Determining addition pathways and stable isomers for CF$_3$ functionalization of endohedral Gd@C$_{60}$

Chris Ewels$^1$, Jeremy Rio$^1$, Hiroyuki Niwa$^2$, Haruka Omachi$^2$, Hisanori Shinohara$^{2,3}$, Mark Rayson$^4$ and Patrick Briddon$^4$

$^1$Institut des Materiaux Jean Rouxel (IMN), Universite´ de Nantes, CNRS UMR6502, 2 Rue de la Houssiniere, BP32229, Nantes 44322, France
$^2$Department of Chemistry, Graduate School of Science, Nagoya University, Chikusa 464-8602, Nagoya, Japan
$^3$Institute for Advanced Research, Nagoya City University, Chikusa 464-8602, Nagoya, Japan
$^4$School of Engineering, University of Newcastle, Newcastle upon Tyne NE1 7RU, UK

Using density functional theory approaches, we follow the sequential addition of CF$_3$ functional groups to the surface of the metallic endofullerene species Gd@C$_{60}$. The presence of gadolinium in the interior of the cage strongly influences the addition sequence. The calculations are able to successfully identify end points in the addition sequence at Gd@C$_{60}$(CF$_3$)$_n$, $n = 3$ and two isomers at $n = 5$, in predictive agreement with experiment. Inverting the algorithm to determine the most labile groups also identifies the correct positively charged Gd@C$_{60}$(CF$_3$)$_4$ isomer, as confirmed by experimental mass spectra. The importance of surface mobility, notably at later stage addition, is discussed.

1. Introduction

Functionalization of carbon fullerene molecules is now a mature field, with chemical group addition to the cage surface used to impart new functionality and resulting in species with application in diverse areas, notably light harvesting [1]. Another motivation for functionalization of fullerene cage surfaces is to stabilize the cages, notably for carbon cages that do not obey the isolated pentagon rule (IPR). In these cases, functionalization removes dangling bond or gap states, opening the energy gap between highest occupied (HOMOs) and lowest unoccupied (LUMOs) molecular orbitals and stabilizing the species [2–5]. In a similar
way, functionalization has been proposed as a route to stabilization of certain metallic endofullerenes [5,6]. Introduction of a metallic atom within the cage can sometimes result in a locally reactive cage surface, which results in aggregation and hinders isolation and separation of the metallofullerene species. This is particularly important when there is covalent orbital interaction between the metallic species and the p$_d$-orbitals of the surrounding cage.

An important recent case in point is the encapsulation of a gadolinium atom within the $I_h$-$C_{60}$ fullerene cage. This species has only recently been successfully isolated and its crystal structure identified [1] through a process of in situ trifluoromethylation during arc-discharge synthesis of the metallofullerenes [7]. The resulting CF$_3$ functionalized species have wide HOMO–LUMO gaps not much reduced from the parent C$_{60}$, enabling their crystallization and characterization. In that study, we briefly reported supporting density functional theory (DFT) calculations exploring isomer formation. Insertion and stabilization of heavy atoms like gadolinium inside the C$_{60}$ molecule is of interest for a variety of applications, for example in superconductivity [8] or biomedical applications, in which its very great magnetic moment makes it a promising candidate as a contrast agent for magnetic resonance imaging [9].

In the current publication, we explore the functionalization process via such DFT calculations in significantly more depth, and discuss the importance and limitations of the results and methods employed, as well as extending them to removal of labile surface species through positive ionization. The idea here is to use DFT modelling techniques to determine the position of the functional groups used to stabilize the structure without experimental input. A full exploration of all possible isomers of Gd@C$_{60}$(CF$_3$)$_n$, $n = 1$…6 (or more), would be impossible with currently available computing resources. Simply assuming that different combinations of surface carbon sites are functionalized by CF$_3$ groups results in an enormous number of potential isomers. Additional variability is present in this case owing to the potential rotational variance of the CF$_3$ group, as well as the spatial location of the Gd in the cage interior. Clearly, such a problem is not tractable with a simple ‘brute force’ approach to calculating all possible variants.

We developed a potential solution to this problem several years ago by making the assumption that, in most cases, addition occurs sequentially, and is not followed, at least at low surface coverages, by surface rearrangement [2,3]. The current publication is the first time that the approach has been attempted as a predictive tool with no prior experimental data available. Assuming sequential addition, we first model all possible structures with the addition of a single CF$_3$ group, selecting the most stable from these 60 calculations (symmetry can be used to reduce this number slightly, but the presence of Gd limits this compared with pristine $I_h$-$C_{60}$). The most energetically favourable structure for Gd@C$_{60}$(CF$_3$) is then used to add a second CF$_3$ at all possible addition sites, generating a further 59 possible structures, all of which were optimized. Again the most stable structure was taken for Gd@C$_{60}$(CF$_3$)$_2$ and the process continued for further CF$_3$ addition. This approach is detailed further in [10], and has previously been successfully applied to determine fluorinated isomers of C$_{60}$ [11] and chlorinated isomers of C$_{70}$ [12] in good agreement with experiment. In cases where two isomers were closer than 2.3 kcal mol$^{-1}$ (0.1 eV) in energy and not related by symmetry, both were used as starting points for subsequent CF$_3$ addition. In this way, it is possible to trace a full additional sequence ‘tree’. In total, around 1000 isomers were geometrically optimized for the study. Furthermore, we present supporting experimental mass spectrometry data on charged Gd@C$_{60}$(CF$_3$)$_n$ species.

