Graphene field-effect transistors as bioanalytical sensors: design, operation and performance†

Anouk Béraud,a,b Madline Sauvage,a,c Claudia M. Bazán,a Monique Tie,a,d Amira Bencherifa,e and Delphine Bouilly*a,b

Graphene field-effect transistors (GFETs) are emerging as bioanalytical sensors, in which their responsive electrical conductance is used to perform quantitative analyses of biologically-relevant molecules such as DNA, proteins, ions and small molecules. This review provides a detailed evaluation of reported approaches in the design, operation and performance assessment of GFET biosensors. We first dissect key design elements of these devices, along with most common approaches for their fabrication. We compare possible modes of operation of GFETs as sensors, including transfer curves, output curves and time series as well as their integration in real-time or a posteriori protocols. Finally, we review performance metrics reported for the detection and quantification of bioanlytes, and discuss limitations and best practices to optimize the use of GFETs as bioanalytical sensors.

1. Introduction

Bioanalytical sensors, engineered at the interface between physics, chemistry, biology and nanotechnology, are a class of instruments designed for quantitative analyses of biologically-relevant molecules (e.g. nucleic acids, proteins, metabolites, drugs, etc.). Such biosensors have numerous applications in a variety of areas including biomedicine,1–3 environmental monitoring4,5 and public health.6,7 Analyte detection and transduction into signal can be mediated by different mechanisms, including optical, electrochemical, electrical or mechanical. In the past decades, advances in the field of nanotechnology have catalyzed remarkable innovation in these different subclasses of bioanalytics sensors, especially through the discovery and production of new nanomaterials. For

Anouk Béraud is an M.Sc. student in Physics at Université de Montréal, Canada. She received her B.Sc. degree from Université de Montréal in 2019, with a dual specialization in Physics and Computer Science. Her master research focuses on instrumentation development to probe surface interactions in graphene field-effect transistor biosensors.

Madline Sauvage is a Ph.D. candidate in Molecular Biology at Université de Montréal, Canada. Previously, she completed a B.Sc. degree in Molecular and Cellular Biology in 2017 followed by a M.Sc. degree in Systems Biology in 2018, both at Université de Montréal. In her thesis, she researches new approaches for the detection and suppression of genetic mutations in breast cancer causing resistance to treatment.
example, gold nanoparticles (AuNPs) and inorganic quantum dots (QDs) have been used in the design of ultrasensitive electrochemical and optical biosensors. Materials engineering at the nanoscale has enabled artificial nanopores in solid-state membranes (e.g., Si$_3$N$_4$ membrane, SiO$_2$, SiC and Al$_2$O$_3$ films) capable of registering the translocation of individual DNA molecules. Nanocarbon materials such as carbon nanotubes (CNTs) and graphene have also stimulated improvements in optical, electrochemical as well as in MEMS/NEMS (micro/nanoelectromechanical systems) bioanalytical sensors.

A specific class of nanomaterial-enabled bioanalytical sensors are field-effect transistors (FETs). In FET biosensors, or bioFETs, the interaction with biological analytes is transduced as a change in the electrical conductance of the sensor. The use of FETs for bioanalytical sensing purposes first appeared around 1980, usually adapted from ion-sensitive field-effect transistors (ISFETs) made for pH sensing. For example, Caras and Janata introduced a penicillin-sensitive bioFET assembled by immobilizing specific enzymes on the surface of an ISFET. Early FET sensors were made using traditional semiconductors (e.g., Si) and oxides (e.g., Ta$_2$O$_5$ or Al$_2$O$_3$), and were often limited in sensitivity due to their low surface-to-bulk ratio. The discovery of low-dimensional semiconductors with extremely high surface-to-bulk ratios prompted the design of various highly-sensitive FET sensors for the detection of ions and molecules. Among these, silicon nanowire FETs (Si-NWFETs) and carbon nanotube FETs (CNTFETs) have both been extensively demonstrated as bioanalytical sensors and even ultimately miniaturized into single-molecule FETs with biomolecules. Despite good sensing performance, the development of 1D-FETs remains hindered today by practical challenges in the synthesis, manipulation and scalable integration of 1D nanomaterials. On the other hand, 2D semiconductor nanomaterials also benefit from extreme surface-to-bulk ratio, but are much more compatible with established microfabrication processes. While a plethora of van der Waals materials have been discovered in the past few years, graphene is by far the most available and well-studied specimen among them. Since the isolation of individual graphene sheets from graphite by Novoselov and Geim in 2004, graphene has received much attention for its exciting mechanical, thermal and optoelectronic properties. In particular, graphene was found to exhibit extremely high charge carrier mobility, as well as remarkable sensitivity to electrostatic changes in its near environment, making it a promising material for sensing applications.

In this review, we focus specifically on graphene field-effect transistors (GFETs) as bioanalytical sensor. GFETs have been demonstrated as sensors in physics and chemistry, for instance as photodetectors, gas sensors (e.g., NO$_2$, NH$_3$, H$_2$O) or pH sensors. More recently, GFETs have been introduced as biosensors: for instance, Mohanty et al. reported in 2008 a GFET biosensor able to detect the hybridization between a tethered single strand of DNA and its complement sequence. Since then, intensive research has been focused on developing GFETs for biomolecular detection. In the bioanalytical field, GFETs have generated interest as ion sensitive field-effect transistors (ISFETs), especially for the detection of toxicity-relevant ions such as heavy metal ions (e.g., Hg$^{2+}$, Pb$^{2+}$). They have also been shown to detect multiple biologically-relevant molecules such as glucose, various biomarkers for diseases including cancer, DNA sequences with single-nucleotide mismatch specificity, pathogens such as bacteria and viruses, or drugs like opioids or antibiotics. GFETs are often described as having key advantages for biosensing applications, including easy operation, fast response, real-time monitoring, high specificity and sensitivity with detection limits down to the femtomolar range, and microfluidic integration and multiplexing capability.

In recent years, there have been several reviews discussing the latest research on graphene and its applications as biosensors. However, there is still a lack of a comprehensive review about GFETs focusing on key parameters for assesse-
singing their design, operation and performance, which is essential to progress towards the standardization of this technology and its uptake in industrial, commercial and/or clinical applications. Here we present a critical review of these three aspects of bioanalytical GFET sensors. We cover specifically experiments focusing on the detection of proteins, nucleic acids, bacteria, viruses, small molecules such as glucose, antibiotics or drugs, and heavy ions such as lead, mercury or potassium. We did not investigate pH sensors as they represent a whole field of study by themselves. In the first part of this review, we briefly explain the fundamentals of GFET operation and review reported approaches for the design and fabrication of such devices. In the second part, we discuss and compare the possible modes of operation of GFETs for the detection and quantitation of bioanalytes. Finally, we review the state of performance metrics reported for this technology and discuss limitations and best practices to optimize the design and performance of GFETs as bioanalytical sensors.

2. Design and fabrication

The design of GFET sensors includes four key components: (1) a graphene layer responsible for the transport of electrical current and the transduction of biosensing events, (2) a set of at least three electrodes as required to operate a transistor, (3) a delivery system allowing tested samples to reach the graphene layer, and (4) a layer of biorecognition elements on the graphene surface allowing for the specific capture of targeted analytes. Fig. 1a illustrates a typical layout for these elements. In the following, we review the role and design principles for each of them.

2.1. Graphene material

Graphene is an atomically-thin material made of a two-dimensional hexagonal lattice of carbon atoms. This structure, with each carbon atom sharing three of its four electrons in covalent bonds with its nearest neighbors (sp² bonds), is at the root of the robust mechanical properties of graphene. At the same time, the remaining fourth electrons are delocalized over the two-dimensional lattice in a H orbital responsible for most of the material’s optoelectronic properties. In the context of GFET sensors, we focus on the electrical and electrostatic properties of the material. Graphene is known for its extremely high mobility surpassing that of excellent metals. Being a semi-metal, its electrical conductance is moderately modulated by local electrostatic fields, allowing to operate the material in a field-effect transistor configuration. Because of this moderate ON-OFF modulation, graphene FETs are typically not considered competitive in pure electronics, compared to state-of-the-art 3D semiconductors such as silicon, or even to its 1D counterpart carbon nanotubes. However, their sensitive electrical conductance combined with their extremely high surface-to-bulk ratio provides them with significant advantages for chemical and biochemical sensing.

Graphene can be produced by several different methods before integration in a FET device. First, graphene can be exfoliated from graphite, a material formed of multiple stacked atomic layers of graphene: the process consists in carefully extracting one monoatomic layer from the bulk graphite. Exfoliation can be achieved by various techniques, including chemical exfoliation, ball milling method, or more commonly micromechanical exfoliation, often referred as the “scotch-tape method”. The scotch tape method was the first reported to isolate graphene, and typically provides the best electrical properties, including the highest mobilities and least density of defects. However, it is difficult to obtain large-area flakes with exfoliation, which makes this approach less suitable for large-scale fabrication of devices. Graphene can also be grown by chemical vapor deposition (CVD), most commonly on metallic substrates like Cu or Ni. In this approach, a hydrocarbon precursor is introduced at high temperature, leading to graphene nucleation on the metal surface. Epitaxial growth on insulating SiC is also possible, in which case gra-

Amira Bencherif is a Ph.D. candidate in Biomedical engineering at Université de Montréal, Canada. She has a bachelor degree in Engineering Sciences from Phelma/Institut Polytechnique de Grenoble in 2014 and a joint master degree in 2016 from Phelma, EPFL and Politecnico Di Torino, in Micro- and Nanotechnologies for Integrated Circuits. Her current research focuses on nanoscale architectures based on graphene field-effect transistors for single-molecule measurements.

Dr Delphine Bouilly is a faculty member in Physics as well as a principal investigator at the Institute for Research in Immunology and Cancer (IRIC), both at Université de Montréal. She has a Ph.D. degree in Physics from Université de Montréal in 2013 and was a postdoctoral fellow in Chemistry at Columbia University. Her research group is interested in bionanoelectronics, more specifically in the interactions between biological molecules and nanoelectronic circuits, as well as in biomedical applications of nanoelectronic sensors.
Graphene nucleates following sublimation of the Si atoms. Graphene grown by CVD is often favored in recent works because it is practical to generate large-area graphene layers, making it the best candidate for scalable GFET production. On the other hand, the mobility may be lower than in mechanically exfoliated graphene and the transfer process following growth can damage the graphene and leave impurities. Finally, another form of graphene is reduced graphene oxide (rGO), often used for its low cost and solution-processability. To produce rGO, a strong oxidation solution is used to separate graphite layers into suspended graphene oxide flakes, which are then chemically reduced back into graphene. The oxidation/reduction process tends to leave a high density of defects, which typically causes lower mobilities than in other types of graphene. Independently of the type of graphene used, most GFET sensor studies report working with a single layer of graphene. Some specifically confirm the presence of a single layer with Raman spectroscopy, as single-layer and few-layer graphene can be difficult to distinguish. Others use few-layers graphene, but single-layer has been reported to enhance the sensing performance.

