Multiple Flexible Alignment with SEAL: A Study of Molecules Acting on the Colchicine Binding Site

Miklos Feher* and Jonathan M. Schmidt

Nanodesign Inc., Suite 300, Research Park Centre, 150 Research Lane, Guelph, Ontario, Canada N1G 4Y5

Received July 2, 1999

An extension of the steric and electrostatic alignment alignment (SEAL) method (MultiSEAL) is described that allows the overlay of multiple molecules and conformations. The method is well-suited for the systematic study of possible alignments, also revealing information about the conformational energies associated with a given overlay. It has been tested on three examples: angiotensin II antagonists, 5-HT_3 antagonists, and dopaminergic compounds. The utility of the method is further demonstrated in an analysis of molecules that putatively bind to the colchicine site of tubulin. On the basis of its overlay with colchicine, allocolchicine, 2-methoxy-5-(2′,3′,4′-trimethoxyphenyl)tropone, and combretastatin A-4, it appears that 2-methoxyestradiol (2-ME) is unlikely to fit the colchine site properly. The weak antimitotic activity of 2-ME may be explained by its partial fit in the site.

INTRODUCTION

Aligning molecules is generally a very important step in computational drug design.[1] It is a prerequisite for most 3D quantitative structure−activity relation (QSAR) calculations[2] and can help to identify the structural basis for the common action of drug molecules. There are a number of different approaches to the alignment of molecules,[3−8] among them the powerful SEAL method (steric and electrostatic alignment).[9] In this process, the overlay proceeds by calculating an alignment function that sums a set of atomic functions containing electronic terms (partial atomic charges) and spatial factors (van der Waals atomic radii). Possible alignments are identified by randomly rotating and translating structures with respect to the other and then minimizing the alignment function for each orientation. The major advantage of the method is that it is automatic and hence objective. Despite the fact that it is atom-based, the resulting alignment does not depend on an exact overlay of specific atoms. The SEAL method is generally able to produce alignments that agree well with experimental observations,[9] and the resulting alignment scores have been shown to carry important information on 3D properties of molecules that can also be applied to activity prediction.[10]

The SEAL method was initially developed for the alignment of two rigid molecules, in contrast to the usual needs of practical drug design where a number of molecules may need to be overlayed. A flexible variant of the SEAL procedure (TORSEAL) has been developed by Klebe et al.[11] It combines random perturbation and subsequent force-field minimization steps with the overlay process. Its main limitation is the restriction of the conformational and alignment space that is searched.[11] Although this problem has been reduced by the incorporation of a limited conformational search,[11] an initial assumption for the alignment still appears to be necessary. Our aim has been the development of a fully automatic and systematic multiconformational overlay procedure for multiple molecules that makes no assumptions concerning initial alignment. We term our method MultiSEAL. This process introduces no further approximations to the ones in the SEAL method. We hoped that the absence of prealignment would facilitate the identification of possible multiple binding modes.

To test the applicability of MultiSEAL, we applied the method to a number of known alignment examples. In addition, we tested the approach by evaluating the alignments generated for a series of compounds believed to bind to a common site on tubulin. Colchicine is a well-known antimitotic agent that inhibits cell division by interfering with microtubule assembly.[12] Drugs with this effect have been applied in cancer therapy and hence are of great therapeutic interest. There are several known binding sites on tubulin, of which the colchicine binding site is one of the most extensively studied. A number of review papers summarize available information on tubulin binding kinetics, mechanism of action, and properties of colchicine derivatives.[12−15]

It has been established that a structurally diverse set of molecules, among them combretastatin and certain colchicine derivatives, bind to the colchicine site and competitively inhibit colchicine binding to tubulin.[15] A molecular modeling study[16] using the CASE and MULTICASE methods has identified eight possible biophores for colchicine analogues and four for combretastatin. Using this method, two possible orientations of combretastatin with respect to colchicine have been found, indicating possible multiple binding modes.[16]

A natural metabolite of estradiol, 2-methoxyestradiol (2-ME), has been found to inhibit proliferation of endothelial cells in vitro.[17] In vivo testing has demonstrated its antiangiogenic activity and potential utility as a cancer therapy.[17] Like several other estrogenic compounds, including diethylstilbestrol, 2-methoxyestradiol inhibits tubulin polymerization and is a weak competitive inhibitor of colchicine binding to tubulin.[17,18] On the basis of these observations, it has been suggested that 2-methoxyestradiol and related estrogenic compounds act directly on microtubules to affect cell division. Because of the low cytotoxicity of 2-methoxyestra-

10.1021/ci9900682 CCC: $19.00
Published on Web 02/26/2000

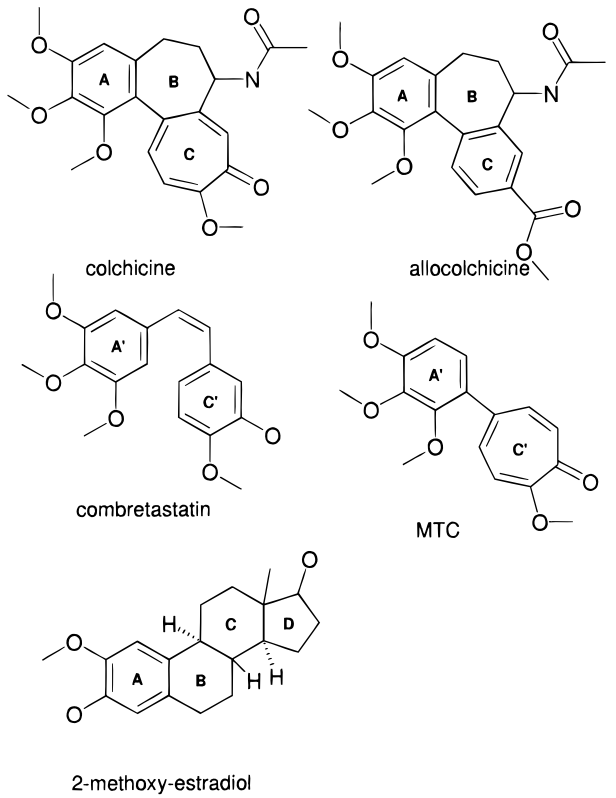

Figure 1. Colchicine analogues studied in this work.

