Signatures of molecular recognition from the topography of electrostatic potential[†]

DHIMOY K ROY, P BALANARAYAN and SHRIDHAR R GADRE*

Department of Chemistry, University of Pune, Pune 411 007
e-mail: gadre@chem.unipune.ernet.in

Abstract.   The recognition of interaction between two molecules is analysed via the topography of their molecular electrostatic potentials (MESP). The point of recognition between two species is proposed to be the geometry at which there is a change in the nature of the set of MESP critical points of one of the molecules *vis-a-vis* with its MESP topography at infinite separation. These results are presented for certain model systems such as pyridine and benzene dimers, cytosine–guanine and adenine–thymine base pairs in various orientations of approach of the two species.

Keywords.   Molecular electrostatic potential (MESP); molecular recognition; topography; critical point (CP).

1.   Introduction

The term molecular recognition generally refers to the specific interaction between two or more molecules through non-covalent bonding such as hydrogen bonding, metal coordination, van der Waals forces, π–π interactions, and/or electrostatic effects.[1] The host and guest involved in molecular recognition exhibit complementarity of certain properties with respect to each other.[1,2] Molecular recognition plays an important role in biological systems and has been observed for receptor-ligand, antigen-antibody, DNA-protein, sugar-lectin interactions, to list a few. Chemists have designed artificial supramolecular systems that exhibit specific molecular recognition. One of the earliest known examples is the class of crown ethers which are capable of selectively binding specific cations.

Molecular recognition is generally classified into static and dynamic one. Static molecular recognition is likened to the interaction between a key and keyhole; it is a 1 : 1 type complexation reaction between a host molecule and a guest molecule to form a host–guest complex. To achieve advanced static molecular recognition, it is necessary to make recognition sites that are specific for guest molecules. In the case of dynamic molecular recognition, the binding of the first guest to the first binding site of a host affects the association constant of a second guest with a second binding site.[3] The dynamic nature of this type of molecular recognition is particularly important since it provides a mechanism to regulate binding in biological systems. Dynamic molecular recognition is also being studied for application in highly functional chemical sensors and molecular devices.

Molecular recognition is thus a term that has been widely used in physics, chemistry and biology. The term essentially refers to the onset of interaction between two molecular species. The interaction here could be a precursor to covalent bond formation or van der Waals type or even a weak hydrogen bonding interation. The idea that molecular recognition lies in the complementarity of interacting surfaces was first clearly formulated by the German chemist Emil Fisher,[4] who proposed in 1894 that the enzyme and substrate fit together 'like lock and key'.

Molecular electrostatic potential (MESP) has been found to be a useful entity for describing reactivity in chemistry.[5–7] Being long range in nature, MESP is found to be useful for exploring molecular recognition.

Previous studies by Náráy Szábo and coworkers[8,9] have tried to understand molecular recognition in terms of electrostatic potentials of the two involved species. The shape complementarity has often been associated with the isosurface shapes of functions such as electron density.[10] Although a quantification of molecular recognition has been given in terms of the parts of the molecule endowed with positive

MESP in regions of negative MESP, a further analysis and understanding of the phenomenon is required. Mapping the topography of a scalar field such as molecular electron density[11] or MESP[12,13] offers a succinct tool for this purpose. In this context, the current work examines the onset of recognition between two molecules via the topography of their MESP's.

The point of recognition between two species is proposed to be the geometry at which there is a change in the nature of the set of MESP critical points (CP's) of one of the molecules in comparison with its MESP topography at infinite separation. These results are presented for certain model systems such as pyridine- and benzene dimers, as well as adenine–thymine and cytosine–guanine base pairs in various modes of approach of the two species.

2. Methodology

The test molecules are subjected to a geometry optimization at the Hartree–Fock (HF) level of theory and 6-311++G($2d$, $2p$) basis set from the Gaussian suite of programs.[14] The electrostatic potentials of these systems were evaluated and their respective topography is mapped using the in-house developed property package UNIPROP.[15,16] The visualization is done using the package UNIVIS-2000.[17]

The MESP is defined by[5,7] (in a.u.):

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

Here, $\{Z_A\}$ are the nuclear charges located at $\{R_A\}$ and $\rho(\mathbf{r})$ is the continuous electron density.

