis(trifluoromethyl)phenyl	*n*-Octyl	4.43
7E	3,5-Bis(trifluoromethyl)phenyl	Benzyl	4.48
7F1	3,5-Bis(trifluoromethyl)phenyl	4-*tert*-Butylbenzyl	4.53

Table 1 (*continued*)

No.	R^1	R^2	pIC_{50}
7G1	3,5-Bis(trifluoromethyl)phenyl	3-Trifluoromethylbenzyl	4.00
7H1	3,5-Bis(trifluoromethyl)phenyl	2-Naphthylmethyl	4.00
7I1	3,5-Bis(trifluoromethyl)phenyl	3-Phenoxypropyl	4.00
7J1	3,5-Bis(trifluoromethyl)phenyl	H	4.00
8A1	4-Phenoxyphenyl	Isobutyl	5.23
8B1	4-Phenoxyphenyl	Cyclopropylmethyl	4.00
8C1	4-Phenoxyphenyl	*n*-Hexyl	4.00
8E	4-Phenoxyphenyl	Benzyl	5.03
8G1	4-Phenoxyphenyl	3-Trifluoromethylbenzyl	4.00
8H1	4-Phenoxyphenyl	2-Naphthylmethyl	4.00
8I1	4-Phenoxyphenyl	3-Phenoxypropyl	4.00
8J1	4-Phenoxyphenyl	H	4.00

[a] Determination of pIC_{50} is given in Ref. 19.
[b] Cysmethynil.

objective of this work was to develop suitably predictive QSAR models that can achieve these aims.

To develop the QSAR, several descriptors representative of size, electronic and lipophilic characteristics were used to characterize the compounds. Area, volume (including polar surface area and polar volume) and the Sterimol parameters (L1, B1, B5)[21] of the substituted phenyl ring and the group attached to the indole nitrogen was chosen to characterize the size component. Lipophilicity was represented by C$\log P$ and Hansch π values of the substituted phenyl ring (π_{Ph2}) and the N-substituent (π_N). The electronic character of the compounds were captured by HOMO, LUMO, molar refractivity, dipole moment, Hammett constant ($\sigma_{aromatic}$) of the phenyl substituents and the inductive constant (σ^*) of the group attached to the indole nitrogen. The parameters were determined from forcefield minimized structures of the compounds using commercially available softwares[22] and are listed in Supplementary Information (Table 1).

Our first approach was to employ multivariate tools (principal component analysis PCA and projection to latent structures PLS) to analyze the structure–activity relationship.[23] PCA, a pattern recognition technique, serves to summarize information in a form that can reveal relationships between compounds and the parameters used to characterize them.[24] It is particularly useful in situations where many of the parameters are correlated, as in this case. PCA will then transform the correlated variables to a new set of uncorrelated variables called principal components. When the data set (72 compounds, 20 parameters) was analyzed by PCA, a significant three-component model ($r^2 = 0.72$, $q^2 = 0.54$) was obtained, which implies that the properties captured by these components can account for 72% of the observed variation, at a predictability level of 54%. An examination of the loading plot of this model (Fig. 1a) shows that the first component receives input from size (volume, area, molar refractivity, Sterimol parameters of N) and lipophilicity (C$\log P$, π_N). The parameters contributing to the second component are HOMO, lipophilicity (π_{Ph2}) of the substituted phenyl ring and the Sterimol length (PL1) and width (PB1, PB5) parameters of the substituted phenyl ring. Interestingly, the first component receives steric and lipophilicity inputs from the N-substituent in contrast to the less important second component which is characterized by the same parameters of the phenyl ring. The score plot which depicts the distribution of the 72 compounds based on the properties captured by the first and second principal components is shown in Figure 1b. Most of the compounds are distributed uniformly within the ellipse, with two exceptions, namely compounds with no N-substituent (lower left quadrant) and compounds with 3,5-bis(trifluoromethyl) substitution on the phenyl ring (lower right quadrant). The isolation of these two subsets, which incidentally are among the weakest inhibitors of the series, can be attributed to their less than optimal steric and lipophilic properties which are encoded in the first and second principal components.

