Journal of Computer-Aided Molecular Design, 9 (1995) 457–462
ESCOM

J-CAMD 301

The atom assignment problem in automated de novo drug design. 5. Tests for envelope-directed fragment placement based on molecular similarity

M.T. Barakat and P.M. Dean*

Drug Design Group, Department of Pharmacology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QJ, U.K.

Received 1 February 1995
Accepted 23 May 1995

Keywords: De novo drug design; Molecular electrostatic potential; Molecular complementarity; Molecular similarity

Summary

The fragment placement method has been successfully extended to the problem of envelope-directed design. The atom assignment paradigm was based on molecular similarity between two molecular structures. A composite supersurface is defined to form the surface onto which the molecular fields are projected. The assignment process is then determined by using molecular similarity in the objective function to be optimized. In principle, this procedure is closely similar to that outlined in the previous paper for site-directed design. The rationale has been extensively tested on two benzodiazepine antagonists believed to bind to the same site.

Introduction

In computer-aided drug design, very often no information is available about the site. In that circumstance, the rational development of new drugs can only take place by considering the properties of known ligands that are selective for the site and mimicking their properties in an attempt to produce new leads. The determination and use of molecular similarity is currently of intense interest in drug design [1]. The main difficulties of design, based on similarity, are associated with recognizing the important common properties of several molecules from a structurally dissimilar set. The structures must be superposed in their optimal orientations, taking into account their individual flexibilities. If a superposition can be obtained, the molecules may be compared by CoMFA to determine regions of similarity in molecular fields [2,3]. A method for producing different superpositions from flexible structures has been devised [4]. This procedure, coupled with CoMFA, should be useful for identifying pharmacophoric groups, and therefore should provide a basis for de novo envelope-directed design and lead optimization. An optimistic scenario of this type must be tempered by the fact that numerous alternative superpositions may be possible with comparable similarity values. This problem, depending on the user's viewpoint, either plagues the use of CoMFA or provides a fascinating wealth of options.

Previous research in this series of papers [5–8] has led to the development of a method of atom assignment onto 3D molecular graphs. The assignment is directed by molecular fields projected from the site onto the molecular graphs. An optimized assignment is made by a combinatorial placement of fragments on the graphs using simulated annealing [7,8]. Rigorous and semi-rigorous tests with known problems suggest that the algorithm performs well in creating structures with optimal complementarity to a target potential. The objective in this paper is to extend the work to envelope-directed de novo design and make use of molecular similarity to direct the atom assignment.

Here we focus attention only on the atom assignment step and attempt to test the procedure with two competitive benzodiazepine antagonists. We take the two superposed structures, construct the enveloping supersurface and use the potential of one molecule to create an atom assignment on the 3D molecular graph of the other molecule, such that the new potential is optimally similar to the target potential. The creation of novel 3D molecular

*To whom correspondence should be addressed.

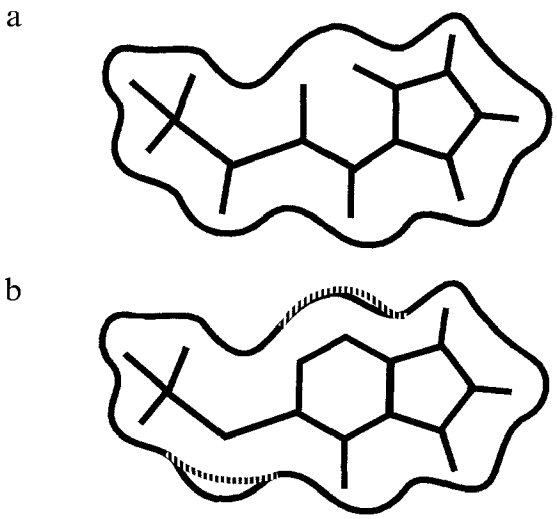

Fig. 1. Diagrammatic representation of the 'supersurface' used in fragment placement for envelope-directed drug design. (a) The lead structure with its van der Waals surface. (b) The molecular graph generated within the surface of the lead (heavy line). Where the standardized surface of the molecular graph itself differs from that of the lead, it is shown in a broken line. The supersurface is taken as the outermost of the two surfaces.