2.2. Substrate and electrodes

In order to form a GFET device, graphene must be transferred on a planar substrate that provides physical support to the thin nanomaterial as well as to the electrodes and sample delivery system. The substrate, or at least its top layer, is normally made of a dielectric or other insulating material to avoid unwanted electrical connections between the different electrodes placed on its surface. The most popular substrate for GFETs is degenerately-doped Si covered with a layer of SiO₂ which is common in the field of electronics and enables the use of the lower layer as a gate electrode (see Fig. 1d). However, SiO₂ surfaces tend to trap charges and impurities, especially during the transfer process. Other materials are investigated as substrates, for example sapphire on which graphene can be grown directly, leading to enhanced mobilities. Research on more flexible and low-cost substrates is ongoing, for example with materials like flexible polyethylene terephthalate, silk fibroin or paper.

Multiple techniques are used to place graphene on its operating substrate, depending on the graphene source. Graphene flakes obtained by mechanical exfoliation can be directly transferred on the substrate from the adhesive tape used for extraction, by stamping the tape on the target substrate. This straightforward method provides clean, uncontaminated graphene, but is typically incompatible with large-scale FET production. Graphene growth by CVD is done on metal substrates, then the graphene is transferred onto a dielectric substrate using either wet or dry transfer methods. In wet transfer, graphene is protected on one side with a soft polymer layer, typically polymethylmethacrylate (PMMA), and the metal substrate on the other side is dissolved in an etching solution. The protected graphene is then rinsed and picked up onto the target substrate. Alternatively, protected graphene can be separated from the metal by electrochemical delamination. Dry transfer techniques include hot pressing and roll-to-roll methods based on thermal release tape (TRT) applied on the graphene. Pick-up and stamping with PDMS can also be used for dry transfer of graphene. In the case of rGO, the flakes can be transferred from solution onto the substrate of choice via a number of methods, such as drop-casting, dip coating or vacuum filtration on a membrane which is then
stamped on the substrate. Graphene oxide flakes are either reduced before transfer, or first transferred and then reduced to rGO.

GFET design includes at least three electrodes, in order to operate as a field-effect transistor. The first two electrodes, called source (S) and drain (D), make direct contact with the graphene and enable the flow of electrical current in the graphene through the application of a difference of electrical potential between them (Fig. 1b). Source and drain electrodes are made of conductive material, typically a metal: most studies report using Au evaporated on top of a thin adhesion layer of Ti, Cr or Ni. Conductive silver paint can sometimes be used as the electrode material, especially on large graphene. The third electrode, called the gate (G), is placed in close proximity to the graphene but not in direct contact. A potential difference is applied between the gate and the drain (or source) to modulate the density and polarity of charge carriers in the graphene; this mechanism is detailed in section 3.1 on electrical transfer curves.

Multiple configurations have been used for the shape and position of the gate electrode: these can be classified in three main categories illustrated in Fig. 1d-f. The choice of gate configuration depends on the experimental protocol selected for analyte delivery and detection. When the sensor is operated in air or other gaseous atmosphere, a back-gate configuration is usually favored (Fig. 1d). In this layout, the conductive lower layer of the substrate acts as the gate electrode, separated from the graphene and drain–source electrodes by a dielectric layer. Most often, this configuration is achieved using degenerately-doped silicon covered by a layer of thermal silicon oxide. The dielectric thickness determines the capacitance of the gate electrode, as discussed in section 3.1. In the case of SiO2, its thickness can be as large as the order of a micrometer, or as thin as approximately ~10–100 nm, this lower bound being to limit the occurrence of pinholes between the backgate and graphene. However, in biosensing experiments, GFETs are most often directly operated in an electrolyte solution. In such configuration, the gate voltage is applied using either a reference electrode immersed in the medium (Fig. 1e) or a coplanar electrode patterned on the substrate (Fig. 1f). Reference electrodes made of Ag/AgCl represent a common choice since their use in electrolyte buffer is well calibrated. Others have reported immersed gate electrodes made of silver or platinum wires. Coplanar gate electrodes are patterned on the substrate in a similar approach as for source and drain electrodes, using deposition of metals such as platinum, silver or gold. In both cases, the gate electrode is coupled with the graphene via an electrical double layer formed by the redistribution of ions in the electrolyte medium; this is discussed in more depth in sections 3.1 and 4.1. These gate configurations are frequently referred to as “top-gate” or “liquid-gate”, but such terminology can be confused with solid-state planar electrodes placed on top of the graphene and with gating using an ionic liquid, respectively. For configurations described here as in Fig. 1e and f, we recommend using “electrolytic” or “electrochemical” to qualify the gate electrode.

2.3. Analyte media and delivery

Biological analytes (nucleic acids, proteins, ions, drugs) are normally found in physiological samples (blood, serum, plasma, urine), i.e. complex solutions containing multiple species as well as specific salinity and pH conditions. In calibration and detection experiments using bioanalytical GFETs, a variety of media types are reported, with different levels of similarity with actual physiological conditions. The choice of media also influences GFET sensitivity and signal strength, especially by its degree of screening of electrostatic charges: this property of the medium is characterized by the Debye length, which is discussed in more details in sections 3.1 and 4.1. In the following, we review different media types used in GFET experiments, as well as delivery methods used to expose the graphene surface to analyte-containing samples.

The majority of reported GFET experiments are done in saline buffer, in which the purified target molecule is diluted at known concentrations. This approach allows to calibrate quantitation curves over a controlled range of analyte concentrations, and the saline environment is necessary to maintain the proper conformation of macromolecules (nucleic acids and proteins). However, high salinity environments create increased screening, which can make detection by GFETs more challenging (see section 4.1). In DNA detection, different saline buffers are reported; the most common is phosphate buffered saline (PBS) either at its physiologically-equivalent 1× ionic strength (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na2HPO4, 1.47 mM KH2PO4) or diluted at 0.1× or 0.01×. Lower salinity enables longer screening distances, allowing to detect hybridization in parts of the sequence furthest from graphene, but if the ionic concentration is too low (for example in water), strand repulsion can destabilize the double helix conformation. Other studies report using other buffers such as hybridization buffer (10 mM PB, 150 mM NaCl, 50 mM MgCl2, 12.5 mM MgCl2 and 30 mM Tris buffer, known to be equivalent to PBS 1× for DNA helix stabilization. Protein detection experiments also commonly use PBS, Tris-HCl, or PBS diluted with or without competing ions, either in aqueous solution, Hepes buffer, Tris-HCl buffer, or PBS buffer.

Some GFET experiments have reported the detection of analytes in more complex biological samples. For example, An et al. achieved the detection of mercury ions in real samples derived from mussels, and Wang et al. tested blood samples from children for lead ions. Thakur et al. detected the pathogen E. coli in river water samples. For proteins, Kim et al. captured the alpha-fetoprotein biomarker on the surface of GFETs by immersing directly in patient plasma, followed by electrical characterization in PBS after washing steps. Recently, Hajian et al. demonstrated DNA detection directly in genomic DNA extracted and purified.
from cell culture and in human genomic samples, whereas Ganguli et al.124 used loop-mediated isothermal amplification (LAMP) followed by detection of primer (ssDNA) on GFET sensors.

A few experiments completely evacuate the medium before electrical characterization. For example, Ping et al.65 exposed GFETs with solutions of DNA before drying and performing electrical measurements. Similarly, Islam et al.89 reported a back-gated GFET immunosensor for the detection of the human chorionic gonadotrophin (hCG) protein, in which the devices were exposed to probe and target in buffer solution followed by vacuum dry before characterization. In most experiments, however, measurements with GFETs are done directly in the analyte solution, which requires a method to contain the sample over graphene. The minimalistic way to achieve this is by placing a droplet of sample on the GFET substrate to cover the graphene areas.125 Most often, a reaction cell is secured on the GFET substrate, enabling containment and delivery of the sample (e.g. Fig. 1c). Due to the small sensing area of GFETs, such cells are frequently made to contain low sample volumes of the order of tens of microliters.40,54,61,102,109 Because of their size, these are often referred to as microfluidic cells, although they do not necessarily use microscale flow control capabilities characteristic of microfluidic systems.126 Polydimethylsiloxane (PDMS) is one of the most popular materials for cell fabrication due to its chemical inertness, mechanical flexibility, transparency, easy processing and low cost.126,127 GFETs integrated with a PDMS cell have been used for the detection of various targets such as proteins,61,102,115 DNA,53,63,105,109 viruses47,48,128 and small molecules.40,50 Other cell materials have been reported, for example poly(methyl methacrylate) (PMMA)53,83 and silicon rubber.70 The two most common cell designs used with GFET biosensors are the open cell and the flow cell: the first one consists of a simple top-open reservoir in which samples can be pipetted in and out.40,42,54,88,105,109,115,129 Flow cells generally consist of a small enclosed channel with tubing for sample inlet and outlet,50,53,61,63,64,130 allowing minimized evaporation and mixing between samples, lower sample volumes (few µL) as well as controlled fluid flow. This minimizes the consumption of reagents and samples, and lessens signal perturbations such as commonly observed during the loading/emptying of open cells.

In recent years, integration of GFETs into advanced microfluidic systems has been proposed to create versatile lab-on-a-chip miniaturized platforms. In particular, integration of GFETs in multichannel microfluidics enables multiplexing, i.e. the ability to parallelize the detection of multiple targets in the same sample. Several studies have demonstrated multiplexed GFET analysis for protein130 and DNA.53,63,109 Microfluidics integration can also enable GFET measurements under stable flow, instead of in static media. For example, Xu et al.53 quantified the kinetics and affinity of DNA hybridization using a high flow rate of 60 ml min⁻¹ through the PMMA microfluidic channel. Similarly, Wang et al.61 presented a GFET integrated with a PDMS microfluidic flow cell to study the binding kinetics and thermodynamic properties of human immunoglobulin E (IgE) by means of time-resolved measurements performed under a flow rate of 5 µl min⁻¹. Temperature-dependent binding kinetics measurements were possible due to the closed flow cell enabling minimal sample evaporation. Measurements in flow mode also ensured a steady concentration of analyte available for binding, thus decreasing detection times.53,60

2.4. Surface functionalization and passivation

GFETs can be used as sensors because the electrical conductance of graphene is sensitive to electrostatic changes in its environment; however the affinity between graphene and other molecules is not specific. For instance, graphene is known to interact with most proteins and nucleic acids, especially through hydrophobic domains of proteins111 and either the backbone152 or aromatic bases of nucleic acid.133 To engineer specificity in GFET sensors, it is necessary to functionalize the graphene surface with molecules able to specifically recognize and capture the target analyte; these biorecognition molecules are henceforward referred to as probe molecules. The coverage of graphene with probe molecules is often incomplete, in which case passivation strategies can be used to block non-specific interactions with graphene. In the following, we discuss the choice of probe molecules as well as strategies for probe immobilization and for passivation.

