Prediction of nanographene binding-scores to trout

Connolly, M. C., Navas, J. M. and Coll, J.

Department of Environment. Instituto Nacional de Investigaciones y Tecnología Agraria e Alimentaria, INIA. Madrid, Spain.

Abstract:

To address the increasing concerns surrounding possible impacts of graphene-related materials on the aquatic environment, this study focused on computational predictions of binding between models of graphenes in the nm size range (nanographenes, nGs) and the aryl-hydrocarbon receptor (AHR) and P450 cytochromes (ICYPs) of rainbow trout (Oncorhynchus mykiss). The AHR plays a key role in the induction of detoxifying and early immune responses and ICYPs are essential for detoxifying planar hydrophobic chemicals such as nGs. After 3D modelling of those trout proteins, docking algorithms predicted the size-dependance profiles of nGs binding-scores to tAHR and ICYPs in the low nM range (high binding-affinities). Virtual oxidations of nGs to nGOs (carboxy-, epoxy- and/or hydroxy-oxidations) further lowered the corresponding binding-scores in level/type-oxidation manners. Among all the ICYPs, the ICYP3A4 (the equivalent to human CYP3A4) was identified as a potential key interaction enzyme for nGs because of its lower binding-scores. These results implicate a possible processing pathway to be further probed through in vitro and in vivo experimentation. Together the information generated can be pivotal for the design of safer graphene-related materials for a variety of applications and help to understand their detoxification in aquatic vertebrates.

Introduction

Graphene-related materials (GRMs) have potential applications in a variety of human activities, from aeronautics to biomedicine, to electronics and in agriculture. This is causing a constant increase in the production of GRMs that will likely lead to releases to the environment during production, use and/or disposal. Although the current concentrations and toxicological hazards of GRMs in the aquatic media are yet minimal, they are expected to increase with a rise of graphene production. Usually, graphene is not used in its pristine form but with a certain degree of oxidations (GOs). GOs have been shown to interact with other environmental pollutants, such as polyaromatic hydrocarbons (PAHs) causing an increase in cytotoxicity as estimated in fish cell lines. Once in the aquatic environment, residues of GOs will likely interact with a variety of biota but there is only a limited knowledge of the bio-reactions they might induce. Among these bio-reactions, perturbations in detoxifying mechanisms and immune responses would have important impacts on vertebrates. To better understand the possible interactions of graphene-related materials with GOs in aquatic systems.

Graphene-related materials (GRMs) have potential applications in a variety of human activities, from aeronautics to biomedicine, to electronics and in agriculture. This is causing a constant increase in the production of GRMs that will likely lead to releases to the environment during production, use and/or disposal. Although the current concentrations and toxicological hazards of GRMs in the aquatic media are yet minimal, they are expected to increase with a rise of graphene production.

Abstract:

To address the increasing concerns surrounding possible impacts of graphene-related materials on the aquatic environment, this study focused on computational predictions of binding between models of graphenes in the nm size range (nanographenes, nGs) and the aryl-hydrocarbon receptor (AHR) and P450 cytochromes (ICYPs) of rainbow trout (Oncorhynchus mykiss). The AHR plays a key role in the induction of detoxifying and early immune responses and ICYPs are essential for detoxifying planar hydrophobic chemicals such as nGs. After 3D modelling of those trout proteins, docking algorithms predicted the size-dependance profiles of nGs binding-scores to tAHR and ICYPs in the low nM range (high binding-affinities). Virtual oxidations of nGs to nGOs (carboxy-, epoxy- and/or hydroxy-oxidations) further lowered the corresponding binding-scores in level/type-oxidation manners. Among all the ICYPs, the ICYP3A4 (the equivalent to human CYP3A4) was identified as a potential key interaction enzyme for nGs because of its lower binding-scores. These results implicate a possible processing pathway to be further probed through in vitro and in vivo experimentation. Together the information generated can be pivotal for the design of safer graphene-related materials for a variety of applications and help to understand their detoxification in aquatic vertebrates.