Next, we went on to develop a projection model for predicting biological activity from the principal components. This is achieved with PLS, a regression extension of PCA. The first PLS model was obtained with all compounds and parameters. This model accounted for 68% of the variation in biological activity ($r^2 = 0.68$) at a predictability level of 49% ($q^2 = 0.49$). It was improved by omitting two outliers which exhibited large differences between the observed and predicted values (compounds **3E** and **8D**) and removing parameters that did not make significant contributions to activity (dipole moment, σ_N, σ_{Ph}) or are duplicated by existing parameters (volume, surface area, π_{Ph2}). The resulting model (A) had an r^2 of 0.72 and q^2 of 0.60. A test set of 12 compounds was arbitrarily selected from the 70 compounds and used to predict the activities of the remaining 58 compounds. The level of predictability, as measured by the root mean square of prediction (RMSEP), was just satisfactory at 0.48. Figure 1c is the coefficient plot of the final PLS model A. This plot identifies the parameters that contributed most to activity (as reflected by the length of the bar), and the nature of the correlation (direct = positive coefficient or inverse = negative coefficient). Thus, the most significant parameters (in order of decreasing importance) are polar surface area and polar volume > lipophilicity of substituted phenyl ring (π_{Ph2}) > width parameters B1, B5 of the N-substituent and π_N > Sterimol L1 of N-substituent. The coefficient plot nicely illustrates the contrasting