2.4.1. Probe molecules for biorecognition

Nucleic acid targets are typically detected via hybridization with their complementary sequence immobilized on the graphene surface. Most DNA hybridization studies directly use single-stranded DNA (ssDNA) as probes, with a nucleotide sequence complementary to that of the targeted DNA.57,58,65,81,104,105,109,129,134,135 The length of ssDNA probes is generally comprised between 12 nt and 50 nt, in order to achieve sequence specificity while avoiding folding and formation of secondary structures in the probe. More complex probe designs have also been explored to improve sensitivity and sequence specificity, in particular to distinguish between single nucleotide polymorphisms (SNPs). For example, Cai et al.44 reported high sensitivity using probes made of single-stranded peptide nucleic acid (PNA) which has a neutral backbone, in opposition to the negatively-charged backbone of DNA, thus enabling the minimization of electrostatic repulsion with the target DNA strand. Hwang et al.43 demonstrated the detection of SNPs using probes based on strand displacement. In this design, probes were made of double-stranded DNA (dsDNA), with the tethered strand perfectly complementary to the target and the other weakly hybridized to the first; the target sequence, when present in the sample, was shown to bind to the probe by displacing the weaker strand. Similarly, Gao et al.57 demonstrated the use of hairpin-folded ssDNA as probe: unfolding of the hairpin was detected when binding the target. Finally, a recent study by Hajian et al.55 used a single-guide RNA inserted in a deactivated CRISPR associated protein 9 (dCas9) to detect a target sequence in amplicons or within intact genomic DNA.
For protein detection, the most common strategy is the use of antibodies as probes, due to their high specificity and affinity for their antigen. For instance, GFETs functionalized with antibodies have been used to detect proteins identified as cancer biomarkers: Kim et al.\textsuperscript{41} immobilized monoclonal antibodies against the prostate specific antigen (PSA) on a GFET biosensor, demonstrating highly sensitive detection of this biomarker of prostate cancer. In a similar way, monoclonal antibodies on GFETs were used to detect alpha-fetoprotein (AFP), a biomarker of hepatocellular carcinoma (HCC), in patient plasma.\textsuperscript{113} Other studies have used GFETs with antibody probes for biomarkers to other conditions, such as human Chorionic Gonadotrophin (hCG), a common pregnancy indicator.\textsuperscript{89} Antibodies on GFETs have also been shown to detect surface proteins of bacteria\textsuperscript{46,84,90} or viruses.\textsuperscript{47,48,136,137} For example, Chang et al.\textsuperscript{84} and Thakur et al.\textsuperscript{46} used anti-\textit{E. coli} antibodies in order to detect the bacteria, and more recently Ono et al.\textsuperscript{90} used immunoglobulin G (IgG) to immobilize the gastric pathogen \textit{H. pylori} on GFETs. Similarly, Liu et al.\textsuperscript{47} used specific antibodies to achieve rotavirus detection. Recently, GFETs with antibodies were also used to detect the SARS-CoV-2 virus responsible for COVID-19.\textsuperscript{136} Antibody probes were also used for the detection of larger complexes such as exosomes\textsuperscript{42} as well as small molecules such as the pesticide chlorpyrifos.\textsuperscript{35}

Aptamers are another type of probe molecules used in GFETs; these are folded single-stranded DNA or RNA oligonucleotides that can bind a target protein or small molecule with high affinity and specificity. Saltzgaber et al.\textsuperscript{64} functionalized graphene with aptamers designed to bind specifically to human thrombin proteins. Farid et al.\textsuperscript{102} reported a GFET functionalized with aptamers for detection of the cytokine interferon-gamma (IFN-gamma) associated with tuberculosis susceptibility. Recently, Wang et al.\textsuperscript{61} studied the binding kinetics of human immunoglobulin E (IgE) to its specific aptamer, allowing the determination of thermodynamic properties of their interaction. In addition, the use of RNA aptamers has been reported for the detection of small molecules, such as the antibiotic tobramycin.\textsuperscript{50}

### 2.4.2. Strategies for probe immobilization

By far the most popular approach to immobilize probe molecules is through graphene functionalization with the linker molecule 1-pyrenebutanoic acid succinimidyl ester (PBASE).\textsuperscript{42–44,48,53,57,61,64,65,81,83,88,90,115,125} On one end, this molecule contains an aromatic pyrene group that binds to the graphene surface through non-covalent \(\pi-\pi\) interactions. The other end is made of a succinimidyl ester group, which is prone to form a covalent bond with amine groups \textit{via} nucleophilic substitution.\textsuperscript{138} Probes made of DNA are often immobilized with PBASE, usually \textit{via} an amine-terminated modifier attached at the 3’ or 5’ extremity of the strand. This approach has been reported for simple ssDNA probes,\textsuperscript{53,65,81,139} and also for more complex ssPNA probes,\textsuperscript{43,125} hairpin-ssDNA probes,\textsuperscript{57} and dsDNA probes based on strand displacement,\textsuperscript{44} as described in the previous section. Aptamers can also be immobilized with the same approach.\textsuperscript{61,64} It should be noted that Kim et al.\textsuperscript{108} reported the immobilization of ssDNA without terminal modifier \textit{via} covalent coupling of the PBASE directly with the amine of nucleobases (adenine, cytosine, and guanine), and of dsDNA \textit{via} non-covalent interactions between the phosphate groups in the DNA backbone and the succinimidyl ester moiety of PBASE. Instead of directly using PBASE, graphene can be functionalized with 1-pyrenebutyric acid, which is then activated using EDC/NHS chemistry into an NHS-ester.\textsuperscript{140} In a different approach leading to the same construct, the pyrene moiety is sometimes directly functionalized to the ssDNA as a modifier to the 3’ or 5’ termination, and the pyrene-DNA complex is then linked to the graphene; this approach was used in Farid et al.\textsuperscript{102} to immobilize aptamer probes, and in Fu et al.\textsuperscript{129} to immobilize ssDNA probe.

The PBASE approach is also frequently used to immobilize proteins, by covalently reacting the succinimidyl ester group with the amine-terminated residue of an amino acid (e.g. lysine) available at the surface of the protein. For instance, this approach was successfully applied to immobilize various antibodies,\textsuperscript{90,115} as well as the dCas9 enzyme used for detection in genomic DNA in Hajian et al.\textsuperscript{55} Some groups use biotin-streptavidin as an intermediary to immobilize protein probes:\textsuperscript{63,90} for example in Ono et al.,\textsuperscript{90} amine sites on the urease probes are functionalized with biotin linkers which are then coupled to streptavidin molecules immobilized on graphene with PBASE. A common aspect of these approaches with proteins is that there are frequently multiple available amine sites on a protein, and thus targeting these provides little control on the orientation of the probe on the sensor surface. This distribution can actually be an advantage for sensing by positioning part of the target-binding sites closer to the graphene surface below the screening limit (see section 4.1).\textsuperscript{141}

Graphene can also be functionalized with covalent moieties, which can then be conjugated with biomolecules. A common reaction to do so is through the use of aryldiazonium salts, in which highly reactive radicals formed from reduced diazonium can directly bind to the carbon lattice of graphene.\textsuperscript{142} The functionality of the aryl group is chosen for further bioconjugation with biomolecule probes: for instance, 4-carboxybenzenediazonium tetrafluoroborate (CBDT) creates stable carboxyphenyl anchor groups on the graphene surface. These –COOH moieties can then be activated using EDC-NHS chemistry into a stabilized NHS-ester ready for coupling to an amine group on the probe, as described with PBASE above. Lerner et al.\textsuperscript{49} used this approach based on CBDT covalent functionalization followed by EDC-NHS reaction to immobilize an opioid receptor protein for naltrexone detection. Others have reported using the EDC-NHS reaction directly on carbonylated defects spontaneously present on the graphene material.\textsuperscript{117} In a reverse configuration, the functionalization of graphene with primary amines (~NH\textsubscript{2}) was shown using electron beam-generated plasmas produced in Ar/NH\textsubscript{3}; amine-terminated ssDNA were coupled with the amine-functionalized graphene using glutaraldehyde as a bifunctional linker.\textsuperscript{143}

Covalent and non-covalent immobilization approaches have different impacts on GFET sensors. Covalent functionalization causes a significant structural change in graphene: it trans-
forms the hybridization of carbon atoms at the functionalization site from sp2 to sp3. These point defects disrupt the conjugation of π electrons, and are known to alter the electronic properties of graphene, including its electrical conductance. However, covalent moieties are extremely stable on the graphene surface, which can be useful for sensors used repeatedly or with high flow rates. On the other hand, non-covalent functionalization such as PBASE does not alter the structural integrity of graphene and therefore its electrical properties. Hence, non-covalent functionalization, usually with PBASE, is largely favored for the immobilization of probe molecules on GFETs. Occasionally, some reports on GFET sensors use no graphene functionalization to immobilize the probes, for example by relying on non-specific interactions between DNA and the graphene. Other works have reported using metallic nanoparticles (such as Pt, Au) as intermediary between graphene and probes. 46,107,146

2.4.3. Strategies for passivation. Passivation of exposed surfaces of the sensor is important to avoid non-specific interactions with species other than the targeted analyte, particularly in complex biological samples such as clinical serum or plasma. One strategy broadly employed is the adsorption of a blocking agent during or after immobilization of the probes. These molecules fill spaces between probe molecules, thus preventing other molecules to make contact with exposed graphene. Different blocking agents have been employed for such purpose, such as bovine serum albumin (BSA), polyethylene glycol (PEG), or mixtures of BSA with Tween20, i.e. a nonionic surfactant made of polyoxyethylene (20) sorbitan monolaurate. Following the coupling of probes with functionalized graphene, some functional groups of the linkers may also remain uncoupled, which can lead to undesirable coupling with non-targeted species. To avoid this, passive adsorption of glycerine has been reported to terminate unreacted NHS groups on PBASE molecules. Similarly, ethanolamine has been used to deactivate and block unbound carboxylic acid reactive groups on the graphene surface. In some sensor designs, the graphene is physically separated from the probes or sensing layer by a thin layer of dielectric, such as Al2O3 or parylene. This protects the graphene and its electrical characteristics by preventing its direct contact with the sample media and the various molecules contained in it. Finally, GFET electrodes can also be passivated with dielectric films (e.g. SiO2/SiNx) either to block interaction with biomolecules and buffer solution, or to eliminate parasitic current.

3. Electrical measurements and metrics

FET-based sensors rely constitutively on electrical measurements, specifically measurements of the electrical current in the device channel – here the graphene layer (see Fig. 1). The general working principle of FET sensors is that the density of charge carriers in the channel (and hence the current) is modulated by the local electrostatic field, which is itself altered by physical or chemical changes in the environment around the channel. Alternate mechanisms to the field effect can include the generation of charge carriers (e.g. in photosensors) or changes in the scattering rates of charge carriers in the channel (e.g. due to increased disorder). In all cases, the detection principle of FET sensors is based on a change in electrical metrics induced by changes in the environment of the sensor. In FET biosensors, this principle is used to detect the capture of biomolecular species at the surface of the sensor. Graphene is a particularly good choice for FET sensors because its atomically-thin geometry makes its electrical conductance remarkably responsive to environmental effects, such as the capture or accumulation of biological analytes near the surface.

In practice, the electrical current of FETs is also controlled by voltages applied to the source, drain and gate electrodes (see Fig. 1). The potential applied between source and drain generates the flow of charge carriers along the channel, while the gate voltage modulates the electric field across the channel – and thus the charge carrier density contributing to the current. FET devices are characterized using three standard curves: transfer curves (current vs. gate bias), output curves (current vs. drain–source bias) and time series (current vs. time) with fixed drain and gate voltages. In sensing applications, the effect of the analyte on such electrical curves can be monitored either a posteriori, by comparing a given metric before and after exposure to the sample, or in real-time by recording dynamic time series of the electrical current.

