(Submitted to *J. Comput. Chem.*)

Low-energy structures of benzene clusters with a novel accurate potential surface

M. Bartolomei[a)], F. Pirani[b)] and J.M.C. Marques[c)]

[a)] Instituto de Física Fundamental, Consejo Superior de Investigaciones Científicas (IFF-CSIC), Serrano 123, 28006 Madrid, Spain.
E-mail: maxbart@iff.csic.es

[b)] Dipartimento di Chimica, Biologia e Biotecnologie, Università di Perugia, Perugia, Italia.
E-mail: pirani@dyn.unipg.it

[c)] Departamento de Química, Universidade de Coimbra
3004-535 Coimbra, Portugal
E-mail: qtmarque@ci.uc.pt

(Received June , 2015)

Abstract

The benzene-benzene (Bz-Bz) interaction is present in several chemical systems and it is known to be crucial in understanding the specificity of important biological phenomena. In this work, we propose a novel Bz-Bz analytical potential energy surface which is fine-tuned on accurate *ab initio* calculations in order to improve its reliability. Once the Bz-Bz interaction is modelled, an analytical function for the energy of the Bz_n clusters may be obtained by summing up over all pair potentials. We apply an evolutionary algorithm (EA) to discover the lowest-energy structures of Bz_n clusters (for $n = 2 - 25$), and the results are compared with previous global optimization studies where different potential functions were employed. Besides the global minimum, the EA also gives the structures of other low-lying isomers ranked by the corresponding energy. Additional *ab initio* calculations are carried out for the low-lying isomers of Bz_3 and Bz_4 clusters, and the global minimum is confirmed as the most stable structure for both sizes. Finally, a detailed analysis of the low-energy isomers of the $n = 13$ and 19 magic-number clusters is performed. The two lowest-energy Bz_{13} isomers show S_6 and C_3 symmetry, respectively, which is compatible with the experimental results available in the literature. The Bz_{19} structures reported here are all non-symmetric, showing two central Bz molecules surrounded by 12 nearest-neighbour monomers in the case of the five lowest-energy structures.

1 Introduction

Clusters are intermediate entities between a single atom or molecule and the bulk matter. In the last decades, they became ubiquitous in several domains of chemistry and physics,

developing an important role for understanding phenomena such as catalysis, crystal growth, self-assembling, and solvation at the molecular level. This fundamental research has included a detailed exploration of the energy landscapes of several systems, ranging from small atomic and molecular clusters to colloids, glasses and biopolymers (see Ref. 1 and references therein). To achieve this endeavour, one needs a sufficiently accurate potential energy surface (PES), that establishes the relation between the atomic and molecular interactions and the particular shape of the cluster, and a state-of-the-art global optimization method to search for the relevant minima on the PES. The low-energy structures so obtained may be then employed as starting geometries for molecular dynamics simulations or for a post-optimization at a high ab initio level; see, for instance, Refs. 2,3. Although many-body interactions must be taken into account for a rigorous description of the PES, many studies use only the sum of pair-potentials to model the interaction energy in clusters, since it may be considered adequate to capture essential trends of structural properties. Among many analytical pair-potentials proposed in the literature for non-bonding interactions, the Lennard-Jones (LJ) function [4,5] is probably the most commonly employed though it overestimates both the short range repulsion and the long-range attraction energy. In addition, three-body interactions contribute to establish the global minimum structure of metal clusters (see Ref. 6 and references therein) and then can affect also molecular complexes governed by van der Waals forces (see Ref. 7 and references therein).

Interactions involving delocalized π electrons are present, among other, in the compounds which have aromatic rings. These are responsible for the non-polar environment of biological systems and their study is crucial to understand the selectivity of several biochemical phenomena at the molecular level. Because of this, many experimental and theoretical studies on the benzene dimer (Bz_2) have appeared in the literature; see, *e.g.*, Refs. 8–10 and references therein. However, Bz_2 is a very difficult system to treat theoretically, since it shows a shallow

PES mainly governed by the dispersion attraction whose calculation at a correlated level of theory requires computationally demanding resources. Perhaps the first complete *ab initio* potential for Bz_2 was calculated by Cacelli *et al.* [11] at the MP2 level with a double-ζ basis set. Other attempts [8, 9, 12] to obtain a complete PES for the benzene dimer rely on the DFT-based symmetry-adapted perturbation theory [13] (*i.e.*, the SAPT(DFT) approach) with triple-ζ basis set, which allows for relatively accurate results at a low computational cost. These [8, 9, 12] and more accurate CCSD(T) calculations [14, 15], at the complete basis set (CBS) limit, indicate that the T-shape conformer of the benzene dimer (with a center-to-center distance of 4.96Å) is more stable than the displaced-stacked structure; this is in agreement with experimental results [16–18]. However, since both conformers are almost isoergic, it has been observed that the energetic order is quite dependent on the accuracy of the method employed in the calculation; a comprehensive review on benzene dimer calculations involving different methods is given in a recent paper by Xantheas and collaborators [19], where they also present a detailed study about the displaced-stacked conformer at a high level of theory.

Larger benzene clusters have been studied experimentally [20–22] by employing resonance-enhanced two-photon ionization (R2PI) spectroscopy. It became apparent from these data [21, 22] that the structures of the benzene trimer and tetramer are compatible with a C_3 and S_4 symmetry, respectively. In addition, Easter and collaborators [23–27] have employed R2PI spectroscopy to study benzene clusters with up to 19 molecules. For $n = 13$, the $B_{2u} \leftarrow A_{1g}$ 0_0^0 spectrum shows no spectral evidence of the absorption due to the central molecule [23], which may indicate the presence of a high symmetric structure, since such transition is forbidden for C_3 or, even, higher symmetry. In contrast, a doublet arises for the corresponding absorption in the (C_6H_6)-$(C_6D_6)_{12}$ $B_{2u} \leftarrow A_{1g}$ 6_0^1 spectrum, where the two peaks are separated by $1.8\,cm^{-1}$ and have about the same intensity [25]; these results point to the possible

coexistence of two quasi-degenerated isomers with nearly equal populations [25]. Furthermore, Easter [28,29] has carried out Monte Carlo simulations with various isotropic potential energy models for the Bz-Bz interaction in order to establish the main structures arising for Bz_{13} in free jet-expansion experiments. Structures belonging to the S_6, C_3 and C_i symmetry point groups have been identified in the same neighbourhood of the configurational space [28], whereas another C_3 low-energy minimum was located [29] in a distinct region of the PES. Since the interconversion between structures from distant regions of the PES may be hampered due to the energy barriers that are likely to occur, the Monte Carlo results are, thus, in agreement with the hypothesis of a coexistence of two Bz_{13} isomers under experimental conditions [23, 25]. More recently, Chakrabarti *et al.* [30] have also studied the Bz_{13} cluster by employing the anisotropic PES of Totton and co-workers [12] and discovered distinct pairs of low-energy structures with C_3, C_i and S_6 symmetry. In turn, theoretical studies involving large benzene clusters usually employ simple analytical models, since accurate *ab initio* methods cannot be applied to treat so many degrees of freedom; for instance, *ab initio* calculations for benzene clusters up to $n = 10$ have been performed [31, 32] at the MP2/6-31++G** level of theory. Accordingly, the all-atoms Optimized Potentials for Liquid Simulations (OPLS-AA) function [33] has been employed by Takeuchi [10] in a recent systematic global optimization study of $(Bz)_n$ clusters with up to $n = 30$. It is shown that several structures differ significantly from the corresponding ones obtained previously by the same author [34] with the William-Starr potential [35]. Thus, given that such benzene clusters are expected to exhibit properties that are not so close to the liquid state, as described by the OPLS-AA model, other significant structural differences are likely to arise for more adequate potentials.

It is well known that searching for the global minimum, or even other low-energy minima, of chemical systems is a very difficult task [1]. Essentially this can be attributed to the

number of local minima, that are expected to increase exponentially with the size of the system [36–38], and in many cases to the roughness of the PES (*i.e.*, an energy landscape with high barriers connecting the metastable structures). The existence of several minima having similar energies may also make difficult the discovery of the most stable structure. For these reasons, many researchers have been involved in the development of global optimization strategies to discover putative global minima of atomic and molecular clusters [39–48]; see Ref. 49 for a comprehensive review on these methods. In particular, evolutionary algorithms (EAs) have shown to be amongst the most successful global optimization methods. Recently, Marques and collaborators [50] proposed an efficient EA to search for low-energy structures of molecular clusters, which was able to reproduce, among other, all putative global minima of benzene clusters up to $n = 30$.

In this work, we investigate the low-energy landscapes of benzene clusters by exploiting a novel analytical PES for the dimer that is fine-tuned on high-level *ab initio* calculations. Thus, we have calculated a representative set of *ab initio* points for the Bz-Bz interaction by employing the MP2C and CCSD(T) methods at the CBS limit. The obtained counterpoise corrected energies were used to optimize a potential model based on the approach developed by Pirani and collaborators [51–53], suitable to describe consistently both more and less stable configurations of the system. Once the analytical model was obtained for the benzene dimer, we have built up the PES for Bz_n by summing over all Bz-Bz interactions. Then, the EA was employed to discover the low-energy minima (including the global one) of Bz_n clusters. Finally, the interaction energy of the minima discovered for the smaller clusters have been also computed at the MP2C *ab initio* level, and used to check the predictions of EA obtained by exploiting the new potential model.

The plan of the paper is as follows. In Section 2 we describe the methodology applied in this work, including the *ab initio* methods employed for the calculation of the Bz-Bz

interaction energy, the analytical function used to model the Bz_n PES, and the EA for performing the global optimization of the clusters. The energetics and structural features of the Bz_n clusters are presented and discussed in Section 3. Finally, the main conclusions are given in Section 4.

