NanoImpact 38 (2025) 100563

Contents lists available at ScienceDirect

NanoImpact

journal homepage: www.elsevier.com/locate/nanoimpact

Research paper

In silico exploration of graphene nanoflakes: From DFT simulations to machine learning-driven toxicity predictions

Nuria Aguilar [a], Patricia de la Fuente [b], Natalia Fernández-Pampín [b], Sonia Martel [b], Laura Gómez-Cuadrado [b], Pedro Angel Marcos [b,c], Alfredo Bol [b,c], Carlos Rumbo [b], Santiago Aparicio [a,c,*]

[a] Department of Chemistry, University of Burgos, 09001 Burgos, Spain
[b] International Research Center in Critical Raw Materials for Advanced Industrial Technologies (ICCRAM), University of Burgos, 09001 Burgos, Spain
[c] Department of Physics, University of Burgos, 09001 Burgos, Spain

ARTICLE INFO

Editor: Bernd Nowack

Keywords:
Graphene nanoflakes
Density Functional Theory (DFT)
In silico toxicity
Machine learning
Nano-bio interactions

ABSTRACT

The present theoretical work provides a ground-breaking and comprehensive study of graphene nanoflakes integrating Density Functional Theory (DFT) simulations, toxicity predictions and a machine learning approach. The properties of graphene nanoflakes as a function of size, shape, and symmetry are systematically analysed using DFT calculations. The interaction of these nanoflakes with human proteins and cell membranes, considered as Molecular Initiating Events for diverse Adverse Outcome Pathways, is explored to infer potential toxicity effects. Leveraging the generated data, machine learning models were developed to predict flake properties and biological interactions. A single score representing the biological interaction or impact of graphene nanoflakes on both proteins and plasma membranes is assigned to each evaluated nanoflake to infer its potential toxicity. Our multiscale approach bring valuable insights into the structure-property-toxicity relationships of graphene nanoflakes, paving the way for their safe and efficient design and application.

1. Introduction

Graphene, a two-dimensional carbon allotrope with a honeycomb lattice structure, has captivated the scientific community since its isolation in 2004 (Novoselov et al., 2004). Its exceptional properties, including high electrical and thermal conductivity, mechanical strength, and large surface area, have positioned graphene as a promising material for various applications ranging from electronics to biomedical devices (Avouris and Dimitrakopoulos, 2012; Seabra et al., 2014). Graphene nanoflakes (GNFs), finite-sized graphene sheets typically less than 100 nm in lateral dimensions, have emerged as particularly interesting structures due to their unique size-dependent properties and enhanced edge effects (Son et al., 2006; Yazyev, 2010).The properties of GNFs are heavily influenced by their size, shape, and edge structure, which can significantly affect their electronic, magnetic, and chemical characteristics (Güçlü et al., 2014). Understanding these structure-property relationships is crucial for tailoring GNFs for specific applications. Density Functional Theory (DFT) simulations have proven to be a powerful tool in elucidating the fundamental properties of GNFs (Hod et al., 2008; Ezawa, 2007). Recent studies have employed DFT to investigate the electronic structure of GNFs with various edge configurations, revealing the emergence of localized edge states and magnetism in certain geometries (Zou et al., 2011; Lee et al., 2005). The size-dependent properties of GNFs have been explored through DFT calculations, demonstrating how the energy gap and magnetic moment evolve with increasing flake dimensions (Fernández-Rossier and Palacios, 2007). Triangular GNFs exhibit size-dependent magnetic properties, with the total spin increasing linearly with the edge length (Sharma et al., 2014). Similarly, hexagonal GNFs have been found to display size-dependent electronic properties, transitioning from semiconducting to metallic behavior as the flake size increases (Zarenia et al., 2010).

While the unique properties of GNFs offer exciting possibilities for applications, concerns about their potential toxicity and environmental impact have arisen (Bianco, 2013; Ou et al., 2016). The high surface area-to-volume ratio and reactive edges of GNFs can lead to enhanced interactions with biological systems, potentially causing adverse effects (Chng and Pumera, 2013). *In vitro* and *in vivo* studies have reported varying degrees of toxicity for graphene-based materials, depending on

* Corresponding author.
E-mail address: sapar@ubu.es (S. Aparicio).

https://doi.org/10.1016/j.impact.2025.100563
Received 30 January 2025; Received in revised form 24 April 2025; Accepted 24 April 2025
Available online 28 April 2025

their physicochemical properties and the biological model used (Zhang et al., 2011; Schinwald et al., 2012). To address these concerns and facilitate the safe development of GNF-based technologies, there is a growing need for robust methods to predict the biological interactions and potential toxicity of GNFs. *In silico* approaches, particularly molecular dynamics (MD) simulations and quantitative structure-activity relationship (QSAR) models, have emerged as valuable tools in this regard (Burello and Worth, 2011; Shao et al., 2019). MD simulations have been used to study the interactions between graphene sheets and lipid bilayers, providing insights into the membrane penetration mechanisms and potential cellular uptake of GNFs (Tu et al., 2013; Guo et al., 2013). Recent studies have also explored the interactions between GNFs and proteins, shedding light on potential biological effects and biomedical applications. For instance, MD simulations were considered to investigate the adsorption of various proteins on graphene sheet surfaces, revealing the important role of π-π stacking interactions and hydrophobic effects (Gu et al., 2015). Similarly, MD simulations revealed the conformational changes of proteins upon adsorption to nanographene, providing valuable insights into the potential impacts on protein function and cellular processes (Zuo et al., 2013).

The complexity of GNF-bio interactions and the vast parameter space of possible GNF structures make it challenging to rely solely on experimental or traditional computational methods for comprehensive toxicity assessments. Machine learning (ML) approaches have shown great promise in addressing this challenge by enabling rapid and accurate predictions of material properties and biological interactions (Cai et al., 2018; Ramprasad et al., 2017). Recent studies have demonstrated the power of Machine Learning (ML) in predicting the properties of graphene-based materials, including electronic structure (Chen and Gu, 2019), mechanical properties (Gu et al., 2018), and thermal conductivity (Wei and Yao, 2019). Particularly the use of combined DFT – ML have successfully considered for the development and design of new materials such as catalysts (Fu et al., 2024), highlighting the potential of both approaches for understanding material properties and optimizing catalytic performance.

In the context of toxicity prediction, ML models have been developed to assess the potential hazards of various nanomaterials, including carbon-based structures (Winkler et al., 2013; Choi et al., 2018). Nano-QSAR models using machine learning algorithms have been developed to predict the biological activity of diverse nanomaterials, including graphene derivatives (Fourches et al., 2010). ML techniques are applied to predict the cytotoxicity of carbon nanotubes based on their physicochemical properties (Choi et al., 2019), demonstrating the potential of this approach for other carbon nanomaterials like GNFs. The integration of DFT simulations, MD-based toxicity predictions, and ML approaches offers a powerful framework for comprehensively studying GNFs and their potential impacts. This multiscale modeling approach can provide insights across multiple length and time scales, from electronic structure to cellular interactions (Barnard, 2016; Puzyn et al., 2011). By combining these methods, researchers can develop more accurate and efficient tools for predicting GNF properties and biological effects, ultimately guiding the design of safer and more effective graphene-based materials (Yan et al., 2011; Mu et al., 2012). Despite these advances, several challenges remain in the *in silico* study of GNFs. These include the accurate representation of edge effects and defects in DFT simulations (Cocchi et al., 2010), the development of force fields that can capture the unique properties of GNFs in MD simulations (Rajasekaran et al., 2016), and the generation of large, diverse, and high-quality datasets for training ML models (Butler et al., 2018).

In this study, a comprehensive *in silico* investigation of GNFs is presented, integrating DFT simulations, toxicity predictions and machine learning approaches, Fig. 1a. This work aims to: *i*) systematically analyze the properties of GNFs as a function of size, shape, and symmetry using DFT calculations; *ii*) explore the interactions of GNFs with human proteins and cell membranes to predict potential toxicity; *iii*) develop ML-based models for the rapid and accurate prediction of GNF properties and biological interactions and *iv)* establish a single score using the predicted biological interaction (or impact) of graphene nanoflakes on both proteins and plasma membranes to infer its potential toxicity. By addressing these objectives, we seek to advance the

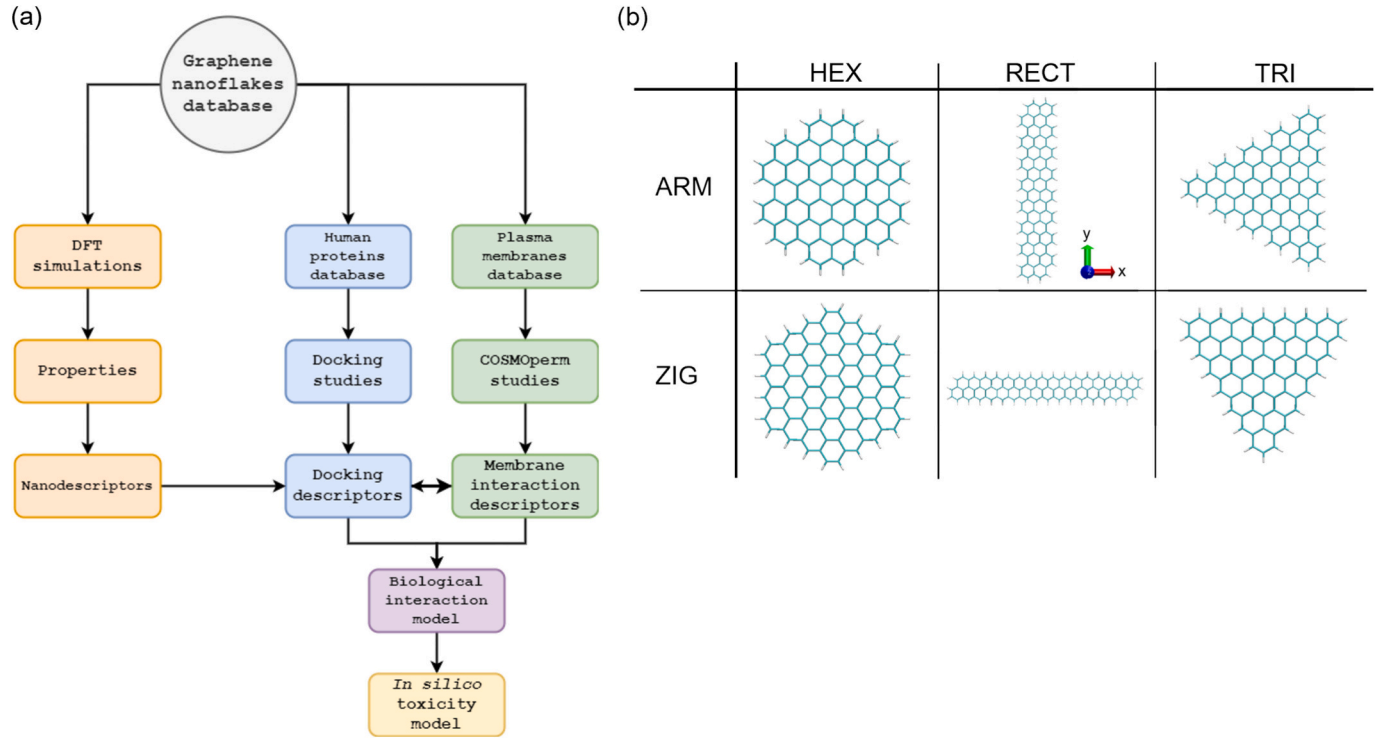

Fig. 1. (a) Scheme of the in silico approach reported and (b) scheme of the three different shapes – hexagonal (HEX), rectangular (RECT) and triangular (TRI) – and two types of edges – armchair (ARM) and zigzag (ZIG) – of graphene flakes considered in this work.

understanding of structure-property-toxicity relationships in GNFs and provide valuable tools for the rational design and safe application of these promising nanomaterials.

