

Article

# Methodological Investigation for Hydrogen Addition to Small Cage Carbon Fullerenes

Yuri Tanuma <sup>1,2</sup> , Toru Maekawa <sup>2,3</sup> and Chris Ewels <sup>1,\*</sup> 

<sup>1</sup> Institut des Matériaux Jean Rouxel, CNRS/Université de Nantes, BP32229, 44322 Nantes, France; yuri.tanuma@cnrs-imn.fr

<sup>2</sup> Graduate School of Interdisciplinary New Science, Toyo University, 2100 Kujirai, Kawagoe, Saitama 350-8585, Japan; maekawa@toyo.jp

<sup>3</sup> Bio-Nano Science Research Centre, Toyo University, 2100 Kujirai, Kawagoe, Saitama 350-8585, Japan

\* Correspondence: chris.ewels@cnrs-imn.fr; Tel.: +33-(0)2-40-37-64-07

**Abstract:** Hydrogenated small fullerenes ( $C_n$ ,  $n < 60$ ) are of interest as potential astrochemical species, and as intermediates in hydrogen-catalysed fullerene growth. However, the computational identification of key stable species is difficult due to the vast configurationally space of structures. In this study, we explored routes to predict stable hydrogenated small fullerenes. We showed that neither local fullerene geometry nor local electronic structure analysis was able to correctly predict subsequent low-energy hydrogenation sites, and sequential stable addition searches also sometimes failed to identify most stable hydrogenated fullerene isomers. Of the empirical and semi-empirical methods tested, GFN2-xTB consistently gave highly accurate energy correlations ( $r > 0.99$ ) to full DFT-LDA calculations at a fraction of the computational cost. This allowed identification of the most stable hydrogenated fullerenes up to 4H for four fullerenes, namely two isomers of  $C_{28}$  and  $C_{40}$ , via “brute force” systematic testing of all symmetry-inequivalent combinations. The approach shows promise for wider systematic studies of smaller hydrogenated fullerenes.

**Keywords:** fullerene; hydrogenation; DFT; functionalisation; growth; xTB; REBO; interstellar



**Citation:** Tanuma, Y.; Maekawa, T.; Ewels, C. Methodological Investigation for Hydrogen Addition to Small Cage Carbon Fullerenes. *Crystals* **2021**, *11*, 1334. <https://doi.org/10.3390/cryst11111334>

Academic Editor: Shujun Zhang

Received: 20 October 2021

Accepted: 28 October 2021

Published: 1 November 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

As well as in the laboratory and on earth, fullerenes have been found in a range of nonterrestrial environments [1–10]. Recent confirmation that  $C_{60}^+$  is the source of some of the strong diffuse interstellar bands (DIBs) [5,11] suggests that fullerenes may play a key astrochemical role. However, the formation mechanisms of fullerenes in such extreme environments are still a debated question. Interestingly, while one might intuitively expect hydrogen to stabilise dangling bonds at edge sites and hence promote the formation of polyaromatic hydrocarbons rather than fullerenes, astrochemists have recently found that  $C_{60}$  and  $C_{70}$  are abundant in hydrogen-containing stars [8,9,12,13]. This suggests that hydrogen may play a role in fullerene growth. As such, stable hydrogenated small fullerenes are potentially of interest in astrochemistry, both as species in their own right and/or as degradation products from polyaromatic hydrocarbons [14], and as intermediates in the formation processes of larger fullerenes.

The stability of fullerenes strongly depends on the arrangement of pentagons and hexagons. According to the isolated pentagon rule (IPR), structures composed of isolated pentagons surrounded by five hexagons are favoured by the strain delocalisation and bonding [15]. The smallest fullerene to fully obey the IPR is the highly stable  $C_{60}$ , Buckminster fullerene. Fullerenes smaller than  $C_{60}$  are increasingly unstable due to steric strain originating within the non-IPR structures, and bond frustration in areas of high nonplanarity. One route to potentially stabilise local strain and under-coordination is chemical functionalisation of the most reactive carbon sites. For example, while pure  $C_{50}$  appears unstable,  $C_{50}Cl_{10}$  has been isolated and characterised [16], and the stability and lifetime

of  $C_{20}$  is greatly enhanced when fully hydrogenated to  $C_{20}H_{20}$  [17]. This suggests that the relative stability of smaller fullerenes may be modified in the presence of hydrogen. While there has been extensive experimental [13,18–21] and theoretical [13,18,22] studies of hydrogenated  $C_{60}$ , less attention has been given to smaller hydrogenated fullerenes. It was theoretically proposed that  $T_d-C_{28}H_4$ , a tetrahedral fullerene with triply fused pentagons on each corner of the tetrahedron, might be stable [15] and indeed has since been experimentally identified [23]. For smaller fullerenes, an *in silico* investigation suggested that fused-pentagon sites are preferred in general for single hydrogen addition [24].

Exploring the landscape of smaller hydrogenated fullerenes is potentially a highly complex and computationally intensive task. Very little can necessarily be determined in advance: neither the hydrogenation sites on the fullerene surface, nor the number of hydrogen additions—nor even the most stable carbon isomer for hydrogenation. This results in a vast and generally intractable number of potential calculations.

Considering first the number of pentagon and hexagon carbon fullerenes smaller than  $C_{60}$ , there is only 1 isomer for each of  $C_{20}$ ,  $C_{24}$ , and  $C_{26}$ , and 2 for  $C_{28}$ . However, this rapidly increases, reaching 40 isomers for  $C_{40}$ , and 437 for  $C_{52}$ . In total, there are 3958  $C_x$  pentagon-hexagon fullerenes with  $x < 60$  [25].

Considering now hydrogenation, if symmetry is not taken into account, then for a single fullerene  $C_n$ , there are  $n$  possible addition sites for a first hydrogen atom. However, this increases rapidly as more sites are hydrogenated. For two hydrogen atoms, there are  $n(n-1)/2$  arrangements (divided by two as the hydrogens are interchangeable), and in general, for  $m$  hydrogen atoms, there are  $n!/(n-m)!m!$  arrangements. Thus, for  $C_{28}H_5$ , this gives 98,280 possible structures, and for larger hydrogenated species such as  $C_{50}H_4$ , there are 230,300 possibilities. While symmetry can help reduce these numbers, they remain highly challenging for standard density functional theory (DFT) calculations, where it can take several minutes on a state-of-the-art desktop PC to geometrically optimise even the smallest fullerene  $C_{20}$ . As we do not, in principle, know how many hydrogen atoms are likely to bind to a given isomer, we would be required to start with single hydrogenation and then increase the number of hydrogen atoms sequentially until the hydrogen binding energy becomes too low.

As the time cost for a brute force approach using DFT to this problem is not possible, it was the aim of the current study to explore, test, and validate different possible methods that might be used to reduce the complexity of this problem and render it accessible computationally. This includes assumptions on growth sequence, the use of geometric and electronic parameters as guides towards selecting the most reactive sites for functionalisation, and the use of different computational approximations such as empirical potentials and tight binding methods to increase calculation speed. We started with a sequence of DFT calculations and used these thereafter as our benchmark energies for the subsequent studies.

## 2. Method

Coordinates for the fullerene cage structures were taken from Yoshida's fullerene library [26]. Geometric analysis of the fullerenes was performed using the *pychemcurv* python libraries discussed further in Ref [27].

All DFT calculations were carried out under the local density approximation (LDA) using the AIMPRO code [28–30]. Pseudo-potentials were given by Hartwigsen, Goedecker, and Hutter (HGH) [31]. Wavefunctions were handled using a basis set of Gaussian-type orbitals [32,33], in which 38 (12) functions were used for carbon (hydrogen) atoms. Charge density was handled using plane waves. Spin averaging was used for systems with an even number of electrons. Spin polarisation was used with a spin excess of 1.0 for systems with an odd number of electrons. An additional spin excess of either 0.1 or 1.0 was added when charge was added while exploring frontier orbitals. Charged systems were handled by introducing a uniform homogenous compensating background charge to the cell. The smearing of electronic occupations using a Fermi function with an effective temperature of

0.04 eV was set during the geometry optimisation procedure. All molecules were modelled in face-centred cells with  $a_0 = 15.875 \text{ \AA}$  to ensure negligible inter-fullerene interactions. Structures were geometrically optimised until the system energy was converged within  $10^{-5} \text{ Ha}$ , and positions to within  $10^{-4} a_0$ .

The LDA was selected based on previous experience with graphite, where it regularly outperformed GGA [34]. Quantum Monte Carlo calculations (including London dispersion) showed that errors are closely correlated to the Laplacian (neglected in density functionals), rather than the gradient of the electron density [35]. Model systems showed some cancellation of errors for the LDA; however, GGA construction can magnify rather than reduce this error. LDA provides a better description of interlayer binding in graphite than GGAs [34], supported by random phase approximation (RPA) calculations [36], and shows better G-band frequency shifts with doping than the GGA [37]. In practice, in the current systems, this choice between LDA and GGA-PBE does not matter, as, when we geometrically optimised over 700 hydrogenated fullerene structures with both functionals and compared energies, the result was a fully linear fit (see Supplementary Materials Figure S4). As the energies are thus directly proportional, the comparative results in the current study will also be directly transferable to GGA studies.

Empirical potential calculations were performed using the LAMMPS code [38], using the REBO [39], AIREBO [40], and AIREBO-M [41] potentials. In each case, the structures were optimised with a  $10^{-10}$  convergence threshold (ratio of energy change to total energy magnitude), at zero Kelvin. The reactive empirical bond order (REBO) is a potential developed by Brenner [39]. REBO is handled by a Tersoff-type potential [42,43], which deals with the formation and deformation of covalent bonds during a calculation and was developed for hydrogen [44] and carbon [44,45]-containing systems. The adaptive intermolecular reactive empirical bond order (AIREBO) potential [40] was developed for carbon-hydrogen systems. It uses REBO to describe nonbonded interactions. In this study, we included its torsion term and the Lennard–Jones (LJ) cut-off radius set at 2.5 (in  $\sigma$  scale factor). AIREBO-M is a hybrid potential of AIREBO and Morse [41]. In this potential, a LJ term in AIREBO is replaced with a Morse potential parameterised to Møller–Plesset method (MP2) calculations [46,47]. In this study, we included its torsion term and the Morse cutoff radius set at 3.0 (in  $\sigma$  scale factor).