2 Computational procedure

Computational techniques of different type have been used and relevant codes have been exploited to carry out the present investigation in a sequence of steps. As indicated in the previous section, at the beginning the attention has been addressed to extensive theoretical calculations, using state-of-the-art *ab initio* methods, in order to characterize the energy and structure of the benzene dimer in its basic configurations. The following step focused on the development of a semiempirical representation of the intermolecular potential energy surface for the benzene dimer; the comparison between the predictions of the semiempirical model and the *ab initio* results has been a crucial test, which allowed to optimize and generalize the semiempirical potential. Its analytical formulation, providing an internally consistent description of the interaction involved in both the more and less stable configurations of the considered chemical systems (which is a crucial condition to perform any type of molecular dynamics simulation), has been exploited to carry out the search of global minimum structure of Bz_n clusters with the EA. We should emphasize that, since state-of-the-art *ab initio* calculations are computationally demanding, the joint application of the EA with the new analytical potential constitutes an efficient methodology for studying the low-energy landscape of large benzene clusters; all calculations with the EA took several weeks in four Core2Quad computers used on an exclusive basis. Finally, *ab initio* calculations have been also extended to Bz_3 and Bz_4 clusters for a further control of the adopted methodology.

2.1 Ab initio methods

Ab initio calculations of the interaction potential for the benzene clusters were performed at the "coupled" supermolecular second-order Møller-Plesset perturbation level of theory (MP2C) [54]. In order to check the performance of the MP2C results, supermolecular coupled-cluster with single, double and perturbative triple excitations (CCSD(T)) calculations have been also performed for a reduced number of intermolecular distances corresponding to the limiting rigid configurations of the benzene dimer reported in Figure 1. The counterpoise method [55] was applied to correct for the basis set superposition error and all calculations were carried out with the MOLPRO2012.1 package [56]. The reported results refer to the complete basis set (CBS) limit for which the hybrid extrapolation scheme of Sherrill et al. [57] and Jurečka and Hobza [58] is applied. It consists in properly combining counterpoise corrected MP2C (CCSD(T) energies obtained with the aug-cc-pVDZ basis set [59]) with those at the Hartree-Fock (HF) and MP2 levels obtained with the aug-cc-pVDZ and aug-cc-pVTZ basis sets. In Figure 1 a quite good agreement between the MP2C and CCSD(T) results can be observed which confirms the versatility [60–63] of the former approach. The MP2C method was therefore employed to obtain complexation energies also for some larger benzene clusters at the corresponding minimum configurations as predicted by the evolutionary algorithm (see section 3.3).

2.2 Potential model

In molecular simulations addressed to evaluate static and dynamical properties of systems at increasing complexity, an important request is the adoption of an analytical formulation of the PES, defined in terms of a limited numbers of parameters having a physical meaning, in order to guarantee the reliability of the PES both in the most and less stable configurations of the system of interest.

In the present potential formulation Bz is assumed to be a rigid molecule. For the mo-

momer the experimental C-C bond lengths and C-C-C angles were employed (1.390 Å and 120^o, respectively, whereas C-H bond lengths and C-C-H angles were chosen to be 1.09 Å and 120^o, respectively.

For the benzene clusters the interaction potential has been formulated as follows:

$$V = \sum_{k=1}^{n(n-1)/2} (V_{el,\,k} + V_{nel,\,k}) \tag{1}$$

where the first (second) term describes the electrostatic (non-electrostatic) contribution. According to this model, the calculation of $V_{el,\,k}$ assumes a charge distribution that is formed by 12 negative charges $q_C = -0.04623$ a.u. placed on the carbon atoms, above and below of the aromatic ring and separated by 1.905 Å, and 6 positive charges of q_H=+0.09246 a.u. on each hydrogen atom. This reproduces the quadrupole moment of Bz (see for instance Ref. 64) and ensures an accurate description of the quadrupole-quadrupole interactions at long-range Bz-Bz separations. Thus, the electrostatic energy for each Bz-Bz interaction is calculated by employing the Coulomb law through the expression:

$$\begin{aligned} V_{el} \;=\; & \sum_{i=1}^{12}\sum_{j=1}^{12} \frac{q_C q_C}{4\pi\varepsilon_o r_{(Ch)_i(Ch)_j}} + \sum_{i=1}^{12}\sum_{j=1}^{6} \frac{q_C q_H}{4\pi\varepsilon_o r_{(Ch)_i H_j}} \\ & + \sum_{j=1}^{12}\sum_{i=1}^{6} \frac{q_C q_H}{4\pi\varepsilon_o r_{(Ch)_j H_i}} + \sum_{i=1}^{6}\sum_{j=1}^{6} \frac{q_H q_H}{4\pi\varepsilon_o r_{H_i H_j}} \end{aligned} \tag{2}$$

where $r_{(Ch)_i(Ch)_j}$, $r_{(Ch)_i H_j}$ and $r_{H_i H_j}$ represent the distances between two charges of different molecules, respectively, associated to C atoms, to C and H atoms and to H atoms.

In turn, the Bz-Bz non-electrostatic energy is described as a sum of "effective atom"-"effective atom" contributions involving interaction pair-potentials between interaction centers on different Bz molecules, *i.e.*,

$$V_{nel} = \sum_{i=1}^{6}\sum_{j=1}^{6} (V_{C_i - C_j} + V_{C_i - H_j} + V_{H_i - C_j} + V_{H_i - H_j}) \tag{3}$$

Such equation involves the sum of "effective" contributions since each pair potential depends on the behavior of each atom within the Bz molecule whose effective electronic po-

larizability is different from that of the isolated atom since their sum must be compatible with that of the Bz molecule. Each term in Eq. (3) is described by an improved Lennard-Jones (ILJ) potential function [53, 65, 66], which depends on the distance between the two interacting centers (r) according to the expression:

$$\begin{aligned} V_{\mathrm{ILJ}} &= \varepsilon\left[\frac{m}{n(r)-m}\left(\frac{r_0}{r}\right)^{n(r)}\right. \\ &\quad \left.-\frac{n(r)}{n(r)-m}\left(\frac{r_0}{r}\right)^{m}\right] \end{aligned} \tag{4}$$

where ε and r_0 are, respectively, the well depth and equilibrium distance of the related atom-atom pair potential. The m parameter depends on the interaction type, being equal to 6 for neutral-neutral interactions. The $n(r)$ exponent, defining simultaneously the falloff of the atom-atom repulsion and the strength of the attraction, is expressed as,

$$n(r) = \beta + 4.0\left(\frac{r}{r_0}\right)^2 \tag{5}$$

The first term on Eq. (4) describes the size repulsion while the second one is the combination of dispersion and induction attraction. Due to the additional β parameter, the ILJ function becomes more flexible than the traditional LJ model. Hence, the overestimation of the short range repulsion and of long-range attraction by the LJ potential is corrected in the ILJ function, which appears to be important especially when several partners are involved in the formation of aggregates; for instance, it is sufficient to modify the structure of binary rare-gas clusters [67].

Moreover, the relevant well depth (ε) and equilibrium distance (r_0) parameters, directly related to polarizability and charge of involved partners, assume a transferable character when the non-covalent nature of the interaction is maintained, as found in several cases (see Refs. 64–66, 68 and references therein). The effective C and H atomic polarizabilities are nearly the same for hydrocarbon molecules of different kind [65], but the potential parameter

transferability can be partially broken if additional component of chemical nature is operative, as the charge transfer (CT). Such component become effective between the hydrogen atoms (positive centers) on a benzene molecule and carbon atoms(negative centers) on another molecule. Therefore, in the case of the $V_{C_i-H_j}$ and $V_{H_i-C_j}$ terms of Eq. (3) a further contribution has been added to that of Eq. (4) and it has the following exponential form

$$V_{CT} = -Ae^{-\alpha r} \tag{6}$$

as suggested by a detailed comparison of high resolution scattering experiments with *ab initio* calculations on systems giving weak intermolecular hydrogen bond [69]. The inclusion of this additional term allows a better agreement between the *ab initio* reference results and the analytical Bz-Bz potential (see Figure 1), especially for those dimer configurations far from the parallel ones.

Zero order parameters for all pairs of interest have been presented in Ref. 70, where a study of ion-benzene clusters was successfully performed. In the present investigation only the parameters of C-H pairs have been slightly changed (ε increased of about 16%, r_0 decreased of about 6% and β lowered to 6.5 in order to take into account for the effect of Eq. (6) and to optimize the comparison with the *ab initio* results). The other parameters have been taken fixed and considered transferable. Note that the present main target has been not the fitting of *ab initio* results, obtained for some most stable limiting configurations of the system, but, as stressed above, the improvement of the formulation of the full PES. The highest deviations between model predictions and *ab initio* results for energy and equilibrium position of the most stable configurations is here confined within about 15-20% and 4-5% respectively (see Figure 1 for the dimer and subsection 3.3 for trimer and tetramer). The improved PES formulation, given in terms of few parameters, all having a defined physical meaning, is expected to provide a correct representation of the intermolecular interaction in the full configuration space, necessary condition to perform meaningful molecular dynamics

simulations.

The parameters adopted for the atom-atom interactions are given in Table 1.

