

Nanomaterial functionalisation modulates hard protein corona formation: Atomistic simulations applied to graphitic materials

Mohamed Ali al-Badri,^{*,†} Paul Smith,[†] Khuloud T al-Jamal,[‡] and Christian D Lorenz^{*,†}

[†]*Department of Physics, King's College London, London WC2R 2LS, UK.*

[‡]*Institute of Pharmaceutical Science, King's College London, London SE1 9NH, UK.*

E-mail: mohamed.al-badri@kcl.ac.uk; chris.lorenz@kcl.ac.uk

Abstract

The protein corona is an obstacle to exploiting the exotic properties of nanomaterials in clinical and biotechnology settings, with potential applications in DNA sequencing, point of care testing and drug delivery vehicles. The formation of the protein corona is driven by dynamic atomic scale interactions at the bio-nano interface, which are often impenetrable using conventional experimental techniques. Here, we use molecular dynamics simulations to study the effect of graphene-oxide (GO) functionalisation on apolipoprotein-c3 (apo-c3) adsorption. Here we develop an analysis pipeline, encompassing binding free energies at the bio-nano interface to protein structure analyses such as UMAP dimensionality reduction and machine learning clustering. These collectively extract experimentally verifiable features of protein-nanomaterial adsorption. We find that apo-c3 is denatured by adsorption on GO — largely driven by the large energetic contributions of electrostatic interactions, such as π - π stacking of aromatic amino acids to pristine graphene regions — losing its secondary structure

and forming β -turn binding motifs for protein-protein interactions that drive further protein aggregation. When adsorbing on double-clickable azide- and alkyne-double functionalised graphene oxide (C2GO), apo-c3 largely retains its secondary structure through more diverse interactions with the nanomaterial functional groups. By retaining the structure of the C-terminus, C2GO-adsorbed apo-c3 retains its lipid binding capabilities that may be related to experimentally observed increased cellular uptake of the C2GO hard corona relative to the GO hard corona.

1 Introduction

Graphene-oxide (GO) is a semi-ordered 2D material that shares many novel mechanical and electronic properties with graphene. It is utilised in applications in ion trapping, desalination, electronics, chemistry and biomedicine.^{1–3} GO is a promising platform for *ex* and *in vivo* biological applications — such as bio-sensing and therapeutics. Its electrochemical properties makes GO an attractive contender for bio-sensing applications such as single nucleotide polymorphism detection in DNA,⁴ next generation nanopore DNA sequencing⁵ and point of care (PoC) applications for early virus detection using covalently linked immobilised monoclonal antibodies.⁶ Therapeutically, the large surface area to volume ratio of 2D materials results in a large loading capacity for targeting molecules, fluorescent dyes and drug molecules for intravenous administration as well as photothermal therapy for cancer treatment.^{7–10}

Unfortunately the transition from *in vitro* to clinical settings is held back by a chemically driven adsorption of serum proteins — referred to as a protein corona — upon their introduction to a biological medium.^{11,12} A hard corona (HC) is a tightly bound monolayer of proteins at the nanoparticle interface. Subsequent protein layers adherent to the HC are referred to as the soft corona (SC).¹³ The character and composition of the HC define the nanomaterial's biological identity and therefore its biological fate, circulation time, cellular uptake and cytotoxicity.^{14,15} HC formation has revealed highly variant protein composition profiles, sensitive not only to inter-species biological media¹⁶ but also disease-specific influences in human GO coronas.¹⁷ Patient-specific or ‘personalised’ protein corona profiles have since been utilised as a diagnostic tool for high-throughput, inexpensive and highly accurate early cancer detection.¹⁸

Modulating the HC character to overcome the disadvantages of 2D materials in biological applications through nano-functionalisation remains challenging.¹⁹ Extrapolating the sensi-

tivity of nano-functionalisation to the aforementioned highly variant identity of the corona from experimental studies is incomplete and requires a better understanding of GO-HC interfacial behaviour to atomistic precision. Such an understanding is paramount to designing the next generation of biosensors and nanomedicines.

Azide- and alkyne-double functionalised graphene oxide (C2GO) has previously been proposed as a cancer targeting nanovector,²⁰ where click reactions conjugate targeting moieties on azide and trimethylsilyl (TMS)-alkyne functional groups.²¹ In vitro, both C2GO and GO show varying HC character when exposed to serum proteins, which subsequently impacts their biological identities.¹⁵ An experimental quality-by-design screening platform was used to unpick the HC character and evaluate its relationship with biological fate, cellular uptake and cytotoxicity. Using this, proteins that reduced material dependent toxicities were identified to play a role in diminishing cytotoxicity, including apolipoprotein C-III (apo-c3).¹⁵

Protein coronas are known to temporally evolve, leaving the total amount of protein constant but varying their composition according to binding affinity, known as the Vroman effect.²² Upon the introduction of a nanomaterial to a biological medium, highly abundant proteins such as albumin, immunoglobulin G and fibrinogen bind to the nanoparticle surface in the early stages of the Vroman effect, followed by their replacement with high binding affinity proteins such as apolipoproteins and coagulation factors in the late stages of the Vroman effect.²²⁻²⁶ The Vroman effect has previously been observed in computational and experimental studies of peptide, cellulose and fatty-acid binding on GO.²⁷

Previously, MD has been used to study the interactions of GO with peptides to understand conformational transitions of amyloid-beta during adsorption,²⁸ the hydration pattern of an adsorbed toy-model alpha-helix²⁹ and verification of enzyme active-site deformation following adsorption.³⁰ To the best of our knowledge, no MD simulations have so far studied

protein adsorption on accurate models of GO, and have instead randomly placed oxidised functional groups on the GO surface. Accurate modelling of GO should reflect the semi-ordered structure of GO; composed of inhomogeneous regions of oxidised and unoxidised domains, where amorphous alcohol and epoxy groups make up the oxidised regions.³¹ *Ab initio* MD simulations of GO show that semi-ordered models of GO are the most stable structures in vacuum as well as in liquid water.³² Furthermore, we have recently shown the importance of accurate functionalisation and accounting for steric strain and edge functional groups in large scale MD simulations of GO, using generalised and bespoke electronic structure MD forcefield design.³³ In this work, we use molecular dynamics (MD) simulations to study protein denaturing through adsorption on GO and C2GO, to understand HC conformation-activity relationships through binding free energy, contact map, protein structure and solvent exposure analyses. We apply machine learning dimensionality reduction techniques to decode the spatio-temporal MD data that is hard to interpret into distinct protein secondary structure conformations.

