Hydrogenation of C 24 Carbon Clusters: Structural Diversity and Energetic Properties

Paula Pla, Clément Dubosq, Mathias Rapacioli, Evgeny Posenitskiy, Manuel Alcamí, Aude Simon

Hydrogenation of C_{24} carbon clusters : structural diversity and energetic properties

Paula Pla,[†] Clément Dubosq,[‡] Mathias Rapacioli,[‡] Evgeny Posenitskiy,[¶] Manuel Alcamí,[†,§,‖] and Aude Simon[*,‡]

[†]*Departamento de Química, Universidad Autónoma de Madrid, Módulo 13, 28049 Madrid, Spain*

[‡]*Laboratoire de Chimie et Physique Quantiques (LCPQ), Fédération FeRMI, Univ. Toulouse UT3 & CNRS, UMR5626, 118 Route Narbonne, F-31062 Toulouse, France*

[¶]*Laboratoire Collisions Agrégats et Réactivité (LCAR), Université de Toulouse (UPS) and CNRS, UMR5589, 118 Route de Narbonne, F-31062 Toulouse, France*

[§]*Instituto Madrileño de Estudios Avanzados en Nanociencia (IMDEA-Nanociencia), 28049 Madrid, Spain*

[‖]*Institute for Advanced Research in Chemical Sciences (IAdChem), Universidad Autónoma de Madrid, 28049 Madrid, Spain*

E-mail: aude.simon@irsamc.ups-tlse.fr

Phone: +123 (0)123 4445556. Fax: +123 (0)123 4445557

Abstract

This work aims at exploring the potential energy surfaces of $C_{24}H_{n=0,6,12,18,24}$ using the genetic algorithm in combination with the density functional based tight binding (DFTB) potential. The structural diversity was analysed using order parameters, in particular the sum of the numbers of five- and six-carbon rings $R_{5/6}$. The most abundant and lowest energy population was designated as the flakes population (isomers of variable shapes, large $R_{5/6}$ values), characterized by an increasing number of spherical isomers when n_H/n_C increases. Simultaneously, the fraction of the pretzels population (spherical isomers, smaller $R_{5/6}$ values) increases. The fraction of cages population (largest $R_{5/6}$ values) remains extremely minor while the branched population (smallest $R_{5/6}$ values) remains the highest energy population for all n_H/n_C ratios. For all $C_{24}H_{n=0,6,12,18,24}$ clusters, the evolution of the carbon ring size distribution with energy clearly shows the correlation between the stability and the number of 6-carbon rings. The average values of the ionization potentials of all populations were found to decrease when n_H/n_C increases, ranging from 7.9 downto 6.4 eV. This trend was correlated to geometric and electronic factors, in particular to carbon hybridization. These results are of astrophysical interest especially regarding the role of carbon species in the gas ionization.

1 Introduction

Among cosmic elements, carbon is the second most abundant one (after oxygen) among those able to form bonds of various natures and thus to participate in cosmic molecular complexity through top-down [1] (destruction) and bottom-up [2] (growth) processes. Given its rich allotropy, a diversity of interstellar carbonaceous molecules and dust grains were shown to contain a large fraction of the carbon budget in the galaxy. We can cite nano-diamonds, fullerenes, polyaromatic structures or (hydrogenated)- amorphous carbon structures. [3] These species were identified thanks to the interplay between spectroscopic observations, experi-

mental and theoretical studies. However, more studies are mandatory to refine the assignment of interstellar features attributed to carbon compounds but whose precise structures remain elusive.

First, the need to identify the carriers of the aromatic infrared bands (AIBs), a set of mid-IR emission bands observed ubiquitously in the interstellar medium (ISM)[4,5] motivated many experimental and theoretical studies. In particular, quantum chemistry revealed a powerful tool to compute many IR spectra of individual polyclic aromatic hydrocarbons (PAHs) of different sizes and shapes, neutral and charged, incorporating heteroatoms or complexed to a metal atom.[6,7] However, the AIBs are constituted of bands but also of broad plateaus. In particular, IR spectroscopic observations of fullerene-rich planetary nebulae (PNe) exhibit a broad plateau with substructure in the 6–9 µm range.[8–12] The origin of this plateau remains elusive and various hypotheses were formulated about its possible carriers, including hydrogenated amorphous carbon (HAC) compounds,[12] very small grains, that can possibly be PAH clusters.[13–15] Similarly, the 8 and 12 µm plateau features of some proto-planetary nebulae (PPNe) were assigned to the vibrations of alkanes or alkyle side groups pertaining to large carbon particles,[16,17] whose details also remain unclear. Very recently, quantum calculations showed that the isomers' population of C_{60} containing the most spherical isomers with a high fraction of sp^2 carbon atoms was a good candidate to carry the 6–9 µm plateau[18] features whereas 8 and 12 µm plateau could hardly be reproduced. These observations and their various interpretations therefore motivate further studies to unravel the molecular nature of the carriers of these spectral features, in particular addressing the possible contribution of hydrogenated carbon compounds.

Second, the UV bump[19] is a broad ultraviolet absorption bump observed on the ISM extinction curves in systems ranging from the Milky Way to high-redshift galaxies, with a position centered at 217.5 nm (5.77 eV or 4.65 μm^{-1}) stable to ∼2 % with relatively large variations in width (0.75 to 1.3 μm^{-1}) in different lines of sight.[20–22] Interestingly, the variations in the UV bump width and shape have been related to the presence of defects within the sp^2

conjugated carbon network. Several theoretical and experimental studies were subsequently carried out to unravel the structural origin of these variations. The effects of disorder on the optical absorption spectra of graphitic grains were modeled,[23] introducing disorder by replacing sp^2 carbons by sp^3 and sp^1 atoms. Interestingly, the authors showed that introducing a moderate degree of such disorder was sufficient to cover the vast majority of astronomical width variations. More generally, the contribution of more disordered species was further elaborated under the form of hydrogenated amorphous carbon (HAC)[24,25] and soot particles [see Ref.[26] and references therein]. In an attempt to establish correlations between carbon grain morphology and spectral data, Rotundi *et al.*[27] proposed that carbon nanostructures ordered on the micrometer scale could be better candidates to interpret the UV bump rather than graphitic particles. For instance, recent quantum modeling showed that a population of carbon spherical isomers with a high fraction of conjugated sp^2 atoms could contribute to the astronomical UV bump.[28]

Finally, large gas phase carbon-based molecules (10-100 atoms) have been proposed to carry some of the Diffuse Interstellar Bands (DIBs),[29] a series of weak absorption bands observed in the near IR-visible range on the extinction curve of our Galaxy.[30] Among about 500 bands defined as DIBs, only five were assigned to C_{60}^+. As the assignment of the others remains elusive [see ref.[31] and references therein], it motivates the search for other candidates, and among carbonaceous-based clusters and molecules in particular. For instance, long carbon chains,[32] polyacenes,[31] PAHs[33] and fullerene-based compounds[34] have been proposed.

In their previous studies, Dubosq et al.[18,28] studied the structural diversity of pure carbon clusters $C_{n=24,42,60}$ as well as their infrared and optical properties in order to investigate their contribution to astronomical IR and UV-visible features. As put forward by the authors, some astronomical features could not be accounted for, revealing the need to consider the presence of hydrogenated carbon clusters. This motivated us to explore and analyse the structural diversity of hydrogenated carbon clusters, starting with the smallest number of carbons studied by Dubosq et al.[18,28] $C_{24}H_{(n=0,6,12,18,24)}$, focusing on the lowest energy

isomers. Such clusters are relevant for astrophysics as they may result from the destruction of very small carbon grains in photodissociation regions, such processes leading to small carbonaceous species.[35] They can also be invoked as possible intermediates during top-down (dissociation) or bottom-up (growth) processes concerning the chemical evolution of large stable molecules, such as PAHs or fullerenes. Indeed, in a top-down scenario, the formation of carbon clusters could result from the photochemical evolution of PAHs that would undergo dehydrogenation followed by carbon skeleton rearrangement and C_2 loss.[36–38] The formation of hydrogenated carbon clusters may also result from C or C_2 accretion as proposed in the case of fullerene.[39] In a bottom-up scheme, hydrogenated carbon clusters may grow by C or C_2 accretion in a similar pattern as the one proposed in the case of fullerene[39] and by H accretion, as suggested to possibly lead to the formation of fullerane from fullerene[34], H being highly reactive and the most abundant element in the ISM. Globally, there is a need for characterizing energetic properties of hydrogenated carbonaceous clusters to state if they should be regarded, in addition to intermediates between PAHs and fullerenes,[36] as significant compounds of the interstellar medium resulting from an equilibrium between destruction and formation processes. We present the theoretical procedure in Section 2, its application to a test-case system, namely C_{24}, together with a discussion about appropriate order parameters in Section 3. Finally, the results are reported in Section 4.

2 Theoretical methods

2.1 Electronic structure calculations

We have used the density functional based tight binding (DFTB)[40–42] method to describe the electronic system. It results from a compromise between accuracy and computational cost. This scheme allows to describe carbonaceous systems with various hybridization orders and C-H chemical bonding, while preserving a computational cost much lower than *ab initio*

methods. In the present work, we performed global exploration using the zeroth order DFTB (DFTB0) hamiltonian[40,41] and subsequent local optimization using the second order (self consistent charge or SCC-) DFTB with the mio set of parameters,[42] a Fermi temperature of 500 K to avoid some SCC convergence issues and an empirical dispersion correction contribution.[43] One should note that several dispersion corrections have been developed for the DFTB scheme[43–45] and that the chosen one does not present significant short range contribution at bond distance that may affect hybridization degres, mostly driven by the other contributions to the DFTB energy. All DFTB calculations were performed with the deMonNano code.[46]

2.2 Genetic algorithm exploration

The structures of C_{24}, $C_{24}H_6$, $C_{24}H_{12}$, $C_{24}H_{18}$ and $C_{24}H_{24}$ clusters obtained in this work were generated using a genetic algorithm (GA). GAs, which are based on the principles of evolution by natural selection,[47,48] are known to be efficient methods for the search of the global minimum structure of clusters and nanoparticles.[49] In this work, the Atomic Simulation Environment (ASE) software package[50] has been interfaced with the deMonNano code.[46] The GA module of ASE has been used to obtain the minimum energy structures of different clusters such as metal clusters in MOFs,[51] supported nanostructures[52] or mixed metal halide ammines.[53]

The idea to explore the potential energy surface (PES) used herein was to benefit from the stochastic nature of genetic algorithms to create a population of structures of each cluster stoichiometry. Structures are relaxed at the DFTB0 level during the run of the GA as stated below. The final structures obtained by the GA/DFTB0 algorithm are further locally optimized at the SCC-DFTB level.

The details of a GA run are the following. First, we have to generate a set of random starting candidates (initial population) which was fixed at 20 structures. Each of these

structures was generated by randomly positioning a set of carbon and hydrogen atoms in a box of 11.2 by 11.2 by 5 Å. Second, the structures are relaxed using forces computed at the DFTB0 level and relaxation stops when gradient value is lower than 0.01 eV/Å or when the number of steps reaches 1000. This initial population constitutes the first parent population. The GA run then consists in n_c cycles. At each cycle, this population is updated by considering n_o offsprings. To generate an offspring, two structures are randomly selected from the parent population and each of them is cut into two fragments. Two fragments originating from two different parent structures are combined and relaxed into a new structure. In addition, three types of mutation can occur during the generation of the offspring candidate with a selected probability of 0.3. One of these mutations is a mirror mutation in which one of the fragments is mirrored. The second one is a rattle mutation in which 40% of the atoms in the structure are translated by a random distance below 0.8 Å in a random direction. The third one which works only when different types of atoms compose the structure is permutation of the atomic number of one third of the atoms in the structure. By repeating these processes at each cycle, the population that will serve to produce the new generation of structures, is updated with the most stable structures. We keep track of all the DFTB0 relaxed geometries at each cycle to include them in the final population. In practice, we performed 200 independent GA runs using the DFTB0 potential for each hydrogenated cluster (50 for C_{24}) each of them consisting of n_c =36 cycles (30 for C_{24}) of n_o = 30 new structures. This number of cycles was set to be a good parameter to reach the global minimum structure but not to perform many more cycles once the minimum structure is reached. For each size, this leads to 200*36*30=216000 structures (45000 for C_{24}).

Some of the structures obtained with the GA are fragmented structures which are removed from the final set. The MolMod python library was used to eliminate the fragmented structures based on their corresponding molecular graph.[54] The non-fragmented structures were subsequently optimized at the SCC-DFTB level. Finally, redundant isomers were re-

moved based on an energetic criterion. All isomers were ranked by energy and each structure differing by less than 10^{-6} a.u. from the previous one was considered as redundant and discarded. We may miss some structures in case of exact degeneracy between two structures of different topologies, which seems however quite improbable given the tightness of the energy criterion. We note however that with this approach, mirror symmetric images are considered to belong to the same structural basin.

