Collective dynamics of Ca atoms encapsulated in C\textsubscript{60} endohedral fullerenes†

Mihaela Cosinsi\textsuperscript{ci,} \textsuperscript{id} Amanda T. Preda, \textsuperscript{id} C. A. Pantis-Simut, \textsuperscript{id} N. Filipoiu, \textsuperscript{id} I. Ghitiu, \textsuperscript{id} M. A. Dulea, \textsuperscript{id} L. Ion, \textsuperscript{id} A. Manolescu and G. A. Nemnes\textsuperscript{id}*

Endohedral C\textsubscript{60} fullerenes with up to four encapsulated Ca atoms were investigated by \textit{ab initio} molecular dynamics simulations (AIMD). The relatively long runs allow us to describe the correlated movement of the Ca atoms inside the fullerene cage. For the systems with one or two Ca atoms a relatively unimpeded rotation was conjectured by earlier nuclear magnetic resonance experiments and supported by previous \textit{ab initio} calculations used to sample the potential energy landscape. Here, by AIMD calculations, we confirm not only the circular motion, but also the correlated movement of the two Ca atoms, which is due to electric dipole interactions on the inner surface of the C\textsubscript{60} molecule. Furthermore, systems with three and four Ca atoms present highly symmetric configurations of the embedded atoms, which are shown to rotate consistently within the fullerene cage, while more complex charge density patterns emerge. Employing artificial neural network models we perform a force-field mapping, which enables us to reproduce the main characteristics of the actual dynamics, such as the circular motion and the correlated movement of the Ca atoms.

1 Introduction

Over the last few decades the discovery of hundreds of carbon allotropes led to significant advances in nanotechnology, broadening the spectrum of applications from electronics and optoelectronics to electro-catalysis, bio-sensing, water treatment and many more.\textsuperscript{1–4} A particular focus is on low dimensional carbon structures like graphynes,\textsuperscript{5} carbon nanowalls,\textsuperscript{6} nanotubes\textsuperscript{7} and fullerenes.\textsuperscript{8–10} Amongst the quasi-zero dimensional systems, modified fullerenes known as endohedral\textsuperscript{11,12} and exohedral fullerenes\textsuperscript{13} have shown distinguished physicochemical properties compared to the corresponding pristine systems.

Endohedral metallofullerenes (EMFs) have a quite long history, starting with the first encapsulation of La atoms,\textsuperscript{14–16} followed by other rare earth metals for magnetic and optical applications.\textsuperscript{11,12} Later, applications were found for uranium encapsulation,\textsuperscript{17} electronic switching by making use of the induced polarization,\textsuperscript{18} while inserting lighter atoms like lithium enables multi-state molecular switching.\textsuperscript{19} On the other hand, endohedral fullerenes can also embed non-metallic noble gas atoms or small molecules and provide important insights concerning the quantised energy level structure of the trapped species using spectroscopic analysis,\textsuperscript{20} while recent studies indicate the possibility to design of binary endohedral fullerene lattices.\textsuperscript{21}