1.1. C$_{60}$(CF$_3$)$_n$

Trifluoromethylation has been previously used as a route to functionalize pristine C$_{60}$, representing some of the most studied fullerene derivatives [13–19]. There are around 30 fullerene-(CF$_3$)$_n$ species with measured X-ray structures and many others with different physicochemical characterizations, of which nearly 30 C$_{60}$(CF$_3$)$_n$ derivatives have been described ([12] and references therein, and the electronic supplementary material of [20]). All of these involve the addition of an even number of trifluoromethyl groups, from $n = 2$ to $n = 18$, with multiple isomers identified, notably for $n = 8$ and $n = 10$. Functionalization of C$_{60}$ is typically performed post-growth, for example through exposure of C$_{60}$ to gaseous CF$_3$I [12] or silver(I) trifluoroacetate [17] at elevated temperatures (greater than 300°C, the decomposition temperature of CF$_3$I), although there are also studies showing the formation of species up to $n = 8$ from arc-electric fullerene growth using a polytetrafluoroethylene (PTFE)-doped graphite powder electrode [21].

The multiple $n = 2$–18 isomers show different addition pattern behaviours [12]. For $n = 2$, CF$_3$ adds in the para-position across a hexagon [13–17]. The $n = 6$ isomers include 1,6,9,12,15,18-C$_{60}$(CF$_3$)$_6$ a
structure in which a pentagon is isolated from the rest of the cage with CF₃ groups on all back-bond sites, and a sixth CF₃ group on one of the pentagon sites [18]. Structure 1,6,11,18,24,27-C₆₀(CF₃)₆ bears some similarity to the CF₃-functionalized endofullerene structures discussed below, namely CF₃ groups occupy four of the five back-bond sites around a pentagon [16,17]. In the case of 1,6,11,18,24,27-C₆₀(CF₃)₆, the fifth and sixth CF₃ functionalization sites lie close nonetheless, the fifth CF₃ group sitting on a carbon neighbouring the unfunctionalized back-bond site and the sixth CF₃ group sitting on the para-site of the hexagon shared with the fifth group. This motif, of four CF₃ groups sitting on back-bond sites around a shared pentagon, is in general repeated in most of the higher-order isomers (n = 8–18), and results in two double bonds sitting on the central pentagon [12].

A more recent theoretical study has analysed these in terms of thermodynamics and kinetics using a genetic protocol to construct regio-isomers [20]. The authors show that two functionalization processes appear to compete: direct addition/abstraction of CF₃ groups, for which a purely thermodynamic model is sufficient to identify most products, and slower surface addend migration processes for which non-equilibrium kinetic modelling of addition sequences is required.

As we will see below, the introduction of a metal atom within the fullerene sphere significantly modifies this behaviour. There are two computational studies of trifluoromethylation of metallofullerenes in the literature to our knowledge, the first concerning Y@C₆₀(CF₃)₃n [22], the second La@C₆₀(CF₃)ₙ [23]. The first study, combined with experimental mass spectra studies, modelled both Y@C₆₀(CF₃)₃ and Y@C₆₀(CF₃)₄, the latter structure having three CF₃ groups on back bonds around the same shared pentagon (equivalent to the top-right Schlegel diagram in figure 6), for which they found low to moderate HOMO–LUMO gaps of 0.30 eV and 0.81 eV, respectively. The second study explored La@C₆₀(CF₃)ₙ for n = 0–5, with 6–10 regio-isomers selected by hand for each n. At n = 3 the authors identified an equivalent CF₃ arrangement to Y@C₆₀(CF₃)ₖ above as the thermodynamically preferred configuration, along with a high-symmetry C₂ᵥ configuration for n = 4. The structures for n = 3 and n = 4 are not identified as stable addition end points in our calculations, for reasons discussed below.