Introduction

Graphene-related materials (GRMs) have potential applications in a variety of human activities, from aeronautics to biomedicine, to electronics and in agriculture. This is causing a constant increase in the production of GRMs that will likely lead to releases to the environment during production, use and/or disposal. Although the current concentrations and toxicological hazards of GRMs in the aquatic media are yet minimal, they are expected to increase with a rise of graphene production.

Although the current concentrations and toxicological hazards of GRMs in the aquatic media are yet minimal, they are expected to increase with a rise of graphene production. Usually, graphene is not used in its pristine form but with a certain degree of oxidations (GOs). GOs have been shown to interact with other environmental pollutants, such as polyaromatic hydrocarbons (PAHs) causing an increase in cytotoxicity as estimated in fish cell lines. Once in the aquatic environment, residues of GOs will likely interact with a variety of biota but there is only a limited knowledge of the bio-reactions they might induce. Among these bio-reactions, perturbations in detoxifying mechanisms and immune responses would have important impacts on vertebrates. To better understand the possible interactions of graphene-related materials with GOs in aquatic systems.

Graphene-related materials (GRMs) have potential applications in a variety of human activities, from aeronautics to biomedicine, to electronics and in agriculture. This is causing a constant increase in the production of GRMs that will likely lead to releases to the environment during production, use and/or disposal. Although the current concentrations and toxicological hazards of GRMs in the aquatic media are yet minimal, they are expected to increase with a rise of graphene production.
Materials and Methods

Nanographene ligands

Graphene-like ligands were obtained by a similarity search in PubMed (https://pubchem.ncbi.nlm.nih.gov/) using C1=CH2C=C(=C(CH3))2 as query. Among the 8 nanographene-like molecules (SMILES C1=CC2=C3C4=CC=C5C6=C7C=C8C=C(C=C(C=C1)C2=C8C7=C3C4=C5) corresponding to Zinc ID 109432703 (https://zinc15.docking.org/). The corresponding downloaded sdf file manually curated, contained 86 small molecular weight graphene-like compounds.

To mimic graphene platelet degradation products, pdb models of nanographene (called nGs here, because of their 10–20 nm lateral side range) were designed and obtained using the open-source Phython GOPY tool (https://github.com/Iourarum/GOPY). The nGs were obtained in different sizes (4 to 156 rings), and with different levels of oxidation (carboxylation-COOH, epoxy oxidation -O, hydroxylation -OH and/or combinations). Oxygen atoms were added to a maximum of ±1.6 Carbon/Oxygen C/O ratio, similar to those described in previous experimental work. While carboxylation can occur only at the C edges, both epoxidation and hydroxylation may occur at each of the CIs inside the rings to be distributed randomly between the two faces of the planar nG sheet (see some depictions in Section S3). In all cases, steric hindrance may reduce the number of theoretical possibilities. Ten random designs for each of the oxidation scenarios were computationally constructed for each oxidation state by modifications of the Phython GOPY tool, which took into account all the requirements mentioned above. The resulting nanographene structures visualized in PyMOL showed a variety of numbers and intramolecular locations approximating the input numbers provided to the GOPY tool because of steric restrictions. Additionally, docking analysis rejected 10-20% of the structures specifically when 2 oxidations were too close in the same ring.

