PCCP

ARTICLE TYPE

Cite this: DOI: 00.0000/xxxxxxxxxx

Aggregation of coronene: the effect of carboxyl and amine functional groups†

C. F. O. Correia,[a] J. M. C. Marques,*[b], M. Bartolomei,[c] F. Pirani[d], E. Maçôas,[a] and J. M. G. Martinho,[a]

Received Date
Accepted Date

DOI: 00.0000/xxxxxxxxxx

The aggregation of coronene is relevant to understand the formation of carbon nanomaterials, including graphene quantum dots (GQDs) that show exceptional photophysical properties. This article evaluates the influence of carboxyl and amine substituting groups on the aggregation of coronene by performing a global optimization study based on a new potential energy surface. The structures of clusters with substituted coronene are similar to those formed by un-substituted monomers, that is, stacked (non-stacked) motifs are favoured for small-size (large-size) clusters. Nonetheless, the presence of carboxyl and amine groups leads to an increase of the number of local minima of comparable energy. The clusters with substituted monomers have also shown to enhance the attractive component interaction, which can be attributed to weak induction and charge transfer effects and to stronger electrostatic contributions. Moreover, the calculated height of magic-number structures of the clusters in this work is compatible with the morphology of the GQDs reported in literature.

1 Introduction

In the last decades, polycyclic aromatic hydrocarbons (PAHs), such as pyrene, coronene, ovalene or circumcoronene, have emerged as relevant chemical systems in several environments. In combustion, PAHs have been recognized as precursors of soot particles [1–5], while their presence in the interstellar medium has been detected through ultraviolet-irradiated regions in our and external galaxies [6]. The theoretical study of clusters formed by PAH monomers assumes an important role in assessing the physico-chemical evolution of small carbonaceous particles under realistic astrophysical conditions [7]. Additionally, clusters of different types of PAH molecules have been subjected to a great number of investigations [8–11] by employing global optimization techniques. This type of optimization methods, that are especially adequate for exploring the energy landscapes of molecular clusters, have been developed by many research groups all over the years [10,12–21]. Among the global optimization studies on PAH systems, those involving coronene clusters have received special attention. Indeed, coronene clusters are considered as prototype systems for the study of strong $\pi-\pi$ interactions in carbon nanomaterials [22,23]. Due to the intrinsic $0D$ (zero dimensional) nature of the clusters, they are particularly relevant as models to understand the physico-chemistry of crystalline carbon dots (Cdots). These nanoparticles of carbon typically have dimensions below 10 nm and they can be produced as a single layer or as a quasi-spherical stack of few layers of sp^2 carbons, also known as graphene quantum dots (GQDs) [24]. Cdots are usually decorated with oxygen and nitrogen containing functional groups that provide solubility in aqueous media, reduce the non-radiative energy relaxation pathways due to surface passivation and afford a rich chemical reactivity for customization towards specific application. However, our understanding about the structural effects on the physico-chemical properties of Cdots is clearly lagging behind the excitement of the scientific community about their potential applications in biomedical, photocatalytic and electronic applications [25–29]. Despite the overwhelming number of yearly publications dealing with Cdots, there are several open questions yet: Are the optical properties determined by the carbon core or do they have a molecular nature? Can their photophysics be explained by a cocktail of aggregated PAHs? Given the heterogeneity of the materials produced, are the properties determined by the presence of truly monolayered dots, or stacks of sp^2 carbons or even molecular by-products?

Actually, most of the computational studies available are based on isolated PAHs to simulate the carbons dots providing a useful, though limited, understanding on the structure-properties correlation [30–32]. The study of coronene clusters functionalized with

[a] Centro de Química Estrutural (CQE), Instituto Superior Técnico, University of Lisbon, 1049-001 Lisboa, Portugal. E-mail: catia.correia@tecnico.ulisboa.pt; ermelinda.macoas@tecnico.ulisboa.pt; jgmartinho@tecnico.ulisboa.pt
[b] CQC, Department of Chemistry, University of Coimbra, 3004-535 Coimbra, Portugal. E-mail: qtmarque@ci.uc.pt
[c] Instituto de Física Fundamental, Consejo Superior de Investigaciones Científicas (IFFCSIC), Serrano 123, 28006 Madrid, Spain. E-mail: maxbart@iff.csic.es
[d] Dipartimento di Chimica, Biologia e Biotecnologie, Università di Perugia, 06123 Perugia, Italy. E-mail: pirani.fernando@gmail.com.
† Electronic Supplementary Information (ESI) available:

polar groups at the edge can provide a deeper insight on the structure of Cdots. Functionalization with carboxylic and amine groups, typically found in Cdots, allows us to investigate the concomitant contribution of $\pi-\pi$, electrostatic, van der Waals and hydrogen bonding interactions to the properties of the corresponding clusters. In addition, COOH moieties are relevant as key intermediates appearing during the oxydation of coronene in the context of soot combustion[33]. The functionalization of coronene with imide has shown to be a potential route for tuning optoelectronic properties[34].

In this work, we have employed an evolutionary algorithm (EA) to perform a global optimization study on clusters of coronene molecules substituted with carboxyl and amine groups. Based on accurate density functional theory (DFT) calculations, an analytical potential energy surface (PES) has been also proposed for the interactions involving pairs of monomers. By applying the EA to the new PES, we have discovered putative global minimum and low-energy structures of the clusters up to 15 monomers, which allows for the assessment of the effect of substituting groups from the comparison with the results on pure coronene aggregates. Accordingly, the plan of the paper is as follows. The methodology employed in the construction of the PES is fully described in Section 2, where we also overview the EA used to discover the low-energy structures of the clusters. In section 3, the global optimization results are presented and discussed, while section 4 summarizes the main conclusions of this study.

2 Methodology

According to the guidelines followed in the recent investigation of pure coronene clusters[9], also the present study has been performed starting from extensive theoretical calculations, addressed to characterize in detail structures, and energy sequence of the coronene substituted dimer in its most stable configurations. The subsequent objective has been the formulation of a semiempirical representation of the multi-dimensional intermolecular PES for the same dimer, whose potential parameters have been defined in terms of fundamental chemical-physical properties of monomers. On general ground, this objective is crucial to describe the main features of dimers, as well as of larger clusters, formed by substituted coronene monomers under a variety of conditions. Moreover, the detailed comparison of the semiempirical model predictions with results of ab initio calculations has been extensively exploited to test, improve, and generalize the semiempirical formulation of the intermolecular potential. The analytical PES here proposed provides an internally consistent description of the intermolecular interaction in the full space of the relative configurations of the investigated systems. As stressed above, this condition has been considered basic to represent the main features of the low-energy landscape, as well as to carry out any type of molecular dynamics simulation. The analytical PES so obtained has been then exploited, by employing an EA method, to search the global minima (and other low- energy structures) of substituted coronene clusters of increasing size.