A significant attention has been devoted to fullerene systems embedding Ca atoms. Endohedral Ca@C\textsubscript{60} fullerenes were extracted using pyridine and aniline solvents\textsuperscript{22–24} and the electronic properties were investigated by ultraviolet and X-ray photoelectron spectroscopy. In another study, which employed nuclear magnetic resonance (NMR), the possibility of a site-hopping motion of a Ca atom inside the cage was indicated.\textsuperscript{25} Lattice type systems of Ca\textsubscript{2}C\textsubscript{60} compounds were also synthesized exhibiting superconductive behavior.\textsuperscript{26–27} Theoretical studies involving alkali or other lighter elements like Li,\textsuperscript{28} Be\textsuperscript{29} and Ca\textsuperscript{30,31} were undertaken, describing chemical stability and electronic properties of the fullerene systems. Polarisation charge switching was inferred from the motion of Ca atoms trapped in fullerene cages.\textsuperscript{32} Such motion has been also investigated for one La atom inside the C\textsubscript{74} cage using relativistic density functional theory.\textsuperscript{33} Further studies included more than one encapsulated atoms, like Sm and Gd ions\textsuperscript{34} or La.\textsuperscript{35} Also, other fullerene based
systems were investigated, like the dimetallofullerene (M$_2$@C$_{100}$) or carbide cluster fullerene M$_2$C$_2$@C$_{98}$ (M = La, Y, and Sc)$_{36}$ Ce$_2$@C$_{90}^{37}$ and bimetallic endohedral fullerenes, PrSc@C$_{60}$. The question of whether the two encapsulated atoms inside the fullerene cage are moving is critical for designing tailored electronic properties. The first evidence for a circular motion of metal atoms at room temperature in La$_2$@C$_{80}$ endohedral dimetallofullerenes was inferred from NMR analysis, where the internal barriers were thought to be small enough. Later, the possibility to control the circular motion of the two La atoms by exohedral addition was investigated. The rapid rotation of the trapped species was also confirmed by NMR in a bimetallic endohedral fullerene, PrSc@C$_{60}$, as well as in Ce$_2$@C$_{90}$ and Lu$_2$@C$_{76}$ ($T_d$) systems. Furthermore, using ab initio methods the transition states and the inferred trajectory were described for a Lu pair inside a C$_{76}$ ($T_d$) cage.

Molecular dynamics studies mostly focused on the structure and dynamics of the fullerene systems confining noble gases and small molecules, showing the possibility to chemically shield water and fluoroacid, e.g., by encapsulation inside a C$_{60}$ cage. Also the changes in the mechanical properties induced by compressive stress were investigated in M@C$_{60}$ (M = Si, Ge) fullerene molecules. Besides, molecular dynamics studies concerning fullerene systems revealed important aspects for nanoscience applications, like impact shields, polarization effects due to ionic liquid interactions or fullerene–substrate interaction for nano-machine applications.

The aforementioned studies conjectured the rapid motion of the embedded species, which could rotate freely inside the carbon cages at room temperature. This is argued by the averaged motion in NMR experiments and by mapping the transition states and possible trajectories in ab initio calculations, in some cases making use of symmetries found in idealized static cages. In this paper, we perform extensive ab initio molecular dynamics (AIMD) simulations, which clearly evidence a correlated movement of up to four calcium atoms inside a C$_{60}$ fullerene. Our results capture the microscopic mechanism, which relies on the interacting dipoles on the inner surface of the cage. We further show that the force field can be mapped on an artificial neural network and reasonable predictions can be achieved for previously unseen dynamics.

2 Structures and methods

The investigated structures are C$_{60}$ fullerenes with up to $N_{Ca} = 4$ embedded Ca atoms, which we label as Ca$_i$@C$_{60}$, with 1 ≤ $i$ ≤ 4, as depicted in Fig. 1. The Ca atoms are initially placed inside the cage at sites given by the coordinates $r_0 = (x,y,0)$, expressed in Å units: (1.5,0,0), (−1.5,0,0), (0,1.5,0), (0,−1.5,0). These coordinates correspond to the corners of a square with the edge length of $1.5\sqrt{2}$ Å. The center of the mass of the fullerene is placed in the origin of the coordinate system.

The molecular dynamics was simulated using the ORCA package. The structural optimizations and the AIMD simulations were performed in the framework of density functional theory (DFT) using B3LYP hybrid functional and the 6-31G* basis. The trajectories were propagated for 4.75 ps with a timestep of 0.5 fs at 300 K, using a Berendsen thermostat with a time constant of 10 fs. Subsequent runs were performed for each system for another 4 ps using the computationally less-expensive GGA–PBE functional. For each atom and time step we collected the positions, velocities and forces. In addition, we performed structural optimizations of the endohedral fullerene systems using GGA–PBE functional with a double-$\zeta$ polarized (DZP) basis implemented in the SIESTA package, which makes use of numeric atomic orbitals. In this approach, norm-conserving Troullier–Martins pseudopotentials were employed and the real space grid was determined by a mesh cutoff parameter of 1.50 Ry. The convergence criterium for the density matrix convergence was set to $4 \times 10^{-4}$.