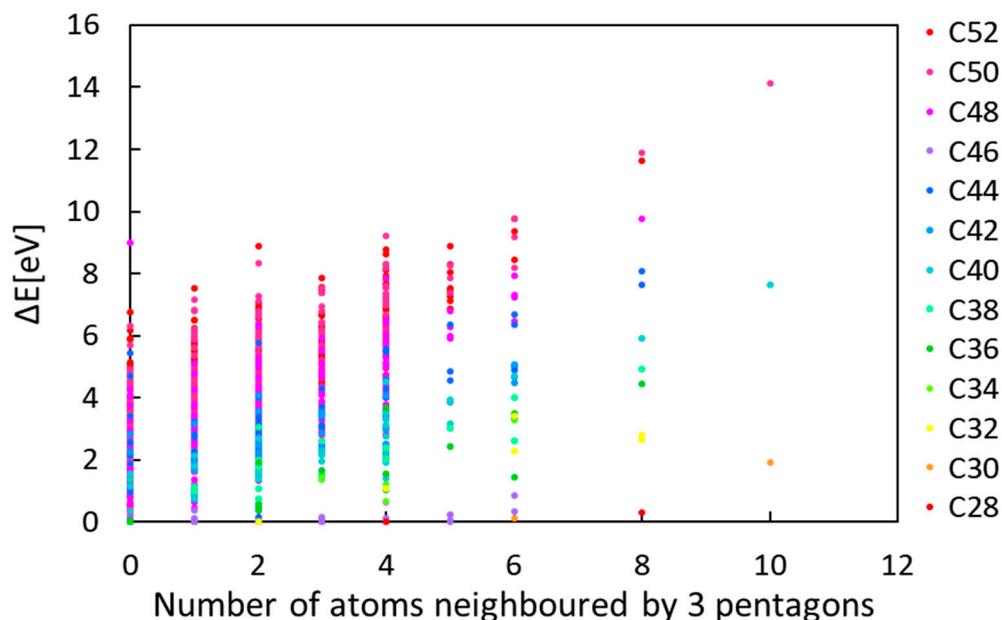
Semi-empirical molecular orbital calculations were performed using MOPAC2016 [48] as implemented in the OpenMOPAC code. Default options were taken, except the MOZYME option that was selected, a method developed for quick calculations of organic molecules that replaces the self-consistent field (SCF) method with a localised molecular orbital (LMO) method. Different semi-empirical parameterisations were tested, namely PM6 + D3, PM7, and RM1. PM6 [49] was fitted to reference data from DFT calculations and was combined here with Grimme’s D3 method for longer-range dispersion interactions [50]. For PM7, geometric and enthalpy errors were improved compared to PM6 [49] by reconsidering new reference data [51]. In this method, a D2 correlation was combined to take into account the dispersion. The correlation term was based on the exchange–correlation functional of the generalised gradient approximation (GGA). Finally, RM1 [52] was developed from AM1 [53] by improved parameters. It utilises different equations from PM6 to consider the core–core interactions.

The semi-empirical extended tight binding method (xTB) as implemented in GFN2-xTB is a quantum calculation method for gas and condensed phases developed by Grimme and co-workers [54]. It covers a large number of elements up to radon and is based on DFT perturbation expansion of the electron density, which makes it applicable to electronically complicated systems. For a large number of isomers of fullerene  $C_{60}$ , xTB has already shown excellent structure–energy agreement with DFT results with Pearson’s correlation coefficient  $r = 0.998$  [55]. Again, we used default settings including for convergence cut-offs.

### 3. Results

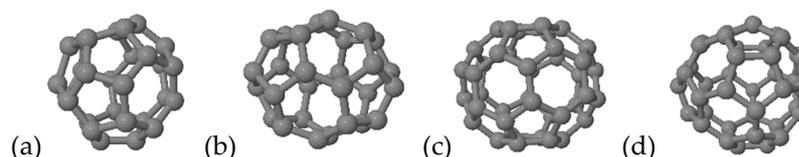
#### 3.1. Pure Carbon Fullerenes

Using DFT-LDA, we first optimised all pure-carbon isomers of  $C_x$  fullerenes from  $x = 28$  to  $x = 52$ , a total of 1246 structures. Figure 1 shows on the y-axis the energy difference between a given isomer and the most stable isomer for the same number of carbon atoms. On the x-axis, we give the number of atoms in the isomer neighbouring three pentagons. This plot confirms the general tendency that the stability of a given fullerene is negatively proportional to the number of triple-fused pentagons, although there are a number of interesting exceptions.



**Figure 1.** Plot of the number of atoms neighbored by triple pentagons versus energy difference compared to the ground state isomer. The structure instability is, in general, proportional to the number of triple pentagons. For the optimal lowest energy isomer of  $C_x$ ,  $\Delta E = 0$ .

For the following hydrogenation study, we focused on only four representative structures,  $C_{28}$  ( $1-T_d$  and  $2-D_2$ , index of Yoshida's library [26]-symmetry symbol) and  $C_{40}$  ( $38-D_2$  and  $40-T_d$ ) (see Figure 2). These were chosen as the  $T_d$  tetrahedral structures have four, the maximum number, triple-fused pentagons, and the other structures are the lowest-energy pure-carbon structures in each isomer group.



**Figure 2.** Structures for all site test. (a)  $C_{28}$  ( $1-T_d$ ), (b)  $C_{28}$  ( $2-D_2$ ), (c)  $C_{40}$  ( $38-D_2$ ), and (d)  $C_{40}$  ( $40-T_d$ ).

#### 3.2. First Hypothesis—Sequential Hydrogen Addition

In order to tackle the problem of the calculation time cost, our first hypothesis is that hydrogen atoms bind sequentially, in each case, to the most stable site, and thereafter do not move. Such an approach was shown to successfully predict stable fluorinated fullerene [56], fullerene chlorination [57], and functionalisation with  $CF_3$  groups [58,59]. This reduces the number of calculations substantially. To determine the stable structure of  $C_nH_m$ , we first geometrically optimised the system  $C_nH$  with hydrogen in each of the  $n$  possible sites. We took the most stable solution and fixed the first hydrogen at this site, tested

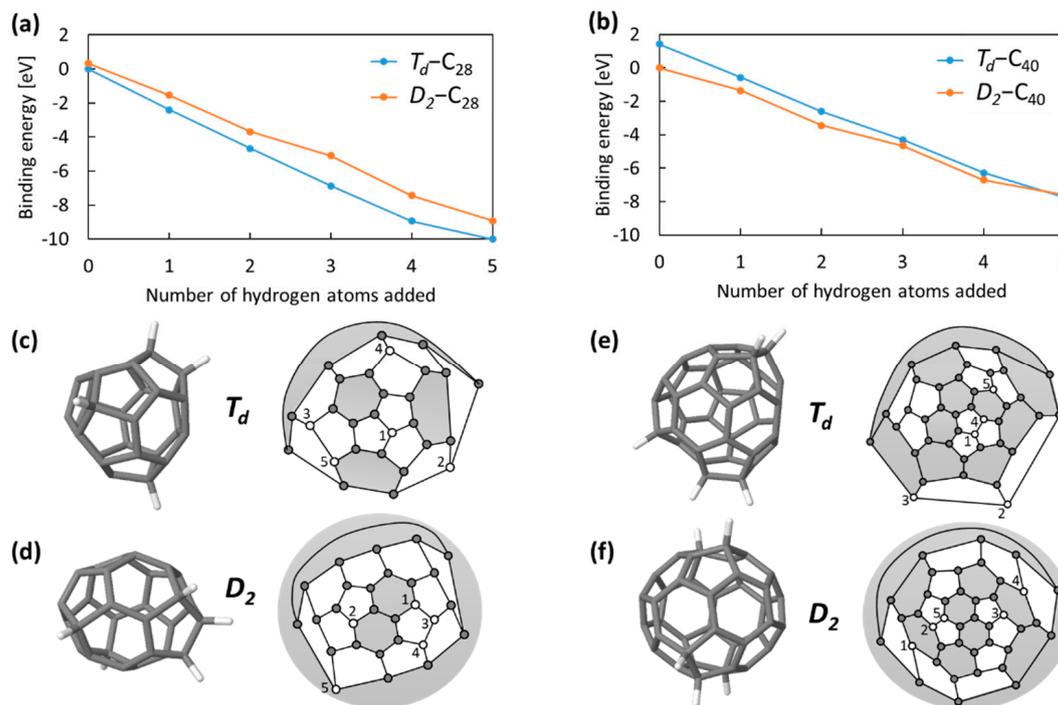
the  $n - 1$  possible sites for the next hydrogen, and so on. As an example, for  $C_{28}H_5$ , this reduces the number of calculations from 98,280 to 130, and for  $C_{50}H_4$  from 230,300 to 194, with the additional benefit that the stable structures for  $C_nH$ ,  $C_nH_2$ , etc. are automatically determined as part of the process.

While the method works well for halogens where the binding energies are high, its validity for hydrogenation is less clear. This hypothesis for sequential hydrogenation is tested in Sections 3.3 and 3.4 below. In the current section, we employ sequential addition for our benchmark DFT calculations, exploring hydrogenation up to  $m = 5$ , i.e.,  $C_{28}H_5$  and  $C_{40}H_5$ .

