An *ab initio* topographical investigation on the molecular electrostatic potential of some chemical mutagens

Apurba K. Bhattacharjee, Savita S. Pundlik* and Shridhar R. Gadre*

Department of Chemistry, Lady Keane College, Shillong 793 001, India
*Department of Chemistry, University of Poona, Pune 411 007, India

A detailed topographical investigation on the molecular electrostatic potentials (MESPs) of different conformers of acetaldehyde, nitrous acid and hydroxylamine has been carried out at the *ab initio* SCF level using TZ2p, 6-31G* and STO-3G basis sets. In general, large regions of negative potential have been observed. An attempt has been made to correlate these potentials with biological activities of the molecules. Mutagenic and toxicological properties appear to be related to the presence of these large negative zones.

MOLECULAR electrostatic potential (MESP) is now an established tool for the study of the stereoelectronic

attributes and properties of binding molecules which would significantly contribute to the elucidation of biochemical phenomena[1–3]. Even long-range electrostatic interactions have now been shown to be important for a number of biological reactions including folding of proteins, determination of pKa's of the ionizable residues etc.[4]. The affinity of a particular molecule for a specific receptor is shown to depend on the MESP characteristics[5]. This ability of a molecule for interaction with the receptor is essentially known as the 'recognition' process, in which the receptor 'recognizes' the characteristic features of the molecule in order to promote their interaction. MESP can, therefore, be used for predicting structure–activity relationships when a receptor-mediated response is supposed to give rise to toxicological, mutagenic or carcinogenic effects even without a detailed knowledge of the receptor[6]. MESP provides the exact value for the electrostatic component of the interaction energy between an unperturbed charge distribution and a proton[7]. Moreover, MESP gives a general picture of the reactive properties of the molecule with a reasonable computational cost and this explains its large number of applications in very different fields of physics, chemistry, biochemistry and pharmacology[8]. In a recent report Zhu *et al.*[9] have shown that radiosensitizers can be made based on MESP data, thereby indicating its possible applications in the study of molecular toxicology. Efficient radiosensitizers are characterized by a wide and deep negative potential in their MESP profiles, while this potential is narrow for compounds devoid of significant activity[9].

The MESP $V(r)$ at a point r due to nuclear charges $\{Z_A\}$ at $\{R_A\}$ and the electronic density $\rho(r)$ of the molecule is given by

$$V(r) = \sum_A \frac{Z_A}{|R_A - r|} - \int \frac{\rho(r')}{|r - r'|}\, d^3r'. \qquad (1)$$

The first term on the right-hand side of this equation represents the bare nuclear potential while the second one represents the electronic density function of the molecule. The sign of $V(r)$ in any particular region depends on whether the nuclear or the electronic effects are dominant there.

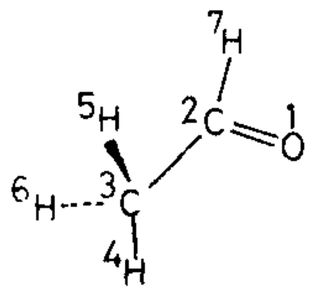
The topographical analysis of $V(r)$ involves identification and location of the critical points (CPs)[10,11], the points at which $\nabla V(r) = 0$. A CP in the MESP topography specifies a point at which the internal electric field (molecular electric field or all the partial derivatives of MESP) is exactly zero. This implies that the numerical value of the molecular electric field can be used for locating CPs in the MESP topography. The rank of the CP is given by the number of nonzero eigenvalues of the Hessian matrix A, whose elements are given by

$$A_{ij} = \frac{\partial^2 V}{\partial r_i\, \partial r_j}\Bigg|_{r = r_{cp}} \qquad (2)$$

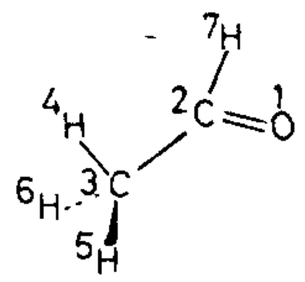
A CP of rank 3 (known as a nondegenerate CP) is characterized by the algebraic sum of the signs of eigenvalues. Thus, there are four possible types of CPs: (3, –3), (3, –1), (3, +1) and (3, +3). The (3, –3) CP cannot exist in $V(r)$ except at the nuclear locations[11]. A (3, –1) or (3, +1) CP would correspond to saddle points whereas a (3, +3) CP is due to the local minimum of $V(r)$ beyond the van der Waals surface of the molecule.