3 Benchmark of the GA/DFTB algorithm and definition of structural families

Before generating sets of carbonaceous clusters for different H/C ratios, we performed preliminary simulations to check the efficiency of the GA in terms of PES exploration and thus diversity of structures. We chose C_{24} as its PES exploration followed by structural and spectral analysis was achieved in our previous works[18,28]. In those studies, the PES exploration was performed using parallel tempering molecular dynamics (PTMD)[55] with the REBO force field[56], followed by local SCC-DFTB optimization. Despite the fact that, in both studies [Ref.[18] and the present work], the final optimizations were conducted at the same level of theory, they are expected to present differences arising from the two global exploration schemes (PTMD-REBO vs GA-DFTB0).

In the present work, we finally generated 15173 isomers for C_{24}. In Figure 1 (i, a) is reported the 2D distribution of isomers as a function of order parameters based on asphericity β obtained from the Hill-Wheeler analysis,[57] and percentage of sp^2 hybridized carbon that we determined in our previous work in order to define the isomers' families of C_{60}[18]. This 2D distribution reported in Figure 1 (i, a) is continuous, therefore we cannot strictly define distinct families. However, we will use these parameters and the families' names to describe the evolution of the structures on the 2D map. The families' characteristics are reported in Table 1 (a). The ones with the highest sp^2 ratio are called cages. Among the least spherical

isomers ($\beta > 0.3$), those with a still high but lower sp^2 ratio are called flakes while more disordered isomers (fewer sp^2 carbon atoms) are named branched. Isomers with more spherical structures ($\beta < 0.3$) but sp^2 ratios below 75 % are called pretzels. Following these definitions 7472 flakes, 7606 branched, 92 pretzels and 4 cages isomers were generated using the GA algorithm. Interestingly, fully dehydrogenated coronene (coronene is the $C_{24}H_{12}$ compact PAH) was found as the most stable structure for C_{24} as expected[56,58,59] but we have not found the fullerene structure, which is not the most stable isomer but should compete.[60] An example of isomer for each family is reported in Figure 2.

Using the PTMD/REBO exploration followed by local SCC-DFTB optimization, 44341 isomers were found for C_{24}. The 2D distribution for these isomers is reported in Figure 1 (ii, a) and 714 flakes, 37309 branched, 6307 pretzels and 11 cages were obtained (see Appendix in reference[18]). So branched structures are by far the most abundant, characterized by an extended gyration radius distribution (see Figure 1 (ii, d)).

The energy profile shown in Figure 1 (i,b) shows that the flakes generated by the GA algorithm are more stable than the branched isomers (maximum of the energy profiles shifted to the red), while in the case of PTMD/REBO simulations (Figure 1 (ii,b)), practically all structures correspond to branched and pretzel structures that are less stable. It thus seems that the GA algorithm tends to explore more efficiently the lowest energy part of the PES for C_{24}. The GA algorithm is extremely different from the PTMD algorithm due to both its exploration scheme and the choice of the moveclass. In the GA algorithm, at each generation, new structures present patterns very different from those of the previous generation. This appeared to be very efficient to visit distant basins of the PES, in particular as jumps between basins are not affected by the heights of the barriers connecting them. On the other side, in the mentioned PTMD scheme, the chosen moveclass (random displacements of atoms) leads to continuous structural evolutions. As a consequence, more steps are required to connect distant basins. The issue of high energetic barriers crossing is in principle

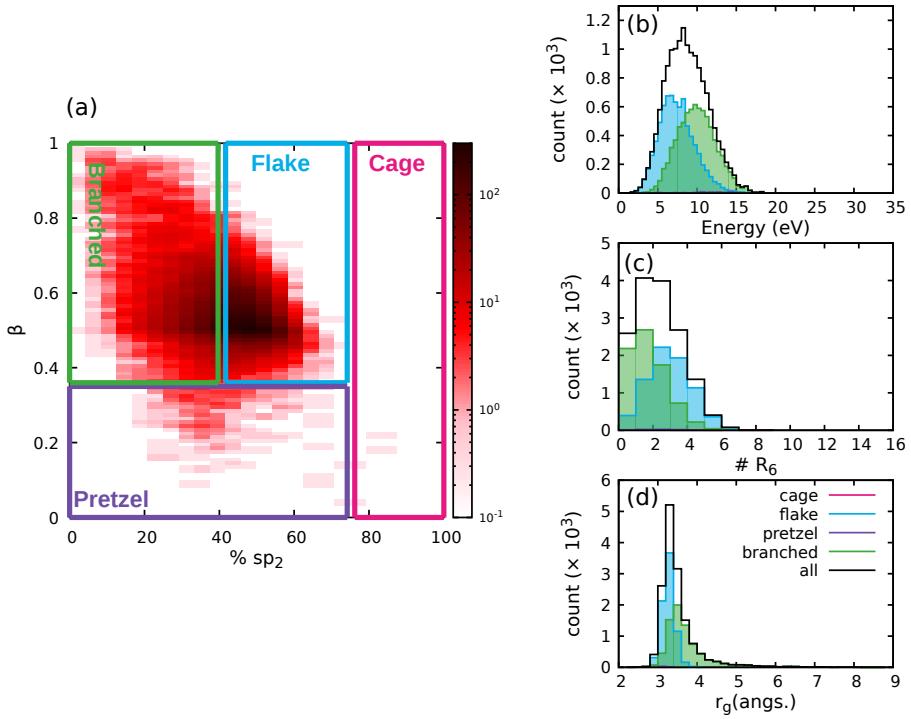
overcome by the use of high temperatures, but this is at the computational price of running many trajectories in parallel. Let us finally mention, that we only compare here results from two specific strategies but that a general discussion about the efficiency of PTMD and GA approaches is out of the scope of the present work, as it would require the two algorithms to be performed with the same PES and to deeply investigate the role of the various parameters (number of temperatures, moveclasses, ...). In the present work, the GA algorithm was adapted to favor the most stable structures as at each GA run, geometries are relaxed at the DFTB0 level and the population is updated with the most stable structures at each GA cycle (see Section 2.2). This better exploration of the lowest energy part is also reflected in the number of structures presenting 6-carbon rings, which should contribute to stabilize the system. When representing the number of structures as a function of the number of 6-carbon rings (Figure 1 (c)), the maximum appears at around 2 rings for GA simulations while it is close to 0 for PTMD/REBO simulations.

Table 1: Definitions of families based either on asphericity β and percentage of sp^2 carbon atoms (a) or on asphericity β and number of 5 and 6-carbon rings $R_{5/6}$ (b).

	(a)		(b)	
	% sp^2	β	$R_{5/6}$	β
Cages	> 100-75%	0.0-1.0	≥ 9	0.0-1.0
Flakes	75-45%	0.3-1.0	3-8	0.3-1.0
Pretzels	75-0%	0.0-0.3	0-8	0.0-0.3
Branched	45-0%	0.3-1.0	0-2	0.3-1.0

The sp^2 hybridization ratio appeared to be a suited parameter for pure carbon clusters. However, in the case of hydrogenated structures, the sp^2 parameter does not univocally describe a structural order. Indeed, in the case of pure carbon clusters, the carbon atoms which are not sp^2 are sp^1, whereas for higher hydrogenation ratios, sp^2 atoms can also be replaced by sp^3 atoms. Therefore we decided to use the sum of the numbers of 5 and 6-carbon rings per structure (designed hereafter as $R_{5/6}$) as a more general order parameter, combined with

(i) GA/DFTB structures



(ii) PTMD/REBO structures

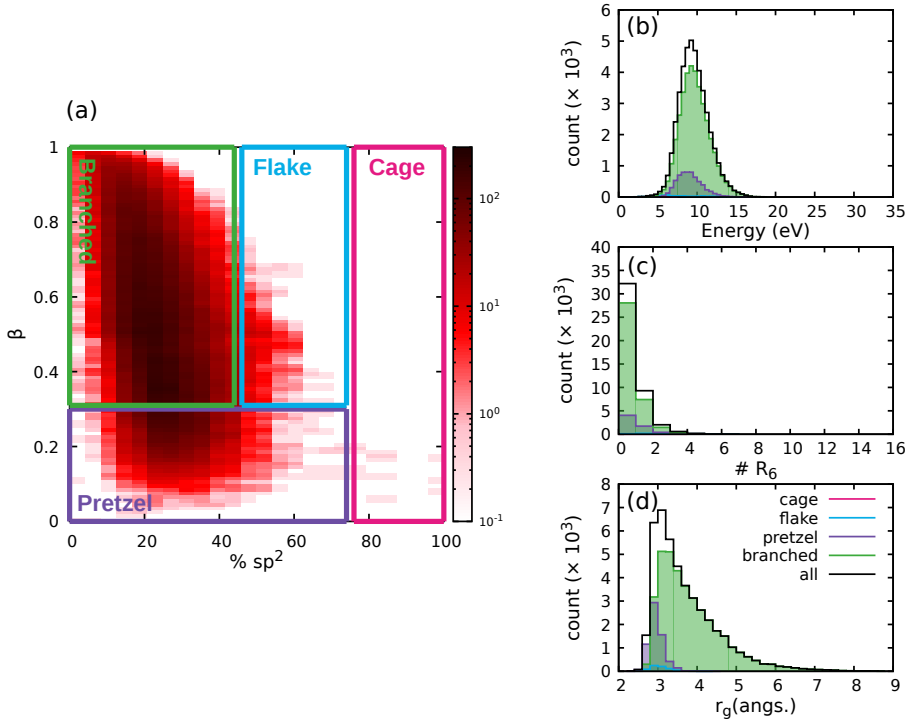


Figure 1: Distributions of isomers of C_{24} obtained from the GA/DFTB (i) and PTMD/REBO (ii) structures. On the left: (a) 2D distribution as a function of the sp^2 hybridization percentage (% sp^2) and the asphericity parameter (β). On the right: 1D distributions as a function of (b) energy, (c) number of 6-carbon rings (#R_6), and (d) gyration radius (r_g).

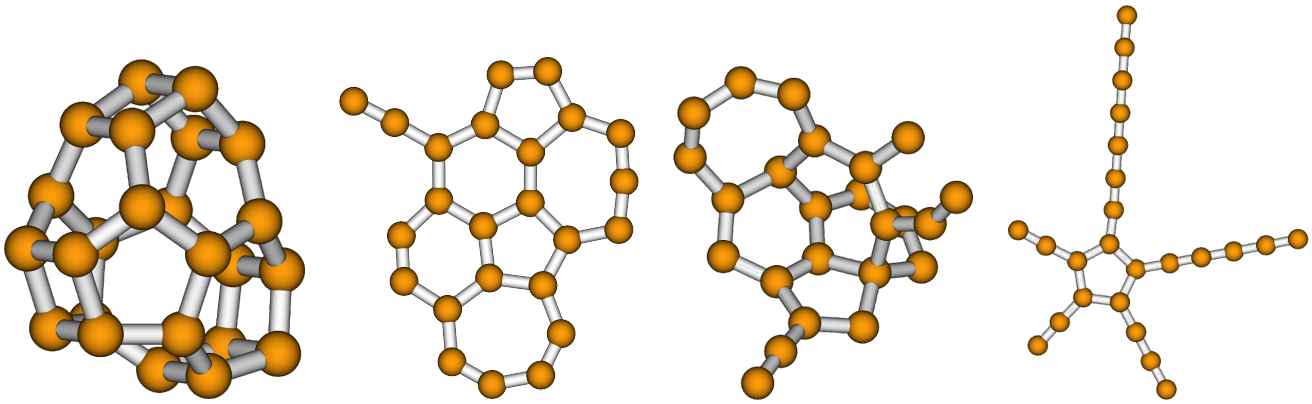


Figure 2: Examples of geometry per family for C_{24}. These correspond to the average of β and sp^2 values for each family.

the asphericity parameter β. The distribution of isomers as a function of these two parameters ($R_{5/6}$, β) for C_{24} can be found in Figure 3 (a) and the new definitions of the families are shown in Table 1 (b).

Using this definition, flakes are the most abundant (62%, see Table 2 1^{st} column), followed by branched structures (37%). The highest density of flakes correspond to mostly planar structures ($\beta \sim 0.5$) with more than two 5 or 6-carbon rings. Branched structures possess less than two 5 or 6-carbon rings. Pretzels, more spherical, are far less abundant (0.6%), and the most ordered structures (cages) are rare (3 structures). In the following, families for hydrogenated clusters are refered to using the ($R_{5/6}$, β) couple of order parameters.

