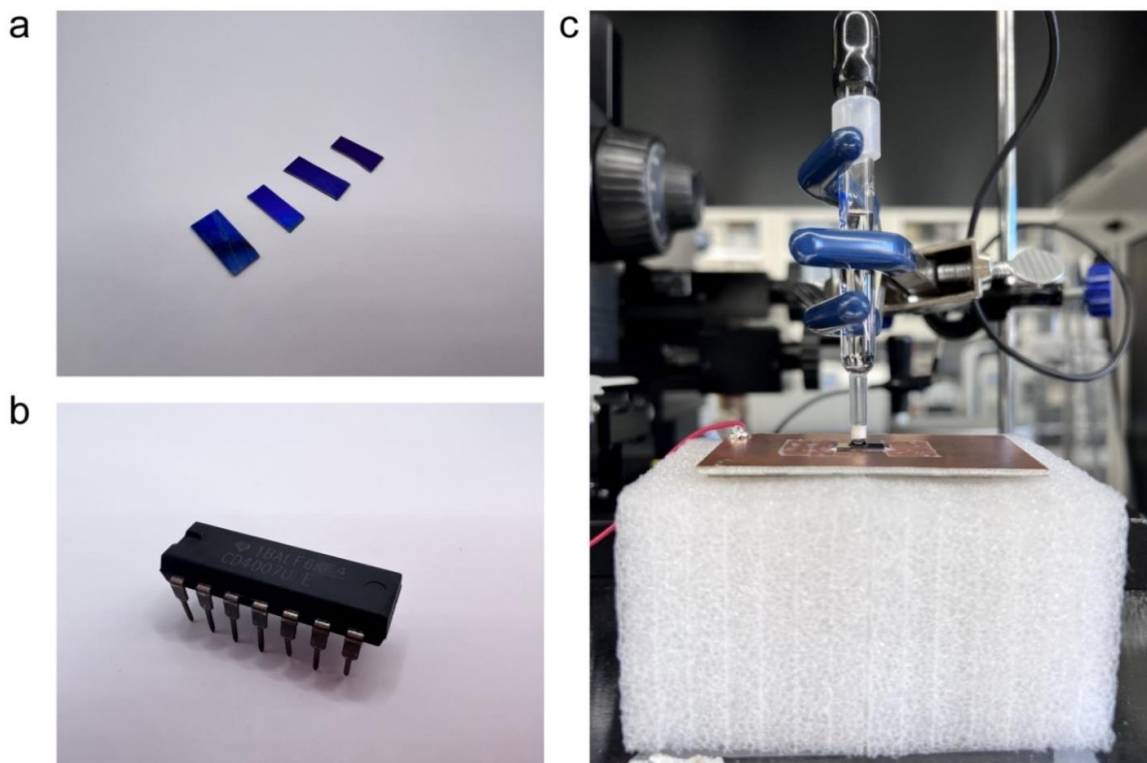


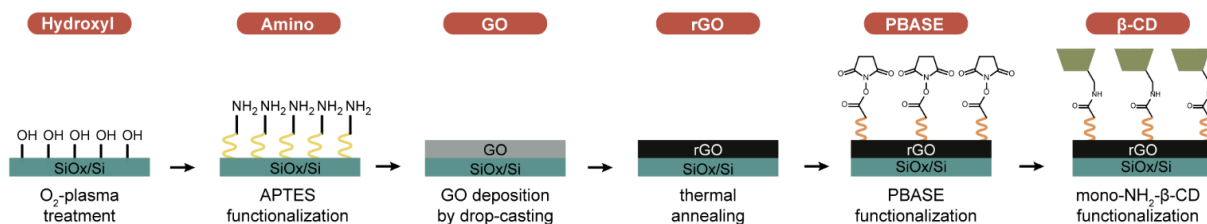
# **Reversible ppt-Level Detection of Perfluorooctane Sulfonic Acid in Tap Water using Field-Effect Transistor Sensors**

## Table of Contents

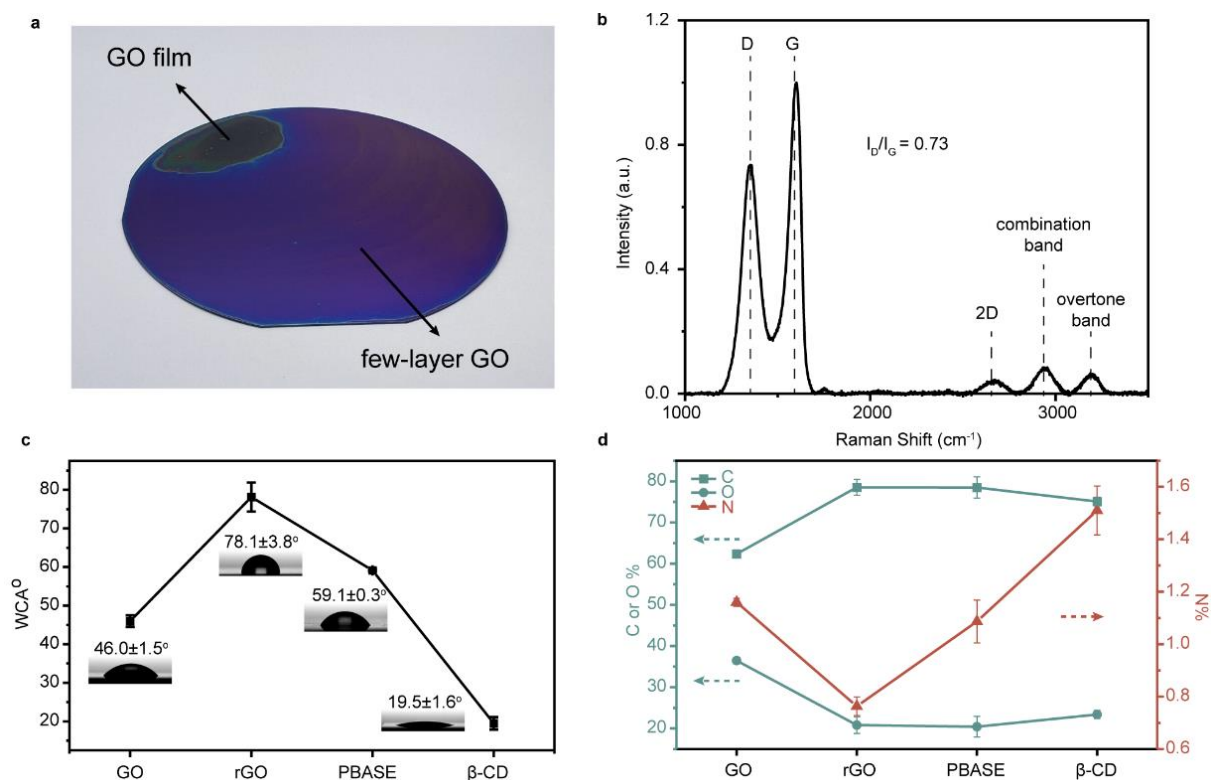
Supplementary Figures 1-9.....	Page S3-S11
Supplementary Tables 1-2.....	Page S12
Supplementary Discussions 1-2.....	Page S13-S16
Supplementary Tables 3.....	Page S17-S18
Supplementary Figures 10-21.....	Page S19-S32
Supplementary Tables 4-5.....	Page S33
Supplementary Figures 22-23.....	Page S34-S35
Supplementary Notes 1-3.....	Page S36-S47
Supplementary References.....	Page S48-S50



**Supplementary Figure 1. Photos of real sensor platform.** **a**, Photo of RG electrodes (cut into small pieces used for testing). **b**, Photo of the commercial MOSFET used in our RGFET sensing platform. **c**, Photo of the whole remote gate consisting of the RG electrode (few-layer rGO on Si wafer), reference electrode, and a sample droplet in between to contact both electrodes (note: the foam and the copper board serve as the physical support for the RG electrodes; the copper board also connects the RG electrodes with the wire to the gate electrode of the commercial MOSFET).



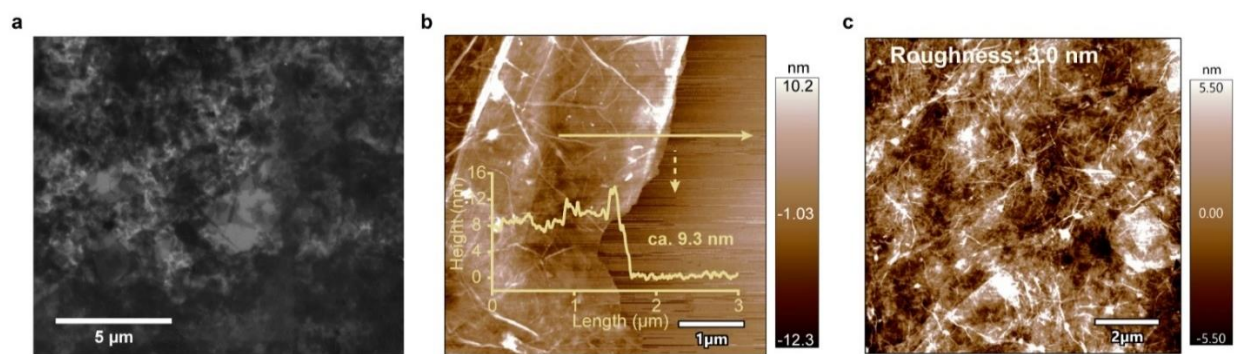
**Supplementary Figure 2. Step-by-step fabrication process of RG electrodes (the top labels represent the chemistry of the outmost layer of the electrodes).** Firstly, a Si wafer was treated with oxygen plasma to generate active hydroxyl groups on the surface for subsequent silanization with (3-aminopropyl)triethoxysilane (APTES), which was added to enhance the adhesion between the Si substrate and the graphene oxide (GO) coatings in the next step. GO layer was deposited by drop casting a GO aqueous dispersion on a Si wafer. Thermal annealing was then conducted to reduce GO to rGO to gain a higher electrical conductivity and more aromatic structures providing  $\pi$ - $\pi$  interactions for PBASE immobilization<sup>1</sup>. PBASE molecules served as the linkers between the rGO surface and  $\text{NH}_2$ - $\beta$ -CD probes. The pyrene end of PBASE bound with rGO, while the N-hydroxysuccinimide (NHS) ester end was later replaced by the primary amine group of mono- $\text{NH}_2$ - $\beta$ -CD.