diol, a number of its analogues have been investigated for antimitotic and antiangiogenic activity.[19−21] In the present study, our objective was to use the MultiSEAL procedure to identify common features in a diverse set of ligands, including colchicine analogues and estradiol metabolites. It was hoped that this analysis would shed light on the mechanism of action of these molecules and aid in the identification of additional active analogues. The molecules considered in this work are shown in Figure 1. They include colchicine (COL), allocolchicine (ALLO), combretastatin A-4 (COMB), 2-methoxy-5-(2′,3′,4′-trimethoxyphenyl)-tropone (MTC), and 2-methoxyestradiol, all of which have been shown to inhibit tubulin polymerization and colchicine binding to tubulin.[15,18]

PROCEDURE

In a simplistic way, one could think of the flexible alignment of two molecules as an overlay problem involving a pairwise fitting of conformers from corresponding databases.

In the original SEAL alignment function,[9] the alignment score was defined for a single pair of molecules. To align more than two molecules, it was necessary to derive the corresponding overlap function. The original SEAL similarity score, A_F, is calculated in the following manner:

$$A_F = -\sum_{i=1}^{m}\sum_{j=1}^{n}(w_E q_i q_j + w_S v_i v_j)e^{-\alpha r_{ij}^2} \qquad (1)$$

where m and n are the total numbers of atoms comprising the two molecules, r_{ij} is the distance between the ith and jth atoms located on the two different molecules, α defines the distance dependence, w_E, w_S, etc., are user-defined weighting values for steric, electrostatic, and other properties, and q_i and q_j are the partial charges at the ith and jth atoms. The parameters v_i and v_j are arbitrary powers (the default is 3) of the van der Waals radii of the ith and jth atoms.

Suppose that we have three molecules A, B, and C, containing m, n, and o numbers of atoms, respectively. If we overlay molecule A and B first, then the score is given by

$$A_{F1} = -\sum_{i=1}^{m}\sum_{j=1}^{n}(w_E q_A q_B + w_S v_A v_B)e^{-\alpha r_{AB}^2} \qquad (2)$$

When A and B are overlayed, this entity can be viewed as a supermolecule with $n + m$ atoms that can in turn be overlayed onto molecule C. (In practice, the transformation of the atoms of molecules A and B to the supermolecule needs to be performed first.) Because of the properties of summation, we can break up the score from the second overlay to two components as follows:

$$A_{F2} = -\sum_{i=1}^{m}\sum_{j=1}^{o}(w_E q_A q_C + w_S v_A v_C)e^{-\alpha r_{AC}^2} - \sum_{i=1}^{n}\sum_{j=1}^{o}(w_E q_B q_C + w_S v_B v_C)e^{-\alpha r_{BC}^2} \qquad (3)$$

The sum of the scores from the two overlays contains all possible cross-terms between the three molecules. Hence the total score is characteristic of the entire process; i.e.

$$a_{F,TOT} = A_{F1} + A_{F2} \qquad (4)$$

By inspection of the summation, it can be seen that the process is commutative and associative. It is also easily shown that the same relationships also hold when additional molecules are added to the overlap. Because of these commutative and associative relations, molecules or sets of molecules can be fitted in pairwise fashion in any order, while the score of the overlap will simply be the sum of the individual overlap functions. In MultiSEAL multiple conformations are considered by explicitly considering conformers from a conformational analysis of all studied molecules. The process has a reasonable speed for drug-sized molecules even with the unoptimized code. The consideration of 10 000 pairwise overlays in the case of colchicine−combretastatin required just over 22 h of CPU time on a 195 MHz SGI Octane, or about 51.5 h on a 300 MHz PC with a Pentium II processor running under Windows NT. The process appears to be approximately linearly dependent on the number of heavy atoms. In terms of computer speed, a previously overlapped set of molecules (i.e. a supermolecule) behaves almost identically to a molecule with a number of atoms equal to the sum of those in the individual molecules.

In principle, as many conformations as possible need to be considered. On the other hand, it may be necessary to select a subset of conformations when analyzing large, flexible molecules. This can be achieved by using an appropriate clustering method. For molecules with at most a few thousand conformers, applying MultiSEAL on the cluster representatives is rapid and provides good results. Difficulties arise only when the conformational clusters contain too wide a range of conformer geometries (resulting

from either too many conformers or an inappropriately chosen clustering algorithm). In this case, it is still possible to fine-tune the alignment. This may be achieved by further clustering those conformational clusters that contributed to the best scoring solution, selecting a set of cluster representatives from these and repeating the last step of the overlay with these representatives. In the case of a single molecule with *n* new clusters, this introduces only *n* additional overlay steps. This procedure was tested in the present study for the combretastatin molecule (3630 conformers). It was found that the procedure led to no major improvement in its overlap with colchicine, probably because the number of combretastatin conformations was still sufficiently low. The improvement in the SEAL score was less than 0.5%, and the change in root mean square distance as a result of this step was less than 0.4 Å. Nevertheless, this procedure is expected to be important for molecules that have an even greater number of conformers.

The original SEAL procedure incorporates terms describing steric and electrostatic interactions. It has been proposed[11] that extra terms for physicochemical parameters and hydrophobicity may improve the ranking of alignments. However, the findings in this study indicate that the inclusion of physicochemical factors and hydrophobicity does not lead to any statistically significant improvement in the overlays.[11] (The best rms_{mean} deviation between the experimental X-ray structure and the calculated alignment was 1.21 with the original steric and electrostatic terms, 1.18 with the simultaneous inclusion of hydrophobicity, and 1.19 with an additional refractivity term.[11]) The effect of the hydrophobicity term on the discrimination of the two best scoring solutions in the quoted examples varied and did not change their relative order (see Table 1 of ref 11). These results indicated to us that the inclusion of an extra hydrophobic function would be unlikely to improve the quality of the fits significantly. The current study was conducted using only the steric and electrostatic terms. In contrast, the quality of the overlays critically depends on the magnitude of the terms. These were kept at the optimized values recommended by Klebe et al.[11] ($\alpha = 0.2$, $w_S = 16$, and $w_E = 2$).