2.1 Electronic-structure calculations

As for pure coronene clusters[9], also for present coronene substituted clusters the interaction energies were computed within the generalized gradient approximation of the Perdew, Burke, and Ernzerhof (PBE)[35] exchange-correlation functional using the cc-pVTZ basis set[36]. In order to correct the obtained interaction energies for required dispersion contributions we have added those from the DFT-D3(BJ) method of Grimme et al[37]. Moreover, the counterpoise method[38] was applied to correct for the basis set superposition error. The reliability of the obtained interaction energies has been checked through the comparison with corresponding benchmark MP2C[39] estimations (exploiting a complete basis set (CBS) extrapolation scheme[40,41]), as reported in the Supplementary Information (see Fig. S1 in there) for the interaction profiles corresponding to a couple of parallel configurations of the coronene substituted dimer. The PBE-D3(BJ) method was therefore safely employed to identify the low lying structures of the coronene substituted dimer and to obtain the proper sequence of the related complexation energies. The same approach was also used to verify the relative stability of the low-lying minima discovered with the EA for cluster up to pentamers, as shown in Table 4. The partial atomic charges involved in the electrostatic term of the analytical potential function formulation (see next section) have been determined from electronic structure calculations using the B3LYP hybrid functional[42] and the cc-pVQZ basis set[36] within the Charge Model 5 (CM5) approach[43], which represents an extension of the Hirshfeld population analysis and in general guarantees reliable estimations. All calculations were carried out with the Gaussian 09 package[44].

2.2 Analytical potential function

As stressed above, extensive molecular simulations, especially for systems of increasing complexity, require the use of an analytical formulation of the multidimensional PES. A careful analysis of the dimer behavior has been considered crucial to achieve such purpose. In particular, an important condition adopted is that the parameters defining the multidimensional PES must have a physical meaning in order to guarantee the representation correctness of the intermolecular potential in the full space of the relative configurations of involved partners. In this study, we assume that coronene (Cor) molecules substituted with -COOH and -NH_2 groups (hereafter designated as R-Cor-R') are rigid bodies, in which the C-C bond lengths and C-C-C angles were set to 1.420 Å and 120°, respectively, whereas C-H bond lengths and C-C-H angles were chosen to be 1.090 Å and 120°, respectively. Such values are very close to those of pure coronene. However, in the evaluation of the interaction parameters we assumed that the size of each C atom in substituted coronene is slightly smaller with respect to that in pure coronene because of the aromatic π electron cloud polarization promoted by the two functional groups involving N and O atoms with high associated electronegativity value. The main features of the Cor monomer are shown in Figure 1.

For clusters formed by n (R-Cor-R') monomers, the total inter-

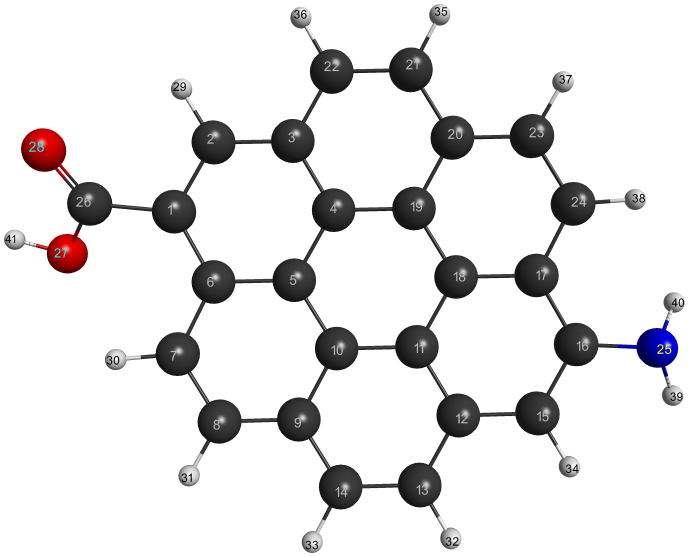

Fig. 1 Representation of the 3D structure of the (R-Cor-R') monomer. The numbers on the atoms indicate the order followed to describe the interaction potential in this work.

action potential is defined as:

$$V_{cluster} = \sum_{k=1}^{n(n-1)/2} \left(V_{el,k} + V_{nel,k}\right) \quad (1)$$

where the first (second) term in the summation refers to the electrostatic (non-electrostatic) component arising as combination of all independent interacting pairs distributed on the frame of monomers involved in the cluster. This model has been developed as a natural evolution of those previously proposed for the interaction involving benzene-benzene[45] and coronene-coronene[9] systems. The calculation of $V_{el,k}$, *i.e.*, between monomers A and B in vacuum (permitivity ε_o) forming the k pair, assumes the following expression:

$$V_{el} = \sum_{i=1}^{41} \sum_{j=1}^{41} \frac{q_i q_j}{4\pi\varepsilon_o r_{ij}} \quad (2)$$

where i (j) runs over all charge points distributed on the monomer A (B) frame and r_{ij} is the distance between two charge points located in distinct monomers; i (or j) values are here assigned to the atomic centers, as in Figure 1, while the corresponding charges are given in Table 1. Charge values distributed on the coronene-substituted molecular frame have been computed as indicated above and are considered here as fixed parameters. All possible deviations (expected to be small) from the used values are accounted for in the optimization of the parameters of the non-electrostatic component of the interaction (see below).