2. Method

DFT calculations were performed under the local spin density approximation, implemented in the AIMPRO code [24–27]. Relativistic pseudopotentials were included via the Hartwigsen–Goedecker–Hütter scheme [28]. The basis consisted of Gaussian function sets multiplied by polynomial functions including all angular momenta up to maxima p (l = 0, 1), d (l = 0, 1, 2) and f (l = 0–3) [29]. For carbon a pdppd basis set was used, resulting in 38 independent functions, for Gd an fffff basis (90 independent basis functions) and for fluorine ddpp (28 independent functions). Calculations were fully spin polarized with spin relaxation. Periodic boundary conditions at the gamma point were applied, with cell size large enough to avoid interaction between neighbouring fullerenes. A system-dependent plane wave energy cut-off of 300 Ha (Ha, Hartree energy) and a non-zero electron temperature of kT = 0.04 eV for electronic-level occupation were taken. Atomic positions were geometrically optimized until the maximum atomic position change in a given iteration dropped below 10⁻⁶ a₀ (a₀, Bohr radius). Diffusion barrier calculations were obtained via the climbing nudged elastic band between fully optimized end points. Charge and spin states of individual atoms were obtained using Mulliken population analysis. Symmetry was deliberately broken, and none imposed during optimization, to ensure the structures were not symmetrically over-constrained. In total around 1000 structures were geometrically optimized for this study. For comparison we also optimized key stable structures at the unrestricted GGA level of theory (using PBE [30] and Krack pseudopotentials [31]), obtaining almost identical structures and HOMO–LUMO gaps with both functionals.

The enthalpies of reaction are calculated using (CF₃)₂ as the reference state for the trifluoromethyl groups:

\[ \text{Gd@C}_{60} (\text{CF}_3)_n + \frac{1}{2} (\text{CF}_3)_2 \rightarrow \text{Gd@C}_{60} (\text{CF}_3)_{n-1} \]

We take the enthalpy difference as an indication of the relative stability in large part due to computational expediency. This is something of an assumption, although, given the similarity in structural motifs, we might expect the difference in entropy between isomers to be small. We note also that the HOMO–LUMO gap in our stable structures is in general much larger than that in other isomers (as discussed below), even though this is not used as a selection criterion. Kinetic stability is also discussed further below.
3. Results

3.1. Gd@C₆₀

In Gd@C₆₀, the Gd atom drifts away from the fullerene centre, sitting instead below a hexagon with a Gd–C average bond length of 2.407 Å (figure 1). There is a net calculated spin on Gd of 7.14 μB and total system spin of 6.71 μB. This is consistent with previous theoretical calculations [32], which also found that Gd is stable when adjacent to a hexagon centre with Gd–C distances of 2.38–2.41 Å and spin density of 6.74 μB on Gd. We also tested Gd inside the C₂ᵥ–C₆₀ isomer no. 1809 (structurally obtained from Ih-C₆₀ through a 90° bond rotation of a single C–C shared hexagon bond about the bond centre, resulting in two paired pentagons on the fullerene surface). While this isomer is 36.9 kcal mol⁻¹ less stable than Ih-C₆₀ when empty, this energy difference drops to only 20.45 kcal mol⁻¹ in the presence of encapsulated Gd. Hence, while Gd appears able to stabilize non-IPR isomers, the energy difference is still sufficient such that we do not expect a significant population of Gd@#1809-C₆₀ metallic endofullerene in the sample. This is consistent with the experimental X-ray data, which show only Gd@Ih-C₆₀ in the trifluoromethylated species [4].

3.2. Gd@C₆₀(CF₃)ₙ, n = 1–5

The CF₃ addition sequence in the following section is predicated on a number of assumptions. Firstly, we assume the CF₃ groups add sequentially, and, once the stable addition site has been found, they are not subsequently surface mobile. Secondly, we compare the CF₃ binding energy with that of a (CF₃)₂ molecular species, on the assumption that, once this comparative binding energy becomes positive, there will be a net energetic driving force for CF₃ species to combine and leave the fullerene surface, and hence addition will cease. Finally, we assume that, unless the two most stable symmetry-distinct isomers have calculated energies within 0.1 eV of each other (approx. 15–20% of the CF₃ binding energy), the second most stable isomer will be statistically so unlikely to form that it can be ignored. When two isomers are close in energy, we take both as potential starting points for subsequent CF₃ addition, allowing us to trace multiple addition pathways. We will show below that these assumptions lead to a remarkable predictive agreement with experimentally determined isomers.

Figure 2 shows a plot of the binding energies for CF₃ in the most stable addition sequence to Gd@C₆₀. Note that the axis does not show the average binding energy for each CF₃ group at this surface density, but rather the energy change involved in the addition of a new CF₃ group.

The plot shows that CF₃ addition is stable up to (CF₃)₃ before above this addition becomes endothermic. Odd-numbered addition is energetically the most favourable, and notably addition to form Gd@C₆₀(CF₃)₃ is extremely favourable, with (CF₃)₄ acting as a barrier before the addition sequence completes at (CF₃)₅. Thus these calculations suggest that n = 3 and n = 5 are likely to be stable ‘magic number’ species.