The potential correlations in the trajectories of the embedded atoms ($N_{Ca} \geq 2$) were monitored by looking at the angle, $x_{ij}$, defined by the center of the mass and the positions of Ca atoms $i$ and $j$. This is determined from $\cos(x_{ij}) = \frac{r_i \cdot r_j}{||r_i|| \cdot ||r_j||}$.

As the computational burden of the AIMD simulations is relatively high, we employ artificial neural networks (ANNs) in order to predict the subsequent dynamics, which follows after the time interval for which the reference dynamics has been established. To this end, we map the force field experienced by the Ca atoms inside the C$_{60}$ cage, providing a direct relation between their coordinates and force components, $\{F_x,F_y,F_z\}_i \rightarrow \{r_i,\theta,\phi\}_i$, with $i \leq 4$. Then, using the force-field mapping, a predicted dynamics is simulated and compared with the actual one, focusing on reproducing the overall features of the trajectories, like the correlated movement of the Ca atoms.

The ANN architecture comprises 4 dense hidden layers, with 128, 1024, 512, 64 neurons. For all neurons, we employed the ReLU activation function, except for the last layer where tanh was employed, while a dropout of 0.3 was applied to each of the hidden layers. The input layer contains $3 \times N_{Ca}$ entries,
corresponding to the spherical coordinates of the embedded atoms. The output layer contains an equal number of neurons as the input, reflecting the set of force components in cartesian coordinates. Rather than predicting the force components individually, we perform a multi-target regression, as the correlations between the different components are expected to improve the overall accuracy.

Although quite extended as time interval, the AIMD simulations cover only a relatively small fraction of the configurational space. The embedded Ca atoms do not cover the entire sphere as they rotate. This could lead to erroneous predictions if the generated trajectory reaches a space domain which is far from similar sets of coordinates such as those obtained from the AIMD simulations. Therefore, we use data augmentation in order to have a more complete covering of the configurational space. We generate a number of 1 886 500 configurations by simultaneously applying rotations of an ideally rigid Ca subsystem. This set results by considering 5500 AIMD steps multiplied to $7^3$ rotations around the Cartesian axes, as the angular grid is set by step size of $\pi/6$. Further, we obtain a reduced set for training and validation by sampling 1/50 of the total set. The test set contains distinct 2000 configurations from the AIMD simulations.

After the force field mapping is obtained, the predicted dynamics is generated. The new trajectories are found by the integration of the equations of motion, using the force field from the ANN model. In this case, we used a simplified thermostat based on velocity scaling. The aim is to compare whether the most important features observed in the reference dynamics are recovered in the predicted one. In particular, observing the correlated movement of Ca atoms is one target. To this end, we compare the $\alpha$ and $\beta$ angles and the mean rotation radius in real vs. predicted MD.

3 Results and discussion

3.1 AIMD-based dynamics

The relatively long AIMD runs (4.75 ps) enable a detailed description of the endohedral fullerene dynamics, with a particular focus on the movement of the Ca atoms inside the C_{60} cage. Fig. 2 shows the trajectories of the Ca atoms for each of the four systems, covering domains shaped as rather thin spherical shells. For Ca_{2}@C_{60}, a number of complete rotations can be observed, with the rotation plane varying up to $\pi/2$ from the initial orientation. For $N_{Ca} \geq 2$, the color patches, representing trajectories, have little overlap suggesting that the Ca atoms tend to fill rather equally the space inside the fullerene. Although the movement of the Ca atoms is reduced as the their number increases, one still finds significant variations in the polar and azimuthal angles, $\theta$ and $\varphi$.