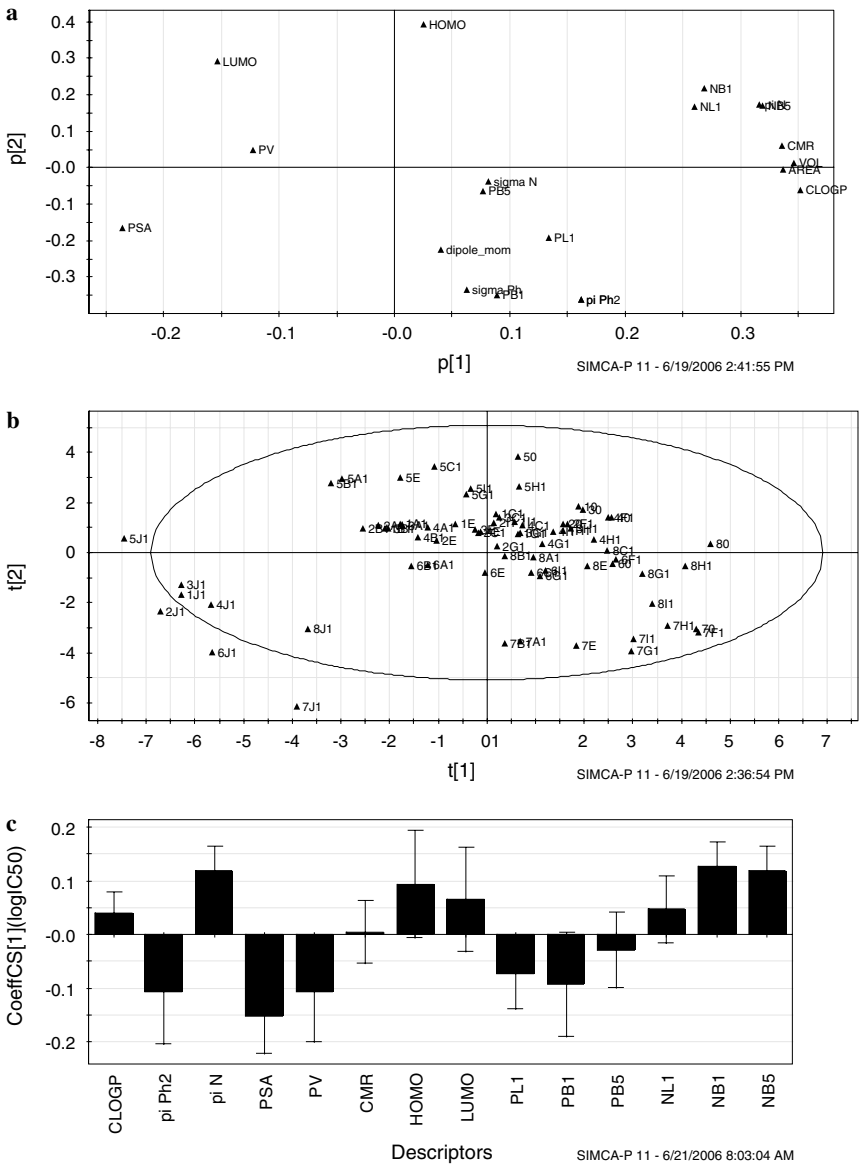


Figure 1. (a) Loading plot of first and second principal components (p[1], p[2]) of 72 compounds and 20 descriptors. (b) Score plot of principal components t_1 versus t_2 for compounds ($n = 72$, 20 descriptors). Compounds are identified in Table 1. The ellipse corresponds to the confidence region based on Hotelling T^2 (0.05). Compounds identified by '7' have 3,5-bis(trifluoromethyl) groups on the phenyl ring and compounds identified by 'J' have no substituent on the indole N. Both sets of compounds are isolated from the rest. (c) Coefficient plot for PLS model A derived from 70 compounds and 14 descriptors, based on the first component. Descriptors with positive coefficients are directly related to activity, while those with negative coefficients are inversely related. Magnitude of parameter indicates its relative contribution to the model. PSA = polar surface area; PV = polar volume; pi Ph2 = Hansch constant π of substituted phenyl ring; pi N = Hansch constant π of N-substituent; CMR = molar refractivity; PL1, PB1, PB5 = sterimol parameters of phenyl ring; NL1, NB1, NB5 = sterimol parameters of N-substituent; HOMO = highest occupied molecular orbital; LUMO = lowest unoccupied molecular orbital.

requirements of the N and phenyl ring substituents. For good activity, the N group must be lipophilic and have large width dimensions. In contrast, the substituted phenyl ring should be less lipophilic and have sterically smaller dimensions.

To further validate PLS model A, the same database was used to derive a stepwise multiple linear regression (MLR) equation.[25] The best Equation 1 has four descriptors (polar surface area PSA, polar volume PV, Sterimol parameter B1 of the substituted phenyl ring and lipophilicity contribution of the substituted phenyl ring π_{Ph2}). When the coefficients in Eq. 1 are standardized, the relative contribution of each parameter is of the order PSA > PV > PB1 > π_{Ph2}. The parameters are inversely related to activity, with the first three parameters significant at $p < 0.001$ and π_{Ph2} significant at $p = 0.001$.

$$\begin{aligned} \mathrm{pIC_{50}} = {} & 10.23(\pm 0.55) - 0.016(\pm 0.003) * \mathrm{PSA} \\ & - 0.010(\pm 0.002) * \mathrm{PV} - 0.835(\pm 0.195) \\ & * \mathrm{PB1} - 0.190(\pm 0.056) * \pi_{\mathrm{Ph2}}. \end{aligned} \quad (1)$$

$n = 70$, $r = 0.811$, $r^2_{adjusted} = 0.636$, standard error of estimate = 0.32, $F = 30.69$.

The Sterimol parameters and lipophilicity contribution of the N-substituent are not found in Eq. 1, but are encoded in polar surface area, to which they are significantly correlated (Supporting Information, Table 1). The inclusion of PB1 in Eq. 1 is notable because PB1 did not contribute significantly to the PLS model. PB1 is also not strongly correlated to the other Sterimol parameters (PB5, PL1) but received a significant contribution from HOMO.

At this juncture, some deductions can be made for the structural requirements for inhibitory activity of the indoloacetamides. It is evident that there is a requirement for compounds with small polar surface areas, a feature that is largely determined by the dimensions and lipophilicity of the N-substituent. The relationship is inverse, so good activity is associated with lipophilic N substituents. The bulk of the N-substituent is also critical and in this context, the width parameters of the group (B1, B5) are just as important as the length of the group (L1). Thus, the poor activity of compounds with *N*-3-phenoxypropyl substituent may be traced to its lack of bulk (long and narrow). This group is almost as long as the *N*-octyl substituent found in cysmethynil but the *N*-octyl group is more flexible and thus has a larger width (NB5) than the *N*-3-phenoxypropyl group. Another significant observation is that the substituted phenyl ring had a lesser impact on activity compared to the N group. In contrast to the preference for more lipophilic and bulkier N substituents, the substituted phenyl ring must be kept small (width and length) and less lipophilic. Thus, the poor activities of compounds with 3,5-bis (trifluoromethyl)phenyl and *N*-3-phenoxyphenyl substituents can be traced to the unfavourable width and length of these groups, respectively.