2. Methods

2.1. Nanoflakes models

The objective of the work is to analyze the effect of GNFs properties (size, shape and edges; Fig. 1b) on their interaction with biological toxicity targets. For this purpose, GNFs have been selected from the Graphene Nanoflake Data Set published by Barnard et al. (Barnard et al., 2016). This Data Set contains a total of 650 final graphene nanoflake structures, including GNFs with different sizes (from 16 to 2176 carbon atoms), shapes (hexagonal, rectangular and triangular) and edges (armchair, *ac*, and zigzag, *zz*). From this Data Set, a total of 98 GNFs have been selected, being the prime the criterion that the maximum number of carbon atoms for a graphene nanoflake should not exceed 300, as a DFT geometry optimization would be applied, but including all the relevant structural features to be explored in this work (size, shape and edges). The selected 98 GFNs are listed in Table S1 (Supplementary Information), including the name of the graphene flake, the total number of carbon atoms (*nC*), the total number of hydrogen atoms (*nH*), the shape – HEX for hexagonal, RECT for rectangular and TRI for triangular –, the type of edge – armchair, ARM, and/or zigzag, ZIG – and the size of the nanoflake, in Å, being *L* the largest size and *l* the smallest. The 98 GNFs contains a total of 8 hexagonal GFNs, 3 with armchair (ARM) edge and 5 showing a zigzag (ZIG) edge. A total of 77 GNFs are rectangular, and as this shape presents two edges, the number of graphene rings along ZIG edge is written first, followed by the number of rings along the ARM edge, i.e., RECT_zz2ac2_C16H10: this graphene nanoflake is rectangular (RECT), has two rings along the ZIG edge (zz2) and 2 along the ARM edge (ac2) and has a total of 16 carbon atoms and 10 hydrogen atoms. The remaining 13 GFNs are triangular, 6 of them with ARM edge and 7 with ZIG edge.

2.2. DFT calculations and nanodescriptors

The 98 selected GNFs were optimized in terms of geometry using Density Functional Theory (DFT). The geometry optimization was carried out with ORCA v.4.2.1 software (Neese, 2012) employing ωB97X as functional, def2-SVP as the basis set and def2/J as the auxiliary basis set. An extra step of electronic optimization was taken using def2-TZVP as the basis set, using the optimized structured from def2-SVP. The final relaxed structures were imported into BIOVIA Materials Studio software (Meunier and Robertson, 2021) to obtain the nanodescriptors for all the 98 GNFs. A total of 73 descriptors from 6 different families –Atom Volumes and Surfaces, Atomistic Descriptors, DMol3 Molecular, Forcite Energetics, Spatial Descriptors and VAMP Electrostatics – were considered (Table S2, Supplementary Information). The nanodescriptors are based on quantum (Dmol3), semiempirical (VAMP) and classical (Forcite) methods and predict structural properties, i.e., atom and element count, areas and volumes; electronic properties such as energy, orbitals and dipole moments and solvation properties as energy or cavity volume; therefore, capturing the most relevant features of the nanoflakes. To obtain the Atom Volumes and Surfaces descriptors, the resolution of the calculation was set to ultra-fine, keeping the van der Waals scalar factor as 1.0, the solvent radius as 5.0 and the Connolly radius as 1.0. For Dmol3 descriptors, water was chosen as solvent and the calculations were running using PBE (Perdew et al., 1997) functional, considering that the GNFs are not charged and keeping the multiplicity of the nanoflake as spin restricted. For Forcite Energetics, the force field chosen for the ultra-fine (quality) calculations was COMPASS III. Finally, for VAMP Electrostatics descriptors, water was chosen as solvent while the calculations were running using AM1 Hamiltonian within the spin state restricted Hartree-Fock (RHF) formalism for non-charged GNFs. The atomistic and spatial descriptors were calculated using the default input given by BIOVIA QSAR Plus module.

2.3. Protein docking studies

To predict the interaction of the GNFs with human proteins, a molecular docking study has been conducted considering a total of 100 human proteins and 40 graphene nanoflakes. The 100 human proteins were chosen to represent some of the principal protein families and main targets for developing molecular initiating events (MIE) under the adverse outcome pathway (AOP) framework (Lizano-Fallas et al., 2023). The proteins considered in this work, sorted by family, with their identifying number (Protein ID, PID) are listed in Table S3 (Supplementary Information). The pdb structure for all the proteins were downloaded from the Protein Data Bank website (Berman et al., 2000).

The molecular docking study has been carried out with VINA v.1.2.3 (Trott and Olson, 2010) software. To prepare the proteins as receptors for docking in VINA, a two-step procedure was applied: (*i*) the proteins were cleaned using the PrepareProtein tool from Play Molecule website (Martínez et al., 2017) and (*ii*) the proteins were prepared using MGL Tools v.1.5.7 (Morris et al., 2009), Gasteiger charges were applied, non-polar hydrogens were merged and a simulation box containing the whole protein was defined. The GNFs were prepared as ligands following the (*ii*) step, omitting the construction of a simulation box around the nanoflake. As the non-polar hydrogens are merged, only the carbon atoms from the graphene optimized structures were considered. For running the calculations, the exhaustiveness of the docking was set to 32 and the number of poses considered in the output to 9. The binding affinity reported in the results section was obtained as the mean of the 9 binding affinity energies for the 9 poses for each GNF.

To study the interaction region, MOE software was employed.

2.4. Plasma cell membrane interactions: COSMOperm approach

To predict the interaction of GNFs with model plasma cell membranes, four types of cell membranes – DMPC, DOPC, POPC and SOPC – solvated with water were built, Fig. S3 (Supplementary Information) to applied the cosmomic/cosmoperm approach (Schwöbel et al., 2020) based on COSMO-RS methodology (Klamt et al., 1998). Four model membranes were built composed of dmpc, dopc, popc or sopc lipids. Each membrane contains 64 lipids per leaflet, were saturated with water in a water/lipid ratio of 31. Membranes were previously equilibrated using NPT Molecular Dynamics simulations for 100 ns using CHARMM36 forcefield (Huang and Mckerell, 2013). To obtain the *cosmo* files for the 42 GNFs to run COSMOperm calculations, a DFT optimization at BP86 def-TZVP theoretical level, required for COSMO-RS method, was carried out using Turbomole software (Franzke et al., 2023). The thermodynamics properties – free energy and entropy – and kinetic properties – diffusion and permeability – for the interaction of the 42 GNFs with the four model plasma cell membranes were predicted at a temperature range from 298.15 K to 310.15 K with 1 K step.

2.5. Machine learning development of QSAR models for biological interactions

2.5.1. Dimensionality reduction

The main dataset contained 98 GNF structures and 78 variables (predictors) describing different spatial, atomistic, electronic, and molecular properties of the GNFs. Four of these descriptors are categorical variables: geometry (hexagonal, triangular, and rectangular), two edge configurations (Edge 1 and Edge 2), and Molecular Point Group (VAMP electrostatics). The remaining variables are continuous numerical variables with different scales. To explore the data distribution, reduce dimensionality, and identify the most influential variables in the data variability, a Factor Analysis of Mixed Data (FAMD) was performed. This technique combines elements of Principal Component Analysis (PCA)

and Multiple Correspondence Analysis (MCA) to handle datasets containing both numerical and categorical variables. The results of the FAMD were summarized using a scree plot and bar plots of the loadings, which highlight the most significant variables for each Principal Component (PC). The scree plot was used to visualize the variance explained by each component, while the bar plots of the loadings showed the contribution of each variable to the selected components. These visualizations allowed the identification of key variables that capture the primary variability in the data, making further analysis more efficient and interpretable. After applying FAMD, a reduced dataset was obtained. Correlations between the selected variables were assessed using the autocorrelation function and Spearman's rank correlation coefficient, revealing significant multicollinearity. Since the selected variables from FAMD were all numerical, Principal Component Analysis (PCA) was applied as an additional check to verify the contribution of the selected variables to the overall variance in the numerical data space.

2.5.2. Pre-processing and model optimization

2.5.2.1. Definition of the response variables. Simulating interactions of GNFs with biological targets is computationally intensive, and in this study, obtaining the necessary files for simulating interactions with membranes for all 98 graphene flakes proved problematic. Therefore, to develop the predictive models, the dataset was reduced to a subset of 33 GNFs (Table S4, Supplementary Information), selected to be analysed more efficiently within a reasonable time frame, while avoiding the issues encountered. The subset was carefully chosen to ensure a variety of shapes, edges, and sizes. This diversity is crucial for representing a broader range of possible physical and chemical properties, which enhances the model's ability to generalize when predicting GNFs interactions with biological targets (e.g., cell membranes and proteins). Of the original 98 NGFs, 77 were rectangular (RECT), which skewed the shape diversity. Nonetheless, the final selection of 33 GNFs maintained the original proportion of rectangular flakes, preserving a balance between diversity and practical computational constraints. This subset contains the following data corresponding to 33 observations (GNFs structures, selected from the general dataset containing all the relevant size, shape and edge features): Descriptors data, binding affinity of docking to 100 proteins (kcal mol^{-1}), and different thermodynamic and kinetic properties: Free Energy (kcal mol^{-1}), Entropy (cal K^{-1}), Diffusion (10^{-9} m^2 s^{-1}) and Log Permeability (cm s^{-1}) measured at 310.15 K (37 °C) for the membrane types: DMPC, DOPC, POPC and SOPC, measured at the center of the membrane (*depth* = 0 Å). This reference point has been chosen as the penetration of GNFs through membranes showed no barriers, thus being the probability of finding the graphene nanoflake at the non-polar region the highest. The average of each thermodynamic and kinetic property for each membrane type was calculated. Since the thermodynamic variables have different scales, they were standardized using min-max normalization (Eq. (1), where x is the original data, *min(x)* is the minimum value of the variable's data set, and *max(x)* is the maximum value). This method was chosen because some variables did not follow a normal distribution. However, it is important to note that this technique may mitigate the influence of a few outliers identified in the data distribution (outliers were determined using the interquartile range (IQR) and visualized with boxplot graphics), as these outliers represent real phenomena. After obtaining the average Free energy, average Entropy, average Diffusion, and average Log Permeability for the four membranes and standardized to the same scale (0–1), an overall average was computed to create a single dependent variable representing the biological interaction with cell membranes (named hereafter Cell Membrane Interaction).