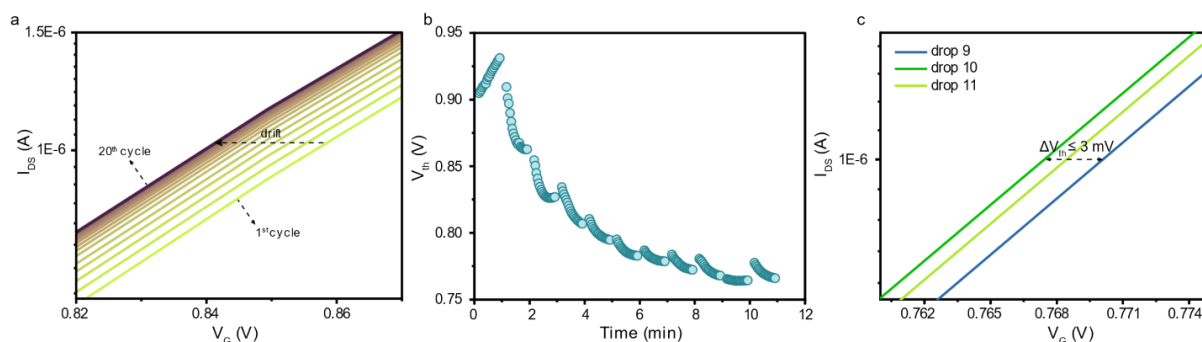
**Supplementary Figure 3. Characterization of RG electrodes.** **a**, Photo of a whole piece of GO-functionalized Si wafer. **b**, Raman spectra of GO RG electrode. **c**, Water contact angles (WCA) change of RG surfaces during fabrication process. **d**, Atomic percentages of C, N, O elements of GO, rGO, PBASE, β-CD RG surfaces calculated from XPS results. The data in **c** and **d** are presented as mean values ± the standard error of the mean (S.E.M.) for  $n = 3$ .

The successful deposition of GO could be observed by the naked eye (Supplementary Fig. 3a)<sup>2</sup>. The characteristic D band (~1,352 cm<sup>-1</sup>), G band (~1,591 cm<sup>-1</sup>), 2D band (~2,668 cm<sup>-1</sup>), combination band (~2,942 cm<sup>-1</sup>), and overtone band (~3,188 cm<sup>-1</sup>) in Raman spectra (Supplementary Fig. 3b) indicated the presence of GO<sup>3</sup> with an  $I_D/I_G$  ratio around 0.73. WCA measurements gave direct evidence of surface hydrophilicity change during fabrication (Supplementary Fig. 3c). The increase in the WCA from GO surface (46.0 ± 1.5°) to rGO surface (78.1 ± 3.8°) was caused by the removal of oxygen-containing groups during thermal annealing.

The WCA then decreased to  $59.1 \pm 0.3^\circ$  upon deposition of PBASE linkers and further dropped to  $19.5 \pm 1.6^\circ$  after the mono-NH<sub>2</sub>- $\beta$ -CD immobilization owing to the abundant hydroxyl groups in  $\beta$ -CD probes. XPS provided chemical evidence of surface modifications (Fig. 1d,e and Supplementary Fig. 3d). The changes in carbon, oxygen, and nitrogen atomic compositions of GO, rGO, PBASE, and  $\beta$ -CD RG surfaces were compared (Supplementary Fig. 3d). The rise in C% and fall in O% from GO to rGO further evidenced the reduction of GO. The increasing N% from rGO to  $\beta$ -CD was a synergistic result of the nitrogen atoms in PBASE or  $\beta$ -CD and the changing membrane thickness, as XPS has a certain penetration depth<sup>4</sup>. The observable elevation of O% in the  $\beta$ -CD RG compared to PBASE or rGO RG was ascribed to the high oxygen content in mono-NH<sub>2</sub>- $\beta$ -CD.



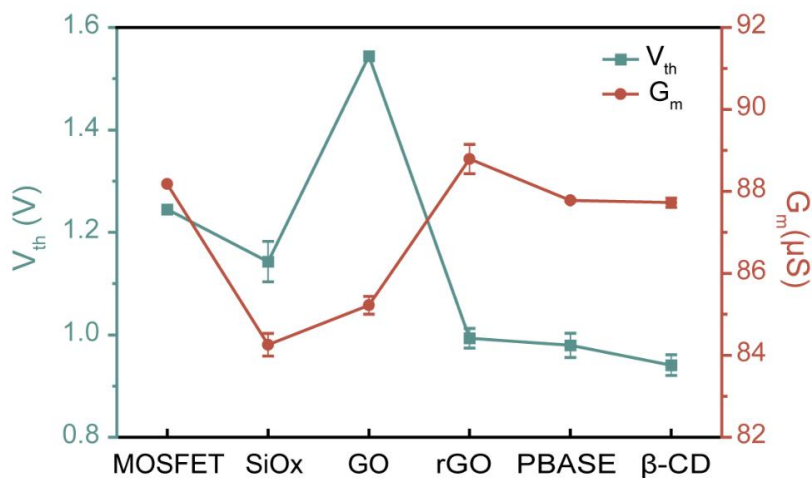
**Supplementary Figure 4. Surface morphology, thickness, and roughness of  $\beta$ -CD RG electrodes.** **a**, SEM image. **b**, AFM image at the edge (inset: height profile along the yellow arrow). **c**, AFM image in the bulk (label at the top: RMS roughness).



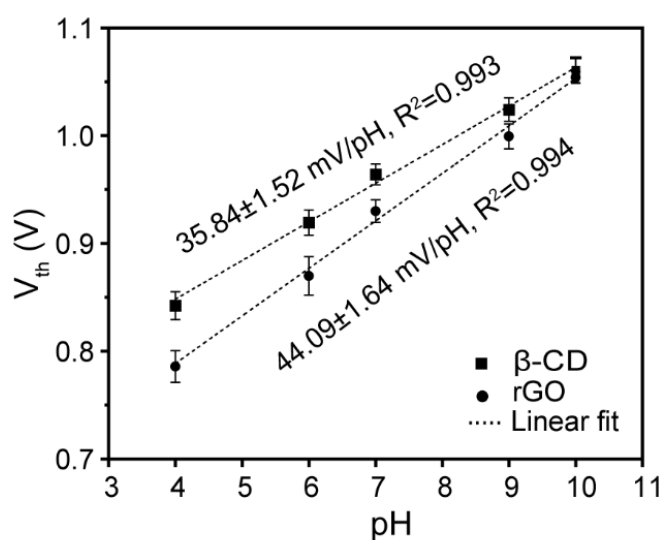
**Supplementary Figure 5. Stabilization process to minimize background drift. a,** Representative 20-cycle transfer curves for one droplet of sample during one measurement. **b,** A typical stabilization process using the baseline solution during one measurement. **c,** Comparison of the 20<sup>th</sup> cycle transfer curves of the last three drops of baseline solution in panel **b**.

For each sample droplet, 20 cycles of transfer curves were recorded in order to minimize the drifts during measurement (Supplementary Fig. 5a) and the  $V_{th}$  of the 20<sup>th</sup> cycle was used for data comparison. An additional stabilization process was implemented by continuously adding and removing baseline solutions (blank buffers without analytes) from the RG surfaces until the  $\Delta V_{th}$  (the difference of  $V_{th}$  at the 20<sup>th</sup> cycle) for the last 3 drops is within 3 mV (Supplementary Fig. 5b,c).