Figure 3 shows the structures obtained this way for  $C_{28}H_5$  and  $C_{40}H_5$  for the two  $T_d$  and  $D_2$  isomers. The energy axis shows the total binding energy of the hydrogen, calculated as

$$E_{\text{binding}} = E(C_nH_m) - E(C_n') - m/2 E(H_2)$$

where  $C_n$  is the fullerene in question and  $C_n'$  is the energetically most stable isomer fullerene. The projected Schlegel diagrams are marked numerically with the hydrogen addition sequence.  $T_d$ - $C_{28}$  is sequentially hydrogenated at each triple pentagon on the tetrahedral corner sites, with the fifth hydrogenation occurring with a lower binding energy visible as a change in the total energy gradient. At each stage, this is the stable  $C_{28}$  isomer. In contrast, the stable  $C_{40}$  isomer is the lower-symmetry  $D_2$ - $C_{40}$ . Subsequent hydrogenation occurs around the “caps” of this isomer. When the binding energies are normalised by the number of hydrogen atoms added, we see a general convergence towards hydrogen binding of  $\sim 1.6$  eV for the  $D_2$ - $C_{40}$  isomer, with higher binding ( $\sim 2.0$  eV) for the  $T_d$ - $C_{28}$ , consistent with its smaller size and lower aromaticity.



**Figure 3.** DFT-LDA hydrogen binding energy relative to the most stable isomer and isolated  $H_2$  (eV) for (a)  $C_{28}$  and (b)  $C_{40}$  obtained by sequential hydrogenation. The  $T_d$  ( $D_2$ ) isomer is given in blue (orange); symmetry notation refers to nonhydrogenated fullerene cage. (c–f) 3D molecular structures and Schlegel diagrams for  $C_nH_5$ . Grey and white represent carbon and hydrogenated carbon atoms, respectively. In the Schlegel diagrams, white circles indicate hydrogenated sites, and hexagons are shaded in grey. Numbering in Schlegel diagrams indicates the hydrogenation order.

Hydrogen binds more strongly to the  $T_d$ - $C_{40}$  isomer than the  $D_2$ - $C_{40}$  (although it does not, this time, adopt the triple-pentagon tetrahedrally symmetric sites). This lowers the

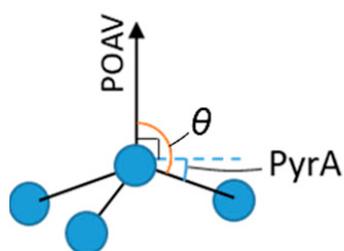
relative energy difference between the isomers until, at  $C_{40}H_5$ , there is a transition and the  $T_d-C_{40}H_5$  isomer becomes the most stable. This demonstrates that the most stable pure carbon isomer cannot be taken as a guide in general for which the hydrogenated isomer will be stable, i.e., wider isomer testing is required.

### 3.3. Second Hypothesis—Predicting Reactive Sites via Local Geometry

The first hypothesis of sequential addition still requires every possible addition site to be geometrically optimised for each fullerene isomer, and it is desirable to reduce this further. In the section hypothesis, we postulate that the most reactive site for hydrogenation on a given fullerene might be determined in advance from the *local geometry of the fullerene before the hydrogen is added*. If this hypothesis is correct, it allows the prediction in advance of which site will be hydrogenated next, in principle, reducing the number of calculations for  $C_{28}H_5$ , for example, from 130 to simply 5.

There are a number of geometric parameters that have been explored in the literature as a way of indicating local site reactivity in fullerenes, most of which have been discussed at a good level of detail in Reference [27]. In general, they all try to quantify in some way the degree to which the  $p_z$ -orbital of the carbon atom is able to form a strong  $\pi$ -bond with one or more of its neighbours, typically through some measure of local curvature. In the following, we selected one geometric criterion and calculated its value for all nonhydrogenated sites on a given fullerene cage. The most extreme value (minimum or maximum, as discussed below) was taken as the next site to hydrogenate. The new structure was then geometrically optimised using DFT, and the process repeated. In this way, a hydrogenation sequence can be built up with geometric site selection at each step, and this sequence can be compared to the “lowest-energy” sequence derived in Section 3.2.

We first tested the pyramidalization angle (PyrA). This is defined as  $\theta - \pi/2$ , where  $\theta$  is the angle between the  $\pi$ -orbital axis vector (POAV) and a vector from a central atom to a side atom (except the atom along POAV), as shown in Figure 4 [60–62]. The POAV is defined as a vector at constant angle to the three carbon-carbon bonds [63]. Thus, in a flat graphene sheet, the pyrA is zero. A further derivative from pyrA is the hybridisation value defined by Haddon [60–62] and explored further by Sabalot-Cuzzubbo and colleagues [11]. This is essentially a more chemical reworking of the pyrA in terms of the amount of  $s^x p^y$  character in the system. For both of these parameters, we calculated the values for all potential binding sites and selected the maximum for hydrogenation at each step.

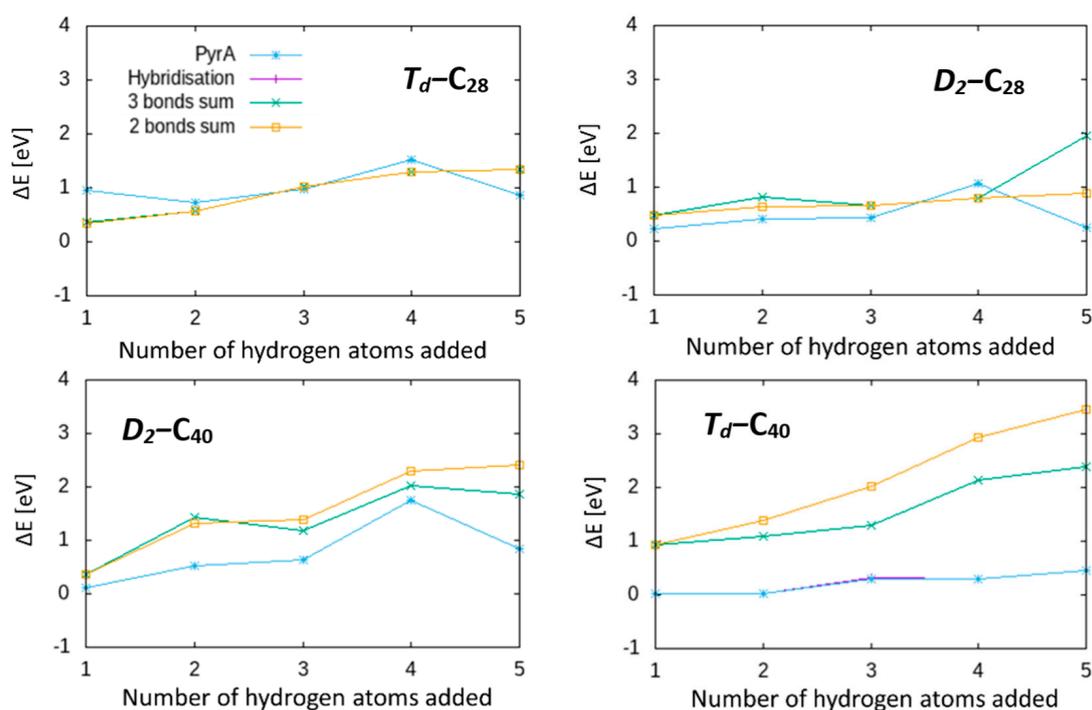


**Figure 4.** Schematic diagram of pyrA, defined as  $\theta - \pi/2$ . POAV is the vector along the atom  $p_z$  orbital and is defined as having a constant angle with all three carbon-carbon bonds. Blue spheres represent carbon atoms.

An alternative geometric approach is to use bond lengths rather than angles. The idea behind this simple geometric test is that decreasing the  $\pi$ -bond character of an atom will tend to result in longer, purely single-character bonds, and hence a highly reactive radical should have three long single bonds. In the three-bond sum geometric parameter, we simply sum the three bond lengths between a central atom and its adjacent atoms and hydrogenate the atom with the largest three-bond sum. A refinement of the three-bond sum model is a two-bond sum. In principle, in a strongly localised but nonaromatic  $\pi$ -bonded atom, it is possible to have one short double bond and two long single bonds in aromatic bonds, which may nonetheless result in a high three-bond sum. The two-bond sum model

tries to avoid this by eliminating the shortest bond and hydrogenating the atom with the lowest sum of the remaining two bond lengths.

In each of these four cases, we followed the hydrogen addition sequence and compared the energy of the DFT-optimised structure with that found from testing all available sites at each addition step from Section 3.2 (Figure 5, Schlegel projections showing hydrogenation sites shown in Figure S1). As expected, hybridisation and pyrA give the same hydrogenation sequence. While these two parameters find the lowest-energy structures for  $T_d$ -C<sub>40</sub>H and  $T_d$ -C<sub>40</sub>H<sub>2</sub>, in general, none of these geometric methods correctly select the lowest-energy hydrogenation sites, that is, none are capable of predicting the DFT-predicted most-stable structures.



**Figure 5.** Plots of the difference in calculated total energy via DFT-LDA between hydrogenated fullerene structures, with hydrogen site selection based on geometric parameters (pyrA and hybridisation, and sum of two- or three-bond lengths), compared to the lowest-energy hydrogenated isomer found through sequential addition from Section 3.2. Structures are C<sub>28</sub> (1- $T_d$  and 2- $D_2$ ) and C<sub>40</sub> (38- $D_2$  and 40- $T_d$ ).