4 Results

4.1 Fragmentation

As previously mentioned, during the optimization procedure some fragmented structures are removed from the final population. As can be seen in Figure 4, the fraction of non-fragmented structures decreases when the hydrogenation rate increases. This highlights the

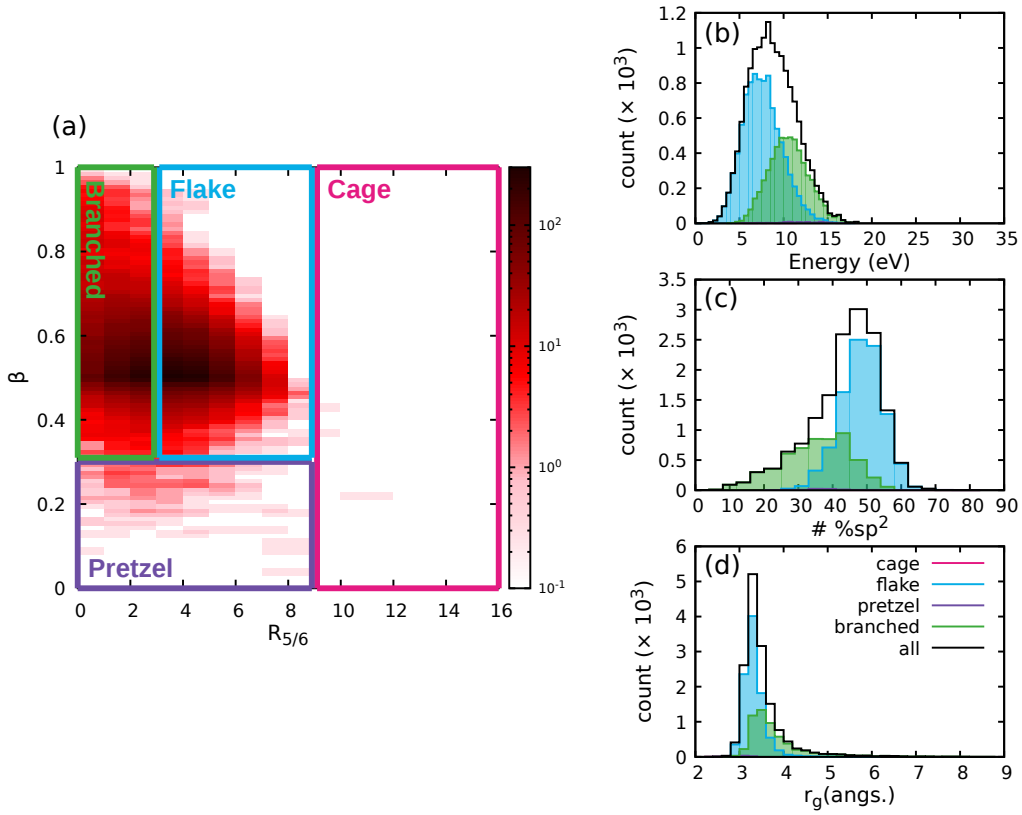


Figure 3: Distributions of C_{24} isomers. Left: 2D distribution as a function of $R_{5/6}$ and β. Right: 1D distributions as a function of energy (b), sp^2 hybridization ratio (c), and gyration radius r_g (d).

difficulty in generating structures with a high H/C ratio.

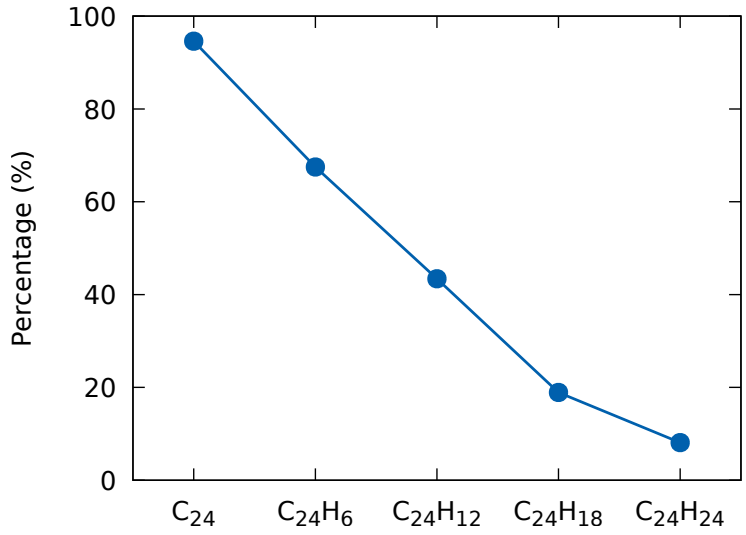
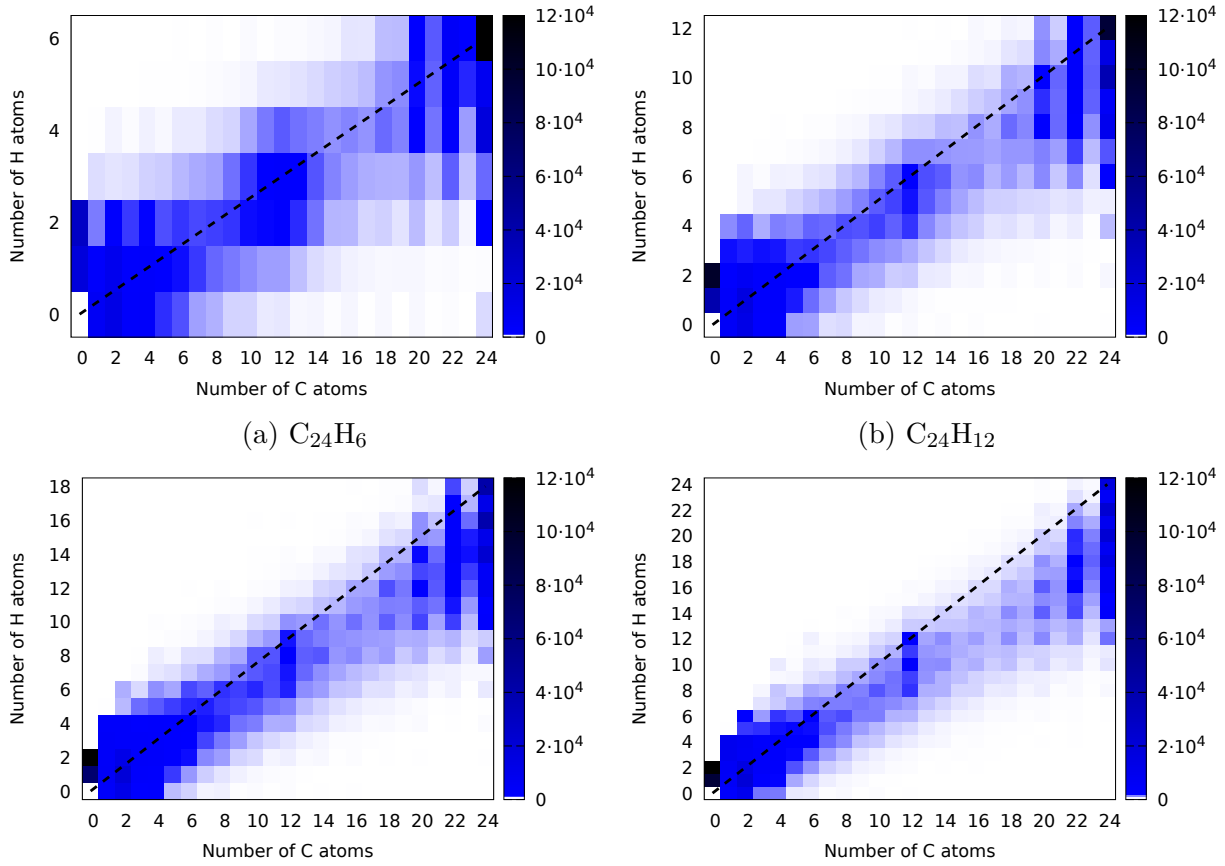


Figure 4: Percentage of non-fragmented structures of $C_{24}H_n$ (n = 0, 6, 12, 18, 24) obtained using the GA.

To get insights into the nature of the fragments we analyzed all structures obtained during simulations. The procedure consisting in counting, for a given hydrogenation degree $C_{24}H_{n=6,12,18,24}$, the fragments with a defined number of C and H atoms. Figure 5 shows the distribution of the fragments for each hydrogenation rate.



(a) $C_{24}H_6$ (b) $C_{24}H_{12}$

(c) $C_{24}H_{18}$ (d) $C_{24}H_{24}$

Figure 5: Distribution of fragments as a function of the number of C and H atoms for $C_{24}H_6$(a) , $C_{24}H_{12}$ (b), $C_{24}H_{18}$ (c) and $C_{24}H_{24}$ (d). The dashed line was added to guide the eye into structures with the same H/C ratio as the parent structure.

Interestingly, as indicated by the shapes of the distributions in Figure 5, the H/C ratio of the fragmented structures remains similar to the initial set. We observe three areas with a higher number of fragments. The first one corresponds to small molecules which must be linear or small rings such as a pentagon or a hexagon, since the maximum number of carbon atoms is 6. These molecules present also a few H atoms. The second one corresponds to structures with a high number of C atoms mainly 24, 22 and 20. For simulations of $C_{24}H_{24}$, the fragments $C_{12}H_{12}$ and $C_{12}H_{10}$ are especially abundant, with again a similar H/C ratio as the parent. In the GA algorithm, a particular fragment is more abundant because its formation is favored in the first offspring generations of the simulation. Afterwards, the

probability for it to be replaced by another more stable structure in the subsequent generation steps becomes low. It seems that the GA algorithm favors fragments of similar stoichiometry as the parent structure, which can be an artefact of the algorithm. Interestingly, in particular for fragments with more than 12 C atoms, those with an even number of C atoms are major than those with an odd number of C atoms. For these fragments, an even number of H atoms is more favorable, leading to a stable closed-shell configuration. Many fragments observed in the simulations correspond to H atoms and H_2 molecules and their number increases as the number of H atoms increases in the simulation. Although the color bar in the plots is set to a maximum of 120000 in all panels to simplify comparison, the number of H_2 molecules is higher than that one in some cases. We have obtained 28366 H_2 molecules for $C_{24}H_6$, 100539 for $C_{24}H_{12}$, 232583 for $C_{24}H_{18}$ and 410496 for $C_{24}H_{24}$. The number of isolated H atoms is always smaller but it also increases with more H atoms in the simulation. The number of H_2 molecules found is larger than the number of non-fragmented structures for $C_{24}H_{12}$, $C_{24}H_{18}$ and $C_{24}H_{24}$ but lower than the number of non-fragmented structures for $C_{24}H_6$.

To summarize this section, the fragment analysis shows that the final number of stable structures decreases when the hydrogenation degree increases and that fragments with even numbers of carbon and hydrogens are favored.

4.2 Evolution of structural diversity as a function of hydrogen ratio

4.2.1 Overview

The total number of non fragmented and non redundant structures for $C_{24}H_{n=6,12,18,24}$ as well as isomers per family, are reported in Table 2. We must specify here that the total number of isomers found by this technique remains orders of magnitude smaller that the number of reasonable structures that can be formed for each stoichiometry, all the more as the number of hydrogen atoms increases. This is probably due to the sampling limitation

inherent to the algorithm, and certainly to the increase of the fragmentation efficiency when the hydrogenation rate increases (Figure 4). The 2D distributions of these isomers as a function of the order parameters β and $R_{5/6}$ for all hydrogenation degrees are reported in Figure 6, with examples of isomers for each family in Figure 7. The distributions of isomers as a function of energy are reported in Figure 8.

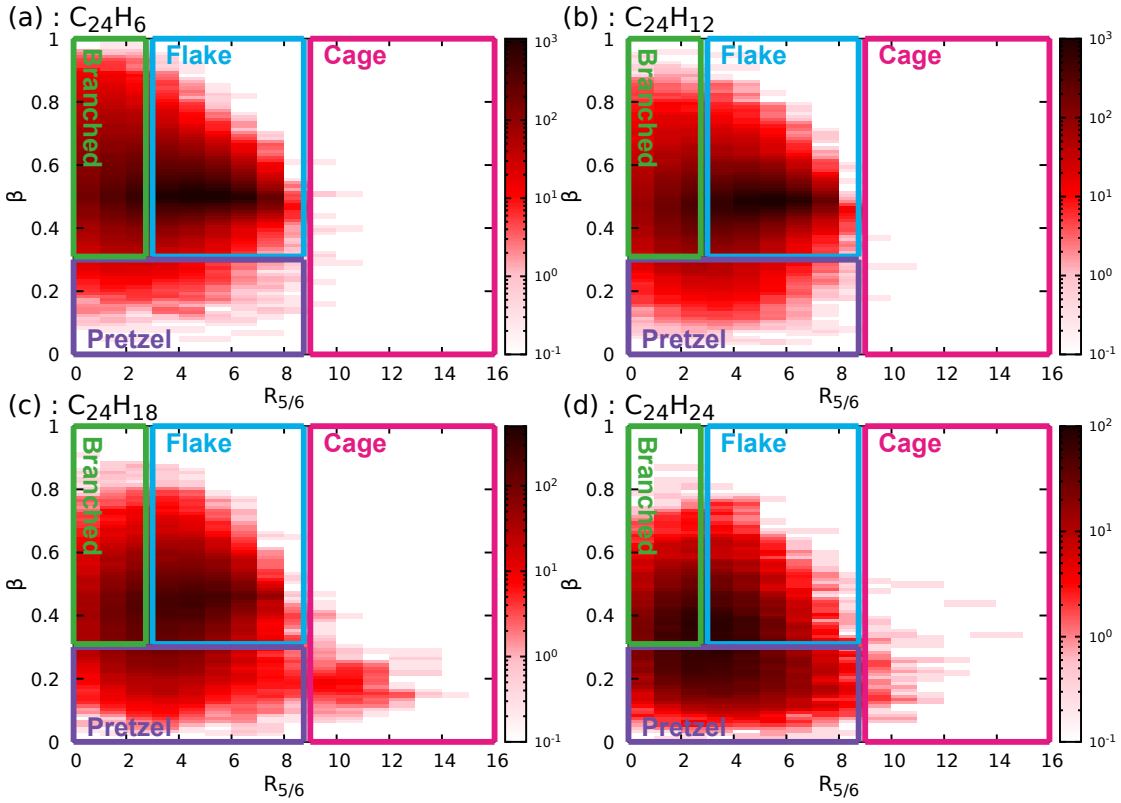


Figure 6: 2D distributions of $C_{24}H_6$ (a), $C_{24}H_{12}$ (b), $C_{24}H_{18}$ (c) and $C_{24}H_{24}$ (d) isomers as a function of the order parameters $R_{5/6}$ and β.