graphs is not the subject of this paper; there are a large number of different procedures available for this step, which have been reviewed recently by Lewis and Leach [9]. This strategy allows us to determine how closely the technique will re-create an original structure from a similarity paradigm. Unlike the self-placement procedure, this is not a perfect test of assignment but it is equivalent to the site-directed placement tests elaborated in the previous paper. The method should be capable of providing some interesting alternative atom assignments, with closely similar field properties, on the same molecular graph.

Computing methods

In paper 4 of this series [8], atom assignment by fragment placement onto a molecular graph, using molecular complementarity, was applied to the case of site-directed drug design, where the potential (electrostatic or hydrophobic) from the site was projected onto the surface of the graph. Fragment placement proceeded according to the optimization of the complementarity (for electrostatic potentials) or similarity (for hydrophobic potentials) of the potentials from the site and the new putative ligand. The model can be extended to the problem of envelope-directed drug design by maximizing the similarity in the potential function between two ligands, instead of optimizing the complementarity of the electrostatic potential function between a ligand and its site. This is achieved by certain sign changes in the annealing algorithm. Thus, the optimization procedure is almost identical to that described earlier [7] and does not need further description.

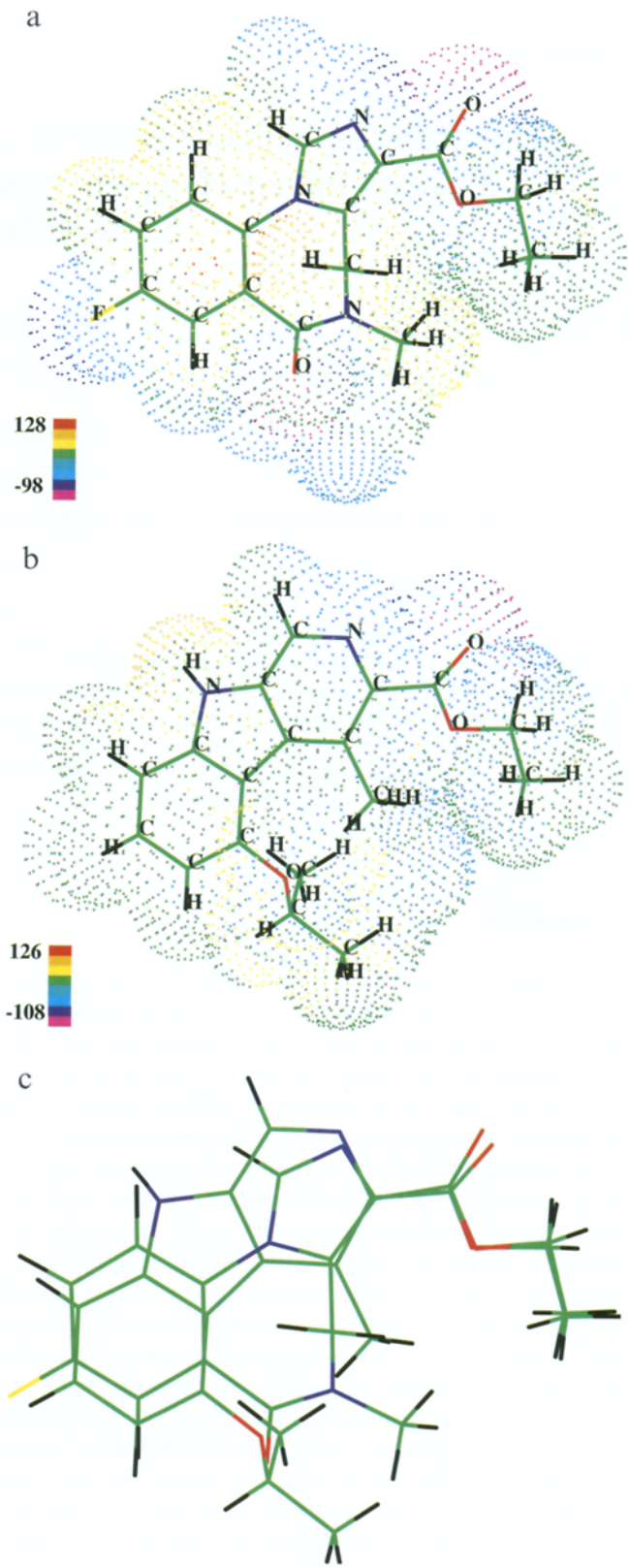

Fig. 2. Two benzodiazepine receptor antagonists Ro 15-1788 (a) and ZK 93426 (b), in their optimal superposition orientations. The electrostatic potential of each is displayed on the supersurface (values in kJ mol^{-1}). A tessellation frequency of 5 is used, giving 3563 points. (c) An overlay of the two structures.

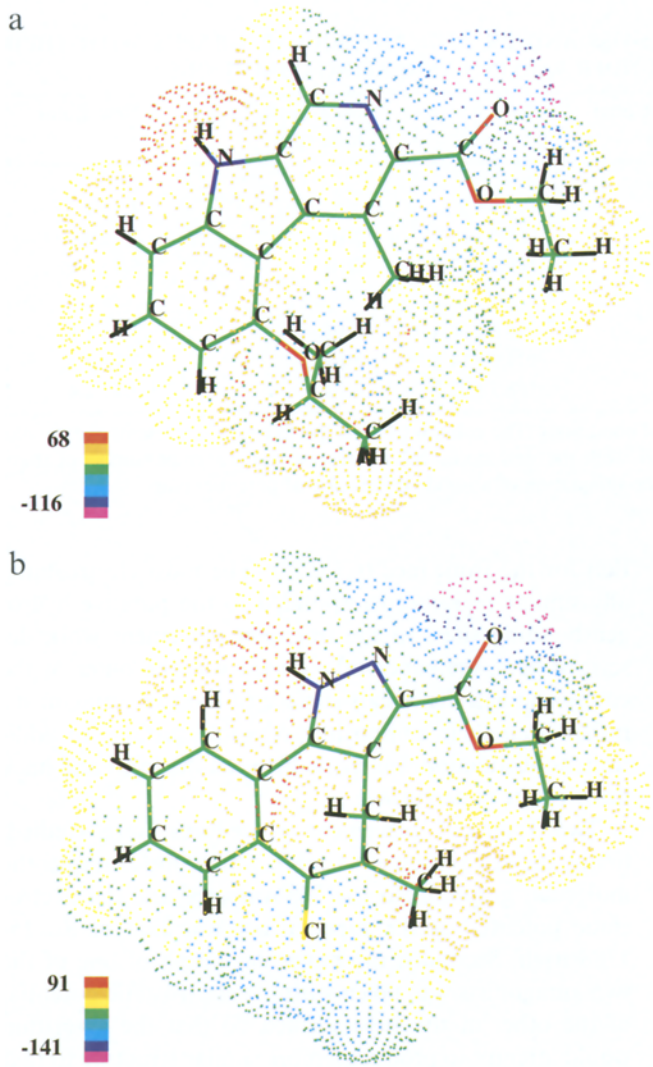

Fig. 3. (a) The structure of ZK 93426 with its electrostatic potential projected onto the supersurface. (b) A structure that is electrostatically similar to ZK 93426 (pyrazole derivative of Ro 15-1788), which was produced by fragment placement onto the molecular graph of Ro 15-1788. The electrostatic potential is displayed on the supersurface (values in kJ mol^{-1}). A tessellation frequency of 5 is used, giving 3395 points.