### 3.4. Third Hypothesis—Predicting Reactive Sites via Local Electronic Structure

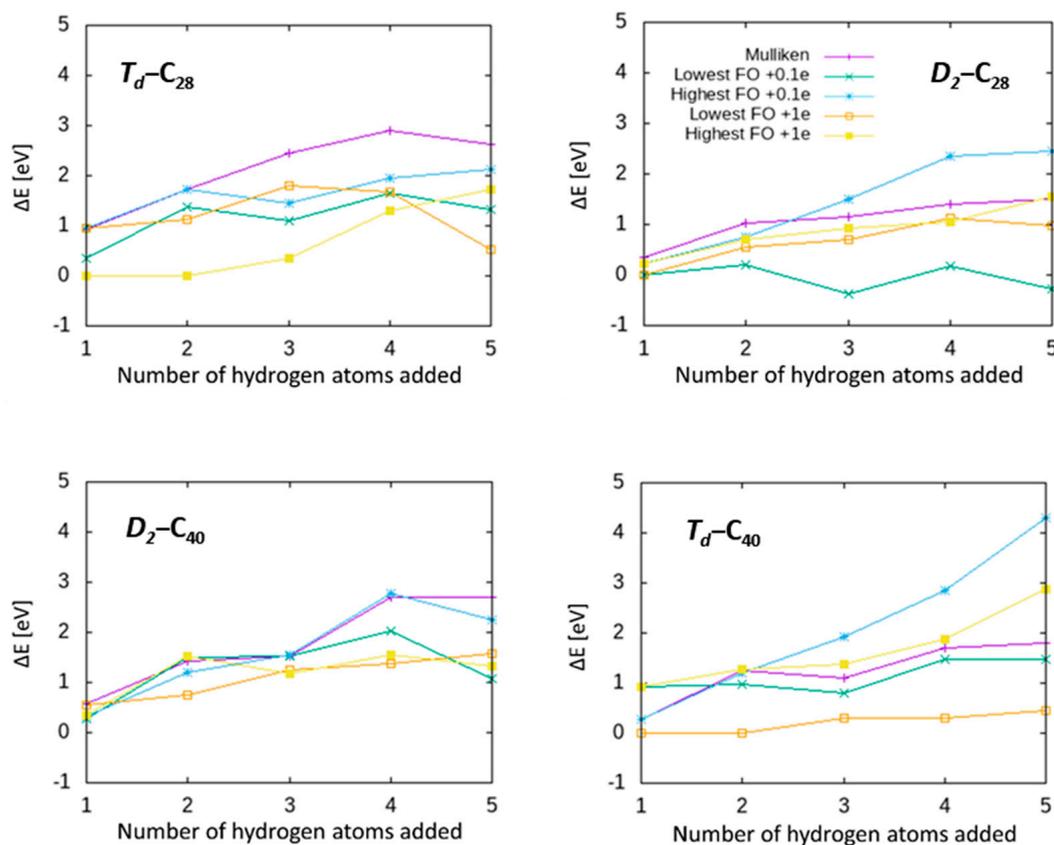
Instead of using local geometry, we next considered potential hydrogenation site selection rules based on the local electronic structure of the parent molecule. It might be imagined that the site with the most local  $p_z$ -orbital character will release the most energy when hydrogenated. Alternatively, in cases where there is charge redistribution (notably on partially hydrogenated cages), sites showing excessive charging may also be targets for hydrogenation. We tested a number of electronic possibilities to select the next site for hydrogenation.

The simplest is to consider local Mulliken charges as determined by a Mulliken population analysis, as an estimate of atomic charge distribution in a molecule [64]. Hydrogenating the atom showing the highest Mulliken charge results in the sequence shown in Figure 5 (blue trace, Schlegel projections in Figure S1), where the lowest energy pathway is clearly not selected. A more sophisticated approach is to use frontier orbitals [65]. In this case, we add or remove some charge to the system, and by taking the difference in Mulliken values between the charged and neutral molecule, we identify where this charge is localising and can use that as an indication of local site reactivity. Such an approach doubles the

number of geometric optimisations required, as for each fullerene, a second optimisation of the charged species is required, but this remains a tractable approach compared to the more “brute force” testing approaches described above. In essence, adding some charge gives an indication of the localisation of the lowest unoccupied molecular orbital (LUMO), with the benefit of geometric optimisation allowing the system to respond to the charge state change.

In the following, we tested two charging values, adding either  $0.1e$  or a full electron to the system. Frontier orbital values for individual atoms were determined as the difference between the calculated Mulliken charge for the atom when the molecule was charged, and the Mulliken charge for that atom in the neutral system. Thus, the lowest Frontier orbital value in the system indicates that for the atom that accumulates the most charge when the  $0.1/1e$  is added to the whole system, the highest Frontier orbital value is the atom that accumulates the least (or indeed loses the most charge if the frontier orbital value is positive). Once the site is chosen, we hydrogenate this atom and continue with geometric optimisation in the neutral charge state as before. We tried the following sequences of both the lowest and highest frontier orbital values.

The results are shown in Figure 6 (Schlegel projections in Figure S1). In general, the frontier orbital values perform better than the simple Mulliken charges as a guide, and most give more stable isomers than the geometric parameters tested above. Nonetheless, none of them are able to systematically reproduce the low-energy structures obtained through sequential addition in Section 3.2. Thus, we have demonstrated that the hypotheses of predicting hydrogenation sites on small fullerenes using the geometric and electronic parameters tested are not valid, i.e. these do not provide reliable routes to predicting hydrogenated structures directly from the parent fullerenes.



**Figure 6.** Plots of the difference in calculated total energy via DFT-LDA between hydrogenated fullerene structures, with hydrogenation sites selected based on the electronic parameters (Mulliken charges or Frontier orbital values) compared to the hydrogenated structures obtained in Section 3.2, for  $C_{28}$  ( $1-T_d$  and  $2-D_2$ ) and  $C_{40}$  ( $38-D_2$  and  $40-T_d$ ).

It is worth examining the Frontier orbital data for  $D_2$ - $C_{28}$  in more detail, as it reveals another important result. It can be seen that when considering the lowest Frontier orbital atoms for hydrogenation, in the case of adding only  $0.1e$ , for  $D_2$ - $C_{28}H_3$  and  $D_2$ - $C_{28}H_5$ , the addition sequence actually predicts hydrogenated isomers that are *more stable* than those obtained in Section 3.2. This is important as it invalidates our first hypothesis, namely that sequential hydrogen addition always generates the lowest-energy hydrogenated fullerenes. The implication is that if hydrogen can rearrange on the fullerene surface as further hydrogen atoms are added, this has the potential to lower the system energy. It is difficult to say whether kinetic barriers will limit this process under laboratory conditions; while calculations for single hydrogen migration suggest relatively high barriers [24], in practice, a range of barriers will exist depending on local geometry, degree of hydrogen surface coverage, and local environment. In interstellar environments where UV-photon absorption is the primary interaction event [14], structural rearrangement is likely.

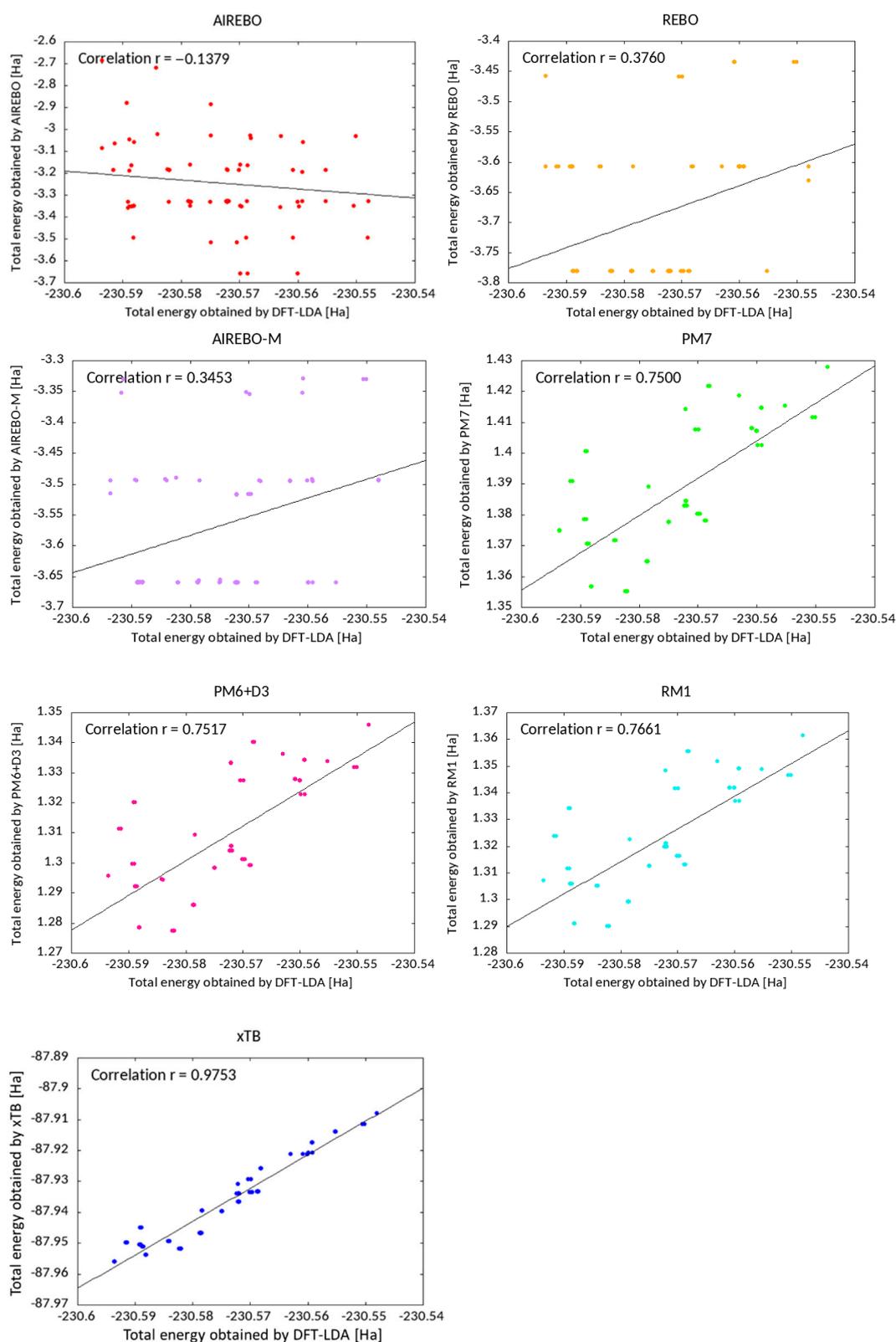
Thus, we have demonstrated that none of the three proposed hypotheses provide a valid route to simplifying structure selection for hydrogenated fullerenes. This leaves a potentially intractable problem given the large number of hydrogenated fullerenes that need testing, and the associated time for each DFT calculation. The next approach is therefore to try and reduce the computational cost of the calculations.