The MESP minima can give an idea of the deprotonation sites as well as the interaction possibilities at the electrophilic centres of the receptor. Since MESP reflects the molecular interaction energy with a test positive point charge, the (3, +3) CPs would correspond to their proton affinities[12]. This would imply a threshold interaction energy of the atoms in the molecule to extract a proton from the proton donor site on the receptor. Good correlations between proton affinity or hydrogen bond acceptor parameters and MESP have been established in the recent past[13].

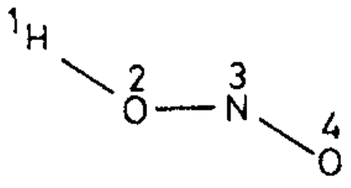
Chemical mutagens act by modifying the bases of DNA. Nitrous acid, hydroxylamine and acetaldehyde are all well-known mutagens[14]. Nitrous acid causes AT ↔ GC transitions whereas hydroxylamine is a highly specific mutagen which reacts almost exclusively with cytosine to give a derivative that pairs with adenine rather than with guanine and so it produces a unidirectional transition of CG to AT[14]. Acetaldehyde has also been reported[15] to be highly toxic, comparable to nitrous acid and hydroxylamine. These molecules, apart from their toxicities, are also known to be efficient proton removers and deaminating agents[16]. Further, these molecules are also interesting because there is a strong evidence that they can act as mutagens most likely at the same site[12].



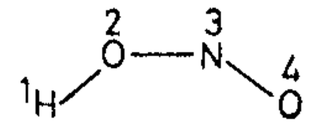
eclipsed
1a



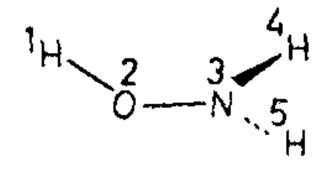
staggered
1b



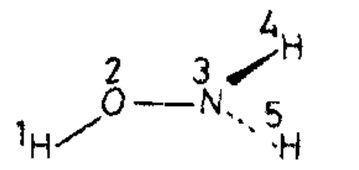
trans
2a



cis
2b



cis-staggered
3a



trans-staggered
3b

The objective of the present study is to assess the importance of three-dimensional MESP topographical analysis of acetaldehyde (1a, 1b), nitrous acid (2a, 2b) and hydroxylamine (3a, 3b) at a higher level of theory in the mutagen–receptor interaction process. Although electron density maps of some of these molecules[17] have been reported earlier, to our knowledge detailed topographical analysis at this level of theory has not been reported so far in the literature.

We have used *ab initio* optimized geometries of three molecules in their different conformations[17–19] and obtained the wave functions (with TZ2p, 6-31G* and STO-3G basis sets) by using the INDMOL package[20] on a 64-node transputer-based machine called PARAM[21]. The MESP calculations were carried out using these wave functions on the same computer with the help of a program developed by Shirsat *et al.*[22]. The computed electrostatic potentials are at a distance of about 3 a.u. surrounding the molecule, which corresponds approximately to the van der Waals radius of the largest atom in this group of molecules (oxygen)[23]. The potential at this surface would usually be encountered by the receptor during the 'recognition' process. The basis set dependence of the MESP was checked by using different basis sets in these calculations. The CPs were located using the analytical expressions for the gradient of $V(r)$ given by Gadre *et al.*[10,24].

In order to have a quantitative measure of the negative region typical for biological recognition interaction[3], we have calculated the volume inside a closed surface having an MESP energy of –15.6 kcal/mol. This is done by calculating MESP values on a 64-node PARAM machine[21] inside a rectangular parallelopiped enclosing an isosurface of this energy and sampling the points through a random search (Monte Carlo method). The volumes thus obtained for all the molecules studied are indicated in Table 1.