There are four types of non-degenerate critical points (CP's), for a three-dimensional function such as MESP, at which $\nabla V = 0$. The CP's are characterized by the eigenvalues of the Hessian of the matrix at the CP and denoted in the (Rank R, Signature S) notation. The four types of non-degenerate CP's are $(3, -3)$ (maxima), $(3, -1)$ (saddle which is a minimum in only one direction) $(3, +1)$ (saddle which is a maximum in only one direction) and $(3, +3)$ (minima). The CP's found using the program UNIPROP satisfy the necessary condition which is expressed by the Poincaré–Hopf relation,[18,19] viz. $n_{-3} - n_{-1} + n_{+1} - n_{+3} = n_- - n_+$. Here n_S denotes the number of given type of CP (e.g. n_{-3} stands for the number of $(3, -3)$ CP's, etc.) and n_- and n_+ denote the number of asymptotic maxima and minima respectively.

For specifically examining the problem of recognition, a few selected orientations of the interacting species (called monomers) A and B are chosen and subjected to a single point wavefunction evaluation at that geometry. The wavefunction obtained is thereafter utilized for topography mapping. Subsequently the monomers are brought close to each other (in a weakly bound form $A...B$) from a large distance and the topography mapped for the MESP at different steps along the path.

3. Results and discussion

3.1 *Pyridine dimer*

The pyridine dimer has recently been studied in the literature for understanding its weak interaction patterns.[20] Hence, this system has been chosen as the first model system for exploring molecular recognition. In the analysis presented here, the first step involves the mapping of the topography of the separated monomer. The MESP maps together with the topography of pyridine are presented in figure 1. For an individual system A the topography is denoted by $N_{CP}[V_A] = [n_{+3}, n_{+1}, n_{-1}, n_{-3}]$. For pyridine, the criticalities of MESP are given by [5, 7, 13, 11]. The most negative minimum has an ESP value of -0.113 a.u.

As a first test case, a T-shaped pyridine dimer is considered, three geometries of which are shown in figure 2. The distance between the center of masses of the two monomers is varied and the topography mapped at each point. A new set of CP's appears
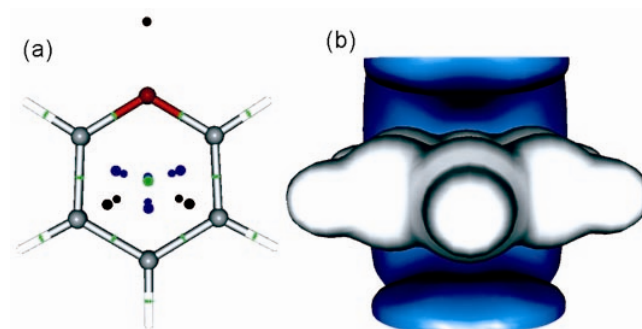
Figure 1. (a) Pyridine topography at HF/6-311++G($2d$, $2p$) level of theory with $(3, +3)$, $(3, +1)$ and $(3, -1)$ CP's denoted by small black, blue and green spheres respectively. (b) Molecular electrostatic potential of pyridine at HF/6-311++G($2d$, $2p$) depicted by a blue isosurface of value -0.01 a.u. and a white isosurface of value $+0.1$ a.u. See text for more details.
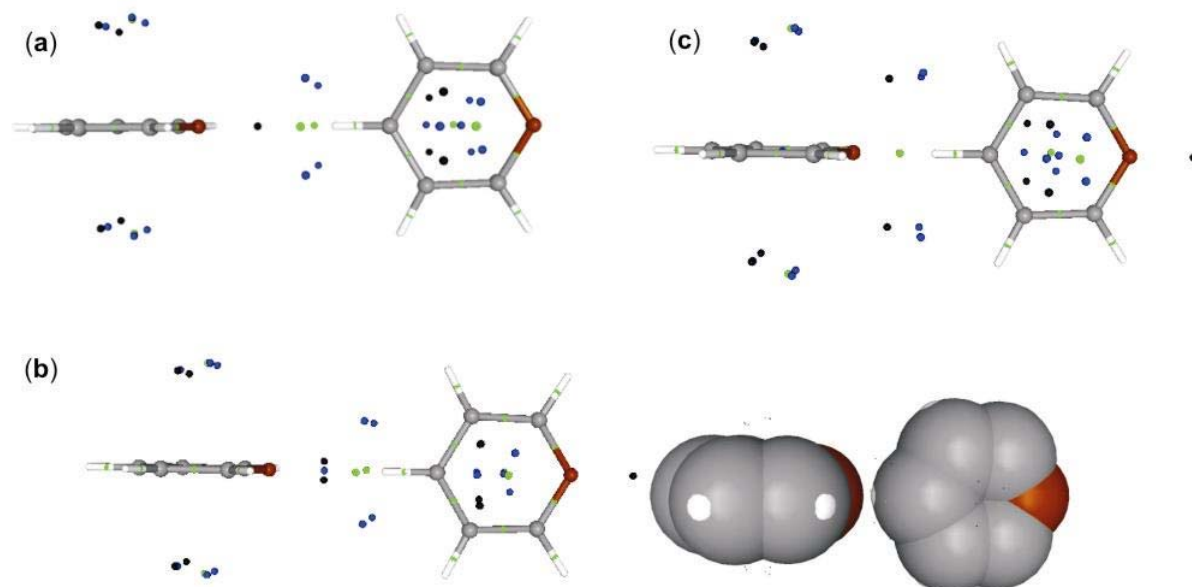
Figure 2. T-shaped geometries of pyridine dimer. The (3, +3), (3, +1) and (3, −1) CP's are denoted by small black, blue and green spheres respectively. See text for more details. Geometries with distances between center of masses of monomers at (**a**) 15·00 a.u., (**b**) 11·88 a.u., (**c**) 10·00 a.u.