As for the non-electrostatic energy component, associated to the same k pair, consistently with previous PES formulations[9,45], it is described again as a sum of partial contributions, *i.e.*,

$$V_{nel} = \sum_{i=1}^{41} \sum_{j=1}^{41} V_{X_i - Y_j} \quad (3)$$

where $V_{X_i - Y_j}$ represents the non-electrostatic interaction between

Table 1 Partial charges assigned to the atoms of the monomer. Each atom belongs to the group indicated in parenthesis.

i	atom (group)	charge/e
1	C (Cor)	-0.014067
2	C (Cor)	-0.074214
3	C (Cor)	-0.014482
4	C (Cor)	0.007266
5	C (Cor)	0.002901
6	C (Cor)	-0.003732
7	C (Cor)	-0.102700
8	C (Cor)	-0.093862
9	C (Cor)	-0.017734
10	C (Cor)	0.002825
11	C (Cor)	-0.006821
12	C (Cor)	-0.013269
13	C (Cor)	-0.099765
14	C (Cor)	-0.095018
15	C (Cor)	-0.113695
16	C (Cor)	0.102275
17	C (Cor)	-0.014972
18	C (Cor)	0.004031
19	C (Cor)	0.004350
20	C (Cor)	-0.011434
21	C (Cor)	-0.096995
22	C (Cor)	-0.092226
23	C (Cor)	-0.097937
24	C (Cor)	-0.097875
25	N (-NH_2)	-0.641022
26	C (-COOH)	0.273383
27	O (-COOH)	-0.367484
28	O (-COOH)	-0.351177
29	H (Cor)	0.113885
30	H (Cor)	0.099767
31	H (Cor)	0.105742
32	H (Cor)	0.102366
33	H (Cor)	0.104263
34	H (Cor)	0.102495
35	H (Cor)	0.104652
36	H (Cor)	0.106880
37	H (Cor)	0.104842
38	H (Cor)	0.102726
39	H (-NH_2)	0.311651
40	H (-NH_2)	0.303564
41	H (-COOH)	0.360169

the atomic center X_i of monomer A and the atomic center Y_j of monomer B. The adoption of an effective atomic electronic polarizability, different from that of the isolated atom, takes into account of its participation in stable chemical bonds. Moreover, the sum of all effective atomic values must be consistent with the polarizability of the R-Cor-R' molecule. In equation 3, each term is described by an improved Lennard-Jones (ILJ) potential function V_{ILJ}[46–48], which depends on the separation distance, r, between the two interacting centers according to the expression:

$$V_{\mathrm{ILJ}} = \varepsilon \left[\frac{m}{n(r)-m} \left(\frac{r_0}{r}\right)^{n(r)} - \frac{n(r)}{n(r)-m} \left(\frac{r_0}{r}\right)^{m} \right] \qquad (4)$$

where ε and r_0 are, respectively, the well depth and equilibrium distance associated to each specific interaction pair formed by two "effective" atoms. As usual for neutral-neutral systems, the m parameter is here chosen equal to 6 in all cases. The first term in Eq. 4 describes the size (or Pauli) repulsion, while the second one represents the attraction, mostly depending on dispersion contributions. The $n(r)$ exponent, defining simultaneously the falloff of the effective atom-effective atom repulsion and the strength of the attraction, is expressed as[46]

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

For typical non covalent interactions the additional β parameter assumes standard values falling in the 7-9 range[45–49] being related to the hardness of involved partners. Therefore, the β value and the proper $n(r)$ dependence make the ILJ function more flexible and more realistic than the classical LJ model, removing most of its inadequacies due to an excessive attraction and to a too strong repulsion. Moreover, in the tuning and final choice of the parameters, whose zero order values were estimated from the effective atomic polarizabilities, the following modulation criteria have been adopted: 1) Each interaction pair, involving a positive charged H atom of a monomer (see Table 1) and a negative charged (essentially O and N, *cf.* Table 1) of another monomer, can promote the formation of intermolecular hydrogen bond (HB). Accordingly, the r_0 parameter has been decreased of about 10-11% and, consequently, ε has been increased in order to maintain constant the long range dispersion attraction contribution, defined as $\varepsilon \cdot r_0^6$. For such a pair also β has been lowered to 6.5, since the variation within limiting ranges of β allows also to indirectly include additional interaction components emerging at intermediate and short range, as perturbative-stabilizing induction and charge transfer effects associated to the weak HB formation. 2) For pairs involving less positive H atoms, as those linked to the coronene plane, that can promote weak HB, r_0 parameter has been decreased of only 2% and ε has been slightly increased in order to maintain always constant the long range dispersion attraction contribution, while β has been maintained at 6.5. 3) As anticipated before also the size, and consequently related r_0 of all pairs involving C(Cor), has been decreased with respect to values adopted for pure coronene clusters, in order to account for polarization effects of the π electron cloud. The final parameters employed for the non-electrostatic atom-atom interactions are given in Table 2.

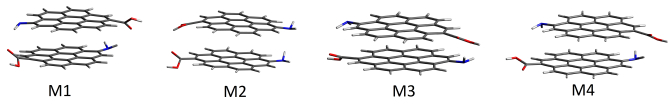

Fig. 2 Four dimer structures optimized at the PBE-D3(BJ)/VTZ level of theory. The corresponding energies are given in Table 3.

Note also that in the adopted potential formulation the parameters, which maintain the correct physical meaning, have been optimized to obtain a proper comparison with the results of the *ab initio* estimations. The analytical PES so obtained provides binding energies which represent the best compromise with values predicted by both MP2C and DFT methods and the correct stability sequence of low lying dimer minima (see Figure 2 and Table 3). Moreover, the same formulation reproduces related binding energy values which appear to be significantly stronger than those for the coronene dimer[9].

2.3 Global optimization method

A hybrid evolutionary algorithm (EA) developed in our group[20] has been employed to search for low-energy structures of the clusters (R-Cor-R')$_n$ (with R≡NH_2 and R'≡COOH) studied in this work; the (R-Cor-R') monomers are treated as rigid molecules. The EA combines an efficient global optimization strategy to explore promising structures at the basins of attraction of the potential energy surface with the limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm[50,51] to obtain the local minimum geometry at the bottom of each basin. This latter procedure reduces the space to be searched, since the global optimization runs over structures already minimized and, thus, avoids to deal with a number of bad-quality intermediate solutions. On following, we describe the main features of the EA, while further details were published in the original paper[20] and revised in a recent work on molecular clusters[52]. Additionally, the EA has been successfully applied to the global optimization of clusters involving coronene[9,49].

Basically, a generational approach is adopted in the present EA, meaning that the pool of structures forming the population of possible solutions is replaced from one generation to the other. Since each (R-Cor-R') molecule is assumed to be a rigid-body structure, the geometry of the cluster is defined by specifying the center-of-mass coordinates and the orientation Euler-angles of the corresponding monomers. For a cluster with n (R-Cor-R') monomers, a solution in the pool is, thus, constituted by n six-tuples, each one encoding three Cartesian coordinates of the center-of-mass and three Euler angles.