Tridimensional trout AHR and CYP 3D models

The trout aHR (tAHR) was modeled using the SWISS-MODEL homology modeling (https://swissmodel.expasy.org/interactive). Structural similarities were expressed in Angstroms A / number of common atoms, as estimated by the similarity index of tAHR and hAHR proteins including those from similar transcription factors. Among them, 6 models of tAHR were submitted to the SWISS-MODEL homology modeling, which automatically selected 7 possible templates with the highest sequence identity (Table S1). The hAHR-1 (highlighted grey background) was selected as model template for tAHRa and b. tHRAa and b sequences showed 97% SI. hHRA (residues 1 to 420) were submitted to the SWISS-MODEL homology modelling, which automatically selected 7 possible templates with the highest sequence identity (Table S1). hHRA-1 (residues 107-261 of human AHR) were crystallographically solved and proteins including those from similar transcription factors. Among them, 6 models of human AHR were submitted to the SWISS-MODEL homology modeling, which automatically selected 7 possible templates with the highest sequence identity (Table S1). The tCYPs were modeled using the open-source Phyton GOPY tool (https://github.com/Iourarum/GOPY). The nGs were obtained in different sizes (4 to 156 rings), and with different levels of oxidation (carboxylation-COOH, epoxy oxidation -O, hydroxylation -OH and/or combinations). Oxygen atoms were added to a maximum of ±1.6 Carbon/Oxygen C/O ratio, similar to those described in previous experimental work. While carboxylation can occur only at the C edges, both epoxidation and hydroxylation may occur at each of the CIs inside the rings to be distributed randomly between the two faces of the planar nG sheet (see some depictions in Section S3). In all cases, steric hindrance may reduce the number of theoretical possibilities. Ten random designs for each of the oxidation scenarios were computationally constructed for each oxidation state by modifications of the Phython GOPY tool, which took into account all the requirements mentioned above. The resulting nanographene structures visualized in PyMOL showed a variety of numbers and intramolecular locations approximating the input numbers provided to the GOPY tool because of steric restrictions. Additionally, docking analysis rejected 10-20% of the structures specifically when 2 oxidations were too close in the same ring.

3D Modeling of trout AHR

One of the requirements for this study was to obtain a 3D model of the PAS domain of trout AHR (tAHR) based on the two amino acid sequences of hAHR (hAHR) available (97% identical) and some corresponding 3D solved structures. Since only residues 107-261 of human AHR (hAHR) were crystallographically solved and those were not enough to cover all the corresponding tAHR ligand binding domains, a template search was undertaken for a better template. Seven hAHR models were selected by the SWISS-model program using template PAS domains from diverse proteins including those from similar transcription factors. Among them, 6 models showed < 2 Å RMSD differences in their 3D structure and hAHR binding pockets. The hAHR-1 model covering the highest number of amino acids of 357 residues was selected (Table 1, highlighted in grey background) to model the two tAHRs. The resulting modeled tAHRa and b structures were very similar (Table 1, last entries).

Results

3D Modeling of trout CYPs

Because of the absence of crystalization studies that solved the structures of trout CYPs (tCYPs), their hypothetical 3D structures were modeled for the first time (Table 2, Figure S1). Despite the differences in their amino acid sequences (Table S1), all the tCYP models have similar overall 3D structures as shown by their < 3 Å RMSD. The tCYP 3D structures have several α-helices, one β-sheet, and one highly conserved domain (Table S1, red amino acids). Heme binding implicating residues ~100-140 and ~300-320 were in the pM range, as shown by seeSAR. An amino-terminal tunnel-like access (Figure S1, grey arrow) may be the main route for substrates/inhibitors to penetrate nearby the heme for optimal oxidation. A binding pocket predicted by seeSAR, was identified around the heme in all tCYPs (Figure S1, yellow pocket).

Computational binding of ng-like compounds to hAHR and tAHR

Most but not all ng-like compounds showed similar binding-scores to hAHR and tAHR despite their differences in amino acid sequences (Figure 1A). Nevertheless, their corresponding ng binding-poses were different (Figure 1B and C).

Among the ng-like compounds there was a reduction of binding-scores to hAHR and tAHR despite their differences in amino acid sequences (Figure 1A). Nevertheless, their corresponding ng binding-poses were different (Figure 1B and C).
Computational binding of nGos to tAHR

Carboxy (–COOH), epoxy (–O-) and hydroxy (–OH) oxidations were mimicked by adding them to the 9 and 25-ring nGs to randomly generate their corresponding nGos using 1.6 Carbon/Oxygen C/O ratios (see some of their structures in Figure S3). In this test, docking conditions were adjusted to yield binding-scores of ~10^11 nM and ~10^10 nM for 25- and 9-rings nGs, respectively. Results with nGos predicted 10^10–10^11 nM binding-scores for 25-rings (Figure 3, red and grey-edged circles, respectively). In particular, epoxidation (–O-) alone or in combination with carboxylation (–COOH), was the most efficient process to reduce the binding-scores of 25-ring nGos. In contrast, hydroxylation (–OH), and its combinations increased their binding-scores. Mapping of the corresponding 3D binding-poses predicted that the amino terminal end of the PAS domain of tAHR was targeted by most 25- and 9-rings nGos (Figure 3C and 3G, respectively).

