Supporting Information

Quantitative Structure-Activity Relationship (5D-QSAR) Study of Combretastatin-Like Analogues as Inhibitors of Tubulin Assembly

Sylvie Ducki,[†,§] Grant Mackenzie,[†] Nicholas J. Lawrence,[‡] James P. Snyder*[,§]

[†] *Centre for Molecular Drug Design, Cockcroft Building, University of Salford, Salford M5 4WT, UK.* [§] *Department of Chemistry, Emory University, 1515 Dickey Drive, Atlanta, GA 30322, USA. Department of Chemistry,* [‡] *Cardiff University, P.O. Box 912, Cardiff, CF10 3TB, UK.*

Data Set. The 3D molecular structures of the 47 ligands were generated using *Macromodel 6.5*[1] and optimized with *MMFF* force field complemented by the continium water model (GBSA/H_2O).[2]

Monte Carlo (MC) conformational searches. These were performed within *Macromodel 6.5*[3] using the *MMFF* force field and 1,000 step searches within 20.9 kJ/mol (5 kcal/mol) of the global minimum, as described by Jansen.[4] All structures were subjected to the truncated Newton conjugate gradient (*TNCG*) minimization method to a convergence criterion of 0.01 kJ/(Å.mol). The energy of the lowest energy structure, the number of conformations found and the frequency with which the simulation visited the lowest energy structure were monitored to assure an exhaustive search (Table 2). Within the MC framework, if the global minimum conformation is found 15-20 separate times, it can be assumed that the search is complete.

APOLLO alignment. Input to the alignment procedure consists of a set of MC-derived conformations for each ligand. The module *RMSFIT* from the program *APOLLO*[4-6] was used to identify those conformations of the different ligands that exhibited the best overall least-squares fit to specified atomic positions on a conformer of colchicine: the three oxygen atoms in the A-ring (weight 1.0), and the two the oxygens of ring C (weight 0.1). The specific conformer of colchicine selected as template was derived from a full conformational search using *MMFF/Macromodel*. It corresponds to a boat conformation for ring-B with the amido side chain in the equatorial position. The latter two features are found in colchicine's global minimum. The latter also orients the amide's NH group toward ring-A. However, the conformer used in the QSAR alignment involves a 180° rotation about the CH-NH bond, directing the amide C=O toward ring-A. This orientation of

the amide was chosen to match the colchicine conformer best suited to an explicit tubulin-colchicine model under development.[7]

The energies of the conformations were used together with the root mean square (RMS) deviations between conformer and target to score the matches. Suitable superpositions based on the score and diversity in orientation were extracted using the *MMDFIT* module.[5,6] For the 47 ligand molecules defining the data set, the inclusion of multiple torsional isomers resulted in a total of 160 conformers.

5D-QSAR Analysis. The data input for *Quasar 3.5* was prepared with the *PrGen* software[8] and included 1-6 conformations for each ligand, atomic partial charges, ligand solvation energy (E_{solv}), ligand entropy correction ($T\Delta S$), internal energies relative to the lowest-energy conformer (ΔE_{int}) and free energies of binding ($\Delta G°_{exp}$). Atomic charges for each structure were obtained from MOPAC and correspond to MNDO ESP charges scaled to HF/6-31G* values.[9] A solvation free energy (ΔG_{solv}) for each structure was derived with the AMSOL AM1/SM5.4 procedure and falls in range from -7 to -25 kcal/mol.[10,11] Free energies of binding ($\Delta G°_{exp}$) were calculated from experimental IC_{50} values (Concentration of ligand required to inhibit 50% of tubulin assembly) as follows: $\Delta G°_{exp}$ (kcal/mol) = RT ln IC_{50} (M), where R = 1.987×10^{-3} kcal/mol.K and T = 298 K. Manipulation of *Quasar 3.5* to generate a QSAR correlation as described below ultimately produces a predicted free energy of binding ($\Delta G°_{pred}$).

The mean envelope was generated about the ligands of the training set using all the induced-fit models available in *Quasar 3.5*: a linear induced fit scaled at 0.75, four field-based modes (steric, electrostatic, H-bond and lipophilicity), and minimization along the steric lines scaled at 1.00. Individual envelopes of the ligands of the dataset were then generated. Points on the receptor surface were initially randomly populated with atomistic properties (hydrophobicity, H-bonds, salt bridges) to furnish a starting population of receptor models. Using a genetic algorithm, this family of receptor models was allowed to evolve with cross-validation. The following *Quasar*-specific settings were used, unless otherwise specified: Equalization = Both (Equalising solvation energies and atomic partical charges); polarization = None (No ligand-receptor polarisation); Attenuation factors = 1.0 (weight of solvent), 1.0 ($T\Delta S$), 1.0 (Internal energy), 1.0 (Polarisation), 1.0 (Lack-of-fit), 2.0 (Cross-validation); Static surface area = None; Cross-validation groups n = 3

(Internal splitting of training set for cross-validation pruposes); H-bond function shift = gsm (g.s.mean H-bond radii); Induced-fit weight = 1.0 (Weight of envelope-adaptation energy); Speed-up factor = 0.9 (Selection criterion for low-frequent induced fit models); Dynamic surface area = None (Using a generic algorithm to identify solvent-accessible or "open" regions"); Size of parent family = 200 (genetic pool); Number of cross-overs (varied according to termination criterion); Mutation rate = 0.02 (Transcription error rate during crossover events); Target q^2 = 0.95 (Termination criterion, usually > 0.9); Experimental error = 0.200 (Termination criterion), Crop collection = a 1.0 (asymmetric selection); and Functional-group analysis = Yes (Contribution of functinal group to ΔG).

References

(1) http://www.schrodinger.com/Products/macromodel.html

(2) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **1990**, *112*, 6127-6129.

(3) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. Macromodel - an Integrated Software System for Modeling Organic and Bioorganic Molecules Using Molecular Mechanics. *J. Comput. Chem.* **1990**, *11*, 440-467.

(4) Jansen, J. M.; Copinga, S.; Gruppen, G.; Molinari, E. J.; Dubocovich, M. L.; Grol, C. J. The high affinity melatonin binding site probed with conformationally restricted ligands.1. Pharmacophore and minireceptor models. *Bioorg. Med. Chem.* **1996**, *4*, 1321-1332.

(5) Koehler, K. F.; Spangler, D. P.; Snyder, J. P. Pharmacophore Identification through Molecular Similarity. *Abstr. Pap. Am. Chem. Soc.* **1988**, *196*, 35-COMP.

(6) Snyder, J. P.; Rao, S. N.; Koehler, K. F.; Pellicciari, R. APOLLO. *Trends in Receptor Research*; Angeli, P.; Gulini, U.; Quaglia, W.: Amsterdam, 1992; pp 367-403.

(7) Nettles, J. H.; Mackenzie, G.; Ducki, S.; Snyder, J. P. **2004**, in preparation.

(8) Zbinden, P.; Dobler, M.; Folkers, G.; Vedani, A. PrGen: Pseudoreceptor modeling using receptor-mediated ligand alignment and pharmacophore equilibration. *Quant. Struct.-Act. Relat.* **1998**, *17*, 122-130.

(9) Besler, B. H.; Merz, K. M.; Kollman, P. A. Atomic Charges Derived from Semiempirical Methods. *J. Comput. Chem.* **1990**, *11*, 431-439.

(10) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium. *J. Phys. Chem.* **1996**, *100*, 19824-19839.

(11) AMSOL; http://comp.chem.umn.edu/amsol/