The system spin fluctuates between 7.0 μB for all odd-numbered additions (including the stable n = 3 and n = 5) and 6.4–7.3 μB for even-numbered additions. We note that figure 2 only shows a zoom around the lowest energy isomers, with significantly less stable structures found with enthalpies of reaction of up to +32 kcal mol⁻¹.
For the initial \( n = 1 \) addition, there are around 20 structures with energies within 2 kcal mol\(^{-1}\) of each other. This is because Gd is not constrained during optimization and alters its position within the cage, rendering all structures roughly equivalent when the Gd atom sits beneath a hexagon with a shift towards a neighbouring pentagon, and addition occurs on one of the back bonds of this pentagon. To test our initial assumption that CF\(_3\) groups will not be surface mobile, we calculated the diffusion barrier for CF\(_3\) to move to a neighbouring site, using a nudged elastic band method. The results (figure 3) show that the barrier is surprisingly high at 52.7 kcal mol\(^{-1}\), with the CF\(_3\) essentially de-bonding and re-bonding to the neighbouring site, confirming that CF\(_3\) will not be surface mobile.

For the \( n = 2 \) second addition, two distinct sites can be identified. The most stable addition is at one of the two symmetrically equivalent sites, continuing CF\(_3\) addition around the back bonds of the pentagon neighbouring the Gd (figure 4a). However, a second addition site, shown in figure 4b, is only 1.0 kcal mol\(^{-1}\) less stable, where Gd migrates once more towards the central hexagon. The plot in figure 2 shows continued addition to the structure shown in figure 4a.
Before continuing with the primary addition sequence, we first explore further CF\textsubscript{3} addition to the \( n = 2 \) isomer shown in figure 4\textit{b}. Once again there are two addition sites quite close in energy for \( n = 3 \). The most stable site is shown as an empty circle in figure 4\textit{b} and results in a C\textsubscript{3} symmetry isomer Gd@C\textsubscript{60}(CF\textsubscript{3})\textsubscript{3} as shown in figure 5. Further addition to this structure is endothermic, showing that it is a stable end structure for addition and a predicted magic number. The second possible addition site to figure 4\textit{b}, \( 0.93 \text{ kcal mol}^{-1} \) less stable than that shown in figure 5, has only one stable subsequent CF\textsubscript{3} addition site for \( n = 4 \), which returns it to the primary addition sequence (shown in figure 6 below). Thus in conclusion, this side branch at \( n = 2 \) offers two alternatives, one of which results in a new predicted stable C\textsubscript{3} symmetry isomer of Gd@C\textsubscript{60}(CF\textsubscript{3})\textsubscript{3}, while the other continues to add CF\textsubscript{3} and returns to the primary addition pathway.

The entire addition pathway, including the bifurcating path at \( n = 2 \), is shown in figure 6. The primary addition pathway, shown around the right edge of the diagram, continues with CF\textsubscript{3} group addition to the back bonds of a single pentagon. The exception to this occurs for the final addition step at \( n = 5 \), where, instead of completing the pentagon enclosure, the final CF\textsubscript{3} group adds at a more distant site corresponding to the secondary site favoured at \( n = 2 \). An alternative \( n = 5 \) isomer is also close in energy, where the CF\textsubscript{3} group sits slightly further away. Together these two structures represent the predicted stable ‘magic number’ structures for Gd@C\textsubscript{60}(CF\textsubscript{3})\textsubscript{5} and are shown in figure 7. We note that it is extremely endothermic to attempt to complete the sequence of five CF\textsubscript{3} groups around the pentagon; of all possible addition sites for \( n = 5 \) on the cage, this is the second least stable site and is 20.54 kcal mol\textsuperscript{-1} less stable than isomer I shown in figure 6.

The calculations are in excellent agreement with experimental observation. The three primary observed species after CF\textsubscript{3} functionalization of Gd@C\textsubscript{60} are Gd@C\textsubscript{60}(CF\textsubscript{3})\textsubscript{3} and the two isomers of Gd@C\textsubscript{60}(CF\textsubscript{3})\textsubscript{5} identified in figure 7 [4]. X-ray crystallography confirms that the experimental isomers indeed correspond to those identified here, including the underlying location of Gd. Additionally, the most stable isomer (figure 7\textit{a}) is found in roughly double the quantities of the isomer shown in figure 7\textit{b}, consistent with the calculations, which both show it to be the energetically most stable isomer, and also the isomer representing the end point of the majority of the addition sequences. We note that the calculated CF\textsubscript{3}
migration barrier between the Gd@C_{60}(CF_3)_{5}^{(I)} and Gd@C_{60}(CF_3)_{5}^{(II)} isomer is only 24.7 kcal mol^{-1}, and hence it is feasible at room temperature that these two isomers may interchange over a time scale of seconds.