2.3 Evolutionary algorithm

The evolutionary algorithm (EA) employed for the global optimization of Bz_n clusters was developed by Marques and collaborators [50]. It has been benchmarked for the discovery of the global minima of water, benzene and cation benzene clusters. More recently, good results have been also obtained [2] for $Li^+(H_2O)_n$ and $Li^+(CH_3OH)_n$ (with $n = 1 - 20$) clusters, as modelled with semi-empirical potentials; in fact, new lowest-energy structures were discovered [2] for $Li^+(H_2O)_{17}$, $Li^+(CH_3OH)_6$, and $Li^+(CH_3OH)_7$. In addition, the EA has been also used by two of the present authors to study the microsolvation of the alkali-ions Na^+, K^+ and Cs^+ with either benzene [70] or hexafluorobenzene [71]; global minimum structures of the alkali-ions microsolvation clusters formed by up to 21 benzene molecules were achieved. It is worth noting that the EA was able to distinguish [70] between the global minimum of such microsolvation clusters and several low-energy structures separated by no more than $1\,\mathrm{kJ\,mol^{-1}}$.

Since our EA was fully detailed in the original paper [50], we just give here the main features of the global optimization method. It begins with the random generation of a set of cluster structures, which forms the initial population of individuals (*i.e.*, the tentative solutions). These are relaxed to the nearest local minima by means of the quasi-Newton L-BFGS method [72, 73], that is an efficient gradient-driven local optimization procedure. We note that each individual encodes the location of the n Bz molecules of the cluster by maintaining n sixtuples of the form: $(x_i, y_i, z_i, \alpha_i, \beta_i, \gamma_i)$, where the first three variables define the Cartesian coordinates of the ith Bz center of mass and the remaining three are Euler angles that specify the corresponding orientation. Stochastic tournament selection is applied to choose the most promising individuals (*i.e.*, those with lower interaction energy)

from the current population to be the parents of the next generation. Such parents are, then, subjected to the application of genetic operators that allow to generate new solutions (*i.e.*, the offspring). The EA employs state-of-the-art operators for numeric optimization with prespecified probabilities: first, the simulated binary crossover [74] (SBX) is applied to pairs of parents that, hence, exchange components between them (see Ref. 50 for a detailed description of the arithmetic SBX operator); second, sigma mutation [75] is applied to the clusters that result from crossover, so that the coordinates and/or orientation of a few Bz molecules are slightly modified. Thus, mutated variables are updated with values drawn from a Gaussian distribution with mean 0 and standard deviation σ; note that σ controls the magnitude of the mutation. Before evaluating the quality of the offspring, the L-BFGS method [72, 73] is again applied to drive these clusters to nearest optimum structures. Then, the new generation of individuals replaces the previous population, though an elite operator is applied to prevent the disappearance of the best parent structure. This means that the parent with the lowest-energy replaces the worst offspring in the new population, whenever the quality of the best structure deteriorates from one generation to the other. Such generational iterative process (hereafter designated as a run) proceeds until a prespecified number of potential-function evaluations is reached, and it is repeated several times in order to obtain statistically meaningful results. Following our previous works [50, 70], we have used here the same robust EA setting: number of runs, 30; at least, 1×10^4 function evaluations per run; population size, 100; tourney size, 5; crossover rate, 0.75; mutation rate, 0.1; $\sigma = 0.1$.

The implementation of the method employed in this work uses both energetic and structural dissimilarity criteria to select a pool of distinct low-energy minima for a more detailed analysis. Nonetheless, we do not attempt an exhaustive and systematic search of all minima for each cluster size.

3 Results and discussion

3.1 Benzene dimer

In Figure 1, we show the main features of the benzene dimer interaction for four different geometries which can be considered as the main limiting configurations for this system. They are referred to as "sandwich", "graphite-like", "perp-1" and "perp-2". In the first two the benzene planes are parallel to each other and the interaction energies are depicted as a function of the interplane distance R. The last ones correspond instead to two different configurations of C_{2v} symmetry in which one benzene plane is perpendicular to the other one and in this case R corresponds to the distance between the related centers of mass. It can be seen that a quite good agreement is obtained between that two sets of *ab initio* calculations: MP2C interaction energies almost perfectly reproduce the gold standard CCSD(T)/CBS results and both of them also match the results of Ref.9 (not reported for the sake of simplicity), obtained at the density functional theory-symmetry adapted perturbation theory (SAPT(DFT)) level. The MP2C results have been therefore used as a guide to fine-tune the parameters of the Bz-Bz potential model here employed. Figure 1 also presents the optimized model PES estimations and it can be seen that they reproduce well the size and position of the *ab initio* energy curves even if slight differences can be noticed. In particular, for the present model PES the "perp-2" configuration is close to the global minimum (see next section) and it is predicted to be slightly more stable than "perp-1" one, while the opposite is found for the first principles estimations with a difference of about 10 meV ($\sim$ 1 kJ/mol). Actually we have found that, in agreement with previous studies [9], present *ab initio* findings give a slightly deformed version of the "perp-1" configuration (see the tilted M2 structure in Figure 1 of Ref. 9) as the most stable geometry, being the energy difference between them of about 4.5 meV ($\sim$ 0.45 kJ/mol). Despite these slight discrepancies we believe that present Bz-Bz model PES is accurate enough to provide a reliable description of larger benzene clusters.

3.2 Global minima and structural motifs

In Figure 2, we represent the structures of Bz_n ($n = 2 - 25$) clusters obtained with the EA by employing the new PES proposed in this work; also represented are the symmetry point group of each structure. For Bz_2, the global minimum shows a T-shaped structure (similar to the "perp-2" geometry of Fig. 1) with only one plane of symmetry that contains the top Bz molecule, while the bottom benzene assumes a perpendicular position, though slightly displaced to the left of the C_2 axis (see also Section 3.1). This belongs to the C_s symmetry point group in spite of the perfect T-shaped motif which corresponds to a C_{2v} structure. It is apparent from Figure 2 that, as the cluster increases, the structure tends to maximize the number of T-shape pair-interactions. In some cases, this leads to high-symmetric structures: Bz_4 (S_4), Bz_8 (S_4) and Bz_{13} (S_6); also elements of symmetry other than the identity are present in Bz_3 (C_3) and Bz_6 (C_2). Nonetheless, most of the global minima are non-symmetric C_1 structures.

We represent in Figure 3 the distance ($d_{Bz^c-centre}$) that separates the most central molecule (hereafter designated as Bz^c) from the geometrical centre of the cluster as a function of n. Only for Bz_{13}, one has a global minimum with one molecule placed exactly in the centre of the cluster (see below). This structure has an icosahedral shape, where the central molecule is surrounded by 12 benzene rings. Almost central molecules also appear in the global minimum structures of Bz_{12}, Bz_{14} and Bz_{15}, where $d_{Bz^c-centre}$ is less than 1 Å. In contrast, smaller Bz_n clusters (*i.e.*, with $n \leq 11$) show larger values of $d_{Bz^c-centre}$, since the most favourable orientation of molecules leads to some empty space in the central part of the structure (see Figure 2). As displayed by the crosses for $n = 8 - 11$ in Figure 3, we have found other structures with more central Bz molecules, but they correspond to higher-energy minima (see also below). Also for $n > 13$, $d_{Bz^c-centre}$ tends to increase slower than for small-size clusters and its highest value is less than 3 Å for $n = 25$.

Another relevant piece of information related to the solvation of the most central molecule by the other ones is shown in Figure 4. We represent in this figure a scatter plot of the distances ($d_{\mathrm{Bz-Bz^c}}$) between the most central ring and the remaining molecules of the cluster. For all clusters, there is a set of distances that are less than 7 Å, which correspond to the nearest-neighbour molecules. Larger distances that may be associated to a second shell appear for $n = 8 - 11$ and, also, for $n > 14$. Curiously, this second shell is missing for $n = 12$ and 13. We should emphasize that the icosedral-type geometry of Bz_{13} with a molecule in the centre of the cluster is expected to close the first solvation shell, which indicates an "abnormal" second-shell pattern observed for Bz_8, Bz_9, Bz_{10}, and Bz_{11}. Besides the global minimum, the EA was also able to discover other low-energy structures (corresponding to the above mentioned points displayed by the crosses in Figure 3) with no second-shell molecules. All such structures show C_1 symmetry and their energies fall above the the corresponding global minimum by 2.188 kJ mol^{-1}, 1.102 kJ mol^{-1}, 1.535 kJ mol^{-1}, and 1.781 kJ mol^{-1}, respectively. These structures are the second (third) lowest-energy ones for Bz_8 and Bz_9 (Bz_{10}). By contrast, the EA has discovered for Bz_{11} eight lower-energy C_1 structures, besides the global minimum, that show molecules in the second-shell. It is worth noting that the presence of such a great number of low-symmetry isomers has been previously advanced to explain a "notable absence of definition" in the line shape spectrum of Bz_{11} cluster [25]. Finally, it is apparent from Figure 4 that $d_{\mathrm{Bz-Bz^c}}$ values greater than 9.5 Å appear for $n = 17 - 25$, since it becomes difficult to accomodate more molecules in the first and second shells of these clusters. Although this may be considered the onset of a third shell, the range of distances is not so distinct as for the first and second ones.