In this section, we examine specifically how GFET devices are electrically operated for bioanalytical sensing purposes. First, we review the characteristics of operating curves (transfer curves, output curves and time series) and the associated electrical metrics in GFETs. We then compare and discuss the use of these metrics for before–after or real-time detection of biological analytes.

3.1. Transfer curves

Transfer curves of transistors are obtained by sweeping the gate voltage $V_g$ while maintaining a fixed bias $V_{ds}$ between the source and drain electrodes. The resulting current $I_{ds}$ (or resistance $R_{ds} = V_{ds}/I_{ds}$, or conductance $G_{ds} = I_{ds}/V_{ds}$) is plotted as a function of the gate bias. In GFETs, this plot typically results in a V-shaped curve, as illustrated in Fig. 2a. This shape translates an exchange in the polarity of the majority charge carriers in the graphene layer when sweeping the gate voltage: the left branch (or p-branch) represents an increasing density of positive charge carriers (holes), while the right branch (or n-branch) represents negative charge carriers (electrons). Between the two branches, the density of charge carriers – and thus the current – reaches a minimum with equal populations of both positive and negative carriers, referred to as the Dirac point or charge neutrality point (CNP). The p- and n-branches extend linearly from the charge neutrality point such that

$$I_{ds} = g_m (V_g - V_{CNP})$$

(1)
where $V_{\text{CNP}}$ is the gate voltage at the charge neutrality point. The slope $g_m$ is called the transconductance

$$g_m = \frac{W}{L} \mu C_g V_{ds}$$

which depends on the width $W$ and length $L$ of the graphene, $\mu$ the mobility of charge carriers and $C_g$ the gate capacitance. Transconductances for holes and electrons are not necessarily the same, in which case the transfer curve is asymmetrical.

Transfer curves can be obtained using any of the three gate electrode configurations described in section 2.2 and illustrated in Fig. 1. The gate capacitance – and thus the transconductance – is highly dependent on this layout. In a back-gate configuration, the gate capacitance is dominated by that of the insulating layer separating graphene from the planar gate electrode, typically an oxide with a thickness $t$ ranging from $\sim10$ nm to a few $\mu$m. The capacitance of this insulating layer is inversely proportional to its thickness: $C_{\text{ox}} = \varepsilon_{\text{ox}}/t$, with $\varepsilon_{\text{ox}}$ the electric permeability of the dielectric.

The capacitance of this insulating layer is inversely proportional to its thickness: $C_{\text{ox}} = \varepsilon_{\text{ox}}/t$, with $\varepsilon_{\text{ox}}$ the electric permeability of the dielectric. The case of immersed or co-planar gate configurations, the shape and position of the gate electrode can vary considerably, but the capacitance is mostly determined by the electrical double layer (EDL) formed at the graphene surface by the reorganization of ions in the electrolyte media. This EDL acts similarly as a very thin dielectric layer – in the range of angstroms to a few nanometers. The resulting gate capacitance is much larger than that of back-gate dielectrics, and can reach levels comparable to the quantum capacitance $C_Q$. The gate capacitance is then determined by combining the quantum and EDL capacitances in series: $C_g = C_Q^{-1} + C_{\text{EDL}}^{-1}$. Gate potentials applied across the EDL can be over two orders of magnitude more efficient than through the back gate: consequently, the sweeping range of gate voltage required to capture the linear p- and n-branches is much smaller for immersed or coplanar gates, typically in the order of $\pm 1$ V, compared to $\pm10$ V for thin oxides, going up to $\pm 100$ V for thick insulators in the back-gate. In electrolyte media, the range of gate bias sweep must also be restricted to avoid unwanted hydrolysis reactions and other electrochemically-driven reactions at the electrodes.

The choice of gate configuration for a GFET sensor depends on the application. The capture of biomolecular analytes (nucleic acids and proteins) normally occurs during immersion of the probe-functionalized graphene layer in the sample, either an analyte-enriched buffer or a biological sample, such as biomedical (blood, serum, urine, etc.), food or environmental. Analyte detection by electrical measurements, though, can occur directly in the same media or after its removal. Immersed or co-planar gate configurations allow electrical measurements directly in electrolytic samples, and are thus usually favored in GFET bioanalytical experiments. The back-gate configuration is generally not used when the GFET interface is immersed with electrolytes, because screening by the EDL can lessen the back-gate voltage. Back-gated GFET sensors are more frequently used for the detection of volatile analytes in gaseous media, for example in applications such as...
the detection of pollutants. Nevertheless, back-gated GFETs have been recently reported to detect exosomes directly in buffer using the back-gate by exposing only part of the graphene surface to the sample, and they also have been used to detect DNA or naltrexone by immersing the device for exposure following by drying before measurement. Drying the sample is limited to a posteriori detection and can result in non-specific adhesion of various species on the sensor surface, so particular attention to specificity should be exerted in this approach. Finally, let’s note that the electrical interaction between analyte and graphene could also differ between dry and immersed conditions, as difference in environment are expected to alter screening effects as well as intramolecular charge transport properties.

From transfer curves, several electrical metrics can be used for sensing, as illustrated in Fig. 2b–d and discussed in the following:

### 3.1.1. Change in CNP voltage.

The most commonly used electrical metric in GFET sensing is a change in the CNP voltage value, i.e. the gate voltage associated with the minimum of the transfer curve, as illustrated in Fig. 2b. The CNP voltage depends on the doping level of the graphene; for intrinsic graphene at low drain–source bias, it is expected at values close to 0 V, but in reality, it can be either positive if the graphene is p-doped (indicating a larger density of holes) or negative if n-doped (larger density of electrons). The doping level depends on many factors, including the nature of the interface between graphene and other materials (substrate, electrodes, media) and the distribution of charged species and impurities in these materials. Consequently, the choice of materials in device design, their quality and the different processing steps during fabrication of the GFETs have influence on the initial doping state of the graphene layer. In particular, for GFETs made from CVD-grown graphene, the quality of the fabrication process is sometimes associated with the magnitude of the doping, as it can reflect the quantity of impurities located between graphene and the substrate following the transfer process (see section 2.2). Efforts in reducing contaminants in the transfer process has been shown to bring the closer to 0 V. Biosensing experiments based on the change in can be carried out regardless of the initial doping of the graphene, as long as it is moderate enough to have the CNP visible in the gate voltage sweep at every step of the experiment; otherwise another metric must be used.

In biosensing experiments, the interaction between biological targets and biorecognition elements at the surface of graphene can alter the doping state of graphene, thus creating a shift in the CNP voltage from its initial value. This CNP shift is by far the most common metric for biosensing using GFETs. For example, Fakih et al. used the shift in CNP voltage as the sensing metric for K+ ions: they measured transfer curves for a wide range of concentrations of the target ion, as illustrated in Fig. 2e, showing a systematic shift of the curve with analyte concentration. In this experiment, the detection appears to be purely mediated by a doping mechanism, since the whole transfer curve is shifted without altering its amplitude and slope between measurements. From these transfer curves, a clear linear correlation between the CNP voltage and the log of analyte concentration was demonstrated, also shown in Fig. 2e. The change in is also used as a detection metric for complex macromolecular analytes such as DNA oligomers. For example, Gao et al. used the shift of the CNP as a sensing metric for 22 nt single-stranded DNA targets binding to hairpin DNA probes. They reported high sensitivity and specificity with this metric, using it to detect single nucleotide mismatches in the target. Finally, the change in CNP voltage is also frequently used to monitor intermediary steps in the assembly of the biorecognition layer, such as graphene chemistry or immobilization of biomolecular probes.

The polarity of the CNP voltage shift raises interesting questions. Polarity represents the direction of the change on the voltage axis: p-doping when the CNP shifts to more positive voltage, n-doping when it shifts to more negative voltage. The polarity depends on the interaction between analyte molecules and the functionalized graphene layer. Polarities of the change in CNP voltage are reported in Table 1 for different types of analytes: cations, glucose and DNA. All cation sensors report a negative doping, which is consistent with an electrostatic gating model: the capture of positively-charged targets attracts negative charge carriers in the graphene, generating n-doping and a negative shift of the . Oppositely, negatively-charged target molecules would increase the density of holes in graphene and generate a positive shift. This electrostatic gating effect is usually postulated as the mechanism also involved in the detection of molecules; however observations are often inconsistent with this model. For instance, various experiments of GFET sensors for DNA and glucose present opposite polarities in the change of CNP voltage, as compiled in Table 1. For DNA sensors, this discrepancy is associated with at least two opposite effects. Studies observing a p-shift often attribute it to a chemical gating effect, in which the deprotonation of the phosphate backbone of the captured target DNA leaves it negatively charged in buffer, leading to the positive shift. On the other hand, observations of n-doping are explained by non-electrostatic stacking interactions between nucleotides and graphene, or donor effect, which is supported by DFT calculations. These differences may arise from experiment-specific differences in the graphene–analyte–solution interactions when immersed in

### Table 1: Polarity of reported changes in CNP voltage for different analytes in published studies

<table>
<thead>
<tr>
<th>Target</th>
<th>Doping polarity</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cations (K+, Hg^{2+}, Pb^{2+})</td>
<td>n–</td>
<td>87, 101, 119–121, 155–158</td>
</tr>
<tr>
<td>Glucose</td>
<td>p+</td>
<td>40, 91 and 92</td>
</tr>
<tr>
<td></td>
<td>n–</td>
<td>159</td>
</tr>
<tr>
<td>DNA</td>
<td>p+</td>
<td>57, 65, 81, 83 and 92</td>
</tr>
<tr>
<td></td>
<td>n–</td>
<td>43, 44, 59, 60, 63, 104, 105, 107, 109, 125 and 160–162</td>
</tr>
</tbody>
</table>
electrolyte solution, including differences in DNA adsorption, DNA conformation and distribution of counter ions. In the case of glucose sensing, the mechanisms explaining the inconsistencies between experiments exposed in Table 1 have not been investigated in literature. The case of proteins is more complex, as their polarity changes with the pH of solution. In their work, Kim et al. observed the effect of pH on the $V_{\text{CNP}}$ shift for a PSA-ACT complex with an isoelectric point of 6.8: a negative shift of the $V_{\text{CNP}}$ was observed at pH 7.4 when the protein is negatively charged, and oppositely at pH 6.2 when the protein is positively charged. Considering that the density of proteins on graphene is typically too small to generate such a shift via direct charge transfer, the observed shift in $V_{\text{CNP}}$ was explained in this case by asymmetrical scattering due to charged impurities. To summarize, the mechanisms behind the polarity of the $V_{\text{CNP}}$ shift seem to depend not only on the nature of the target, but also on design or environmental factors considering diverging responses reported from similar competing mechanisms have been suggested for DNA to explain those discrepancies and a mechanism has been suggested for proteins, but this topic calls for further investigation.