The dynamics is described in more detail in Fig. 3, showing the trajectories in spherical coordinates. The motion of the Ca atom in Ca_{1}@C_{60} system shows a circular motion with a roughly constant radius $R_1 = 0.97 \text{Å}$, around the fullerenes’ center of mass (CM). Inside the fullerene cage a sombrero-like potential was found. As the initial radial position of the Ca atom at $r_0 = (1.5,0,0) \text{Å}$ is relatively far from the equilibrium position, oscillations become visible in the first 2000 steps (1 ps). These weaken by the action of the thermostat and, subsequently, the Ca atom has only small oscillations around $R_1$. The time evolutions of $\theta$ and $\varphi$ angles indicate complete rotations around CM.

For the Ca_{2}@C_{60} system we observe slower dynamics of the Ca atoms. This results from the charge transfer between the Ca atoms and the fullerene cage. The metal atoms become
positively charged, while the opposite charge is localized on the fullerene, in their proximity. In an oversimplified view, the interaction between the two dipoles, which move on the inner surface of the fullerene cage, induces a correlated movement, to a good approximation, on diametrically opposite points. This can be seen from the behavior of the $\theta$ angle, which shows a clear anticorrelation between the two Ca atoms, while the $\varphi$ angles are correlated, indicating that both atoms rotate in the same way around the $z$-axis. Furthermore, the equilibrium distance is increased compared to Ca$_2$@C$_{60}$ at $R_2 = 1.41$ Å. This trend is seen for Ca$_3$@C$_{60}$ and Ca$_4$@C$_{60}$ systems, which show that the average radius increases to $R_3 = 1.6$ Å and $R_4 = 1.71$ Å respectively. The $\theta$ and $\varphi$ angles of different pairs of Ca atoms show distinct regions of correlation and anticorrelations, suggesting a collective behavior of the embedded species.

In order to check the correlated moves of the Ca atoms inside the C$_{60}$ cage, we monitored the angle $\varphi_j$ with the vertex in the CM of the fullerene and its sides defined by the positions of two Ca atoms. Fig. 4 describes the evolution of the $\varphi_j$ angle versus time-step in the AIMD simulation. For the Ca$_2$@C$_{60}$ system, we use a different approach to calculate the time-averaged angle between the two Ca atoms, as it gets close to 180°. Rather than using the standard definition of the dot product we first average the position vectors of the two Ca atoms and then we determine the angle. In this way, we obtain the time-averaged $\langle \varphi_j(t) \rangle = 179.3^\circ$, which proves that indeed the two Ca atoms move on almost diametrically opposite points. Taking into account Fig. 2 and 3 one can see that the two Ca atoms have a considerable movement, which is correlated, with rather small fluctuations of the radial coordinates. Furthermore, for Ca$_3$@C$_{60}$ system, all three angles, $\alpha_{12}$, $\alpha_{13}$, $\alpha_{23}$, present a very similar behavior and their average, $\langle \varphi_j(t) \rangle = 119.93^\circ$, is close to the ideal value of 120°, corresponding to angles formed by the medians in an equilateral triangle. In this case, large oscillations are visible within the first 3000 steps, as the system is equilibrated. Here, in contrast to Ca$_2$@C$_{60}$ where the initial angular condition is matching the subsequent dynamics, the three Ca atoms are further away from equilibrium positions. A similar process takes place in the case of Ca$_4$@C$_{60}$, where the initial planar configuration of Ca atoms, corresponding to the four corners of a square, is distorted towards a regular tetrahedron. This shape is maintained in the subsequent MD simulation, with the tetrahedron rotating inside the fullerene cage. The resulting average value of the $\varphi_j$ angles is $\langle \varphi_j(t) \rangle = 109.69^\circ$, very close to the ideal value of $109.5^\circ$ in a regular tetrahedron. The Ca clusters have highly symmetric configurations, $D_2$ for $N_{Ca} = 3$ and $T_d$ for $N_{Ca} = 4$, while the entire systems, which include the carbon atoms, do not have identifiable symmetries. This is due to the deformations of the C$_{60}$ cages as seen from the MD runs and structural relaxations. One exception is Ca$_4$@C$_{60}$ which is consistent with $C_{3v}$ similar to the Li@C$_{60}$ system.$^{54}$