2 Results

2.1 Binding on the bio-nano interface

In both GO and C2GO systems, apo-c3 readily adsorbs to the substrate and reaches a stable conformation over the course of the adsorption trajectory. To investigate the binding of apo-c3 to the GO and C2GO interface, we compute contact maps and free energies of binding that underpin the changes in protein structure following adsorption. These analyses probe interfacial dynamics that may play a large role in the formation of a protein corona around a nanomaterial upon its introduction to a biological medium. To identify adsorption of apo-c3 to the graphitic sheets we calculated the minimum distance over time between any heavy (non-hydrogen) atom of the protein to any heavy atom of the sheet (Fig. S1). We also compute the minimum distance of each residue to the graphitic sheets over time. These

are combined to produce a heatmap of contact probability between apo-c3 and GO/C2GO (Fig. 1B).

The average binding free energy components for the GO-apo-c3 and C2GO-apo-c3 systems are given in Table 1. The binding free energy components are decomposed to per-residue contributions (Fig. 1A). In this way we can acquire a better understanding of each amino-acid's contribution to the binding free energy of the protein-nanomaterial complex. Per-residue binding free energy contributions are colour-coded by their interaction strength (blue to red) in the protein structure of the final configuration (Fig. 1C) for GO and C2GO, where each largely contributing amino acid residue has been illustrated and labelled explicitly.

Table 1: Binding energy components from MM-PBSA calculations performed on the MD trajectories of adsorbed apo-c3 on GO and C2GO sheets. Stronger binding components have been highlighted in bold.

	ΔE_{vdW} (kJ/mol)	ΔE_{elec} (kJ/mol)	ΔG_{polar} (kJ/mol)	ΔG_{apolar} (kJ/mol)	$\Delta G_{binding}$ (kJ/mol)
GO	-338.8±0.8	-208.3±1.5	422.2±5.5	-31.9±0.1	-156.8±5.3
C2GO	-463.2±0.9	-135.7±1.0	420.5±3.0	-47.5±0.1	-225.9±2.8

To understand how nano-functionalisation affects access of protein residues to groups/regions on the nanosheet, post adsorption structural fluctuations are visualised as contact probabilities (Fig. 1B, Fig. S2 and Fig. S3). Instead, there is a strong correlation between protein residue contact with azide/TMS functional groups and oxygen-containing surface functional groups, where these residues are likely being anchored to the nanosheet via azide/TMS groups. As a result, the protein is bound proportionately over the entire bio-nano interface (Fig. S1) by a more diverse set of interactions. This stabilises the adsorbed protein and acts as an energetic barrier to denaturing the protein as is the case with GO adsorption (Fig. 1). These reveal that, contrary to what has been proposed before,¹⁵ graphitic nanosheet surface functional groups such as azides do not sterically hinder serum protein residues from the

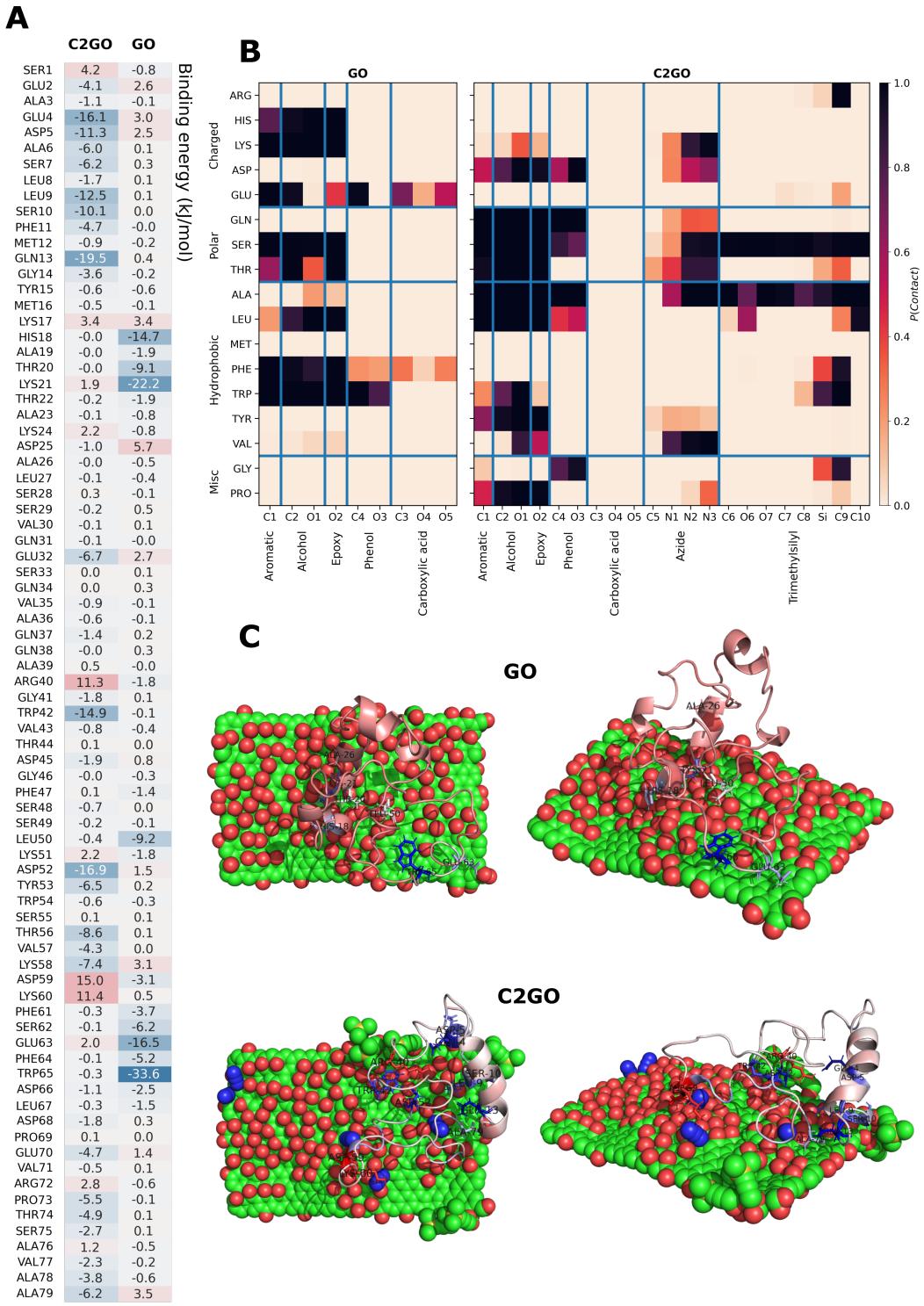


Figure 1: MM-PBSA binding energy contributions per apo-c3 amino acid residue for GO and C2GO sheets, colour coded by magnitude (A), heat map showing contact probability of apo-c3 amino acid type with GO and C2GO functional group atoms (B) and adsorbed apo-c3 structure on GO and C2GO, protein amino acids at the graphitic interface are coloured by MM-PBSA binding energy contribution (C).

graphitic surface and oxygen containing functional groups

Contact heatmaps may not distinguish residues that happen to be in close proximity to graphitic functional group atoms from proximal driving forces of adsorption. Binding free energy calculations in the form of MM-PBSA calculations are required to complement contact maps. Combining both binding free energies (Fig. 1A) and contact maps (Fig. 1B) we can deconvolve the affinity to proximity relationship. While TYR53 shows a high contact probability with C2GO surface functional (aromatic, alcohol, epoxy and azide) groups and ASP52 shows none (Fig. S3), the binding energy strength of TYR53 and ASP52 are -6.5 kJ/mol and -16.9 kJ/mol, respectively (Fig. 1A), suggesting ASP52 plays a fundamental role in the binding of TYR53 to the C2GO surface. By extension, contact maps can complement MM-PBSA results; TRP42 has a high binding affinity of -14.9 kJ/mol to C2GO, which is translated to a high contact probability of ALA39-TRP42 residues, specifically to the C2GO TMS group.