Calculations. All calculations in this work were performed using the fully programmable Molecular Operating Environment program suite.[22] Before applying the MultiSEAL process, a conformational analysis was performed using the RIPS procedure (random incremental pulse search),[23] which also identifies ring conformers. In each step of this method, all rotatable bonds are rotated randomly. The atom positions are then perturbed by a fixed amount (±2 Å in this work). During this process, the absolute configuration of each molecule was preserved. The structures were subsequently optimized using the Merck force field (MMFF94[24]). Each resulting conformation was checked to determine if it had already been generated by comparing all atom positions using a predefined rms tolerance on the heavy atoms (0.2 Å in this work). The procedure was terminated when the number of failures to find new conformations exceeded 500 in a row. To ensure that the conformational space was sufficiently explored, the conformational analysis was repeated with the same parameters. Fewer than 10 new conformers were produced in all cases, making it likely that the great majority of the conformers within the 10 kcal/mol window were identified.

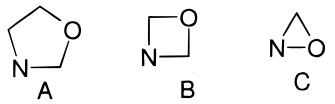

Figure 2. Hypothetical set of molecules used to test the accuracy of the multiple overlay.

Accounting for the relative energy of the conformers is an important issue in conformational analyses. In many cases, only the lowest energy conformer is taken into account.[25] Frequently conformers within an energy window are considered according to their Boltzmann populations,[25] but there is little justification for doing this.[26] First, the lowest energy conformer is not necessarily the biologically active one. Second, there is great difference in relative conformational energies from different force fields, showing that the energies are often highly inaccurate. The absolute accuracy may be even worse, as solvation is generally not taken into account. Furthermore, all low-energy conformations are populated, and these are in dynamic equilibrium with all other conformations. For these reasons, all conformers within the arbitrary energy window were considered as potentially active forms. When considering the size of this window, it should also be born in mind that molecular superposition is an attempt to understand the common binding pattern of ligands. It is the expectation in this process that the aligned ligands will be as similar as possible to the receptor-bound form. However, it is well-documented in the literature[27,28] that ligand conformations in the active site may have free energies by as much as 10 kcal/mol above the lowest free-energy solution structure. Hence a sufficiently large conformational energy window must be defined in order to accommodate the active conformations.

In some examples, clustering was performed to speed the calculations. Conformer clustering was achieved on the basis of the pairwise rms displacement of corresponding heavy atoms after optimal rigid-body superposition, as implemented in MOE.[22] In this approach, clusters are created using the distance matrix, obtained after pairwise superposition and based on the specified rms displacement. This parameter was chosen so that the number of clusters stays below 20. In all cases, the specified rms distance between clusters was between 0.9 and 1.4 Å. This also made sure that the structures within each cluster were not too diverse. The lowest energy conformation was selected from each group as a cluster representative.

Although in the MultiSEAL process the sum of scores from the individual steps characterizes the entire overlay process (see eqs 1−4 above), the individual scores describe the quality of the fit for the given step. Hence in the overlay of colchicine analogues, the 200 best scoring overlays from each step were always kept, to avoid biasing the end results at an intermediate step.

Validation. The multiple overlap procedure was first tested on a hypothetical set of molecules, as shown in Figure 2. The overlay was carried out in the three possible orders. The best scores, as a function of the order of selection, are given in Table 1. This test was repeated five times, and the best alignment, as well as the scores, remained identical. This experiment also indicated the kind of numerical accuracy that can be expected if the sequence of overlays is changed.

The method was next tested using three examples: two angiotensin II receptor antagonists, four 5-HT_3 receptor

Table 1. Results of the Overlay of the Three Hypothetical Molecules from Figure 2

first overlay	$A_{F,1}$	second overlay	$A_{F,2}$	$A_{F,TOT}$
A−B	−69.7739	(A−B)−C	−98.5440	−167.9916
A−C	−52.0824	(A−C)−B	−115.8433	−167.9257
B−C	−47.1693	(B−C)−A	−120.8312	−168.0005

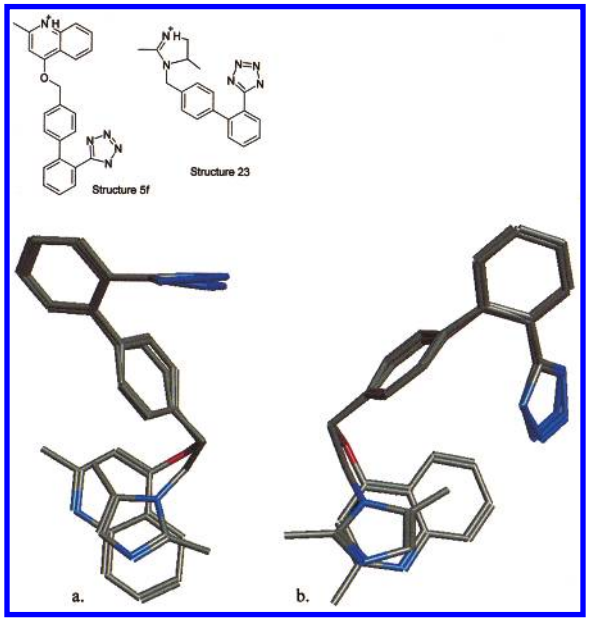

Figure 3. Best alignments of two angiotensin II antagonists. The structure names used are from ref 5.

antagonists, and four flexible dopaminergic compounds. The results were compared with those obtained using a genetic algorithm-based (GA-based) approach.[5]

For the overlay of the two angiotensin II antagonists (see Figure 3) 71 conformers were generated by the RIPS procedure for structure 5f and 95 for structure 23. These conformations were grouped into five clusters. The cluster representatives were then overlaid, with the two best solutions shown in Figure 3. Comparing these with the overlays depicted in Figure 4 in Jones et al.,[5] Figure 3a appears to correspond to the overlay shown in the top left corner, while Figure 3b corresponds to the alignment at the bottom right. Overlays similar to the other two solutions with GA[5] were also generated in this work, although their score was substantially (over 10%) lower than the highest scoring solution. Although this was a relatively simple problem that could have been handled directly by considering all conformations, this result helped to show that the method can be made more efficient by using conformational clustering and that alignment results could be obtained in a matter of minutes.