Further confirmation of the QSAR developed so far was sought from comparative molecular field analysis (CoMFA).[26] Based on a preliminary analysis, three compounds (8A1, 3E and 8D) were omitted because of large residual values. The remaining 69 compounds were divided into a training set ($n = 56$) and a test set ($n = 13$). Compounds in both sets had the same range of biological activity and were well represented in terms of the type of substituents on the indole nitrogen and phenyl rings. The best CoMFA model for the training set had a cross-validated q^2 of 0.646 (seven principal components), a non cross-validated r^2 of 0.868 and standard error of estimate (SEE) of 0.209 (Table 2). The steric and electrostatic field contributions were 68% and 32%, respectively, indicating a greater steric influence on activity. The training set was able to predict the activity of the test set compounds with an r^2_{pred} of 0.601. A plot of predicted versus observed values is given in Figure 2.

Visualization of the steric contours of this model shows a swathe of green distributed at the proximal and distal ends of the N-substituent. A yellow patch is interspersed between these green zones (Fig. 3a). The inference is that there are two optimal lengths of the N-substituent corresponding to the green contours—a short group that is within reach of the proximal green patch and a longer group that coincides with the distal green patch. The shorter *N*-3-trifluoromethylbenzyl and longer *N*-octyl groups, both of which are associated with good activity, fall within these green zones. On the other hand, groups linked to poor activity like *N*- (2-naphthyl)methylene and *N*-(4-*tert*-butylbenzyl) protruded into the disfavoured yellow zone. Interestingly, only CoMFA revealed the presence of two optimal lengths of the N-substituent. The other models emphasized the preference for a bulky N-group that presumably extended into the green zones. Earlier, the poor activity of the *N*-3-phenoxypropyl group was attributed to its narrow dimensions. The CoMFA steric contour map shows that the phenyl ring of this substituent extended into the

Table 2. Summary of CoMFA analysis of training set ($n = 56$)

Cross-validated correlation coefficient q^2	0.646
Standard error of prediction (SEP)	0.342
Number of Components[a]	7
Non cross-validated correlation coefficient r^2	0.868
Standard error of estimate (SEE)	0.209
F value	44.91
Field Contributions	
Steric	0.68
Electrostatic	0.32
r^2_{pred} [b]	0.601

[a] Optimum number of components obtained from cross-validated partial least square analysis. This number of components is also used in the final non cross-validated analysis.

[b] Predictive r^2 is based only on the 13 compounds not included in the training set. It is determined from r^2_{pred}=[SD − PRESS]/SD where SD = sum of the squared deviations between IC_{50} of compounds in test set and mean IC_{50} of training set molecules, and PRESS is sum of the squared deviations between predicted and actual IC_{50} for every member in the test set.

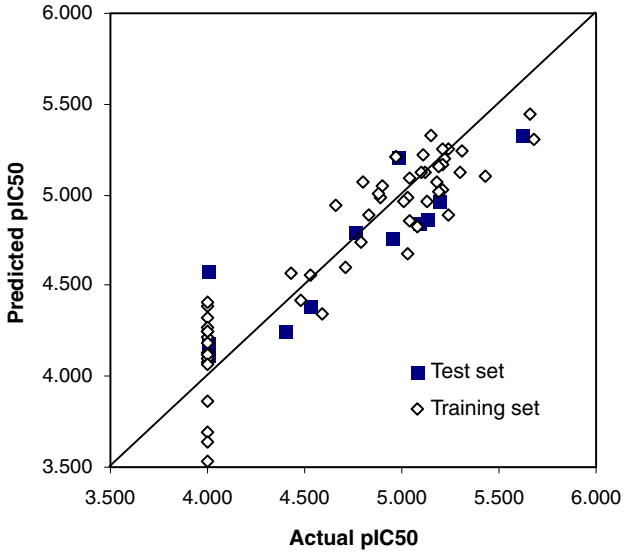


Figure 2. Plot of the predicted pIC_{50} versus observed pIC_{50} values of $n = 69$ compounds (compounds omitted are 3E, 8A1 and 8D) based on best CoMFA model. r^2 for training set = 0.868, r^2_{pred} for test set = 0.601.