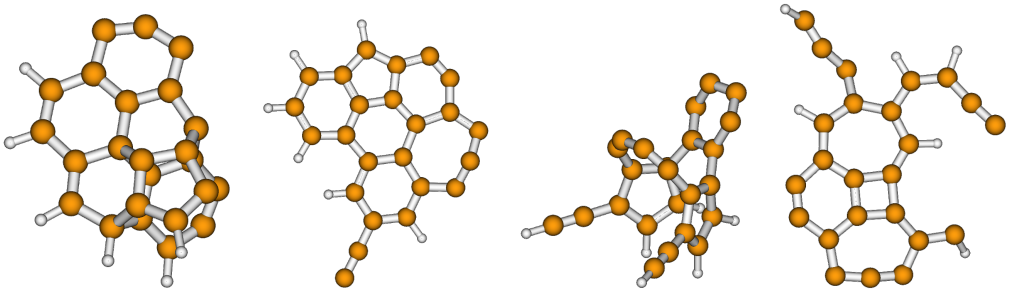
As can be seen in Table 2, up to n_H=18, flakes are major (62% to 79%), followed by branched (19% to 37%). In terms of stability, flakes are also the most stable structures (see Figure 8), followed by pretzels and branched. When n_H=24, the flakes' population remains major but to lesser extent (45%) while the pretzel's population increases to become similar to the branched population (28% *vs* 27%). Interestingly, flakes and pretzels's populations have

Table 2: Total number of non fragmented and non redundant structures and numbers of isomers per family obtained with the GA algorithm for all $C_{24}H_n$ systems.

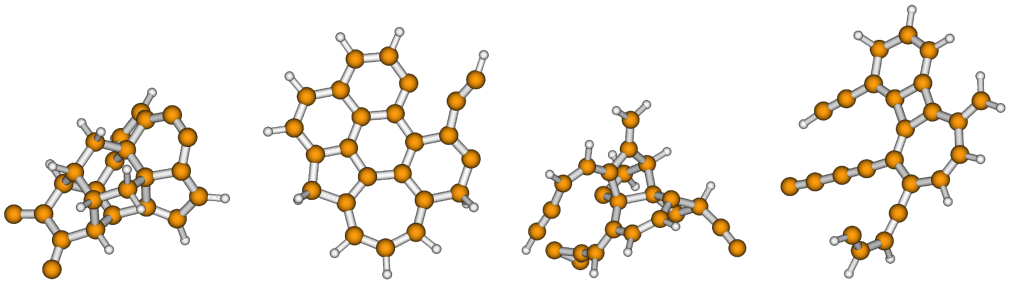
n	0	6	12	18	24
total	15173	68896	52213	26019	10894
flakes	9449 (62%)	49989 (73%)	41264 (79%)	17889 (69%)	4924 (45%)
branched	5626 (37%)	18191 (26%)	9722 (19%)	5578 (21%)	2889 (27%)
pretzel	95(0.6%)	704 (1 %)	1224 (2.3%)	2293 (9%)	3009 (28%)
cage	3	12	5	259 (1%)	72

similar energy distributions (Figure 8) while branched structures remain higher in energy. For all hydrogenation degrees, cages, i.e. the most ordered and spherical isomers, represent a tiny fraction of the population.

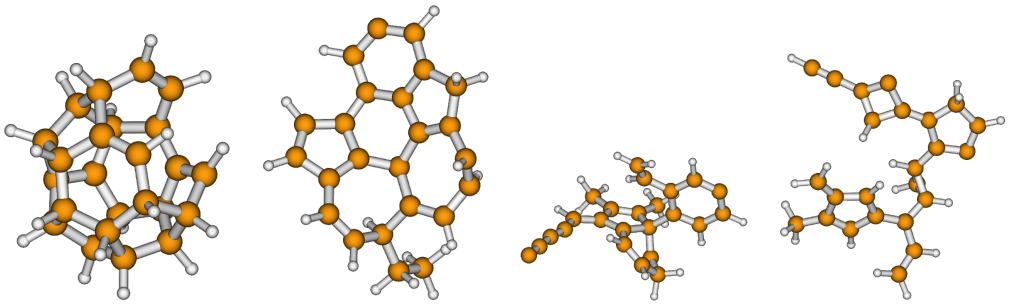
As mentioned in Section 3, fully dehydrogenated coronene, is the most stable structure for C_{24} with the GA algoritm, as expected.[56,58,59] However, the fullerene structure[60] has not been found. It is interesting to underline that the most stable structures found for hydrogenated clusters correspond to derivatives of coronene (flakes with seven hexagonal rings) except for $C_{24}H_{24}$ where one of the rings has broken in favor of two methyl groups (see Figure 9 and Figure S1 of the Supplementary Information (SI) where the six most stable structures for the flakes families for all stoichiometries are reported). As mentioned hereabove, in previous works[56,58,59], the most stable isomer found for C_{24} was dehydrogenated coronene except when using explicitly correlated methods where the cage structure is found to be a few kcal/mol more stable than dehydrocoronene.[60] For $C_{24}H_6$, three isomers are found very close in energy (less than 0.01 eV) which correspond to isomers in which the hydrogen atoms are placed by pairs (see Figure S1 of the SI). These dehydrogenated coronene derivatives containing hydrogen atoms placed by pairs have been previously studied as they are more likely to be found in interstellar conditions.[61] These three isomers are followed by one structure possessing a five-carbon ring, located 0.8 eV above the most stable isomer. Coronene is the most stable isomer of $C_{24}H_{12}$ in agreement with the results by Kosimov



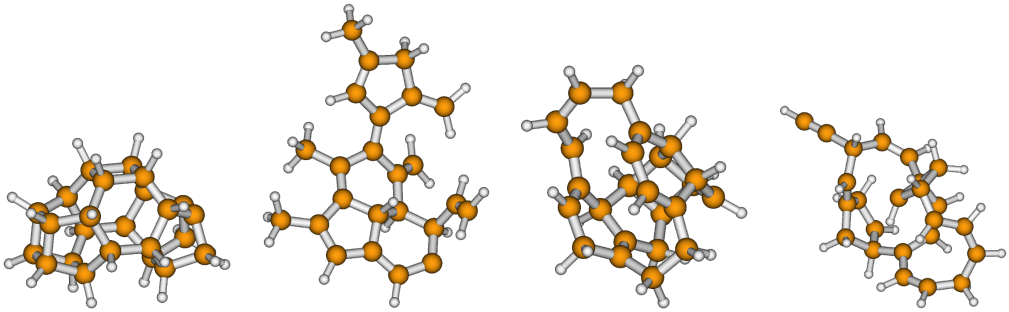
(a) $C_{24}H_6$



(b) $C_{24}H_{12}$



(c) $C_{24}H_{18}$



(d) $C_{24}H_{24}$

cage flake pretzel branched

Figure 7: Examples of isomers for each family for $C_{24}H_{6,12,18,24}$

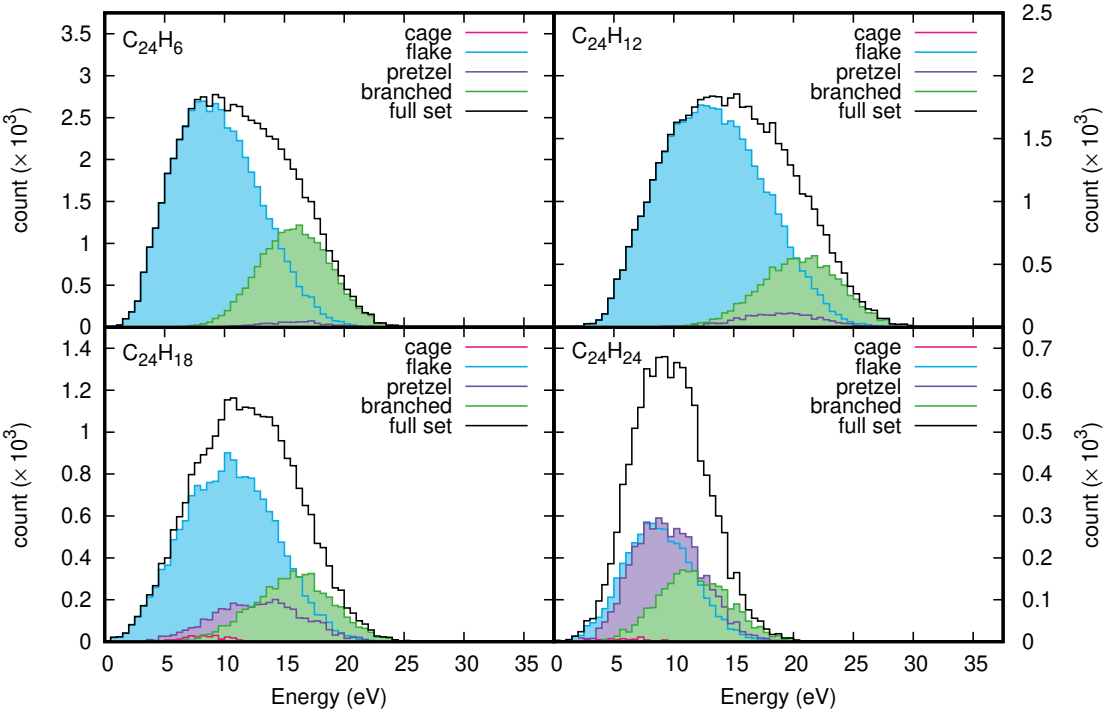
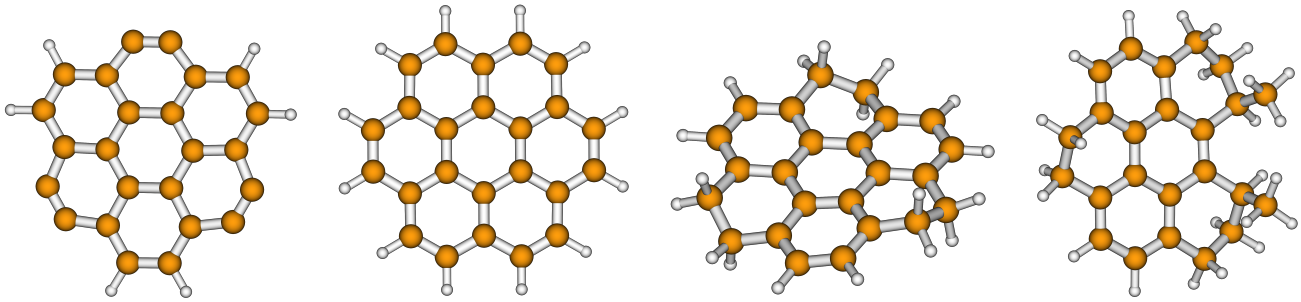


Figure 8: Distributions of $C_{24}H_6$ (a), $C_{24}H_{12}$ (b), $C_{24}H_{18}$ (c) and $C_{24}H_{24}$ (d) isomers as a function of energy

et al..[59] Next isomer obtained by the GA algorithms is located 1.33 eV above. It possesses different ring arrangements, and includes a five-carbon ring as well as the following four isomers (in increasing energetic order (see Figure S1 (b)). Interestingly, the 6^{th} isomer located 1.71 eV above coronene presents a fulvene function. Finally, it is worth noticing that the most stable isomer of $C_{24}H_{18}$ corresponds to the most stable isomer of hydrogenated coronene with 6 additional H atoms.[62] This is also the case of other low energy isomers found by the GA algorithm (three among five within 0.53 eV (see Figure S1 (c)), the two other isomers possessing methyl groups. Interestingly, regarding $C_{24}H_{24}$, the most stable isomer found by the GA algorithm possesses two methyl groups (see Figure S1(d) in the SI) whereas the most stable isomer of hydrogenated coronene with 12 additional H atoms[62] was found 0.31 eV higher in energy after local SCC-DFTB optimisation, and it has not been found with the GA algorithm. The energetic order between these two isomers was confirmed at the DFT level, with an energy difference of 0.11 eV at the B3LYP/6-31G(d,p) level for instance. Interestingly, other similar hydrogenated coronene isomers were found with the GA algorithm (see Figure S1(d), 4^{th} and 6^{th} isomers above the most stable one). The pool of structures obtained with the GA algorithm is certainly not exhaustive, but it should be representative of the structural diversity of the studied systems. Besides, interestingly, it seems that some low energy structures that could hardly have been intuitively guessed, such as the most stable isomer of $C_{24}H_{24}$, are found. The quality of the PES exploration could be investigated tuning the parameters of the GA algorithm. For instance, increasing the size of the box in the Z plane may favor the formation of fullerene-type structures. However, increasing the box size will also favor the formation of fragmented structures, which is not desirable in this study. It is worth mentioning that the same parameters were used for all stoichiometries.