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1)$$

A single dependent variable was also created to represent the interaction with proteins by averaging the docking energy data for the 100 proteins. Additionally, this dependent variable related to the proteins was also normalized following Eq. (1) (named hereafter Protein Interaction) to allow for its subsequent combination with the membrane data in the scoring analysis, ensuring consistency in the scale of the dependent variables.

2.5.2.2. Recursive Feature Elimination (RFE). Given the collinearity present in the predictor variable set, a variable selection process was conducted using supervised machine learning methods. In this study, the Random Forest algorithm was utilized through the Recursive Feature Elimination (RFE) technique. This method aims to identify the optimal set of variables for predicting the dependent variable by measuring performance with Root Mean Square Error (RMSE) and ensuring reproducibility. Random Forest was chosen due to its robustness against multicollinearity, atypical values, and overfitting—issues likely given the fewer observations than predictor variables. Additionally, it effectively captures potential nonlinear or complex relationships that may be present in the data. RFE helps mitigate issues related to collinearity by systematically eliminating the least important features, thereby enhancing model interpretability. The RFE was performed on the entire dataset to select relevant features, identifying which of the descriptors selected are truly necessary for the models. This selection process was complemented by bootstrapping to ensure robust estimates of feature importance, generating multiple samples of the dataset to evaluate the stability of the selected features across different subsets. Subsets of features ranging from 3 to 25 were analysed to observe how model performance changes as features are added or removed, with 50 bootstrapping repetitions to ensure sufficient samples for assessing stability.

The RFE procedure was executed with a list of random seeds for feature selection variability, ensuring reproducibility in each RFE repetition. Each repetition had its own set of seeds, which is important for managing variability in the data and obtaining reliable estimates. Another seed was established for the execution of the RFE model to ensure that the model is trained and evaluated consistently each time the code is run, allowing for direct and objective comparison of results. The performance of the model was evaluated using the following metrics: Root Mean Square Error (RMSE), R-square (R^2), Mean Absolute Error (MAE), and their corresponding standard deviations (RMSE SD, Rsquare SD and MAESD). The optimal number of decision trees was determined using a learning curve based on RMSE metrics, identifying the most suitable configuration. A range of tree values was explored (from 10 to 200 trees), and the performance was assessed to find the configuration that minimized the RMSE.

2.5.3. Random Forest predictive models

To assess the interaction between GNFs and proteins and cell membranes, two regression models using Random Forest were developed. Random Forest was chosen due to its robustness against multicollinearity, atypical values, and overfitting—issues likely given the fewer observations than predictor variables. Additionally, it effectively captures potential nonlinear or complex relationships that may be present in the data.

The first predictive model predicts the relationship between Cell Membrane Interaction and the corresponding descriptors selected in the previous RFE step. The second predictive model predicts the relationship between Protein Interaction and the corresponding descriptors selected in the previous RFE step. Both models were trained with the complete dataset of 33 GNFs (Table S4, Supplementary Information) and optimized using grid search for the hyperparameters: the number of variables to be considered at each split, exploring a range of values around the value corresponding to approximately one-third of the total number of independent variables used for each model (Hastie et al., 2009; Breiman, 2001). To ensure reproducibility, random seeds were set for

the feature selection process and the model training and evaluation phases. The number of trees was assessed using a learning curve (from 10 to 200 trees). The robustness of the models was evaluated using repeated cross-validation (with 5 folds and 10 repetitions for both models), with RMSE as the evaluation metric. The best models were selected based on a combination of metrics: RMSE, MAE and $R.^2$ The models were tested using an external dataset comprising seven GNF structures (Table S4, Supplementary Information), along with the corresponding descriptors. Model predictions were evaluated both graphically and quantitatively through the calculation of RMSE (Root Mean Square Error). Additionally, a regression analysis was performed between the predicted and real values, where the R^2 and p-value were calculated to assess the goodness of fit and the statistical significance.

2.6. Score. Assessing the impact of graphene nanoflake interactions with biological targets

The predictions generated by the models during both the training and validation phases are used to assess a score according to Eq. (2). The score is calculated as 0.5 times the predictions for Cell Membrane Interaction plus 0.5 times the predictions for Protein Interaction. This means that each component contributes equally to the final score, with both predictions being weighted at 50 %. The score reflects the level of biological interaction between GNFs and biological targets (proteins and plasma membranes), which are considered Molecular Initiating Events in Adverse Outcome Pathways. Since the variables are normalized, the score ranges from 0 to 1, with lower values indicating a lower biological impact, from which minor toxicity may be inferred.

$$Score = 0.5 \times (pred.membrane) + 0.5 \times (pred.protein) \quad (2)$$

The rationale behind this criterion is that smaller Graphene flakes, and therefore with fewer carbon atoms, have a greater interaction with proteins and membranes (energy per number of atoms, not the total energy value). We assume that a higher score indicates greater interaction and, consequently, a higher potential toxicity.

2.7. Software and packages

The statistical analysis (including FAMD, autocorrelation, and PCA), Random Forest regression models, calculations, and graphical outputs were performed using R (R Core Team, 2024) and RStudio (Posit team, 2023) The Factor Analysis of Mixed Data (FAMD) and the PCA were performed with "FactoMineR" and "factoextra" packages (Le et al., 2008; Kassambara and Mundt, 2020). The autocorrelation function was estimated using "corrplot" package (Wei and Simko, 2024). The Recursive Feature Elimination method and the Random Forest regression models were developed using the "caret", "doMC" and "Metrics" packages (Kuhn, 2008; Analytics and Weston, 2018; Hamner and Frasco, 2018). Data manipulation and graphical visualization were performed using the "tidyverse" suite of packages (Wickham et al., 2019). Reading and manipulation of spreadsheet data were carried out using the "openxlsx" package (Schauberger and Walker, 2024).

3. Results and discussion

3.1. Protein docking: insights from mechanistics and energetic viewpoints

A DFT-level characterization of the graphene nanoflakes is carried out before running the protein docking calculations. The protein docking calculations are classical simulations (*Monte-Carlo (MC) iterated search combined with the BFGS17 gradient-based optimizer*) and the docking performed in this work only considered rigid-rigid molecular docking, meaning that both the protein and the ligand (GNFs) are not being optimized in terms of geometry. The reason to choose this type of molecular docking is to maintain the structure of the GNF after the adsorption in the protein cavity (or protein surface) as the GNF planar and rigid structure is characteristic of the material itself and only the affinity (via the interaction energy) of the GNF with the protein is studied. Because of the rigid-rigid molecular docking calculations, the GNF after adsorption is the same in terms of bond distances and angles (torsions), thus, the vibration frequencies are calculated only for the GNFs before adsorption (docking). The scaling factor recommended for functional and basis ωB97X/def2-TZVP is 0.98358 (Alecu et al., 2010) and the scaling factor automatically applied by ORCA (v. 6.0.1) software when using the aforementioned functional/basis theoretical level is 1.0. Rosi et al. (Rosi et al., 2010) reported the vibrational frequencies and intensities of two GNFs, containing 30 and 32 carbon atoms, calculated using LDA (DFT). The most intense frequencies for both GNFs are 985.95 cm^{-1} (299.51 km/mol) for the model with 30C atoms and 979.48 cm^{-1} (1060.2 km/mol) for the 32C atom model. These frequencies are in good agreement (taking into consideration that those models contain only C atoms while our models contain H atoms and that the level of theory employed in the paper was LDA, not GGA as in this work) with the results reported in Table S5 (Supplementary Information), where the most intense frequencies are: 973.42 cm^{-1} (172.55 km/mol) for HEX_ZIG_C54H18, 901.41 cm^{-1} (105.68 km/mol) for RECT_zz2ac4_C28H014 and 814.93 cm^{-1} (152.79 km/mol) for TRI_ARM_C60H24. Additionally, the Electrostatic potential maps (ESP) for six GNFs are reported in Fig. S1 (Supplementary Information), showing no relevant changes in ESP considering neither the geometry (HEX, RECT, TRI) nor the size of the GNFs.

A first biological effect was analysed via the interaction of the GNFs with the considered human proteins dataset (Table S3, Supplementary Information). The binding affinity of The GNFs with the proteins for the optimal poses was normalized per number of carbon atoms, *nC*, in the GNFs, Table S6 (Supplementary Information), and plotted in Fig. 2a. The reported evolution of binding affinity with the *nC* follows a clearly non-linear trend as the size of the GNF increases, with the smaller GNFs showing very strong interactions with the considered proteins whereas as the GNF size increases, an asymptotic value (roughly 0.1 kcal mol^{-1} per C atom) is reached. Therefore, the reported results point to very strong interaction of small GNFs as a result of the insertion of the GNFs in the protein available cavities, whilst as the GNF size increases, the nanomaterial – protein interactions mainly occured in the outer surfaces of the protein, leading to an additive effect with increasing GNF size.

Fig. S2 (Supplementary Information) shows the mechanism of interaction with two archetypical proteins (1i1B and 2hi4) of GNFs with hexagonal, rectangular and trigonal geometries for small flakes (24, 16, and 22 C atoms, respectively), which are able to fit into the corresponding protein active sites. The reported results confirms that the poses with stronger GNF – protein interaction correspond to the GNFs inserted into the protein cavity, where efficient π – π interaction with available aromatic rings in the active cavity is allowed, being independent of the geometry of the GNF shape and the type of protein.

The mechanism of GNF – protein interaction for small GNFs is analysed in detail in Fig. 3. The reported results show interaction via the available Phe molecular sites in the protein cavities for all the cases, with very short contacts for the π – π interactions (roughly in the 3.5 to 4.0 Å range), which points to very efficient interactions among the GNFs aromatic moieties and the protein cavity. Likewise, the π – π interactions are accompanied by interaction with neighbour polar and acidic sites, and also to GNF and protein relevant exposure. These mechanisms are almost independent of the shape of the GNFs, as the size of the nanoparticle controls the interaction with the protein cavity size, with a critical value of the GNFs (roughly $nC = 150$) in which the GNFs nanoparticle size does not allow penetration into the proteins cavity and the mechanism of interaction evolves toward adsorption of the outer protein surface and thus to *i*) weaker interaction and *ii*) lower perturbation of protein properties. Therefore, those GNFs with $nC < 150$ will have larger effects on the considered proteins and thus possible larger toxicological effects.