To finish this section, we note that another possible route for functionalisation site prediction based on electronic structure is via local aromaticity, in particular, through the calculation of nucleus-independent chemical shifts (NICS) for each ring in the system [66]. This has previously been successfully applied to fullerene hydrogenation [67] and fluorination [56]. Positive (negative) NICS values indicate less (more) aromatic rings, and the values for three rings surrounding a given atom can be used as an indication of the degree of aromaticity of that site. We did not follow that approach here, as we have shown above that systematic addition routes do not necessarily lead to system ground state structures, but it would be interesting for a future study to explore whether aromaticity can be a useful tool for determining hydrogenation magic numbers.

### 3.5. Second Strategy, Semi-Empirical and Empirical Routes

Rather than applying selection rules to simplify the problem and then using high-level, but slow, DFT tools, an alternative approach is to not simplify the problem but instead tackle every possible hydrogenated fullerene and simplify instead the tool used to make the calculation time acceptable. We used our extensive collection of DFT calculations from Section 3.2 as a comparative benchmark. We tested the four representative structures using the following methods: empirical potentials using LAMMPS [38], semi-empirical molecular orbital approaches using MOPAC [48], and a relatively new semi-empirical extended tight binding method (GFN2-xTB). GFN2-xTB has previously shown excellent structure–energy agreement with DFT results for  $C_{60}$  isomers, with Pearson’s correlation coefficient  $r = 0.998$  [55].

In each case, we geometrically optimised the structure using the method described and then plotted the calculated total energy against that calculated using DFT-LDA. The results are shown for  $C_{40}H_5$  in Figure 7, along with the corresponding Pearson’s correlation coefficient. Plots for two other fullerenes ( $C_{28}H_5$  and  $C_{40}H$ ) are shown in Figures S1 and S2.



**Figure 7.** Total energy correlation plots for isomers of  $C_{40}H_5$ . Energy obtained by DFT-LDA is plotted along the x-axis and by empirical and semi-empirical methods along the y-axes as labelled. Linear correlation in each model is shown as  $r$  in each plot.

The empirical potential-based methods (REBO, AIREBO, and AIREBO-M) show essentially no correlation ( $r < 0.4$ ) with the DFT-LDA energies. In contrast, the MOPAC

semi-empirical-based methods show the same overall energetic stability tendency as DFT, with correlation in the range  $r = 0.75\text{--}0.77$ . All three methods show a very similar distribution of points. However, in the all three cases, the most stable structure by DFT does not correspond to that obtained by the semi-empirical methods, and indeed lies somewhat off (for  $C_{40}H_5$ , using PM7, PM6 + D3, and RM1, the structure predicted by DFT-LDA to be the most stable structure is only 20th on a list of 72). Thus, even if these methods were used to pre-screen suitable structures and a cut-off were applied (with full DFT-LDA geometric optimisation of only the most stable RM1 structures), the cut-off would necessarily have to be quite high, at least 30% of structures, in order to ensure that the lowest-energy DFT-LDA structure was included.

In strong contrast to the above methods, GFN2-xTB shows outstandingly good correlation ( $r = 0.98$ ) in energetic stability compared to DFT-LDA. Enlarging the lowest-energy range, there is a data cluster, and the lowest-energy structure in DFT shows the lowest energy also by the GFN2-xTB calculation isolated from the other close data points. For other simpler fullerenes (less hydrogenated and a smaller cage), the correlation is even better (Figures S1 and S2),  $r = 0.993\text{--}0.998$ . Table 1 shows the calculated energy difference between the ground-state structure and the next most stable (symmetry nonequivalent) structure for  $C_{28}H_n$  and  $C_{40}H_n$ ,  $n = 1 \dots 5$ , calculated using DFT-LDA and GFN2-xTB. In each case, GFN2-xTB predicts the same structure ordering amongst the isomers tested, with very close energy differences (standard deviation of 0.087 eV).

**Table 1.** Comparison of energy difference between lowest energy and second-lowest energy isomer in the calculated series, for  $C_{28}H_n$  and  $C_{40}H_n$ ,  $n = 1 \dots 5$ , calculated at DFT-LDA and GFN2-xTB levels of theory (eV). The difference between these two values and the associated standard deviation is also given.

	DFT-LDA [eV]	GFN2-xTB [eV]	Difference [eV]
$C_{28}H$	0.3494	0.4033	0.0539
$C_{28}H_2$	0.3570	0.4026	0.0456
$C_{28}H_3$	0.3617	0.4553	0.0936
$C_{28}H_4$	0.3319	0.5459	0.2140
$C_{28}H_5$	0.3210	0.2832	−0.0378
$C_{40}H$	0.1137	0.0834	−0.0303
$C_{40}H_2$	0.4171	0.4094	−0.0078
$C_{40}H_3$	0.0375	0.0247	−0.0127
$C_{40}H_4$	0.4699	0.3674	−0.1025
$C_{40}H_5$	0.0008	0.0000	−0.0008
	Standard Deviation =		0.0869

If we compare calculation times for these different methods, the potential advantages are clear. The average run-time of each calculation for pure  $C_{20}$  on a 16-core desktop PC is summarised in Table 2 (along with the corresponding  $r$ -values from Figure 7). As expected, the empirical potentials are fastest but show no correlation to the DFT-LDA total energies. Of the semi-empirical methods, GFN2-xTB shows by far the best run time, as well as gives the best correlation to DFT-LDA. These results show that we can safely switch to using GFN2-xTB, as it is able to successfully generate DFT-LDA accuracy in total energies of hydrogenated smaller fullerenes, but at over 600 times less computational time.

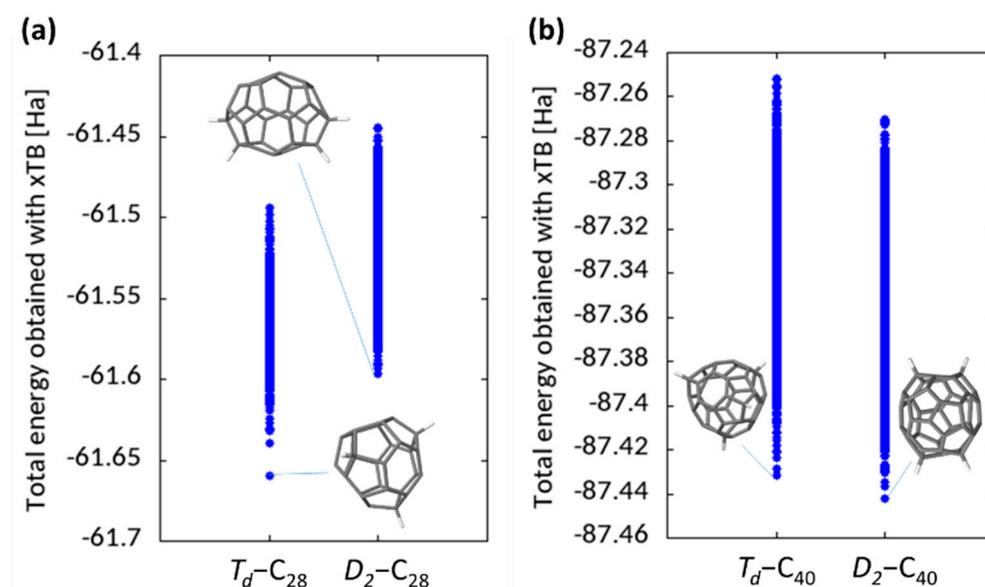
**Table 2.** Comparison of empirical and semi-empirical methods with DFT. Run time is the average of 10 runs for pure  $C_{20}$ .  $r$  is Pearson's correlation coefficient from.