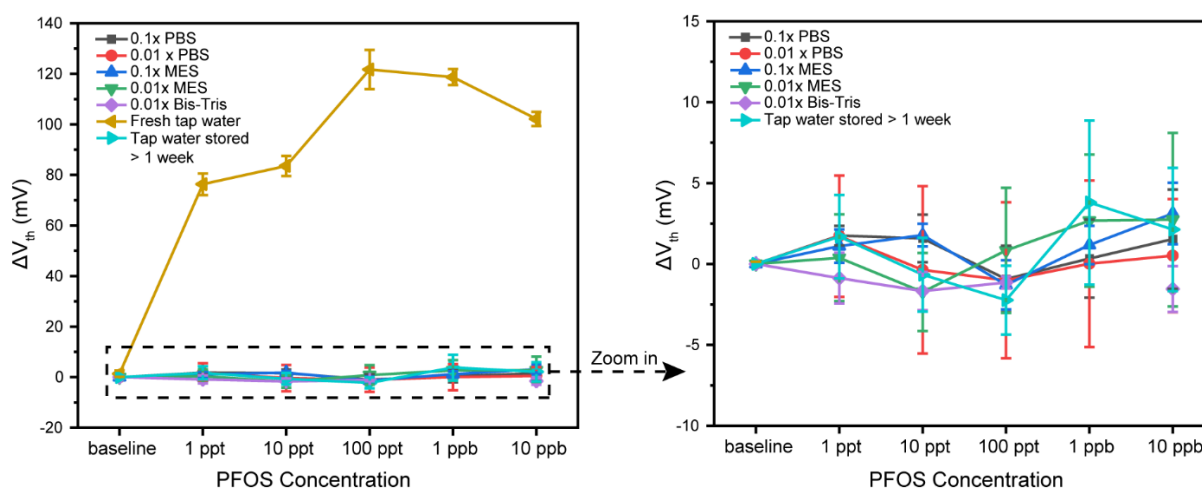




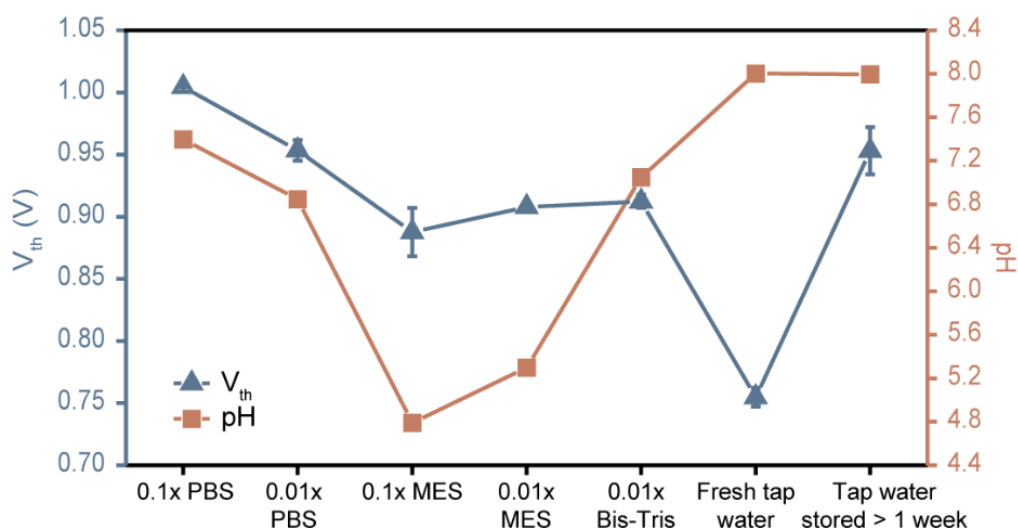
**Supplementary Figure 6. Comparison of  $V_{th}$  and transconductance ( $G_m$ ) of different RG surfaces, measured in pH=7 buffer.** The data are presented as mean values  $\pm$  S.E.M. for  $n = 3$ .



**Supplementary Figure 7. pH Sensitivity of rGO and  $\beta$ -CD RG electrodes.** The data are presented as mean values  $\pm$  S.E.M. for  $n = 3$ .



**Supplementary Figure 8. Comparison of sensor responses to PFOS solutions prepared in different buffers including 0.1× phosphate-buffered saline (PBS), 0.01× PBS, 0.1× 2-(N-morpholino)ethanesulfonic acid (MES), 0.01× MES, 0.01× bis(2-hydroxyethyl)amino-tris(hydroxymethyl)methane (Bis-Tris), fresh tap water (note: in our findings, tap water is a natural buffer with a stable pH around 8, which will not change with different spiked analyte concentrations or during sample storage), and tap water stored >1 week. The data are presented as mean values  $\pm$  S.E.M. for  $n = 3-6$ .**



**Supplementary Figure 9. Comparison of  $V_{th}$  (stabilized) and pH of different buffers on  $\beta$ -CD RG surfaces.** The data are presented as mean values  $\pm$  S.E.M. for  $n = 3$ .

The stabilized  $V_{th}$  values of different blank buffers were measured and their correlation with the buffer pH was studied. The results indicate our sensors are also sensitive to different buffers and the differences in  $V_{th}$  values for these buffers are not solely caused by pH variation. For example,  $0.01 \times$  MES and  $0.01 \times$  Bis-Tris buffers have distinct pH (5.30 and 7.05) but similar  $V_{th}$  ( $\sim 0.91$  V); fresh tap water and tap water stored for  $>1$  week have identical pH ( $\sim 8.00$ ) but different  $V_{th}$  (0.75 V and 0.95 V). This suggests that besides the influences of buffer pH, our sensors may also be responsive to specific charged components in various buffers. But the interactions between the buffers and the sensing surface had already been equilibrated during the stabilization process, thereby not disturbing the analyte response.

**Supplementary Table 1. ICP results on metal ions in tap water used in this work.** The data are presented as mean values  $\pm$  S.E.M. for n = 4-6. The unit is ppm (mg/L).

Element	Ca <sup>2+</sup>	Mg <sup>2+</sup>	Na <sup>+</sup>	K <sup>+</sup>	Al <sup>3+</sup>	Ba <sup>2+</sup>	Zn <sup>2+</sup>	Sr <sup>2+</sup>	Cu <sup>2+</sup>	Fe <sup>2+</sup> /Fe <sup>3+</sup>	Li <sup>+</sup>	Ni <sup>2+</sup>
Fresh TW	46.68 $\pm$ 0.12	11.88 $\pm$ 0.46	12.87 $\pm$ 0.1	1.31 $\pm$ 0.0	0.12 $\pm$ 0.01	0.02 $\pm$ 0.0	0.07 $\pm$ 0.0	0.13 $\pm$ 0.0	0.19 $\pm$ 0.06	0.09 $\pm$ 0.04	0.01 $\pm$ 0.00	0.01 $\pm$ 0.00
TW stored >1 week	46.82 $\pm$ 0.30	12.27 $\pm$ 0.14	11.46 $\pm$ 0.3	1.30 $\pm$ 0.0	0.09 $\pm$ 0.01	0.02 $\pm$ 0.0	0.09 $\pm$ 0.0	0.12 $\pm$ 0.0	0.15 $\pm$ 0.07	0.03 $\pm$ 0.01	0.03 $\pm$ 0.01	0.01 $\pm$ 0.00

**Supplementary Table 2. TOC results for tap water used in this work.** The data are presented as mean values  $\pm$  S.E.M. for n = 5-6. The unit is ppm (mg/L).

TW	TC	IC	TOC	NPOC
Fresh	32.51 $\pm$ 0.15	16.54 $\pm$ 0.09	15.98 $\pm$ 0.20	2.05 $\pm$ 0.39
stored >1 week	37.38 $\pm$ 0.29	30.41 $\pm$ 0.14	6.98 $\pm$ 0.16	6.34 $\pm$ 0.64

### **Supplementary Discussion 1. Exploration of different pH buffers for PFOS sensing and identification of fresh tap water as an effective buffer.**

We explored the potential applications of seven buffers, including 0.1× phosphate-buffered saline (PBS), 0.01× PBS, 0.1× 2-(*N*-morpholino)ethanesulfonic acid (MES), 0.01× MES, 0.01× bis(2-hydroxyethyl)amino-tris(hydroxymethyl)methane (Bis-Tris), fresh tap water (note: in our findings, tap water is a natural buffer with a stable pH around 8, which will not change with different spiked analyte concentrations or during sample storage), and tap water stored >1 week. We tested the RGFET sensor response against PFOS spiked in these buffers at selected concentrations (Supplementary Fig. 8) and found only PFOS samples prepared in fresh tap water could give strong sensor responses, suggesting the importance of testing freshly-collected tap water samples. We speculate this could be related to the distinctly low  $V_{th}$  of fresh tap water (0.75 V) compared to other buffers (~0.90-1.00 V) (Supplementary Fig. 9), since the sensing signal ( $\Delta V_{th}$ ) depends on both the baseline  $V_{th}$  value and the final  $V_{th}$  value caused by analyte adsorption (see Supplementary Discussion 2).

The molecular mechanism behind the  $V_{th}$  of the fresh tap water is hard to investigate as the composition of tap water is complex and dynamic. We explored the behavior of fresh tap water vs. tap water stored >1 week by comparing their cation and carbon concentrations, which we suspected would have an impact either on the PFOS adsorption or the surface charge properties. Cation concentrations are generally similar for fresh tap water and stored tap water, while we observed small but noticeable decreases in sodium ion concentration after tap water was stored over one week (Supplementary Table 1). The changes in the carbon content including inorganic carbon (IC), total organic carbon (TOC), and non-purgeable organic carbon (NPOC) were more obvious, where the IC rose significantly (16.54 ppm to 30.41 ppm), TOC decreased by ~50% (15.98 ppm to 6.98

ppm), and NPOC increased by >3 times (2.05 ppm to 6.34 ppm) (Supplementary Table 2). The NPOC content could partially reflect the amount of dissolved organic carbon (DOC)<sup>5, 6</sup> and particulate organic carbon (POC), where there are many anionic or non-conductive organic carbons<sup>7</sup> that could lead to a rise in the  $V_{th}$ . The increased IC could be related to the dissolved carbon dioxide during the storage, which may potentially lead to negative doping of the sensing membrane and thus cause an increase in the  $V_{th}$ . However, these observations are insufficient to definitively explain the distinct  $V_{th}$  of fresh tap water because there are many other species in tap water that we have not screened but still have the possibility to alter the surface charge or the adsorption behavior. What matters most for an effective buffer is not any specific component, but the resulting  $V_{th}$  caused by synergistic interactions between everything in the buffer and the sensing surface. In other words, using other buffers that have a relatively lower level of  $V_{th}$  like that of the fresh tap water would potentially produce sensing responses for PFOS, as long as they do not prevent PFOS adsorption.