2.2 Protein structure

Changes to protein structure during adsorption — leading to protein denaturing or deformation — are evaluated using secondary structure, intramolecular hydrogen bonding and solvent-accessible surface area (SASA) analyses. Protein native contacts indicate changes in secondary structure due to adsorption, their temporal evolution is calculated using the define secondary structure of proteins (DSSP) algorithm⁷ and protein intra-molecular hydrogen bonds.

2.2.1 DSSP and hydrogen bonds

To understand the dynamic progression of denaturing in GO-adsorbed apo-c3, DSSP and hydrogen bond analyses indicate the sequential loss of secondary structure and number of intra-molecular hydrogen bonds respectively. The number of protein intramolecular hydro-

gen bonds over time (Fig. 2B) complement the DSSP results (Fig. 2A) at a higher resolution and are invariant to the DSSP algorithm criteria for defining strands, helices or coils. It shows that over time, unlike C2GO-adsorbed apo-c3, GO-adsorbed apo-c3 has a dynamic character, sporadically gaining or losing hydrogen bonds. These high spatio-temporal hydrogen bond frequencies stabilise intrachain contacts by recovering the loss of hydrogen bonds, indicating structural reorganisation is taking place during adsorption.

Loss of secondary structure in C2GO-adsorbed apo-c3 is transient, with temporary loss of α -helix character spanning residues 30-40 (Fig. 2A). In contrast, GO-adsorbed apo-c3 α -helix character is lost irreversibly (Fig. 2B), concomitantly with the formation of β -turns whose lifetime range tens to hundreds of ns and persist throughout adsorption. A decrease of α -helices accompanied by an increase of β -strands elsewhere in the protein is a characteristic observed in experimental secondary-structure analysis of denatured protein libraries using vacuum-ultraviolet circular dichroism.³⁴ Within the initial 150 ns of adsorption, α -helices spanning residues 1-10, 30-33 and 40-50 transition to coils (Fig. 2A). Subsequently at 200 ns, the C-terminus loses α -helix character in residues 55-65, the lost hydrogen bonds (Fig. 2B) are recovered through forming β -turn motifs (Fig. 2A) and stabilising the apo-c3 protein backbone (Fig. 2C) for the remainder of adsorption until 400 ns. β -turns remain stable in a solvated state, potentially serving as recognition motifs of protein-protein interaction,³⁵ until binding takes place with other proteins. Both L4-S7 and W54-T56 transitioned from an alpha-helix to type-I and type-II β -turns, respectively, whereas S48-K51 transitioned from a coil to a type II β -turn.

2.3 Clustering

UMAP dimensionality reduction is a useful approach to map high dimensional data of MD protein backbone coordinates trajectories to a lower dimensional protein configuration space. This is due to the better performance of UMAP in preserving local and global structure rela-

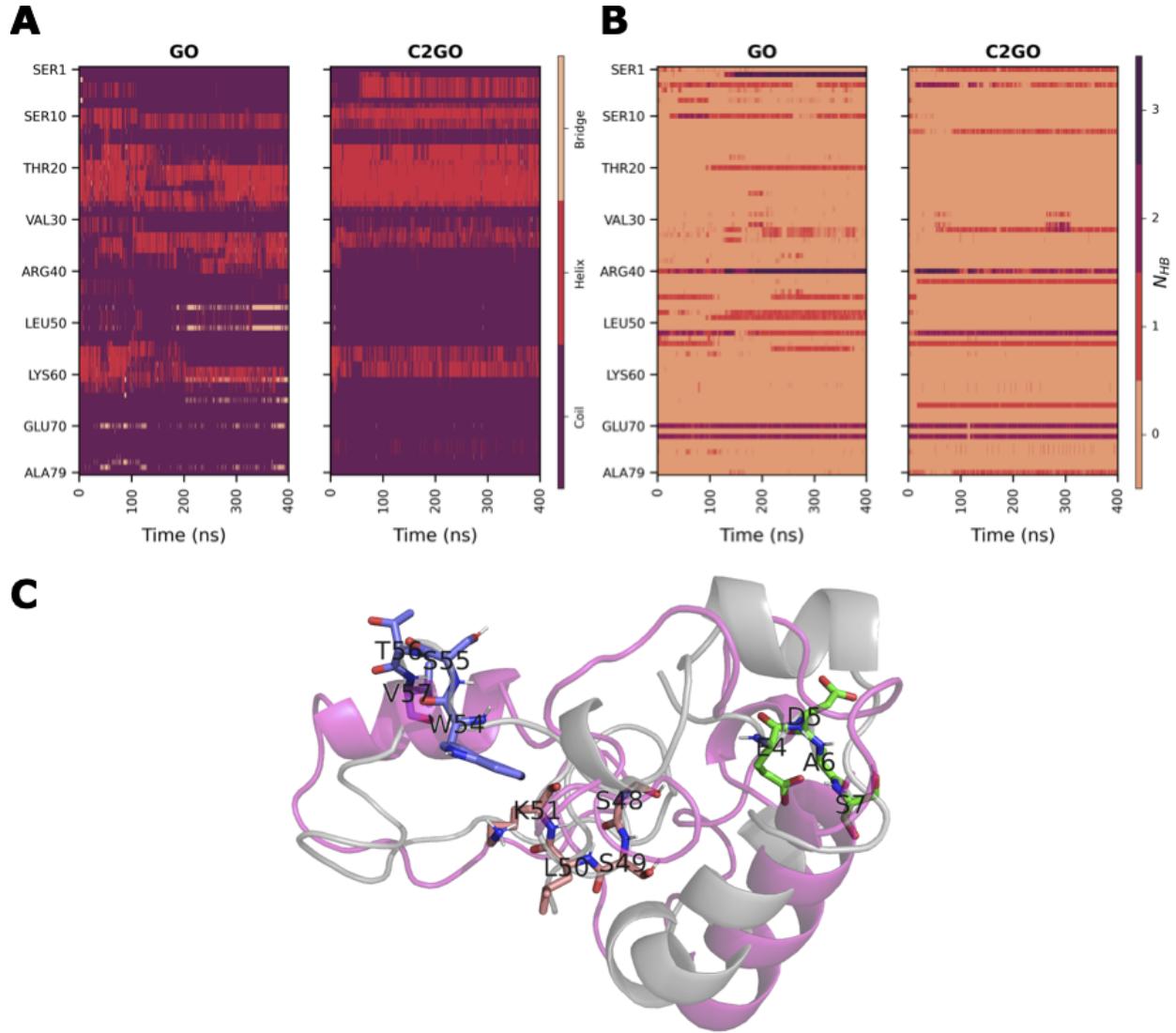


Figure 2: Define Secondary Structure of Proteins (DSSP) algorithm applied to the MD trajectory of apo-c3 adsorption with GO and C2GO (A), the number of intramolecular hydrogen bonds per amino acid residue throughout adsorption (B) and illustration of β -turns induced in GO-adsorbed apo-c3 following denaturing (residues E4-S7, S48-K51 and W54-T56), pink and grey structures respectively correspond to the initial and final adsorption conformations (C).