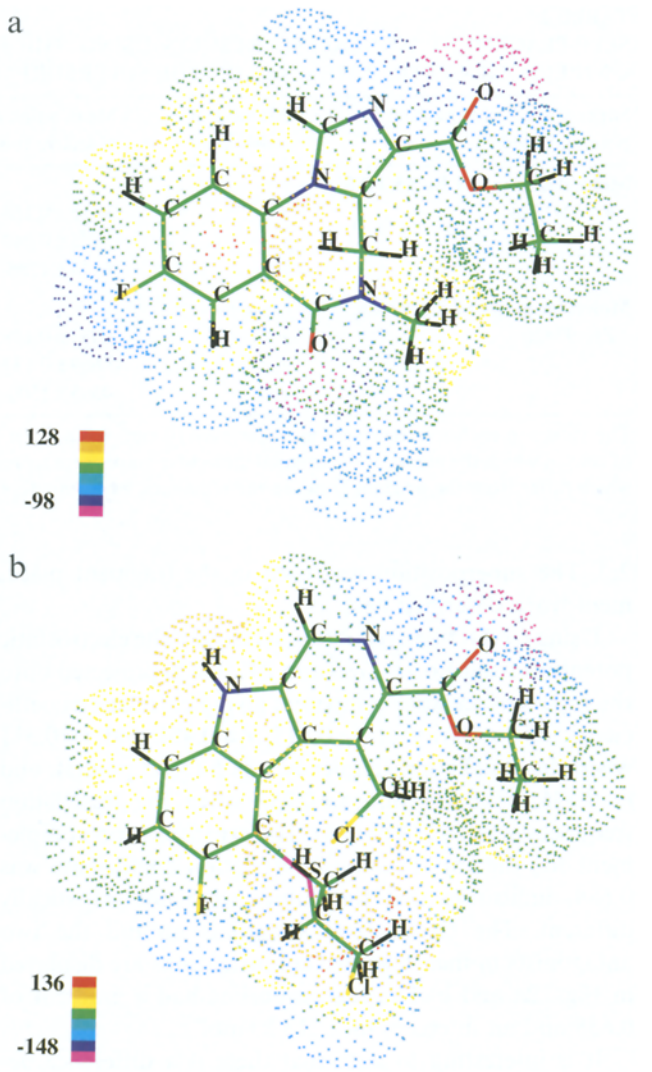

Fig. 4. (a) The structure of Ro 15-1788 with its electrostatic potential projected onto the supersurface. (b) A structure that is more electrostatically similar to Ro 15-1788 than ZK 93426, which was produced by fragment placement onto the molecular graph of ZK 93426. The electrostatic potential is displayed on the supersurface (values in kJ mol^{-1}). A tessellation frequency of 5 is used, giving 3661 points.

One difference between this approach and the site-directed method of fragment placement [8] concerns the definition of the surface around the molecular graph. In envelope-directed atom assignment, the surface must be a 'supersurface' which is the outer surface of both the lead and the target molecular graph (see Fig. 1). This prevents the surface points from falling within the van der Waals radius of any atom in the lead or within a standardized radius of any vertex in the molecular graph, which would result in spuriously high potential values at those points. Also, the pair-*sum* method is modified to become the minimization of pair *differences*. The hydrogen-bonding options are suitably modified to create donor or acceptor similarity at designated vertices, rather than complementarity.

Testing the fragment placement procedure

To investigate the performance of the algorithm when dealing with envelope-directed drug design, two competitive benzodiazepine receptor antagonists have been chosen, the benzodiazepine Ro 15-1788 and the β-carboline ZK 93426. The structure of Ro 15-1788 was taken from its crystal [10], while that of ZK 93426 was obtained from the modified structure of another β-carboline [11], followed by optimization using MM2 [12]. Low-energy conformers of each antagonist were determined by Borea et al. [13]. Pairs of conformers were matched combinatorially by atom positional matching with null correspondences [14]. A similar match, taking into account flexibility by using a combination of simulated annealing and cluster analysis [4], gave an identical superposition (Fig.