**Supplementary Discussion 2. The effect of the baseline  $V_{th}$  and the potential cause for the coming back sensor signals after the saturation point.**

The relation of the surface potential and surface charge density can be described by the Grahame equation below<sup>8</sup>.

$$\sigma = \sqrt{8C_0\epsilon_r\epsilon_0k_B N_A T} \sinh\left(\frac{ze\varphi_s}{2k_B T}\right)$$

where,

$\sigma$  = surface charge density

$C_0$  = the bulk concentration of the electrolyte ions

$\epsilon_r$  = relative permittivity of the medium

$\epsilon_0$  = the permittivity of free space

$k_B$  = Boltzmann constant

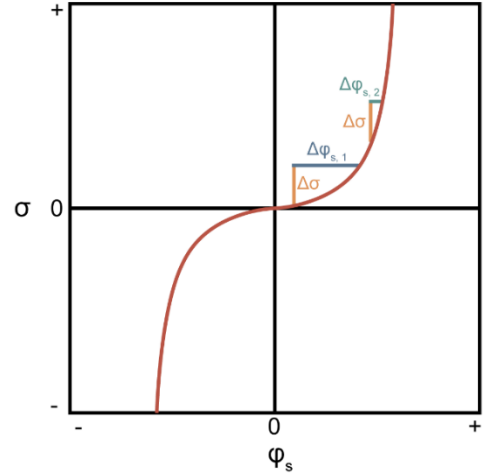
$N_A$  = Avogadro constant

$T$  = temperature

$z$  = the valence of the electrolyte ion

$e$  = the elementary charge

$\varphi_s$  = surface potential



**(1) The effect of baseline  $V_{th}$  on the final  $\Delta V_{th}$**

During the sensor testing, the  $\Delta\sigma$  is related to the amount and the charge of surface adsorbed molecules and the  $\Delta\varphi_s$  is the parameter we quantify as the sensing signals. As indicated by the Grahame equation, the relation between  $\Delta\sigma$  and  $\Delta\varphi_s$  is non-linear but follows a  $f(x) = a \cdot \sinh(b \cdot x)$  function where  $a$  and  $b$  are coefficients related to the solution properties. For a certain  $\Delta\sigma$  value, if the initial  $\varphi_s$  ( $\varphi_i$ ) is different, the resulting final  $\varphi_s$  ( $\varphi_f$ ) and the calculated  $\Delta\varphi_s$  ( $\varphi_f - \varphi_i$ ) could be very different as illustrated in the figure above (for the same  $\Delta\sigma$ ,  $\Delta\varphi_{s,1} > \Delta\varphi_{s,2}$ ). For the PFOS prepared in different buffers, if we assume the coefficients related to the solution properties are the same and the adsorbed PFOS amount by the RG surface is similar at same PFOS concentrations, because different buffers have different  $V_{th}$  values and especially, the fresh tap water has a uniquely low  $V_{th}$  (Supplementary Fig. 9), which are related to the initial  $\varphi_s$  values, the resulting  $\Delta\varphi_s$  (related to  $\Delta V_{th}$ ) could also be varied. Actually, different buffers may also have different  $C_0$ ,

$\varepsilon_r$ , or  $z$  values, which would influence the shape of  $\sigma - \varphi_s$  curve and affect the resulting  $\Delta V_{th}$ .

Besides, the adsorbed analyte amount at the same analyte concentration but in different buffers could also be different as indicated by the QCM results (Fig. 3b,c), which could result in different  $\Delta\sigma$  to affect the observed  $\Delta V_{th}$ .

(2) Potential reasons for the “coming back” signals of RGFET sensors after the saturation point

In the Grahame equation, the  $\sigma$  and  $C_0$  have opposite effects on  $\varphi_s$ . At low target analyte concentrations, the  $C_0$  is more determined by the bulk solution and less influenced by the analyte concentrations, so the changing  $\sigma$  is the major factor that determines the final  $\varphi_s$ . While after reaching a certain high concentration (near the saturation point), the increasing analyte concentration may lead to an increase in  $C_0$ , which would potentially cancel part of the  $\Delta\varphi_s$  brought by the  $\Delta\sigma$ . Moreover, at the high concentrations, the analyte adsorption could have entered the multilayer adsorption regime, which would bring less  $\Delta\sigma$  limited by the Debye length, thus having smaller effects on increasing the  $\Delta\varphi_s$ .



**Supplementary Table 3. LC-MS/MS results for blank tap water and PFOS solution prepared in tap water.**

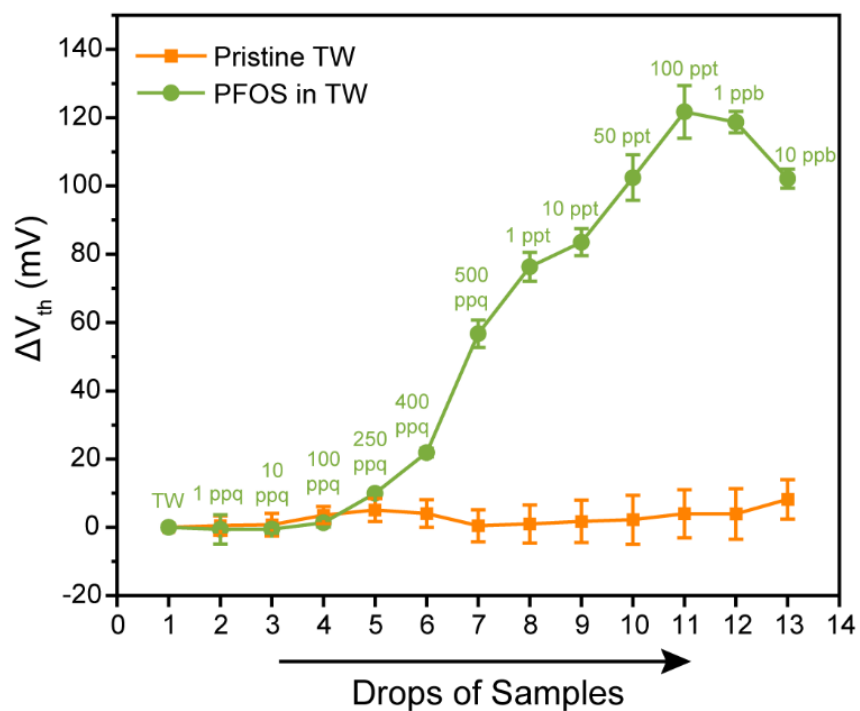
Spiked PFOS Concentration	Water Type	Original Sample Volume	Solid Phase Extraction (SPE)	Time between preparation and analysis	LC-MS/MS Results
Blank tap water	Tap Water	500 mL	Yes	4 days	2.36 ppt*
1 ppt	Tap Water	500 mL	Yes	4 days	5.06 ppt*
1 ppt	Tap Water	500 mL	Yes	7 days	6.34 ppt*
1 ppt	Tap Water	125 mL	Yes	14 days	5.53 ppt
10 ppt	Tap Water	125 mL	Yes	14 days	22.7 ppt
100 ppt	Tap Water	125 mL	Yes	4 days	108 ppt
100 ppt	Tap Water	125 mL	Yes	7 days	116 ppt
100 ppt	Tap Water	125 mL	Yes	14 days	101 ppt
1 ppm	DI Water	5 mL	No	14 days	1.65 ppm
100 ppm	DI Water	5 mL	No	4 days	119 ppm
100 ppm	DI Water	5 mL	No	7 days	116 ppm
100 ppm	DI Water	5 mL	No	14 days	123 ppm
1000 ppm	DI Water	5 mL	No	4 days	957 ppm
1000 ppm	DI Water	5 mL	No	7 days	981 ppm
1000 ppm	DI Water	5 mL	No	14 days	1096 ppm

\*Laboratory Reagent Blank (LRB) result is >50% of method reporting limit (MRL).