Table 1. Topographical changes for pyridine dimer T-shaped approach. $N_{CP}[V_A] = [n_{+3}, n_{+1}, n_{-1}, n_{-3}]$ denotes the number of minima, numbers of (3, +1) and (3, −1) saddles and (3, −3) maxima respectively. V_A and V_B correspond to the monomer MESP in the composite MESP V_{AB}. See text for details.

Distance	$N_{CP}[V_{AB}]$	$N_{CP}[V_A]$	$N_{CP}[V_B]$	Change in CP
T-shaped dimer				
20·00	10, 16, 28, 22	5, 7, 13, 11	5, 7, 13, 11	No change
15·00	10, 16, 28, 22	5, 7, 13, 11	5, 7, 13, 11	No change
12·50	10, 18, 30, 22	5, 7, 13, 11	5, 7, 13, 11	No change
11·88	11, 19, 30, 22	4, 8, 13, 11	5, 7, 13, 11	+3 → +1
11·25	11, 19, 30, 22	6, 7, 13, 11	5, 7, 13, 11	+3 → −1
10·00	11, 18, 29, 22	6, 7, 13, 11	5, 7, 13, 11	+3 → −1
Antiparallel dimer				
20·00	10, 16, 28, 22	5, 7, 13, 11	5, 7, 13, 11	No change
15·00	10, 16, 28, 22	5, 7, 13, 11	5, 7, 13, 11	No change
10·00	10, 16, 28, 22	5, 7, 13, 11	5, 7, 13, 11	No change
7·50	11, 19, 30, 22	4, 7, 14, 11	5, 7, 13, 11	+3 → −1
5·00	8, 16, 30, 22	4, 7, 14, 11	4, 8, 13, 11	+3 → −1
Antiparallel displaced				
20·00	10, 15, 28, 22	5, 7, 13, 11	5, 7, 13, 11	No change
15·00	10, 16, 28, 22	4, 5, 13, 11	4, 6, 13, 11	+1 with −1
10·00	8, 11, 26, 22	4, 5, 12, 11	4, 5, 12, 11	+1 with −1

even at a very large distance of 20·00 a.u. These new CP's have an MESP value that is negative but close to zero. The CP's close to the pyridines still retain characteristics of the individual monomer. Hence the basic topographical features of the pyridines in the dimer complex are still preserved.