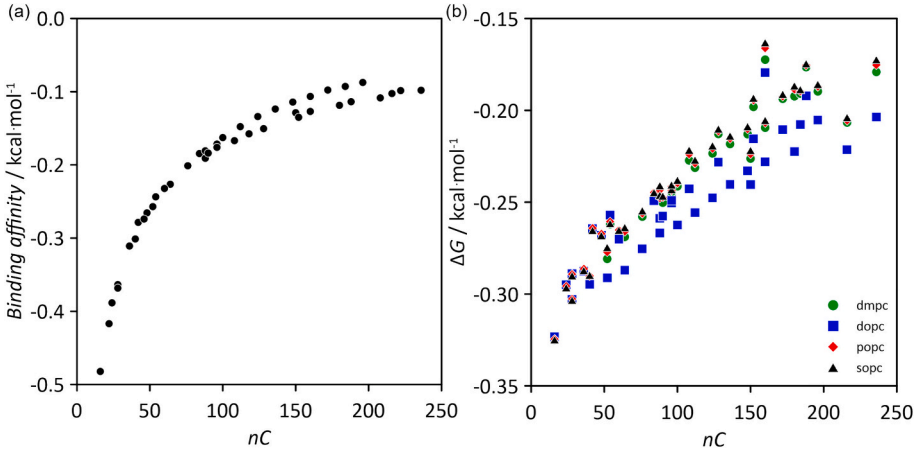

Fig. 2. (a) Interaction energy – binding affinity – per number of carbon atoms for protein interaction and (b) free energy, ΔG, profile through the four membrane types considered in this work – DMPC, DOPC, POPC and SOPC–, measured at the center of the membrane; energy expressed per number of carbon atoms.

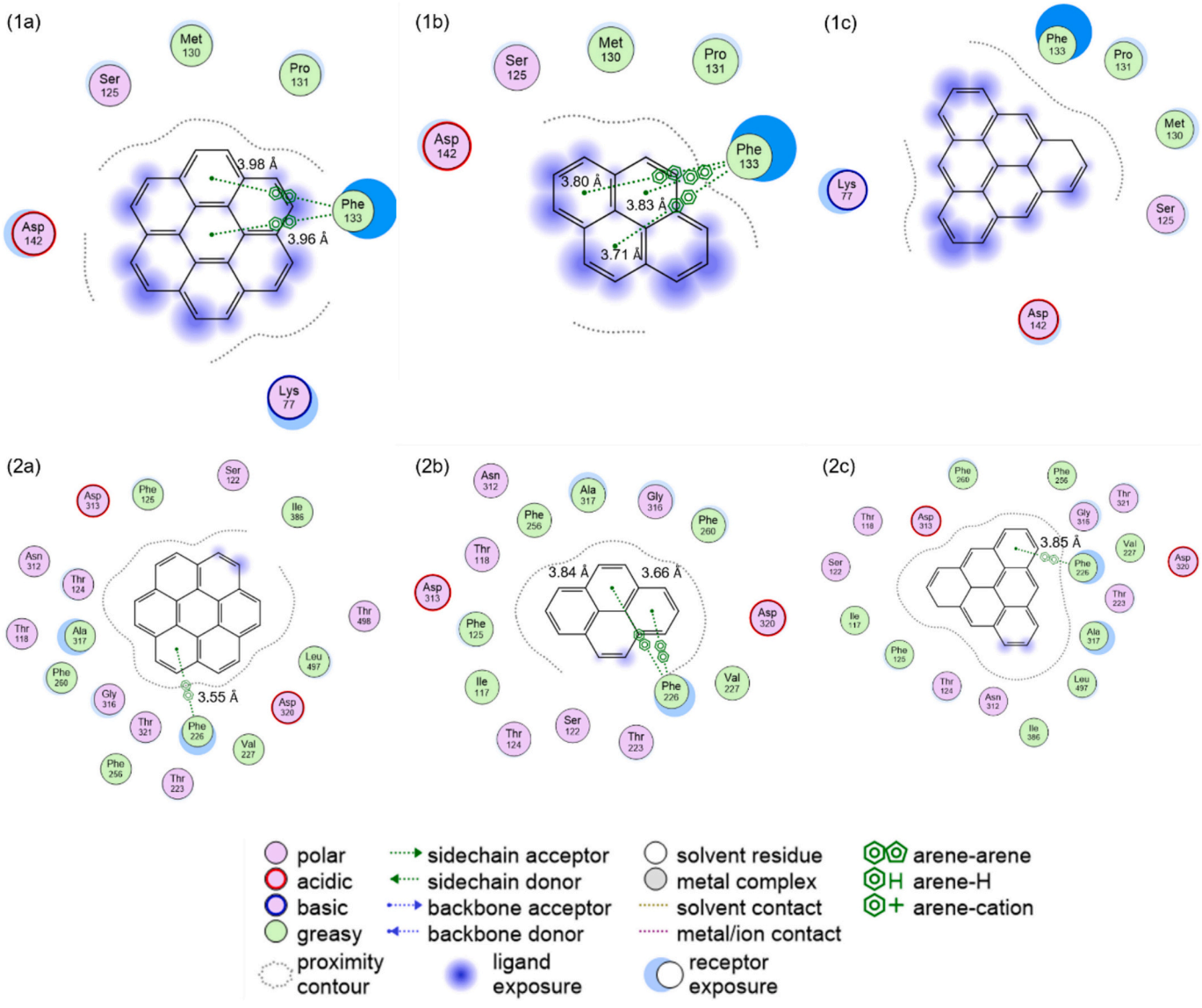

Fig. 3. Ligand interaction diagram for (1) 1i1b protein and (2) 2hi4 with three graphene flakes (a) HEX_ZIG_C24, (b) RECT_zz2ac2_C16 and (c) TRI_ZIG_C22.

3.2. Plasma membrane interactions: structural, thermodynamic and dynamics

The second considered biological effect of GNFs is the interaction with model (POPC, DOPC, DMPC, SOPC) plasma membranes, Fig. S3 (Supplementary Information). The first analysis conducted considers the thermodynamics of the penetration of the GNF along the model plasma membrane, quantified via the corresponding free energy, ΔG. The ΔG was calculated at the center of the bilayer (membrane), Table S7 (Supplementary Information), to quantify the GNFs stabilization upon full insertion into the membrane, and normalized per *nC*, Fig. 2b and Fig. S4 (Supplementary Information). The reported results show $\Delta G < 0$ at the membrane center for all the considered GNFs, i.e., all the GNFs are stabilized at the membrane center. Nevertheless, a non-linear behavior with *nC* is inferred, in parallel to those reported for the interaction with the considered proteins (Fig. 2a), with the larger (in absolute value) ΔG for the smaller GNFs although and being roughly constant (-0.20 kcal mol^{-1}) for $nC > 150$, with very minor effects on the type of lipid considered for the formation of the bilayer (slightly stronger interactions with DOPC bilayer).

The evolution of ΔG when GNFs difufse along the membrane, i.e., the possible energy barriers for GNFs penetration into the membranes, is reported in Fig. 4a for POPC bilayer and for all the considered lipids in Fig. S5 (Supplementary Information). It is clear that for all the types of lipid bilayers and GNFs sizes and shapes, the GNFs are able to cross the membrane without any thermodynamic barrier, evolving spontaneously toward the central apolar region of the lipid bilayer where they are largely stabilized. Therefore, all the GNFs penetrate the membrane and remain in the bilayer apolar central region.

The dynamic properties of the GNFs once penetrating in the different regions of the membranes are quantified via the corresponding self-diffusion coefficients, *D*. The reported results reveal that for the top region of the lipid bilayers, which are largely polar regions and thus for which GNF, considering their lipophilicity, show very minor affinity, have large *D* coefficients, i.e., GNFs cross these regions very quickly, Fig. 4b, without any thermodynamic barrier, Fig. 4a. This behavior is analogous for all the considered membranes and independent of the size and shape of the GNFs (Fig. S5, Supplementary Information). Likewise, as a consequence of the strong stabilization of the GNFs in the central regions of the membrane, the GNFs mobility is largely decreased, being almost frozen for some of the GNFs, which would lead to a large disruptive effect on the lipid bilayer lateral mobility, changing dynamic properties of the membrane upon insertion of the GNFs. The entropic effects upon GNFs penetration into the bilayer are analysed in Fig. S5 (Supplementary Information), although the entropic profiles show a complex pattern, presenting two maxima in the regions immediately below the polar region of the bilayer, which would corresponds to the GNFs rearrangement upon crossing the polar headgroups to adopt a parallel configuration with lipid non-polar chains to improve lipophilic interactions.

The penetration ability of the GNFs into the membranes was quantified via the predicted permeability coefficients, Figs. 4c and S4 (Supplementary Information). Regarding the shape effect on permeability, HEX > TRI > RECT, although for all the cases large permeability is inferred. Likewise, increasing GNFs size decreases permeability. Nevertheless, for all the GNFs cases and membraned lipidic composition, the nanoparticles show large permeability, thus being able to penetrate the membranes.

A deeper analysis on the mechanism of GNFs-membrane interaction is reported in Fig. 5. Firstly, as the GNF *nC* increases, the nanoparticle enlarge its stabilization at the membrane center, independently of the considered membrane lipidic composition. Secondly, for large nanoparticles, they are almost stacked with close to null mobility as a consequence of the strong interaction with the lipidic apolar chains, whereas small GNFs have certain mobility even after stabilization at the membrane center. Thirdly, the permeability evolves through a maximum for GNFs with $nC = 50$, showing the large penetration ability as a result of the combination of entropic and energetic effects upon penetration. Nevertheless, all the GNFs are able to penetrate the considered membranes independently of their size and lipid composition.

The effect of the GNFs shape is analyzed in Fig. S6 (Supplementary Information), with the shapes of nanoflakes evolving in a similar trend as the nanoparticle size grows. Therefore, GNFs shape has a minor effect on their penetration into the membranes. The possible effect of the type of considered GNFs edges is analysed in Fig. S7 (Supplementary Information) for armchair/zig zag configuration into RECT GNFs. Again, the edge effect seems to be negligible in comparison with GNFs size. Therefore, all the considered GNFs in the studied size range are able to penetrate the considered model cell membranes, being stabilized at the membrane centers, with the main effect controlling the thermodynamic and kinetic effects of the penetration rising from the GNFs size. The penetration is totally spontaneous and evolves through non-barrier intermediate situations with nanoparticles reorienting upon crossing the polar head groups of the membranes to remain at the bilayer center, especially for large GNFs. These results, points to large effects of GNFs on cell membranes, which would lead to large disruption of membrane properties and toxicity effects, as inferred from the strong and efficient interaction both with human proteins and cell membranes.