The initial population of structures forming the pool of possible solutions is randomly generated and optimized to local minima by employing the L-BFGS algorithm[50,51]. The next generation is, then, obtained by applying the genetic operators. First, tournament selection is applied to choose pairs of structures from the population that undergo crossover to originate offspring solutions. Our EA relies on the arithmetic simulated-binary crossover[53] to combine the selected pairs of structures. Second, both center-of-mass and orientation of monomers have a certain

Table 2 Analytical Parameters of the (R-Cor-R')-(R-Cor-R') Interaction Potential: Well Depth[(a)] (ε), Equilibrium Distance[(b)] (r_0) and β Parameters of the ILJ Potential for the Atom-Atom Pair Interactions.

ILJ interaction	ε/meV	r_0/Å	β
C(Cor)–C(Cor)	3.670	3.900	7.0
C(Cor)–H(Cor)	2.200	3.505	6.5
H(Cor)–H(Cor)	1.610	3.099	9.0
C(COOH)–C(Cor)	4.460	3.810	7.0
C(COOH)–C(COOH)	4.755	3.709	7.0
C(COOH)–H(Cor)	2.427	3.464	8.0
C(COOH)–O(COOH)	4.484	3.603	7.0
C(COOH)–H(COOH), C(COOH)–H(NH_2)	2.298	3.451	8.0
O(COOH)–C(Cor)	4.050	3.720	7.0
O(COOH)–H(Cor)	2.790	3.250	6.5
O(COOH)–O(COOH)	4.465	3.481	7.0
O(COOH)–H(COOH), O(COOH)–H(NH_2)	5.140	2.900	6.5
C(COOH)–N(NH_2)	5.020	3.709	7.0
O(COOH)–N(NH_2)	4.773	3.603	7.0
H(COOH)–H(Cor), H(NH_2)–H(Cor)	1.560	3.076	9.0
H(COOH)–C(Cor), H(NH_2)–C(Cor)	2.470	3.470	6.5
H(COOH)–H(COOH), H(NH_2)–H(NH_2), H(COOH)–H(NH_2)	1.509	3.053	9.0
N(NH_2)–N(NH_2)	5.317	3.709	7.0
N(NH_2)–H(NH_2), N(NH_2)–H(COOH)	4.610	3.100	6.5
N(NH_2)–C(Cor)	4.680	3.810	7.0
N(NH_2)–H(Cor)	2.857	3.400	6.5

[(a)] 1 meV=0.0964853 kJ mol^{-1}; [(b)] 1 Å=0.1 nm

Table 3 Interaction Energies of Four Minimum Structures Obtained For the Dimer at the PBE-D3(BJ)/VTZ Level of Theory And That Are Displayed in Fig. 2. Also Represented Are the Corresponding Single-Point MP2C/CBS and the Analytical PES Energies.

structure	E_{DFT} (kJ mol^{-1})	E_{MP2C} (kJ mol^{-1})	E_{PES} (kJ mol^{-1})
M1	-115.31	-121.44	-112.28
M2	-107.11	-119.40	-111.83
M3	-107.53	-115.05	-108.03
M4	-99.33	-108.68	-94.85

Table 4 Interaction Energies for the Global Minimum (GM) and Other Relevant Structures of the (R-Cor-R')$_n$ ($n = 3-5$) Clusters Obtained by applying the EA to the Analytical PES and by Performing Single-Point PBE-D3(BJ)/VTZ.

cluster size (n)	structure	E_{PES} (kJ mol^{-1})	E_{DFT} (kJ mol^{-1})
3	GM	-230.44	-197.48
	M1	-229.68	-196.73
	M2	-229.66	-184.64
	M3	-177.79	-160.25
4	GM	-348.71	-295.68
	M1	-343.40	-295.64
	M2	-339.14	-284.30
	M3	-336.87	-278.61
	M4	-315.23	-271.92
5	GM	-464.67	-397.44
	M1	-458.56	-394.55
	M2	-447.48	-359.91
	M3	-438.64	-370.37
	M4	-420.14	-358.99

probability to be modified by applying sigma mutation[54]. Third, the structures so obtained are subjected to local optimization by using the L-BFGS algorithm[50,51], which leads to the formation of the new generation of solutions. As mentioned above, this new population replaces the old one in the pool of possible solutions. Such replacement strategy, however, includes the application of an elite operator in order to avoid lost the lowest-energy structure.

The procedure just described for the formation of a new generation of possible solutions of the (R-Cor-R')$_n$ cluster is repeated until the corresponding potential energy function is evaluated by 30000 times; this iterative process is usually designated as a run. As we are dealing with a stochastic approach, the EA has to be run several times so that statistically meaningful results are achieved. Here, we have employed 30 runs for each cluster size; other relevant settings of the EA are the same as those given in a previous work[45].

Finally, the putative global minimum for a given cluster size is the structure with the lowest-energy that has been obtained after the completion of 30 runs. The EA also saves a given number of other low-energy structures besides the global minimum, which are relevant for the subsequent analysis. We further note that the average CPU time estimated for (R-Cor-R')$_3$ is $\langle t_{CPU} \rangle = 8742\,s$ per run of the EA in a computer Intel(R) Core(TM) i7-3770 CPU @ 3.40 GHz with x86_64 architecture.

3 Results and discussion

The presence of functional groups in substituted coronene increases the number of structures of comparable energy in the formed clusters with respect to the case of equivalent aggregates with un-substituted monomers. In particular, several energy minima are expected to occur in the multidimensional PES of each substituted coronene cluster. The interaction energies associated with the low lying structures of (R-Cor-R')$_n$ clusters with size $n = 3-5$ and characterized by the EA are reported in Table 4.