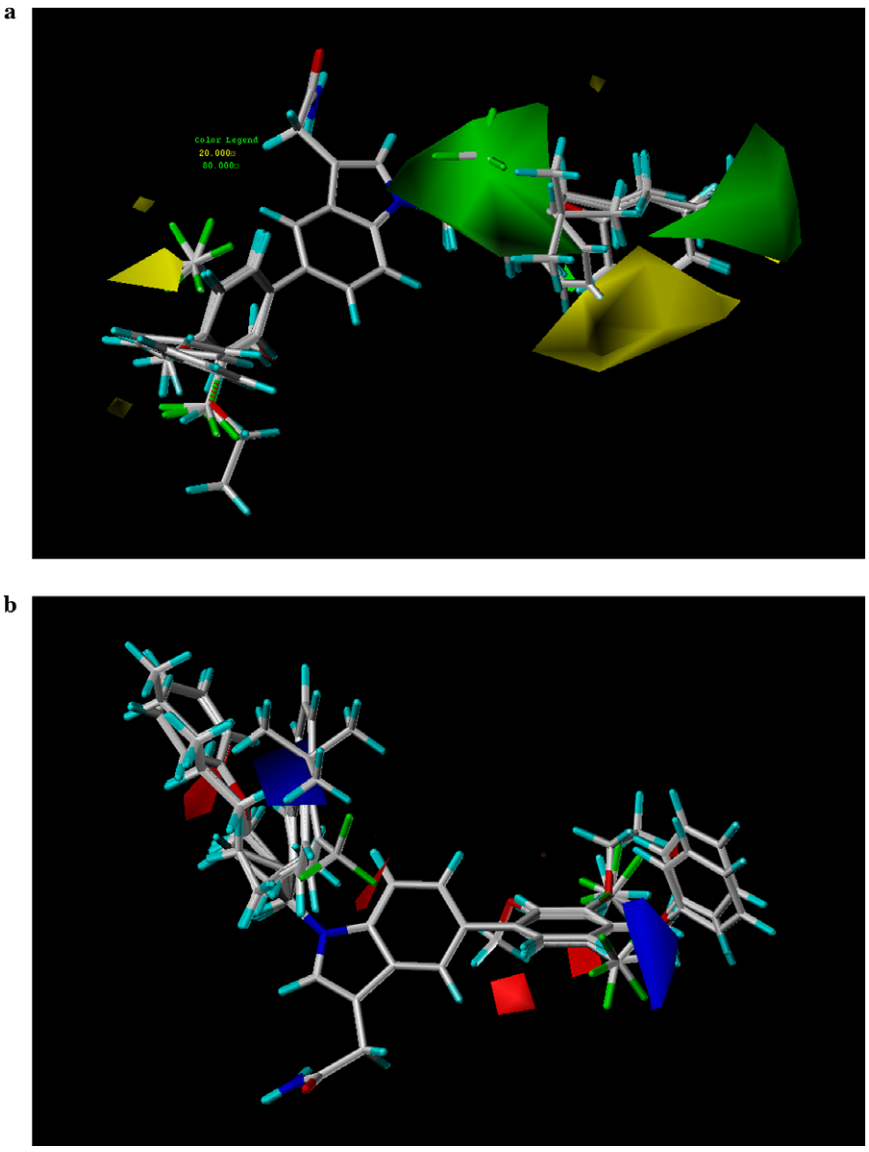


Figure 3. (a) Steric map from the CoMFA model showing the alignment based on the indole ring. Green contours (contribution level of 80%) represent areas where steric bulk will enhance activity, and yellow contours (contribution level of 20%) highlight areas which should be kept unoccupied for increased activity. (b) Electrostatic map from the CoMFA model showing the same alignment as in (a). Blue contours (contribution level of 85%) represent regions where an increase in positive charge will enhance activity, and red contours (contribution level of 15%) highlight areas where more negative charge is favoured.

sterically disfavoured yellow zone. Only yellow steric contours were found in the vicinity of the phenyl ring, but these were sparse compared to the yellow zones around the N-group. The inference is that the steric characteristics of the phenyl ring contributed less to activity and there is a preference for less bulk at this position.

None of the models investigated so far has identified electrostatic characteristics to be important contributors to activity, and this was further confirmed by the CoMFA model. The main feature in Figure 3b that can be correlated with the findings from PLS and MLR was the red contours around the phenyl ring, which implied that the ring should be electron-rich for activity. As shown in Eq. 1, the width parameter PB1 of the phenyl ring is inversely correlated to activity. PB1 is significantly correlated to HOMO (Pearson correlation coefficient—0.616, $p < 0.0001$). Molecules with large HOMO imply electron-rich structures. Thus, the poor activity associated with the 3,5-bis(trifluoromethyl) substituted phenyl ring may be attributed to its steric and electrostatic features. Sterically, it is too wide (B1) and the electron-withdrawing trifluoromethyl groups render the phenyl ring electron deficient (low HOMO).