The topographical changes for pyridine dimer at various distances are shown in table 1. It can be seen that first change of topography of pyridine in the T-shaped homodimer complex occurs when the distance between the center of masses (c.m.) is about 11·88 a.u. At this distance, the (3, +3) minimum

over the nitrogen of one of the pyridines changes to a (3, +1) saddle. This point at which the first topographical change occurs to the MESP of a monomer in the complex, as compared with that of the free monomer is proposed to be the onset of the recognition process. This first topographical change is observed near to the contact point (cf. figure 2b) of the van der Waals surfaces of the pyridine monomers. Such a connection of MESP topography with the van der Waals radii of atoms has been brought out previously by Bhadane and Gadre.[21]

Yet another orientation that is possible for pyridine dimer is that of antiparallel face-to-face case where the nitrogen atom of one pyridine moiety may interact with the hydrogen of the other along the C_2 axis. In this particular orientation, a change in the topography is observed only at a c.m. to c.m. distance of 7·5 a.u. as against the point of recognition at 11·88 a.u. in the case of the T-shaped geometry. Figure 3a displays the geometry and topographical change described. The change that is seen here is also given in table 2, and it can be seen that the topography of one of the pyridines changes from [5, 7, 13, 11] to [4, 7, 14, 11]. In this case a minimum of pyridine is converted to (3, −1) saddle.

In the third orientation, viz. the antiparallel displaced one, a first change in the topography of the monomer is noticed at a large distance of 15 a.u. In this case, the topography changes from [5, 7, 13, 11] to [4, 5, 12, 11] and this change occurs for both the monomers. Figure 3b shows the geometry and topographical change described. This antiparallel displaced geometry, when brought closer together finally gives that orientation which is energetically more stable than the antiparallel and the T-shaped geometry.

3.2 *Benzene dimer*

Yet another interesting case of interaction is often seen in the case of benzene. Many gas phase calculations of benzene dimer by Hobza *et al*[22] have found their place in the literature and the competition between the H...π and the π...π interactions has been a primary object of their study. Similar to the case of pyridine dimer, here a T-shaped geometry and face to face approach is considered here in order to study the recognition phenomenon. The isolated benzene dimer topography is described by [12, 13, 14, 12], where one finds 6 minima each above and below the planes of the ring. For the T-shaped interaction, at a distance of 20.0 a.u. a set of new (3, −1) and (3, +1) CP's appears in the distribution. However, these are of magnitude close to zero. The first noticeable change in the topography of benzene in the benzene dimer occurs when the distance between the center of masses of the monomers is 10.3 a.u. Here the change in the nature of the topography is a combination of two (3, +1) saddles and a minimum of benzene disappearing to give a single minimum. The topography of the benzene appearing horizontal in the T-shaped geometry at this point happens to be [12, 11, 14, 12].

In a face-to-face approach of benzene which is reminiscent of a direct π–π interaction, at a distance of 20.0 a.u. two low valued (3, −1) CP appear on the sides together with a (3, +1) CP of similar magnitude, on the C_6 axis which is the axis of approach. At this distance there is no change in the nature of the CP's of benzene in its dimer form. This continues till a distance 8·5 a.u. between the center of masses. Although many other new CP's appear for the composite MESP, V_{AB}, the individual benzenes retain their topographical similarity to the isolated benzene till a distance of 8·5 a.u. Beyond this distance there is a dramatic change in the topography of both of the monomers in the dimer complex, with the topography changing from [12, 11, 14, 12] to [9, 9, 14, 12]. Here two (3, +3) CP's combine with a
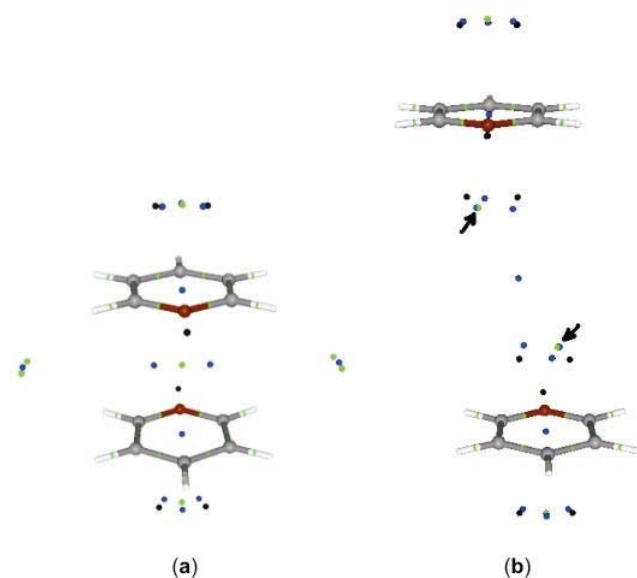
Figure 3. Antiparallel (**a**) and antiparallel displaced (**b**) geometries of pyridine dimer. The (3, +3), (3, +1) and (3, −1) CP's are denoted by small black, blue and green spheres respectively. See text and table 1 for more details.