The four 5-HT_3 receptor antagonists used as a second example are shown in Figure 4.[29] The RIPS conformational analysis generated 19 conformers for structure 37, 915 for 44, 4066 for 45, and 420 for 47. These conformations were then clustered with the criterion that two conformations were considered different if the maximum heavy atom rms of their appropriate atoms was over 0.8 Å. The molecules were overlaid sequentially, and the best 100 overlays were used for the next fit. From the solutions in the last step, the top scoring 50 overlays contained about the same number of two basic types of overlap patterns. These are shown in Figure 4a,b. In both cases the common points shared among the

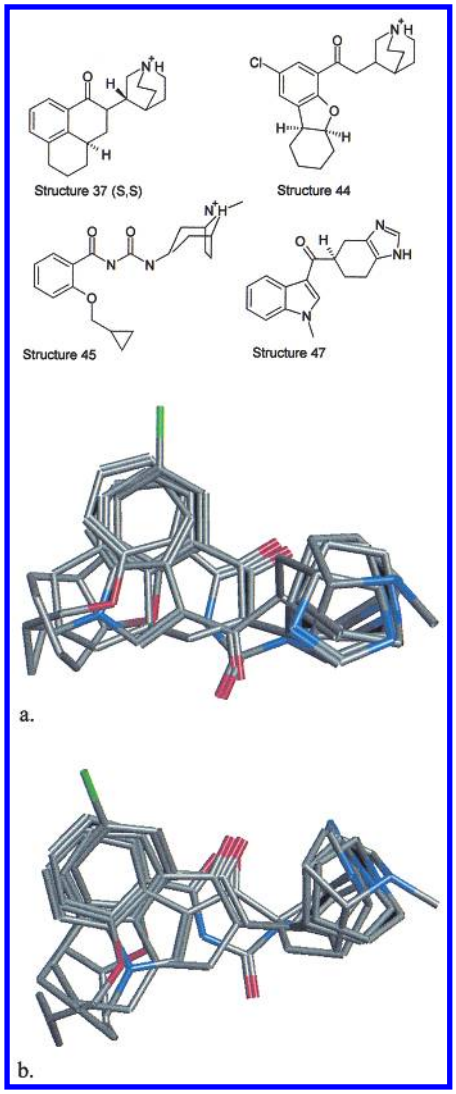

Figure 4. Two highest scoring overlay types from the alignment of four 5-HT_3 receptor antagonists. The structure names used are from ref 5.

four molecules are the aromatic rings and the sp^2 oxygens. Nitrogen atoms that can act as hydrogen bond donors are also located close to each other, similarly to previous findings.[5] The essential difference between the overlay patterns shown in Figure 4a,b lies in the way the sp^2 oxygens are aligned. While in Figure 4b the sp^2 oxygens of all four molecules are overlaid and the second sp^2 oxygen of structure 45 remains unpaired, the alignment depicted in Figure 4a displays a different orientation of structure 47 in which its sp^2 oxygen overlays with the second amide oxygen of molecule 45. Hence the two sets of solutions can be viewed as corresponding to two possible binding orientations of structure 47, while the other essential properties of the fits are approximately the same. The differentiation between the two overlay types is marginal: the best 50 overlays correspond to a score spread of less than about 5%. In essence, the second overlay (Figure 4b) is qualitatively similar to that obtained with GA alignments,[5] except for the torsional angle at the central amide bond in structure 45. Crystallographic data indicate[30] that the central amide has a cis configuration, whereas the GA overlay could only consider the trans form. In this work, the top scoring overlays only contained the cis central amide bond, in accordance with the crystal structure

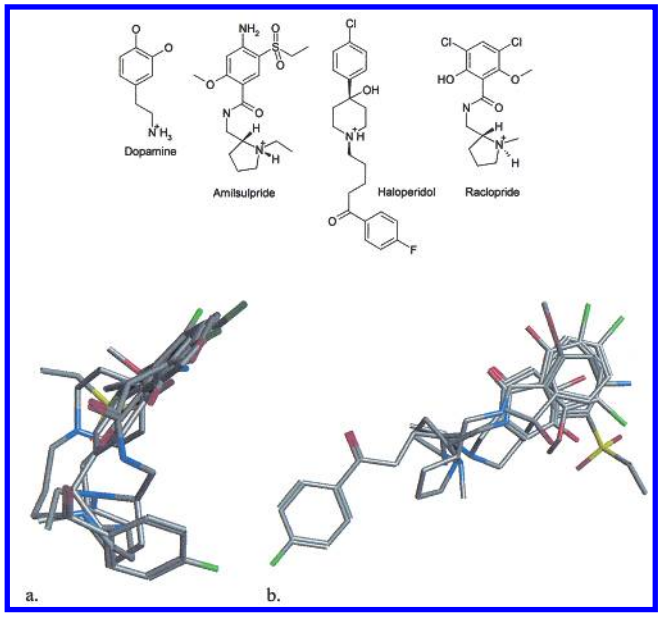

Figure 5. Two high-scoring alignments from the overlay of four dopaminergic compounds. The structure names used are from ref 5.

of the unbound ligand. It must be added that the MultiSEAL method allows the screening of the solutions for specific bonding characteristics, so if necessary, solutions with specific geometric features, such as a trans amide bond, can also be selectively searched and identified.