TABLE 1
SELF-PLACEMENT ANNEALING RESULTS USING THE Ro 15-1788 AND ZK 93426 MOLECULAR GRAPHS WITH THEIR OWN ELECTROSTATIC POTENTIAL ON THE SUPERSURFACE UNDER VARYING ANNEALING CONDITIONS

Super-surface electrostatic potential used	Method of assessment	Mean value of the final objective function	Time (s)	Ratio perfect runs/ total runs	New atoms
Molecular graph of Ro 15-1788					
Ro 15-1788	r	0.815±0.168	72.9±7.2	0/10	3.3±1.6
	r_s	0.845±0.040	88.5±15.9	0/10	2.7±0.9
	eo	6144±1788	109.8±17.7	0/10	5.4±1.6
Molecular graph of ZK 93426					
ZK 93426	r	0.837±0.158	96.6±11.3	1/10	2.3±1.5
	r_s	0.669±0.137	134.9±14.2	0/10	3.7±2.3
	eo	4595±1500	168.7±21.7	0/10	2.7±2.4

The values in each row represent averages over 10 runs for a given set of conditions. The column labelled 'perfect runs' represents the number of runs in which the atoms of the original molecular graph were reproduced, while the 'new atoms' column represents the average number of atoms which differ from the original graph in the remaining runs. *eo* refers to the summation of absolute differences of potential pairs.

2c). The superposition was used in the fragment placement work described here.

Using CNDO-derived residual charges, the electrostatic potentials from Ro 15-1788 and ZK 93426 projected onto the supersurface were compared, giving correlation coefficients for similarity of $r=0.617$, $r_s=0.451$ and $\tau=0.311$ when the tessellation frequency was 5 (3563 points), and $r=0.596$, $r_s=0.419$ and $\tau=0.283$ when the tessellation frequency was 1 (172 points). The error function complement for the two Pearson's correlation coefficients was 0.684, indicating that the values were not statistically different. The electrostatic potentials around the two antagonists in their matched conformations are displayed in Figs. 2a and b. The regression line had a gradient of 0.658 and an intercept of 2.718 kJ mol^{-1}.

It is interesting to note that there is a difference between the electrostatic potential comparison on the front and back faces of the superposed molecules [13]. The front and back surfaces were defined relative to the plane of the ring system of Ro 15-1788, which lies in the plane of the paper for Fig. 2a. The Spearman's rank correlation coefficient for the back face was better ($r_s=0.67$) than that for the front face ($r_s=0.23$). The most electrostatically similar regions corresponded to the putative hydrogen-bonding sites. In the fragment placement work described here, the supersurface used was the outer of the composite van der Waals surface of the two antagonists, rather than their composite accessible surface. In addition, no distinction was made between the front and back faces of the supersurface.

Two fragment placement procedures were studied: firstly, the self-placement test, in which, for example, the molecular graph of Ro 15-1788 was used and the electrostatic potential at the supersurface was that of Ro 15-1788 itself. Secondly, the molecular graph of one of the two antagonists was taken with the electrostatic potential of the other at the supersurface, so that the algorithm would attempt to produce a more similar structure to that whose electrostatic potential was projected onto the supersurface. For each of the three methods of assessing similarity, 10 runs were performed (Pearson's and Spearman's rank correlation coefficients, and pair differences).

Computer times are quoted for a Sun SPARCstation IPX for FORTRAN 77 programs.

TABLE 2
ANNEALING RESULTS USING EITHER Ro 15-1788 OR ZK 93426 AS A MOLECULAR GRAPH WITH THE ELECTROSTATIC POTENTIAL OF THE OTHER ON THE SUPERSURFACE UNDER VARYING ANNEALING CONDITIONS

Supersurface electrostatic potential used	Method of assessment	Mean value of the final objective function	Time (s)	Ratio perfect runs/ total runs	New atoms
Molecular graph of Ro 15-1788					
ZK 93426	r	0.703±0.064	79.3±8.5	0/10	4.1±1.5
	r_s	0.414±0.091	113.3±16.3	0/10	6.5±3.7
	eo	6019±988.4	116.3±9.3	0/10	6.0±1.4
Molecular graph of ZK 93426					
Ro 15-1788	r	0.568±0.113	101.6±20.2	0/10	4.5±2.2
	r_s	0.595±0.036	142.6±26.3	0/10	8.2±1.0
	eo	12960±13720	146.6±16.3	0/10	5.4±1.8

The values in each row represent averages over 10 runs for a given set of conditions. See the footnote to Table 1 for an explanation of the parameters.