3.1.2. Change in transconductance. A second type of metric in sensing with transfer curves is a change in the transconductance $g_m = \partial I_{ds}/\partial V_g$, which is the slope of the linear part of the p- or n-branch of the transfer curve. This metric is illustrated in Fig. 2c. The transconductance is an important indicator of transistor performance, i.e. its ability to convert a small change in voltage in a large change in current. For a given device geometry (graphene surface and gate electrode), the transconductance is mainly dependent on the mobility $\mu$ of charge carriers, as described in eqn (2). The mobility of charge carriers in graphene is usually an indicator of its structural and electronic quality, since it is considered to be inversely proportional to the number of impurities in the sample. In a biosensing experiment, the introduction of analytes can alter the transconductance by introducing additional scattering sites and hence increasing disorder, which results in a change in the transfer curve slope as illustrated in Fig. 2c. A limited number of studies explicitly use the change in transconductance as a metric for biosensing, although many experiments show a variation of the slope in at least one of the branches, as can be observed in experiments by Chen et al. shown in Fig. 2f and by Farid et al. shown in Fig. 2g. The polarity of the added scattering sites influences which charge carrier is affected and the slopes often vary asymmetically, as seen in Fig. 2f and g.

Using the change in transconductance directly as a metric requires a linear fit of the p/n branches; the required postprocessing for such analysis and the subjectivity involved in determining the lower and upper limits of the linear range can be considered as limitations of this method. Some groups use this metric indirectly by measuring the change in current at a fixed gate bias. This is robust in cases where the change of transconductance is the only observed variation. For example in Fig. 2f by Chen et al., exposure to the analyte changes only the transconductance of the p-branch without affecting the CNP voltage; authors thus calibrated the current variation at $V_g = -0.5$ V against analyte concentration. In cases where the CNP voltage changes simultaneously to the transconductance, this indirect method would be problematic because it would then aggregate both variations, as discussed further in the following section. Finally, in other cases, the absence of change in the transconductance is explicitly reported and used to interpret the underlying mechanism of the biosensor. For example, Okamoto et al. observed positive doping without any variation in transconductance after the binding of negative antigen fragments, allowing them to hypothesize that antigen capture only changed the negative carrier density without introducing scattering effects.

3.1.3. Change in current. The last type of reported electrical metric with transfer curves is a change in the amplitude of the electrical current, either in the p/n branches or at the CNP, or both, as illustrated in Fig. 2d. For example, Chen et al. observed a change of current only in the p-branch (see Fig. 2f), while Farid et al. reported a change in the electrical current at the CNP (see Fig. 2g). Electrical current being determined by the product of carrier charge, density and velocity ($I = qnmv$), a change in current indicates either a change in the density of charge carriers, in scattering processes, or a combination of both. As discussed in the two previous sections, interactions of biological analytes with the sensor can indeed result in both these effects: a lateral shift of the CNP results from a modification in charge carrier density, and a decrease in transconductance in one or both the branches reflect an increase of scattering. A change of current at a given gate voltage, although easiest to measure, is thus difficult to interpret as it may add or subtract contributions from two different mechanisms. Following the current amplitude at the CNP should in principle control for any shift in the doping state, but even this can be convoluted with an asymmetrical change in transconductance. For example, in Fig. 2g, Farid et al. reported a shift of $V_{\text{CNP}}$ accompanied with a decrease of the CNP current, but the latter may actually be driven by the asymmetrical change in transconductance. In addition, the current minimum of the transfer curve can be affected by the rate of the gate sweep, which can vary with changes in current amplitude, unless constant integration time is specified. Overall, the change in the CNP current is an experimentally-practical metric to be used empirically, but one should keep in mind that it is co-dependent on both the two other metrics.

3.2. Output curves. Apart from transfer curves, GFETs can be operated to measure output curves, in which the drain–source current $I_{ds}$ is recorded as a function of drain–source voltage $V_{ds}$ for a fixed gate voltage $V_g$. The typical output curve of a GFET is represented in Fig. 3a: as the applied bias increases from zero, the amplitude of the current increases with the same polarity as the applied bias. The curvature of the output curve is generally considered a good indicator of the quality of the contacts.
between graphene and source/drain electrodes, and of charge transport along the graphene. With good graphene and electrical contacts, a linear ohmic regime is usually expected at low bias. In practice though, a positive curvature or superlinear regime is sometimes observed due to potential barriers created by non-ideal contacts or defect sites. The output conductance is defined as the slope of the output curve. Its amplitude is evidently function of the gate voltage, which modifies the carrier density, as seen in Tsang et al. A change in the output conductance is occasionally used as a detection metric in GFET sensing experiments. For example, this is done by Huang et al. in Fig. 3b: on the left, they show output curves taken at $V_g = 0$ V after a fixed incubation time in increasing concentrations of bacteria. On the right, the variation in current at $V_{ds} = 100$ mV is used for quantitation of the bacteria. The increase in output conductance with increasing concentrations suggests either $p$-doping or an increase of the transconductance (decrease of disorder). To disambiguate between the two, transfer curves were acquired at concentrations 0 cfu mL$^{-1}$ and 100 cfu mL$^{-1}$, showing a $p$-shift and no transconductance change. This allowed the authors to attribute the variation of output curves to an increase of negative carriers in the system, due to the negatively charged bacteria through electrochemical gating. As this example demonstrates, output curves as a sensing metric should be paired with at least a pair of transfer curves in order to distinguish a change in carrier concentration from a change in disorder.

### 3.3. Time series

In time series, the evolution of the drain–source current (or conductance, or resistance) is collected as a function of time at fixed drain–source and gate voltages. Time series typically start recording before the introduction of reagents and follow the evolution of biochemical interactions between analyte and sensor. These interactions induce variations in current as a function of time, due to changes in charge carrier concentration or scattering effects via mechanisms discussed in earlier sections. Detection of the analyte, and sometimes its quantitation, is assessed from the change in electrical current after injection of the analyzed sample. Since the gate voltage is fixed, this curve is akin to following the evolution of a single point of the transfer curve in time. The choice of gate voltage has a direct influence on the amplitude and the polarity of the signal. It is generally expected for the signal amplitude to be maximized when choosing a gate voltage corresponding to a high transconductance region of the transfer curve. Actually, the interplay between gate voltage and signal amplitude can be quite complex, as illustrated in subsets a to e of Fig. 4. Subset a represents a system undergoing a $p$-doping shift between times $t_1$ and $t_2$, and subsets c to e schematize the resulting times series taken at three different gate voltages ($V_A$, $V_B$, and $V_C$). Even though they result from the same analyte–sensor interaction, time series obtained at gate voltages $V_A$, $V_B$, and $V_C$ exhibit current changes of different amplitudes ($\Delta I_A > \Delta I_B$) or even different polarities ($\Delta I_B < 0$ and $\Delta I_C > 0$). We see here how a slight change of gate voltage, especially close to the CNP, can result in a significantly different profile of the time series. This was experimentally demonstrated by Sudibya et al., who observed both an increase and a decrease of current with increasing concentration of Ca$^{2+}$ ions, depending on the chosen $V_g$. These results highlight the fact that a variation in electrical current cannot be associated to a specific doping polarity without characterization of the transfer curve profiles before and after interaction with the analyte. Moreover, current variation in time series cannot be interpreted as a specific mechanism by itself: for example, the time series represented in Fig. 4d could equivalently be generated by $p$-doping (Fig. 4a at $V_B$) or by a decrease of transconductance in the $p$-branch (Fig. 4b at $V_D$). Insight from transfer curves is thus also necessary in order to correctly identify the mechanism generating current variations in time series.

Time series most often directly present the value of the current as a function of time, as in experiments of Fakih et al. in Fig. 4f and of Saltzgaber et al. in Fig. 4g (top part). Sometimes, the current is converted as a change in voltage such as in Saltzgaber et al. in Fig. 4g (bottom part). An effective voltage shift representation was also used by Xu et al. to study the kinetics of DNA hybridization events and extract binding constants for several concentrations of target...
In this approach, the current change $\Delta I_{ds}$ is converted to a voltage change with the relation $\Delta V_{CNP} = \frac{\Delta I_{ds}}{g_m}$.

It's important to note that this approach is only valid if the transconductance remains constant before and after the addition of targets. As previously mentioned, transfer curves should be provided to confirm that doping is the only mechanism at play. Signal in time series is sometimes normalized as a relative change from a baseline current. Use of normalization can help in assessing signal strength despite sensor-to-sensor variations and effects associated to the medium. For example, Chen et al. used a simple normalization $I_{ds}/I_0$ with $I_0$ the initial current in deionized water and Liu et al. showed...
a relative current \((I_{ds} - I_0)/I_0\) with \(I_0\) the stabilized current after immobilization of the probe molecules. When normalized signal is presented, the conditions used for the baseline should be specified and it is good practice to make available the original time series of the baseline and of the experiment before normalization.

It is possible to avoid the limitations of time series following a constant gate voltage, by implementing a more sophisticated acquisition protocol based on an oscillating gate voltage. In this approach, the gate voltage is continuously swept back and forth over a defined range while the drain–source current is recorded. This results in a two-dimensional mapping of the electrical current as a function of both gate voltage and time. Ideally, the range of the gate voltage sweep is chosen to cover the CNP, which allows to follow the doping state of the graphene at each time point. For example, Zuccaro et al.\(^{54}\) applied this approach of continuous gate sweeps to produce 2D maps of the low-bias resistance as a function of gate voltage and time, as shown in Fig. 4i (left). This approach allows to extract a time series of the \(V_{\text{CNP}}\), as illustrated in Fig. 4i (right), which is a powerful way to quantify the kinetics of the change in doping state.

### 3.4. Comparison of electrical metrics in “before–after” vs. “real-time” protocols

An important consideration when designing a biosensing experiment with GFET sensors is deciding which type of electrical measurements and metrics to use, and in which sequence to collect them. The design of the acquisition protocol depends on the nature of the scientific question or application for which the biosensor is used. We can divide protocols into two categories: “before–after” and “real-time”. The former refers to experiments comparing the value of a metric, at a specific time point after exposure to the sample, to its baseline value before exposure. This is suitable if the goal is to assess the presence of a target (yes/no type of result). It is also relevant for applications requiring quantification of an analyte: the amplitude of the change in the chosen metric is then compared to a previous calibration of the sensor. However if the application or the scientific question requires information about the kinetics of the biochemical interaction, then a “real-time” protocol recording the evolution of a metric over a relevant period of time is necessary. Among the previously described metrics, some focus on the state of the system at a specific time point, while others allow to monitor the evolution of the system, which makes them naturally more or less convenient for each protocol type.