As a final comparative example, the overlay of the four dopaminergic compounds shown in Figure 5 was studied. Because of the flexibility of these ligands and the issues concerning proper coverage of conformational space, this task presented some difficulties to both GA[5] and MultiSEAL. To speed up the conformational search, the fast bond rotation search algorithm of MOE was applied.[22] This approach identifies the rotatable bonds, locates all local minima with respect to single bond rotations, and identifies unique conformers. This method ignores ring conformations. The only flexible ring, the saturated piperidine ring of haloperidol, was considered in a chair conformation. Although this search method is fast (it required less than 10 min/molecule on a 195 MHz SGI Octane), its major disadvantage is that it does not ensure full conformational coverage. With this method we obtained 6 dopamine, 94 amilsulpride, 332 haloperidol, and 29 raclopride conformations. As before, the first 100 solutions from each fit were used at the next step of the fit. The top scoring solutions belonged to two major classes. Due to the flexibility of the molecules, dozens of solutions for both classes have been found. The alignment displayed in Figure 5a had the highest score. The high score arises from a sterically very compact fit, rather than from an overlay of pharmacophoric groups. Although the score for the alignment displayed in Figure 5b is about 8% lower, it shows the alignment of all four protonated nitrogens and the aromatic rings. The three sp^3 donor oxygens are also aligned. Hence this alignment is essentially similar to the best solution obtained with the GA technique.[5] It appears likely that the higher score for the sterically compact solution in the present work arises as a result of the high steric weight (w_S in eq 3 is 8 times greater than w_E). This explanation was indeed proven in this work for pairwise overlays: on increasing the weight of w_E in the haloperidol−raclopride fit, the relative score obtained from the proper overlap of pharmacophores increased. Although the value of these constants may not be ideal for the specific case of dopaminergic compounds, it was judged that the optimization of α, w_E, and w_S in each project would ruin the objectivity of the method. Instead, the procedure followed in this work was to apply only the optimized values of these constants[11] for reasons of reproducibility, establish families or clusters among the highly scoring solutions, and place less emphasis on the actual value of the overlap score.

Table 2. Total Scores and Relative Conformational Energies of the First 20 Best Fits from the MultiSEAL Procedure[a]

	total score	$E_{rel,colch}$	$E_{rel,allo}$	$E_{rel,comb}$	$E_{rel,mtc}$	orient[b]
1	−1780.3	5.37	3.92	7.29	0.00	a
2	−1733.5	8.53	6.23	7.29	0.67	a
3	−1728.1	5.37	3.92	7.29	0.59	a
4	−1727.0	5.37	3.92	4.13	0.00	a
5	−1725.6	8.53	6.23	5.95	0.67	b
6	−1721.5	5.37	9.12	2.80	0.00	a
7	−1717.2	8.53	1.23	5.95	0.67	a
8	−1716.0	5.22	3.92	7.29	0.00	a
9	−1710.0	8.53	6.23	5.73	0.67	a
10	−1704.3	8.53	6.23	3.52	0.67	b
11	−1699.5	5.37	6.58	7.29	0.00	a
12	−1699.4	8.53	6.23	0.45	0.67	b
13	−1698.2	5.37	3.92	2.80	0.00	a
14	−1698.0	5.37	3.92	0.00	0.00	a
15	−1695.4	5.37	3.92	4.13	0.59	a
16	−1693.8	5.19	3.12	7.29	0.00	a
17	−1692.5	5.37	9.12	2.80	0.59	a
18	−1691.5	6.94	5.65	5.95	1.62	b
19	−1691.2	6.94	5.65	2.35	1.62	a
20	−1687.2	8.53	1.23	7.29	0.67	a

[a] Conformational energies from obtained by molecular mechanics in vacuo using the MMFF94 force field. See text for further details. [b] Orientation of combretastatin with respect to the other three molecules, as depicted in Figure 7.

RESULTS AND DISCUSSION

The conformational analysis of colchicine and its analogues using RIPS generated a large number of conformers (866 for COL, 1382 for ALLO, 3630 for COMB, and 232 for MTC). A total of 12 conformers was generated for 2-methoxyestradiol. The four molecules suspected to interact most similarly with the colchicine binding site (COL, ALLO, MTC, and COMB) were fitted first. The results are summarized in Table 2. The best scoring overlap is shown in Figure 6. Close inspection of the best 25 overlaps reveals that there are two basic patterns. In both, colchicine, allocolchicine, and MTC always overlap in a similar fashion, while combretastatin can assume two different orientations. The molecules COMB and COL were extracted from these overlays and are shown in Figure 7. In the majority of the best fits (18 out of the best 25, including the best scoring fit), the rings marked as A or A′ are aligned with each other. In the other pattern of orientation the C′ ring of COMB is aligned with the A or A′ rings of the other three molecules. These two orientations of combretastatin may correspond to multiple binding modes in the binding process, although the first would be more likely to occur. The overlap of combretastatin is the poorest among the four molecules, with the rms distance between its functional groups and those of colchicine 2−3 times greater than the corresponding distances in the case of ALLO and MTC. Nonetheless, even for combretastatin, these distances are below 1.2 Å.

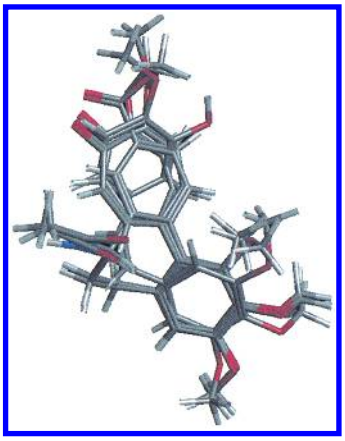

Figure 6. Best overlap of colchicine, allocolchicine, combretastatin, and MTC.