Method	AIREBO	REBO	AIREBO-M	PM7	PM6 + D3	RM1	xTB	DFT
Run time	0.039 s	0.039 s	0.043 s	1.33 s	0.95 s	0.91 s	0.34 s	210 s
$r$	−0.1379	0.3760	0.3453	0.7500	0.7517	0.7661	0.9753	-

### 3.6. All Combination of Hydrogenation Sites Testing by xTB

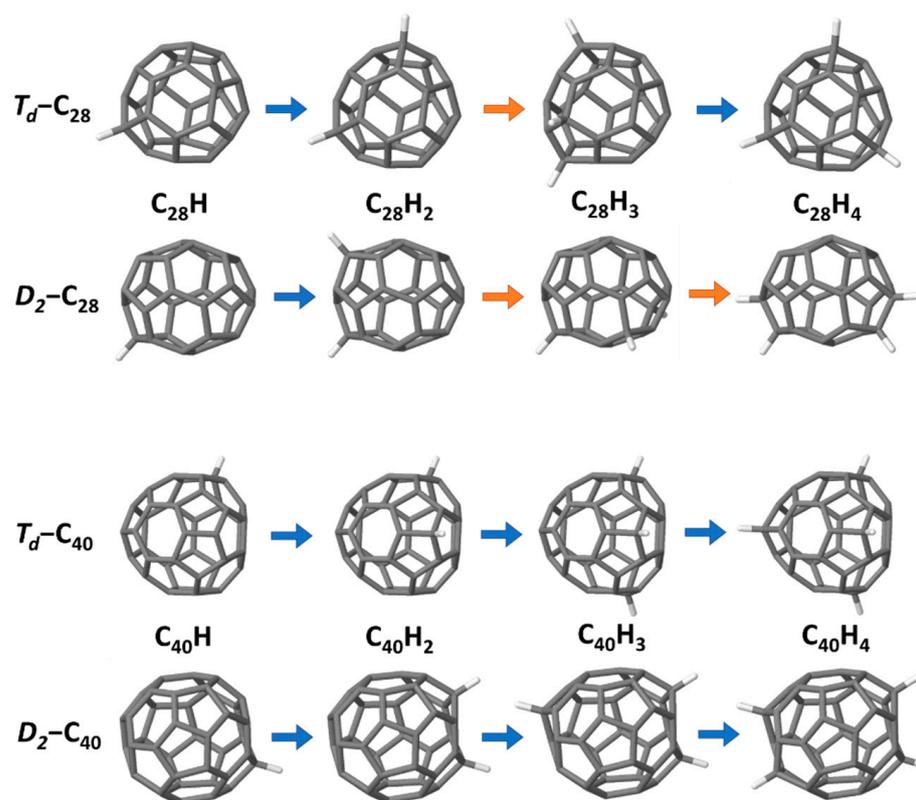
Next, we tested all the possible combinations of hydrogenation sites up to the 4th hydrogenation using GFN2-xTB for the four fullerene structures discussed above,  $T_d$ -C<sub>28</sub>,  $D_2$ -C<sub>28</sub>,  $T_d$ -C<sub>40</sub>, and  $D_2$ -C<sub>40</sub>. This brute force approach calculates the total energy for all hydrogen isomers of fullerenes C<sub>n</sub>H<sub>m</sub>, unlike the sequential test described in Section 3.2. As such, it provides a definitive lowest-energy hydrogenation sequence for C<sub>n</sub>H<sub>m</sub>,  $m = 0 \dots 4$ .

For C<sub>28</sub>H<sub>4</sub>, the most stable structure found is the  $T_d$ -C<sub>28</sub>H<sub>4</sub> shown in Figure 8 (Schlegel projection shown in Figure S1). This has tetrahedral symmetry, and corresponds to the structure found in the literature [23]. This is also the most stable structure found through sequential addition (using DFT-LDA, GFN2-xTB, PM7, PM6 + D3, and RM1). Energies of all the possible isomers are summarised in Figure 8. The single data point at the lowest energy lies significantly below the others, i.e., for C<sub>28</sub>H<sub>4</sub>, there is a single stable hydrogenated structure. The  $D_2$ -C<sub>28</sub> isomer is significantly less stable; within the  $D_2$ -C<sub>28</sub>H<sub>4</sub> subset of results, the most stable isomer consists of two hydrogen pairs at opposing ends of the molecule. This can be understood in terms of a localisation of strain, allowing a maximum flattening out and aromaticity in the two hexagon pairs along the fullerene sides. This structure is different, and more stable, than that found for  $D_2$ -C<sub>28</sub>H<sub>4</sub> by sequential addition in Section 3.2 above (confirmed with DFT-LDA to be more stable by 0.21 eV).



**Figure 8.** Calculated total energy of all the possible (a) C<sub>28</sub>H<sub>4</sub> and (b) C<sub>40</sub>H<sub>4</sub> for  $T_d$  and  $D_2$  isomers obtained by GFN2-xTB, with lowest-energy structures labelled.

The lowest-energy C<sub>n</sub>H<sub>m</sub> structures for H = 1 ... 4 for the two isomers are shown in Figure 9. The  $T_d$  sequence for C<sub>28</sub> is largely sequential functionalisation of the triple pentagons, the exception being C<sub>28</sub>H<sub>3</sub>. In this case, the sequential structure has only a single unfunctionalised triple pentagon, which means the radical is highly localised, and hence, it is more stable to redistribute the hydrogen atoms. In contrast, the  $D_2$ -C<sub>28</sub> lowest-energy series is not at all sequential, with different hydrogenation combinations balancing the relief of localised curvature at the fullerene “ends” with radical redistribution. This flux of different structures is also presumably indicative of the relative instability of this isomer compared to the  $T_d$ .



**Figure 9.** The lowest-energy  $C_nH_m$  structures for  $n = 28$  and  $40$ , and  $m = 1 \dots 4$  for the  $T_d$  and  $D_2$  isomers calculated using GFN2-xTB, testing all symmetry-inequivalent possibilities. Grey and white in molecular structures represent carbon and hydrogen atoms, respectively. Blue arrow indicates hydrogenation in a sequential route, while orange arrow indicates nonsequential hydrogenation.

The situation with  $C_{40}$  is somewhat different. In this case, the most stable structures match exactly the series found through sequential addition in Section 3.2. For the  $T_d$ -isomer, this is presumably because, in this case, the curvature at the triple pentagon is less localised than in  $C_{28}$  due to the larger cage size, allowing more delocalisation of the radical for  $C_{40}H_3$ . For the  $D_2$  isomer, hydrogenation localises the curvature at the fullerene ends, resulting in a cylindrical structure somewhat resembling a small-diameter capped carbon nanotube. The  $D_2$  isomer is the most stable for all species up to  $C_{40}H_4$ . The relative energies show that hydrogenation is increasingly stabilising the  $T_d$  isomer (and indeed, as shown above, at  $C_{40}H_5$ , there is an inversion in stability and the  $T_d$  is more favoured). Further testing will be needed to determine whether, in larger fullerenes such as  $C_{40}$  and above, the lowest-energy hydrogenated fullerenes always match the sequential addition patterns. If so, this would represent a significant computational benefit.

#### 4. Discussion

This study draws out several key points when designing a methodology to explore the vast configurational space of hydrogenated isomers of small fullerenes.

The first is that, unfortunately, it does not appear possible to systematically predict subsequent lowest-energy hydrogenation sites based on either the geometry or electronic structure of a given fullerene or hydrogenated fullerene, at least not with the parameters that we have tested here. There are a number of potential reasons for this, including substantial strain relaxation of the structure, and charge and aromaticity redistribution after hydrogenation.

It has also been demonstrated that sequential hydrogen addition to the lowest-energy site at each step does not always result in the global lowest-energy hydrogenated fullerene isomer. Nonetheless, this procedure generates quite low-energy pathways and, in many

cases, effectively did identify the lowest-energy isomers. It may be possible to use a modified version of this procedure (for example, allowing hydrogen atoms a limited number of site changes to neighbouring sites before subsequent hydrogen addition) in order to improve the robustness of the algorithm while maintaining a tractable number of structures to explore. Additionally, under laboratory hydrogenation conditions, where kinetic barriers for surface migration may be difficult to overcome, this hypothesis may still result in predicting experimentally detectable isomers, which may not necessarily be those with the lowest total energy.

Considering both the calculation costs and the correlation to full DFT-LDA results, GFN2-xTB appears to give reliable and rapid results for hydrogenation of small fullerenes. The predictability is supposed to originate with the DFT-based parameters for the electron density in the reference data.

Comparison with higher levels of theory highlights several points. First, the agreement between xTB and DFT-LDA found here will be fully reproduced with xTB and GGA functionals. Plotting LDA total energies against GGA, GGA-D2, and GGA-D3(BJ) energies for 741 hydrogenated fullerenes shows an  $R^2$  value of  $\sim 1$  (see Supplementary Figure S4). This also validates a choice of either LDA or GGA-Dx for the energetics of these systems. When we compare against hybrid functionals (B3LYP using CAM-B3LYP with the Def2TZVP basis set and GD3BJ dispersion corrections [68]) for the most and second-most stable isomers of  $T_d$ -C<sub>28</sub>H<sub>4</sub> and  $D_2$ -C<sub>28</sub>H<sub>4</sub>, we find the same qualitative energy ordering, but a slight drop in the order of 0.1 eV in the absolute energy differences (see Table S1 in Supplementary Materials). We have chosen to ignore thermal corrections to the energies in the current study as fullerene formation conditions in nonterrestrial environments are far from clear, as well as due to the high additional associated computational cost for vibrational analysis. Calculating thermal contributions with the hybrid functionals for these four fullerenes show that the thermal contribution to the relative stability between isomers is also of the order of  $\sim 0.1$  eV, and again does not change the qualitative stability ordering of the isomers.

In the case of C<sub>28</sub>, the lowest-energy hydrogenated structure is  $T_d$ -C<sub>28</sub>H<sub>4</sub> with tetrahedral hydrogenation; each triple pentagon is hydrogenated. This relaxes the angle strain localised on the tetrahedral corners and is consistent with literature observations. For the  $D_2$ -C<sub>28</sub> isomer, there are two six-fused pentagons along its sides, and the curvature localised at the caps is partially relieved by pairwise hydrogenation at each cap. Nonetheless, the four hydrogen atoms appear insufficient to relax it and the structure is much less stable than the tetrahedral solution is.

In contrast, in the case of C<sub>40</sub>, the  $T_d$  isomer is less stable than the  $D_2$  isomer. Instead, the  $D_2$  isomer, in which there are no triple-fused pentagons but two six-pentagon chains, shows the lowest energy. When we look at the structure of C<sub>40</sub>H<sub>4</sub>, fused pentagon (pentagon chains) sites are indeed hydrogenated [3], not only on the triple pentagons. We suppose that for the larger fullerene cages, pentagon chain structures may play an important role to stabilise the whole structure of non-IPR species.