TABLE 3
COMPARISON OF CORRELATIONS FOR ATOM-PLACED VARIANT STRUCTURES WITH THE ORIGINAL STRUCTURES

Molecular structure	r	r_s	τ	Gradient	Intercept (kJ mol^{-1})
Molecular graph of Ro 15-1788 (or variant) and potential of ZK 93426					
Ro 15-1788	0.617	0.451	0.311	0.65	2.72
Ro 15-1788 variant	0.677	0.361	0.244	1.49	−6.70
Molecular graph of ZK 93426 (or variant) and potential of Ro 15-1788					
ZK 93426	0.617	0.451	0.311	0.65	2.72
ZK 93426 variant	0.685	0.613	0.435	1.44	−6.59

Results

Self-placement

The results of the self-placement are shown in Table 1. The values in all tables refer to the direct output of the fragment placement program, without further processing.

The self-placement runs for the Ro 15-1788 graph did not reproduce Ro 15-1788 in any of the runs. This was attributed to the difficult configurational landscape presented to the algorithm, giving rise to many deep local minima which trapped the algorithm before it reached the global minimum. The objective function with some of these minima was sometimes better than that for Ro 15-1788 itself, suggesting that the fragment residual charges on this molecular graph had imperfect transferability. This can be illustrated by comparing the perfectly placed preprocessed correlation coefficients of $r=0.916$, $r_s=0.849$ with those of one of the better final objective function values of $r=0.924$, $r_s=0.882$. This better minimum structure actually contained a furan instead of the imidazole of Ro 15-1788. Other variations on Ro 15-1788 contained a pyrazole ring instead of the imidazole.

When the ZK 93426 molecular graph was used in self-placement, only one run out of 10 (with Pearson's correlation coefficient as the method of assessment) reproduced ZK 93426. The other structures were mostly variants on ZK 93426 in which the ether link was replaced with a thio-ether, giving rise to better preprocessed correlation coefficients ($r=0.907$, $r_s=0.780$) than for the perfectly placed graph ($r=0.903$, $r_s=0.768$). The algorithm was therefore locked towards the minimum of the thio-ether rather than that of ZK 93426 itself. The runs were repeated with a slower cooling rate (the acceptance scaling constant, C, was set to 1 instead of 2), but no improvement to the results was achieved.

Results using a different electrostatic potential on the supersurface

Table 2 shows the comparison of the similarity in the electrostatic potential of the alternate antagonist on the super-surface. The self-placement results showed that the algorithm was performing well, but that the main difficulty was with the transferability of fragment charges in two structures which were principally aromatic. The results of the placements using an electrostatic potential on the supersurface taken from the alternative antagonist revealed several structures which had better preprocessed correlation coefficients than those for Ro 15-1788 and ZK 93426. Some of the best of these new structures had their residual charges recalculated, and the new electrostatic potential correlations were obtained using a supersurface with a tessellation frequency of 5. They are compared in Table 3. When the molecular graph was that of Ro 15-1788 and the electrostatic potential on the supersurface was that of ZK 93426, the pyrazole variant of Ro 15-1788 had a Pearson's correlation coefficient which was better (erfc $=1.46\times10^{-5}$, indicating significance) than that for the Ro 15-1788/ZK 93426 comparison. Figure 3 shows the structure of this Ro 15-1788 analogue and how it compares electrostatically with ZK 93426. It may be helpful to the reader to compare this figure with the original structures in Fig. 2.