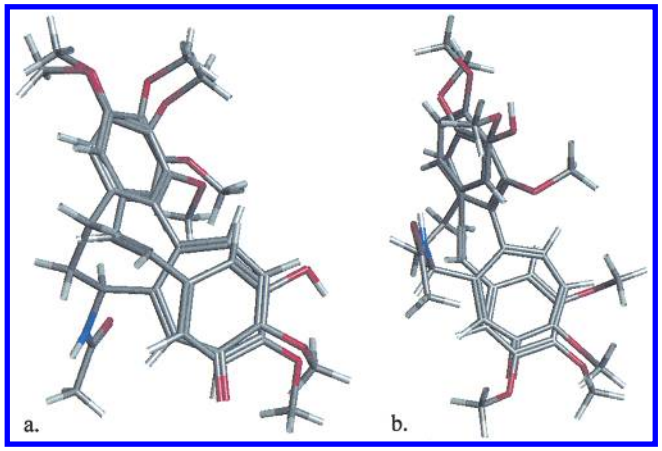

Figure 7. Two orientations of combretastatin with respect to colchicine. Orientation a in this figure depicts overlap 1 in Table 2, while orientation b shows overlap 5.

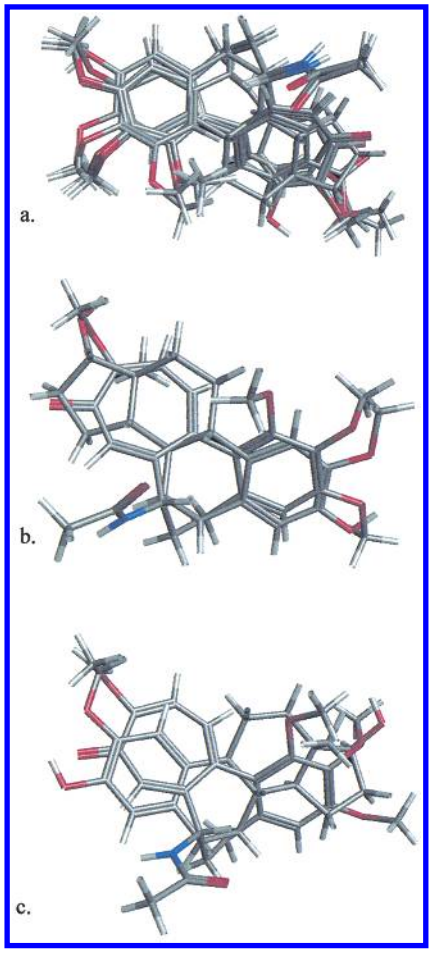

Figure 8. Best overlap of 2-methoxy estradiol with (a) the other four molecules, (b) colchicine with the A rings overlapping, and (c) colchicine with the A and C rings overlapping.

The two overlays in Figure 7 can be compared to the results of a pharmacophore search on colchicine, combretastatin, and podophyllotoxin.[16] In that work the two possible overlays (i.e. ring overlaps A to A′ and A to C′) of colchicine and combretastatin were also identified. However, in these pharmacophore overlays the hydrophobic surfaces of the molecules were not aligned. The pharmacophore overlay shown in Figure 6b of Haar et al.[16] can be directly compared to the overlap depicted in Figure 7a of the present work. Whereas MultiSEAL aligned both aromatic rings of the two molecules, the pharmacophore overlay[16] aligned only one ring at a time.

A unique advantage of MultiSEAL, compared to other flexible alignment methods, is that it can unambiguously identify the conformers contributing to a given overlay and their relative energies (with the caveat regarding force field energies, discussed previously). This helps to avoid fits involving high-energy conformations, as may happen, for example, when the Catalyst approach is used.[30] The relative energies of the conformers used in the present study are given in Table 2. In the best fit, the conformational energy contribution of combretastatin is 7.3 kcal/mol. The remainder of the total energy is distributed among the other molecules. Combretastatin is the most flexible of the molecules in the set, and the inaccuracies due to estimated absolute force field energies and the neglect of solvation terms may play a more significant role in its case. The higher conformational energy itself is probably a good indicator of the decreased ability of COMB to overlap well with the other molecules. Nonetheless, it is important to mention that no meaningful rigid overlap was possible using only the lowest energy conformers of the four molecules within the given model.

Once it was established that these four molecules could be aligned, they were overlaid on 2-methoxyestradiol. The best-scoring result is shown in Figure 8a. Parts b and c of Figure 8 illustrate the two possible overlaps of colchicine and 2-methoxyestradiol. Figure 8b was obtained by only depicting colchicine and 2-ME from Figure 8a. In this overlay it is the A rings of the two molecules that overlap each other. The result of the best direct fit of the two molecules is shown in Figure 8c, in which the C ring of colchicine is seen to overlap with the A ring of 2-ME. The terms used by the Multiseal procedure (see eqs 1−4 above) to generate overlaps incorporate information about all five molecules simultaneously and produce the best overall fit of any two structures with respect to all others in the set. In contrast, the overlap shown in Figure 8c is based on only two molecules, colchicine and 2-ME, and neglect information concerning other structures in the data set. In either case the overlay of 2-methoxyestradiol is considerably worse than that of the other molecules under consideration. Corresponding functional groups are typically displaced by 1−2 Å. Hence it is unlikely that 2-methoxyestradiol fits the entire colchicine binding site well, although interaction with a portion of the binding site is not precluded. A better overlay might be found

by relaxing the conformer energy cutoff, but would be unreasonable due to excessive distortion of the ligand.

It has been shown experimentally that both colchicine and 2-methoxyestradiol inhibit tubulin polymerization (IC_{50} for colchicine 0.8 ± 0.07 μM, for 2-ME 1.9 ± 0.2 μM [18]). It has also been established that 2-ME is a weak competitive inhibitor (50% inhibition is achieved only with a 2-ME concentration of 40 μM to displace 2 μM colchicine [18]). The results obtained with MultiSEAL provide some explanation for these observations. The pattern of substitutions on the A ring of 2-ME is similar to the substitutions on the A ring of colchicine, and this may be sufficient to allow it to partially interact with the colchicine binding site. This may be a crucial factor for inhibition of tubulin polymerization by both compounds. In the case of estradiol, in which the 2-methoxy group is missing, the A ring cannot overlap properly with the other structures. As a result, estradiol is expected to be a poor inhibitor of tubulin polymerization (IC_{50} 30 ± 6 μM [18]). Higher affinity binding to the colchicine site is likely to require the overlap of functional groups at both ends of the molecule. The poor overlay of 2-ME on colchicine indicates that 2-ME is likely to be a weak competitive inhibitor. Since estradiol fits the binding pocket even more poorly, it is not surprising that it is an even weaker competitive inhibitor of colchicine binding than 2-ME (30% at a concentration of 100 μM [18]). The poor ability of other compounds, such as estrone, estriol, 4-methoxyestradiol, estradiol 3-*O*-methyl ether, and 4-methoxyestradiol 3-*O*-methyl ether, to bind to the colchicine site[18] can be understood using the same arguments based on partial and full overlap with colchicine and its analogues.