Computational binding of nGos with different oxidation levels

To study the influence of the levels of nGO oxidation on the binding to tAHR, different numbers of oxygen-containing molecules were added to 25-ring nGos. Results showed that increasing the level of peroxidation (–O-) to 70 % caused a continuous reduction to 10^9 nM binding-scores (Figure 4, green circles). Increasing the level of carboxylation (–COOH) or hydroxylation (–OH) to ~25 % reduced their binding-scores to ~10^8 nM, but increased them at higher oxidation levels (Figure 4, red and blue circles). Both nG and most nGO binding-scores to tAHR remained < 10^7 nM, which may be significant compared to the 10^11 nM binding-score of the AHR prototypical agonist TCDD under the same docking conditions.

Modeling of 20x20 nm nGs resulted in nGs of 72 rings down to binding-scores of >10^-4 nM, compared to those of 25 nG or 37 circumcoronene rings which bound to the tAHR / tAHR with binding-scores in the ~ 10^-1 nM range (Figure 1A). To further investigate the influence of nG size in the binding to tAHR and CYPs, different sizes were designed and computationally tested by blind docking (PAS tAHR domain) or binding-pocket docking (CYPs).

The results predicted that the lower binding-scores to tAHR corresponded to nGs > 100-rings. Larger nGs maintained these binding-scores. Binding to CYPs showed a similar size-profiles as to tAHR, but in the 10^1–10^0 binding-score ranges (Figure 2, blue, green and red lines). Most tAHR binding to nGs mapped to its PAS ligand binding site. In contrast, most nGs binding to CYPs were localized on their surfaces away from their heme sites, suggesting some unspecific bindings (Figure S2). Only 4-5-ring nGs were mapped nearby to their heme site, most probably because only such small compounds could penetrate through the CYP tunnels to such internal location to be specifically oxidized (Figure S2).

Models of 9 (7x7 nm, 32 Carbons) and 25 (13x13 nm, 77 Carbons)-ring nGs were selected for further studies, because significant binding-scores could be obtained in minimal computational times.
Because of the nGs tendency to bind non-specifically to ICYPs (Figure S2), the size of the nGs were reduced to 5×5 nm (4 rings) and grids of 25×30 x25 Å surrounding the ICYPs heme were used for docking. Under the above mentioned conditions, the nG binding-scores varied for each ICYP molecular species from 10^2 to 10^4 nM (Figure 5, black-dashed horizontal lines). By increasing the level of nGO oxidation to ~25 % most binding-scores decreased to different minimal levels for different ICYPs. However, increasing the level of most oxidations >25 %, resulted in increased binding-scores for most ICYPs, except ICYP3AR. Thus, with >25 % oxidations, ICYP3AR binding-scores were further reduced to 10^3 by carboxylation (Figure 5, 3AR, red circles), and to a lower extent by peroxidation and hydroxylation.

**Discussion**

Lateral sizes of commercial powered graphene are of ~10 nm while the so called nanoplatelets range between 20-100 nm for oxidized graphene (GO) and between 100-1000 nm for carboxylated graphene 10. Our initial working hypothesis was that larger graphene fragments (e.g., platelets) are broken down and taken up by trout cells transformed into smaller graphene pieces of ~10-100 rings (graphenes in the nm size range, here called nGs). We hypothesized that once internalized by cells, nGs/nGOs may interact with microsomal tAHR activating phase I detoxification mechanisms, such as those implicating peroxidase-dependent reactive oxygen 13, 14 and/or tICYPs enzymatic activities/tICYP gene transcription, to induce other bioeffects. All these possibilities remain largely unexplored.

To test such hypothesis by in vitro and in vivo experimentation, given the large numbers of different nGs/nGO derivatives, their numbers need to be reduced and possible candidates identified for further experimental studies. To do that, we proposed to use in silico screening to select for those nGs with the highest binding affinities (lower binding-scores) to tAHR and ICYPs. We discovered that the maximal number of nG rings that tAHRs/tCYPs could bind is up to 25-rings (graphenes in the nm size range, here called nGs). We hypothesized that once internalized by cells, nGs/nGOs may interact with microsomal tAHR activating phase I detoxification mechanisms, such as those implicating peroxidase-dependent reactive oxygen 13, 14 and/or tCYPs enzymatic activities/tCYP gene transcription, to induce other bioeffects. All these possibilities remain largely unexplored.