Table 2. Topographical changes for adenine–thymine approach in figure 3. $N_{CP}[V_A] = [n_{+3}, n_{+1}, n_{-1}, n_{-3}]$ denotes the number of minima, numbers of $(3, +1)$ and $(3, -1)$ saddles and $(3, -3)$ maxima respectively. V_A and V_B correspond to the monomer MESP in the composite MESP V_{AB}. See text for details.

Distance	$N_{CP}[V_{AB}]$	$N_{CP}[V_A]$	$N_{CP}[V_B]$	Change in CP
14·50	9, 14, 33, 29	5, 8, 19, 16	4, 5, 13, 13	No change
9·50	10, 14, 33, 29	5, 8, 19, 16	4, 5, 13, 13	No change
5·88	10, 16, 35, 29	5, 8, 19, 16	4, 4, 14, 13	No change
5·00	9, 14, 34, 29	5, 9, 17, 16	4, 4, 15, 13	$+3 \rightarrow +1$
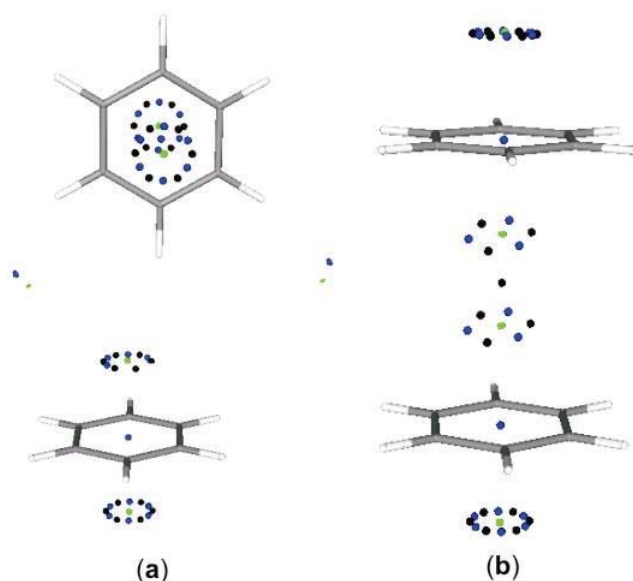
Figure 4. Benzene T-shaped (**a**) at 10 a.u. separation and benzene parallel (**b**) at 8 a.u. c.m. to c.m. separation . The $(3, +3)$, $(3, +1)$ and $(3, -1)$ CP's are denoted by small black, blue and green spheres respectively. See text for more details.

$(3, +1)$ saddle to give a single minimum, on the side in which the benzenes face each other, while on the other sides, the topographical similarity with the isolated benzene is retained.

3.3 *Adenine–thymine and cytosine–guanine base pairs*

A similar approach is adopted for analysing the nature of interaction in the case of the DNA base pairs adenine–thymine (*A–T*) and cytosine–guanine (*C–G*). In the case of the adenine–thymine base pair, the orientation was chosen to be the one usually found in the DNA helical structure and that has been reported in the literature to be the most favoured orientation in a gas phase calculation. This orientation

involves two weak hydrogen bonds and the corresponding approach is given in figure 4, together with the varying topography. Adenine in the complex has been observed to show the first change in topography with respect to its separated state as given in table 2. The topography changes from [5, 8, 15, 15] to [5, 7, 16, 15]. In contrast to the changes occurring for the pyridine dimer case (where the minima reminiscent of the lone pair of nitrogen changed to a saddle point), here a $(3, +1)$ CP changes to a $(3, -1)$ CP. All this occurs when the distance between the center of masses of the two monomers is about 6 a.u. The distance between N…H and H…N forming the two weak hydrogen bonds of the adenine thymine base pairs are still beyond the usual hydrogen bonding distance.

