Molecular Similarity in Terms of Valence Electron Density

Philippa E. Bowen-Jenkins and W. Graham Richards

Physical Chemistry Laboratory, South Parks Road, Oxford OX1 3QZ, U.K.

Ab initio comparison of the valence electron density of pairs of molecules provides a quantitative measure of similarity which conforms with qualitative ideas of bioisosterism; the technique should provide a useful criterion in molecular design.

Much of the area of molecular design, particularly in the pharmaceutical industry, depends on the substitution of a molecular fragment with a group of atoms having similar steric and electronic properties, bioisosteric replacement. This procedure has traditionally depended on the chemist's knowledge,[1] but attempts are being made to quantify it. Carbo[2] introduced an index of similarity, R_{AB}, between molecules (or molecular fragments) A and B, in terms of their electron densities, ρ_A and ρ_B, equation (1), where integrations extend

$$R_{AB} = \frac{\int \rho_A \rho_B d\tau}{(\int \rho_A^2 d\tau)^{\frac{1}{2}} (\int \rho_B^2 d\tau)^{\frac{1}{2}}} \quad (1)$$

$$R_{AB} = \frac{\sum_{\mu}\sum_{\nu}\sum_{\rho}\sum_{\sigma} D^{A}_{\mu\nu} D^{B}_{\rho\sigma} \int \chi^{A}_{\mu} \chi^{A}_{\nu} \chi^{B}_{\rho} \chi^{B}_{\sigma} \, d\tau}{\left(\sum_{\mu}\sum_{\nu}\sum_{\rho}\sum_{\sigma} D^{A}_{\mu\nu} D^{A}_{\rho\sigma} \int \chi^{A}_{\mu} \chi^{A}_{\nu} \chi^{A}_{\rho} \chi^{A}_{\sigma} \, d\tau\right)^{\frac{1}{2}} \left(\sum_{\mu}\sum_{\nu}\sum_{\rho}\sum_{\sigma} D^{B}_{\mu\nu} D^{B}_{\rho\sigma} \int \chi^{B}_{\mu} \chi^{B}_{\nu} \chi^{B}_{\rho} \chi^{B}_{\sigma} \, d\tau\right)^{\frac{1}{2}}} \quad (2)$$

over all space. This function will be unity for pairs of identical molecules and has the range of values 0 to 1. Previously[3] we have presented an *ab initio* method for the computation of this function based on an expansion of equation (1) in terms of molecular wavefunctions, equation (2) where D is the charge density–bond order matrix (first order density matrix) and χ the atomic orbitals.

The well known bioisosteric replacement of $-CH_2-$ by $-S-$ led us to consider the series $MeCH_2Me$, MeOMe and MeSMe. The accurate calculation indicated that propane showed a greater similarity to the ether than to the sulphide, contrary to the normal ideas of bioisosterism. Consideration of the formula and results suggested that this was attributable to the core electrons of the sulphur atom. Since valence electrons are of primary importance in determining chemical activity, it seems reasonable to exclude core electron contributions when using index values in attempts to predict drug activity. There is also a hidden advantage in this valence approximation. It has been shown that similarity index values are highly sensitive to the relative positions of the molecules compared.[3] This is most pronounced when there is a centre of high electron density in one molecule. To attain a maximum in the similarity index, the relative positions of the molecules will, in limiting cases, be determined by the necessity of maximising overlap between the electron density of the whole of the second molecule with the highly negative part of the first. An important consequence of this is that least-squares matching of nuclear positions (as is standard in X-ray crystallography), which is at present the only reasonable and consistent method of molecular superposition available, is inadequate for providing a reliable approximation to the optimum relative molecular positions in these cases. By excluding core electrons, the least-squares method for molecular superposition provides a much better approximation to the optimum relative molecular positions. This can be seen in both Table 1 and Figures 1 and 2.

Molecular wavefunctions used in the calculations were obtained at the 4–31G level using the Gaussian 80 *ab initio* software.[4] Both this and the index computation program were implemented on a VAX 11/785 computer. The central processor time required for the actual index calculation is typically of the order of one quarter of that required for computation of the molecular wavefunctions.

This technique should have applications to more contentious fragment replacements such as the ring systems in