To test such hypothesis by in vitro and in vivo experimentation, given the large numbers of different nGs/nGO derivatives, their numbers need to be reduced and possible candidates identified for further experimental studies. To do that, we proposed to use in silico screening to select for those nGs with the highest binding affinities (lower binding-scores) to tAHR and ICYPs. We discovered that the maximal number of nG rings that tAHRs/tCYPs could bind is up to 25-rings (graphenes in the nm size range, here called nGs). We hypothesized that once internalized by cells, nGs/nGOs may interact with microsomal tAHR activating phase I detoxification mechanisms, such as those implicating peroxidase-dependent reactive oxygen 13, 14 and/or tCYPs enzymatic activities/tCYP gene transcription, to induce other bioeffects. All these possibilities remain largely unexplored.

To test such hypothesis by in vitro and in vivo experimentation, given the large numbers of different nGs/nGO derivatives, their numbers need to be reduced and possible candidates identified for further experimental studies. To do that, we proposed to use in silico screening to select for those nGs with the highest binding affinities (lower binding-scores) to tAHR and ICYPs. We discovered that the maximal number of nG rings that tAHRs/tCYPs could bind is up to 25-rings (graphenes in the nm size range, here called nGs). We hypothesized that once internalized by cells, nGs/nGOs may interact with microsomal tAHR activating phase I detoxification mechanisms, such as those implicating peroxidase-dependent reactive oxygen 13, 14 and/or tCYPs enzymatic activities/tCYP gene transcription, to induce other bioeffects. All these possibilities remain largely unexplored.

To test such hypothesis by in vitro and in vivo experimentation, given the large numbers of different nGs/nGO derivatives, their numbers need to be reduced and possible candidates identified for further experimental studies. To do that, we proposed to use in silico screening to select for those nGs with the highest binding affinities (lower binding-scores) to tAHR and ICYPs. We discovered that the maximal number of nG rings that tAHRs/tCYPs could bind is up to 25-rings (graphenes in the nm size range, here called nGs). We hypothesized that once internalized by cells, nGs/nGOs may interact with microsomal tAHR activating phase I detoxification mechanisms, such as those implicating peroxidase-dependent reactive oxygen 13, 14 and/or tCYPs enzymatic activities/tCYP gene transcription, to induce other bioeffects. All these possibilities remain largely unexplored.

To test such hypothesis by in vitro and in vivo experimentation, given the large numbers of different nGs/nGO derivatives, their numbers need to be reduced and possible candidates identified for further experimental studies. To do that, we proposed to use in silico screening to select for those nGs with the highest binding affinities (lower binding-scores) to tAHR and ICYPs. We discovered that the maximal number of nG rings that tAHRs/tCYPs could bind is up to 25-rings (graphenes in the nm size range, here called nGs). We hypothesized that once internalized by cells, nGs/nGOs may interact with microsomal tAHR activating phase I detoxification mechanisms, such as those implicating peroxidase-dependent reactive oxygen 13, 14 and/or tCYPs enzymatic activities/tCYP gene transcription, to induce other bioeffects. All these possibilities remain largely unexplored.

To test such hypothesis by in vitro and in vivo experimentation, given the large numbers of different nGs/nGO derivatives, their numbers need to be reduced and possible candidates identified for further experimental studies. To do that, we proposed to use in silico screening to select for those nGs with the highest binding affinities (lower binding-scores) to tAHR and ICYPs. We discovered that the maximal number of nG rings that tAHRs/tCYPs could bind is up to 25-rings (graphenes in the nm size range, here called nGs). We hypothesized that once internalized by cells, nGs/nGOs may interact with microsomal tAHR activating phase I detoxification mechanisms, such as those implicating peroxidase-dependent reactive oxygen 13, 14 and/or tCYPs enzymatic activities/tCYP gene transcription, to induce other bioeffects. All these possibilities remain largely unexplored.