Transfer curves are especially suitable for before–after measurements, as they provide an informative picture of the electronic state of the sensor at a fixed point in time. Indeed, this type of curve provides information on the doping level, through the CNP position, as well as on both carrier mobilities (electrons and holes), through the transconductance of each branch. Transfer curves can be used to assess completion of different steps of sensor assembly, functionalization and biochemical interactions, by collecting a gate sweep after each step and comparing the resulting electrical metrics to the initial curve. Even in experiments focused on real-time measurements for reaction kinetics, it is recommended to collect at least initial transfer curves to assess the performance of the sensors, as done by Cohen-Karni et al.\(^{173}\) In general, before–after analysis of transfer curves is useful for events that have clear before and after states, which are typically before exposure to a reagent and after the reaction with this reagent is considered completed. This type of measurements is commonly used to verify the impact of a passivation layer,\(^{46}\) to confirm the presence of functionalization adducts,\(^{57}\) to assess the linking of the probes\(^{121}\) or the linking of the target to the probes.\(^{44}\) In sensing experiments, this method is most frequently used for quantitation with various types of analytes including ions,\(^{121,122,158}\) proteins,\(^{102,130}\) glucose\(^{40,159}\) and DNA.\(^{63,65,83}\) It is also used for simple yes/no detection, like to assess the presence of a single-mismatched DNA\(^{43,44,57}\) or a specific ion.\(^{158}\) Output curves can be used in the same way as transfer curves in before–after detection schemes, as in Huang et al.,\(^{45}\) but they provide less information on the electrostatic state of the graphene and on physical mechanisms occurring in the system (e.g. change in doping or diffusion). The analysis of transfer or output curves usually requires post-processing, because the extraction of the CNP voltage, transconductances or output conductance can be performed only after completion of the relevant voltage sweep. From the transfer curve, the CNP voltage is often estimated using the point of the transfer curve with the minimum of current.\(^{133}\) Other studies use curve-fitting to extract the voltage associated with the minimum, usually with a quadratic function\(^{61}\) or with more sophisticated models.\(^{85}\) Curve fitting is a more precise method since it is not limited by the width of the gate voltage steps during the measurement. The postprocessing required to determine the transconductance of each branch \((g_h \text{ and } g_e)\) involves subjectivity in determining the lower and upper limits of the linear range, which can be considered a limitation of this method. Real-time measurements are usually performed via unidimensional or bidimensional time series as described in section 3.3. Real-time measurements are of course essential to study the kinetics of a dynamical reaction.\(^{40,53–55,88,174}\) In such experiments, like the study of DNA hybridization by Xu et al.\(^{53}\) illustrated in Fig. 4h, the electrical current is monitored during the introduction of analytes and during washing steps. Time series covering washing steps enable to monitor either the removal of non-specific species and unbound analytes or, in the case of weak probes: target affinities, to observe the dissociation of the analyte from the sensor. Time series can be adjusted with a Langmuir binding kinetics model or similar model to estimate adsorption and dissociation constants.\(^{53,108}\) Time-resolved measurements can also be used for quantification purposes. For example, Fakih et al.\(^{119}\) studied the influence of K\(^+\) concentration with both transfer curves (Fig. 2e) and time series (Fig. 4f), and observed similar correlations with analyte concentration. This type of experiment is especially conclusive when the signal reaches a clean plateau during target
exposure, like in Fig. 4f: such stabilization of the signal facilitated its quantification, which is critical for analyte quantification. A combination of the two purposes, kinetics and quantification, can be done simultaneously, like in Saltzgaber et al.\textsuperscript{64} in Fig. 4g, where the successive introductions of different target concentrations, separated by washing steps, are analyzed to gain insight on the effect of concentration on the kinetics of the reaction. Real-time experiments also allow to assess the reaction rate and the time required to stabilize the interaction between the analyte and the target, which can then be used to determine how to time transfer curves for before–after measurements. For real-time biosensing, experiments need to be done in a saline environment with a coplanar or immersed gate configuration, which requires a flow cell or microfluidic circuitry. Measurements can be done in a static or continuous flow setting. Faster reaction times have been reported for DNA sensing in such settings,\textsuperscript{53,134} but continuous flow was reported to lead to noisier signals because of vibrations due to the water pump.\textsuperscript{50} The choice of flow configuration thus depends on the priorities in the experiment.

For most experiment purposes, both types of measurements are best used together. Standard time series are very instructive about the kinetics of the analyte–sensor interactions, but since they only measure the current at fixed biases, they provide little insight on the physical mechanism underlying these interactions. When time series are coupled with transfer curves, either at specific time points or, even better, continuously in two-dimensional time series,\textsuperscript{54} then the mechanisms behind the evolution of the current can be further investigated. In addition, quantitative analyses based on current changes (either for quantitation or kinetics) often rely on the assumption that the change in current is proportional to the change in graphene doping state, but this is only true in the linear regime of p/n branches and if there is no change of the charge carrier mobilities during the reaction; this needs to be confirmed with transfer curves. Finally, when using before–after experiments without any time series, it is difficult to assess whether the interaction with analytes is stabilized or not; consequently, incubation times are often chosen very long in order to make sure the reaction has occurred. The use of time series, at least during calibration assays, could help optimize the incubation time used in detection assays.

4. Performance assessment

The experiments considered in the scope of this review aim at developing GFETs as a bioanalytical technology, \textit{i.e.} for the detection or quantitation of molecules relevant in biology. In this section, we review the criteria used to evaluate the performance of GFETs as biosensors. In this context, performance include two aspects: quality and reliability.\textsuperscript{175} Quality criteria are established by the performance of the sensor itself with respect to several detection metrics. In the following, we discuss four of these metrics: spatial range of detection, limit of detection, sensitivity to target concentration and response time. Reliability criteria can be assessed by the experimental design; here we will discuss appropriate statistical sampling and analysis, as well as controls experiments.

4.1. Spatial range of detection

For electrolyte-gated GFETs, it is important to take into account charge screening by mobile ions in the medium. According to the Debye–Hückel model, charged molecules in solution are screened by mobile counter-ions such that their electric potential is dampen exponentially with distance, with a decay constant \(\lambda_D\) called the Debye length. This constant represents the screening length and is given by

\[
\lambda_D = \sqrt{\frac{e k_B T}{2 N_A e^2 I}}
\]

where \(\varepsilon\) is the permittivity of the medium, \(k_B\) the Boltzmann constant, \(T\) the temperature, \(N_A\) the Avogadro’s number, \(e\) the electron charge and \(I\) the ionic strength of the solution. The ionic strength is given by

\[
I = \frac{1}{2} \sum_i \rho_i z_i^2
\]

where \(\rho_i\) and \(z_i\) are respectively the density and valence of ion species \(i\). Generally speaking, the Debye length represents the distance at which charges are screened; thus, charges located farther than the Debye length are usually considered out of range for electrostatic detection by a FET sensor.\textsuperscript{176–178} For an aqueous solution at room temperature, this length becomes \(\lambda_D (\text{nm}) = 0.304/\sqrt{I}\) where \(I\) is in mol L\(^{-1}\). For 1× PBS buffer, it is as short as \(-0.7\) nm. Therefore, one must take \(\lambda_D\) into consideration when designing specific probe molecules, as too-long a distance between target binding events and the FET surface may significantly reduce the signal\textsuperscript{146,178–180} or completely screen it out.\textsuperscript{146,176} Such limitations due to Debye length on the spatial range of detection in FET sensors have been experimentally observed in different types of FETs. For example, Sorgenfrei et al.\textsuperscript{178} used an ssDNA probe tethered to a CNTFET to study the effects of shortening the Debye length, \textit{via} changing PBS concentration, on the detection of probe hybridization with a target complementary DNA (cDNA). They found the resistance to decrease significantly (resistance change \(\Delta R/R\) dropping from 80% to 10%) when increasing buffer salinity from 0.1× to 5× PBS (corresponding to a decrease of \(\lambda_D\) from 2.3 nm to 0.3 nm). They also showed that moving the target cDNA further from the surface, by removing two base pairs from the target cDNA (\(-0.66\) nm distance increase), reduced \(\Delta R/R\) from 80% to 20%. This was performed in 1× PBS and, notably, a signal was still detectable, although greatly reduced, even though hybridization occurred at distance of \(-1.36\) nm, exceeding the estimated \(\lambda_D\) of 0.7 nm.

This proximity requirement between the captured analyte and graphene presents a challenge in designing the interface of GFETs, especially in biomedical applications targeting detection in physiological samples. Saline buffers such as 1× PBS or 1× PB, commonly used to emulate physiological environments (\textit{e.g.} human blood), have a very short Debye length of 0.7 nm. In comparison, common probe molecules
such as antibodies for protein detection can be upwards of 10 nm in size. To circumvent this issue, some groups have opted to use solutions with low ionic concentration in order to achieve $A_D$ above 10 nm. Others have designed smaller probe molecules such as antibody fragments or aptamers, allowing to reduce probe length from 10–15 nm to <5 nm, in order to improve device sensitivity. For instance, Kim et al. found that replacing a typical antibody probe (~10 nm) with an aptamer probe (~4 nm) on otherwise similar GFET sensors improved sensitivity to the target protective antigen (PA) by 1000 times (12 aM to 12 fM) in 10 μM PBS ($\lambda_D \sim 23.6$ nm). They also found that signal of PA binding was completely screened out in 1 mM PBS ($\lambda_D \sim 2.3$ nm), even using small aptamer probes, whereas using 10 μM and 100 μM PBS (7.3 nm and 23.6 nm, respectively) showed similar signal intensity (determined by the shift in the charge neutrality point) and limit of detection (smallest concentration detected). Interestingly, the range of PA concentration covered before reaching signal saturation was narrower in 100 μM PBS, indicating lower salt concentration solutions to yield a wider range of detection.

In the case of DNA hybridization experiments, however, many groups have reported detection at very high salt concentrations and even for very long DNA sequences. This reduced limitation to the screening length is likely enabled by the capture of charges close to the FET surface by the first nucleotides of the probe DNA, regardless of its total length. Although high salt concentrations are preferred for stabilizing double-strand DNA, signal and sensitivity can still be improved by decreasing salt concentrations. Additionally, single nucleotide polymorphism have been detected, even if located further along the DNA strand than the Debye length. This is explained by the decreased hybridization stability of single mismatched DNA, leading to partial or complete dissociation of the duplex near the graphene surface.

Many strategies for overcoming Debye length limitations while maintaining physiological environmental conditions have been proposed: such as displacing the screening range away from graphene by covering it with a polymer layer permeable to biomolecules or with charged macromolecules to create a fixed-ion region. Similarly, Chen et al. reported extending the screening length by adding an MoS$_2$ layer on graphene. Other strategies include indirect detection of a target via the products of its reaction on an enzyme, produced outside of the screening length and diffused to the surface of the GFET and using a solution containing 12.5 mM of MgCl$_2$ in a 30 mM Tris buffer, known to provide similar dsDNA stability as in 1× PBS, to increase $A_D$ to 1.6 nm. Finally, issues related to electrolyte screening can be circumvented all together using a backgate and measuring in air conditions.

4.2. Limit of detection and sensitivity

The performance of GFETs as biological transducers is commonly referred to as their sensitivity. Formally, analytical sensitivity describes the ability of the sensor to distinguish between small differences of analyte concentration. Interestingly, this property is actually rarely assessed in bioanalytical GFET studies; rather, the most widely reported performance metric is the limit of detection (LOD), which indicates the lowest concentration at which an analyte can be confidently detected by the sensor. Both of these metrics, sensitivity and LOD, are often conflated, yet they represent distinct standards.