Table 1. Types and locations of negative-valued MESP critical points (CPs) calculated at TZ2p level

Molecule*	CP type	Coordinates (a.u.) x	y	z	MESP value (kcal/mol)	Volume (a.u.)³
Acetaldehyde eclipsed[a]	(3, +3)	–1.42	–1.82	0.0	–52.8	
	(3, +1)	–2.43	0.0	0.0	–46.5	239.3
	(3, +3)	–1.47	1.77	0.0	–51.4	
staggered[b]	(3, +3)	–1.42	–1.82	0.0	–53.0	
	(3, +1)	–2.43	0.0	0.0	–46.8	241.4
	(3, +3)	–1.47	1.77	0.0	–51.6	
Nitrous acid *trans*-[c]	(3, +3)	1.20	4.40	0.0	–21.9	
	(3, +1)	2.87	5.78	0.0	–11.9	
	(3, +3)	5.81	4.63	0.0	–36.9	
	(3, +1)	7.54	2.72	0.0	–25.9	156.0
	(3, +3)	6.25	0.47	0.0	–36.5	
	(3, +1)	4.82	–1.20	0.0	–22.9	
	(3, +3)	3.40	–1.92	0.0	–28.5	
cis-[d]	(3, +3)	–0.56	4.60	0.0	–18.4	
	3, +3)	3.04	6.06	0.0	–30.2	
	(3,+1)	5.72	4.65	0.0	–10.3	34.0
	(3, +3)	5.59	1.96	0.0	–24.4	
	(3, +1)	5.56	–0.93	0.0	–7.5	
	(3, +3)	3.18	–2.06	0.0	–20.7	
Hydroxylamine *cis*-staggered[e]	(3, +3)	4.92	2.44	0.0	–68.7	
	(3, +1)	4.80	–0.18	0.0	–38.0	
	(3, +1)	3.32	–1.77	0.0	–53.1	278.8
	(3, +3)	3.19	–1.64	–0.8	–53.3	
	(3, +3)	3.19	–1.64	0.8	–53.3	
trans-staggered[f]	(3, +3)	3.03	–1.91	0.0	–51.2	
	(3, +3)	1.20	4.00	0.0	–63.0	123.0

*Nuclear coordinates (in a.u.) of some of the atoms are:
[a]O (0,0,0), C (1.18,0,0), C (2.03,1.24,0), H (1.40,2.12,0).
[b]O (0,0,0), C (1.18,0,0). C(2.03,1.25,0), H (1.82,1.85,0.87).
[c]H (0,0,0), O (0.95,0,0), N (1.40,1.23,0), O (2.56,1.29,0).
[d]H (0,0,0), O (0.96,0,0), N (1.59,1.12,0), O (0.9,2.09,0).
[e]H (0,0,0), O (0.95,0,0), N (1.36, 1.26,0), H (0.99,1.71,–0.84).
[f]H (0,0,0), O (0.95,0,0), N (1.47,1.22,0), H (2.05,1.38,0.83).

In Table 1, the CP characteristics of the molecules studied at TZ2p level are presented. The calculated MESP minima (most negative values) are beyond the van der Waals surface of the molecules. The electrostatic potential maps of the molecules are presented in Figure 1.

In acetaldehyde, practically no energy difference is found between the MESP critical points of the eclipsed and the staggered conformations, which indicates that the MESP profile is independent of methyl rotation, although methyl barrier to internal rotation around the C=O bond has been calculated at the 6-31G* level to be 1.06 kcal/mol (ref. 17). The two negative-valued minima are found in the lone pair regions of oxygen in both the rotamers. The TZ2p MESP maps of both the rotamers show the presence of strong negative potential zones laterally around oxygen encompassing a broad region which might be linked to the strong biological activity of the molecule. The negative regions associated with the carbonyl oxygen extend to a fairly large region; the minima representing the lone pairs have a value around –51.0 kcal/mol separated by a saddle corresponding to about –46.0 kcal/mol.

In the case of *trans*-nitrous acid, seven distinct CPs with a deepest minimum of about –37.0 kcal/mol are observed. A large extended area of about –20.0 kcal/mol is observed which is most likely to be the characteristic feature of the 'recognition' interaction. For *cis*-nitrous acid, six CPs were located, with the deepest minimum at –30.0 kcal/mol. A broad extended area of about –8.0 kcal/mol is observed as the saddle energy.