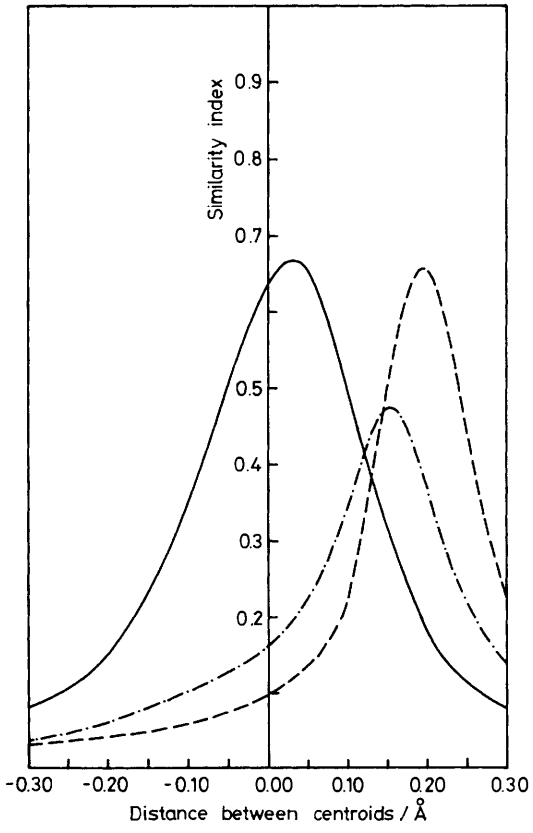

Figure 1. Similarity index of total electron density as a function of the separation between the centroids of the molecules: ——— Me_2O/Me_2CH_2 comparison; ·—·—·— Me_2S/Me_2CH_2 comparison; – – – Me_2O/Me_2S comparison.

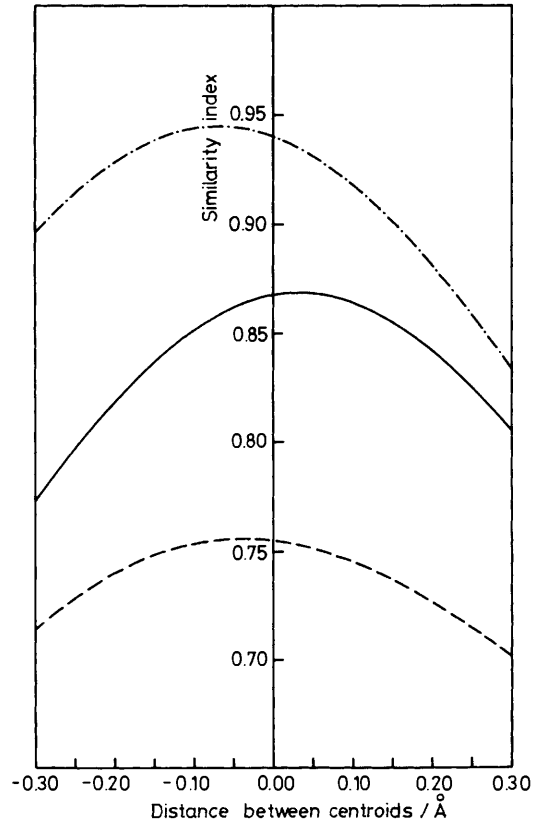

Figure 2. Similarity index of valence electron density as a function of the separation between the centroids of the molecules: ——— Me_2O/Me_2CH_2 comparison; ·—·—·— Me_2S/Me_2CH_2 comparison; – – – Me_2O/Me_2S comparison.

Table 1. Values of the similarity index for isosteric molecules.

Comparison	Method of matching[a]	Full[b] R_{AB}	Full[b] Δ	Valence[b] R_{AB}	Valence[b] Δ
Me_2O/Me_2CH_2	1	0.61	0.05	0.87	0.05
	2	0.63	0.00	0.87	0.00
	3	0.64	0.02	0.87	0.05
Me_2O/Me_2S	1	0.66	0.20	0.73	0.20
	2	0.10	0.00	0.75	0.00
	3	0.66	0.20	0.76	0.05
Me_2CH_2/Me_2S	1	0.48	0.16	0.89	0.16
	2	0.16	0.00	0.94	0.00
	3	0.48	0.16	0.95	0.08

[a] Matching methods are: 1 The central atoms are superposed, and the remaining main atom positions least-squares matched; 2 positions of all main atoms are least-squares matched following centroid superposition; 3 optimisation of index value by relative molecular translation along the symmetry axis. [b] Δ Represents the distance between molecular centroids, in Å.

prostaglandins, and also in developing a gradient method for the location of the maximum of the similarity function based on the adjustment of relative molecular positions.

This work was partially supported by a contract with the National Foundation for Cancer Research. The authors thank Dr. J. Kendrick of ICI New Science Group for helpful discussions.

Received, 26th September 1985; Com. 1400

References

1 C. W. Thornber, *Chem. Soc. Rev.*, 1979, **7**, 563.
2 R. Carbo, L. Leyda, and M. Arnau, *Int. J. Quant. Chem.*, 1980, **17**, 1185.
3 P. E. Bowen-Jenkins, D. L. Cooper, and W. G. Richards, *J. Phys. Chem.*, 1985, **89**, 2195.
4 J. S. Binkley, R. A. Whiteside, R. Krishnan, R. Seeger, D. J. de Frees, H. P. Schlegel, S. Topiol, L. R. Khan, and J. A. Pople, *Quantum Chemistry Program Exchange*, 1981, **13**, 406.