Sensitivity is a crucial performance metric for quantitation applications, in particular when it’s required to identify analyte concentration with great precision. Sensitivity can be assessed from the calibration curve of the sensor, i.e. the evolution of a chosen electrical metric (ex. CNP voltage, current, see section 3) as function of analyte concentration, as illustrated in Fig. 5a. Sensitivity is generally quantified as the slope $S$ of the linear regime of the curve, given by

$$S = \frac{\Delta X}{\Delta N}$$

where $\Delta X$ is the variation of the electrical metric corresponding to a change $\Delta N$ in analyte concentration. As discussed in section 3, analyte–sensor interactions in GFETs are often transduced as a change in CNP voltage, yet the exact relation between $\Delta V_{CNP}$ and analyte concentration is complex and depends on the precise layout of the functionalized graphene interface and of its coupling to analyte and media. Sensitivity can however be assessed empirically. This is straightforward when the CNP voltage is directly measured as the metric, for instance in before-after experiments with transfer curves: Fakih et al. used this approach to assess the sensitivity of their ion sensor in mV per decade. Most often, the measured electrical metric is the drain–source current, especially in real-time experiments, in which case the sensitivity should also be proportional to the transconductance $g_{mo}$ (per eqn (1)). This explains why conventionally, the gate voltage for time series is chosen at the place in the curve with the highest transconductance value. Transconductance itself is proportional to the mobility $\mu$ of the graphene, the gate capacitance $C_g$ and the width-to-length ratio $W/L$ of the graphene surface (eqn (2)). In addition, it is generally understood that the sensitivity of a GFET is also limited by noise: the larger the variance in the signal, and thus in the calibration curve, the more difficult it is to distinguish between close concentrations of analyte. Along this line, Fakih et al. argue that maximizing material mobility, sensor active area and capacitive coupling allows to minimize noise and improve sensitivity. In general however, the sensitivity of GFETs as biosensors has not been extensively studied or reported. A deeper understanding of its underlying factors and of their relative importance would be valuable to better design GFETs, especially for applications requiring precise quantitation of the analyte.

The limit of detection is a different important performance metric, that determines the minimal concentration of analyte detectable by the sensor. The LOD is determined as the concentration at which signal exceeds the baseline by an interval of confidence, as illustrated in Fig. 5a. A low LOD depends...
positively on signal strength and is limited by noise and other sources of variance in the measurements. Although distinct, the LOD and sensitivity can be correlated. For instance, in Fig. 5a, we can see that improving the sensitivity or the slope in the calibration curve is likely to lower the LOD. The LOD is relevant for detection and also for quantitation of analytes: it determines the lower bound of the dynamic range of the sensor, the upper end being limited by the saturation of the signal. In practice, the required LOD depends on the purpose of the sensor: for some applications, a predefined LOD needs to be achieved as prescribed by norms or regulations, for example safe Pb levels in children's blood or glucose levels associated with diabetes. In the development of GFET biosensors, the LOD is the most frequently reported indicator of performance, and is often benchmarked against other detection techniques, such as PCR-based techniques for DNA detection, or ELISA tests for immunoassays. We surveyed the literature on GFETs used as bioanalytical sensors and retrieved 61 studies reporting a value for the LOD (see the ESI† for a list of the studies). An interesting observation from that compilation is that the reported LODs cover a very wide range of concentrations, from 0.1 mM (ref. 91) to 8 zM, which represents a remarkable spread of 16 orders of magnitude. In the following, we discuss the different factors that can influence the LOD in GFETs.

First, we note that methodologies to determine the LOD vary between studies. They can usually be classified in one of two categories: extrapolation or direct measurement. Both approaches first require precise characterization of the sensor baseline, i.e. the value and standard deviation of the chosen electrical metric (ex. CNP voltage, current) in absence of the targeted analyte. The LOD differs from the baseline of the sensor by an interval of confidence, defined as a chosen multiple of the standard deviation. The International Union of Pure and Applied Chemistry defines the LOD to be at three times the standard deviation, and the number of replicas for baseline measurement is recommended at $n = 20$. In the first approach, the electrical metric is measured in presence of different concentrations of analyte, and the resulting response is extrapolated towards lower concentrations. The LOD is then identified as the concentration value at which the extrapolation function intersects with the upper bound of the interval of confidence. For example, Chen et al. used a linear fit to extract a correlation between the relative current change of different concentrations, with replicas, until the observed change in the electrical metric is no longer statistically different from the baseline level. This underlines the importance of assessing the baseline value and its variance. In some specific cases, it is more relevant to use another threshold than the baseline signal to determine the LOD. For example, in Campos et al., the LOD of perfectly-matched DNA is determined as the concentration sharing the same signal value as the highest signal obtained with single-mismatched DNA. In general, the LOD assessment method should be detailed in order to enable proper benchmarking between studies. In our literature survey (see ESI†), the LOD values were taken as reported, without adjustment for the determination method. Extrapolation methods tend to require less measurements, but the extracted LOD can be off the mark if analyte signal deviates from the extrapolation model at low concentrations. In this case the second type of approach is the most reliable as it ensures that...
concentrations above the LOD can really be distinguished from the baseline. When used properly, i.e. with robust statistical analysis and with the right extrapolation function, the two methods should provide equivalent results for the LOD. In studies, the lower concentrations measured and the extrapolated values vary by usually less than an order of magnitude and at most, by two. The data may also include underestimations of the LOD, if the interval of confidence associated with the variance of the baseline is not properly assessed. All things considered, we estimate that discrepancies from the LOD evaluation methods add some variance to the dataset, but are not sufficient to explain the wide spread of LODs across all experiments.

To analyze the distribution of LODs in the literature, we first classified the data by type of analyte, namely ions, small molecules, DNA and proteins. Fig. 5d compiles individual data points and the corresponding box plots characterizing their distribution for each type of analyte. Interestingly, we observe differences in the distributions with analyte type. Ions sensors show the most compact distribution, with a geometric mean and standard deviation of the LOD = 10^(-8.5±2.3) M (n = 14). This is consistent with ions being the most homogeneous category, almost entirely consisting of monoatomic species. Small molecules present a distribution of 10^(-6±3.2) M (n = 10), which is not significantly different from that of ions (p-value = 0.7 for a two-tailed two-sample Student’s t-test assuming non-equal variances). The variance for small molecules is larger than for ions, which may reflect a higher heterogeneity in this category regrouping glucose and various other organic molecules. We note that the charge of the analyte does not appear as a dominant factor in determining the LOD of GFETs: ions present a net charge whereas small molecules don’t – yet they present similar LOD averages. On the other hand, nucleic acids and proteins present LOD distributions with much lower averages, respectively 10^(-13.2±3.4) M (n = 23) and 10^(-13.1±2.4) M (n = 20). These distributions are very similar to each other, and both significantly different from that of ions (p < 1 x 10^-4 and p < 1 x 10^-5, respectively). It appears that there is a distinct segregation between small species (ions, small molecules) and large macromolecules (DNA, proteins). This is likely due to the size difference: macromolecules contain hundreds to thousands of atoms arranged in complex higher-order tri-dimensional structures. Consequently, they are individually more perturbative of the electrostatic environment near the GFET graphene layer, either by carrying more charged sites, by covering a larger area on the graphene and/or by displacing their volume in the saline media. In the same direction, the few GFET studies on the detection of cell-sized analytes such as bacteria and viruses also tend to report very low LODs for these very large analytes.

Within macromolecules, the size of analytes can also vary significantly, depending on the length of the DNA sequence or the mass of the protein. Size effects at this scale have been reported in other techniques, for instance improved detection metrics with longer DNA sequences in surface plasmon resonance or DNA microarrays. In the GFET literature, we looked for a distribution of LODs with size in macromolecules, but observed no significant trend in the available data. For instance, Fig. 5e maps the reported LODs of the 23 DNA studies as function of the length of the target sequence, which reveals no observable correlation. We obtained a similar scattering with protein data points as a function of protein mass (not shown). This suggests that eventual size effects within macromolecules, if they exist, are not dominant compared to variations in sensor and experimental design between studies (discussed below). It should be noted that a series of studies from the Johnson group, using a consistent GFET design, reported lower LODs when increasing the length of the DNA oligomer. In this specific set of experiments, GFETs were rinsed and dried after immersion in the DNA solution, so that detection was performed using back-gated transfer curves. In contrast, most GFET DNA studies use electrolyte-gated designs, in which detection is limited by Debye screening and thus dominated by nucleotides tethered near the graphene surface. This could explain the lack of dependence with DNA length observed overall in the literature data.

Once classified by the type of analyte, LOD values in the literature data still spread over 8 to 12 orders of magnitude per analyte type; this remaining variance must come from differences between studies in the design of GFET sensors (e.g. materials, geometry, surface chemistry, etc.). In particular, it is often suggested that the quality of the graphene is important to produce an efficient FET sensor. We looked into the type of graphene used in GFET assembly (exfoliated, CVD or rGO) as a proxy for graphene quality: it is commonly expected for mechanically exfoliated graphene to present high mobility, and for rGO to present a high density of defects, due to their respective fabrication and transfer methods (see sections 2.1 and 2.2). In the data we collected from the literature, reported mobilities for exfoliated graphene are between 200 cm² (V s)^(-1) (ref. 90) and 4400 cm² (V s)^(-1) and in a similar range for CVD from 605 cm² (V s)^(-1) (ref. 125) to 5000 cm² (V s)^(-1). For rGO, they are between 6 cm² (V s)^(-1) (ref. 196) and 23 cm² (V s)^(-1), so indeed lower by 2–3 orders of magnitude. In Fig. 5d, individual LOD data points in each analyte type were separated in three columns, corresponding to the type of graphene used in the fabrication of the GFETs. From this dataset, we observe no significant correlation between graphene type and LOD values. In particular, LODs from low-mobility rGO-GFETs are not systematically lower than those obtained with exfoliated or CVD graphene with higher mobilities. In protein sensors, we also note that very low LODs (~fM or lower) were obtained with all types of graphene. When keeping all other design parameters constant, some groups have reported a small improvement (one order of magnitude) in the LOD due to graphene type, for instance by changing from rGO to CVD or from few-layer to monolayer graphene, yet such an effect appears eclipsed across multiple studies in which other design parameters are varied. Similarly, the area of the graphene surface has been men-
tioned as a potential influential parameter with respect to the LOD, yet we could not observe a conclusive trend across the reported data.

Rather than graphene properties (type or area), we estimate that variations in the design and assembly of the bio-recognition interface are more likely to explain the large disparity in LOD between studies. As described in section 2.4, there is a wide diversity in the probe molecules used for analyte capture, even for the same type of analyte, in their immobilization chemistry, as well as in blocking agents used to passivate sites around the probes. Additional strategies in the design of the interface have also been reported to enhance signal strength, such as the addition of protective layers or addition of gold nanoparticles. Importantly, the detailed structure of the interface is often weakly characterized: in particular, the density of probe molecules on graphene is usually poorly controlled and often unknown, in large part because it is difficult to measure experimentally. In consequence, there is little quantitative knowledge on the correlation between interface design and the resulting LOD (or any other performance metrics). Better modeling of the interface properties (e.g., density and orientation of probes and blocking molecules, surface defects and impurities, distribution of surrounding ions in the media, analyte docking conformations and probabilities), and especially of their coupling with GFET electrical properties would be highly beneficial to optimize GFET sensors for analytical applications.