tionships compared with conventional dimensionality reduction techniques such as principle component analysis (PCA) or t-distributed stochastic neighbour embedding (t-SNE). A projection of the high dimensional spatio-temporal MD data into a low-dimensional space is used to identify distinct protein structures, bridging pathways and global relationships between distinct configurations. It is also a practical tool for visualising what is otherwise a vast dataset.

The number of clusters in the protein configuration space reflects the number of ensembles the protein backbone adopts, namely the distinct structures of apo-c3 during the entire adsorption process. Using this and the visualisation of protein structures corresponding to each cluster, we can see that apo-c3 has higher structural variance when adsorbing on GO than on C2GO, which have 13 and 10 clusters, respectively (Fig. 3). GO-adsorbed apo-c3 clusters show how the protein sequentially undergoes a process of reorganisation (Fig. 3) driven by interchain hydrogen bonding interactions (Fig. 2B) due to the binding interactions with the GO surface (Fig. 1A).

The temporal evolution of apo-c3 adsorption follows the cluster numbers in the protein configuration space (Fig. 3). The pathways between clusters are synonymous with the DSSP analysis, corresponding to transitions between configuration states where apo-c3 secondary structure either gains or loses α -helix, β -bridge or coil character. In both cases of GO and C2GO adsorption, the reorganisation process requires transitions to discrete intermediary states (clusters) before converging to a stable complex (Fig. 3). The stable complex may or may not recover its secondary structure following these transitions, corresponding to whether it has or has not been denatured via the adsorption process. This is indeed the case in GO-adsorbed apo-c3, which has undergone drastic structural changes and retains some of its N-terminus (Fig. 3) but loses the C-terminus secondary structure. In contrast C2GO-adsorbed apo-c3 does recover most of its secondary structure, where different cluster

pathways transition between the globally highly conserved native backbone (Fig. 3), which is indicative of binding deformations instead of protein denaturing. Note that the images of the protein backbone annotating the clusters only approximate secondary structure character and are not as accurate as the characterisation of the state-of-the-art secondary structure prediction of DSSP analysis (Fig. 2).

2.4 Solvent exposure

As a result of structural changes induced by adsorption, hydration apo-c3 amino acid residues change and contribute to the role of potential protein aggregation *in vivo*.

As the structure of apo-c3 changes during adsorption on GO, solvent exposure is increased significantly over multiple regions of the amino acid sequence (Fig. 4A). Some residues will have a larger solvent-accessible surface area (SASA) after adsorption due to denaturation/change in conformation. Some residue will have a smaller SASA after adsorption either due to a change in conformation leading to new protein-protein contacts or due to interaction of the residue with the nanomaterial in question. There is a correlation with increased solvent exposure in regions where apo-c3 forms β -turns, as analysed by DSSP and illustrated in (Fig. 2). The β -turns remain in this solvated state, potentially serving as recognition motifs for protein-protein interactions until binding to other proteins *in vivo* or *in vitro*.³⁵ As well as increased solvent exposure of the GO-adsorbed apo-c3 C-terminal helix (residues 47-65), the EVRPTSAVAA minimotif (residues 70-79) has a consistently higher solvent exposure than the native state of apo-c3 in solution (Fig. 4A,B). These exposed hydrophobic residues keep the adsorbed complex in a disordered state until it can coalesce with its protein/lipid environment. In contrast, the solvent exposure of C2GO is limited to isolated short sequences leaving most of the C-terminal helix region (residues 47-65) intact (Fig. 4A).

Previous work has found that the C-terminal region of apo-c3 is essential for mediating