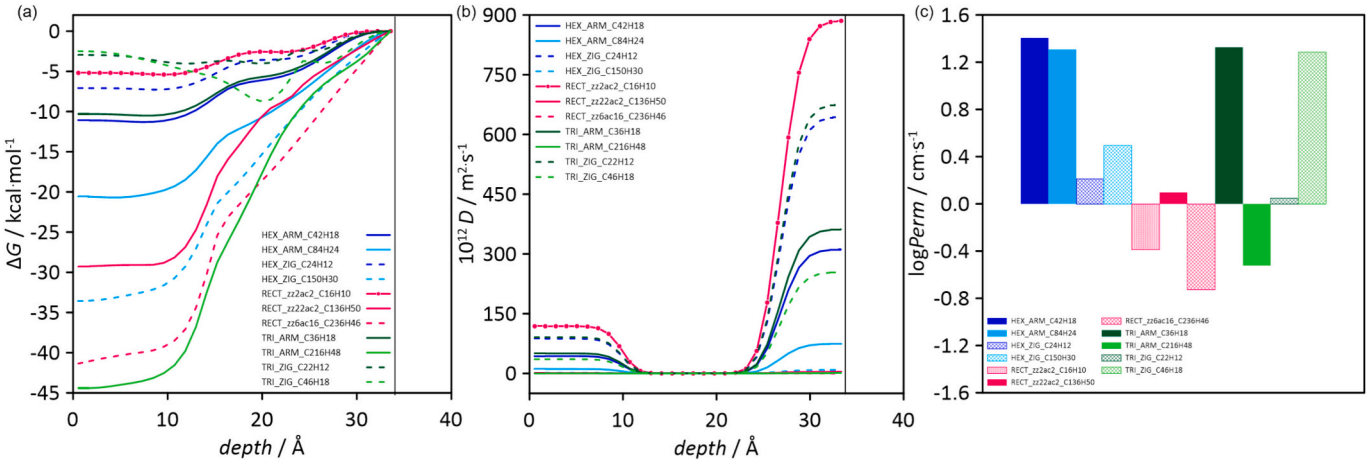

Fig. 4. Prediction of (a) free energy ΔG, (b) diffusion and (c) permeability through POPC membrane for a total of 11 graphene flakes, sampling different shapes, edges and sizes.

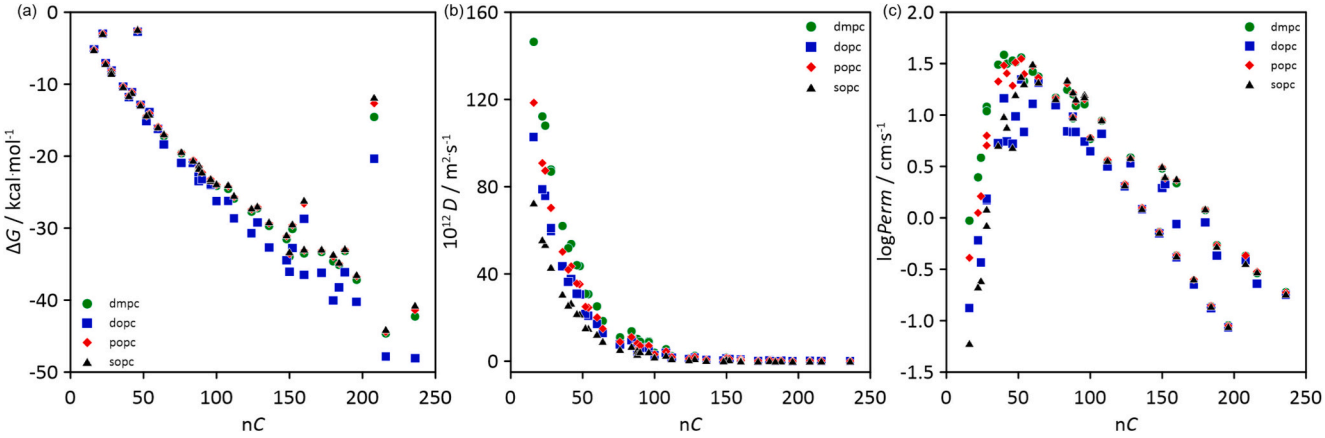

Fig. 5. Prediction of (a) free energy Δ*G*, (b) diffusion and (c) permeability measured at the center of the membrane (DMPC, DOPC, POPC and SOPC) for a total of 40 graphene flakes, sampling different shapes, edges and sizes, sorted by number of carbon atoms.

3.3. QSPR models: development and validation

3.3.1. Dimensionality reduction

The results of the Factor Analysis of Mixed Data (FAMD) indicated that the dimensional space could be reduced to five dimensions (Table S8, Supplementary Information), which together account for 74 % of the variability. The first and second Principal Components (PC1 and PC2) together account for 55.38 % of the variance in the data distribution, with PC1 explaining 41.9 % and PC2 explaining 13.44 %. The third, fourth, and fifth Principal Components (PC3, PC4, and PC5) explain 9.95 %, 5.38 %, and 4.15 %, respectively (Fig. S8, Table S8, Supplementary Information). The loadings of each variable were calculated and represented in absolute values to determine the contribution of each variable to each dimension or Principal Component (Fig. S9, Supplementary Information), which allows for the identification of the variables that have the greatest influence on the underlying structure of the data. To reduce dimensionality, only the variables that exceed a specified threshold for each Principal Component were selected.

The first principal component (PC1) explains less than 50 % of the total variance, suggesting that the remaining components also contain variables that contribute to the data's overall variability. As loadings tend to decrease with diminishing explained variance, the thresholds were set flexibly based on the explained variance of each component in an attempt to reduce dimensionality while retaining the essential information from each component that contributes to the total variance. The threshold for PC1 was set at 0.65, as it explains the largest portion of the variance in the data. For the subsequent components, the thresholds were set lower: 0.5 for PC2 and PC3, 0.45 for PC4, and 0.35 for PC5 (Fig. S9, Supplementary Information). A reduced dataset, consisting of 44 numerical and continuous descriptors (from an original 78), was obtained. The autocorrelation analysis revealed severe multicollinearity between the variables, with 38 out of 44 variables exhibiting pairwise Spearman correlations greater than 0.8. These high correlations were particularly observed within the first Principal Component (PC1), and to a lesser extent, in the second (PC2) and third (PC3) components. Additionally, several variables showed correlations greater than 0.7 and some above 0.6 within PCs 2, 3, 4, and 5 (Fig. S10). This suggests that many of the predictors in the dataset are strongly correlated.

The PCA confirmed that the variables selected from FAMD remained the most influential, except for one, Shadow area: YZ plane (belonging to PC3), which showed a significant drop in its loading compared to the two most influential variables in this dimension (Shadow area fraction: YZ plane and Shadow length: LY) and it was therefore not considered. The final dataset of descriptors contains 43 variables (Table S9, Supplementary Information).

3.3.2. Model optimization: Recursive Feature Elimination (RFE)

The Recursive Feature Elimination (RFE) to the Cell Membrane Interaction model identified that the best set of descriptors comprises all 43 features selected from the dataset. This selection was based on a bootstrapping process that used subsets of features ranging from 3 to 25 descriptors, with 50 bootstrapping repetitions and an optimal ensemble of 200 trees. The results show robust performance, with evaluation metrics including an RMSE of 0.060, an R^2 of 0.71, a MAE of 0.039, and standard deviations of 0.021 for RMSE, 0.157 for R^2, and 0.013 for MAE (Table S10, Supplementary Information).

For the Protein Interaction model, the Recursive Feature Elimination (RFE) analysis, based on a bootstrapping process, utilized subsets of features ranging from 3 to 25 descriptors, with 50 bootstrapping repetitions and an optimal ensemble of 50 trees. This analysis identified 22 descriptors as the optimal set. The selected descriptors include: EE_VAMPE, TE_VAMPE, HF_VAMPE, MP_VAMPE, OXZZ_VAMPE, OYYZ_VAMPE, OXYY_VAMPE, OYZZ_VAMPE, QYZ_VAMPE, DX_VAMPE, ZZDOD_X_VAMPE (electrostatic descriptors); TE_DMOL3M, BE_DMOL3M, SE_DMOL3M, DIE_DMOL3M, CV_DMOL3M, HOMO-E_DMOL3M (molecular descriptors); SAF_YZ_SD, SL_LY_SD (spatial descriptors); TMM_AD, nC (atomistic descriptors); and CSOV_AVS (atom volumes and surfaces). See Table S9 (Supplementary Information) for acronym definitions. The evaluation metrics for the model were as follows: RMSE = 0.0866, R^2 = 0.90, MAE = 0.068, with standard deviations of 0.018 for RMSE, 0.044 for R^2, and 0.015 for MAE (see Table S11, Supplementary Information). These results confirm the effectiveness of the selected descriptor set, demonstrating that the model is consistent across the various bootstrapping samples.

3.3.3. Cell membrane and Protein Interaction models: development and evaluation

The best Cell Membrane Interaction and Protein Interaction models were both obtained using the Random Forest algorithm with 50 trees, as found by the learning curve. The values of *mtry* tested through grid search for the Cell Membrane Interaction model included 10, 12, 14, and 16, while for the Protein Interaction model, the tested values were 3, 5, 7, 9, and 11 (see Methods for details on grid search). The best value of mtry for the Cell Membrane Interaction model was mtry = 10, and for the Protein Interaction model, it was mtry = 7 (Table S12, Supplementary Information). Both models were evaluated using repeated cross-validation with 10 repetitions and 5-fold cross-validation. The evaluation metrics for the best Cell Membrane Interaction model were RMSE = 0.050, R^2 = 0.79, and MAE = 0.035 (Table S12, Supplementary Information), while for the Protein Interaction model, the metrics were RMSE = 0.078, R^2 = 0.92, and MAE = 0.064 (Table S12, Supplementary Information). The best models were selected based on the smallest RMSE

value.

The performance of each optimized model was evaluated on an unseen test set consisting of seven GNFs, yielding the following metrics: RMSE = 0.043, R^2 = 0.94, and MAE = 0.033 for the Cell Membrane Interaction model, and RMSE = 0.099, R^2 = 0.94, and MAE = 0.078 for the Protein Interaction model. For the Cell Membrane Interaction model, both RMSE and MAE in the test phase were slightly lower than those in the training phase, with the R^2 value in the test phase being higher. In contrast, for the Protein Interaction model, while R^2 stays similar across both phases, the RMSE and MAE were slightly higher in the test phase.