To put our models to the test, several compounds with different substituents on the indole nitrogen and phenyl ring were considered and their activities predicted from Eq. 1 and the CoMFA model. For example, 2-(1-octyl-5-*m*-methoxy-1H-indol-3-yl)acetamide has a predicted pIC_{50} of 5.57 based on Eq. 1 and 5.52 based on CoMFA. Noting that the predicted pIC_{50} values of cysmethynil are 5.72 (Eq. 1) and 5.32 (CoMFA), the synthesis of this compound is worth considering.

In conclusion, we have used different approaches to establish the QSAR of indoloacetamides as inhibitors of ICMT. The approaches are complementary and serve to give a more definitive picture of the structural requirements for activity. These are (i) the presence of a lipophilic residue on the indole nitrogen, in contrast to the requirement for a less lipophilic substituted phenyl ring, (ii) the presence of two optimal lengths for the N group which should also have fairly wide dimensions, and (iii) the phenyl ring, which unlike the N-substituent should not have bulky substituents and must be electron-rich. These findings also serve to delineate the pharmacophore for activity but as it stands, remains incomplete because the changes made so far have focused on the N-substituent and the phenyl ring. The primary amide side chain has not been structurally altered and a more comprehensive picture will emerge when concurrent changes are made at this location. In any case, the QSAR discussed in this report will provide directions for future structural modifications of this class of compounds and prioritizing candidates for synthesis.

Acknowledgments

This work was supported by the Medicinal Chemistry Program of the Office of Life Sciences, National University of Singapore (M.L.G.), and National Institutes of Health Grant GM46372 (P.J.C). J.L.L. is supported by a research scholarship from the National University of Singapore.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmcl.2006.11.030.