We now compare the present results with other global optimization studies [10, 34, 50] on Bz_n clusters modelled with either the Williams-Starr (WS) potential [35] or the all-atom OPLS-AA force field [33]. To facilitate the discussion, we designate hereafter the novel

potential as BPM (*i.e.*, after the names of the present authors). For a global comparison between pairs of structures, we apply the previously developed SAICS program [76], which obtains the best overlap of the two clusters and calculates the corresponding root-mean square deviation (RMSD); hence, it is possible to identify whether two structures are equal or different (and how different). In comparison with the similarity descriptor employed in Ref. 10, the present method based on the RMSD calculation is more time-consuming, but it has the advantage to give always reliable results (see Ref. 76 for a detailed discussion on several similarity descriptors). In Figure 5, we represent the RMSD values resulting from the comparison of the global minimum structures obtained in this work with those from OPLS-AA (solid line) and WS (dashed line) potentials as a function of the cluster size. We observe from this figure that the largest structural differences between our global minima and those of both WS and OPLS-AA models occur for Bz_{11}, Bz_{17}, Bz_{21} and Bz_{24} (the latter only for the comparison with WS). These are among the 11 (10) structures with RMSD values greater than 1 Å for the comparison with the WS (OPLS-AA) global minima. More similar structures (with RMSD values less than 0.2 Å) are found instead for 9 (7) global minima, respectively. In addition, Figure 6 performs a local-structure comparison between the three sets of global minima. In this analysis, we have considered two parameters: the maximum number of nearest-neighbour contacts (N_{max}) among the monomers of the cluster and the number of Bz molecules with $N_{max} = 12$ (which is expected to be the highest value of N_{max} for Bz_n structures). It is apparent from Figure 6 that, among the three potentials, small differences in the values of N_{max} occur for some global minima in the cluster size range $6 \leq n \leq 11$; also for $n = 22$ the WS PES leads to a global minimum with $N_{max} = 11$, while the corresponding value for the other two potentials is 12. Concerning the number of Bz molecules with $N_{max} = 12$, differences in the global minima of the three potentials arise only for $n \geq 22$. Whilst the OPLS-AA potential leads to global minima with 3 molecules

surrounded by 12 neighbours for $n \geq 23$, the structures from the present PES only reach that value for $n = 25$. A more striking behaviour is observed in the case of the WS global minima for which the number of molecules with $N_{max} = 12$ is two for Bz_{21}, decreases to zero for Bz_{22}, increases to three for Bz_{23}, and recovers the value of two for Bz_{24} and Bz_{25}. It is worth noting from the comparison between Figure 5 and Figure 6 that a big difference in RMSD values may not be reflected in the local-structure parameters considered here. On the other side, there are clusters with different values of one of the local-structure parameters that have relatively small values of the RMSD. However, we should emphasize that the differences in the local-structure features are always very small.

The binding energy of the Bz_n clusters is calculated as

$$E_n = -\frac{E_{\mathrm{GM},\,n}}{n} \tag{7}$$

where $E_{\mathrm{GM},\,n}$ is the energy of the global minimum at the cluster size n. In the top panel of Figure 7, we compare the binding energies of the Bz_n clusters given by the BPM, OPLS-AA and WS potentials. While the OPLS-AA and WS functions show similar binding energies for all cluster sizes, the BPM model gives always higher values of E_n and the difference for the former potentials tends to increase with n. In turn, the relative stability of the clusters may be estimated by the second energy difference, *i.e.*,

$$\Delta_2 E = E_{\mathrm{GM},\,n+1} + E_{\mathrm{GM},\,n-1} - 2E_{\mathrm{GM},\,n} \tag{8}$$

where $E_{\mathrm{GM},\,n+1}$, $E_{\mathrm{GM},\,n-1}$ and $E_{\mathrm{GM},\,n}$ are the global minimum energies at $n+1$, $n-1$ and n sizes, respectively. Maxima of $\Delta_2 E$ are usually designated as "magic numbers" and correspond to particularly stable structures in comparison with their neighbour sizes. It is apparent in the bottom panel of Figure 7 that a prominent peak appears for the three potentials at $n = 13$. Another magic number also arises for $n = 19$ with the BPM and OPLS-AA potentials; in the case of the WS potential, the maximum is broader. Finally,

another peak appears at $n = 23$, but it is not so prominent for the BPM potential. In the case of OPLS-AA, the magic numbers have been attributed [10] to the stepwise increase of the number of benzene molecules with $N_{max} = 12$ for those cluster sizes (see Figure 6); an increase of such local structures at the magic-number sizes is also observed in Figure 6 for WS potential. A similar trend is observed with the BPM potential for $n = 13$ and 19, but the magic number for $n = 23$ has 2 benzene molecules with $N_{max} = 12$ that are the same number as for the neighbour sizes. Nonetheless, from a more detailed analysis of the structure of the clusters, we observed that some Bz-Bz distances are only slightly larger than the nearest-neighbour criterion followed in the present work (*i.e.*, 6 Å); by taking this into consideration, one gets 2 (3) benzene molecules with $N_{max} = 12$ for $n = 22$ ($n = 23$ and $n = 24$), which is the same pattern as that reported in Ref. 10.

3.3 Ab initio energy order for the structures of Bz_3 and Bz_4

In order to confirm the reliability of the model PES even for clusters larger than dimers we have obtained a validity test by performing ab initio calculations also for trimers and tetramers. In particular, the main structures of the Bz_n ($n = 3$ and 4) aggregates have been also investigated at the MP2C/CBS level of theory and single-point interaction energies starting from the geometries corresponding to the most stable minima predicted by the EA have been computed.

For trimers the lowest minimum corresponds to a cyclic structure of C_3 symmetry (see Figure 2) labelled as GM in Table 2. We have found that this configuration is favoured over linear benzene chains of C_{2v} symmetry in which one monomer is perpendicular to the next one (as in the "perp-2" configuration of Figure 1). These linear configurations correspond indeed to the relative minima LM2, LM3 and LM4 whose interaction energies are reported in Table 2 and they are found to be roughly 100 meV lower (in absolute value) than that of the GM structure. A good agreement between the ab initio and EA interaction energies can

be also seen and the same stability order is obtained.

As for tetramers, none the obtained most stable geometries are linear and the the global minimum corresponds to a cyclic and packed structure of S_4 symmetry in which each monomer is able to closely interact with the next ones in relative configurations which are slightly distorted with respect to the perpendicular "perp-1" and "perp-2" geometries found for the dimer (see Figure 1). We have also checked that for tetramers the MP2C computations are able to confirm the global minimum as the most stable among those predicted by the EA optimization.

3.4 Detailed analysis of $n = 13$ and 19

Since both Bz_{13} and Bz_{19} clusters are magic numbers for the global minima, we have performed a detailed analysis of other low-energy structures in order to evaluate their relative stability. Among the Bz_n clusters (with $n > 2$), Bz_{13} is perhaps the most studied [23,25,28–30] one. The spectroscopic experiments [23, 25] were not totally conclusive about the geometry of Bz_{13}, since they indicated the presence of a C_3 (or higher symmetry) structure, on one side, and a less symmetric isomer by the other side. To explain this structural ambiguity, Easter [28] has hypothesized the coexistence of two isomers with different symmetries in the free jet expansions; hence, it has been advanced [29] that the presence of isomers with C_3, C_i and S_6 symmetry is compatible with the experiments. Although most of the PESs available in the literature predict the existence of only one of those symmetric structures, the OPLS-AA potential proposed by Jorgensen and Severance [33] has two distinct C_3 minima and a C_i structure [29]; a C_3 global minimum structure has been recently confirmed [10] for the OPLS-AA model. In turn, the PES developed recently by Totton *et al.* [12] has shown to present three pairs of symmetric structures belonging to C_3 (which includes the global minimum), C_i and S_6 point groups [30]. These six structures lie within an energy range of $1.1\,\mathrm{kJ\,mol^{-1}}$ and the energy landscape shows a barrier higher than $5\,\mathrm{kJ\,mol^{-1}}$ between a S_6

isomer and the C_3 global minimum that may be able to work as a kinetic trap [30]. Thus, the Totton PES is consistent with the coexistence of two symmetric isomers (C_3 and S_6) which was proposed [28] to explain the experiments [23, 25]. As displayed in Figure 8, the new PES developed in the present work has also three symmetric low-energy minima (S_6, C_3 and C_i) in a range of $\sim 7\,\mathrm{kJ\,mol^{-1}}$. In particular, the energy difference between the global minimum (with S_6 symmetry) and the C_3 structure is about $3\,\mathrm{kJ\,mol^{-1}}$, which is also compatible with the idea of having two isomers with similar populations. In addition, it is shown in Figure 8 that other low-energy C_1 local minima appear for Bz_{13} in the new PES. All such structures have a benzene molecule in the centre and correspond to distorted icosahedra, which is an evidence of the strong stability of this motif for Bz_{13}.

As for Bz_{19}, the six lowest-energy structures discovered with the EA for the BPM potential are shown in Figure 9. Although all minima display C_1 symmetry, the five lowest-energy isomers show two benzene molecules in the middle of the structure that are surrounded by 12 neighbours; conversely, one of the interior benzenes of LM6 is surrounded by 12 molecules, while the other has only 11 neighbours, thus, indicating a more distorted structure. Accordingly, all the five lowest-energy structures fall in a range of about $2.576\,\mathrm{kJ\,mol^{-1}}$, while a gap of $1.292\,\mathrm{kJ\,mol^{-1}}$ is observed between LM5 and LM6. In addition, it is apparent from Figure 9 that global minimum and LM2 structures are very similar. Indeed, the RMSD value calculated for the best overlap between the two structures is 0.39 Å, and the corresponding energy difference is less than $0.1\,\mathrm{kJ\,mol^{-1}}$. The possibility of having such a pair of similar structural isomers may be an explanation for the experimental observation of two equal-intensity singlet peaks in the R2PI spectrum for this cluster size [25]. This experimental achievement has been attributed [10] to the two non-equivalent interior molecules of the OPLS-AA global minimum, which is also very similar to the BPM one (*cf.* Figure 5).