$C_{24}H_6$ $C_{24}H_{12}$ $C_{24}H_{18}$ $C_{24}H_{24}$

Figure 9: Most stable isomers obtained with GA for $C_{24}H_n$ (n = 6, 12, 18 and 24).

4.2.2 Structural evolution as a function of n_H/n_C

In this section, we provide a detailed description of the families' characteristics by ascending order of hydrogen rate in $C_{24}H_n$. The results are discussed in the light of Table 2, Figures 6 and 8, as well as additional Figures in the Supporting Information, that represent the distribution of isomers for each family as a function of the sp^2 ratio (Figure S2) and the gyration radius r_g (Figure S3).

•$C_{24}H_6$

With respect to C_{24}, a larger fraction of flakes structures was obtained for $n_H/n_C = 1/4$ (73% vs 62%) to the detriment of branched structures (26% vs 37%) showing that the presence of a few H atoms favors more ordered structures. The energetics profile (Figure 8) is wider than for C_{24} (Figure 3) indicating a spreading of the structures with less degeneracy. Besides, it is clearly asymmetric with a sharp red increase and a loose blue tail. The maximum density of structures is located at higher energy (9.07 eV vs 8.12 eV in the case of C_{24}). Regarding the electronic and structural properties, adding 6 hydrogen atoms to C_{24} leads to an increase of the sp^2 carbon ratio maximum (68% for $C_{24}H_6$ vs 48% for C_{24}) while the gyration radii distribution remains similar (maximum at 3.3 Å).

•$\mathbf{C_{24}H_{12}}$

When $n_H/n_C = 1/2$, a slightly larger fraction of flakes structures is obtained with respect to $n_H/n_C = 1/4$ (79% vs 73%), the maximum density of structures remaining concentrated in the same region (for a $\beta \sim 0.5$ and $R_{5/6}$ between 3 and 9). The branched population slightly decreases (26% to 19%) whereas the pretzel population slightly increases (1% to 2.3%). Regarding the energy profile, it continues widening, with structures reaching 30 eV above the most stable isomer. The sp^2 fraction maximum increases up to reaching its maximum among all n_H/n_C ratios (85%) while the gyration radius remains at 3.3 Å.

•$\mathbf{C_{24}H_{18}}$

For $n_H/n_C = 3/4$, the fraction of flakes structures decreases (69% vs 79% for $n_H/n_C = 1/2$), that of branched structures remains steady (21% vs 19% for $n_H/n_C = 1/2$), while the fraction of pretzels structures increases (9% vs 2.3% for $n_H/n_C = 1/2$), as well as that of cages (1%) which now appears as a spot on the 2D map (Figure 6 c)). Interestingly, pretzels and cages possess a more spherical character than flakes and branched isomers. Regarding the energy profile, it now tightens, all structures remaining less than 25 eV above the most stable isomer. The maximum of the $\%sp^2$ profile is now located at 77% while the gyration radius decreases down to 3.2 Å. It can be understood as increasing the number of H atoms leads to the existence of a larger number of more spherical and hence more compact isomers.

•$\mathbf{C_{24}H_{24}}$

When $n_H/n_C = 1$, the structural diversity presents a drastic change : the flakes population is still dominant (45% vs 69% for $n_H/n_C = 3/4$) but the fraction of pretzel population increases (28% vs 9% for $n_H/n_C = 3/4$). The energy profiles of both populations become similar with the low energy tail corresponding to the most stable isomers. In line with the

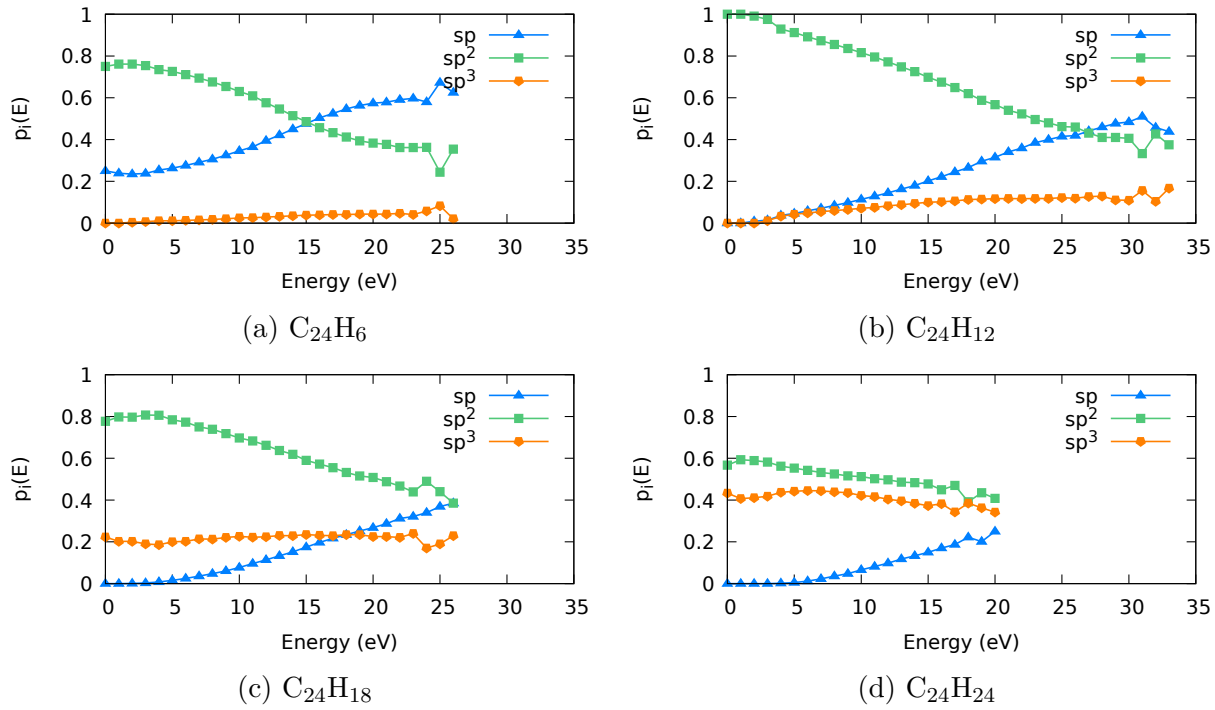
partial conclusions drawn for $C_{24}H_{18}$, increasing the n_H/n_C ratio when it is higher than 1 leads to an increased populations of stable spherical isomers. The fraction of the branched population remains steady while that of the cage decreases with respect to $n_H/n_C = 3/4$. The maximum of the $\%sp^2$ profile is now down to 60%, which is partly due to the presence of an increased fraction of $\%sp^3$ atoms as reported in next subsection. The positions of the gyration radius maximum is stabilized at 3.2 Å corresponding again to a population increase of more spherical and more compact structures due to the high hydrogenation degree.

4.2.3 Evolution of hybridization and rings' distribution as a function of energy

In the following section, we propose an analysis of the evolution of carbon hybridization and rings' distributions as a function of energy for all families of all hydrogenation rates.

•Hybridization and stability

We first analyse the hybridization evolution based on Figure 10. The sp^i (i=1-3) character of the C atoms was obtained using bond order as in our previous work[18]. We however note that the same results were obtained by simply counting the number of neighbouring atoms : a C atom is of sp^i (i=1-3) character if it is covalently bonded to i+1 atoms.



(a) $C_{24}H_6$ (b) $C_{24}H_{12}$

(c) $C_{24}H_{18}$ (d) $C_{24}H_{24}$

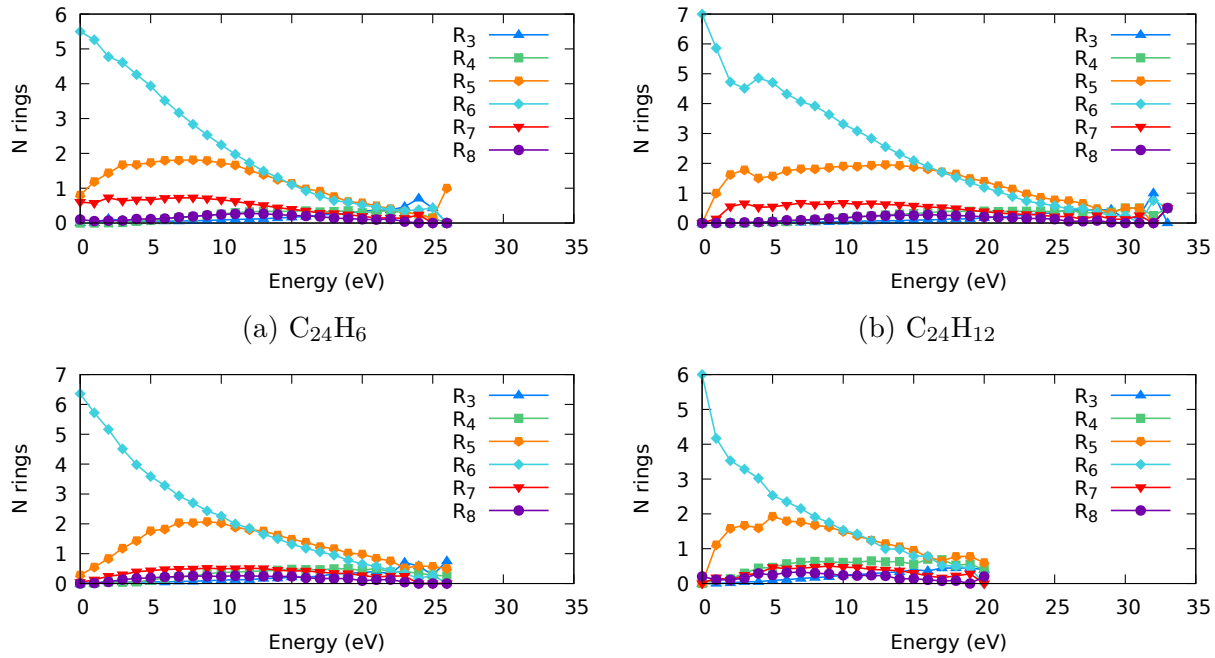
Figure 10: Evolution of the fraction of sp^i (i=1-3) carbon atoms as a function of energy $p_i(E)$ for $C_{24}H_6$ (a), $C_{24}H_{12}$ (b), $C_{24}H_{18}$ (c) and $C_{24}H_{24}$ (d). The data points are obtained averaging the sp^i carbon atoms' fraction for all isomers within 1 eV energy windows. This quantity becomes less valuable in the higher energy region (about 5 eV below the maximum energy) due to the low number of structures in these energy windows.

As can be seen in Figure 10, the sp^2 fraction (p_2) is overall dominant for all hydrogenation rates, which is consistent with the major presence of flakes structures. This predominance is major for n_H=12 where the most stable structures possess 100% of sp^2 carbon atoms. This is not the case for the others: the sp^1 fraction (p_1) amounts to ∼0.25 for n_H=6, while the sp^3 fraction (p_3) is ∼ 0.22 and 0.43 for n_H=18 and 24, respectively. When energy increases, p_2 decreases, which is true for all hydrogenation rates but to various extents. The variation is maximum for n_H=12 (1 down to ∼ 0.38), minimum for n_H=24 (0.57 down to ∼0.41). sp^2 atoms are not major anymore for energies above 15 eV in the case of n_H=6 where p_1 becomes higher (∼ 0.63 vs ∼ 0.35)). In the case of n_H=12, p_1 and p_2 become equal for the very higher energy (above 27 eV). For n_H=18, p_3 is found to be steady along the whole energy range (0.18-0.24) while p_1 increases from 0 to ∼ 0.39. In the case of n_H=24, the smallest variations

are observed. p_3 also remains steady (about 0.40) while p_1 increases from 0.0 to 0.25 at the highest energy (20 eV).