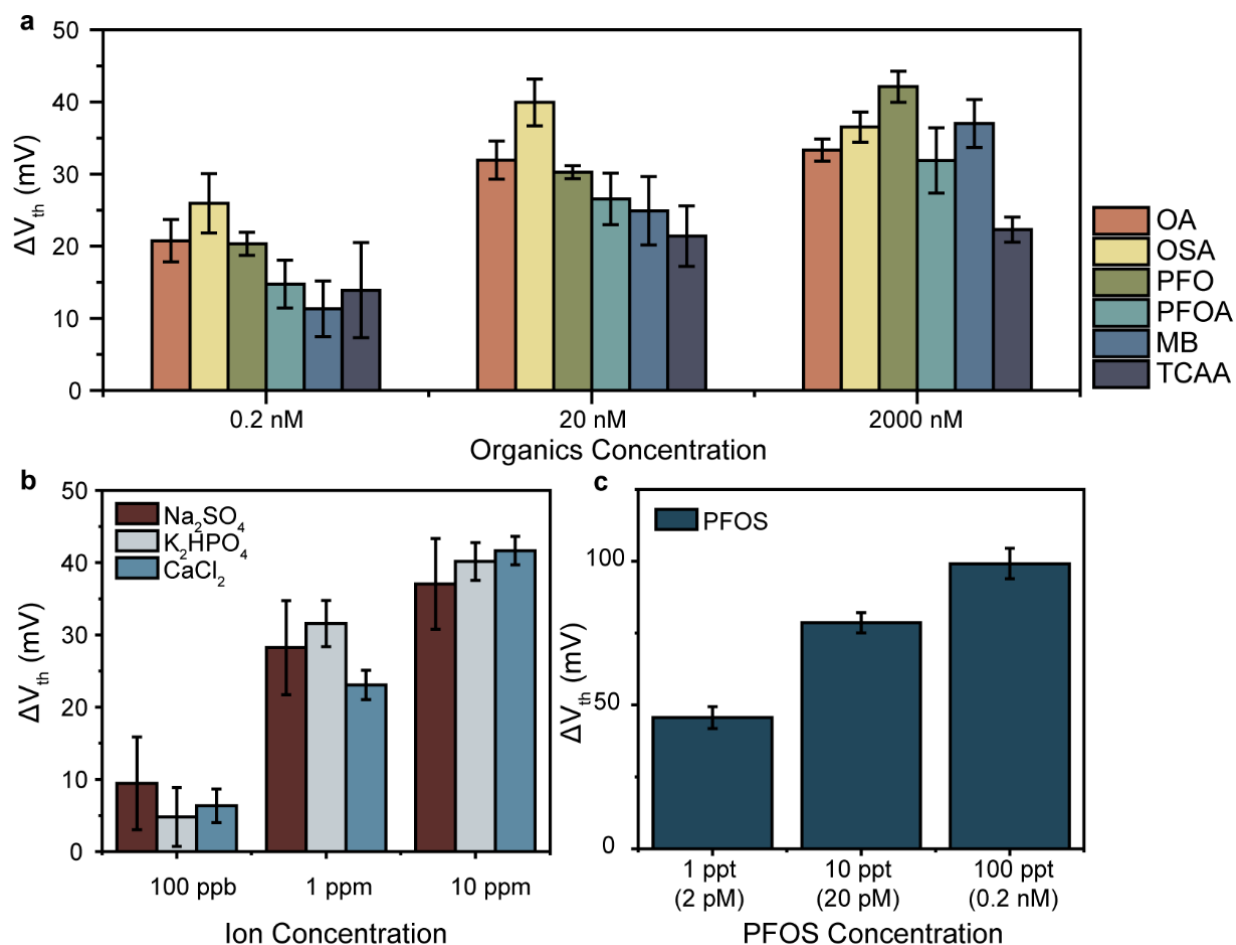
PFOS sample preparation can be affected by many factors including the source of chemicals, preparation methods, container materials<sup>9</sup>, sample storage conditions<sup>10</sup>, and accidental cross-contamination. To ensure the quality of sample preparation, we prepared several PFOS samples at selected concentrations within the application range (ppt to ppm) for our RGFET sensors and the QCM study, quantified their real concentrations using standard EPA methods for drinking water (LC-MS/MS), and compared them to the designed/calculated concentrations (based on the spiked PFOS amount). The measured concentrations of most samples are close to the designed concentrations as shown in Supplementary Table 3, validating the reliability of our sample

preparation protocol. Note that the samples were analyzed without solvent rinsing the sample containers. Recoveries of PFOS for samples may be lower than expected, as the walls of the original bottles for each sample were not methanol rinsed. As a result, PFAS compounds may have not been rinsed of the walls of the original container into the extraction media.

However, we have not been able to confirm sub-ppt concentrations using LC-MS/MS given the background PFOS concentration in blank tap water and the higher reporting limits of PFOS using existing standard methods.



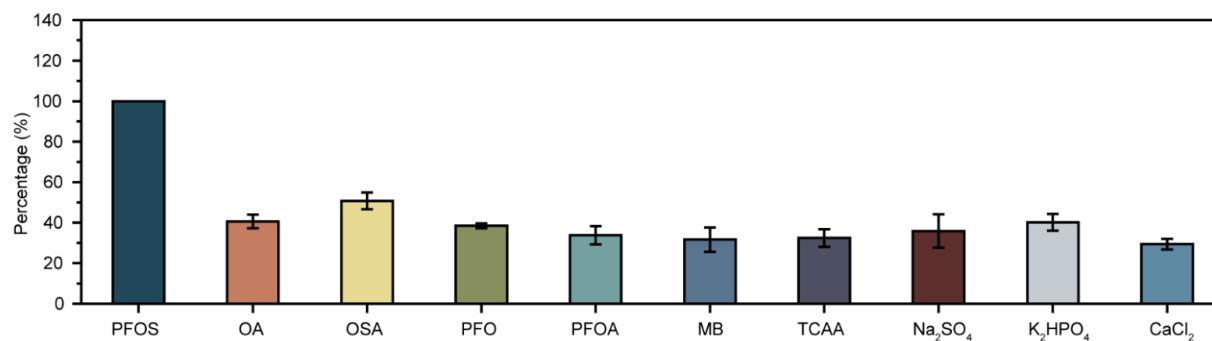
**Supplementary Figure 10. Comparison of sensor response for PFOS solutions in tap water and pristine tap water after stabilization on  $\beta$ -CD-functionalized RG electrodes.** The data are presented as mean values  $\pm$  S.E.M. for  $n = 3$ .



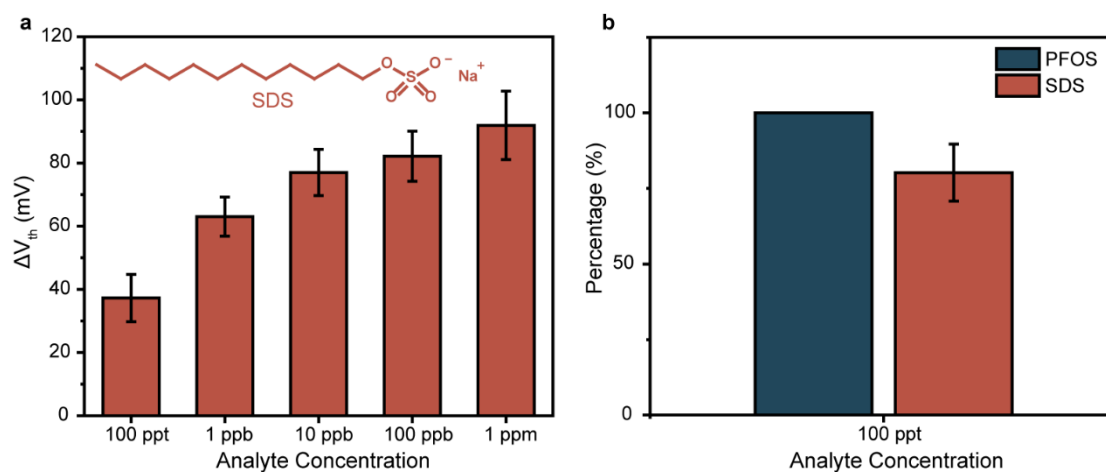
**Supplementary Figure 11. Sensor responses of PFOS, 6 organic interferents and 3 inorganic ions in tap water at different concentrations on  $\beta$ -CD RG electrodes. a,** Sensor response for organic interferents @ 0.2 nM, 20 nM and 2,000 nM. **b,** Sensor response for inorganic ions @ 100 ppb, 1 ppm, 10 ppm. **c,** Sensor response for PFOS @ 1 ppt, 10 ppt, 100 ppt. The data are presented as mean values  $\pm$  S.E.M. for  $n = 3$ .

Three concentrations of each analyte solution were added onto the RG electrodes sequentially from lower concentrations to higher concentrations during the sensor testing. Our sensing results would be influenced by the reaction time (the running time of 20 cycles is  $\sim 1$  min, but samples with higher concentrations may not reach equilibrium state within 1 min, and their sensing response