Cis-staggered hydroxylamine shows three minima, two of which are at –68.0 and the other one is at –53.0 kcal/mol. A saddle energy range of about –40.0 to –50.0 kcal/mol is found to be extended over a large area on only one side of the molecule. In the case of *trans*-staggered hydroxylamine, however, the MESP

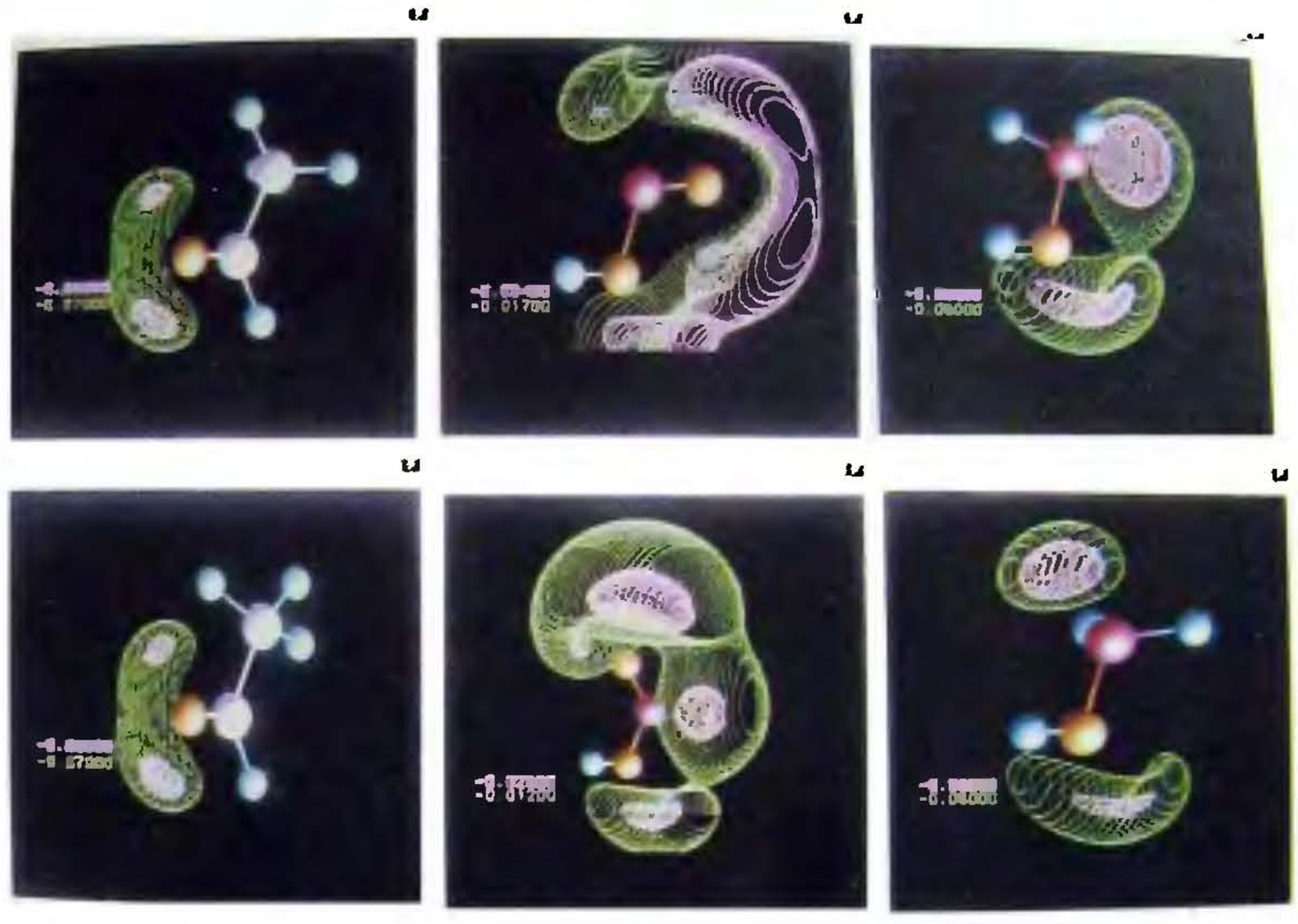


Figure 1. The molecular electrostatic potential contours of. *a*, acetaldehyde eclipsed (upper) and staggered (lower). The green colour corresponds to a value of –44.0 kcal/mol, while the pink one corresponds to –50 2 kcal/mol. *b*, *Trans*-nitrous acid (upper) showing green colour contours of –10 6 kcal/mol and pink colour contours of –21.3 kcal/mol, and *cis*-nitrous acid (lower) showing green colour contours of –7 5 kcal/mol and pink colour contours of –17.5 kcal/mol. *c*, Hydroxylamine *cis*-staggered (upper) and *trans*-staggered (lower). The green colour corresponds to a value of –37 6 kcal/mol, while the pink one corresponds to –50.2 kcal/mol.

minima are localized on both sides of the molecule, one near oxygen and the other near nitrogen, with the potential values of –50.8 and –62.7 kcal/mol, respectively. The saddle energy of about –40.0 kcal/mol is found to be well extended on both sides of this molecule.

Calculations made at the 6-31G* level could reproduce similar MESP profiles for all the molecules. For *cis*-staggered hydroxylamine, however, the two minima off the *xy* plane obtained with TZ2p basis could not be distinguished and they appeared to have converged into a single one in the plane.

Although STO-3G calculations show qualitative similarity of the MESP features of the molecules, the locations and the potential values of the CPs differ significantly. In acetaldehyde, for example, the deepest minimum has a higher value of MESP (–34.0 kcal/mol) and in *cis*-nitrous acid, it is as low as – 49.0 kcal/mol. Thus, the choice of STO-3G basis sets for MESP calculations of such molecules should perhaps be made with caution.

The magnitudes of the negative-valued minima of all the molecules may be correlated with their proton affinities. This is because MESP reflects the molecular interaction energy with a positive point charge; therefore, this MESP minimum threshold would correspond to the proton affinity of the molecules (the amount of energy necessary to extract a proton from the proton donor site on the receptor). Such correlations have been reported earlier for a series of oxygen- and nitrogen-containing molecules[12, 25].