•Rings' distribution and stability

We now discuss the distribution of rings' sizes R_n (n = 3-8 C atoms) as a function of energy. As can be seen on Figure 11, the distributions' evolution looks similar for all hydrogenation rates. $R_{5,6}$ are clearly the largest numbers, all the others R_n (n=3,4,7,8) remaining lower than 1. Interestingly, among these minor ring sizes, 7-carbon rings have more occurrences at lower energies for the two most unsaturated hydrogenated carbon clusters ($n_H = 6, 12$).



(a) $C_{24}H_6$ (b) $C_{24}H_{12}$

(c) $C_{24}H_{18}$ (d) $C_{24}H_{24}$

Figure 11: Evolution rings' distributions as a function of energy for $C_{24}H_6$ (a), $C_{24}H_{12}$ (b), $C_{24}H_{18}$ (c) and $C_{24}H_{24}$ (d). The data points are obtained averaging among all the structures of each energy window set to 1 eV. Again, the high-energy region is poorly described so points outside the trend might appear.

At lower energies, R_6 is clearly major (between 5 and 7) and decreases when energy increases. This indicates that the existence of 6-carbon rings is one of the main stabilizing factors for C_mH_n clusters. At the lowest energy, the decrease of R_6 is correlated to the increase of R_5 and these two rings' sizes disappear at higher energies. For all hydrogenation rates, starting from the minimum energy structures, R_5 increases to reach a maximum value of 2 and then decreases. Interestingly, R_5 remains constant over a wide range of energy ($\sim$ 10 eV for the lowest hydrogenation rates, $n_H = 6, 12$), which is not the case for higher hydrogenation rates ($n_H = 18, 24$). For the latter, R_5 reaches a maximum at $\sim 5\,$eV, before decreasing similarly to R_6.

4.3 Effects of hydrogenation on ionization energies

In this section, we discuss the effect of hydrogenation on the ionization potentials (IPs) of the above discussed families of isomers. Both vertical (VIPs) and "local" adiabatic (LAIPs) ionization potentials were determined for all individual isomers, the former by simply removing one electron of the neutral system and computing the energy difference, the latter by optimizing the cation from the neutral geometry as a starting point, ensuring that it is a minimum, and computing the energy differences. The VIPs' distribution as a function of energy for all families are reported in Figure 12 for $n_H = 0, 12, 24$ and in the Supporting Information (Figure S4) for $n_H = 6, 18$. The corresponding LAIPs' distribution are reported in the Supporting Information, Figures S5 and S6 for n_H=0,6,12,18 and 24 .

The average values (per family and over the total number of isomers) of the VIPs and LAIPs as a function of n_H are reported in Figure 13 (a) and (b) respectively. The corresponding data are reported in Table 3. The energies corresponding to the distribution maxima for both VIPs and LAIPs are reported in Table 4.

Before presenting the results, it is necessary to discuss their accuracy. Indeed, the VIP

of coronene was determined at 7.49 eV, i.e. 0.2 eV above the experimental value (7.29 eV)[63] while that of buckminsterfullerene was determined at 7.4 eV, e. g. 0.2 eV lower than the experimental value (7.6 eV for the most recent value[63]). We computed the VIP of coronene at the DFT level using different functionals and the 6-31G(d,p) basis set, as implemented in the Gaussian16 suit of programs[64]. We found significantly lower energy values than the experimental one and thus than the SCC-DFTB one, i. e. 6.85 eV (B3LYP), 6.77 eV (PBE) and 6.67 eV (B97D3). We computed the VIPs of the most stable isomers of all families found with the GA algorithm at the B3LYP/6-31G(d,p) level of theory and compared them to the SCC-DFTB results. The values are reported in Table S1 of the SI. The DFT values were found to be much lower than the SCC-DFTB values, the amplitude of the discrepancy being very dependent on the hydrogenation degree and on the isomer. However, as verified in the case of coronene, DFT values are not expected to be better than SCC-DFTB ones. To conclude this warning based on the results for coronene and buckminsterfullerene, the absolute IP values reported in the rest of the manuscript should not be regarded with an accuracy better than 0.2 eV. However, the variations of IP's distributions between the different families and the shifts with respect to hydrogenation degrees are expected to provide quantitative trends.

As can be seen in Figure 12 (a), the VIP distributions of C_{24} are dominated by an intense band with maxima at similar energy for the flakes and branched families (8.4 eV). The VIP distribution for the pretzel family is more noisy due to the smaller number of structures, and the maximum was found at 8.6 eV. Regarding the cages, the average VIP was found at lower energy (8.0 eV). Globally, increasing the hydrogenation rate leads to a decrease of the IPs, as detailed below. As can be seen in Figure S4, whatever the family, the VIP distribution of $C_{24}H_6$ is dominated by an intense band, the energy of the maximum depending on the family: it is lower for the flakes (7.6 eV) than for pretzels (7.8 eV) and branched (7.9 eV). Regarding the cages, one finds discrete transitions due to the few number of structures. Similar as for C_{24}, the average value of the VIPs (7.5 eV) is lower than that of the flakes

(7.7 eV). Increasing the hydrogen ratio results in redshifts of the average IPs (both vertical and adiabatic, see Figure 13) and of the IP's distribution maxima. Interestingly, regarding the average values of the VIPs and of the LAIPs over the total number of isomers for each hydrogenation ratio, the decrease is similar from $n_H = 0$ to $n_H = 6$ and from $n_H = 6$ to $n_H = 12$ (-0.6 eV for the VIP), and becomes smaller from $n_H = 12$ to $n_H = 18$ (-0.3 eV for the VIP) up to become negligible from $n_H = 18$ to $n_H = 24$. Besides, as can be noticed in Figure 13, the average value over the flakes populations is very similar to the one over all isomers, which was expected as this population is major. For $n_H = 24$, the most intense band for the flakes and pretzels families becomes clearly divided into two subbands. The energy of the maxima of the lowest energy subbands follow the trends previously observed for lower hydrogenation rates (redshifts) whereas the other one is located above 7.0 eV.

We can make some attempts to interpret these trends using the evolution of the sp^n carbon fraction described in Section 4.2.3. A decrease in the IP may be due to an increase of the sp^2 ratio in conjunction with a decrease of the sp^1 ratio. Indeed, referring to model systems, the IP of ethyne (11.4 eV[63]) is larger than the IP of ethylene (10.5 eV[63]). As previously noticed (Section 4.2.3), the sp^3 ratio increases with the hydrogenation degree. In this case, it can lead to two opposite trends regarding the IPs. In the first case, the sp^3 carbon atoms are under the form of small aliphatic groups such as methyl groups which are electron donor to an aromatic ring (with sp^2 electrons). This tends to decrease the IP. For instance the IP of benzene is 9.24 eV[63] whereas that of toluene is 8.83 eV[63]. This trend was confirmed comparing the VIPs of the $C_{24}H_6$ and $C_{24}H_{24}$ flakes isomers of Figure 7, which were determined to be 7.9 and 6.7 eV at the SCC-DFTB level of theory. In the second case, the sp^3 carbon atoms are part of carbon cycles. This is expected to lead to an increase of the IP. As an example with model systems, the IP of cyclohexane is 9.88 eV,[63] which is 0.62 eV above that of benzene. These trends are confirmed for our systems as the SCC-DFTB VIPs of the $C_{24}H_6$ and $C_{24}H_{24}$ cages isomers of Figure 7 were found to be 7.6 and 8.0 eV, respectively. The two subbands observed in the case of $C_{24}H_{24}$, where the fraction of sp^3 carbon atoms

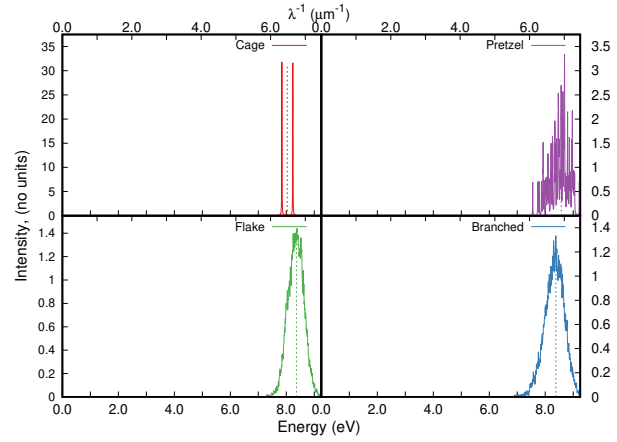
is relatively the most important among all stoichiometries, can be subsequently interpreted using the two types of sp^3 carbon atoms mentionned hereabove : the presence of methyl groups bonded to aromatic cycles lead to a decrease of the IP with respect to the equivalent pure aromatics while non aromatic cycles possessing sp^3 carbon atoms tend to have higher IP values. In the case of the pretzel family, one can also distinguish structures with a majority of cycles, leading to the lower IPs whereas structures with no cycle are also found, leading to higher IP values.

Regarding the LAIP spectra (see Supporting Information, Figures S5 and S6), they are similar to the VIP spectra although : -(i)- they are redshifted, revealing minor geometrical change upon ionization, with a value of the shift of the maxima ranging from -0.1 eV (flakes, $C_{24}H_6$) downto -0.2 eV (branched $C_{24}H_{18}$); -(ii)- additional weaker bands appear at lower energy (below ~ 6 eV).

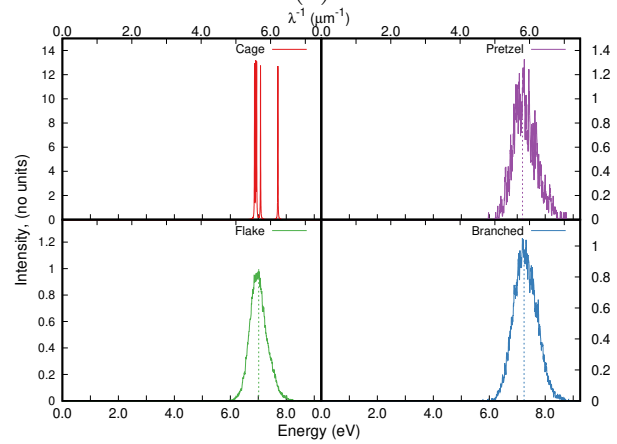
Table 3: Average value of energies (expressed in eV) found for the LAIPs (A columns) and VIPs (V columns) for all families for all hydrogenation rates.

	cages		flakes		pretzels		branched		total	
	A	V	A	V	A	V	A	V	A	V
C_{24}	7.9	8.0	8.0	8.3	8.2	8.5	8.05	8.3	8.0	8.3
$C_{24}H_6$	7.0	7.5	7.4	7.7	7.5	7.8	7.4	7.8	7.4	7.7
$C_{24}H_{12}$	6.5	7.1	6.8	7.0	6.9	7.3	6.9	7.3	6.8	7.1
$C_{24}H_{18}$	6.4	6.6	6.5	6.8	6.6	7.00	6.6	7.00	6.5	6.8
$C_{24}H_{24}$	6.5	7.00	6.4	6.7	6.5	6.8	6.5	6.8	6.5	6.8

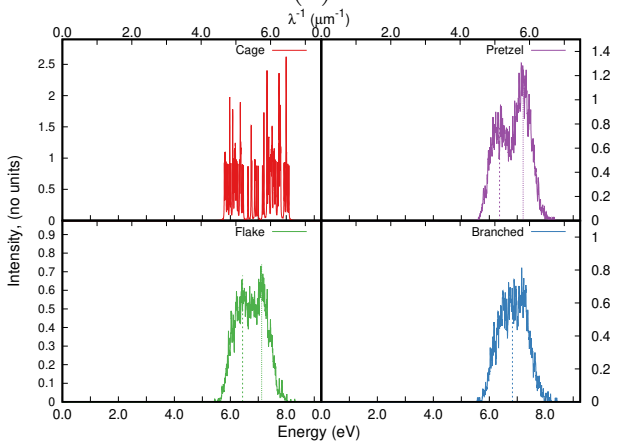
In summary, the ionization energies of the hydrogenated carbonaceous isomers computed in the present work were found lower than those of the most stable well known forms such as coronene or buckminsterfullerene, 7.5 eV and 7.4 eV respectively at the same level of theory, for $n_H/n_C \geq 1/2$. For a given family, the IP tends to decrease when increasing the hydrogen fraction, probably related to the larger number of isomers with increasing fraction of sp^2 carbon atoms with respect to sp^1. In the case of $n_H/n_C = 1$ where a significant number of sp^3 atoms can be found, a second higher energy maximum appears.