For the cytosine–guanine base pair, an orientation that would subsequently result in the formation of three hydrogen bonds is considered. In such an approach the first change in the topography of the monomer is observed for guanine, viz. from [4, 5, 13, 13] to [4, 4, 14, 13]. Here the change occurs for a $(3, +1)$ saddle of guanine which changes to a $(3, -1)$ saddle. This change is found to occur when the distances between the center of masses is 6.9 a.u. which is slightly more than the first change in topography in the case of adenine–thymine.

4. Concluding remarks

As prophesied by Emil Fischer more than 100 years ago,[4] the phenomenon of molecular recognition has carried itself over to a molecular level, and it has been often studied for simple molecular systems instead of just interactions in biology. The present study has analysed molecular recognition at a fundamental level employing the scalar field of molecular electrostatic potential. Based on the numerical results of pyridine and benzene dimers in various modes of approach and the DNA base pairs, it is

proposed that point of molecular recognition occurs when the topography of the monomer in the dimer complex changes with respect to that of the isolated monomers. The topography of MESP used in such a manner quantifies electrostatic complimentarity.[19] It also relates to contact between van der Waals surfaces of the monomers, as the changes in topography occur quite close to but beyond the empirically set van der Waals distances. It may be noted that the proposition in the current work offers a definition of molecular recognition via the topography of a scalar field and not in terms of an integrated quantity.

The long-range nature of intermolecular interactions is captured since the MESP dies off as $1/r$. This is exemplified in the case pyridine dimer where the earliest change occurs when the distance between the center of masses is about 15 a.u. Yet another advantage of the current proposition of molecular recognition is that the change in topography relates to change in shape indices of any electron density derived property as was proposed previously by Mezey *et al*.[10] The results presented here are at the HF level, but it has been noted before that the topography of MESP, when mapped at a reasonable basis set (which includes polarization and diffuse functions) is fairly insensitive to correlation and level of theory.[23] Furthermore, solvation effects could be incorporated by employing continuum models, as has been demonstrated in the earlier studies of MESP of solvated systems.[24]

In summary, the present work offers a clear-cut interpretation of how molecules recognize each other at large intermolecular separations, in terms of the topography of the molecular electrostatic potential. It is hoped that this definition finds applicability for a wide variety of chemically and biologically interesting systems.