The graph showing the relationship between normalized predicted values and actual values (from both the training and validation phases), fitted through a linear model (Y ~ X), shows an adjusted R^2 of 0.94 and 0.96 for Cell Membrane Interaction and Protein Interaction respectively and a p-value <0.0001 for both models (Fig. 6). For the Cell Membrane Interaction model, this strong relationship between predicted and actual values further supports the error metrics, suggesting good generalization to unseen data and reliable predictions. However, in the case of the Protein Interaction model, the strong correlation should be interpreted with caution, as the difference in error metrics shows that this model may have some limitations in generalizing to new data. These performance metrics were achieved following an exhaustive hyperparameter optimization process, which included the choice of optimal values for *mtry*, number of trees, cross-validation configurations, and repetitions. Given the small size of the training and test sets, these results stand for the best performance achievable with the available data. While further data and model refinements may improve results, the current Random Forest models offer valuable insights and are well-suited to the scope of this analysis. Future work could focus on using larger datasets and exploring alternative modeling techniques to enhance predictive accuracy and robustness, and incorporating model interpretability techniques such as SHAP (SHapley Additive exPlanation) analysis framework to better understand the contributions of predictors in the predictions, following the approach developed in a previous work (De-la-Huerta et al., 2025).

3.3.4. *Score. Assessing the impact of graphene nanoflake interactions with biological targets*

From Eq. (2), a score was obtained for each evaluated GNFs. This score, ranging from 0 to 1, is a quantitative measure of the interaction between the GNFs and human cell proteins and membranes. A graphical representation was used to assess the relationship between the score (Y-axis), the size (total number of carbon atoms, X-axis), and the geometry (shape) and edge characteristics, represented by colour (Fig. 7). To facilitate the visualization of the trend, a LOESS (Locally Estimated Scatterplot Smoothing) curve was fitted to the data. LOESS fits a smooth curve to the data using local polynomial regression, considering only the nearest data points to each location. The LOESS curve indicates a smooth decreasing trend between the score and the number of carbon atoms (Fig. 7). This relationship appears to be predominantly driven by the size of the GNFs. The different geometries (and implicitly their edges) do not appear to significantly influence this trend, as the coloured points representing different shapes follow the same overall decreasing pattern. While other structural factors such as geometry and edges may still have some subtle effects, it seems that the primary factor determining the score is the number of carbon atoms in the GNFs.

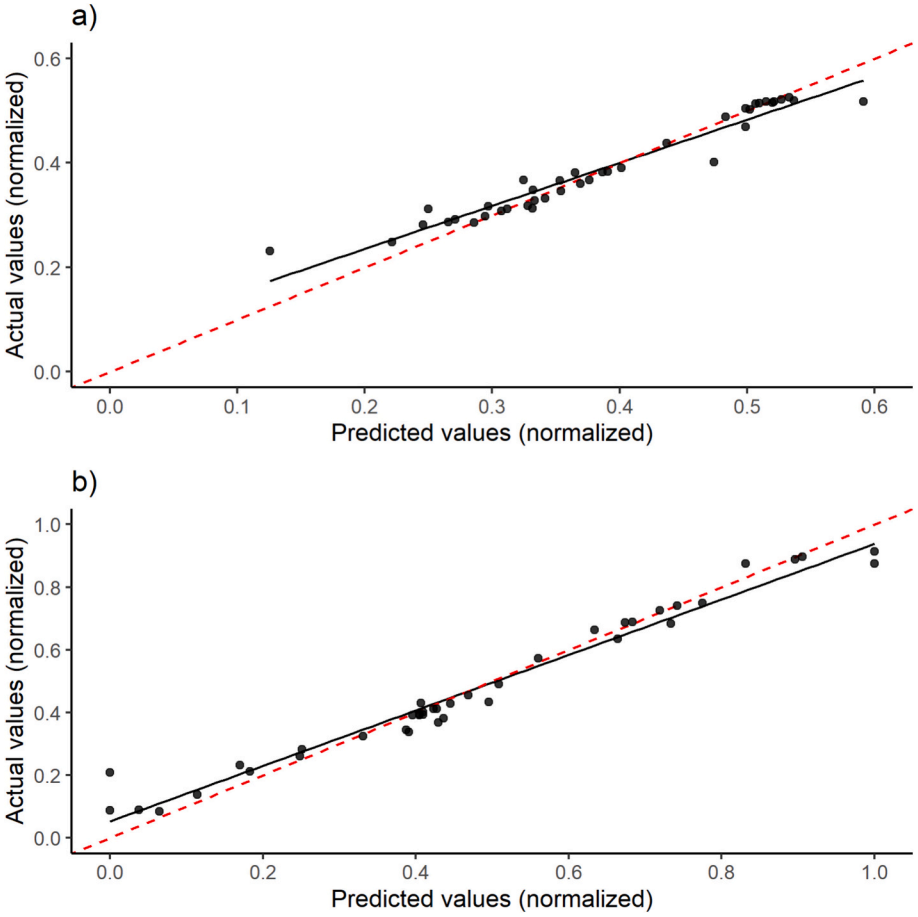

Fig. 6. Relationship between normalized predicted and actual values for a) Cell Membrane Interaction model and b) Protein Interaction model. The black line shows the linear fit (Y ~ X), and the red dashed line shows the 1:1 line. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

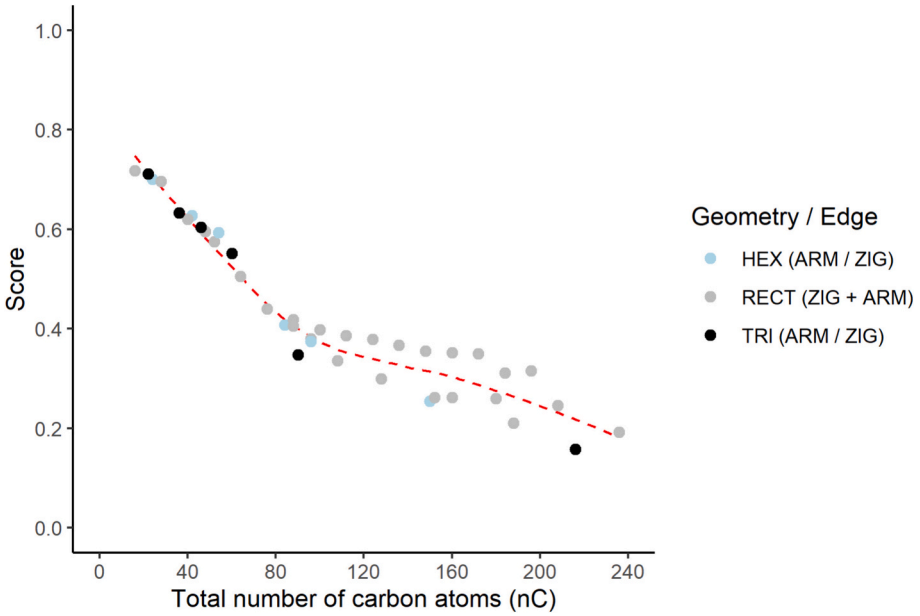

Fig. 7. Relationship between the score (Y-axis), the size (total number of carbon atoms, X-axis), and the geometry (shape) and edge characteristics, represented by colour. Light blue colour represents hexagonal GNFs, with 2 having an armchair (ARM) edge and 4 having a zigzag (ZIG) edge; grey colour represents 28 rectangular GNFs with two edge types, ZIG and ARM; black colour represents triangular GNFs, with 4 having an ARM edge and 2 having a ZIG edge. The dashed red line represents the LOESS curve. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Smaller nanoparticles tend to have a greater biological impact on cells due to their increased cellular penetration capacity and larger surface area, which can result in higher toxicity or stronger interactions with the cells. Based on existing literature, the present study assumes that higher scores, which are associated with GNFs containing fewer carbon atoms, reflect greater interaction and, consequently, a higher biological impact and potential toxicity. Following this rationale, four distinct levels of impact or toxicity (low, medium, high, and very high) were defined, and the evaluated GNFs were classified accordingly. These toxicity levels are determined based on quartiles of the score distribution, with each category representing a range of scores. The "Low" toxicity level corresponds to the lower quartile of the scores. The "Medium" impact or toxicity level corresponds to the middle 50 % of the score distribution. The "High" level corresponds to the 75th percentile, and the "Very High" level corresponds to the upper quartile of the score. This classification allows for an easy comparison of the biological impact or toxicity of the GNFs based on their score. Fig. 8 shows the comparative analysis of the biological impact or toxicity of the GNFs based on their score. Out of the 40 GNFs evaluated, 14 were classified with high toxicity, compared to the remaining GNFs. The GNFs with the highest impact (score = 0.71) has the smallest number of carbon atoms (16). A total of 22 GNFs exhibit medium toxicity, and four of them show lower toxicity than the rest. These GNFs with lower toxicity have the highest number of carbon atoms (above 180). As observed in the chart, the toxicity level tends to increase as the number of carbon atoms decreases. Notably, two GNFs, RECT-zz32ac2-C196H70 and RECT-zz30ac2-C184H66, which have a relatively high number of carbon atoms (196 and 184), could be expected to fall into the low toxicity category based on their carbon atoms content. However, they are classified as medium toxicity, which suggests that other structural factors, such as geometry and edges, may also influence their toxicity, beyond just the number of carbon atoms.

4. Conclusions

In this comprehensive study, we have presented an integrated *in silico* approach to investigate the properties, biological interactions, and potential toxicity of graphene GNFs. By combining DFT simulations, MD studies and ML techniques, we have gained valuable insights into the structure-property-toxicity relationships of these promising nanomaterials. DFT simulations revealed the strong dependence of GNF properties on size, shape, and edge configuration. It was observed that electronic, magnetic, and chemical characteristics can be finely tuned by controlling these parameters, offering a pathway for designing GNFs with tailored properties for specific applications. The systematic analysis of these structure-property relationships provides a solid foundation for future experimental work and the rational design of GNF-based devices.

The exploration of GNF interactions with biological systems, particularly human proteins and cell membranes, has shed light on potential toxicity mechanisms. Simulations demonstrated that the high surface area and reactive edges of GNFs can lead to strong interactions with biomolecules, potentially disrupting cellular processes. However, it was also identified that certain GNF configurations that exhibit reduced biological interactions, suggesting possibilities for designing safer nanoflakes for biomedical applications.

The developed ML models for predicting GNF properties and biological interactions represent a significant advancement in the field. These models, trained on comprehensive dataset of DFT and MD simulation results, offer rapid and accurate predictions of GNF characteristics and potential toxicity. This approach not only accelerates the screening of new GNF designs but also provides a valuable tool for researchers and regulators in assessing the safety of these nanomaterials.

In conclusion, this work demonstrates the power of integrating multiple computational approaches to study complex nanomaterials like GNFs. The findings not only advance the fundamental understanding of GNF properties and interactions but also provide practical tools for their safe and efficient application across various fields, from electronics to biomedicine. As it is necessary to explore the vast potential of graphene nanoflakes, this multifaceted in silico approach may play a role in guiding experimental efforts and accelerating the development of novel graphene-based technologies. The results from this study have led to a software tool (https://enaloscloud.novamechanics.com/diagonal/grapheneimpact/, n.d.) which allow the prediction and classification of GNFs toxicity in terms of size, shape and properties.