In the same table are also given for an useful comparison results obtained performing Single-Point PBE-D3(BJ)/VTZ calculations. The comparison suggests that although the EA values tend to be higher in absolute value than the DFT determinations it is necessary to take into account that the latter refer to single point estimations and that the analytical PES employed in the minimization procedure assumes rigid monomers. In addition, we have checked that for the absolute minimum (GM) of the trimer a DFT optimization with flexible monomers the corresponding interaction energy increases to $-215.9\,\mathrm{kJ\,mol^{-1}}$ and get closer to the EA prediction. Furthermore, it is important to stress that the stability sequence of the low lying structures is provided in a consistent way by the two methods and the structural details of the low-lying minima of Table 4 are addressed in the end of this section. In turn, the putative global minimum structures discovered by the EA are represented in Figure 3. It is apparent from this figure that columnar-type global minimum structures are favored up to $n = 5$. Conversely, larger clusters show multi-stack configurations: two stacks forming a "handshake" structure (for $n = 7-12$) or three stacks (for $n = 13-15$). We observe that large stacks with up to six monomers, which tend to be distorted, can be stabilized as the cluster grows up. Such kind of structures, highlighting the prevalence of $\pi-\pi$ interactions, have been also reported in previous global optimization studies on pure coronene (Cor_n) clusters[9,55]. Indeed, we have found[9] quite similar structures for Cor_n clusters that were modeled with a PES based on the same approach employed in this work for the title system. Nonetheless, some subtle differences arise between the global minimum structures of Cor_n and $(\text{R-Cor-R'})_n$ clusters; a detailed analysis requires the comparison between Figure 3 and Figure 3 of Ref. 9. Mainly, we should emphasize that stacks in $(\text{R-Cor-R'})_n$ clusters tend to be more distorted than in the corresponding Cor_n structures and, in some cases, showing monomers squeezed by the stronger attraction involved (see below). Moreover, the major qualitative difference appears for $n = 6$ whose putative global minimum shows three stacks (with 2+2+2 monomers) for Cor_n and only two stacks (with 4+2 monomers) for $(\text{R-Cor-R'})_n$.

The stability of the growing clusters can be evaluated from the average binding energy, *i.e.*,

$$E_n = -E_{\mathrm{cluster}}(n)/n \qquad (6)$$

where $E_{\mathrm{cluster}}(n)$ is the global-minimum energy at size n. In Figure 4, we represent the average binding energy (E_n) as a function of the cluster size. First of all it can be appreciated that, if compared with the corresponding results for Cor_n clusters reported in Fig. 4 of Ref. 9, present average binding energies are significantly larger of about 40%. This enhancement in the global interaction energy must be associated to the presence of polar functional groups which generates further attractive interaction contributions in the medium and long ranges due to the formation of weak hydrogen bonds as well as to stronger electrostatic effects. In addition, we may observe that the above mentioned change from a one-stack structure to multi-stack motifs when moving from $n = 5$ to $n = 6$ cannot be noticed in the E_n curve. In contrast, the van der Waals (Coulomb) contribution for E_n diminishes (increases)

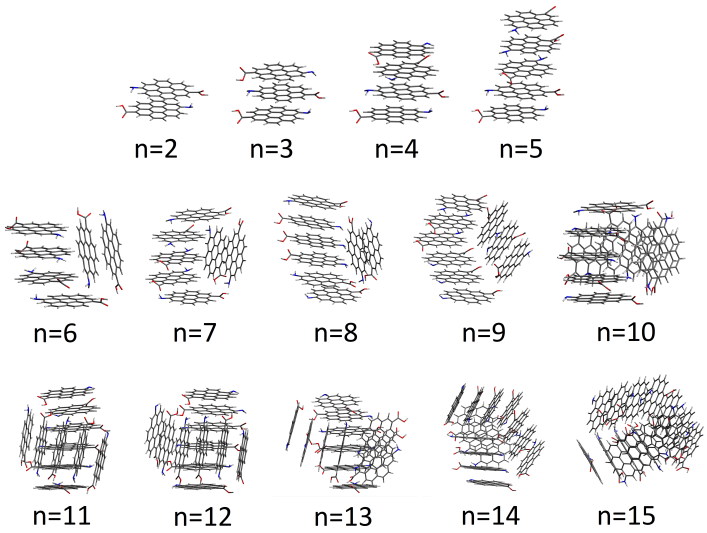

Fig. 3 Putative global minimum structures of the $(\text{R-Cor-R'})_n$ clusters (with $n = 1-15$); R≡-NH_2 and R'≡-COOH. Plots of the structures were obtained with the Visual Molecular Dynamics (VMD) program[56].

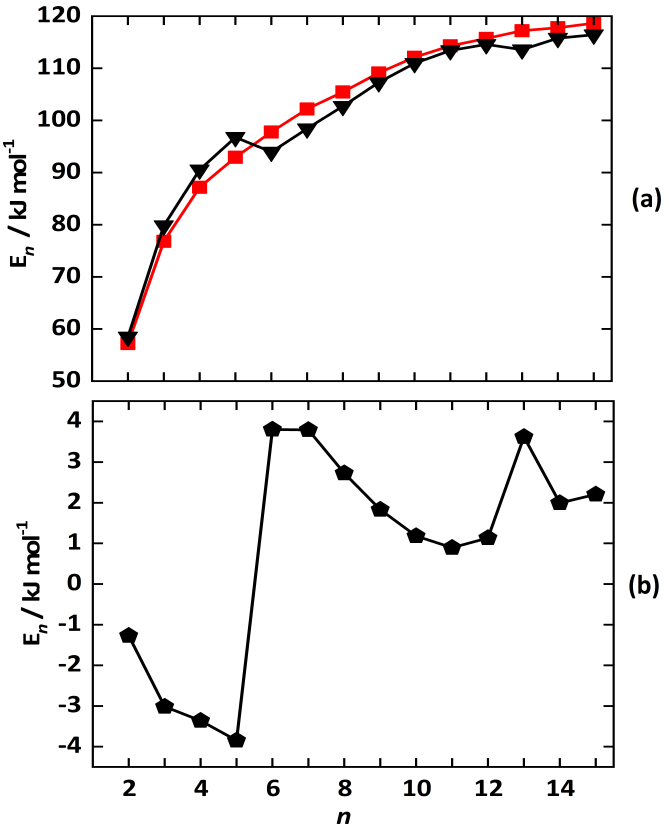

Fig. 4 Energy per monomer for the $(\text{R-Cor-R'})_n$ clusters (with $n = 1-15$): (a) average binding energy of the cluster (red line and squares) and van der Waals component (black line and triangles); (b) average Coulomb-energy contribution.

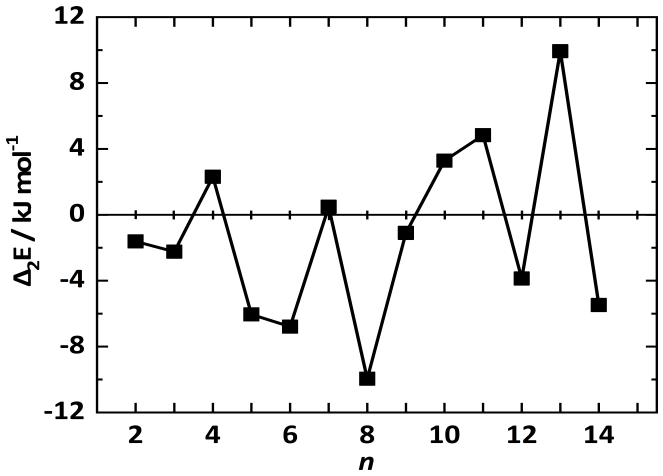

Fig. 5 Second energy difference for the (R-Cor-R')$_n$ clusters (with $n = 2-14$).