There is a clear pattern in the addition behaviour. The bifurcation in addition routes at \( n = 2 \) is linked to the position of the Gd within the cage. The separated addition route (lower line, figure 6) corresponds to Gd remaining located next to its initial hexagonal site. The primary addition route (top line, figure 6) corresponds to Gd migrating closer to the central pentagon and its associated functional groups, in most cases located under the pentagon–hexagon bond, and at the end of the addition sequence, under the neighbouring hexagon–hexagon bond. Thus Gd appears to play a dual role when selecting CF_3 addition sites, both

**Figure 6.** Schlegel projections showing calculated CF_3 addition pathways to Gd@C_{60} (top) to stable trifluoromethylated species (bottom). Previous addition sites are marked with blue circles; new addition sites are marked with red circles. Green shading indicates the location of the Gd atom. Numbers indicate energy release in kcal mol^{-1} (from figure 2). All pathways are shown where isomers are within 2.3 kcal mol^{-1} of the most stable. Fullerenes at the bottom of the figure where further addition is unfavourable have shaded backgrounds and a dark outline, and represent stable end points of a reaction pathway (adapted from Nakagawa et al. [6]).

**Figure 7.** Calculated stable isomer structures for Gd@C_{60}(CF_3)_{5}: (a) Gd@C_{60}(CF_3)_{5}^{(I)} and (b) Gd@C_{60}(CF_3)_{5}^{(II)}. Gd, pink; C, grey; F, yellow.
acting as charge donor and enhancing surface charge localization, but also through the strong hybridization between the Gd valence states and the C 2p states of the fullerene. This suggests that we would expect different addition sequences for non-covalently bound endohedral +3 oxidation species.

It is important to note that the structures predicted by this method do not necessarily correspond to the thermodynamically most stable regio-isomers. For example for \( n = 3 \), the structure top right of figure 6 is 12.81 kcal mol\(^{-1}\) more stable than the structure at the bottom left (in agreement with previous calculations for \( \text{La@C}_{60}(\text{CF}_3)_3 \) [23]), yet it is the bottom-left structure that we predict to be experimentally observable (as confirmed by experiment [6]). In this case, the thermodynamically more stable regio-isomer is not stable because it is simply an intermediate structure en route to further CF\(_3\) functionalization. This shows that kinetic effects during functionalization cannot be ignored, as has also been shown for some pristine C\(_{60}\) trimethylfluorinated species [20].

### 3.3. Role of Gd in the addition sequence: metalloocene behaviour

Inclusion of Gd within the C\(_{60}\) cage significantly alters the functionalization behaviour of the fullerene. Pristine C\(_{60}\) shows a wide range of isomers upon trifluoromethylation, with comparatively high surface densities (up to \( n = 18 \) CF\(_3\) functional groups and higher; see Introduction). By contrast, Gd@C\(_{60}\) shows only the three stable isomers reported here for \( n = 3 \), 5(I) and 5(II). As the functionalization methods are different, we cannot fully exclude the possibility of higher surface coverage species if, for example, Gd@C\(_{60}\)(CF\(_3\))\(_5\) is subsequently exposed to CF\(_3\)I at 550 °C; however, the calculations suggest this should remain unlikely as such processes would be endothermic compared with (CF\(_3\))\(_2\) production.

As might be expected, while pristine C\(_{60}\) is only stable with even-numbered \( n \), Gd@C\(_{60}\) prefers odd-numbered \( n \) due to its +3 charge state and hence initial open-shell configuration of the unfunctionalized metallofullerene [12].

Some insight into the behaviour and bonding of the Gd can be taken from Gd metallocene complexes. Gd is known to form stable metalloocene species, notably the Tris(cyclopentadienyl)gadolinium(III) complex Gd(C\(_5\)H\(_5\))\(_3\) consisting of three pentagonal cyclopentadienyl units surrounding the central Gd atom. Other species include the bromine-stabilized metalloocene dimer (Gd(C\(_5\)H\(_5\))\(_2\))\(_2\)-Br\(_2\) [33]. Unlike sandwich metalloccenes such as ferrocene, this species has the cyclopentadienyl tilted away from the dimer complex centre, opening up access to the Gd atom to allow complexation with bromine. In both cases, Gd sits surrounded by anionic pentagonal C\(_5\)H\(_5\) groups.

In our case, by sp\(^3\) hybridizing the back bonds of the pentagon in Gd@C\(_{60}\)(CF\(_3\))\(_5\), CF\(_3\) functionalization effectively isolates a structural equivalent of a C\(_5\)H\(_5\) unit on the fullerene surface. However, a pair or triplet of these, as energetically preferred by the Gd, is not possible within the confines of the C\(_{60}\) cage structure. The closest it can come to this is by positioning the Gd between this semi-isolated pentagon and its nearest pentagon neighbour, beneath the interconnecting bond. The fullerene curvature allows Gd to interact with both canted pentagons. While the first pentagon is surrounded by CF\(_3\) groups and hence has single back bonds and can reproduce C\(_5\)H\(_5\) anionic behaviour, the second pentagon is not able to form an equivalent isolated anion because CF\(_3\) addition becomes energetically unfavourable after \( n = 5 \). Nonetheless, the remaining single CF\(_3\) addition sits at sites which begin this process.