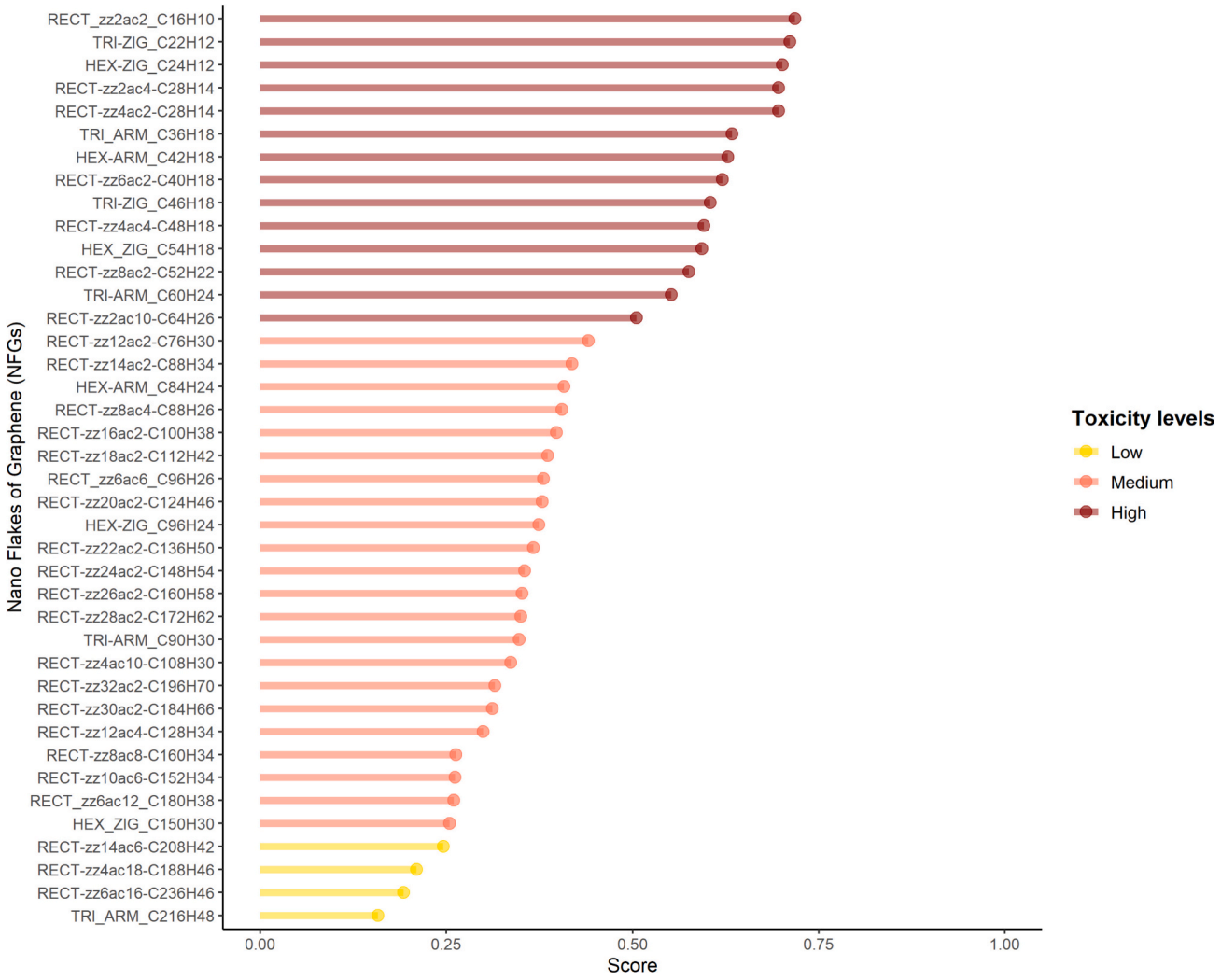

Fig. 8. Lollipop chart. Comparative analysis of the biological impact (toxicity level) of Nano Flakes of Graphene (GNFs) based on their score. Each GNF is represented by a horizontal line segment with a point at the end. The length of the segment represents the score, while the colour of both the segment and the point indicates the toxicity level: Golden Yellow for "Low" toxicity, Light Red-Orange for "Medium" toxicity, and Dark Red for "High" toxicity. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

CRediT authorship contribution statement

Nuria Aguilar: Writing – review & editing, Writing – original draft, Visualization, Validation, Investigation, Formal analysis, Data curation. **Patricia de la Fuente:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation. **Natalia Fernández-Pampín:** Writing – review & editing, Writing – original draft, Visualization, Investigation, Formal analysis, Data curation. **Sonia Martel:** Writing – review & editing, Writing – original draft, Visualization, Validation, Investigation, Funding acquisition, Formal analysis, Data curation. **Laura Gómez-Cuadrado:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Investigation, Funding acquisition, Formal analysis. **Pedro Angel Marcos:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Alfredo Bol:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Conceptualization. **Carlos Rumbo:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Santiago Aparicio:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

Acknowledgements

This work was funded by European Union H2020 Program (H2020-NMBP-TO-IND-2020-twostage-DIAGONAL-GA- 953152) and Junta de Castilla y León (Spain, project NANOCOMP - BU058P20). We also acknowledge SCAYLE (Supercomputación Castilla y León, Spain) for providing supercomputing facilities. The statements made herein are

solely the responsibility of the authors. Authors declare no competing interests.

Appendix B. Supplementary data

Total number of atoms, shape, type of edge and dimensions of each of the 98 GNFs considered in this work (Table S1); BIOVIA Materials Studio descriptors chosen for the 98 GNFs (Table S2); Human proteins considered in this work (Table S3); GNFs used for pre-processing, model optimization, and training of Random Forest predictive models and GNFs used for testing the performance of the models (Table S4); Modes, frequencies and intensities for vibrational spectra for three GNFs and their total Gibbs free energy (Table S5); Docking mean binding affinity and standard deviation of the protein – GNFs interaction (Table S6); Free energy of penetration through the four types of membranes considered in this work (Table S7); Numerical results for Principal Components analysis (Table S8); Relevant descriptors identified through FAMD and PCA analysis (Table S9); Recursive Feature Elimination Analysis (Tables S10–11); Evaluation metrics for the Cell Membrane and Protein Interaction models using Random Forest (Table S12); ESP of six GNFs (Fig. S1); Docking results representing protein-ligand contacts with GNFs (Fig. S2); Representation of the model plasma cell membranes considered in this work (Fig. S3); Free energy for the interaction between graphene nanoflakes and model membranes (Fig. S4); properties for the interaction of selected graphene nanoflakes with model membranes (Fig. S5); Prediction of free energy ΔG, diffusion and permeability of the 4 types of membranes considered in this work, for a total of 40 GNFs (Fig. S6); Prediction of free energy ΔG, diffusion and permeability of the 4 types of membranes considered in this work, for a total of 16 RECT GNFs (Fig. S7); Scree Plot of Principal Components analysis (Fig. S8); Results of Principal Component analysis (Fig. S9); Correlation matrix analysis (Fig. S10).

File Supplementary_Information_xyz_DFT.zip containing the xyz optimized coordinates (DFT wB97x/def2-TZVP) for the 98 GFNs.

CSV file containing nanodescriptors for the considered graphene nanoflakes used for models development along this work. Supplementary data to this article can be found online at https://doi.org/10.1016/j.impact.2025.100563.

Data availability

Data will be made available on request.

References

Alecu, I.M., Zheng, J., Truhlar, D.G., 2010. Computational thermochemistry: scale factor databases and scale factors for vibrational frequencies obtained from electronic model chemistries. J. Chem. Theor. Comput. 6, 2892.

Analytics, R., Weston, S., 2018. _doMC: foreach parallel adaptor for 'parallel'_. R package version 1.3.5/r19. https://R-Forge.R-project.org/projects/domc/.

Avouris, P., Dimitrakopoulos, C., 2012. Graphene: synthesis and applications. Mater. Today 15, 86–97.

Barnard, A.S., 2016. Challenges in modelling nanoparticles for drug delivery. J. Phys. Condens. Matter 28, 023002.

Barnard, Amanda, Sun, Baichuan, Shi, Hongqing (Robyn), 2016. Graphene Nanoflake Data Set. v1. Commonwealth Scientific and Industrial Research Organisation. dataset. https://doi.org/10.4225/08/57998CC4D7891.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The Protein Data Bank. Nucleic Acids Res. 28, 235–242.

Bianco, A., 2013. Graphene: safe or toxic? The two faces of the medal. Angew. Chem. Int. Ed. 52, 4986–4997.

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Burello, E., Worth, A.P., 2011. A theoretical framework for predicting the oxidative stress potential of oxide nanoparticles. Nanotoxicology 5, 228–235.

Butler, K.T., Davies, D.W., Cartwright, H., Isayev, O., Walsh, A., 2018. Machine learning for molecular and materials science. Nature 559, 547–555.

Cai, J., Luo, J., Wang, S., Yang, S., 2018. Feature selection in machine learning: a new perspective. Neurocomputing 300, 70–79.

Chen, C., Gu, G.X., 2019. Machine learning for composite materials. MRS Commun. 9, 556–566.

Chng, E.L.K., Pumera, M., 2013. The toxicity of graphene oxides: dependence on the oxidative methods used. Chem. Eur. J. 19, 8227–8235.

Choi, J.S., Ha, M.K., Trinh, T.X., Yoon, T.H., Byun, H.G., 2018. Towards a generalized toxicity prediction model for oxide nanomaterials using integrated data from different sources. Sci. Rep. 8, 6110.

Choi, J.S., Trinh, T.X., Yoon, T.H., Kim, J., Byun, H.G., 2019. Quasi-QSAR for predicting the cell viability of human lung and skin cells exposed to different metal oxide nanomaterials. Chemosphere 217, 243–249.

Cocchi, C., Ruini, A., Prezzi, D., Caldas, M.J., Molinari, E., 2010. Electronic and optical properties of bilayer graphene nanoflakes. J. Phys. Chem. C 115, 2969–2973.

De-la-Huerta, S., Escobedo-Monge, M.A., Marcos, P.A., Esteban-Ollo, J.A., Montejo-Gil, L., Conde-Rioll, M., Atilhan, M., Bol, A., Aparicio, S., 2025. Nature's tool kit: designing biocompatible and affordable NADES for sustainable extraction of plant bioactives. Sust. Chem. One World 5, 100043.

Ezawa, M., 2007. Metallic graphene nanodisks: electronic and magnetic properties. Phys. Rev. B 76, 245415.