CONCLUSIONS

A robust and automated procedure has been described that allows the direct application of the SEAL method to multiple conformations of multiple molecules. With this modification, MultiSEAL is a potentially useful tool for prealignment in 3D QSAR studies. It can also be used to locate potential pharmacophoric groups in a set of molecules. Because of the moderate speed of the present implementation, it is directly applicable for the overlay of molecules that have at most a few thousand conformations.

Ideally, all conformations of every molecule should be considered and all the overlays taken into account to obtain all possible binding modes and mutual orientations. In this case, the limitations of MultiSEAL are those of the SEAL method itself, i.e., the consideration of only atom-based properties and partial charges. A less limited extension of the approach using pharmacophore groups, similar to the recent pairwise overlay method of Miller et al.,[31] is presently being considered. In addition, it may be necessary to include further terms or to modify the weight of the existing ones in the SEAL overlap expression. This, however, is quite difficult, as the results will be critically sensitive to the selection of good examples. Although the bound conformations, as determined by X-ray crystallography, are probably the best bases for verification, they may also introduce unwanted bias into an alignment model. It is realistic to consider the set of aligned molecules as describing the ligand−receptor recognition event just before the actual binding process itself starts. In contrast, the X-ray crystal structure shows the bound state of the complex in a privileged conformation, one that may be favored in the crystallization process. It appears that instead of trying to improve the discrimination of the preferred solutions on a case-by-case basis, a better solution is to put less emphasis on the actual scores and consider alternative solutions by clustering the obtained overlays. This approach should eliminate any bias introduced by the scoring functions.

For practical reasons, it is necessary to reduce the number of conformers considered. This makes it necessary to apply appropriate clustering algorithms to select representative conformations. Provided the clustering method is chosen sensibly, this will not introduce much bias into the results. Another bias is introduced when the best overlays are selected and fitted to the next molecule. Since it is the sum of scores of all fits that describes the quality of the overall process, there is no guarantee that an inferior fit at an intermediate step will be unable to contribute to a good overall fit at the end. However, the examples studied so far indicate that in practice the best overall fits always involve good quality intermediate fits.

During the multiconformational alignments we observed that the solutions in all cases fall into families or clusters that look essentially similar. In other words, a number of different conformations can be selected without changing the overall appearance of the fit. It is interesting to consider this result in terms of the possibility that a number of similar conformations can bind similarly to a receptor. This would contrast to the generally accepted view that molecules bind to the receptor in a specific (usually the lowest energy) conformation. The latter view is reinforced by results from X-ray crystallography, where a single conformation is usually observed. To find out more about the recognition event, it would be invaluable to have direct experimental information about the ligand structures during the binding process itself in a real time-resolved fashion in the solution phase.

ACKNOWLEDGMENT

Dr. Chris Williams at the Chemical Computing Group Inc. (Montreal, Canada) is gratefully acknowledged for his help in coding the program modifications.

REFERENCES AND NOTES

(1) Klebe, G. Structural Alignment of Molecules; In *3D QSAR in Drug Design*; Kubinyi, H., Ed.; ESCOM: Leiden, The Netherlands, 1993; pp 173−197.
(2) Martin, Y. C. 3D QSAR: Current State, Scope, and Limitations. *Perspect. Drug Discovery Des.* **1998**, *12/13/14*, 3−23.
(3) Martin, Y. C.; Bures, M.; Dahaner, E.; DeLazzer, J.; Lico, I.; Pavlik, P. A Fast New Approach to Pharmacophore Mapping and Its Application to Dopaminergic and Benzodiazepine Agonists. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 83−102.
(4) Lemmen, C.; Lengauer, T. Time-Efficient Flexible Superposition of Medium-Sized Molecules. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 357−368.
(5) Jones, G.; Willett, P.; Glen, R. C. A Genetic Algorithm for Flexible Molecular Overlay and Pharmacophore Elucidation. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 532−549.
(6) Masek, B. B.; Merchant, A.; Matthew, J. B. Molecular Shape Comparison of Angiotensin II Receptor Antagonists. *J. Med. Chem.* **1993**, *36*, 1230−1238.
(7) Nissink, J. W.; Verdonk, M. L.; Kroon, J.; Mietzner, T.; Klebe, G. Superposition of Molecules: Electron Density Fitting by Application of Fourier Transforms. *J. Comput. Chem.* **1997**, *18*, 638−645.
(8) Lemmen, C.; Lengauer, T.; Klebe, G. FlexS: A Method for Fast Flexible Ligand Superposition. *J. Med. Chem*. **1984**, *41*, 4502−4520.