(a)

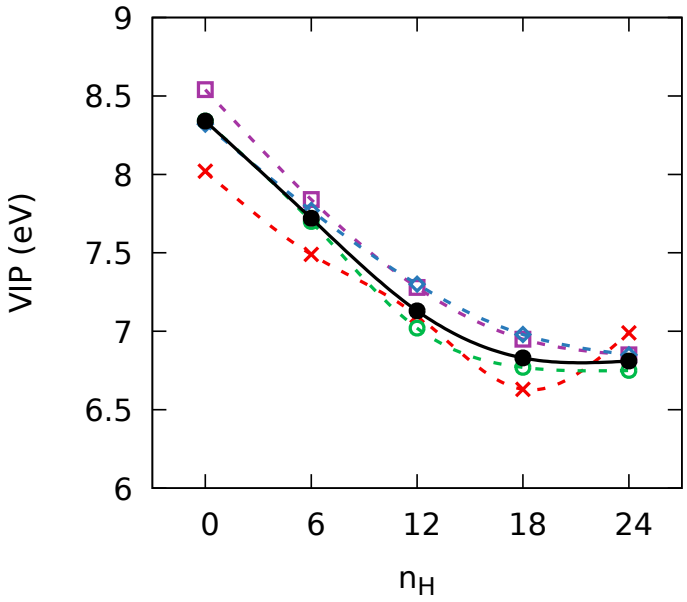


(b)

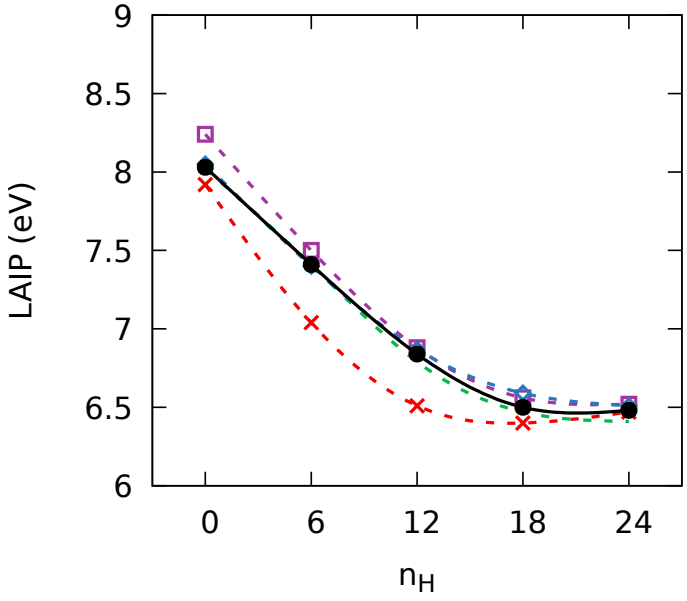


(c)

Figure 12: Vertical ionization potentials' (VIPs) distributions as a function of energy for the four structural families of C_{24} (a) , $C_{24}H_{12}$ (b) and $C_{24}H_{24}$ (c). A lorentzian profile of full width at half maximum (FWHM) of 0.01 eV was applied to the the discrete distribution to improve the appearance. Normalisation with respect to the number of isomers per family was also performed.



(a)



(b)

Figure 13: Evolution of the average values of the VIPs (a) and LAIPs (b) for $C_{24}H_n$) as a function of n_H for the total number of isomers (black full circles) and for each family: cages (red crosses), flakes (green circles), pretzels (purple squares) and branched (blue diamonds). Plain and dashed lines were drawn to guide the eyes.

Table 4: Energies (expressed in eV, and FWHM in parenthesis) of the maxima found for the LAIPs (A columns) and VIPs (V columns) for all families for all hydrogenation rates. The cage family is not present due to the reduced number of isomers producing discrete bands rather than a continuous distribution.

	flakes		pretzels		branched	
	A	V	A	V	A	V
C_{24}	8.2 (0.7)	8.4 (0.7)	8.4 (0.5)	8.6 (0.5)	8.3 (0.8)	8.4 (0.8)
	6.4 (1.3)				6.3 (1.3)	
$C_{24}H_6$	7.5 (0.8)	7.6 (0.8)	7.7 (0.9)	7.8 (0.6)	7.7 (1.0)	7.9 (0.9)
	6.1 (2.4)				5.8 (2.2)	
$C_{24}H_{12}$	6.9 (0.8)	7.0 (0.7)	7.0 (0.9)	7.2 (0.9)	7.1 (1.0)	7.2 (1.0)
$C_{24}H_{18}$	6.6 (0.9)	6.7 (0.9)	6.7 (1.1)	7.00 (1.1)	6.7 (1.1)	6.9 (1.0)
	5.0 (2.6)					
$C_{24}H_{24}$	6.2 (0.8)	6.4 (0.7)	6.2 (0.7)	6.4 (0.6)	6.8 (1.4)	6.8 (1.3)
	7.0 (0.7)	7.1 (0.7)	7.1 (0.8)	7.2 (0.8)		
	4.3 (1.1)		4.6 (1.1)		4.4 (1.3)	

5 Conclusion

In this work, we explored the potential energy surfaces of $C_{24}H_{n=0,6,12,18,24}$ up to 20-25 eV using a genetic algorithm while the electronic structure was determined at the DFTB0 level of theory, followed by further SCC-DFTB local optimization. During these simulations, when n_H/n_C increases, enhanced fragmentation was observed with H_2 as the major fragment. The structural diversity of the non fragmented structures was analysed using order parameters, that we chose as the sum of the numbers of 5 and 6-carbon rings $R_{5/6}$ and the asphericity constant β. The most abundant and lowest energy population correspond to a flakes population, constituted of isomers of variable shapes ($0.3 < \beta \leq 1$) with a large number of 5 or 6-carbon rings ($3 < R_{5/6} \leq 8$). This population is characterized by a larger number of spherical isomers when n_H/n_C increases. Simultaneously, the fraction of pretzels structures, more spherical ($\beta \leq 0.3$) and possessing fewer 5 or 6-carbon ring cycles ($0 < R_{5/6} \leq 8$), increases. For all hydrogenation rates, the fraction of cages ($\beta \leq 1.0$ and $R_{5/6} \geq 9$) remains extremely minor. The branched structures, characterized by the smallest number of 5 or 6-carbon rings ($R_{5/6} \leq 2$), is the highest energy population for all n_H/n_C ratios.

For all hydrogenation rates, we went further into the analysis determining the evolution of

carbon hybridization ratio (sp^n, n=1-3) as well as rings' sizes distribution as a function of energy. We showed that sp^2 carbon atoms overall dominate, although their predominance decreases when increasing energy due to the increase of sp^1 carbon atoms' fraction. Besides, when n_H/n_C increases, the fraction of sp^3 carbon atoms increases and remains quite constant over the entire range of considered energies. For all hydrogenation ratios, the fraction of 6-carbon rings and, to a lesser extent, of 5-carbon rings, is dominant at low energy but decreases when energy increases.

The VIP's distribution of the different populations are shown to present a maximum going from 7.9 eV ($C_{24}H_6$, branched family) down to 6.4 eV ($C_{24}H_{24}$, pretzels). The VIPs of branched structures are systematically higher than that of flakes. When n_H/n_C increases, the VIPs were found to decrease for all families, and we correlated this trend to the evolution of the geometric and electronic factors studied in the present paper.

These results are of astrophysical interest as they should be taken into account in astrophysical models especially regarding the role of carbon species in the gas ionization. In particular, carbonaceous grains such as PAHs are expected to play a role in the heating of the interstellar medium converting stellar radiation to thermal energy due to photoelectric effects[65–69]. In astrophysical models, empirical laws are often used to simulate a decrease of the ionization energies by about 2-3 eV from the smallest PAHs to the largest species.[68,69]. In the present work, we have shown that, for a given size, the distribution of the IPs extends over 2 eV and the shape of the latter depends on the type of PAH family (cage, pretzel, flake or branched). In addition, increasing the hydrogenation degree decreases the IPs by a few eVs, making the hydrogenated species more efficient compounds for interstellar gas heating from photoelectric processes. This work will be followed by a report of the IR and UV-visible spectra of the populations analyzed in the present work.

Acknowledgement

The authors gratefully acknowledge financial support by the Agence Nationale de la Recherche (ANR) Grant No. ANR-16-CE29-0025, Spanish MICINN (PID2019-110091GB-I00) and the Severo Ochoa Program for Centres of Excellence in R&D (SEV-2016-0686). P.P. thanks the Spanish MECD for a FPU grant. E.P. acknowledges the European Union (EU) and Horizon 2020 funding awarded under the Marie Skłodowska-Curie action to the EUROPAH consortium, grant number 722346. The authors also thank the computing mesocenter CALMIP (UMS CNRS 3667) for generous allocation of computer resources (p0059 and p17002).

Supporting Information Available

Definitions of the Hill Wheeler parameters are specified.

Figure S1: Most stable isomers of the flakes family for $C_{24}H_n$ (n=6,12,18,24).

Figure S2: Distribution of isomers as a function of the sp^2 ratio for all hydrogenation rates and all families

Figure S3: Distribution of isomers as a function of the gyration radius r_g for all hydrogenation rates and all families

Figure S4: VIPs' distributions for the four structural families of $C_{24}H_6$ and $C_{24}H_{18}$

Figure S5: LAIPs' distributions for the four families of structure for C_{24}, $C_{24}H_6$ and $C_{24}H_{12}$

Figure S6: LAIPs' distributions for the four families of structure for $C_{24}H_{18}$ and $C_{24}H_{24}$

Table S1: Comparison between SCC-DFTB and DFT (B3LYP/6-31G(d,p)) VIPs of most stable isomers of all families for $C_{24}H_n$ (n=0,6,12,18,24).

The data presented in this work are available in the Zenodo general data repository under the DOI 10.5281/zenodo.4787816.

This material is available free of charge via the Internet at `http://pubs.acs.org/`.

References

(1) Jones, A. P. Dust evolution, a global view: II. Top-down branching, nanoparticle fragmentation and the mystery of the diffuse interstellar band carriers. *R. Soc. Open Sci.* **2016**, *3*, 1–29.

(2) Santoro, G.; Martinez, L.; Lauwaet, K.; Accolla, M.; Tajuelo-Castilla, G.; Merino, P.; Sobrado, J. M.; Pelaez, R. J.; Herrero, V. J.; Tanarro, I. et al. The Chemistry of Cosmic Dust Analogs from C, C-2, and C(2)H(2)in C-rich Circumstellar Envelopes. *Astrophys. J.* **2020**, *895*.

(3) Dartois, E. Interstellar Carbon Dust. *J. Carb. Res.* **2019**, *5*.

(4) Léger, A.; Puget, J. L. Identification of the 'unidentified' IR emission features of interstellar dust? *Astron. Astrophys.* **1984**, *137*, L5–L8.

(5) Allamandola, L. J.; Tielens, A. G. G. M.; Barker, J. R. Polycyclic Aromatic-Hydrocarbons and the Unidentified Infrared-Emission Bands - Auto Exhaust Along the Milky-Way. *Astrophys. J.* **1985**, *290*, L25–L28.

(6) Bauschlicher, J., Charles W.; Peeters, E.; Allamandola, L. J. The Infrared Spectra of Very Large Irregular Polycyclic Aromatic Hydrocarbons (PAHs): Observational Probes of Astronomical PAH Geometry, Size, and Charge. *Astrophys. J.* **2009**, *697*, 311–327.

(7) Simon, A.; Joblin, C. The Computed Infrared Spectra of a Variety of $[FePAH]^+$ Complexes: Mid- and Far-Infrared Features. *Astrophys. J.* **2010**, *712*, 69–77.

(8) García-Hernández, D. A.; Manchado, A.; García-Lario, P.; Stanghellini, L.; Villaver, E.; Shaw, R. A.; Szczerba, R.; Perea-Calderon, J. V. Formation of fullerenes in H-containing planetary nebulae. *Astrophys. J. Lett.* **2010**, *724*, L39.

(9) García-Hernández, D. A.; Kameswara Rao, N.; Lambert, D. L. Are C-60 molecules

detectable in circumstellar shells of R coronae borealis stars ? *Astrophys. J.* **2011**, *729*, 126.

(10) García-Hernández, D. A.; Iglesias-Groth, S.; Acosta-Pulido, J. A.; Manchado, A.; García-Lario, P.; Stanghellini, L.; Villaver, E.; Shaw, R. A.; Cataldo, F. The formation of fullerenes: clues from new C-60, C-70, and (possible) planar C-24 detections in magellanic cloud planetary nebulae. *Astrophys. J. Lett.* **2011**, *737*, L30.

(11) García-Hernández, D. A.; Villaver, E.; García-Lario, P.; Acosta-Pulido, J. A.; Manchado, A.; Stanghellini, L.; Shaw, R. A.; Cataldo, F. Infrared study of fullerene planetary nebulae. *Astrophys. J.* **2012**, *760*, 107.

(12) Bernard-Salas, J.; Cami, J.; Peeters, E.; Jones, A. P.; Micelotta, E. R.; Groenewegen, M. A. T. On the Excitation and Formation of Circumstellar Fullerenes. *Astrophys. J.* **2012**, *757*, 41.