4 Conclusions

We have performed a detailed global geometry optimization by employing our own EA. To describe the Bz-Bz interaction, we have proposed a new PES that is fine-tuned on *ab initio* calculations carried out at the CCSD(T)/CBS level of theory. A T-shaped structure has been established as the global minimum, though the graphite-like structure does not show a large energy separation. This result is in line with previous theoretical investigations [8, 9, 12]. In contrast, the OPLS-AA and WS previous potentials (that have been used for comparison) show distinct global minimum structures for the dimer. Whilst the former presents a tilted T-shaped structure, the two benzene molecules tend to be more stacked for the WS potential. Accordingly, the global minimum of Bz_n ($n = 3 - 25$) modelled with the present potential function have structures that are distinct from the OPLS-AA and WS ones for most of cluster sizes. The binding energy of the Bz_n clusters is always higher in the case of the present potential, while the corresponding values for OPLS-AA and WS global minima are quite similar. Nonetheless, the three potentials show the same magic number structures, namely at $n = 13$, 19 and 23 (though less prominent for the new one). This has been rationalized on the basis of the number of nearest-neighbour molecules that surround the most central monomers. The maximum value of such number (*i.e.*, N_{max}) is 12 for $n \geq 13$, as expected for benzene clusters. Most of the differences in the values of N_{max} among the three potentials occur for $6 \leq n \leq 11$, where competition between structures with and without molecules on the second-shell is observed for some cluster sizes. In addition, we have detected a large number of low-energy minima with C_1 symmetry for Bz_{11}, which is compatible with the absence of definition in the line shape spectrum of this cluster [25].

It is worth noting that the energies of the four low-lying minima of Bz_3 recalculated at the MP2C/CBS level of theory follow the same order as given by the present empirical potential; also, similar calculations for a set of low-energy minima obtained with the EA

for Bz_4 has confirmed the global minimum as the most stable structure. In turn, we have reported 9 (6) low-energy minimum structures of Bz_{13} (Bz_{19}). Whereas all Bz_{19} clusters are non-symmetric C_1 structures, there are three low-energy minima of Bz_{13} that belong to the S_6 (the global minimum), C_3 and C_i point groups of symmetry. The presence of the lowest-energy S_6 and C_3 minima for Bz_{13} reinforces the idea of having a coexistence of two symmetric isomers in the free jet expansions, which has been proposed to explain the R2PI spectra [23, 25]. In summary, the above mentioned agreement with available experimental data confirms the validity of the theoretical methodology employed in this work. However, for selected stable configurations of larger Bz_n clusters, some validation calculations, carried out at MP2C level, can be useful to further test the performance of the present method with respect to previous models. Finally, such a simple methodology may be extended to the study of systems involving poly-aromatic hydrocarbons of increasing complexity in future investigations.

Acknowledgments

J.M.C.M. acknowledges the financial support from the Coimbra Chemistry Centre (UID/QUI/00313/2013). We are grateful for the provision of computational time in the supercomputer resources hosted at Laboratório de Computação Avançada, Universidade de Coimbra. M.B. acknowledges financial support by the Spanish "Ministerio de Ciencia e Innovacion" for the FIS2013-48275-C2-1-P grant. Allocation of computing time by CESGA (Spain) is also acknowledged. F.P. acknowledges financial support from the Italian Ministry of University and Research (MIUR) for PRIN 2010-2011, grant 2010 ERFKXL_002.

References

[1] Wales, D. J., *Energy Landscapes: With Applications to Clusters, Biomolecules and Glasses*, Cambridge University Press, Cambridge, UK, 2003.

[2] Llanio-Trujillo, J. L.; Marques, J. M. C. and Pereira, F. B., *Comput. Theor. Chem.*, 2013, **1021**, 124–134.

[3] Albertí, M.; Aguilar, A. and Marques, J. M. C., *Eur. Phys. J. D*, 2014, **68**, 364.

[4] Lennard-Jones, J. E., *Proc. Roy. Soc. A*, 1924, **106**, 463.

[5] Lennard-Jones, J. E., *Proc. Phys. Soc.*, 1931, **43**, 461–482.

[6] Wales, D. J.; Munro, L. J. and Doye, J. P. K., *Dalton Trans.*, 1996, pages 611–623.

[7] R. A. DiStasio, J.; Gobre, V. V. and Tkatchenko, A., *J. Phys.: Condens. Matter*, 2014, **26**, 213202.

[8] Podeszwa, R.; Bukowski, R. and Szalewicz, K., *J. Phys. Chem. A*, 2006, **110**, 10345–10354.

[9] van der Avoird, A.; Podeszwa, R.; Szalewicz, K.; Leforestier, C.; van Harrevelt, R.; Bunker, P. R.; Schnell, M.; von Helden, G. and Meijer, G., *Phys. Chem. Chem. Phys.*, 2010, **12**, 8219–8240.

[10] Takeuchi, H., *J. Phys. Chem. A*, 2012, **116**, 10172–10181.

[11] Cacelli, I.; Cinacchi, G.; Prampolini, G. and Tani, A., *J. Am. Chem. Soc.*, 2004, **126**, 14278–14286.

[12] Totton, T. S.; Misquitta, A. J. and Kraft, M., *J. Chem. Theor. Comput.*, 2010, **6**, 683–695.

[13] Jeziorski, B.; Moszyński, R. and Szalewicz, K., *Chem. Rev.*, 1994, **94**, 1887–1930.

[14] Lee, E. C.; Kim, D.; Jurecka, P.; Tarakeshwar, P.; Hobza, P. and Kim, K. S., *J. Phys. Chem. A*, 2007, **111**, 3446–3457.

[15] Pitoňák, M.; Neogrady, P.; Rezac, J.; Jurecka, P.; Urban, M. and Hobza, P., *J. Chem. Theory Comput.*, 2008, **4**, 1829–1834.

[16] Kusaka, R.; Inokuchi, Y. and Ebata, T., *J. Chem. Phys.*, 2012, **136**, 044304.

[17] Arunan, E. and Gutowsky, H. S., *J. Chem. Phys.*, 1993, **98**, 4294–4296.

[18] Henson, B. F.; Hartland, G. V.; Venturo, V. A. and Felker, P. M., *J. Chem. Phys.*, 1992, **97**, 2189–2208.

[19] Miliordos, E.; Aprà, E. and Xantheas, S. S., *J. Phys. Chem. A*, 2014, **118**, 7568–7578.

[20] Börnsen, K. O.; Lin, S. H.; Selzle, H. L. and Schlag, E. W., *J. Chem. Phys.*, 1989, **90**, 1299–1306.

[21] Iimori, T.; Aoki, Y. and Ohshima, Y., *J. Chem. Phys.*, 2002, **117**, 3675–3686.

[22] Iimori, T. and Ohshima, Y., *J. Chem. Phys.*, 2002, **117**, 3656–3674.

[23] Easter, D. C.; Whetten, R. L. and Wessel, J. E., *J. Chem. Phys.*, 1991, **94**, 3347–3354.

[24] Easter, D. C.; Li, X. and Whetten, R. L., *J. Chem. Phys.*, 1991, **95**, 6362–6370.

[25] Easter, D. C.; Khoury, J. T. and Whetten, R. L., *J. Chem. Phys.*, 1992, **97**, 1675–1682.

[26] Easter, D. C.; Baronavski, A. P. and Hawley, M., *J. Chem. Phys.*, 1993, **99**, 4942–4951.

[27] Easter, D. C.; Mellot, J. and Weiss, T., *J. Chem. Phys.*, 1998, **109**, 8365–8373.

[28] Easter, D. C., *J. Phys. Chem. A*, 2003, **107**, 2148–2159.

[29] Easter, D. C., *J. Phys. Chem. A*, 2003, **107**, 7733–7742.

[30] Chakrabarti, D.; Totton, T. S.; Kraft, M. and Wales, D. J., *Phys. Chem. Chem. Phys.*, 2011, **13**, 21362–21366.

[31] Mahadevi, A. S.; Rahalkar, A. P.; Gadre, S. R. and Sastry, G. N., *J. Chem. Phys.*, 2010, **133**, 164308.

[32] Yeole, S. D. and Gadre, S. R., *J. Chem. Phys.*, 2011, **134**, 084111.

[33] Jorgensen, W. L. and Severance, D. L., *J. Am. Chem. Soc.*, 1990, **112**, 4768–4774.

[34] Takeuchi, H., *J. Chem. Inf. Model.*, 2007, **47**, 104–109.

[35] Williams, D. E. and Starr, T. L., *Comput. Chem.*, 1977, **1**, 173–177.

[36] Hoare, M. R. and McInnes, J., *Faraday Discuss. Chem. Soc.*, 1976, **61**, 12–24.

[37] Tsai, C. J. and Jordan, K. D., *J. Chem. Phys.*, 1993, **99**, 6957–6970.

[38] Stillinger, F. H., *Phys. Rev. E*, 1999, **59**, 48.

[39] Deaven, D. M. and Ho, K. M., *Phys. Rev. Lett.*, 1995, **75**, 288–291.

[40] Gregurick, S. K.; Alexander, M. H. and Hartke, B., *J. Chem. Phys.*, 1996, **104**, 2684–2691.

[41] Wales, D. J. and Doye, J. P. K., *J. Phys. Chem. A*, 1997, **101**, 5111–5116.

[42] Niesse, J. A. and Mayne, H. R., *J. Comput. Chem.*, 1997, **18**, 1233–1244.

[43] Roberts, C.; Johnston, R. L. and Wilson, N. T., *Theor. Chem. Acc.*, 2000, **104**, 123–130.

[44] Shao, X.; Cheng, L. and Cai, W., *J. Comput. Chem.*, 2004, **25**, 1693.

[45] Pereira, F. B.; Marques, J. M. C.; Leitão, T. and Tavares, J. In *Proceedings of the 2006 IEEE Congress on Evolutionary Computation, Vols. 1-6*, pages 2270–2277, Vancouver, 2006. CEC.

[46] Takeuchi, H., *J. Chem. Inf. Model.*, 2006, **46**, 2066–2070.