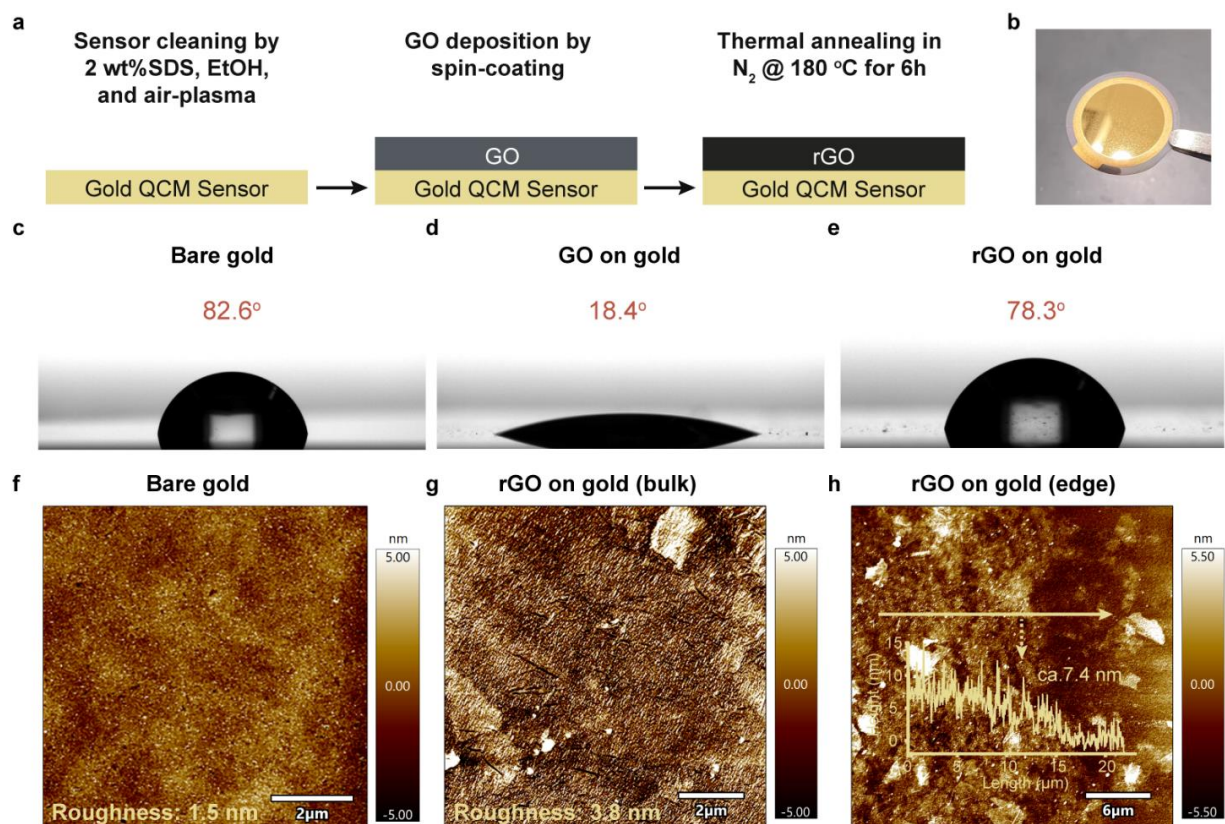
will increase with the time); in order to control the variables, we chose the responses of different analytes with the same contact time for data comparison (e.g., the 2<sup>nd</sup> drop of sample, which means the middle concentration) as shown in Supplementary Fig. 12.



**Supplementary Figure 12. Selectivity at realistic concentrations.** Sensor responses to PFOS and 6 organic interferents and 3 inorganic ions in terms of the percentage response compared to the PFOS response in tap water on  $\beta$ -CD RG surface. The concentration for PFOS is 10 ppt, for organic interferents is 20 nM, and for inorganic ions is 1 ppm. The data are presented as mean values  $\pm$  S.E.M. for  $n = 3$ .



**Supplementary Figure 13. Sensor responses to SDS in tap water. a,** Sensor response to SDS in tap water from 100 ppt to 1 ppm. **b,** Comparison of PFOS and SDS responses at 100 ppt in terms of percentage of PFOS response. The data are presented as mean values  $\pm$  S.E.M. for  $n = 3$ .

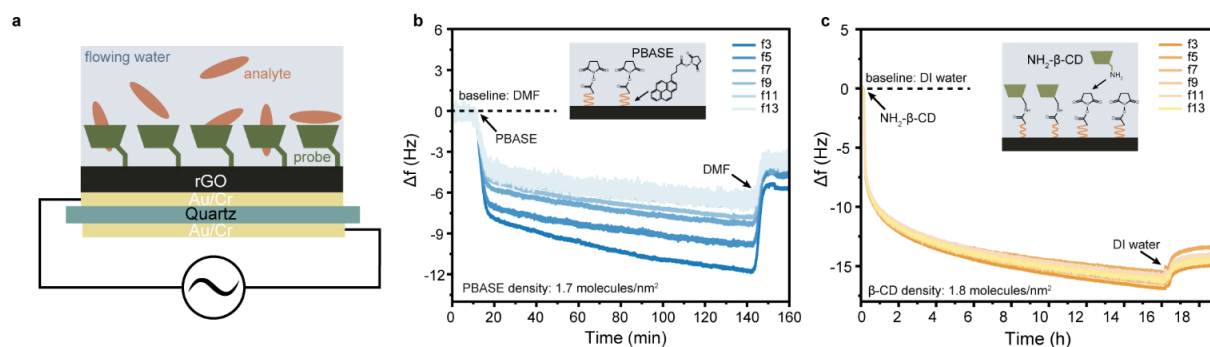


**Supplementary Figure 14. Preparation of rGO-functionalized gold QCM sensor.** **a**, Illustration of surface functionalization process. **b**, Photo of a real rGO-functionalized sensor. **c-e**, Water contact angles for bare gold, GO-coated surface, and rGO-coated surface after each step of surface treatment. **f-g**, AFM images of bare gold and rGO-coated sensor in the bulk (label at the bottom: RMS roughness). **h**, AFM image of rGO-coated sensor at the edge (inset: height profile along the yellow arrow).

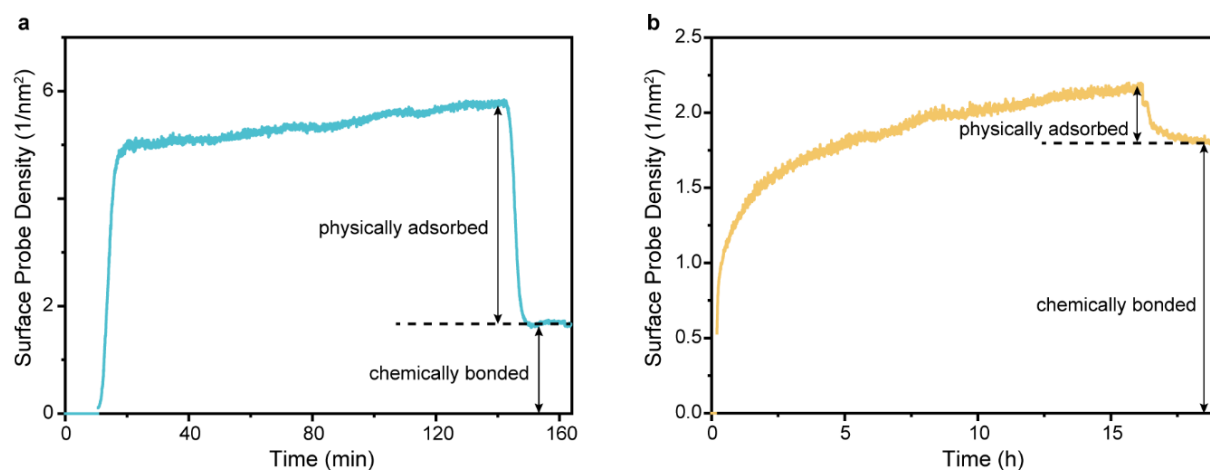
Before the deposition of rGO, the QCM gold sensor surface was cleaned with SDS and ethanol and treated with air plasma. No linker was added between the gold surface and the rGO layer as their adhesion is strong enough to hold rGO layer stably on the gold surface<sup>11</sup>. GO was deposited by spin-coating using ethanol as the solvent, which was later reduced @ 180 °C under N<sub>2</sub> for 6 hours. The preparation protocols of rGO on QCM sensors are different from that on RGFET



sensors due to their different surface properties and the special requirements with QCM sensors (cannot tolerate high temperatures and cannot be cut), but the resulting rGO films are similar as characterized by WCA and AFM.

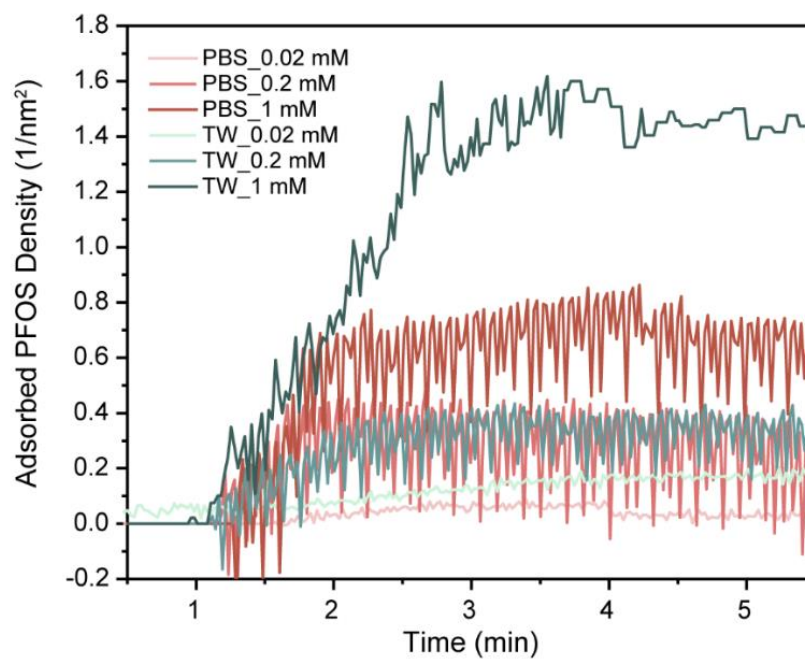


**Supplementary Figure 15. Surface adsorption study by QCM-D.** **a**, Illustration of QCM sensor surface during measurement. **b**, QCM curves at several overtones of the fundamental frequency during PBASE immobilization process (inset: illustration of the surface chemistry) with calculated probe density. **c**, QCM curves at several overtones of the fundamental frequency during  $\text{NH}_2$ - $\beta$ -CD immobilization process (inset: illustration of the surface chemistry) with calculated probe density.