References and notes

1. Casey, P. J.; Seabra, M. C. *J. Biol. Chem.* **1996**, *271*, 5289.
2. Ashby, M. N. *Curr. Opin. Lipidol.* **1998**, *9*, 99.
3. Schmidt, W. K.; Tam, A.; Fujimura-Kamada, K.; Michaelis, S. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 11175.
4. Clarke, S.; Vogel, J. P.; Deschenes, R. J.; Stock, J. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4643.
5. Glomset, J. A.; Farnsworth, C. C. *Annu. Rev. Cell. Biol.* **1994**, *10*, 181.
6. Kloog, Y.; Cox, A. D. *Semin. Cancer Biol.* **2004**, *14*, 253.
7. Winter-Vann, A. M.; Casey, P. J. *Nat. Rev. Cancer* **2005**, *5*, 405.
8. Doll, R. J.; Kirschmeier, P.; Bishop, W. R. *Curr. Opin. Drug Discov. Devel.* **2004**, *7*, 478.
9. Gibbs, J. B. *Cell* **1994**, *77*, 175.
10. Mazieres, J.; Pradines, A.; Favre, G. *Cancer Lett.* **2004**, *206*, 159.
11. Kohl, N. E.; Omer, C. A.; Conner, M. W.; Anthony, N. J.; Davide, J. P.; Desolms, S. J.; Giuliani, E. A.; Gomez, R. P.; Graham, S. L.; Hamilton, K.; Handt, L. K.; Hartman, G. D.; Koblan, K. S.; Kral, A. M.; Miller, P. J.; Mosser, S. D.; O'Neill, T. J.; Rands, E.; Schaber, M. D.; Gibbs, J. B.; Oliff, A. *Nat. Med.* **1995**, *1*, 792.
12. Henriksen, B. S.; Anderson, J. L.; Hycryna, C. A.; Gibbs, R. A. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 5080.
13. Anderson, J. L.; Henriksen, B. S.; Gibbs, R. A.; Hrycyna, C. A. *J. Biol. Chem.* **2005**, *280*, 29454.
14. Donelson, J. L.; Hodges, H. B.; MacDougall, D. D.; Henriksen, B. S.; Hrycyna, C. A.; Gibbs, R. A. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 4420.
15. Winter-Vann, A. M.; Baron, R. A.; Wong, W.; dela Cruz, J.; York, J. D.; Gooden, D. M.; Bergo, M. O.; Young, S. G.; Toone, E. J.; Casey, P. J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 4336.
16. Shi, Y. Q.; Rando, R. R. *J. Biol. Chem.* **1992**, *267*, 9547.
17. Perez-Sala, D.; Gilbert, B. A.; Tan, E. W.; Rando, R. R. *Biochem. J.* **1992**, *284*, 835.
18. Winter-Vann, A. M.; Kamen, B. A.; Bergo, M. O.; Young, S. G.; Melnyk, S.; James, S. J.; Casey, P. J. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 6529.
19. Baron, R. A.; Casey, P. J. *BMC Biochem.* **2004**, *5*, 19, Briefly, the ICMT inhibitory activity assay involved quantification of [^{3}H] methyl incorporation into the substrate biotin-*S*-farnesyl L-cysteine (BFC). Reactions were initiated by addition of Sf9 membranes containing ICMT to an assay mixture containing BFC and [^{3}H]AdoMet in 100 mM Hepes, pH 7.4, and 5 mM $MgCl_2$. Reactions were carried out at 37 °C for 20 min and terminated by addition of 10% Tween 20. Following termination, streptavidin beads were added, and the mixture mixed by gentle agitation overnight at 4 °C. The beads were harvested by centrifugation in a tabletop microcentrifuge (10,000 rpm, 5 min) and washed three times with 20 mM NaH_2PO_4, pH 7.4, containing 150 mM NaCl. The beads were then suspended in the same buffer, transferred to scintillation vials for radioactivity measurements.
20. Weak inhibitors with IC_{50} values greater than 50 μM were not accurately determined. This posed a problem when constructing the PCA and PLS models. Thus, we investigated three models in which weakly active compounds were arbitrarily assigned IC_{50} values of 50, 100 or 200 μM. The best model was obtained when IC_{50} of weak inhibitors were 100 μM. Thus in Table 1, compounds with $IC_{50} > 50$ μM were assigned $pIC_{50} = 4.00$.
21. Verloop, A.; Hoogenstraaten, W.; Tipker, J.. In *Drug Des.*; Ariens, E. J., Ed.; Academic Press: New York, 1976; Vol. 7, pp 165–207.
22. Parameters determined from energy-minimized geometries using the Sybyl 7.0 standard Tripos forcefield (Tripos Associates, St. Louis, Mo) are C log *P*, area, volume, polar surface area, polar volume, molar refractivity and the Hansch π values. π_N is determined from $C\log P_{\text{compound}} - C\log P_{\text{compound without N-substituent}}$ and π_{Ph2} is determined from $C\log P_{\text{compound}} - C\log P_{\text{compound without phenyl ring}}$. The remaining parameters were determined from energy-minimized geometries using the Molecular Modeling Pro Plus Version 6.2.3 (Chemsoftware) forcefield (MM2). The Sterimol parameter L1 measures the substituent length along the axis formed by the bond between the substituent and the atom to which it is attached. B1 is the shortest radius perpendicular to the L1 axis and B5 is the longest possible radius perpendicular to L1.
23. PCA and PLS are analyzed using SIMCA – P + 11 version 11 (2005) (Umetrics AB, Umea, Sweden).
24. Livingstone, D. *Data Analysis for Chemists*; Oxford University Press: Oxford, 1995.
25. SPSS 14.0 for Windows.
26. CoMFA is carried out on LINUX Redhat using Sybyl 7.0 molecular modeling software (Tripos Inc., St Louis, MO) The compounds are built using fragments in the

Sybyl database and fully geometry optimized using the standard Tripos forcefield with a distance-dependent dielectric function until a root mean square (rms) deviation of 0.001 kcal/mol Å is achieved. The partial atomic charges required for the electrostatic interaction are computed using the Gasteiger–Huckel method. Compounds are aligned using the indole core as template. CoMFA steric and electrostatic interaction fields are calculated at each lattice intersection point of a regularly spaced grid of 2.0 Å. The grid pattern, generated automatically by the Sybyl/CoMFA routine, extends 4.0 Å units in *X*, *Y* and *Z* directions beyond the dimensions of each molecule. The steric term, which represents van der Waals (Lennard-Jones) interaction, and the coulombic term, which represents the electrostatic interactions, are calculated using the standard Tripos forcefield. A distance-dependent dielectric constant is used and an sp^3 carbon atom with a van der Waals radius of 1.52 Å and +1.0 charge is used as the probe to calculate the steric and electrostatic fields. Values of the steric and electrostatic fields are truncated at 30 kcal/mol.