(13) Tielens, A. Interstellar Polycyclic Aromatic Hydrocarbon Molecules. *Annu. Rev. Astron. Astrophys.* **2008**, *46*, 289–337.

(14) Buss, J., R. H.; Tielens, A. G. G. M.; Cohen, M.; Werner, M. W.; Bregman, J. D.; Witteborn, F. C. Infrared Spectra of Transition Objects and the Composition and Evolution of Carbon Dust. *Astrophys. J.* **1993**, *415*, 250.

(15) Rapacioli, M.; Joblin, C.; Boissel, P. Spectroscopy of polycyclic aromatic hydrocarbons and very small grains in photodissociation regions. *Astron. Astrophys.* **2005**, *429*, 193–204.

(16) Kwok, S.; Volk, K.; Bernath, P. On the Origin of Infrared Plateau Features in Proto-Planetary Nebulae. *Astrophys. J. Lett.* **2001**, *554*, L87–L90.

(17) Hrivnak, B.; Volk, K.; Kwok, S. 2-45 micron infrared spectroscopy of carbon-rich protoplanetary nebulae. *Astrophys. J.* **2000**, *535*, 275–292.

(18) Dubosq, C.; Falvo, C.; Calvo, F.; Rapacioli, M.; Parneix, P.; Pino, T.; Simon, A. Mapping the structural diversity of C $_{60}$ carbon clusters and their infrared spectra. *Astron. Astrophys.* **2019**, *625*, L11.

(19) Stecher, T. P. Interstellar Extinction in the Ultraviolet. *Astrophys. J.* **1965**, *142*, 1683.

(20) Fitzpatrick, E. L.; Massa, D. An analysis of the shapes of interstellar extinction curves 1. The 2175-A BUMP. *Astrophys. J.* **1986**, *307*, 286–294.

(21) Fitzpatrick, E. L.; Massa, D. An analysis of the shapes of interstellar extinction curves. V. The IR-through-UV curve morphology. *Astrophys. J.* **2007**, *663*, 320–341.

(22) Fitzpatrick, E. L.; Massa, D.; Gordon, K. D.; Bohlin, R.; Clayton, G. C. An Analysis of the Shapes of Interstellar Extinction Curves. VII. Milky Way Spectrophotometric Optical-through-ultraviolet Extinction and Its R-dependence. *Astrophys. J.* **2019**, *886*, 108.

(23) Papoular, R. J.; Yuan, S.; Roldán, R.; Katsnelson, M. I.; Papoular, R. Effects of structural and chemical disorders on the vis/UV spectra of carbonaceous interstellar grains. *Month. Not. Roy. Astron. Soc.* **2013**, *432*, 2962–2974.

(24) Jones, A.; Duley, W.; Williams, D. The structure and evolution of hydrogenated amorphous carbon grains and mantles in the interstellar medium. *Quart. J. Roy. Astron. Soc.* **1990**, *31*, 567–582.

(25) Gadallah, K.; Mutschke, H.; Jäger, C. UV irradiated hydrogenated amorphous carbon (HAC) materials as a carrier candidate of the interstellar UV bump at 217.5 nm. *Astron. Astrophys.* **2011**, *528*, A56.

(26) Gavilan, L.; Le, K. C.; Pino, T.; Alata, I.; Giuliani, A.; Dartois, E. Polyaromatic disordered carbon grains as carriers of the UV bump: Far-UV to mid-IR spectroscopy of laboratory analogs. *Astron. Astrophys.* **2017**, *607*, A73.

(27) Rotundi, A.; Rietmeijer, F.; Colangeli, L.; Mennella, V.; Palumbo, P.; Bussoletti, E. Identification of carbon forms in soot materials of astrophysical interest. *Astron. Astrophys.* **1998**, *329*, 1087–1096.

(28) Dubosq, C.; Calvo, F.; Rapacioli, M.; Dartois, E.; Pino, T.; Falvo, C.; Simon, A. Quantum modeling of the optical spectra of carbon cluster structural families and relation to the interstellar extinction UV bump. *Astron. Astrophys.* **2020**, *634*, A62.

(29) Tielens, A. G. G. M. Diffuse Interstellar Bands: The Way Forward. *Proc. Int. Astron. Union* **2013**, *9*, 399–411.

(30) Herbig, G. H. The Diffuse Interstellar Bands. *Ann. Rev. Astron. Astrophys.* **1995**, *33*, 19–73.

(31) Omont, A.; Bettinger, H. F.; Tanshoff, C. Polyacenes and diffuse interstellar bands. *Astron. Astrophys.* **2019**, *625*, A41.

(32) Maier, J. P.; Walker, G. A. H.; Bohlender, D. A. On the Possible Role of Carbon Chains as Carriers of Diffuse Interstellar Bands. *Astrophys. J* **2004**, *602*, 286–290.

(33) Salama, F.; Galazutdinov, G. A.; Krełowski, J.; Biennier, L.; Beletsky, Y.; Song, I.-O. Polycyclic Aromatic Hydrocarbons and the Diffuse Interstellar Bands: A Survey. *Astrophys. J.* **2011**, *728*, 154.

(34) Omont, A. Interstellar fullerene compounds and diffuse interstellar bands. *Astron. Astrophys.* **2016**, *590*, A52.

(35) Pilleri, P.; Montillaud, J.; O., B.; Joblin, C. Evaporating very small grains as tracers of the UV radiation field in photo-dissociation regions. *Astron. Astrophys.* **2012**, *542*, A69.

(36) Berné, O.; Tielens, A. G. G. M. Formation of buckminsterfullerene (C60) in interstellar space. *Proc. Nat. Acad. Sci.* **2012**, *109*, 401–406.

(37) Berné, O.; Montillaud, J.; Joblin, C. Top-down formation of fullerenes in the interstellar medium. *Astron. Astrophys.* **2015**, *577*, A133.

(38) Zhen, J.; Castellanos, P.; Paardekooper, D. M.; Linnartz, H.; Tielens, A. G. G. M. Laboratory formation of fullerenes from PAHs: top-down interstellar chemistry. *Astrophys. J.* **2014**, *797*, L30.

(39) Dunk, P. W.; Kaiser, N. K.; Hendrickson, C. L.; Quinn, J. P.; Ewels, C. P.; Nakanishi, Y.; Sasaki, Y.; Shinohara, H.; Marshall, A. G.; Kroto, H. W. Closed network growth of fullerenes. *Nature Comm.* **2012**, *3*, 855.

(40) Porezag, D.; Frauenheim, T.; Köhler, T.; Seifert, G.; Kaschner, R. *Phys. Rev. B* **1995**, *51*, 12947–12957.

(41) Seifert, G.; Porezag, D.; Frauenheim, T. Calculations of Molecules, Clusters, and Solids with a Simplified LCAO-DFT-LDA Scheme. *Int. J. Quantum Chem.* **1996**, *58*, 185–192.

(42) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys. Rev. B* **1998**, *58*, 7260.

(43) Rapacioli, M.; Spiegelman, F.; Talbi, D.; Mineva, T.; Goursot, A.; Heine, T.; Seifert, G. Correction for dispersion and Coulombic interactions in molecular clusters with density functional derived methods: Application to polycyclic aromatic hydrocarbon clusters. *J. Chem. Phys.* **2009**, *130*, 244304–10.

(44) Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. Hydrogen bonding and stacking interactions of nucleic acid base pairs: A density-functional-theory based treatment. *J. Chem. Phys.* **2001**, *114*, 5149–5155.

(45) Zhechkov, L.; Heine, T.; Patchkovskii, S.; Seifert, G.; Duarte, H. A. An Efficient *a Poste-*

riori Treatment for Dispersion Interaction in Density-Functional-Based Tight Binding. *J. Chem. Theor. Comput.* **2005**, *1*, 841–847.

(46) Heine, T.; Rapacioli, M.; Patchkovskii, S.; Frenzel, J.; Koster, A.; Calaminici, P.; Duarte, H. A.; Escalante, S.; Flores-Moreno, R.; Goursot, A. et al. *deMonNano,* `http://demon-nano.ups-tlse.fr/` **2009**,

(47) Goldberg, D. E. *Genetic Algorithms in Search, Optimization and Machine Learning*, 1st ed.; Addison-Wesley Longman Publishing Co., Inc.: USA, 1989.

(48) Forrest, S. Genetic algorithms: principles of natural selection applied to computation. *Science* **1993**, *261*, 872 LP – 878.

(49) Johnston, R. L. Evolving better nanoparticles: Genetic algorithms for optimising cluster geometries. *Dalton Trans.* **2003**, 4193–4207.

(50) Hjorth Larsen, A.; Jørgen Mortensen, J.; Blomqvist, J.; Castelli, I. E.; Christensen, R.; Dułak, M.; Friis, J.; Groves, M. N.; Hammer, B.; Hargus, C. et al. The atomic simulation environment—a Python library for working with atoms. *J. Phys. Cond. Mat.* **2017**, *29*, 273002.

(51) Vilhelmsen, L. B.; Walton, K. S.; Sholl, D. S. Structure and Mobility of Metal Clusters in MOFs: Au, Pd, and AuPd Clusters in MOF-74. *J. Am. Chem. Soc.* **2012**, *134*, 12807–12816.

(52) Vilhelmsen, L. B.; Hammer, B. A genetic algorithm for first principles global structure optimization of supported nano structures. *J. Chem. Phys.* **2014**, *141*, 44711.

(53) Jensen, P. B.; Lysgaard, S.; Quaade, U. J.; Vegge, T. Designing mixed metal halide ammines for ammonia storage using density functional theory and genetic algorithms. *Phys. Chem. Chem. Phys.* **2014**, *16*, 19732–19740.

(54) Verstraelen, T. *MolMod Software Library,* `http://molmod.ugent.be/software/`

(55) Sugita, Y.; Okamoto, Y. Replica-Exchange Molecular Dynamics Method for Protein Folding. *Chem. Phys. Lett.* **1999**, *314*, 141–151.

(56) Bonnin, M. A.; Falvo, C.; Calvo, F.; Pino, T.; Parneix, P. Simulating the structural diversity of carbon clusters across the planar-to-fullerene transition. *Phys. Rev. A* **2019**, *99*.

(57) Hill, D. L.; Wheeler, J. A. Nuclear Constitution and the Interpretation of Fission Phenomena. *Physical Review* **1953**, *89*, 1102–1145.

(58) Kent, P. R. C.; Towler, M. D.; Needs, R. J.; Rajagopal, G. Carbon clusters near the crossover to fullerene stability. *Phys. Rev. B* **2000**, *62*, 15394–15397.

(59) Kosimov, D. P.; Dzhurakhalov, A. A.; Peeters, F. M. Carbon clusters: From ring structures to nanographene. *Phys. Rev. B* **2010**, *81*, 195414.

(60) Manna, D.; Martin, J. M. L. What Are the Ground State Structures of C20 and C24? An Explicitly Correlated Ab Initio Approach. *J. Phys. Chem. A* **2016**, *120*, 153–160.

(61) Malloci, G.; Mulas, G.; Cecchi-Pestellini, C.; Joblin, C. Dehydrogenated polycyclic aromatic hydrocarbons and UV bump. *Astron. Astrophys.* **2008**, *489*, 1183–1187.

(62) Pla, P.; Wang, Y.; Martín, F.; Alcamí, M. Hydrogenated polycyclic aromatic hydrocarbons: isomerism and aromaticity. *Phys. Chem. Chem. Phys.* **2020**, *22*, 21968–21976.

(63) Lias, S. G. In *"WebBook de Chimie NIST, Base de Données Standard de Référence NIST Numéro 69"(https://doi.org/10.18434/T4D303)*; P.J.Linstrom„ Mallard, W., Eds.; National Institute of Standard and Technology, Gaithersburg MD, Chapter "Ionization Energy Evaluation".

(64) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H. et al. Gaussian16 Revision B.01. 2016; Gaussian Inc. Wallingford CT.

(65) Watson, W. D.; Salpeter, E. E. Molecule Formation on Interstellar Grains. *Astrophys. J.* **1972**, *174*, 321.

(66) Helou, G.; Malhotra, S.; Hollenbach, D. J.; Dale, D. A.; Contursi, A. Evidence for the Heating of Atomic Interstellar Gas by Polycyclic Aromatic Hydrocarbons. *Astrophys. J.* **2001**, *548*, L73–L76.

(67) Hollenbach, D. J.; Tielens, A. G. G. M. Photodissociation regions in the interstellar medium of galaxies. *Rev. Mod. Phys.* **1999**, *71*, 173–230.

(68) Weingartner, J.; Draine, B. Photoelectric Emission from Interstellar Dust: Grain Charging and Gas Heating. *Astrophys. J. Supp. Ser.* **2001**, *134*, 263–281.

(69) Bakes, E. L. O.; Tielens, A. G. G. M. The Photoelectric Heating Mechanism for Very Small Graphitic Grains and Polycyclic Aromatic Hydrocarbons. *Astrophys. J.* **1994**, *427*, 822.

Graphical TOC Entry