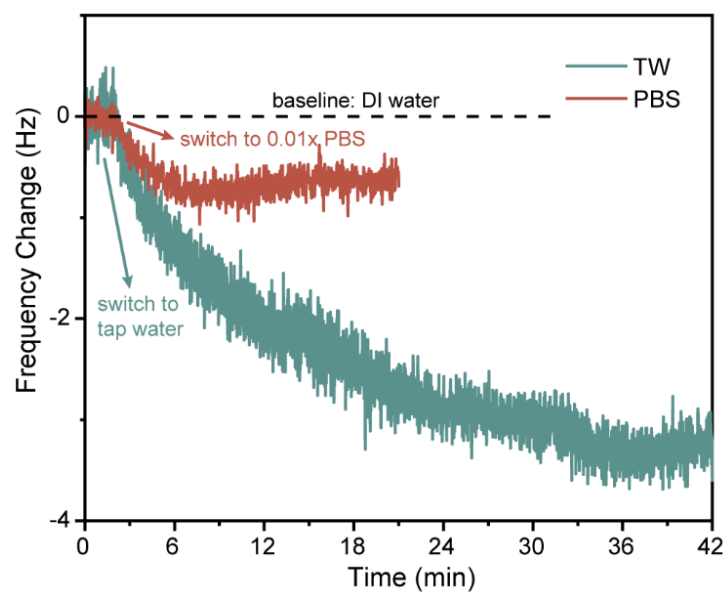


**Supplementary Figure 16. Surface probe density changes during the functionalization process (fitted from frequency change curve).** **a**, PBASE functionalization. **b**,  $\beta$ -CD functionalization.

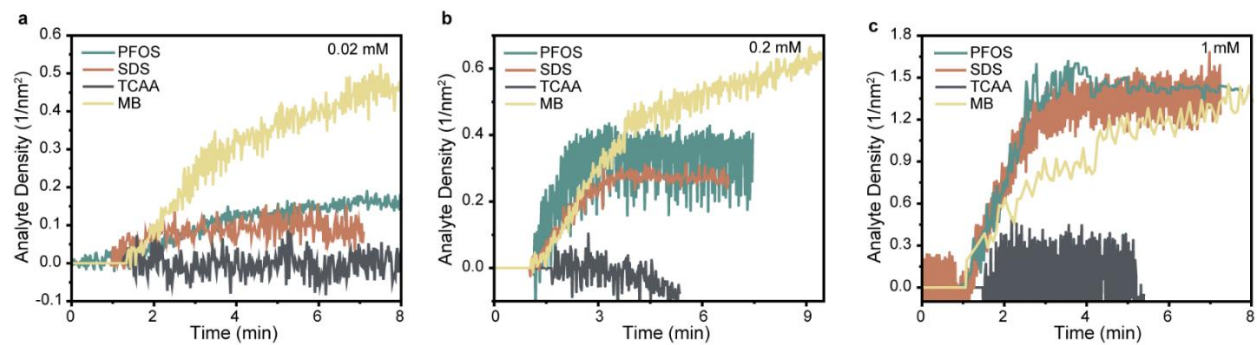
After rinsing with baseline solutions (as labeled in Supplementary Fig. 15b,c) for enough time to remove the weakly adsorbed molecules, the remaining frequency change ( $\Delta f$ ) represents the mass that has been immobilized on the surface. These results confirm the effective PBASE and  $\text{NH}_2$ - $\beta$ -CD functionalization during the RG electrode preparation. The frequency change curves can be converted into probe density change curves (Supplementary Fig. 16) using appropriate models (see Methods in the main text) to calculate the final surface probe density. PBASE and  $\text{NH}_2$ - $\beta$ -CD densities were measured to be 1.7 molecules/ $\text{nm}^2$  and 1.8 molecules/ $\text{nm}^2$ , respectively. The similar surface densities of these two probes suggest that almost all the PBASE linkers reacted with the  $\text{NH}_2$ - $\beta$ -CD probe, indicating excellent  $\beta$ -CD functionalization efficiency.



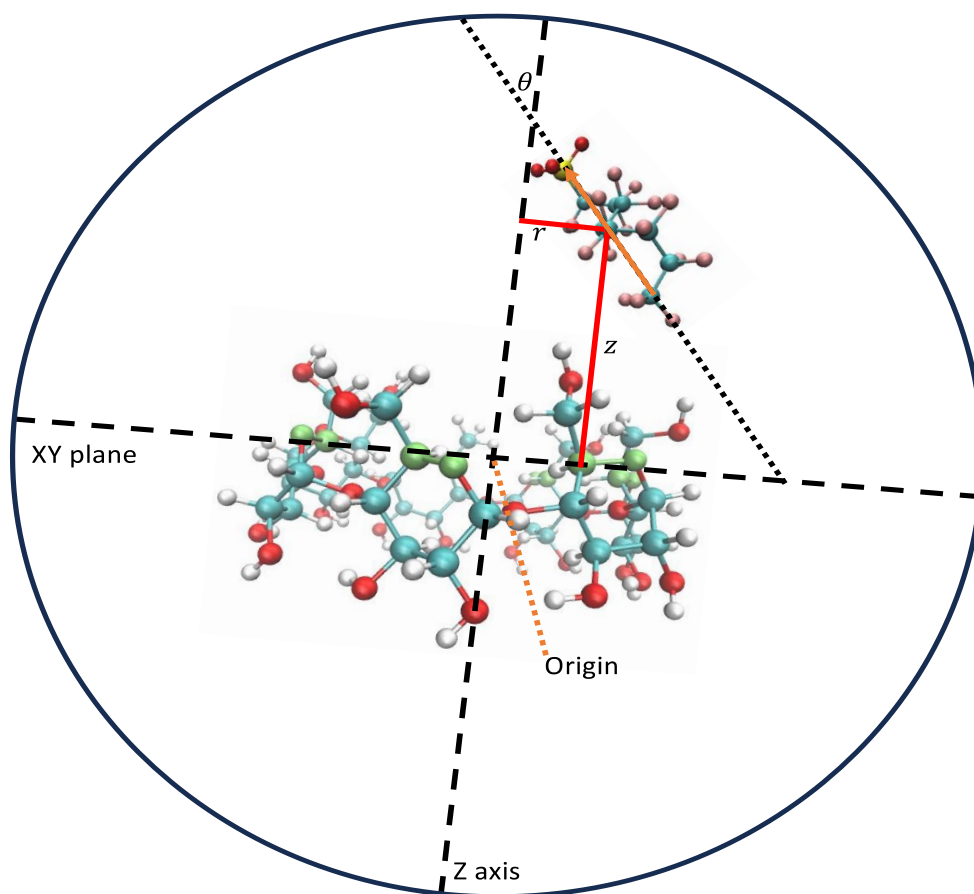
**Supplementary Figure 17. Surface probe density changes for PFOS in tap water and 0.01x PBS during the adsorption process (fitted from frequency change curve).**



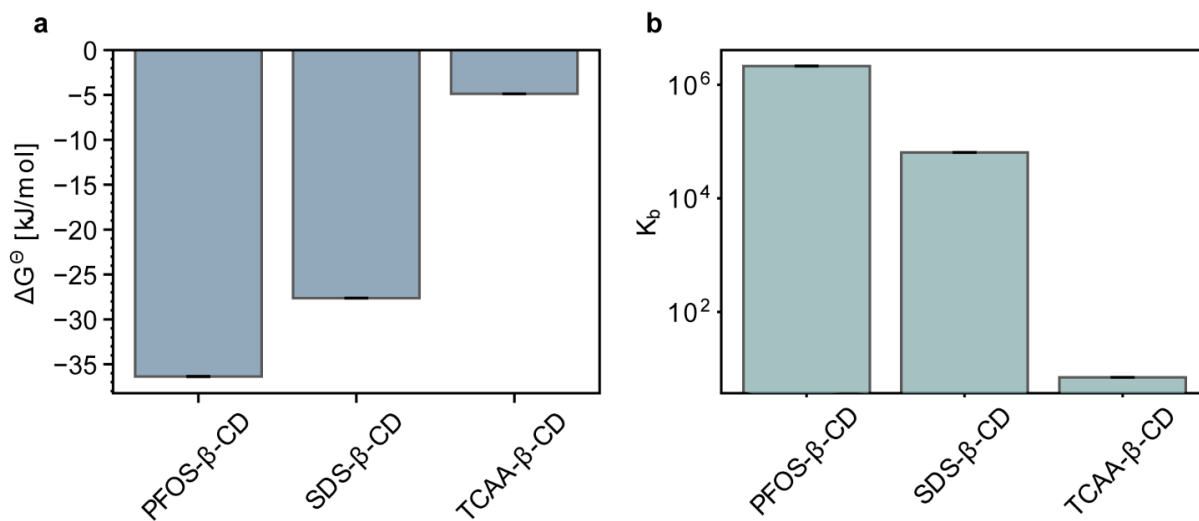
**Supplementary Figure 18. Frequency (f<sub>5</sub>) changes when switching from DI to tap water and 0.01x PBS (fitted from frequency change curve).**



**Supplementary Figure 19. Surface probe density changes for PFOS, SDS, TCAA, and MB in tap water during the adsorption process (fitted from frequency change curve). a, 0.02 mM. b, 0.2 mM. c, 1 mM.**



**Supplementary Figure 20. Illustration of the cyclodextrin-centric coordinate system and PBMetaD Collective Variables (CVs) within this reference frame for  $\beta$ -CD probe and PFOS analyte system.** The XY plane is defined by the carbon and oxygen atoms that are colored in green in the narrow-side plane of each 6-sided D-glucose units in each  $\beta$ -CD probe and the Z axis is defined as the normal vector to that plane. The three PBMetaD CVs correspond to (i)  $z$ : the Z coordinate of the center of mass (COM) of the target molecule, (ii)  $r$ : the radial distance of the COM of the target molecule away from the Z axis in the XY plane, and (iii)  $\theta$ : the planar angle between a vector drawn between the tail and head of the target molecule and the Z axis. Carbon, hydrogen, oxygen atoms are colored in cyan, white, and red, respectively. Molecular renderings were produced with VMD<sup>12</sup>.