Overall, the average radius of the fullerene cage, as well as the radius of the rotating Ca atoms, increase with the number of embedded atoms, as depicted in Fig. 5. The pristine fullerene average radius of 3.55 Å is gradually increased to 3.56 Å, 3.61 Å and 3.64 Å, for the four considered systems. Correspondingly, the spherical volumes increase in the sequence 189.89 Å$^3$, 192.87 Å$^3$, 197.38 Å$^3$, 203.10 Å$^3$. The calculation of the convex hull volumes yields a similar trend: 166.27 Å$^3$, 169.14 Å$^3$, 173.21 Å$^3$, 178.46 Å$^3$. Their evolution in time is shown in Fig. S1 in the ESI.$^1$ The calculated volumes of pristine C$_{60}$ are 162.81 Å$^3$ (B3LYP/6-31G*) and 163.38 Å$^3$ (PBE/DZP), which are similar to the reported values in ref. 55.

Inserting multiple Ca atoms in the C$_{60}$ is less likely as the number of Ca atoms gets larger and this can affect the structural integrity of the endohedral fullerene. Although the systems present stability for the total simulated timescale of 8.75 ps (see MD files in the ESI), which is larger than the equilibration time, under strong thermal fluctuations or other external perturbations, the Ca atoms could be expelled from the cage. To check this in more detail, we determined the energy obtained by expelling one Ca atom from Ca$_4$@C$_{60}$. This was found to be $\sim 6.48$ eV, which is smaller than 7.4 eV, the binding energy of a carbon atom in the C$_{60}$ cage.$^{56}$ There are also other

![Fig. 4](image-url) Correlated movement of the Ca atoms evidenced by the angles $\alpha_j$ formed by the atom positions, with the vertex at the center of the mass. The average values of 179.3°, 119.9° and 109.7°, for $N_{Ca} = 2$, $N_{Ca} = 3$, respectively, correspond within a good accuracy to the diametrically opposite points in the circular motion ($N_{Ca} = 2$), an equilateral triangle ($N_{Ca} = 3$) and a regular tetrahedron ($N_{Ca} = 4$). For $N_{Ca} = 2$, the average of the position vectors for the two atoms was employed (*). The equilibration is achieved after $\sim 3500$ steps for Ca$_3$@C$_{60}$ and $\sim 5000$ steps for Ca$_4$@C$_{60}$ systems, as indicated by arrows.
indicators that the system will remain stable. The volume of \( \sim 163 \text{ Å}^3 \) of the C\(_{60}\) fullerene is comparable with the volume spanned by 4 Ca atoms in the fcc lattice, with lattice constant \( a = 5.59 \text{ Å} \), i.e. 174.67 \( \text{Å}^3 \). The radius increase of the C\(_{60}\) cage is only 2.8% compared to Ca\(_{4}\)@C\(_{60}\). In addition, we note that similar or even larger structures have been embedded in fullerenes,\(^{57}\) like Sc\(_2\)C\(_2\),\(^{58}\) Sc\(_3\)N, Er\(_3\)Sc\(_3\)–N\(_{59}\) and Sc\(_2\)O\(_2\) in C\(_{80}\),\(^{60}\) while using first-principle methods, large deformations of the fullerene cages have been investigated.\(^{45}\) We performed further investigations concerning the conditions that would allow the fullerene cage breaking and the escape of the Ca atoms. This is indeed observed by microcanonical ensemble simulations considering significantly high initial kinetic energies of the embedded atoms. Increasing the radial velocities of the Ca atoms, for kinetic energies exceeding 15 eV, we observe them exiting the cage. Part of this energy is transferred to the nearby carbon atoms, which depart from their equilibrium locations and drift in the radial directions. For smaller energies, the Ca atoms bounce back from the highly deformed fullerene cage. This dynamics is captured for several initial kinetic energies of the Ca atoms in the Ca\(_4\)@C\(_{60}\) system, as snapshot images in Fig. S2 of the ESL.\(^\dagger\)