Gd sits beneath the interconnecting bond between these two pentagons. The two covalent bonds it would normally form with a halide in (Gd(C\(_5\)H\(_5\))\(_2\))\(_2\)-Br\(_2\) are partially replaced by formation of covalent bonds with this core carbon atom pair (replacing what would be a double bond in a conventional fullerene). Thus this model would suggest that CF\(_3\) addition allows Gd to recreate an approximation towards a local metalloocene-like environment. We note that this double bond replacement is similar to the structure seen in C\(_{60}\)(CF\(_3\))\(_2\)O, where, in this case, an epoxide bond is formed on the fullerene exterior of this core carbon atom pair [18].

### 3.4. Positive +1 charge state: Gd@C\(_{60}\)(CF\(_3\))\(_4\)\(^+\)

It is interesting to explore the energy landscape for the +1 charged system. During laser desorption positive-ion mass spectrum measurements, laser excitation excites and removes one electron. We approached this problem in two distinct ways but arrived at similar conclusions. The first was simply to repeat the addition pathway determination as discussed above for the neutral species, but with the cage in a +1 charge state at each step. When we do so, we obtain a similar addition pathway and stable species, with the exception that the second route towards C\(_3\) – Gd@C\(_{60}\)(CF\(_3\))\(_5\) becomes destabilized. Importantly, addition at \( n = 5 \) becomes endothermic. Thus the addition pathway is expected to finish at \( n = 4 \) when in the +1 charge state with Gd@C\(_{60}\)(CF\(_3\))\(_4\), the stable positively
charged isomer. As the second addition route is destabilized, the calculations predict only one stable isomer, where all four CF3 groups sit at back-bond sites to the same pentagon.

A second approach to the problem is to assume that stable neutral Gd@C60(CF3)5 species will be photoexcited and lose an electron. Given that the above calculations show \( n = 5 \) is unstable in the +1 charge state, we removed each of the CF3 functional groups from Gd@C60(CF3)\(^+\)\(^5\) in the +1 charge state to determine which was most labile (figure 8). Consistent with the sequential addition calculations, removal of the CF3 groups around the pentagon is endothermic, and only removal of the most isolated CF3 group is exothermic. Thus both sets of calculations suggest there will be one clearly stable species and isomer, Gd@C60(CF3)\(^+\)\(^4\), in the +1 charge state.

This result can be compared directly to experiment. The full sample preparation method is given in [6]. In brief, a DC arc-discharge chamber with flowing He atmosphere is used to produce raw soot, using graphite rods as electrodes, with Gd impregnation in the anode and PTFE rods held nearby. The resultant soot then undergoes \( \sigma \)-xylene extraction to exclude amorphous carbon. A TiCl4 Lewis acid method [34,35] is then used to exclude empty fullerenes, giving a mixture of metallofullerenes. From these, high-performance liquid chromatography (HPLC) purification using a Buckyprep-M column results in specific Gd@C\(_60\) derivatives, which can be fractionated out using another Buckyprep column. The data presented here come from the highest intensity column fragment.

Figure 9 shows matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectra in both positive- and negative-ion modes for separated derivatives of Gd metallofullerenes. Dithranol was used as a MALDI matrix. It should be noted that CF3- groups are detached from Gd metallofullerenes at the time of laser desorption and ionization, so that only the intact Gd@C\(_6\) metallofullerenes are observed. The negative spectrum clearly shows the preferential formation of Gd@C\(_{60}\)(CF3)\(^+\)\(^n\) \( n = 3 \) and \( n = 5 \) species as discussed above and in [6]. In contrast in the positive-ion mode, corresponding to the calculations we have performed here, it can be seen that one of the CF3 groups in Gd@C\(_{60}\)(CF3)\(^+\)\(^5\) is clearly highly labile, because the concentration of \( n = 5 \) species is significantly diminished, replaced instead with a high concentration of \( n = 4 \) species. Equivalent CF3 loss for the \( n = 3 \) species is not seen, and the \( n = 4 \) is only present in trace quantities in the negative-ion case. This result is entirely consistent with the calculations presented above.

4. Summary and conclusion

In this article, we show that it is possible to successfully predict stable magic number isomers for the trifluoromethylation of Gd@C\(_{60}\) by considering the CF3 addition process as a sequential addition...
process to most stable bonding sites. Stable magic number species are identified when subsequent
addition is endothermic when compared with (CF₃)₂ formation. This approach successfully identifies
two Gd@C₆₀(CF₃)₅ and one Gd@C₆₀(CF₃)₃ stable isomers which have been identified and
characterized experimentally. Equally the method works well for charged species such as the
Gd@C₆₀(CF₃)⁺ cation. The CF₃ bonding arrangement and Gd positioning can be rationalized within
the framework of a Gd metallocene picture.

The trifluoromethyl functionalization behaviour of Gd@C₆₀ is very different from that of pristine C₆₀.
Along with a switch from addition of even to odd numbers of groups, the wide range of isomers for
pristine C₆₀(CF₃)ₙ is reduced to only three for Gd@C₆₀(CF₃)ₙ, and maximum coverage is significantly
reduced to only five CF₃ groups. Different addition patterns are observed, although there is some
motif similarity with functionalization of back-bond sites around a shared pentagon.