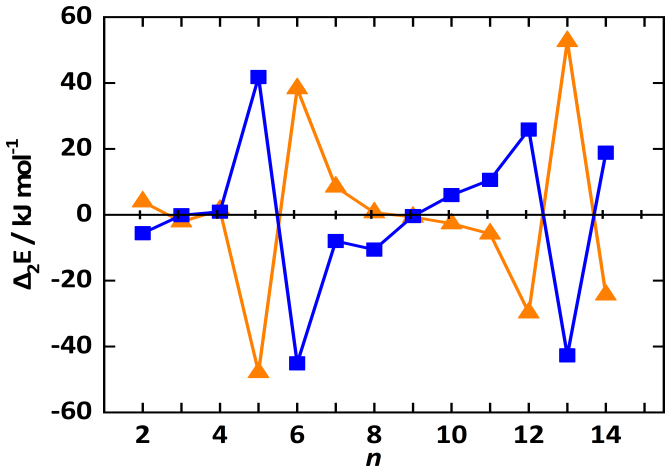

Fig. 6 Contributions for the (R-Cor-R')$_n$ ($n = 2-14$) second-energy difference in Fig. 5: van der Waals (blue curve) and Coulomb (orange curve) interactions.

significantly from $n = 5$ to $n = 6$; notice also that Coulomb interaction provides destabilization effects up to $n = 5$. A similar behavior, though with a smaller magnitude, is also observed when changing from a handshake-type motif (at $n = 12$) to a three-stack structure (at $n = 13$). Thus, it appears that stacking maximizes the van der Waals contribution in relation to the Coulomb-energy component. Although the former constitutes, by far, the major contribution for the binding energy, the latter becomes relevant for changing the structural motif of the global minimum.

Another relevant parameter that accounts for the stability of the clusters is the second energy difference:

$$\Delta_2 E(n) = E_{\text{cluster}}(n-1) - 2E_{\text{cluster}}(n) + E_{\text{cluster}}(n+1) \qquad (7)$$

This equation estimates the relative stability of the (R-Cor-R')$_n$ global minimum structure regarding to the neighbor-size clusters, whose energies are $E_{\text{cluster}}(n-1)$ and $E_{\text{cluster}}(n+1)$. Clusters with high relative stability correspond to maxima of the $\Delta_2 E$ curve and are usually designated as "magic numbers". Figure 5 shows magic numbers for (R-Cor-R')$_n$ clusters at $n = 4$, 11, and 13. In particular, the $n = 11$ magic number is associated to a broad maximum that includes also the (R-Cor-R')$_{10}$ structure, while the most prominent peak arising at $n = 13$ may be related to the transition from the two- to the three-stack structural motifs. It is worth noting that, in the same size range, $n = 10$ has been assigned as the strongest magic number for Cor$_n$ clusters[9], which might be attributed to the compact handshake structure formed by 5+5 stacks[8]. We also note that the height of the magic-number structures ranges from 1 nm ($n = 4$) to 2 nm ($n = 13$). This is in agreement with the morphology of small graphene quantum dots prepared using a bottom up approach, which shows height distributions peaking in the range $1-2$ nm and lateral dimensions < 10 nm[24,57,58]. The existence of more than one stack in larger clusters can explain why it is difficult to collect a clear picture of the graphitic in-plane lattice spacing of the dots using electron microscopy. It is also in agreement with the broad band that characterizes the X-ray diffractogram in GQDs, resembling amorphous material.

A detailed analysis of the contribution of van der Waals and Coulomb energy components for establishing the magic numbers can be extracted from Figure 6. A first inspection of this figure clearly shows that the two energy contributions act in opposite directions concerning the relative stabilization of most of the clusters, and hence the magic numbers arise from the subtle balance between van der Waals and Coulomb components. Accordingly, distinct energetic features contributing to the above mentioned $n = 11$ and $n = 13$ magic numbers (*cf.* Figure 5) are apparent in Figure 6: whereas the former is due to the van der Waals contribution, the latter results from the Coulomb-energy stabilization. In contrast, both energy components of the empirical potential contribute to the small magic number appearing at $n = 4$.

We now move the discussion to the analysis of the low-energy structures whose energies are given in Table 4. The corresponding structures are displayed in Figures 7-9; the global minimum configurations are also included to facilitate the comparison with the other minima. We can observe in these figures and Table 4 that stacked structures are more stable than unstacked ones. For instance, the first unstacked structure for (R-Cor-R')$_3$ is about 52 kJ mol^{-1} above the global minimum, within the proposed model. It is worth noting that this energy difference is significantly higher than the value (~ 23 kJ mol^{-1}) previously reported[9] for the trimer of the pure coronene cluster. As in the case of Cor$_n$ clusters[9], the energy difference between stacked and unstacked structures of (R-Cor-R')$_n$ decreases as the size of the cluster increases. In turn, several stacked structures can arise by rotating monomers around the stack axis and by sliding one monomer over the other. This changes the relative orientation of the substituting groups and the corresponding minima may have distinct values of the energy as is the case of tetramers GM, M1, M2 and M3 (pentamers GM, M1 and M2) shown in Figure 8 (Figure 9). This may be attributed to the strong electrostatic contribution due the presence of the polar functional groups as well as the influence of inter-monomer weak hydrogen bonds. Finally, the formation of different configurations are also achieved by changing the up/down orientation of both the carboxyl and amine groups of one monomer in relation to the other. In general, this leads to a