The oscillations with an average frequency of 15.62 THz obtained by fast Fourier transform, correspond to the breathing mode of C\(_{60}\) for all four systems. Oscillation frequencies of 14.36 THz and 14.9 THz by theoretical approaches and experimental data as indicated in Table 3 of ref. 61. There is a quite small variation of the vibration frequency in all four systems. In contrast, a noticeable variation of the frequency of the radial oscillations of Ca atoms is observed, pointing out an increase: 6.1 THz (N\(_{\text{Ca}}\) = 1), 12.7 THz (N\(_{\text{Ca}}\) = 2), 13.4 THz (N\(_{\text{Ca}}\) = 3), 13.2 THz (N\(_{\text{Ca}}\) = 4). At the same time, the amplitudes of the oscillations tend to get lower with the number of Ca atoms.

The collective movement of the Ca atoms can be understood from the charge distribution in the fullerene cage, illustrated in Fig. 6. We determine the charge difference \( \Delta \rho(\vec{r}) = \rho_{\text{sys}}(\vec{r}) - \rho_{\text{Ca}i}(\vec{r}) - \rho_{\text{Ca}j}(\vec{r}) \), where \( \rho_{\text{sys}} \) is the charge density of the endohedral fullerene, while \( \rho_{\text{Ca}i} \) and \( \rho_{\text{Ca}j} \) are the charge densities associated with the fullerene cage and Ca atoms, respectively. The \( \delta^{\text{Ca}} \) charges are determined with respect to undeformed charge distributions of the Ca atoms and the fullerene cage, respectively, considering a sphere of radius \( R_0 = 2.3 \text{ Å} \), which includes the Ca atoms. First, in the case of Ca\(_1\)@C\(_{60}\) system, a charge transfer occurs from the Ca atom, which becomes positively charged \( \delta^+ = 1.18 e \), to the fullerene cage, which receives the corresponding negative charge \( \delta^- = -\delta^+ \), distributed in the proximity of the Ca atom.

A dipole is formed, which slides on the inner surface of the fullerene. When two Ca atoms are present, in Ca\(_2\)@C\(_{60}\) system, the transferred charge is \( \delta^+ = 0.51 e \) and the dipoles repel each other, resulting in the localization of the Ca atoms on the diameter, a condition which is kept to a large extent during their thermal motions. For Ca\(_3\)@C\(_{60}\) a three-fold symmetry of the charge distribution is observed. The positive charge is shared between pairs of Ca atoms and three patches of negative charge are observed on C\(_{60}\) cage, near the Ca atoms. In addition, a region of confined negative charge is visible in the middle of the fullerene cage. Based on this symmetry, one can infer the correlated movement of the three Ca atoms, which can be found during the MD simulation, up to some thermal fluctuations, at the corners of an equilateral triangle. The Ca\(_4\)@C\(_{60}\) system starts...
from a planar configuration, with the Ca atoms occupying the corners of a square, and it evolves to a good approximation to a tetrahedral configuration. The previous observations concerning the charge transfer are also found in this case. The positive charge resides in the proximity of the four Ca atoms, now reflecting the tetrahedral symmetry. The negative charge is found on the fullerene cage and it is also localized in the middle of the Ca based tetrahedron. These elements suggest that the rotation of the Ca structure occurs as a whole inside the C_60 fullerene. In spite of this simple view, in all four systems, the charge distribution is rigorously determined by the ab initio quantum mechanical calculations. Although in Fig. 6 snapshots are presented, qualitatively, these images and conclusions persist during the entire AIMD simulations.