In order to compare the biological activities, we have scanned volumes enclosed within the negative region having potentials in the neighbourhood of –15.0 kcal/mol for all the molecules. A value of such a magnitude has been reported by Murray *et al.*[3] as optimum potential minimum to explain the toxicities of several dibenzo-*p*-dioxins. In our scan, we too observe (Table 1) a very large area being encompassed with this energy. It may thus be reasonable to infer that both the depth of the MESP minima and the magnitude of the 'volumes' together play a crucial role in the biological interaction process. From an inspection of the 'volume enclosure' of –15.6 kcal/mol, it is apparent that both the rotamers of acetaldehyde (eclipsed 239.8 a.u^3 and staggered 241.4 a.u^3) and *cis*-staggered hydroxylamine (278.8 a.u^3) are more prone to biological recognition interaction than *trans*-staggered hydroxylamine (123.0 a.u^3), *trans*-nitrous acid (156.0 a.u^3) and *cis*-nitrous acid (34.0 a.u^3). It may be summarized from the above results that *cis*-staggered hydroxylamine would be biologically more active, followed by the two acetaldehyde rotamers. Apart from having high negative potential regions the molecules are much smaller in size, which enable them to a greater accessibility to the receptor site. Thus, three factors – (i) the magnitude of the deepest negative potential, (ii) a large optimum negative potential zone and (iii) a

smaller size (consequently, a better steric accessibility) – appear to be crucial for the molecules to form an energetically favourable complex with the receptor which would interfere with the subsequent reactions taking place in the receptor.

The computational results of this theoretical MESP analysis on the different conformers of acetaldehyde, nitrous acid and hydroxylamine indicate that three factors – (i) the magnitude of the deepest negative potential, (ii) a large 'optimum' negative potential zone and (iii) small size – would contribute largely towards their biological activity. The MESP analysis may thus be used as a tool to determine the degree of mutagenic capacity *vis-à-vis* deprotonating as well as deaminating abilities of these molecular mutagens.