[47] Rossi, G. and Ferrando, R., *J. Phys.: Condens. Matter*, 2009, **21**, 084208.

[48] Marques, J. M. C. and Pereira, F. B., *Chem. Phys. Lett.*, 2010, **485**, 211–216.

[49] Marques, J. M. C. and Pereira, F. B., *J. Mol. Liq.*, in press; doi: 10.1016/j.molliq.2015.03.009

[50] Llanio-Trujillo, J. L.; Marques, J. M. C. and Pereira, F. B., *J. Phys. Chem. A*, 2011, **115**, 2130–2138.

[51] Albertí, M.; Aguilar, A.; Lucas, J. M.; Pirani, F.; Cappelletti, D.; Coletti, C. and Re, N., *J. Phys. Chem. A*, 2006, **110**, 9002–9010.

[52] Albertí, M.; Aguilar, A.; Lucas, J. M.; Cappelletti, D.; Laganà, A. and Pirani, F., *Chem. Phys.*, 2006, **328**, 221–228.

[53] Pirani, F.; Brizi, S.; Roncaratti, L. F.; Casavecchia, P.; Cappelletti, D. and Vecchiocattivi, F., *Phys. Chem. Chem. Phys.*, 2008, **10**, 5489.

[54] Pitonák, M. and Hesselmann, A., *J. Chem. Theory Comput.*, 2010, **6**, 168–178.

[55] Boys, S. F. and Bernardi, F., *Mol. Phys.*, 1970, **19**, 553–566.

[56] Werner, H. J.; Knowles, P. J.; Lindh, R.; Manby, F. R.; Schütz, M. and some others, , Molpro, version 2012.1, a package of ab initio programs, 2012.

[57] Sinnokrot, M. O.; Valeev, E. F. and Sherrill, C. D., *J. Am. Chem. Soc.*, 2002, **124**, 10887–10893.

[58] Jurečka, P. and Hobza, P., *Chem. Phys. Lett.*, 2002, **365**, 89–94.

[59] Kendall, R. A.; Dunning, Jr., T. H. and Harrison, R. J., *J. Chem. Phys.*, 1992, **96**, 6796–6806.

[60] Hesselmann, A. and Korona, T., *Phys. Chem. Chem. Phys.*, 2011, **13**, 732–743.

[61] Bartolomei, M.; Carmona-Novillo, E.; Hernández, M. I.; Campos-Martínez, J. and Pirani, F., *J. Phys. Chem. C*, 2013, **117**, 10512–10522.

[62] Bartolomei, M.; Carmona-Novillo, E.; Hernández, M. I.; Campos-Martínez, J.; Pirani, F.; Giorgi, G. and Yamashita, K., *J. Phys. Chem. Lett.*, 2014, **5**, 751–755.

[63] Bartolomei, M.; Carmona-Novillo, E.; Hernández, M. I.; Campos-Martínez, J.; Pirani, F. and Giorgi, G., *J. Phys. Chem. C*, 2014, **117**, 10512–10522.

[64] Albertí, M.; Castro, A.; Laganà, A.; Moix, M.; Pirani, F.; Cappelletti, D. and Liuti, G., *J. Phys. Chem. A*, 2005, **109**, 2906–2911.

[65] Pirani, F.; Albertí, M.; Castro, A.; Teixidor, M. M. and Cappelletti, D., *Chem. Phys. Lett.*, 2004, **394**, 37–44.

[66] Faginas-Lago, N.; Huarte-Larrañaga, F. and Albertí, M., *Eur. Phys. J. D*, 2009, **55**, 75–85.

[67] Marques, J. M. C. and Pereira, F. B., *J. Comput. Chem.*, 2013, **34**, 505–517.

[68] Albertí, M.; Aguilar, A.; Lucas, J. M. and Pirani, F., *J. Phys. Chem. A*, 2010, **114**, 11964–11970.

[69] Cappelletti, D.; Ronca, E.; Belpassi, L.; Tarantelli, F. and Pirani, F., *Acc. Chem. Res.*, 2012, **45**, 1571–1580.

[70] Marques, J. M. C.; Llanio-Trujillo, J. L.; Albertí, M.; Aguilar, A. and Pirani, F., *J. Chem. Phys. A*, 2012, **116**, 4947–4956.

[71] Marques, J. M. C.; Llanio-Trujillo, J. L.; Albertí, M.; Aguilar, A. and Pirani, F., *J. Chem. Phys. A*, 2013, **117**, 8043–8053.

[72] Nocedal, J., *Math. Comp.*, 1980, **35**, 773–782.

[73] Liu, D. and Nocedal, J., *Math. Program. B*, 1989, **45**, 503–528.

[74] Deb, K. and Beyer, H.-G., *Evol. Comput.*, 2001, **9**, 197–221.

[75] Pereira, F. B.; Marques, J. M. C.; Leitão, T. and Tavares, J. In Siarry, P. and Michalewicz, Z., Eds., *Advances in Metaheuristics for Hard Optimization, Springer Natural Computing Series*, pages 223–250, Berlin, 2008. Springer.

[76] Marques, J. M. C.; Llanio-Trujillo, J. L.; Abreu, P. E. and Pereira, F. B., *J. Chem. Inf. Model.*, 2010, **50**, 2129–2140.

[77] Humphrey, W.; Dalke, A. and Schulten, K., *J. Molec. Graphics*, 1996, **14**, 33–38.

Table 1: **Bz-Bz analytical potential parameters: well depth[a] (ε), equilibrium distance (r_0), β and m of the ILJ potential for the involved atom pairs; pre-expoenential A and exponential α coefficients for the additional V_{CT} interaction term to be taken into account for the C-H atom pairs.**

ILJ interaction	ε / meV	r_0/ Å	β	m
C-C	3.340	4.073	9.0	6
C-H	2.000	3.505	6.5	6
H-H	1.610	3.099	9.0	6
V_{CT} interaction	A / meV	α/ Å^{-1}		
C-H	$1.9{\cdot}10^4$	2.9		

[a] 1 meV= 0.0964853 kJ mol^{-1}

Table 2: **Interaction energies for the global minimum (GM) and lowest local minima (LMn) of the Bz trimer as obtained by the EA optimization. Corresponding ab initio estimations are also reported.**

Trimer clusters	EA optimization / meV	Ab initio (MP2C/CBS) / meV
GM	-384.5	-347.5
LM2	-260.3	-226.4
LM3	-258.4	-222.8
LM4	-258.3	-221.6

[a] 1 meV= 0.0964853 kJ mol^{-1}

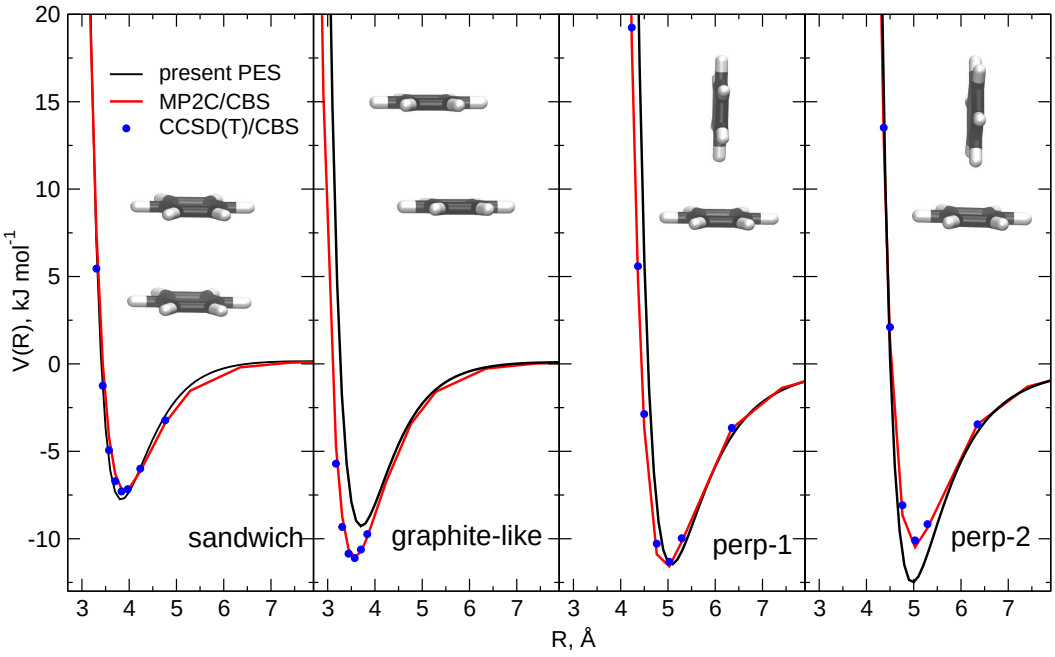

Figure 1: Ab initio and present model results for the interaction energy of the benzene dimer in four different limiting configurations (symmetry point groups in parentheses): "sandwich" (D_{6h}), "graphite-like" (C_{2h}), "perp-1" (C_{2v}) and "perp-2" (C_{2v}). The first two geometries correspond to stacked benzene molecule and R describes the interplane distance. In the "graphite-like" geometry one of the benzene molecule is displaced with respect to the other to mimick graphite equilibrium stacking. "perp-1" and "perp-2" correspond to geometries where one benzene perpendiculary approaches the other one and R describes the distance between the centers of mass.