3.2 Predicted dynamics

Assessing the correlated dynamics of the Ca atoms requires rather long simulation times, which is computationally expensive in the AIMD framework, particularly if the number of embedded atoms increases. Using ANNs, we perform a mapping of the force field for describing the movement of the Ca atoms. This would enable more efficient MD simulations, as long as the essential features can be reproduced. Since the equilibration time increases with $N_{Ca}$ up to $\sim 2.5$ ps, which

![Fig. 7](https://example.com/fig7.png)

**Fig. 7** Reference force components ($F_{x1}, F_{y1}, F_{z1}$) and predicted force components ($\tilde{F}_{x1}, \tilde{F}_{y1}, \tilde{F}_{z1}$) for (a) Ca_1@C_60 and (b) Ca_2@C_60 systems. The evaluation takes place for the real trajectory, corresponding to the last 1 ps.
becomes a consistent fraction of the entire simulation, we focus on predicting the dynamical features of Ca$_2@C_{60}$ and Ca$_3@C_{60}$ systems.

As the initial configurations of the AIMD simulations are relatively far from equilibrium, we discard the first 2000 steps. In this way, we exclude the initial part of the AIMD simulations, where large oscillations are present, and focus on the correlated dynamics, which is well established after $\sim 1$ ps. The training data-set is formed by the next 5500 steps, corresponding to 2.75 ps, while the remaining 2000 steps (1 ps) shall be used for testing. The learning curves are shown in Fig. S3 (ESI†).

We first analyze the prediction accuracy for the force components, ($\tilde{F}_x$, $\tilde{F}_y$, $\tilde{F}_z$), which act on the embedded Ca atoms, compared to the reference dynamics, ($F_x$, $F_y$, $F_z$), as shown in Fig. 7. Starting with the Ca$_3@C_{60}$ system, we obtain that the predicted force values are rather close to the reference ones. They are calculated on the same set of coordinates, corresponding to the original trajectory in the last 2.5 ps. During the ANN architecture optimization it was established that indeed the presence of all force components in the output as a multi-target regression has a beneficial impact on the accuracy. For the Ca$_2@C_{60}$ system, the errors are larger, but still reasonable with respect to reproducing the main features of the real dynamics, as the ANN model captures most of the typical oscillatory behavior, although not with full amplitude. For $N_{\text{Ca}} > 2$, as expected, the accuracy is getting lower with the increase in the complexity of the embedded system, as the number of possible atomic configurations inside the cage grows larger. The Ca atoms interact with each other and with the C$_{60}$ cage, while it is deformed, increasing the difficulty for achieving an accurate mapping of the MD. The force field predictions for $N_{\text{Ca}} = 3$ and $N_{\text{Ca}} = 4$ are shown in Fig. S4 and S5 (ESI†), respectively. To obtain a meaningful description, a larger training set would be required and higher accuracies for the force-field predictions.

Next, we check if the correlated movement of the two Ca atoms can be obtained from the predicted dynamics based on the learned force field. The predicted vs. reference trajectories are depicted in Fig. 8(a). Even for highly accurate force fields, a slight perturbation of the trajectory, induced by thermal fluctuations, can render rather different trajectory pairs. However, a close inspection reveals that two trajectories in one pair, stemming from either real or predicted dynamics, are strongly correlated. For the predicted dynamics, the new trajectories are expected to reproduce the main features, like preserving the correlated move; the average was performed on the position vectors ($\mathbf{r}$). (c) Predicted distance between Ca atoms, $d(Ca_1, Ca_2)$. (d) The predicted evolution of the radial coordinates, $R_{1,2}$, of the two Ca atoms, showing anti-correlations.