Given the rapidity and accuracy of the approach using GFN2-xTB, it is now our intention to pursue the most stable C<sub>n</sub>H<sub>m</sub> structures for other small fullerenes, with the eventual intention of mapping hydrogen-catalysed fullerene growth. For such large-scale studies, other methods beyond brute-force testing, such as genetic algorithms and Monte-Carlo-driven methods, will be required.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/cryst11111334/s1>. Figure S1: Schlegel projections of optimal  $C_{28}H_4$  structures obtained by (a) GFN2-xTB all-combination test, PM7, PM6 + D3, RM1, GFN2-xTB, and sequential all-site testing by DFT, (b) hybridisation value and pyrA, (c) 2-bond-lengths sum and 3-bond-lengths sum, (d) Mulliken value, (e) minimum frontier value with 0.1e addition, (f) minimum frontier orbital value with 1e addition, (g) maximum frontier value with 0.1e addition, (h) maximum frontier value with 1e addition, (i) AIREBO, (j) REBO, and (k) AIREBO-M. Figure S2: Total energy correlation plots for isomers of  $C_{28}H_5$ . Energy obtained by DFT-LDA is plotted along the x-axis and by empirical and semi-empirical methods along the y-axes as labelled. Figure S3: Total energy correlation plots for isomers of  $C_{40}H$ . Energy obtained by DFT-LDA is plotted along the x-axis and by empirical and semi-empirical methods along the y-axes as labelled. Figure S4. Correlation plot of total energy calculated using DFT-LDA against DFT-GGA for 741 hydrogenated fullerene species. Actual  $r = 0.99999987$ . Table S1: Comparison of the calculated total energy (eV) for the most stable  $C_{28}H_4$  isomer with the second-most stable  $T_d$ - $C_{28}$  and most stable  $D_{2h}$ - $C_{28}$ -based  $C_{28}H_4$  isomer using GFN2-xTB and CAM-B3LYP.

**Author Contributions:** Conceptualization, C.E.; methodology, C.E. and Y.T.; validation, C.E. and T.M.; investigation, Y.T.; resources, C.E.; data curation, Y.T. and C.E.; writing—original draft preparation, Y.T. and C.E.; writing—review and editing, T.M. and C.E.; supervision, T.M. and C.E.; project administration, T.M. and C.E.; funding acquisition, T.M., Y.T. and C.E. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Inoue Enryo Memorial Grant of Toyo University, C.E. and Y.T. acknowledge BI-FR/21-22-PROTEUS-003 project for funding, and Y.T. acknowledges financial support from French governmental bursary JP19-00207. Computing support was funded by Region Pays de la Loire “Paris Scientifiques 2017”, Grant Number 09375.

**Data Availability Statement:** The data presented in this study are openly available in Zenodo at 10.5281/zenodo.5615405. <https://doi.org/10.5281/zenodo.5615404> (accessed on 27 October 2021) including the most stable  $C_{28}H_n$  and  $C_{40}H_n$  structures and a complete library of GFN2-xTB optimised xyz files for all fullerene isomers from C20 to C80.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. García-Hernández, D.A.; Manchado, A.; García-Lario, P.; Stanghellini, L.; Villaver, E.; Shaw, R.A.; Szczerba, R.; Perea-Calderón, J.V. Formation of Fullerenes in H-Containing Planetary Nebulae. *Astrophys. J. Lett.* **2010**, *724*, L39. [[CrossRef](#)]
2. 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. [[CrossRef](#)]
3. 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. [[CrossRef](#)]
4. Otsuka, M.; Kemper, F.; Cami, J.; Peeters, E.; Bernard-Salas, J. Physical Properties of Fullerene-Containing Galactic Planetary Nebulae. *Mon. Not. R. Astron. Soc.* **2014**, *437*, 2577–2593. [[CrossRef](#)]
5. Campbell, E.K.; Holz, M.; Gerlich, D.; Maier, J.P. Laboratory Confirmation of  $C_{60}(+)$  as the Carrier of Two Diffuse Interstellar Bands. *Nature* **2015**, *523*, 322–323. [[CrossRef](#)]
6. Sellgren, K.; Werner, M.W.; Ingalls, J.G.; Smith, J.D.T.; Carleton, T.M.; Joblin, C.  $C_{60}$  in Reflection Nebulae. *Astrophys. J. Lett.* **2010**, *722*, L54. [[CrossRef](#)]
7. Berné, O.; Cox, N.L.J.; Mulas, G.; Joblin, C. Detection of Buckminsterfullerene Emission in the Diffuse Interstellar Medium. *Astron. Astrophys.* **2017**, *605*, L1. [[CrossRef](#)]
8. Berné, O.; Tielens, A.G.G.M. Formation of Buckminsterfullerene ( $C_{60}$ ) in Interstellar Space. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 401–406. [[CrossRef](#)] [[PubMed](#)]
9. Candian, A.; Gomes Rachid, M.; MacIsaac, H.; Staroverov, V.N.; Peeters, E.; Cami, J. Searching for Stable Fullerenes in Space with Computational Chemistry. *Mon. Not. R. Astron. Soc.* **2019**, *485*, 1137–1146. [[CrossRef](#)]
10. Clayton, G.C.; Marco, O.D.; Whitney, B.A.; Babler, B.; Gallagher, J.S.; Nordhaus, J.; Speck, A.K.; Wolff, M.J.; Freeman, W.R.; Camp, K.A.; et al. The dust properties of two hot r coronae borealis stars and a wolf-rayet central star of a planetary nebula: In search of a possible link. *Astron. J.* **2011**, *142*, 54. [[CrossRef](#)]
11. Foing, B.H.; Ehrenfreund, P. Detection of Two Interstellar Absorption Bands Coincident with Spectral Features of  $C_{60}+$ . *Nature* **1994**, *369*, 296–298. [[CrossRef](#)]

12. García-Hernández, D.A.; Rao, N.K.; Lambert, D.L. Are C<sub>60</sub> Molecules Detectable in Circumstellar Shells of R Coronae Borealis Stars? *Astrophys. J.* **2011**, *729*, 126. [CrossRef]
13. Palotás, J.; Martens, J.; Berden, G.; Oomens, J. The Infrared Spectrum of Protonated Buckminsterfullerene C<sub>60</sub>H<sup>+</sup>. *Nat. Astron.* **2020**, *4*, 240–245. [CrossRef]
14. Omont, A.; Bettinger, H.F. Intermediate-Size Fullerenes as Degradation Products of Interstellar Polycyclic Aromatic Hydrocarbons. *Astron. Astrophys.* **2021**, *650*, A193. [CrossRef]
15. Kroto, H.W. The Stability of the Fullerenes C<sub>n</sub>, with n = 24, 28, 32, 36, 50, 60 and 70. *Nature* **1987**, *329*, 529–531. [CrossRef]
16. Xie, S.-Y.; Gao, F.; Lu, X.; Huang, R.-B.; Wang, C.-R.; Zhang, X.; Liu, M.-L.; Deng, S.-L.; Zheng, L.-S. Capturing the Labile Fullerene[50] as C<sub>50</sub>Cl<sub>10</sub>. *Science* **2004**, *304*, 699. [CrossRef]
17. Prinzbach, H.; Weiler, A.; Landenberger, P.; Wahl, F.; Wörth, J.; Scott, L.T.; Gelmont, M.; Olevano, D.; Issendorff, B.V. Gas-Phase Production and Photoelectron Spectroscopy of the Smallest Fullerene, C<sub>20</sub>. *Nature* **2000**, *407*, 60–63. [CrossRef]
18. Hall, L.E.; McKenzie, D.R.; Attalla, M.I.; Vassallo, A.M.; Davis, R.L.; Dunlop, J.B.; Cockayne, D.J.H. The Structure of Hydrogenated Fullerene (C<sub>60</sub>H<sub>36</sub>). *J. Phys. Chem.* **1993**, *97*, 5741–5744. [CrossRef]
19. Cataldo, F. Fullerene, the Hydrogenated C<sub>60</sub> Fullerene: Properties and Astrochemical Considerations. *Fuller. Nanotub. Carbon Nanostruct.* **2003**, *11*, 295–316. [CrossRef]
20. Iglesias-Groth, S.; García-Hernández, D.A.; Cataldo, F.; Machado, A. Infrared Spectroscopy of Hydrogenated Fullerenes (Fulleranes) at Extreme Temperatures. *Mon. Not. R. Astron. Soc.* **2012**, *423*, 2868–2878. [CrossRef]
21. Goldshleger, N.F.; Moravsky, A.P. Fullerene Hydrides: Synthesis, Properties, and Structure. *Russ. Chem. Rev.* **1997**, *66*, 323. [CrossRef]
22. Zhang, Y.; Sadjadi, S.; Hsia, C.-H.; Kwok, S. Search for Hydrogenated C<sub>60</sub> (Fulleranes) in Circumstellar Envelopes. *Astrophys. J.* **2017**, *845*, 76. [CrossRef]
23. Veljković, M.; Nešković, O.; Djerić, A.; Veličković, S.; Šipka, V. Hypervalent Molecular Cluster: C<sub>28</sub>H<sub>4</sub>. *Mater. Sci. Forum* **2005**, *494*, 181–186. [CrossRef]
24. EL-Barbary, A.A. Hydrogenation Mechanism of Small Fullerene Cages. *Int. J. Hydrog. Energy* **2016**, *41*, 375–383. [CrossRef]
25. Fowler, P.W.; Manolopoulos, D.E. *An Atlas of Fullerenes*; Clarendon Press: Oxford, UK, 1995; ISBN 0-19-855787-6.
26. Yoshida, M. Yoshida's Fullerene Library. Available online: <http://www.jcrystal.com/steffenweber/gallery/Fullerenes/Fullerenes.html> (accessed on 18 November 2020).
27. Sabalot-Cuzzubbo, J.; Salvato-Vallverdu, G.; Bégué, D.; Cresson, J. Relating the Molecular Topology and Local Geometry: Haddon's Pyramidalization Angle and the Gaussian Curvature. *J. Chem. Phys.* **2020**, *152*, 244310. [CrossRef]
28. Rayson, M.J.; Briddon, P.R. Highly Efficient Method for Kohn-Sham Density Functional Calculations of 500–10000 Atom Systems. *Phys. Rev. B Condens. Matter Mater. Phys.* **2009**, *80*, 205104. [CrossRef]
29. Rayson, M.J. Rapid Filtration Algorithm to Construct a Minimal Basis on the Fly from a Primitive Gaussian Basis. *Comput. Phys. Commun.* **2010**, *181*, 1051–1056. [CrossRef]
30. Briddon, P.R.; Rayson, M.J. Accurate Kohn-Sham DFT with the Speed of Tight Binding: Current Techniques and Future Directions in Materials Modelling. *Phys. Status Solidi* **2011**, *248*, 1309–1318. [CrossRef]
31. Hartwigsen, C.; Goedecker, S.; Hutter, J. Relativistic Separable Dual-Space Gaussian Pseudopotentials from H to Rn. *Phys. Rev. B* **1998**, *58*, 3641–3662. [CrossRef]
32. Kohn, W.; Sham, L.J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**. [CrossRef]
33. Briddon, P.R.; Jones, R. LDA Calculations Using a Basis of Gaussian Orbitals. *Phys. Status Solidi* **2000**, *217*, 131–171. [CrossRef]
34. Latham, C.D.; Heggie, M.I.; Alatalo, M.; Öberg, S.; Briddon, P.R. The Contribution Made by Lattice Vacancies to the Wigner Effect in Radiation-Damaged Graphite. *J. Phys. Condens. Matter* **2013**, *25*, 135403. [CrossRef]
35. Nekovee, M.; Foulkes, W.M.C.; Needs, R.J. Quantum Monte Carlo Investigations of Density Functional Theory of the Strongly Inhomogeneous Electron Gas. *Phys. Rev. B* **2003**, *68*, 235108. [CrossRef]
36. Björkman, T.; Gulans, A.; Krasheninnikov, A.V.; Nieminen, R.M. Are We van Der Waals Ready? *J. Phys. Condens. Matter* **2012**, *24*, 424218. [CrossRef] [PubMed]
37. Hof, F.; Impellizzeri, A.; Picheau, E.; Che, X.; Pénicaud, A.; Ewels, C.P. Chainlike Structure Formed in Iodine Monochloride Graphite Intercalation Compounds. *J. Phys. Chem. C* **2021**. [CrossRef]
38. Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1–19. [CrossRef]
39. Brenner, D.W.; Shenderova, O.A.; Harrison, J.A.; Stuart, S.J.; Ni, B.; Sinnott, S.B. A Second-Generation Reactive Empirical Bond Order (REBO) Potential Energy Expression for Hydrocarbons. *J. Phys. Condens. Matter* **2002**, *14*, 783–802. [CrossRef]
40. Stuart, S.J.; Tutein, A.B.; Harrison, J.A. A Reactive Potential for Hydrocarbons with Intermolecular Interactions. *J. Chem. Phys.* **2000**, *112*, 6472–6486. [CrossRef]
41. O'Connor, T.C.; Andzelm, J.; Robbins, M.O. AIREBO-M: A Reactive Model for Hydrocarbons at Extreme Pressures. *J. Chem. Phys.* **2015**, *142*, 024903. [CrossRef] [PubMed]
42. Tersoff, J. New Empirical Approach for the Structure and Energy of Covalent Systems. *Phys. Rev. B* **1988**, *37*, 6991–7000. [CrossRef]
43. Tersoff, J. Modeling Solid-State Chemistry: Interatomic Potentials for Multicomponent Systems. *Phys. Rev. B* **1989**, *39*, 5566–5568. [CrossRef]
44. Brenner, D.W. Tersoff-Type Potentials for Carbon, Hydrogen and Oxygen. *MRS Online Proc. Libr. (OPL)* **1988**, *141*, 59–64. [CrossRef]