Fernández-Rossier, J., Palacios, J.J., 2007. Magnetism in graphene nanoislands. Phys. Rev. Lett. 99, 177204.

Fourches, D., Pu, D., Tassa, C., Weissleder, R., Shaw, S.Y., Mumper, R.J., Tropsha, A., 2010. Quantitative nanostructure-activity relationship modeling. ACS Nano 4, 5703–5712.

Franzke, Y.J., Holzer, C., Andersen, J.H., Begušić, T., Bruder, F., Coriani, S., Della Sala, F., Fabiano, E., Fedotov, D.A., Fürst, S., Gillhuber, S., Grotjahn, R., Kaupp, M., Kehry, M., Krstić, M., Mack, F., Majumdar, S., Nguyen, B.D., Parker, S.M., Pauly, F., Pausch, A., Perlt, E., Phun, G.S., Rajabi, A., Rappoport, D., Samal, B., Schrader, T., Sharma, M., Tapavicza, E., Treß, R.S., Voora, V., Wodyński, A., Yu, J.M., Zerulla, B., Furche, F., Hättig, C., Sierka, M., Tew, D.P., Weigend, F., 2023. J. Chem. Theory Comput. 19, 6859–6890.

Fu, Q., Xu, T., Wang, D., Liu, C., 2024. Rare earth modified carbon-based catalysts for oxygen electrode reactions: a machine learning assisted density functional theory investigation. Carbon 223, 119045.

Gu, G.X., Chen, C.T., Richmond, D.J., Buehler, M.J., 2018. Bioinspired hierarchical composite design using machine learning: simulation, additive manufacturing, and experiment. Mater. Horiz. 5, 939–945.

Gu, Z., Yang, Z., Wang, L., Zhou, H., Jimenez-Cruz, C.A., Zhou, R., 2015. The role of basic residues in the adsorption of blood proteins onto the graphene surface. Sci. Rep. 5, 10873.

Güçlü, A.D., Potasz, P., Hawrylak, P., 2014. Graphene-based integrated electronic, photonic and spintronic circuits. J. Phys. Chem. Lett. 5, 2430–2435.

Guo, R., Mao, J., Yan, L.T., 2013. Computer simulation of cell entry of graphene nanosheet. Biomaterials 34, 4296–4301.

Hamner, B., Frasco, M., 2018. _Metrics: evaluation metrics for machine learning_. R package version 0.1.4. https://CRAN.R-project.org/package=Metrics.

Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer.

Hod, O., Barone, V., Scuseria, G.E., 2008. Half-metallic graphene nanodots: a comprehensive first-principles theoretical study. Phys. Rev. B 77, 035411.

https://enaloscloud.novamechanics.com/diagonal/grapheneimpact/.

Huang, J., Mckerell, A.D., 2013. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. J. Comput. Chem. 34, 2135–2145.

Kassambara, A., Mundt, F., 2020. _factoextra: extract and visualize the results of multivariate data Analyses_. R package version 1.0.7. https://CRAN.R-project.org/package=factoextra.

Klamt, A., Jonas, V., Bürger, T., Lohrenz, J.C.W., 1998. Refinement and parametrization of COSMO-RS. J. Phys. Chem. A 102, 5074–5085.

Kuhn, M., 2008. Building predictive models in R using the caret package. J. Stat. Softw. 28 (5), 1–26. https://doi.org/10.18637/jss.v028.i05.

Le, Sebastien, Josse, Julie, Husson, Francois, 2008. FactoMineR: an R package for multivariate analysis. J. Stat. Softw. 25 (1), 1–18. https://doi.org/10.18637/jss.v025.i01.

Lee, H., Son, Y.-W., Park, N., Han, S., Yu, J., 2005. Magnetic ordering at the edges of graphitic fragments: magnetic tail interactions between the edge-localized states. Phys. Rev. B 72, 174431.

Lizano-Fallas, V., Carrasco, A., Cristobal, S., 2023. Prediction of molecular initiating events for adverse outcome pathways using high-throughput identification of chemical targets. Toxics 11, 189.

Martínez, G., Giorgino, T., De Fabritis, G., 2017. PlayMolecule ProteinPrepare: a web application for protein preparation for molecular dynamics simulations. J. Chem. Inf. Model. 57, 1511–1516.

Meunier, M., Robertson, S., 2021. Materials Studio. 20th anniversary. Mol. Simul. 47, 537–539.

Morris, G., Huey, R., Lindstrom, W., Sanner, M., Belew, R., Goodsell, D., Olson, A., 2009. Autodock4 and AutoDockTools4: automated docking with selective receptor flexiblity. J. Comput. Chem. 30 (2009), 2785–2791.

Mu, Q., Su, G., Li, L., Gilbertson, B.O., Yu, L.H., Zhang, Q., Sun, Y.P., Yan, B., 2012. Size-dependent cell uptake of protein-coated graphene oxide nanosheets. ACS Appl. Mater. Interfaces 4, 2259–2266.

Neese, F., . The ORCA program system. Wiley Interdiscip. Rev.: Comput. Mol. Sci. *2, 73–78*.

Novoselov, K.S., Geim, A.K., Morozov, S.V., Jiang, D., Zhang, Y., Dubonos, S.V., Grigorieva, I.V., Firsov, A.A., 2004. Electric field effect in atomically thin carbon films. Science 306, 666–669.

Ou, L., Song, B., Liang, H., Liu, J., Feng, X., Deng, B., Sun, T., Shao, L., 2016. Toxicity of graphene-family nanoparticles: a general review of the origins and mechanisms. Part. Fibre Toxicol. 13, 57.

Perdew, P., Burke, K., Ernzerhof, M., 1997. Generalized gradient approximation made simple. Phys. Rev. Lett. 78, 1396.
Posit team, 2023. RStudio: Integrated Development Environment for R. Posit Software, PBC, Boston, MA. URL. http://www.posit.co/.
Puzyn, T., Rasulev, B., Gajewicz, A., Hu, X., Dasari, T.P., Michalkova, A., Hwang, H.M., Toropov, A., Leszczynska, D., Leszczynski, J., 2011. Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles. Nat. Nanotechnol. 6, 175–178.
R Core Team, 2024. _R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
Rajasekaran, G., Narayanan, P., Parashar, A., 2016. Effect of point and line defects on mechanical and thermal properties of graphene: a review. Crit. Rev. Solid State Mater. Sci. 41, 47–71.
Ramprasad, R., Batra, R., Pilania, G., Mannodi-Kanakkithodi, A., Kim, C., 2017. Machine learning in materials informatics: recent applications and prospects. npj Comput. Mater. 3, 54.
Rosi, N.A., Kassim, H.A., Shrivastava, K.N., 2010. Graphene infrared spectroscopy: DFT vibrational frequencies. AIP Conf. Proc. 1250, 269.
Schauberger, P., Walker, A., 2024. _openxlsx: read, write and edit xlsx files_. R package version 4.2.6.1. https://CRAN.R-project.org/package=openxlsx.
Schinwald, A., Murphy, F.A., Jones, A., MacNee, W., Donaldson, K., 2012. Graphene-based nanoplatelets: a new risk to the respiratory system as a consequence of their unusual aerodynamic properties. ACS Nano 6, 736–746.
Schwöbel, J.A., Ebert, A., Bitterman, K., Huniar, U., Goss, K.U., Klamt, A., 2020. COSMOperm: mechanistic prediction of passive membrane permeability for neutral compounds and ions and its pH dependence. J. Phys. Chem. B 124, 3343–3354.
Seabra, A.B., Paula, A.J., de Lima, R., Alves, O.L., Durán, N., 2014. Nanotoxicity of graphene and graphene oxide. Chem. Res. Toxicol. 27, 159–168.
Shao, Y., Wang, Y., Mei, K., Wu, F., Zhou, L., Zhang, L., 2019. Prediction of toxicity of nanoparticles by machine learning. Ecotoxicol. Environ. Saf. 186, 109819.
Sharma, B.R., Manjanath, A., Singh, A.K., 2014. Magnetic properties of triangular graphene quantum dots: comparative study between Hubbard and DFT models. AIP Adv. 4, 127131.
Son, Y.-W., Cohen, M.L., Louie, S.G., 2006. Energy gaps in graphene nanoribbons. Phys. Rev. Lett. 97, 216803.
Trott, O., Olson, A.J., 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. 31, 455–461.
Tu, Y., Lv, M., Xiu, P., Huynh, T., Zhang, M., Castelli, M., Liu, Z., Huang, Q., Fan, C., Fang, H., Zhou, R., 2013. Destructive extraction of phospholipids from *Escherichia coli* membranes by graphene nanosheets. Nat. Nanotechnol. 8, 594–601.
Wei, Taiyun, Simko, Viliam, 2024. R package 'corrplot': visualization of a correlation matrix (version 0.94). Available from. https://github.com/taiyun/corrplot.
Wei, Z.Y., Yao, X.F., 2019. Machine learning approach to predict thermal conductivity of carbon nanotubes filled polymer composites. Results Phys. 13, 102149.
Wickham, H., Averick, M., Bryan, J., Chang, W., LD, McGowan, François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the tidyverse. J. Open Source Softw. 4 (43), 1686. https://doi.org/10.21105/joss.01686. <doi: 10.21105/joss.01686>.
Winkler, D.A., Mombelli, E., Pietroiusti, A., Tran, L., Worth, A., Fadeel, B., McCall, M.J., 2013. Applying quantitative structure-activity relationship approaches to nanotoxicology: current status and future potential. Toxicology 313, 15–23.
Yan, L., Zhao, F., Li, S., Hu, Z., Zhao, Y., 2011. Low-toxic and safe nanomaterials by surface-chemical design, carbon nanotubes, fullerenes, metallofullerenes, and graphenes. Nanoscale 3, 362–382.
Yazyev, O.V., 2010. Emergence of magnetism in graphene materials and nanostructures. Rep. Prog. Phys. 73, 056501.
Zarenia, M., Pereira, J.M., Chaves, A., Peeters, F.M., Farias, G.A., 2010. Simplified model for the energy levels of quantum rings in single layer and bilayer graphene. Phys. Rev. B 81, 045431.
Zhang, X., Yin, J., Peng, C., Hu, W., Zhu, Z., Li, W., Fan, C., Huang, Q., 2011. Distribution and biocompatibility studies of graphene oxide in mice after intravenous administration. Carbon 49, 986–995.
Zou, Y., Li, F., Zhu, Z.H., Zhao, M.W., Xu, X.G., Su, X.Y., 2011. An ab initio study on gas sensing properties of graphene and Si-doped graphene. Eur. Phys. J. B 81, 475–479.
Zuo, G., Kang, S.G., Xiu, P., Zhao, Y., Zhou, R., 2013. Interactions between proteins and carbon-based nanoparticles: exploring the origin of nanotoxicity at the molecular level. Small 9, 1546–1556.