Fig. 8 (a) Reference and predicted trajectories for Ca$_2@C_{60}$ depicted with brightened and faded colors, respectively. The starting points are marked by yellow dots. (b) Predicted average of $\langle R(t)\rangle_1$ and $\langle R(t)\rangle_2$, showing the correlated move; the average was performed on the position vectors ($\mathbf{r}$). (c) Predicted distance between Ca atoms, $d(Ca_1, Ca_2)$. (d) The predicted evolution of the radial coordinates, $R_{1,2}$, of the two Ca atoms, showing anti-correlations.

Finally, we would like to discuss the utility and possible extensions of our ML based approach. Long MD runs can provide an extensive sampling of the configurational space, which, in turn, can be used to determine some targeted electronic and mechanical properties. However, this is far from sampling the configurational space exhaustively. Different initial conditions for multiple embedded atoms may yield rather similar dynamics, which can be predicted by interpolating the previously seen dynamics [Fig. S6 and S7, ESI†]. Once the ML model is trained, it can be used as a base model and additional small perturbations can be applied [Fig. S8–S11, ESI†]. In linear regime, the extended model is expected to provide accurate descriptions for various perturbation setups, without the burden of new AIMD simulations. Therefore, using ML procedures one can derive computationally compact and efficient base-models, which can be further tuned for specific applications. This approach may gain ground as more and more ML models become available for various systems.

4 Conclusions

The dynamical behavior of Ca atoms embedded in C$_{60}$ fullerences was investigated using AIMD simulations. Our results confirm the rotation of the embedded atoms, which was
conjectured in previous studies for Ca\textsubscript{1}@C\textsubscript{60} and Ca\textsubscript{2}@C\textsubscript{60} systems and further evidence their correlated movement inside the carbon cage. The relatively long simulation scales enable an accurate evaluation of the dynamical features, in particular the collective movement of the Ca atoms. Following the relaxation from pre-set initial configurations, the embedded atoms maintain highly symmetric positions, while these structures can rotate within the C\textsubscript{60} cages. To support these findings a quantitative analysis was performed using geometrical parameters suitable to assess the correlated dynamics. The charge transfer analysis reveals the appearance of electrical dipoles sliding on the inner surface for the Ca\textsubscript{2}@C\textsubscript{60} system, while more complex but symmetric charge distributions emerge for \textit{N}\textsubscript{Ca} = 3 and 4. Increasing the number of embedded species in the fullerene cage is however less likely. Furthermore, we developed ANN based models to map the force field exhibited by the Ca atoms, which allows a significantly more efficient description of the dynamical behaviour. This is most accurately described for \textit{N}\textsubscript{Ca} = 1 and 2, quite close to the real dynamics. Thus, we provide a direct observation of the collective movement of the embedded species and provide a consistent ML based description, which can efficiently describe the dynamics of the endohedral fullerene, in particular for new initial conditions and for additional small perturbing fields. Identifying complex dynamical patterns, which can be exploited in molecular electronics applications, remains a current goal and this can be achieved by employing efficient numerical schemes based on AIMD and ML simulations.

Author contributions
M. C., A. T. P., N. F. and G. A. N. performed the \textit{ab initio} calculations and contributed to the writing of the theoretical part. C.-A. P.-S. and I. G. contributed to writing, review and editing. M. A. D., L. I., A. M contributed to review and editing. M. C. and G. A. N performed the numerical analyses and contributed to the writing, review and editing.

Data availability
The data supporting this article have been included as part of the Main Article and ESI.†

Conflicts of interest
There are no conflicts to declare.

Acknowledgements
This work was supported by a grant of the Romanian Ministry of Research, Innovation and Digitalization, CNCS – UEFISCDI, project number PN-III-P4-ID-PCE-2020-1142, within PNCDI III and PN 23210204.

Notes and references