Even greater improvements were found with the ZK 93426 molecular graph in the envelope displaying the electrostatic potential of Ro 15-1788 (Table 3). One of the best structures produced with fragment placement was further processed with recalculated charges and a new supersurface with a tessellation frequency of 5. The improvement in Pearson's correlation coefficient was significant, with erfc $=4.31\times10^{-7}$. A comparison of the structure with Ro 15-1788 is shown in Fig. 4 and can be compared with Fig. 2.

Discussion and Conclusions

The fragment placement method has been successfully extended to the problem of envelope-directed design based on similarity studies of the electrostatic potentials. In the case of the two benzodiazepine antagonists investigated, an analogue of Ro 15-1788 produced by fragment placement has been shown to have improved electrostatic potential similarity (parametrically) with ZK 93426. Similarly, a ZK 93426 analogue was produced which had better electrostatic similarity with Ro 15-1788. Self-place-

ment tests on the two molecules showed non-ergodicity due to problems with fragment transferability. However, the alternative structures showed insignificantly different correlation coefficients from their target potentials. It would appear that these discrepancies are minor.

The optimizations were performed within the whole supersurface. No analysis has been performed on the electrostatic similarity of the front and back surfaces of the structures. This would be informative in light of the findings of Borea et al. [13], who have shown a better correlation for the back than for the front accessible surface of the superposed Ro 15-1788 and ZK 93426. If we had information about the active face of a target molecule, it would be possible to define a penalty function to place greater emphasis on the active face and secondary significance on less important regions of the structure.

The strategy in the previous and current papers has been to use the target 3D molecular graph unchanged. An important development of this work might be to consider the possibility of leaving out certain terminal vertices to improve the similarity, or complementarity in the case of site-directed atom assignment, within the objective function. The null correspondence method might be worth considering to achieve this [14]. Here, terminal vertices would be randomly omitted from the comparison as the optimization proceeds. It can be envisaged that this would create some very different structural variations with improved molecular similarity.

Acknowledgements

The authors would like to thank the Wellcome Trust for financial support through the Wellcome Prize Studentship (M.T.B.) and Wellcome Principal Research Fellowship (P.M.D.) schemes. Part of this work was carried out in the Cambridge Centre for Molecular Recognition.

References

1 Dean, P.M. (Ed.) Molecular Similarity in Drug Design, Blackie Academic and Professional, London, 1995.
2 Cramer III, R.D., Patterson, D.E. and Bunce, J.D., J. Am. Chem. Soc., 110 (1988) 5959.
3 Kim, K.H., In Dean, P.M. (Ed.) Molecular Similarity in Drug Design, Blackie Academic and Professional, London, 1995, pp. 291–331.
4 Perkins, T.D.J. and Dean, P.M., J. Comput.-Aided Mol. Design, 7 (1993) 155.
5 Barakat, M.T. and Dean, P.M., J. Comput.-Aided Mol. Design, 9 (1995) 341.
6 Barakat, M.T. and Dean, P.M., J. Comput.-Aided Mol. Design, 9 (1995) 351.
7 Barakat, M.T. and Dean, P.M., J. Comput.-Aided Mol. Design, 9 (1995) 359.
8 Barakat, M.T. and Dean, P.M., J. Comput.-Aided Mol. Design, 9 (1995) 448.
9 Lewis, R.A. and Leach, A.R., J. Comput.-Aided Mol. Design, 8 (1994) 467.
10 Codding, P.W. and Muir, A.K.S., Mol. Pharmacol., 28 (1985) 178.
11 Bertolasi, V., Ferretti, V., Gilli, G. and Borea, P.A., Acta Crystallogr., C40 (1984) 1981.
12 Burkert, U. and Allinger, N.L., Molecular Mechanics, ACS Monograph Series, Vol. 177, American Chemical Society, Washington, DC, 1982.
13 Borea, P.A., Dean, P.M., Martin, I.L. and Perkins, T.D.J., Mol. Neuropharmacol., 2 (1992) 261.
14 Barakat, M.T. and Dean, P.M., J. Comput.-Aided Mol. Design, 5 (1991) 107.