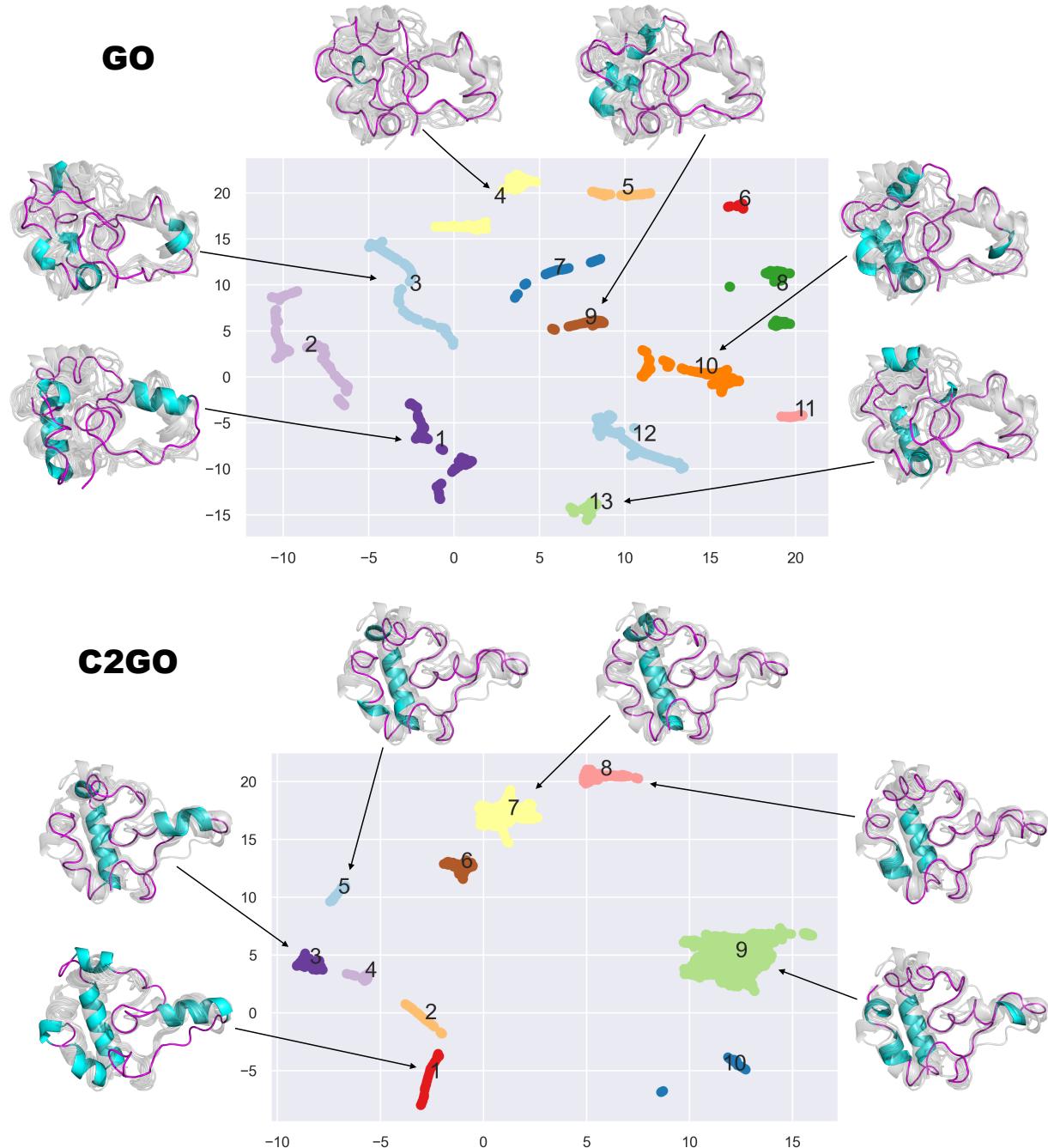


Figure 3: Uniform Manifold Approximation and Projection (t-SNE) dimensionality reduction of protein backbone denaturing during adsorption on GO (top) and C2GO (bottom) nanosheets. Separate clusters show clear separation of distinct protein backbone secondary structures. Protein structures corresponding to each cluster are coloured by secondary structure (helices in blue, loops in pink) on top of an overlay of all cluster conformations (grey).

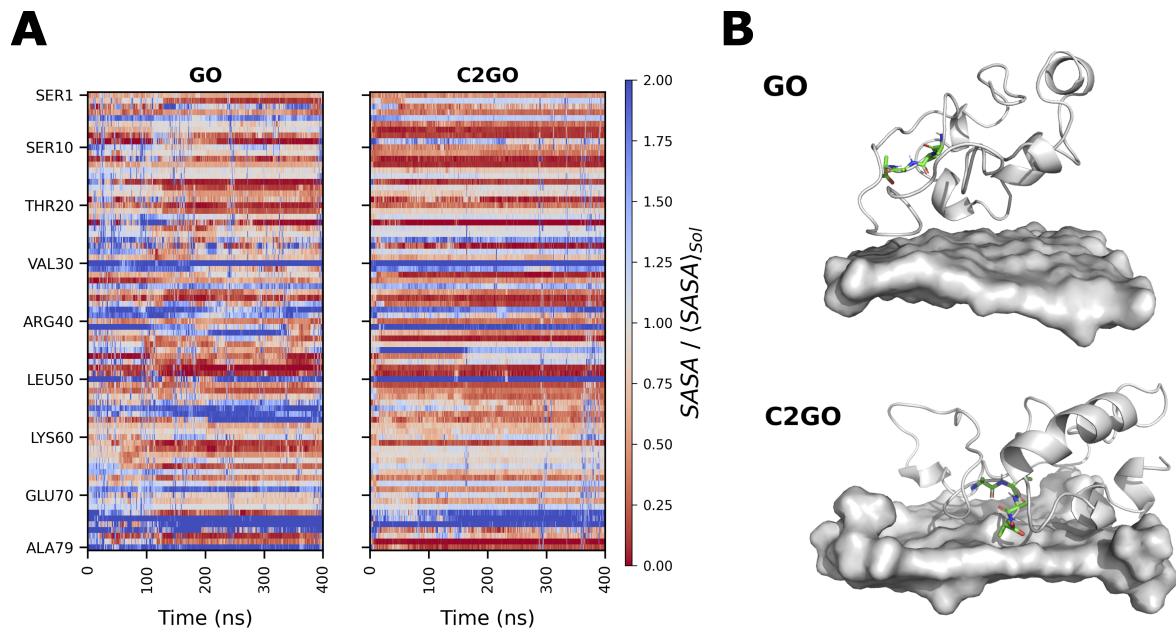


Figure 4: The surface accessible surface area (SASA) of apo-c3 amino acid residues during adsorption to GO and C2GO sheets, normalised by average SASA of apo-c3 in solution (A) and illustration of exposure of the AVAA minimotif in the C-terminal region of apo-c3 to solvent in GO adsorption and contrasting structure in C2GO adsorption, the nanomaterials are represented as a surface for clarity (B).

lipid binding, with the N-terminal region (residues 1-40) playing a limited secondary role.³⁶⁻³⁸ Solvent exposure analysis of the C-terminal helix expresses the conservation of its lipid-binding function from a physio-chemical perspective. These results may in part delineate the experimentally observed preferential uptake of corona-coated C2GO to corona-coated GO by J774 cells.¹⁵

3 Conclusions

This work has dissected the protein structural changes induced by interactions of functional groups and protein residues on the bio-nano interface. Through adsorption on GO, apo-c3 shows denaturing of the secondary structure over large swathes of the protein sequence. Whereas C2GO-adsorbed apo-c3 has limited deformation, with most variance displayed in the N-terminus. The C-terminus of apo-c3 has previously been found to be responsible for lipid binding, with the N-terminus playing a limited secondary role.³⁶⁻³⁸ These findings correlate with experimental evidence of the increased cellular uptake of C2GO to GO hard protein corona.¹⁵

Apo-c3 is found to denature upon adsorption to GO, according to solvent exposure, DSSP and conformational clustering analyses. Following adsorption on GO, apo-c3 forms 3 separate β -turns that serve as binding motifs for protein-protein interactions. These can aid in subsequent aggregation of serum proteins to the corona, contributing to the corona identity in vivo, hence complicating the targetability of the nanomaterial-corona complex.

The functional groups at the binding interface between the nanomaterials and apo-c3 set off a series of dynamic events that result in large-scale secondary structure changes. Contact distances, heatmaps and binding free energy calculations collectively describe the driving forces of the changes in protein structure following adsorption. The introduction

of azide and TMS functional groups was sufficient to stop the denaturing of apo-c3 and the formation of β -turns that serve as protein-protein interaction binding motifs. We find that contact with azide functional groups correlate strongly with contact to other surface functional groups and therefore work cooperatively to maximise the binding surface at the bio-nano interface. Consequently, the C2GO-adsorbed protein is stabilised and is not pushed to structural deformation as is the case in GO-adsorbed protein. The results from this analysis pipeline suggest the causes of protein aggregation and cellular uptake are mediated by the aforementioned changes in protein structure.