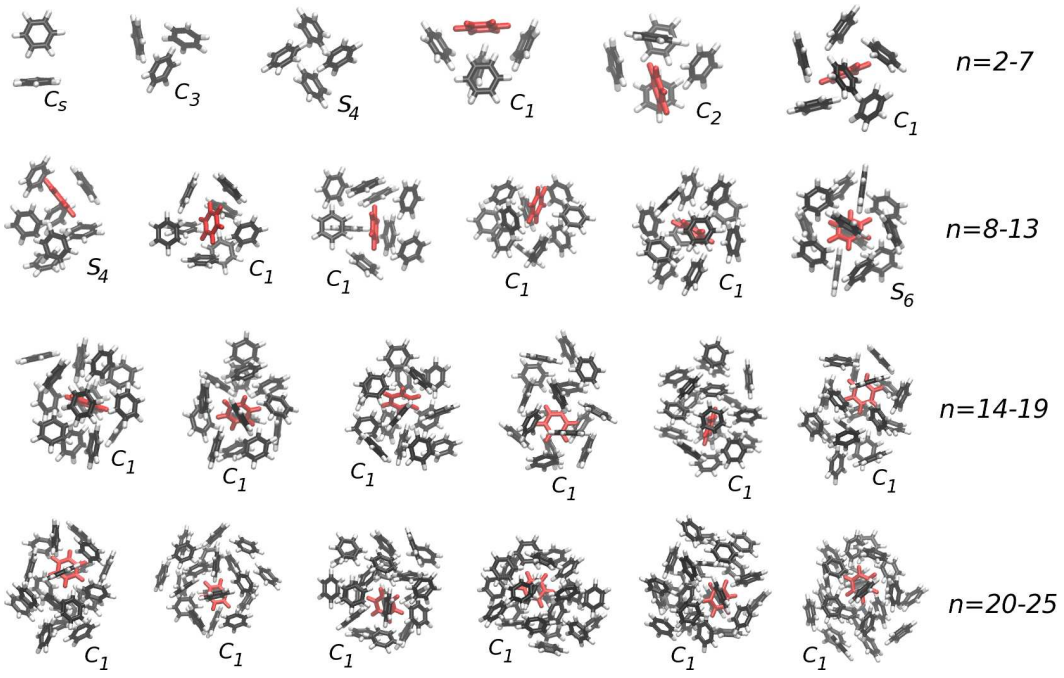

Figure 2: Global minimum structures of the Bz_n clusters. Plots were produced with VMD program [77].

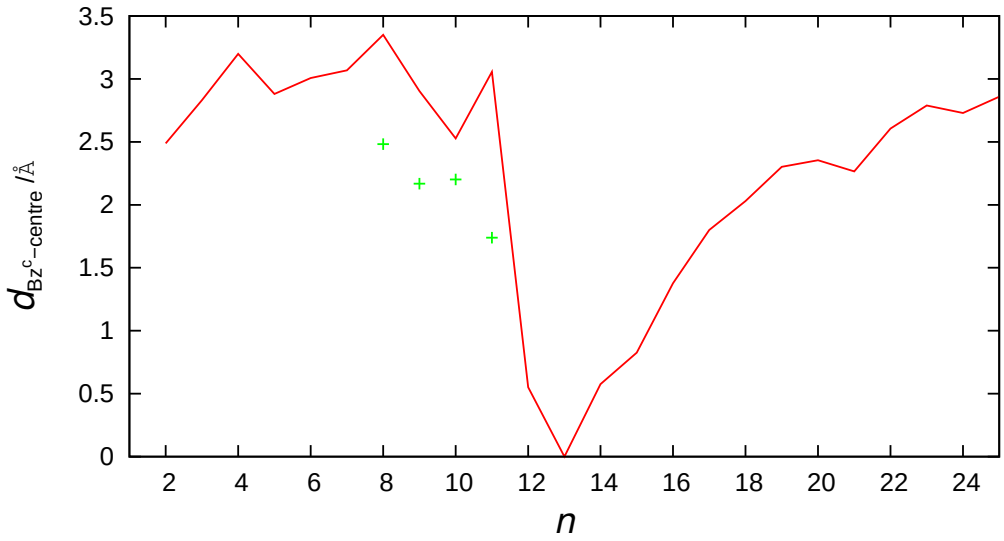

Figure 3: Distance between the most central molecule and the geometrical centre of the cluster. Points represented by crosses for $n = 8 - 11$ correspond to higher-energy structures with one benzene molecule which is more central than in the global minimum. Such structures are also more compact than the corresponding lowest-energy ones and, hence, they show no molecules in the second solvation shell (conversely to the global minimum structures, as displayed in Fig. 4). See the text.

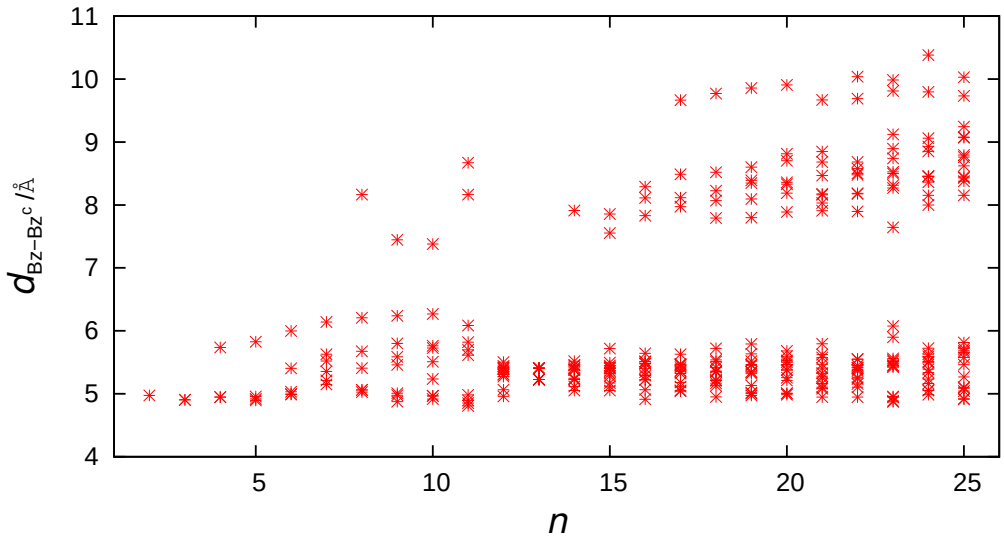

Figure 4: Scatter plot of the distances between the center of each Bz molecule and the most central ring (*i.e.*, Bz^c) of the cluster. See text.

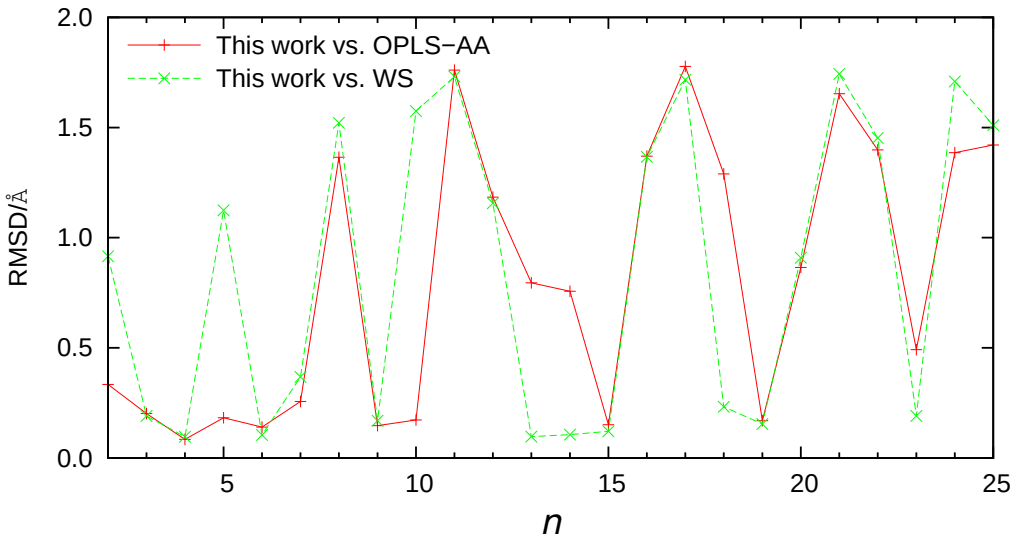

Figure 5: Root-mean square distances (RMSD) for the best overlap between the present structures and the corresponding ones from OPLS-AA (Ref. 10) and William & Starr (Ref. 35) potentials. These values have been calculated by employing the SAICS (superimposing algorithm for the identification of chiral structures) program [76].

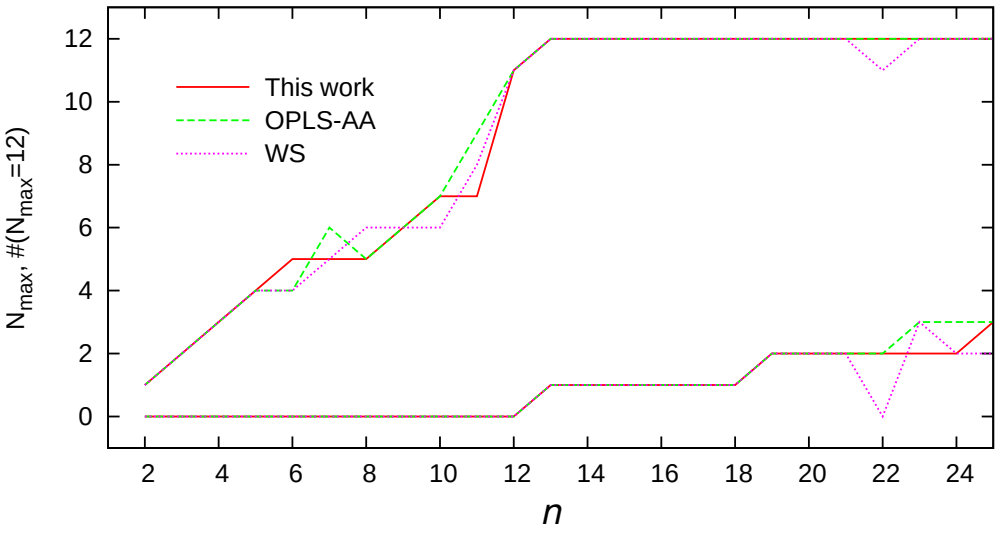

Figure 6: Local structure analysis of Bz_n global minima. Top lines represent the maximum number of nearest neighbors (N_{max}), while the bottom ones are for the number of Bz molecules with $N_{max} = 12$. A distance less than 6 Å between the centre of the rings has been adopted as the criterion for nearest-neighbours.