An interesting future study would be to repeat the CF₃ addition calculations for a sequence of related
systems, namely (i) a neutral C₆₀, (ii) C₆₀ with a +3 charge state but no metal atom, and (iii) C₆₀ with a +3
non-covalently bound endohedral metal atom such as Al. Such a sequence of calculations would allow
separation of the effects associated with charge transfer, charge location and covalent binding with Gd on
the resultant addition sequence. The approach should also be easily adaptable to related systems such as
Gd@C₇₄(CF₃)ₙ discussed elsewhere in this issue [36].

Data accessibility. All structures (in xyz format) discussed in this article are provided in a .tar file in the electronic
supplementary material associated with this article, along with ASCII files tabulating the calculated total energy of
each structure in Hartrees.

Authors’ contributions. C.E. and H.S. conceived the study. C.E. designed and performed the calculations, and wrote the
manuscript. H.N., H.O. and H.S. produced the experimental MALDI-TOF data. J.R. analysed data and produced
figures. M.R. and P.B. developed the density functional code used. All the authors discussed the results, helped
edit the manuscript and gave their final approval for publication.

Competing interests. The authors declare that we have no competing interests.

Funding. J.R. and C.E. acknowledge funding from Region Pays de la Loire ‘Paris Scientifiques 2017’ grant no. 09375 and
CCIPL for computing resources. M.R. thanks the Swedish Research Council (reg. no. 2012-3174) for financial support.

Acknowledgements. We thank the referees for their useful and insightful comments.

References

1. Brabec CJ, Saničíří NS, Hummelen JC. 2001
Plastic solar cells. Adv. Funct. Mater. 11, 15–26.
(doi:10.1002/1616-3028(20010211:1<15:160-ADFM15-3:0.CO;2-A))
2. Guo T, Smalley RE, Scuseria GE. 1993 Ab initio
theoretical predictions of C₂₈, C₂₈H₄, C₂₈F₄,
(Ti⁺)C₂₈H₄ and Mg@C₈₂ (M = Mg, Al, Si, S,
Ca, Sc, Ti, Ge, Zr, and Sn). J. Chem. Phys. 99,
352–359. (doi:10.1063/1.465738)
3. Xie S-Y. 2004 Capturing the labile fullerene[S]₀
as CS₀C₁₀. Science 304, 699. (doi:10.1126/
science.1095567)
4. Shustova NB, Kurychiko IV, Bolskar RD, Sepeht,
K, Strauss SH, Popov AA, Bolotsina OW. 2006
Trifluoromethyl derivatives of insoluble small
HOMO-LUMO-gap hollow higher fullerenes.
NMR and DFT structure elucidation of Ζ₆(C₆₀)
Dₚ(FC₃)₁₂, Ζ₆(C₆₀-Tₐ(2))(FC₃)₁₂, Ζ₆(C₆₀-
Dₚ(5))(FC₃)₁₂, Ζ₆(C₆₀-S)(FC₃)₁₂ and
Ζ₆(C₆₀-C₆₀-S)(FC₃)₁₂. J. Am. Chem. Soc. 128,
15 793–15 798. (doi:10.1021/ja0651781)
5. Ahara J. 1999 Reduced HOMO-LUMO gap as an index of kinetic stability for polyacrylamic hydrocarbon. J. Phys. Chem. A 103, 7487–7495. (doi:10.1021/jp990092b)

6. Nakagawa A et al. 2018 Crystalline functionalized endohedral Ce@C60 metallofullerenes. Nat. Commun. 9, 3073. (doi:10.1038/s41467-018-05496-8)

7. Wang Z, Anyagi S, Omachi H, Kitaura R, Shinohara H. 2015 Isolation and structure determination of a missing endohedral fullerene La(C70) through in situ trifluoromethylation. Angew. Chem. Int. Ed. 55, 199 – 202. (doi:10.1002/anie.201508082)

8. Gannin RH et al. 2010 Polymorphism control of superconductivity and magnetism in Cs3C60 close to the Mott transition. Nature 466, 221 – 225. (doi:10.1038/nature09120)

9. Dom HC, Fatouros PP. 2010 Endohedral metallofullerenes: applications of a new class of carbonaceous nanomaterials. Nanosci. Nanotechnol. Lett. 2, 65 – 72. (doi:10.11611/nnl.2010.1080)

10. Ewels CP, Van Lier G, Geerlings P, Charlier J-C. 2007 Meta-code for systematic analysis of chemical addition (SACHA): application to fluorination of C70 and carbon nanostructure growth. J. Chem. Inform. Model. 47, 2208 – 2215. (doi:10.1021/ci700121h)

11. Van Lier G, Cases M, Ewels CP, Taylor R, Geerlings P. 2005 Theoretical study of the addition patterns of C60 fluorination: C60F6n (n = 1–60). J. Org. Chem. 70, 1565 – 1579. (doi:10.1021/jo0504837u)