To test such hypothesis by in vitro and in vivo experimentation, given the large numbers of different nGs/nGO derivatives, their numbers need to be reduced and possible candidates identified for further experimental studies. To do that, we proposed to use in silico screening to select for those nGs with the highest binding affinities (lower binding-scores) to tAHR and ICYPs. We discovered that the maximal number of nG rings that tAHRs/tCYPs could bind is up to 25-rings (graphenes in the nm size range, here called nGs). We hypothesized that once internalized by cells, nGs/nGOs may interact with microsomal tAHR activating phase I detoxification mechanisms, such as those implicating peroxidase-dependent reactive oxygen 13, 14 and/or tCYPs enzymatic activities/tCYP gene transcription, to induce other bioeffects. All these possibilities remain largely unexplored.
**SUPPORTING INFORMATION**

Table S1. Trout CYPs CLUSTAL O(1,2.4) sequence alignment

| Accession | Description | Sequence |
|-----------|-------------|----------|
| O42563    | CP3AR_ONCMY LDCVLN | LSQAMIFIF |
| Q92088    | CP2M1_ONCMY TDAVIH | KDAFVTQGEE-FSG |
| Q92109    | CP1A3_ONCMY LEAFIL | RQALIKQGED-FAG |
| Q07217    | CP11A_ONCMY VKGALK | AILFKA--EGHYPK |
| O42563    | CP3AR_ONCMY EIPLE-----MDNQGLLMPK-RPIKLRLEARRNTPSNTTATTLKSPTT 518 |
| P30437    | CP17A_ONCMY PLPSL--E---GKFGVVLQP---VKYKVNATPRAGWEKSH--LQTS-- 514 |
| O42563    | CP3AR_ONCMY TAFSVDIDSLNNPSDPFVSNVKKMLK---FDLFNPLFLLVALFPFTGPILEKMKFSFFPT 244 |
| Q92109    | CP1A3_ONCMY MCFGRRY----SHDDQELLGLVNMSDEFGQVVGSG--NPADFIPILRYLPNRTMKRFMD- 258 |
| Q92110    | CP1A1_ONCMY MCFGRRY----SHDDQELLGLVNMSDEFGQVVGSG--NPADFIPILRYLPNRTMKRFMD- 258 |
| O42563    | CP3AR_ONCMY KTVLIKECYNIFTN |
| Q92088    | CP2M1_ONCMY EEI--NIEPACSSFGRLPRS---YD |
| Q92110    | CP1A1_ONCMY PLDMT------PEYGLTMKH---KR |

Figure S1. 2D scheme of trout modelled CYP molecule and its predicted binding pockets. The figure shows the amino acid sequence in grey and the heme in green (up in lines and down in spheres) occupying an internal part of CYPs. A tunnel-like with an opening to the left in the drawing. (Grey arrow) may support a possible way through which substrates and/or inhibitors may get close to the heme for optimal coordination (ligand binding pocket). Most of the crystalized ligands map to that heme binding-pocket (yellow background). Between 6-15 other binding pockets were predicted by seeSAR depending on the CYPs (different color backgrounds in the down figure)

Figure S2. Drawings of complexes different sizes bound to IAHR (up) and ICYP3 (down). AutoDock Vina used a whole molecule grid (blind-docking). The IDs of 4 (5x5 nm), 9 (7x7 nm), 12 (10x10 nm), 20 (12x12 nm), 25 (13x13 nm), 36 (15x15 nm), 72 (20x20 nm), and 100 (25x25 nm) rings were designed using the GDDP tool. Representative carbamates were made by combining one molecule of each chain with several sizes of rings and drawn in PyMOL. The heme at the inner part of the CYP molecules have been removed for clarity.

Figure S3. Cartoons of the xRGs of 25 rings (13x13 nm) designed by the GDDP tool. All the drawings were made in PyMOL.
Acknowledgements Thanks are due to Dra Ana Valdés for her help to obtain additional aEeSAR licenses. Mon Conolly received funding granted by the Community of Madrid (2018-124688-I1132, Mod. 2, Young Doctor Rachet).