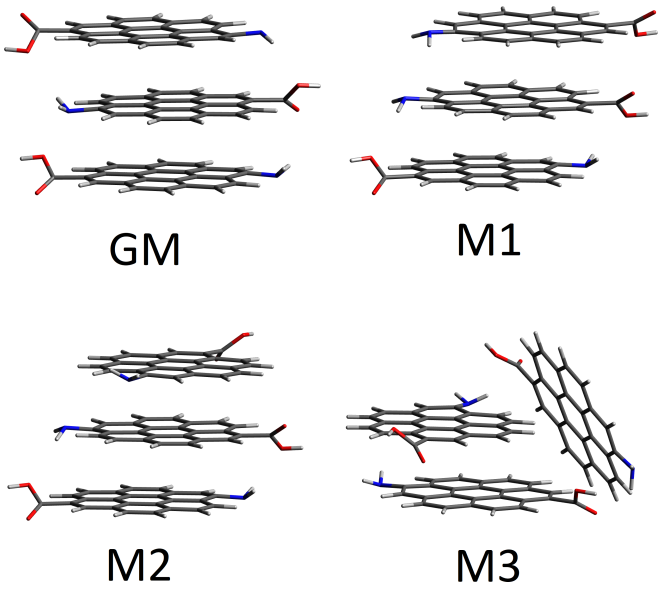

Fig. 7 Minimum structures of the (R-Cor-R')$_3$ cluster whose energies are given in Table 4; R≡-NH_2 and R'≡-COOH. Plots of the structures were obtained with the Visual Molecular Dynamics (VMD) program[56].

small difference in the energy of the two conformers, as shown by the trimers M1 and M2 of Figure 7.

4 Conclusions

We have investigated the energy and structure of (R-Cor-R')$_n$ (n=1-15) clusters, with R≡NH_2 and R'≡COOH, by exploiting a new all-atom analytical PES representing the involved global interaction and validated on accurate electronic structure calculations. The molecular structure and size of the R-Cor-R' monomer can be considered comparable with those of the un-substituted Cor counterpart and this leads to quite similar features in the organization of the related aggregates with a correspondence of stacked and non-stacked motifs. However, the presence of polar functional groups in R-Cor-R' globally enhances the role of the attractive component of the intermolecular interaction in the medium and long ranges due to the influence of weak hydrogen bonds as well as to stronger electrostatic contributions. In fact, the EA approach has allowed to obtain the binding energy and structure of the most stable configurations of the investigated clusters, which present an increase of the number of local minima of comparable energy with respect to correspondent pure coronene aggregates, being the global minima of the former more stable of about 40% than those of the latter. The different role of the electrostatic, V_{el}, and non-electrostatic, V_{nel}, components in determining the cluster structural features has been also evidenced.

The height of the magic-number structures is in agreement with the morphology of GQDs and small bottom-up carbon dots reported in literature. This supports the use of the cluster model of the present work to further study the optical properties of such material duly accounting for interlayer stacking and intrack interactions which are overlooked by simplified single-layer models.

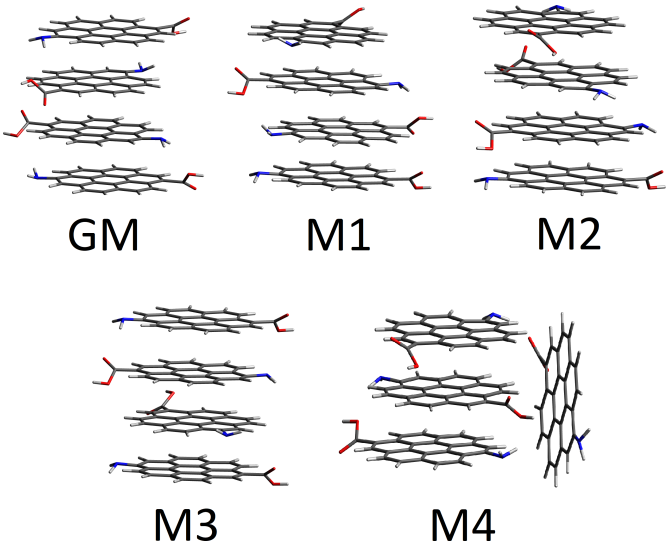

Fig. 8 As in Fig. 7 but for the minimum structures of the (R-Cor-R')$_4$ cluster.

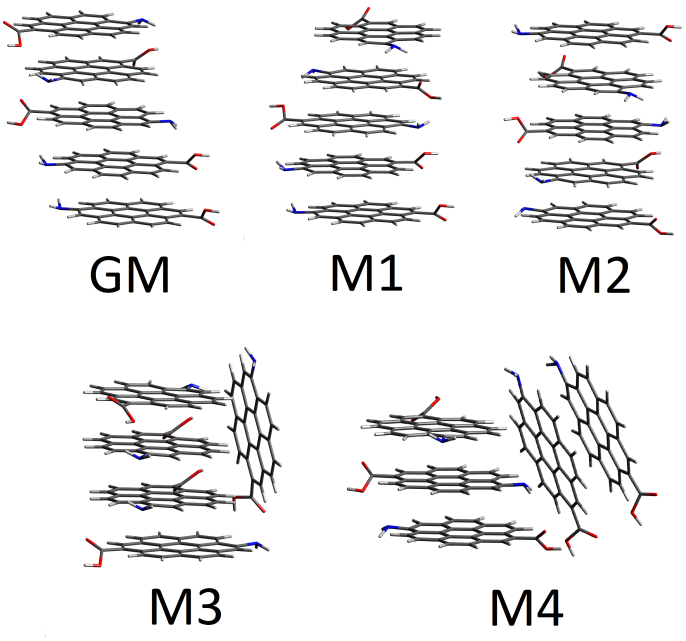

Fig. 9 As in Fig. 7 but for the minimum structures of the (R-Cor-R')$_5$ cluster.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

J.M.C.M. acknowledges the support from the Coimbra Chemistry Centre (CQC), which is financed by the Portuguese "Fundação para a Ciência e a Tecnologia" (FCT) through the programmes UIDB/00313/2020 and COMPETE. We are also 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 FIS2017-84391-C2-2-P grant as well as the allocation of computing time by CESGA supercomputing center (Spain). E.M., J.M.G.M and C.F.O.C. thank FCT for funding: PTDC/NAN-MAT/29317/2017, UIDB/00100/2020 and PD/BD/127805/2016.

Notes and references

1 R. A. Dobbins, R. A. Fletcher, B. A. Benner and S. Hoeft, *Combust. Flame*, 2006, **144**, 773 – 781.

2 M. Alfè, B. Apicella, R. Barbella, J.-N. Rouzaud, A. Tregrossi and A. Ciajolo, *Proc. Combust. Inst.*, 2009, **32**, 697 – 704.