45. Tersoff, J. Empirical Interatomic Potential for Carbon, with Applications to Amorphous Carbon. *Phys. Rev. Lett.* **1988**, *61*, 2879–2882. [[CrossRef](#)] [[PubMed](#)]
46. Møller, C.; Plesset, M.S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622. [[CrossRef](#)]
47. Dunning, T.H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023. [[CrossRef](#)]
48. Stewart, J.J.P. *MOPAC2016*; Stewart Computational Chemistry: Colorado Springs, CO, USA, 2016.
49. Stewart, J.J.P. Optimization of Parameters for Semiempirical Methods V: Modification of NDDO Approximations and Application to 70 Elements. *J. Mol. Modeling* **2007**, *13*, 1173–1213. [[CrossRef](#)]
50. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H–Pu. *J. Chem. Phys.* **2010**, *132*, 154104. [[CrossRef](#)]
51. Stewart, J.J.P. Optimization of Parameters for Semiempirical Methods VI: More Modifications to the NDDO Approximations and Re-Optimization of Parameters. *J. Mol. Modeling* **2013**, *19*, 1–32. [[CrossRef](#)]
52. Rocha, G.B.; Freire, R.O.; Simas, A.M.; Stewart, J.J.P. RM1: A Reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I. *J. Comput. Chem.* **2006**, *27*, 1101–1111. [[CrossRef](#)]
53. Dewar, M.J.S.; Zoebisch, E.G.; Healy, E.F.; Stewart, J.J.P. Development and Use of Quantum Mechanical Molecular Models. 76. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909. [[CrossRef](#)]
54. Bannwarth, C.; Caldeweyher, E.; Ehlert, S.; Hansen, A.; Pracht, P.; Seibert, J.; Spicher, S.; Grimme, S. Extended Tight-Binding Quantum Chemistry Methods. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2021**, *11*, e1493. [[CrossRef](#)]
55. Sure, R.; Hansen, A.; Schwerdtfeger, P.; Grimme, S. Comprehensive Theoretical Study of All 1812 C<sub>60</sub> Isomers. *Phys. Chem. Chem. Phys.* **2017**, *19*, 14296–14305. [[CrossRef](#)]
56. Van Lier, G.; Cases, M.; Ewels, C.P.; Taylor, R.; Geerlings, P. Theoretical Study of the Addition Patterns of C<sub>60</sub> Fluorination: C<sub>60</sub>F<sub>n</sub> (n = 1–60). *J. Org. Chem.* **2005**, *70*, 1565–1579. [[CrossRef](#)]
57. Vlandas, A.; Ewels, C.P.; Lier, G.V. Controlling Fullerene Addition Sequences, Regioselectivity and Magic Numbers via Metal Encapsulation. *Chem. Commun.* **2011**, *47*, 7051–7053. [[CrossRef](#)] [[PubMed](#)]
58. Nakagawa, A.; Nishino, M.; Niwa, H.; Ishino, K.; Wang, Z.; Omachi, H.; Furukawa, K.; Yamaguchi, T.; Kato, T.; Bandow, S.; et al. Crystalline Functionalized Endohedral C<sub>60</sub> Metallofullerides. *Nat. Commun.* **2018**, *9*, 3073. [[CrossRef](#)]
59. Ewels, C.; Rio, J.; Niwa, H.; Omachi, H.; Shinohara, H.; Rayson, M.; Briddon, P. Determining Addition Pathways and Stable Isomers for CF<sub>3</sub> Functionalization of Endohedral Gd@C<sub>60</sub>. *R. Soc. Open Sci.* **2018**, *5*, 180588. [[CrossRef](#)]
60. Haddon, R.C. C<sub>60</sub>: Sphere or Polyhedron? *J. Am. Chem. Soc.* **1997**, *119*, 1797–1798. [[CrossRef](#)]
61. Haddon, R.C.; Scott, L.T.  $\pi$ -Orbital conjugation and rehybridization in bridged annulenes and deformed molecules in general:  $\pi$ -orbital axis vector analysis. *Pure Appl. Chem.* **1986**, *58*, 137–142. [[CrossRef](#)]
62. Haddon, R.C. Hybridization and the Orientation and Alignment of  $\pi$ -Orbitals in Nonplanar Conjugated Organic Molecules:  $\pi$ -Orbital Axis Vector Analysis (POAV2). *J. Am. Chem. Soc.* **1986**, *108*, 2837–2842. [[CrossRef](#)]
63. Haddon, R.C. Comment on the Relationship of the Pyramidalization Angle at a Conjugated Carbon Atom to the  $\sigma$  Bond Angles. *J. Phys. Chem. A* **2001**, *105*, 4164–4165. [[CrossRef](#)]
64. Mulliken, R.S. Electronic Population Analysis on LCAO–MO Molecular Wave Functions. I. *J. Chem. Phys.* **1955**, *23*, 1833–1840. [[CrossRef](#)]
65. Fukui, K.; Yonezawa, T.; Shingu, H. A Molecular Orbital Theory of Reactivity in Aromatic Hydrocarbons. *J. Chem. Phys.* **1952**, *20*, 722–725. [[CrossRef](#)]
66. von Ragué Schleyer, P.; Maerker, C.; Dransfeld, A.; Jiao, H.; van Eikema Hommes, N.J.R. Nucleus-Independent Chemical Shifts: A Simple and Efficient Aromaticity Probe. *J. Am. Chem. Soc.* **1996**, *118*, 6317–6318. [[CrossRef](#)] [[PubMed](#)]
67. van Lier, G.; De Proft, F.; Geerlings, P. Ab Initio Study of the Aromaticity of Hydrogenated Fullerenes. *Phys. Solid State* **2002**, *44*, 588–592. [[CrossRef](#)]
68. 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. *Gaussian 16, Revision C.01*; Gaussian, Inc.: Wallingford, CT, USA, 2016.