Although electrostatic potential patterns shown in this study can offer a rich informative description about the capacity of the molecules to generate stereoelectrostatic forces, they fail to describe their ability to generate hydrophobic bonds and dispersion interactions. Therefore, the MESP profiles in the present investigation would reveal only the stereoelectronic components of intermolecular recognition forces. Thus, further sophisticated calculations including solvent effects are required to prove or disprove the proposed reasons for the biological potency of the molecules.

1. Guha, S., Majumdar, D. and Bhattacharjee, A. K., *J. Mol. Struct (Theochem)*, 1992, **88**, 61–74
2. Majumdar, D., Bhattacharjee, A. K., Das, K. K. and Guha, S., *J Mol Struct (Theochem)*, 1993, **288**, 41–53.
3. Murray, J. S., Evans, P. and Politzer, P., *Int. J. Quant. Chem*, 1990, **37**, 271–275.
4. Nathrup, S., Boles, J. and Reynolds, J., *Science*, 1988, **241**, 67.
5. Thompson, C. and Brandt, R., *Int J. Quant Chem., Quant Biol Symp*, 1983, **10**, 357–361
6. Richard, A. M., *J. Comput Chem.*, 1991, **12**, 959–962.
7. Alhanbra, C., Luque, F. J. and Orozco, M., *J Phys Chem.*, 1995, **99**, 3084–3087.
8. Politzer, P. and Murray, J. S., in *Reviews in Computational Chemistry* (eds Lipkowitz, K. B. and Boyd, D. B.), VCH Publishers, New York, 1991, vol. 2.
9. Zhu, A., Xu, S., Huang, J. and Luo, Z., *Int. J. Radiat. Biol*, 1989, **56**, 893–895
10. Gadre, S. R., Kulkarni, S. A. and Shrivastava, I. H., *J. Chem Phys*, 1992, **96**, 5253–5256.
11. Pathak, R. K. and Gadre, S. R., *J Chem Phys.*, 1990, **93**, 1770–1772.
12. Luque, F. J., Sanz, F., Illas, F., Ponplana, R. and Smeyers, Y. G., *Eur. J. Med Chem*, 1988, **23**, 7–10
13. Murray, J. S., Ranganathan, S. and Politzer, P., *J Org. Chem.*, 1991, **56**, 3734–3738.
14. Stryer, L., *Biochemistry*, 3rd edn, Freeman & Co., New York, 1988, p. 677.
15. Robbins, S. L., Cotran R. S. and Kumar, V., *Pathological Basis of Diseases*, 3rd edn, Saunders, London, 1984, pp. 122, 922.
16. Potapov, V. M. and Tartarinchic, S. N., *Organic Chemistry* (translation), Mir, Moscow, 1979, p. 222
17. Wiberg, K. and Martin, E., *J. Am Chem Soc.*, 1985, **107**, 5035–5040.
18. Gordon, M. S., *J. Am. Chem Soc.*, 1969, **91**, 3122–3125.
19. Cox, A. P. and Kuczkowski, L., *J. Am Chem. Soc.*, 1966, **88**, 5071–5075.
20. Shirsat, R. N., Limaye, A. C. and Gadre, S. R., *J. Comput. Chem.*, 1993, **14**, 445–449.
21. Eknath, P. R., Bhasin, L. and Degwekar, A., in *Advanced Computing* (ed Bhatkar, V. P.), Tata McGraw Hill, New Delhi, 1991, p. 86
22. Shirsat, R. N., Bapat, S. V. and Gadre, S. R., *Chem. Phys. Lett*, 1992, **200**, 373–375.
23. Bondi, A., *J. Phys Chem*, 1964, **68**, 441–445
24. Gadre, S. R. and Shrivastava, I. H., *J. Chem. Phys.*, 1991, **94**, 4384–4390
25. Kollman, P. and Rothenberg, S., *J. Am Chem Soc.*, 1977, **99**, 1333–1336.

ACKNOWLEDGEMENTS. One of us (AKB) thanks the Department of Science and Technology (DST), New Delhi, for a visiting SERC fellowship to carry out the work. The financial assistance from the Council of Scientific and Industrial Research (CSIR) is gratefully acknowledged. We are extremely thankful to the Centre for Development of Advanced Computing (C-DAC), Pune, for providing the computational facilities.

Received 19 September 1994, revised accepted 18 May 1995