Acknowledgements

References

1. Lehn J M 1995 *Supramolecular chemistry – concepts and perspectives* (New York: VHC)
2. Gellman S H 1997 *Chem. Rev.* **97** 1231
3. Shinkai S, Masato I, Sugasaki A and Masayuki T 2001 *Acc. Chem. Res.* **34** 494
4. Fischer E 1897 *Ber. Deutsch. Chem. Gesellschaft* **27** 2955; translated by B Holmstedt and G Liljestrand (eds) 1963 *Readings in pharmacology* p 251; (Quoted verbatim: *Their specific effect on the glucosides might thus be explained by assuming that the intimate contact between the molecules necessary for the release of the chemical reaction is possible only with similar geometrical configurations. To give an illustration I will say that enzyme and glucoside must fit together like lock and key in order to be able to exercise a chemical action on each other. This concept has undoubtedly gained in probability and value for stereochemical research, after the phenomenon itself was transferred from the biological to the purely chemical field*)
5. Scrocco E and Tomasi J 1978 In *Adv. Quantum Chem.* (ed.) P -O Löwdin (New York: Academic) vol 11 p. 115
6. Politzer P and Truhlar D G (eds) 1981 *Chemical applications of atomic and molecular electrostatic potentials* (New York: Academic)
7. (a) Gadre S R and Shirsat R N 2000 *Electrostatics of atoms and molecules* (Hyderabad: Universities Press); (b) Mehta G, Gunasekaran G, Gadre S R, Shirsat R N, Ganguly B and Chandrasekhar J 1994 *J. Org. Chem.* **59** 1953; (c) Mehta G, Khan F A, Gadre S R, Shirsat R N, Ganguly B and Chandrasekhar J 1994 *Angew. Chem.* (*Int. Ed. Engl.*) **33** 1390
8. Náráy-Szabó G 2001 *Commun. Math. Comp. Chem.* **44** 297
9. Náráy-Szabó G 2000 *J. Mol. Struct. Theochem.* **500** 157
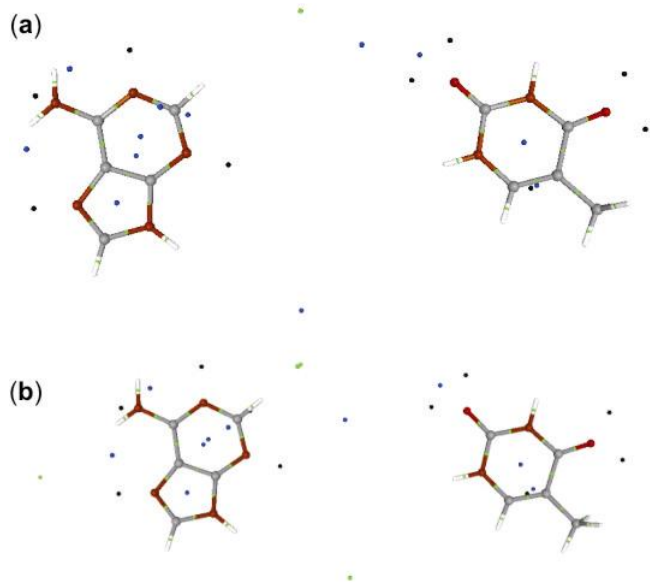
Figure 5. Adenine-thymine geometries along with the (3, +3), (3, +1) and (3, −1) MESP CP's denoted by small black, blue and green spheres respectively. See text and table 2 for more details. Geometries with distances between center of masses of monomers at (**a**) 24·00 a.u., (**b**) 14·00 a.u. respectively.

10. Mezey P G 2001 *J. Math. Chem.* **30** 305
11. Bader R F W 1990 *Atoms in molecules: a quantum theory* (Oxford: Clarendon)
12. Pathak R K and Gadre S R 1990 *J. Chem. Phys.* **93** 1770; (b) Gadre S R, Kulkarni S A and Pathak R K 1991 *J. Chem. Phys.* **94** 8639; (c) Gadre S R and Pathak R K 1990 *Proc. Indian Acad. Sci.* (*Chem. Sci.*) **102** 189
13. Gadre S R 2000 In *Computational chemistry: reviews of current trends* (ed.) J Leszczynski (Singapore: World Scientific) vol 4
14. Frisch M J *et al* 2004 GAUSSIAN 03, Revision C.02, Gaussian Inc., Wallingford CT
15. Program UNIPROP, The molecular property calculation package developed at Theoretical Chemistry Group, Department of Chemistry, University of Pune, Pune. See Bapat S V, Shirsat R N and Gadre S R 1991 *Chem. Phys. Lett.* **200** 373
16. Balanarayan P and Gadre S R 2003 *J. Chem. Phys.* **119** 5037
17. UNIVIS-2000, the molecular properties visualization package, developed at the Theoretical Chemistry Group, Department of Chemistry, University of Pune, Pune; see Limaye A C and Gadre S R 2001 *Curr. Sci.* **80** 1296
18. Leboeuf M, Köster A M, Jug K and Salahub D R 1999 *J. Chem. Phys.* **111** 4893
19. Roy D K, Balanarayan P and Gadre S R 2008 *J. Chem. Phys.* **129** 174103
20. Mishra B K and Sathyamurthy N 2005 *J. Phys. Chem.* **A109** 6
21. Gadre S R and Bhadane P K 1997 *J. Chem. Phys.* **107** 5625
22. Pitoňák M, Neogrády P, Rezak J, Jurecka P, Urban M and Hobza P 2008 *J. Chem. Theory Comput.* **4** 1829
23. Gadre S R, Kulkarni S A, Suresh C H and Shrivastava I H 1995 *Chem. Phys. Lett.* **239** 273
24. Luque F J, Orozco M, Bhadane P K and Gadre S R 1994 *J. Chem. Phys.* **100** 6718