(9) Kearsley, S. K.; Smith, G. M. An Alternative Method for the Alignment of Molecular Structures: Maximizing Electrostatic and Steric Overlap. *Tetrahedron Comput. Methods* **1990**, *3*, 615−633.
(10) Kubinyi, H.; Hamprecht, F. A.; Mietzner, T. Three-Dimensional Quantitative Structure−Activity Relationships (3D QSiAR) from SEAL Similarity Matrices. *J. Med. Chem.* **1998**, *41*, 2553−2564.
(11) Klebe, G.; Mietzner, T.; Weber, F. Different Approaches Toward an Automatic Structural Alignment of Drug Molecules: Applications to Sterol Mimics, Thrombin and Thermolysin Inhibitors. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 751−778.
(12) Engelborghs, Y. General Features of the Recognition by Tubulin of Colchicine and Related Compounds. *Eur. Biophys. J.* **1998**, *27*, 437−445.
(13) Brossi, A.; Yeh, H. J. C.; Chrzanowska, M.; Wolff, J.; Hamel, E.; Lin, C. M.; Quin, F.; Suffness, M.; Silverton, J. Colchicine and Its Analogues: Recent Findings. *Med. Res. Rev.* **1988**, *8*, 77−94.
(14) Hamel, E. Interactions of Tubulin with Small Ligands. In *Microtubule Proteins*; Avila, J., Ed.; CRC Press: Boca Raton, FL, 1990; pp 89−192.
(15) Hastie, S. B. Interactions of Colchicine with Tubulin. *Pharm. Ther.* **1991**, *51*, 377−401.
(16) Haar, E.; Rosenkranz, H. S.; Hamel, E.; Day, B. W. Computational and Molecular Modelling Evaluation of the Structural Basis for Tubulin Polymerization Inhibition by Colchicine Site Agents. *Bioorg. Med. Chem.* **1996**, *4*, 1659−1671.
(17) Fotsis, T. Zhang, Y.; Pepper, M. S.; Adlercreutz, H.; Montesano, R.; Nawroth, P. P.; Schweigerer, L. The Endogeneous Oestrogen Metabolite 2-Methoxyoestradiol Inhibits Angiogenesis and Suppresses Tumor Growth. *Nature* **1994**, *368*, 237−239.
(18) D'Amato, R. J.; Lin, C. M.; Flynn, E.; Folkman, J.; Hamel, E. 2-Methoxyestradiol, an Endogeneous Mammalian Metabolite, Inhibits Tubulin Polymerization by Interacting at the Colchicine Site. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3964−3968.
(19) Cushman, M.; He, H.; Katzenellenbogen, J. A.; Lin, C. M.; Hamel, E. Synthesis, Antitubulin and Antimitotic Activity, and Cytotoxicity of Analogues of 2-Methoxyestradiol, an Endogeneous Mammalian Metabolite of Estradiol That Inhibits Tubulin Polymerization by Binding to the Colchicine Binding Site. *J. Med. Chem.* **1995**, *38*, 2041−2049.
(20) Cushman, M.; He, H.; Katzenellenbogen, J. A.; Varma, R. K.; Hamel, E.; Lin, C. M.; Ram, S.; Sachdeva, Y. P. Synthesis of Analogues of 2-Methoxyestradiol with Enhanced Inhibitory Effects on Tubulin Polymerization and Cancer Cell Growth. *J. Med. Chem.* **1997**, *40*, 2323−2334.
(21) Miller, T. A.; Bulman, A. L.; Thompson, C. D.; Garst, M. E.; Macdonald, T. L. Synthesis and Structure−Activity Profiles of A-Homoestranes, the Estratropones. *J. Med. Chem.* **1997**, *40*, 3836−3841.
(22) *Molecular Operating Environment*, Version 1998.10; Chemical Computing Group Inc.: Montreal, Quebec, Canada, 1998.
(23) Ferguson, D. M.; Raber, D. J. A New Approach to Probing Conformational Space with Molecular Mechanics: Random Incremental Pulse Search. *J. Am. Chem. Soc.* **1989**, *111*, 4371−4378.
(24) Halgren, T. A. Merck Molecular Force Field. I. Basis, Form, Scope, Parametrization and Performance of MMFF94. *J. Comput. Chem.* **1996**, *17*, 490−519.
(25) Kim, K. H.; Greco, G.; Novellino, E. A Critical Review of Recent CoMFA Applications. *Perspect. Drug Discovery Des.* **1998**, *12/13/14*, 257−315.
(26) Yliniemela, A.; Konschin, H.; Neagu, C.; Pajunen, A.; Hase, T.; Brunow, G.; Teleman, O. Design and Synthesis of a Transition State Analogue for the Ene Reaction between Maleimide and 1-Alkanes. *J. Am. Chem. Soc.* **1995**, *117*, 5120−5126.
(27) Nicklaus, M. C.; Wang, S.; Driscoll, J. S.; Milne, G. W. A. Conformational Changes of Small Molecules Binding to Proteins. *Biorg. Med. Chem.* **1995**, *3*, 411−428.
(28) Vieth, M.; Hirst, J. D.; Brooks III, C. L. Do Active Site Conformations of Small Ligands Correspond to Low Free-Energy Solution Structures? *J. Comput.-Aided Mol. Des.* **1998**, *12*, 563−572.
(29) Clark, R. D.; Miller, A. B.; Berger, J.; Repke, D. B.; Weinhardt, K. K.; Kowalczyk, B. A.; Eglen, R. M.; Bonhaus, D. W.; Lee, C.-H.; Michel, A. D.; Smith, W. L.; Wong, E. H. F. 2-(Quinuclidin-3-yl)-pyrido[4,3-b]indol-1-ones and Isoquinolin-1-ones. Potent Conformationally Restricted 5-HT_3 Receptor Antagonists. *J. Med. Chem.* **1993**, *36*, 2645−2657.
(30) Bradley, G.; Ward, T. J.; White, J. C.; Coleman, J.; Taylor, A.; Rhodes, K. F. Novel Antagonists of the 5-HT_3 Receptor. Synthesis and Structure−Activity Relationships of (2-Alkoxybenzoyl)ureas. *J. Med. Chem.* **1992**, *25*, 1515−1520.
(31) Hopfinger, A. J.; Tokarski, J. S. Three-Dimensional Quantitative Structure-Activity Relationship Analysis. In *Practical Application of Computer-Aided Drug Design*; Charifson, P. S., Ed.; Dekker: New York, 1997; pp 105−164.
(32) Miller, M. D.; Sheridan, R. P.; Kearsley, S. K. SQ: A Program for Rapidly Producing Pharmacophorically Relevant Molecular Superpositions. *J. Med. Chem.* **1999**, *42*, 1505−1514.

CI9900682