12. Vlandas A, Ewels CP, Van Lier G. 2011 Controlling fullerene addition sequences, regioselectivity and magic numbers via metal encapsulation. Chem. Commun. 47, 7051. (doi:10.1039/c1cc11072a)

13. Popov AA, Kareev IE, Shustova NB, Stukalin EB, Lebedkin SF, Soppeil K, Strauss SH, Boltalina OV, Dunsch L. 2007 Electrochemical, spectroscopic, and DFT study of C60(CF3)n frontier orbitals (n = 2 – 18): the link between double bonds in pentagons and reduction potentials. J. Am. Chem. Soc. 129, 1155 – 1156. (doi:10.1021/ja071816e)

14. Avent AG, Boltalina OV, Goryunov AA, Darwin AD, Markov YV, Taylor R. 2002 Characterisation of C60(CF3)2: fullerenes. Nanotubes Carbon Nanotext. 10, 235 – 241. (doi:10.1016/S12001473(08)

15. Darwin AD, Abdul-Sada AK, Avent AG, Lyakhovetsky Y, Shilova EA, Taylor R. 2003 Unusual addition patterns in trifluoromethylation of [60]fullerene. Org. Biomol. Chem. 1, 3102. (doi:10.1039/b305741k)

16. Darwin AD, Avent AG, Abdul-Sada AK, Taylor R. 2003 [60]- and [70]fullerenes are trifluoromethylated across 5,6-bonds. Chem. Commun. (12), 1374 – 1375. (doi:10.1039/b303496h)

17. Goryunov AA et al. 2005 Trifluoromethylated [60]fullerenes: synthesis and characterization. Fullerenes, Nanotubes Carbon Nanostruct. 12, 181 – 185. (doi:10.1081/FST-20007154)

18. Aihara J. 1999 Reduced HOMO-LUMO gap as an index of kinetic stability for polyacrylamic hydrocarbon. J. Phys. Chem. A 103, 7487–7495. (doi:10.1021/jp990092b)

21. Fritz HP, Hiemeyer R. 1995 Formation in situ of perfluorolylated fullerenes. Chem. Soc. Rev. 14, 307 – 317. (doi:10.1039/b303496h)

25. Rayson MJ. 2009 Highly efficient separation of metallofullerenes by HPLC. J. Chem. Soc. Commun. 10, 1051 – 1053. (doi:10.1039/b817786a)

27. Bridgdon PR, Rayon MU. 2010 Accurate Kohn-Sham DFT with the speed of tight binding: Current techniques and future directions in materials modelling. Phys. Status Solidi B 248, 1309 – 1318. (doi:10.1002/pssb.201000417)

29. Goss JP, Shaw MJ, Briddon PR. 2006 Marker-method calculations for electrical levels using Gaussian-ortibal basis sets. In Topics in applied physics, pp. 69 – 94. Berlin: Springer. Perdew JP, Burke K, Ernzerhof M. 1996 Generalized gradient approximation made simple. Phys. Rev. Lett. 77, 3865 – 3868. (doi:10.1103/PhysRevLett.77.3865)

30. Rayson MJ. 2010 Rapid filtration algorithm to optimize for gradient-corrected exchange-correlation functional. Theor. Chem. Acta. 114, 145 – 152. (doi:10.1007/s00214-005-0655-y)

31. Lu J et al. 2006 Structure and electronic properties of Gd@C60: all-electron relativistic total-energy study. Chem. Phys. Lett. 425, 82 – 84. (doi:10.1016/j.cplett.2006.04.098)

32. Lamberts W, Lueken H. 1986 Structural chemistry of lanthanide cyclopentadienylhexahalides. Part 1: Two modifications of gadoliniumcyclopentadienylhexahalide, [Gd(C5H5)2Br] and [Lu(C5H5)2Br]. Angor. Chim. Acta. 121, 81 – 87. (doi:10.1002/622395_00122-7)

34. Akiyama K, Hamano T, Nakashima Y, Takuchii E, Noda S, Wang Z, Kubuki S, Shinohara H. 2012 Non-HPLC rapid separation of metallofullerenes and empty cages with TiG Lewis acid. J. Am. Chem. Soc. 134, 9762 – 9767. (doi:10.1021/ja3036257)

35. Akiyama K, Hamano T, Nakashima Y, Takuchii E, Noda S, Wang Z, Kubuki S, Shinohara H. 2012 Non-HPLC rapid separation of metallofullerenes and empty cages with TiG Lewis acid. J. Am. Chem. Soc. 134, 9762 – 9767. (doi:10.1021/ja3036257)

36. Nakagawa A, Anyagi S, Omachi H, Ishino K, Nishino M, Rio J, Ewels C, Shinohara H. 2018 Isolation and structure determination of missing fullerenes Gd@C60(CF3)n through in-situ trifluoromethylation. R. Soc. open sci. Submitted.