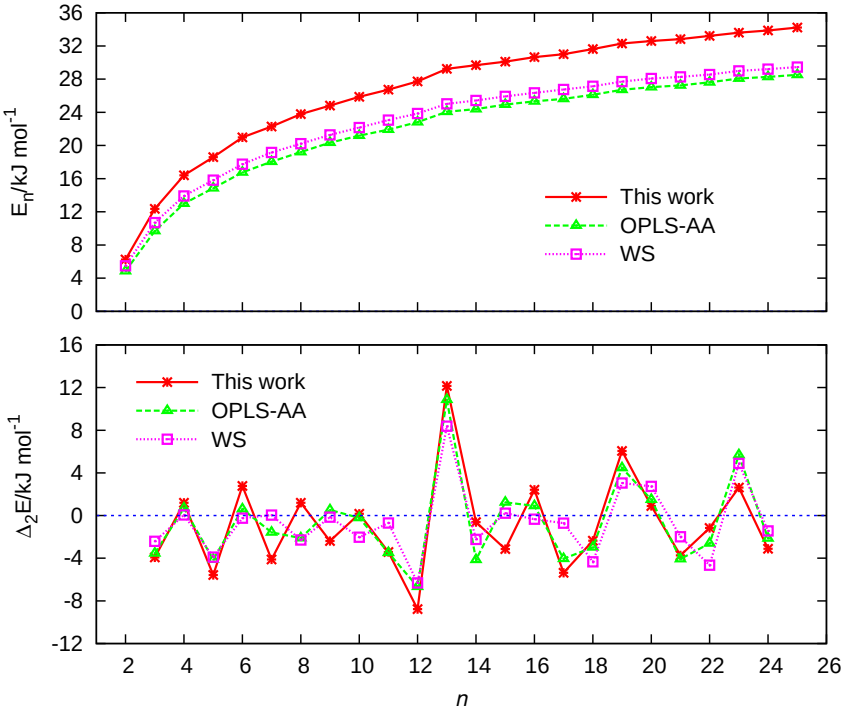

Figure 7: Comparison of the Bz_n energetics for different potential models: (a) average binding energy and (b) second energy difference. The lines represent the results for different potential models: this work (solid line), OPLS-AA structures of Ref. 10 (dashed line), and William and Starr minima of Ref. 50.

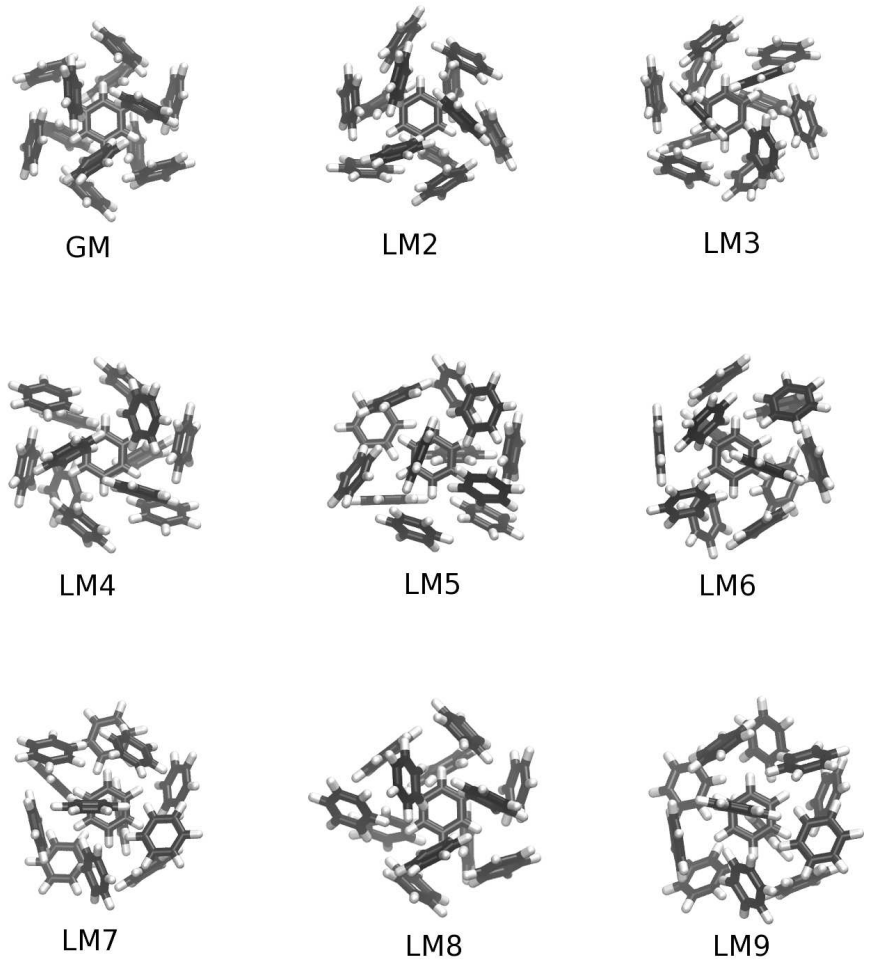

Figure 8: Lowest-energy structures for Bz_{13} cluster obtained with the EA. Labels refer to global minimum (GM) and local minima (LM2 to LM9). The GM structure has been already shown by a different perspective view in Figure 2. The corresponding energies in kJ mol^{-1} (and point groups of symmetry) are: -380.164 (S_6), -377.140 (C_3), -374.266 (C_1), -373.627 (C_i), -373.501 (C_1), -373.320 (C_1), -372.876 (C_1), -372.325 (C_1), and -370.301 (C_1). Plots were produced with VMD program [77].

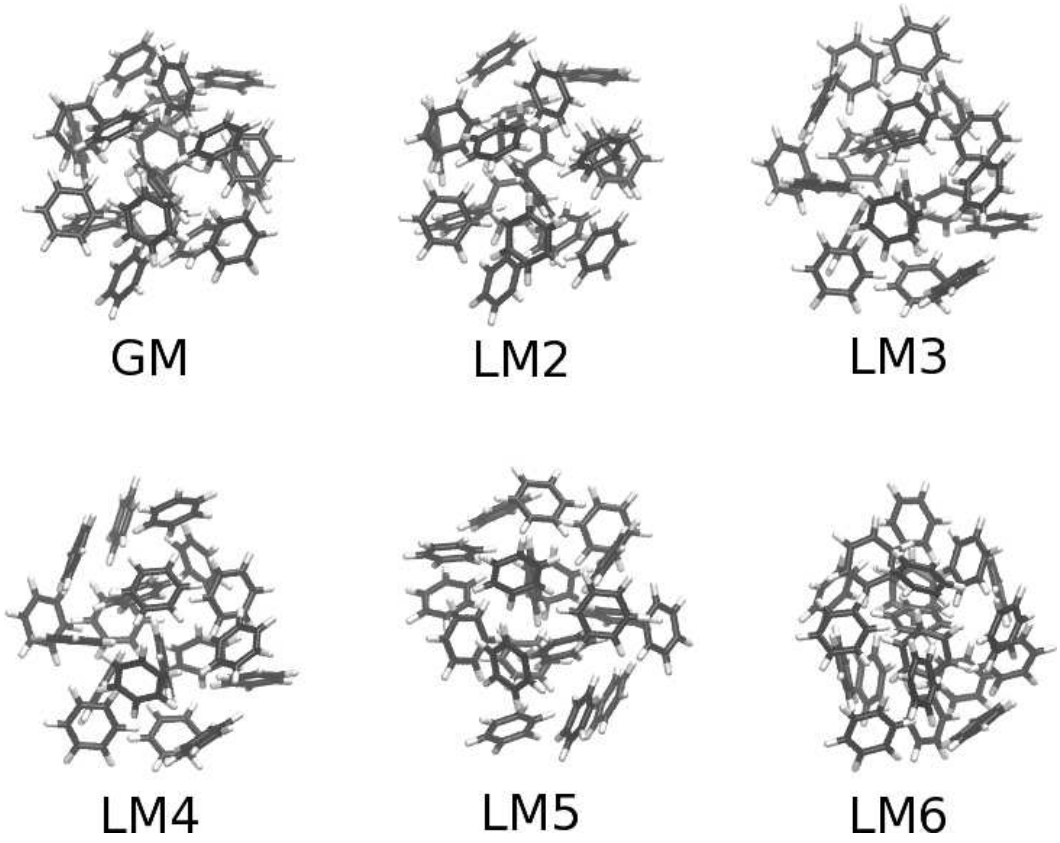

Figure 9: Lowest-energy structures for Bz_{19} cluster obtained with the EA. Labels refer to global minimum (GM) and local minima (LM2 to LM6). The GM structure has been already shown by a different perspective view in Figure 2. The corresponding energies in kJ mol^{-1} (and point groups of symmetry) are: -613.627 (C_1), -613.568 (C_1), -611.762 (C_1), -611.086 (C_1), -611.051 (C_1), and -609.759 (C_1). Plots were produced with VMD program [77].