This work indicates that the dynamic and functional role of adsorbed proteins may be a more important probe to understanding the protein corona, rather than quantifying a static image of its constituent proteins. MD is a valuable tool for such an investigation and our analysis pipeline serves as a transferable method for understanding the structure/function relationship of dynamic protein-nanomaterial adsorption.

4 Methods

4.1 Nanomaterial structure

We have modified a Python package³⁹ to generate azide and trimethylsilyl (TMS)-alkyne functional group conjugated rectangular graphene-oxide flakes.⁴⁰ The package improves on the commonly applied protocol, of randomly placing oxidised functional groups, which we now know is an incomplete model of graphene-oxide structure.^{32,33} Instead, we recreate the two-phase nature of oxidised and unoxidised graphene domains observed in microscopy experiments^{41–44} in accordance with our recent study of accurate large scale modelling of graphene oxide.³³

4.2 Molecular Dynamics

MD simulations were performed in GROMACS version 2020.1 on the ARCHER2 AMD EPYC Zen2 (Rome) 64 core CPUs at 2.2GHz or Nvidia V100 GPUs. The all-atom OPLS forcefield was used to simulate the classical simulations. A position restraint algorithm was applied to all non-solvent atoms during equilibration. The $\sim 130\text{ \AA}$ cubic simulation box was fully solvated with TIP3P water molecules. The system net charge was neutralised by adding four sodium ions into the system. The system was relaxed energetically using steepest-descent energy minimisation for 50000 steps with an energetic step size of 0.01 kJ/mol. The minimisation was terminated after the maximum energetic contribution was lower than a threshold of 1000.0 kJ/mol/nm. NVT and NPT equilibration was performed for 100 ps using two separate velocity-rescaling thermostat coupling temperature to velocities for graphene and solvent molecules (NVT), where a temperature of 300K was maintained and 1 bar using the Parrinello-Rahman barostat (NPT). The Verlet cut-off scheme was employed to generate pair lists and the electrostatic interactions were calculated using the Particle-Mesh Ewald algorithm. Both electrostatic and van der Waals interactions were cut off beyond 1.2 nm. All bonds involving hydrogen atoms were constrained using the LINCS algorithm. Production simulations were run for 400 ns using a timestep of 1 fs. Analysis was performed using the MDAnalysis package⁴⁵ and its analysis modules.^{46,47} We used the MDTraj⁴⁸ implementation of DSSP.

4.3 Contact maps

Contact maps between apo-c3 and the graphitic sheets were calculated using a hard cutoff of 5.0 \AA between any two heavy (non-hydrogen) atoms. Data over the final 50 ns were used for the contact maps, as during this period apo-c3 was stably adsorbed to both GO and C2GO. We constructed contact maps between between GO/C2GO and each specific residue of apo-c3 (Fig. S2 and S3), as well as contact maps between GO/C2GO and each amino acid regardless of its position in the sequence (Fig. 1B).

4.4 PBSA binding energies

We post process the last 10 ns of MD trajectories using MM-PBSA analysis implemented using g_mmpbsa,⁴⁹ to obtain a relative order of binding of apo-c3 to the different graphitic nanosheets. The energetic components of the binding free energy are the changes in the system potential energy *in vacuo*, the polar and non-polar solvation energies. The potential energy accounts for bonded (bond, angle and torsion) and non-bonded (van der Waals and electrostatic) energetic terms. Polar solvation energy is the electrostatic contribution to the solvation free energy and is estimated using the Poisson-Boltzmann equation. The non-electrostatic contribution to the solvation free energy accounts for forces between solute and solvent, which are calculated using the solvent accessible surface area (SASA).⁴⁹

4.5 UMAP dimensionality reduction

We used the atomic coordinates of apo-c3 to obtain an ensemble of distinct structural conformations when adsorbed to the graphitic sheets. We first centre and align the structures along their backbone atoms, using every tenth frame (50 ps) from the trajectory. We then use the uniform manifold approximation and projection for dimension reduction (UMAP⁵⁰) algorithm to embed the atomic coordinates of heavy atoms of apo-c3 into a two-dimensional space. UMAP constructs a graph of the points in the high-dimensional space, then optimises a low-dimensional representation of the graph such that the topological distance is preserved to a degree in the embedding.⁵⁰ Therefore, similar conformations that are close in the high-dimensional space will also be close in the reduced space. We set the `n_neighbours` and `min_dist` hyperparameters to 15 and 0.0, respectively, with the latter being a requirement if the points in the reduced space are to be clustered. We use HDBSCAN,⁵¹ implemented in scikit-learn,⁵² to cluster the apo-c3 conformations in the embedded space. We identify representative structures of each conformation by calculating the mean structure of a conformation, then finding the structure with the smallest RSMD from this mean structure.

4.6 SASA solvent exposure

To quantify the solvent exposure of amino acids during adsorption, we calculated the solvent accessible surface area of each residue of apo-c3 over time. To understand how the conformational changes lead to exposure of residues that are buried in the native state, we normalised the SASA time-series by the mean SASA of each residue of the protein in solution. We used the final 50 ns of apo-c3 solution for determining the mean SASA of each residue. Values greater than and less than 1 indicate, respectively, increased and decreased exposure to the solvent (Fig. 4). The SASA was calculated using the “rolling-ball” Shake-Rupley algorithm,⁵³ implemented in MDTraj.⁴⁸ This approximates the SASA by effectively rolling a ball over each atom of the protein to define a surface, then examining how much of the each atom’s surface is exposed to solvent as opposed to overlapping with surfaces of neighbouring atoms. The radius of the probe is typically set to 1.4 Å — approximately the radius of a water molecule.

Acknowledgement

The authors acknowledge Kuo-Ching Mei for helpful discussions. This work was supported by BBSRC (grant BB/M009513/1) and EPSRC (grant EP/N509498/1). Via our membership of the UK's HEC Materials Chemistry Consortium, which is funded by EPSRC (EP/L000202, EP/R029431), this work made use of the ARCHER2 UK National Supercomputing Service (<http://www.archer2.ac.uk>). This work was supported by The Alan Turing Institute under EPSRC grant EP/N510129/1, through the use of time on Tier 2 HPC facility JADE.