3 B. Shukla and M. Koshi, *Phys. Chem. Chem. Phys.*, 2010, **12**, 2427–2437.

4 H.-B. Zhang, X. You, H. Wang and C. K. Law, *J. Phys. Chem. A*, 2014, **118**, 1287–1292.

5 M. Keller, T. de Bruin, M. Matrat, A. Nicolle and L. Catoire, *Energy & Fuels*, 2019, **33**, 10255–10266.

6 C. Puzzarini and V. Barone, *Phys. Life Rev.*, 2020, **32**, 59 – 94.

7 M. Rapacioli, F. Calvo, C. Joblin, P. Parneix, D. Toublanc and F. Spiegelman, *Astron. Astrophys.*, 2006, **460**, 519–531.

8 M. Rapacioli, F. Calvo, F. Spiegelman, C. Joblin and D. J. Wales, *J. Phys. Chem. A*, 2005, **109**, 2487–2497.

9 M. Bartolomei, F. Pirani and J. M. C. Marques, *J. Phys. Chem. C*, 2017, **121**, 14330–14338.

10 K. Bowal, P. Grančič, J. W. Martin and M. Kraft, *J. Phys. Chem. A*, 2019, **123**, 7303–7313.

11 F. Calvo and E. Yurtsever, *Phys. Chem. Chem. Phys.*, 2020, **22**, 12465–12475.

12 B. Hartke, *J. Phys. Chem.*, 1993, **97**, 9973–9976.

13 Y. Xiao and D. E. Williams, *Chem. Phys. Lett.*, 1993, **215**, 17 – 24.

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

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

16 D. J. Wales and M. P. Hodges, *Chem. Phys. Lett.*, 1998, **286**, 65–72.

17 B. Hartke, *Z. Phys. Chem.*, 2000, **214**, 1251.

18 R. L. Johnston, *Dalton Trans.*, 2003, 4193–4207.

19 H. Takeuchi, *J. Chem. Inf. Model.*, 2007, **47**, 104–109.

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

21 D. Ferro-Costas and A. Fernández-Ramos, *Front. Chem.*, 2020, **8**, 16.

22 R. Podeszwa, *J. Chem. Phys.*, 2010, **132**, 044704.

23 E. M. Pérez and N. Martín, *Chem. Soc. Rev.*, 2015, **44**, 6425–6433.

24 X. T. Zheng, A. Ananthanarayanan, K. Q. Luo and P. Chen, *Small*, 2015, **11**, 1620–1636.

25 S. Y. Lim, W. Shen and Z. Gao, *Chem. Soc. Rev.*, 2015, **44**, 362–381.

26 K. J. Mintz, Y. Zhou and R. M. Leblanc, *Nanoscale*, 2019, **11**, 4634–4652.

27 Y. Xiong, J. Schneider, E. V. Ushakova and A. L. Rogach, *Nano Today*, 2018, **23**, 124 – 139.

28 Z. Peng, X. Han, S. Li, A. O. Al-Youbi, A. S. Bashammakh, M. S. El-Shahawi and R. M. Leblanc, *Coord. Chem. Rev.*, 2017, **343**, 256 – 277.

29 M. Semeniuk, Z. Yi, V. Poursorkhabi, J. Tjong, S. Jaffer, Z.-H. Lu and M. Sain, *ACS Nano*, 2019, **13**, 6224–6255.

30 S. Chen, N. Ullah, T. Wang and R. Zhang, *J. Mater. Chem. C*, 2018, **6**, 6875–6883.

31 M. A. Sk, A. Ananthanarayanan, L. Huang, K. H. Lim and P. Chen, *J. Mater. Chem. C*, 2014, **2**, 6954–6960.

32 M. Vatanparast and Z. Shariatinia, *Struct. Chem.*, 2018, **29**, 1427–1448.

33 F. Zasada, W. Piskorz, P. Stelmachowski, P. Legutko, A. Kotarba and Z. Sojka, *J. Phys. Chem. C*, 2015, **119**, 6568–6580.

34 S. Sanyal, A. K. Manna and S. K. Pati, *J. Phys. Chem. C*, 2013, **117**, 825–836.

35 J. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.

36 R. A. Kendall, T. H. Dunning and R. J. Harrison, *J. Chem. Phys.*, 1992, **96**, 6796–6806.

37 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.

38 S. Boys and F. Bernardi, *Mol. Phys.*, 1970, **19**, 553–566.

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

40 A. Halkier, T. Helgaker, P. Jorgensen, W. Klopper, H. Koch, J. Olsen and A. K. Wilson, *Chem. Phys. Lett.*, 1998, **286**, 243–252.

41 A. Halkier, T. Helgaker, P. Jorgensen, W. Klopper, J. Olsen and A. K. Wilson, *Chem. Phys. Lett.*, 1999, **302**, 437–446.

42 P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623–11627.

43 A. V. Marenich, S. V. Jerome, C. J. Cramer and D. G. Truhlar, *J. Chem. Theory Comput.*, 2012, **8**, 527–541.

44 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin,

V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09 Revision E.01*, Gaussian Inc. Wallingford CT 2009.

45 M. Bartolomei, F. Pirani and J. M. C. Marques, *J. Comput. Chem.*, 2015, **36**, 2291–2301.

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

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

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

49 M. Bartolomei, F. Pirani and J. M. C. Marques, *Phys. Chem. Chem. Phys.*, 2019, **21**, 16005–16016.

50 J. Nocedal, *Math. Comp.*, 1980, **35**, 773–782.

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

52 J. Marques, F. Pereira, J. Llanio-Trujillo, P. Abreu, M. Albertí, A. Aguilar, F. Pirani and M. Bartolomei, *Phil. Trans. R. Soc. A*, 2017, **375**, 20160198.

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

54 F. B. Pereira, J. M. C. Marques, T. Leitão and J. Tavares, Advances in Metaheuristics for Hard Optimization, Springer Natural Computing Series, Berlin, 2008, pp. 223–250.

55 J. Hernández-Rojas, F. Calvo and D. J. Wales, *Phys. Chem. Chem. Phys.*, 2016, **18**, 13736–13740.

56 W. Humphrey, A. Dalke and K. Schulten, *J. Molec. Graphics*, 1996, **14**, 33–38.

57 C. I. M. Santos, I. F. A. Mariz, S. N. Pinto, G. Gonçalves, I. Bdikin, P. A. A. P. Marques, M. G. P. M. S. Neves, J. M. G. Martinho and E. M. S. Maçôas, *Nanoscale*, 2018, **10**, 12505–12514.

58 F. Jiang, D. Chen, R. Li, Y. Wang, G. Zhang, S. Li, J. Zheng, N. Huang, Y. Gu, C. Wang and C. Shu, *Nanoscale*, 2013, **5**, 1137–1142.