4.3. Replicas and controls

In sensor experiments, the proper use of replicas and controls is essential to ensure the reliability of the results. The required number of replicas depends on two types of variability: intra- and inter-device variability. Intra-device variability represents the variation between repeated measurements on the same device. In particular, GFET electrical measurements often drift over time, especially until the equilibration of analyte–probe interactions and of ionic distributions after a change in media or applied electrical potentials. To take account of this, Campos et al. repeated each measurement ten times to reach a stable state, and then kept the tenth curve as representative of the stabilized system. In a similar manner, Xu et al. repeated five times each of their transfer curve measurements and used the average. Moreover, they repeated their experiments on 5 and 8 different devices to account for inter-device variability, allowing to assess the reproducibility of results between sensors. Inter-device variability is inconsistently assessed in the literature. When reported it is often with a limited number of devices, likely due to limitations in the scalability of fabrication and measurements. Some groups implemented fabrication methods to characterize GFET properties over hundreds of devices, enabling sensor experiments with ensembles of 15–30 active devices and 12–20 control devices. Such approaches based on large arrays of GFET sensors enable statistically-robust analyses of GFET performance, which is essential towards clinical and commercial applications. In general, methods for the characterization of both types of variability (intra- and inter-device) should be explicitly reported in GFET studies.

Another important aspect in experimental design is the use of proper positive and negative controls. Negative controls verify that the sensor gives no signal in absence of the specific probe: target pairing, which is important to validate the specificity of the assay to the targeted analyte. Oppositely, positive controls serve to validate that the sensor functions as expected, i.e. that a signal is detected in presence of the targeted analyte.

A common class of negative control consists in exposing probe-functionalized GFETs to non-specific analytes, i.e. species that are not the targeted analyte. It is particularly informative for specificity assessment to test sensors against molecules that are either similar to the targeted analyte (ex. DNA sequences with a few mismatches for DNA sensors) or expected to be present alongside the target in real biological samples. As examples, Kim et al. exposed GFETs prepared with antibodies against Alpha-fetoprotein (AFP) to other protein biomarkers, specifically human chorionic gonadotropin (hCG) and carcinoembryonic antigen (CEA), and Islam et al. exposed GFETs with anti-hCG probes to BSA proteins and small molecules glucose, uric acid and ascorbic acid. For the detection of small molecules such as the opioid naltrexone, authors used flumazenil as negative control, which is a similar compound known not to bind the immobilized receptors. The same idea holds for bacteria: as control for E. coli detection, Huang et al. used the unrelated specie P. aeruginosa, whereas Chen et al. used Listeria. In the case of ion sensing, other ions are most of the time favored as negative controls, for example using Na⁺, Ca²⁺, Mg²⁺ and NH₄⁺ as controls against K⁺ ions. In GFETs targeting DNA hybridization, the most common negative control is to test the sensors with DNA sequences non-complementary to the probe. There are many possible variations of non-complementary sequences derived from the targeted sequence, going from a fully-non-complementary sequence to a very similar single-nucleotide mismatched sequence. Cai et al. validated the specificity of their PNA-functionalized GFETs using both non-complementary and single-nucleotide mismatched sequences. Campos et al. also used a single-mismatched sequence as negative control, and tested the response of their sensors to different concentrations of this control sequence. Other groups tested random one- and two-base mismatched DNA and studied how the position of mismatches affected the sensor response.

In addition to non-specific analytes, it is also relevant to measure the response of the sensor to injection of blank media, in order to assess and eliminate signal due to solution exchange and related perturbations in ion distributions. For example, Zuccaro et al. explicitly measured time series while injecting blank reaction buffer, to compare with injection of the same buffer containing the analyte: the first showed a small short-lived bump due to solution exchange, while the second showed a large, stable shift that can thus be assigned specifically to the analyte.
Another class of negative controls is to test the sensors against the targeted analyte, but without the appropriate probe, to ensure that the response is due to analyte:probe binding and not to non-specific adsorption of the analyte on the sensor. Several groups have reported on the non-specific interaction of their targeted analyte with pristine graphene without probes.\textsuperscript{78,79,121,123,125} A limitation of such experiments is that non-specific interactions of the analyte with the sensor surface are likely to be very different between pristine and probe-functionalized graphene. A strategy to better simulate surface are likely to be very different between pristine and probe-functionalized graphene. A strategy to better simulate this requires multiple steps: sample injection, incubation, washing, all repeated for a number of times.

The duration of each of these steps can be informed by real-time measurements with time series. In particular, the duration of each of these steps can be informed by real-time measurements with time series. In particular, the time required to obtain the result of an analysis is critical to potential applications and to determine how the sensors will be packaged, deployed and used. GFETs are often praised as fast detection tools,\textsuperscript{53,54,120,200} however there is often a lack of clarity as to which steps of their operating protocol are included in this assessment. Specifically, the process of analyte detection or quantitation using GFETs usually require several steps: sample injection, incubation, washing, all repeated for a number of times.

Response times reported for GFETs in the literature vary significantly between experiments, even for similar analytes. For ions, response times are usually small, from almost instantaneous\textsuperscript{51,122,123} to approximately 100 seconds.\textsuperscript{39} For proteins, there is a wider variability: for example, Lei \textit{et al.}\textsuperscript{174} recorded signal reaching a plateau 10 s after insertion of the analyte brain natriuretic peptide, whereas Kim \textit{et al.}\textsuperscript{151} assessed a response time to the prostate antigen PSA of approximately 10 minutes. For DNA hybridization, most studies focus on quantitation experiments using before/after measurements. Incubation times between the two are typically long to maximize hybridization density: from 30 minutes\textsuperscript{65,161} to many hours.\textsuperscript{44} Outside that range, Hajian \textit{et al.}\textsuperscript{52} reported a detection of 1.7 FM in 15 minutes, due to the dCas9 system which actively improves the processing of DNA strands in the sample. The experimental process used to determine these incubation times is seldom described and likely to be based on trial and error, although a few real-time experiments have looked specifically at hybridization and denaturation kinetics.\textsuperscript{53,57}

Some groups notice a significant variation of the response time with target concentration,\textsuperscript{53,57} such as observed by Xu \textit{et al.}\textsuperscript{53} in Fig. 2h. Others define the response time as the time required to reach current saturation in all tested concentrations, in order to determine an incubation time independent of target concentrations.\textsuperscript{41,42} Finally, continuous flow settings appear a promising solution to minimize incubation times, as they report faster response times than standard configurations based on injection followed by static incubation periods. With a 30 μL min\textsuperscript{-1} flow, Stine \textit{et al.}\textsuperscript{134} reported a saturation of signal in less than 800 seconds for the hybridization of fully complementary ssDNA targets at 1 μM concentration. Xu \textit{et al.}\textsuperscript{53} obtained a stabilized hybridization signal in less than a minute for the same target concentration at 60 μL min\textsuperscript{-1} flow.

4.4. Response time

An important practical aspect of sensor performance is the test duration, \textit{i.e.} the time required to obtain the result of an analysis. For a given sensor technology, estimating this metric is critical to identify potential applications and to determine how
consider in assessing their suitability and potential performance. Here, we merely raise some of the considerations that have been explored in the development of GFET sensor technology. Scalable production is an important aspect towards commercialization, in order to achieve competitive production rates and costs: scalable processes for GFET assembly have been developed, usually based on CVD synthesis techniques.\textsuperscript{65,201} Reusability, i.e. the possibility to make successive analyses on the same device, as well as shelf-life of the sensors are other important parameters in technology maturation, and have been tested in some recent studies. For example, reusability has been tested by Wu et al.\textsuperscript{198} by performing successive binding–unbinding cycles with the target and measuring the evolution in the strength of signal, finding a conservation of signal of 94% when comparing the first to the last cycle. Some groups also made assessments of shelf-life by measuring the drift of current over time for multiple concentrations of target\textsuperscript{119} or by repeating experiments after storage time.\textsuperscript{64} Reported shelf-life values vary from one week\textsuperscript{119,156} to a few months.\textsuperscript{119,156} The use of flexible substrates is also an ongoing area of investigation for wearable or skin-implanted devices, for example to detect glucose levels in sweat directly on the skin\textsuperscript{209} or for other health-monitoring purposes.\textsuperscript{202} Silk fibroin,\textsuperscript{201} paper substrates\textsuperscript{92,202} or polyimide\textsuperscript{179} have been successfully tested as flexible substrates.

5. Conclusions

Graphene field-effect transistors have demonstrated promising performance as bioanalytical sensors, including low limits of detection and fast response times in a miniature footprint. Their core feature is the use of graphene conductance as transducer, which provides high sensitivity to the capture of biomolecular species at its surface due to its monatomic thinness. For the past decade, GFET sensors have been prototyped for a wide variety of biologically-relevant analytes: ions, small molecules, nucleic acids and proteins. We reviewed this literature (see ESI†) to discuss best practices in sensor assembly, experimental design and performance assessment, in particular towards the detection, quantitation and kinetic analysis of biomolecules. In sensor design, the type of graphene (exfoliation, CVD, rGO) does not appear as a dominant factor for performance: very low LODs have been reported for high- and low-quality graphene. Two more critical features are the configuration of the gate electrode and the assembly of the bio-recognition interface: both would benefit from better modeling of their effect on graphene conductance transduction. In particular, the specifics of the surface chemistry (e.g. coverage, orientation, stability of immobilized probes, blocking species and captures analytes, and their respective interactions) are often not well known or controlled. On the other hand, limitations due to media screening appear well-understood and modeled by the Debye length, and strategies have been successfully proposed to increase the range of detection. The transduction of analyte capture in electrical conductance can appear as a change in the density of charge carriers by doping (i.e. shift of the charge neutrality point), in the scattering processes (i.e. change in the transconductance), or a combination of them. Transfer curves, alone or combined with time series, are the most appropriate way of studying the physics of this interaction. Some interrogations remain about the coupling mechanism between analyte capture and graphene, as shown by the diverging CNP shift polarities reported for similar analytes. Time series of electrical current alone are not sufficient to interpret interactions mechanisms, but they can provide a robust empirical assessment of target presence or quantification, and they are essential for kinetic studies. Two-dimensional time series combining gate voltage and time sweeps provide both mechanistic and kinetic information in the same measurement. In all cases, in order to produce a reliable and reproducible experiment, intra- and inter-device variabilities need to be assessed and managed using sufficient replicas and appropriate controls. Finally, scalability and cost of fabrication, electronics and fluidics packaging for practical use with samples, as well as reproducibility and stability of the sensor response are important aspects to optimize in order to move the technology forward.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors acknowledge financial support from the Canada Research Chairs, Canada’s Natural Science and Engineering Council (NSERC) and New Frontier Research Fund – Exploration, Fonds de recherche du Québec - Nature et Technologie (FRQNT), and the Institute for Research in Immunology and Cancer (IRIC) of Université de Montréal. A. Béraud and A. Bencherif acknowledge scholarships from IRIC, and M. Sauvage from the TransMedTech Institute. The authors also thank Richard Martel (U. Montréal) for generous support.

References

Analyst Critical Review


69 S. Szuneritis and R. Boukherroub, Interface Focus, 2018, 8, 20160132.