References

- (1) Yuan, Y.; Gao, X.; Wei, Y.; Wang, X.; Wang, J.; Zhang, Y.; Gao, C. Enhanced desalination performance of carboxyl functionalized graphene oxide nanofiltration membranes. *Desalination* **2017**, 405, 29–39.
- (2) Zhu, Y.; Murali, S.; Cai, W.; Li, X.; Suk, J. W.; Potts, J. R.; Ruoff, R. S. Graphene and graphene oxide: synthesis, properties, and applications. *Advanced materials* **2010**, 22, 3906–3924.
- (3) Chung, C.; Kim, Y.-K.; Shin, D.; Ryoo, S.-R.; Hong, B. H.; Min, D.-H. Biomedical applications of graphene and graphene oxide. *Accounts of chemical research* **2013**, 46, 2211–2224.
- (4) Bonanni, A.; Chua, C. K.; Zhao, G.; Sofer, Z.; Pumera, M. Inherently electroactive graphene oxide nanoplatelets as labels for single nucleotide polymorphism detection. *ACS nano* **2012**, 6, 8546–8551.
- (5) Heerema, S. J.; Dekker, C. Graphene nanodevices for DNA sequencing. *Nature nanotechnology* **2016**, 11, 127–136.
- (6) Afsahi, S.; Lerner, M. B.; Goldstein, J. M.; Lee, J.; Tang, X.; Bagarozzi Jr, D. A.;

- Pan, D.; Locascio, L.; Walker, A.; Barron, F., et al. Novel graphene-based biosensor for early detection of Zika virus infection. *Biosensors and Bioelectronics* **2018**, 100, 85–88.
- (7) Robinson, J. T.; Tabakman, S. M.; Liang, Y.; Wang, H.; Sanchez Casalongue, H.; Vinh, D.; Dai, H. Ultrasmall reduced graphene oxide with high near-infrared absorbance for photothermal therapy. *Journal of the American Chemical Society* **2011**, 133, 6825–6831.
- (8) Liu, J.; Cui, L.; Losic, D. Graphene and graphene oxide as new nanocarriers for drug delivery applications. *Acta biomaterialia* **2013**, 9, 9243–9257.
- (9) Sun, X.; Liu, Z.; Welsher, K.; Robinson, J. T.; Goodwin, A.; Zaric, S.; Dai, H. Nano-graphene oxide for cellular imaging and drug delivery. *Nano research* **2008**, 1, 203–212.
- (10) Zhang, L.; Xia, J.; Zhao, Q.; Liu, L.; Zhang, Z. Functional graphene oxide as a nanocarrier for controlled loading and targeted delivery of mixed anticancer drugs. *small* **2010**, 6, 537–544.
- (11) Casals, E.; Pfaller, T.; Duschl, A.; Oostingh, G. J.; Puntes, V. Time evolution of the nanoparticle protein corona. *ACS nano* **2010**, 4, 3623–3632.
- (12) Ke, P. C.; Lin, S.; Parak, W. J.; Davis, T. P.; Caruso, F. A decade of the protein corona. *ACS nano* **2017**, 11, 11773–11776.
- (13) Röcker, C.; Pötzl, M.; Zhang, F.; Parak, W. J.; Nienhaus, G. U. A quantitative fluorescence study of protein monolayer formation on colloidal nanoparticles. *Nature nanotechnology* **2009**, 4, 577–580.
- (14) Nierenberg, D.; Khaled, A. R.; Flores, O. Formation of a protein corona influences the biological identity of nanomaterials. *Reports of Practical Oncology & Radiotherapy* **2018**, 23, 300–308.

- (15) Mei, K.-C.; Ghazaryan, A.; Teoh, E. Z.; Summers, H. D.; Li, Y.; Ballesteros, B.; Piasecka, J.; Walters, A.; Hider, R. C.; Mailänder, V., et al. Protein-Corona-by-Design in 2D: A Reliable Platform to Decode Bio–Nano Interactions for the Next-Generation Quality-by-Design Nanomedicines. *Advanced Materials* **2018**, 30, 1802732.
- (16) Solorio-Rodríguez, A.; Escamilla-Rivera, V.; Uribe-Ramírez, M.; Chagolla, A.; Winkler, R.; García-Cuellar, C.; De Vizcaya-Ruiz, A. A comparison of the human and mouse protein corona profiles of functionalized SiO₂ nanocarriers. *Nanoscale* **2017**, 9, 13651–13660.
- (17) Hajipour, M. J.; Raheb, J.; Akhavan, O.; Arjmand, S.; Mashinchian, O.; Rahman, M.; Abdolahad, M.; Serpooshan, V.; Laurent, S.; Mahmoudi, M. Personalized disease-specific protein corona influences the therapeutic impact of graphene oxide. *Nanoscale* **2015**, 7, 8978–8994.
- (18) Papi, M.; Palmieri, V.; Digiacomo, L.; Giulimondi, F.; Palchetti, S.; Ciasca, G.; Perini, G.; Caputo, D.; Cartillone, M. C.; Cascone, C., et al. Converting the personalized biomolecular corona of graphene oxide nanoflakes into a high-throughput diagnostic test for early cancer detection. *Nanoscale* **2019**, 11, 15339–15346.
- (19) Rampado, R.; Crotti, S.; Caliceti, P.; Pucciarelli, S.; Agostini, M. Recent Advances in Understanding the Protein Corona of Nanoparticles and in the Formulation of “Stealthy” Nanomaterials. *Frontiers in Bioengineering and Biotechnology* **2020**, 8.
- (20) Mei, K.-C.; Rubio, N.; Costa, P. M.; Kafa, H.; Abbate, V.; Festy, F.; Bansal, S. S.; Hider, R. C.; Al-Jamal, K. T. Synthesis of double-clickable functionalised graphene oxide for biological applications. *Chemical Communications* **2015**, 51, 14981–14984.
- (21) Rubio, N.; Mei, K.-C.; Klippstein, R.; Costa, P. M.; Hodgins, N.; Wang, J. T.-W.; Festy, F.; Abbate, V.; Hider, R. C.; Chan, K. L. A., et al. Solvent-free click-

- mechanochemistry for the preparation of cancer cell targeting graphene oxide. ACS applied materials & interfaces **2015**, 7, 18920–18923.
- (22) Vroman, L.; Adams, A.; Fischer, G.; Munoz, P. Interaction of high molecular weight kininogen, factor XII, and fibrinogen in plasma at interfaces. **1980**,
- (23) Foroozandeh, P.; Aziz, A. A. Merging worlds of nanomaterials and biological environment: factors governing protein corona formation on nanoparticles and its biological consequences. Nanoscale Research Letters **2015**, 10, 1–12.
- (24) Ehrenberg, M. S.; Friedman, A. E.; Finkelstein, J. N.; Oberdörster, G.; McGrath, J. L. The influence of protein adsorption on nanoparticle association with cultured endothelial cells. Biomaterials **2009**, 30, 603–610.
- (25) Harnisch, S.; Müller, R. H. Adsorption kinetics of plasma proteins on oil-in-water emulsions for parenteral nutrition. European journal of pharmaceutics and biopharmaceutics **2000**, 49, 41–46.
- (26) Göppert, T. M.; Müller, R. H. Polysorbate-stabilized solid lipid nanoparticles as colloidal carriers for intravenous targeting of drugs to the brain: comparison of plasma protein adsorption patterns. Journal of drug targeting **2005**, 13, 179–187.
- (27) Radic, S.; Geitner, N. K.; Podila, R.; Käkinen, A.; Chen, P.; Ke, P. C.; Ding, F. Competitive binding of natural amphiphiles with graphene derivatives. Scientific reports **2013**, 3, 1–8.
- (28) Baweja, L.; Balamurugan, K.; Subramanian, V.; Dhawan, A. Effect of graphene oxide on the conformational transitions of amyloid beta peptide: A molecular dynamics simulation study. Journal of Molecular Graphics and Modelling **2015**, 61, 175–185.
- (29) Baweja, L.; Balamurugan, K.; Subramanian, V.; Dhawan, A. Hydration patterns of

- graphene-based nanomaterials (GBNMs) play a major role in the stability of a helical protein: a molecular dynamics simulation study. *Langmuir* **2013**, 29, 14230–14238.
- (30) Sun, X.; Feng, Z.; Hou, T.; Li, Y. Mechanism of Graphene Oxide as an Enzyme Inhibitor from Molecular Dynamics Simulations. *ACS Applied Materials & Interfaces* **2014**, 6, 7153–7163.
- (31) Sinclair, R. C.; Coveney, P. V. Modelling nanostructure in graphene oxide: inhomogeneity and the percolation threshold. *Journal of Chemical Information and Modeling* **2019**, 59, 2741.
- (32) Mouhat, F.; Coudert, F.-X.; Bocquet, M.-L. Structure and chemistry of graphene oxide in liquid water from first principles. *Nature communications* **2020**, 11, 1–9.
- (33) al Badri, M. A.; Smith, P.; Sinclair, R. C.; al Jamal, K. T.; Lorenz, C. D. Accurate large scale modelling of Graphene Oxide: ion trapping and chaotropic potential at the interface. *Carbon* **2020**,
- (34) Matsuo, K.; Sakurada, Y.; Yonehara, R.; Kataoka, M.; Gekko, K. Secondary-structure analysis of denatured proteins by vacuum-ultraviolet circular dichroism spectroscopy. *Biophysical journal* **2007**, 92, 4088–4096.
- (35) Tyndall, J. D.; Pfeiffer, B.; Abbenante, G.; Fairlie, D. P. Over one hundred peptide-activated G protein-coupled receptors recognize ligands with turn structure. *Chemical reviews* **2005**, 105, 793–826.
- (36) Sparrow, J. T.; Pownall, H. J.; Hsu, F.-J.; Blumenthal, L. D.; Culwell, A. R.; Gotto, A. M. Lipid binding by fragments of apolipoprotein C-III-1 obtained by thrombin cleavage. *Biochemistry* **1977**, 16, 5427–5431.
- (37) Meyers, N. L.; Larsson, M.; Vorrsjö, E.; Olivecrona, G.; Small, D. M. Aromatic residues

in the C terminus of apolipoprotein C-III mediate lipid binding and LPL inhibition. *Journal of lipid research* **2017**, 58, 840–852.

- (38) Lambert, D. A.; Catapano, A. L.; Smith, L. C.; Sparrow, J. T.; Gotto Jr, A. M. Effect of the apolipoprotein C-IIIC-III1 ratio on the capacity of purified milk lipoprotein lipase to hydrolyse triglycerides in monolayer vesicles. *Atherosclerosis* **1996**, 127, 205–212.
- (39) Sinclair, R. C. , 2020; <https://github.com/velocirobbie/make-graphitics>, Accessed: March 5, 2021.
- (40) al Badri, M. A.; Sinclair, R. C. <https://github.com/maalbadri/Accurate-large-scale-modelling-of-GrapheneOxide>, 2020; Accessed: March 5, 2021.
- (41) Pacilé, D.; Meyer, J.; Fraile Rodríguez, A.; Papagno, M.; Gómez-Navarro, C.; Sundaram, R.; Burghard, M.; Kern, K.; Carbone, C.; Kaiser, U. Electronic properties and atomic structure of graphene oxide membranes. *Carbon* **2011**, 49, 966–972.
- (42) Cai, W.; Piner, R. D.; Stadermann, F. J.; Park, S.; Shaibat, M. A.; Ishii, Y.; Yang, D.; Velamakanni, A.; An, S. J.; Stoller, M.; An, J.; Chen, D.; Ruoff, R. S. Synthesis and solid-state NMR structural characterization of ¹³C-labeled graphite oxide. *Science (New York, N.Y.)* **2008**, 321, 1815–7.
- (43) Saxena, S.; Tyson, T. A.; Negusse, E. Investigation of the Local Structure of Graphene Oxide. *The Journal of Physical Chemistry Letters* **2010**, 1, 3433–3437.
- (44) Erickson, K.; Erni, R.; Lee, Z.; Alem, N.; Gannett, W.; Zettl, A. Determination of the Local Chemical Structure of Graphene Oxide and Reduced Graphene Oxide. *Advanced Materials* **2010**, 22, 4467–4472.
- (45) Gowers, R. J.; Linke, M.; Barnoud, J.; Reddy, T. J. E.; Melo, M. N.; Seyler, S. L.;

- Domanski, J.; Dotson, D. L.; Buchoux, S.; Kenney, I. M., et al. MDAnalysis: a Python package for the rapid analysis of molecular dynamics simulations; 2019.
- (46) Araya-Secchi, R.; Perez-Acle, T.; Kang, S.-g.; Huynh, T.; Bernardin, A.; Escalona, Y.; Garate, J.-A.; Martínez, A. D.; García, I. E.; Sáez, J. C., et al. Characterization of a novel water pocket inside the human Cx26 hemichannel structure. Biophysical journal **2014**, 107, 599–612.
- (47) Smith, P.; Ziolek, R. M.; Gazzarrini, E.; Owen, D. M.; Lorenz, C. D. On the interaction of hyaluronic acid with synovial fluid lipid membranes. Physical Chemistry Chemical Physics **2019**, 21, 9845–9857.
- (48) McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernández, C. X.; Schwantes, C. R.; Wang, L.-P.; Lane, T. J.; Pande, V. S. MD-Traj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. Biophysical Journal **2015**, 109, 1528 – 1532.
- (49) Kumari, R.; Kumar, R.; Consortium, O. S. D. D.; Lynn, A. g_mmmpbsa A GROMACS tool for high-throughput MM-PBSA calculations. Journal of chemical information and modeling **2014**, 54, 1951–1962.
- (50) McInnes, L.; Healy, J.; Saul, N.; Grossberger, L. UMAP: Uniform Manifold Approximation and Projection. The Journal of Open Source Software **2018**, 3, 861.
- (51) McInnes, L.; Healy, J.; Astels, S. hdbscan: Hierarchical density based clustering. The Journal of Open Source Software **2017**, 2, 205.
- (52) Pedregosa, F. et al. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research **2011**, 12, 2825–2830.
- (53) Shrake, A.; Rupley, J. Environment and exposure to solvent of protein atoms. Lysozyme and insulin. Journal of Molecular Biology **1973**, 79, 351–371.