**Supplementary Figure 21. Calculated values of the standard state. a,** binding free energy  $\Delta G^\ominus$ . **b,** binding constant  $K_b$  for a  $\beta$ -CD probe with the PFOS, SDS and TCAA molecules.



**Supplementary Table 4. Binding constant  $K_b$  computed under different choices for the cutoff**

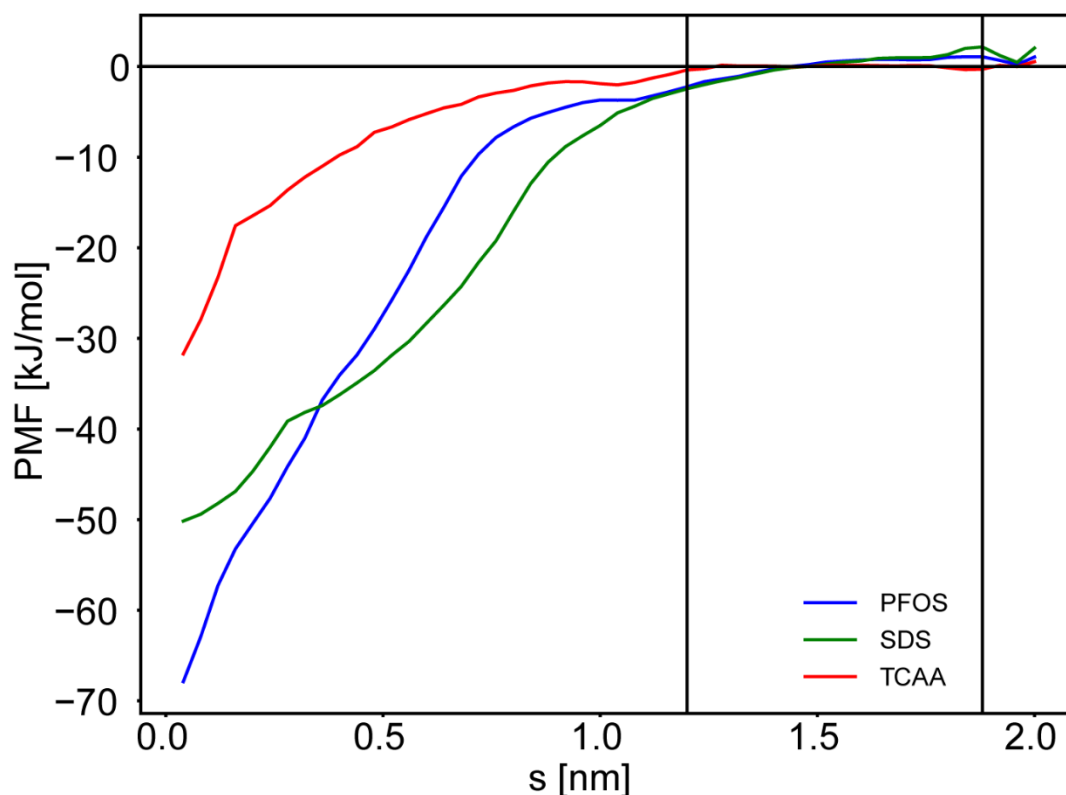
**volume delimiting the bound state**  $s \leq s_{\text{bound}} = \sqrt{r_{\text{bound}}^2 + z_{\text{bound}}^2}$ . We employ a value of  $s_{\text{bound}} = 1.2$  nm in our development above (cf. Supplementary Fig. 22), but the absolute values and trends are quite robust to the particular choice.

$K_b$	$s_{\text{bound}} \leq 0.8$ nm	$s_{\text{bound}} \leq 1.0$ nm	$s_{\text{bound}} \leq 1.2$ nm	$s_{\text{bound}} \leq 1.4$ nm	$s_{\text{bound}} \leq 1.6$ nm
PFOS- $\beta$ -CD	$(1.82 \pm 0.04) \times 10^6$	$(1.92 \pm 0.06) \times 10^6$	$(2.15 \pm 0.05) \times 10^6$	$(2.41 \pm 0.03) \times 10^6$	$(2.37 \pm 0.03) \times 10^6$
SDS- $\beta$ -CD	$(4.46 \pm 0.11) \times 10^4$	$(5.45 \pm 0.08) \times 10^4$	$(6.46 \pm 0.04) \times 10^4$	$(7.86 \pm 0.12) \times 10^4$	$(9.13 \pm 0.10) \times 10^4$
TCAA- $\beta$ -CD	$(1.86 \pm 0.04) \times 10^0$	$(3.74 \pm 0.01) \times 10^0$	$(7.04 \pm 0.05) \times 10^0$	$(1.21 \pm 0.06) \times 10^1$	$(1.83 \pm 0.06) \times 10^1$

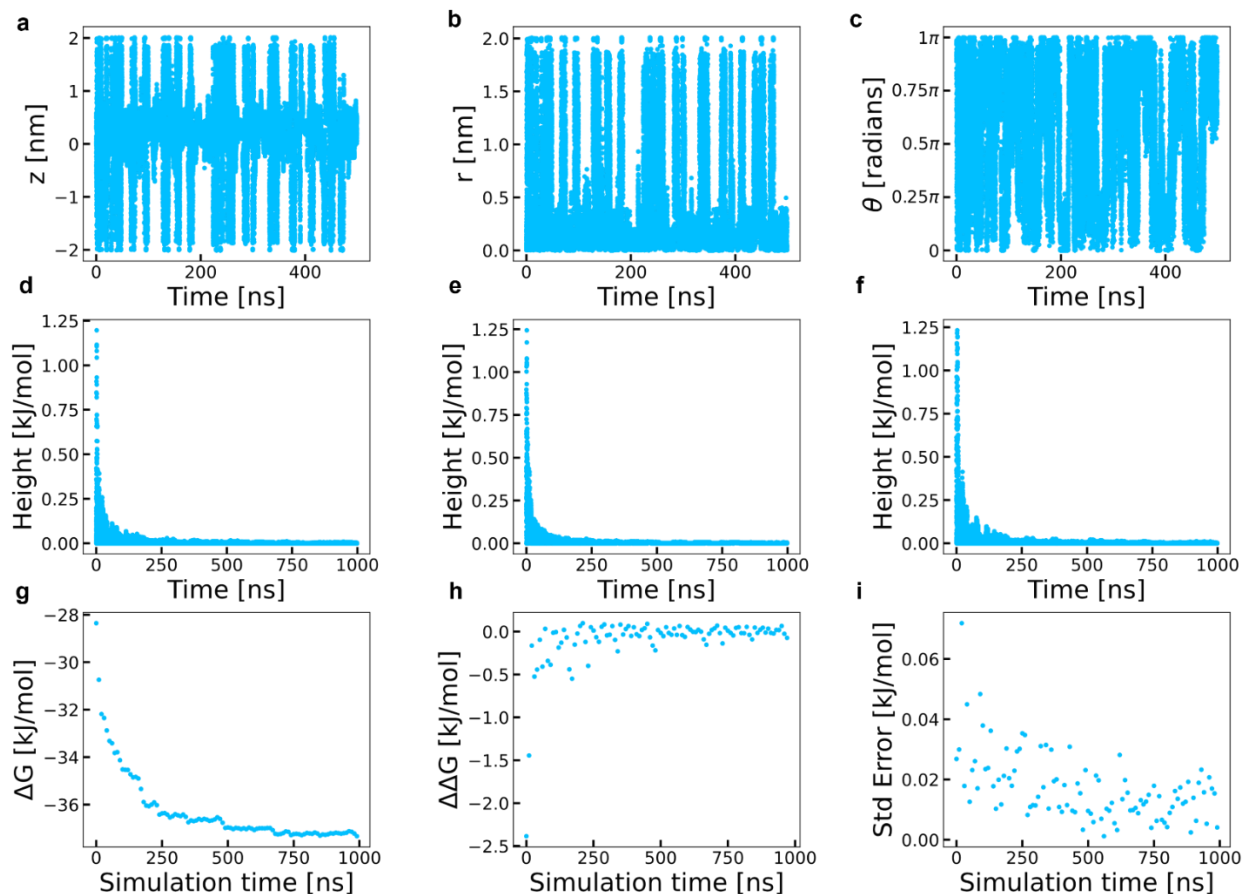
**Supplementary Table 5. Binding free energies  $\Delta G^\theta$  computed under different choices for the**

**cutoff volume delimiting the bound state**  $s_{\text{bound}} = \sqrt{r_{\text{bound}}^2 + z_{\text{bound}}^2}$ . We employ a value of  $s_{\text{bound}} = 1.2$  nm in our development above (cf. Supplementary Fig. 22), but the absolute values and trends are quite robust to the particular choice.

$\Delta G^\theta$ [kJ/mol]	$s_{\text{bound}} \leq 0.8$ nm	$s_{\text{bound}} \leq 1.0$ nm	$s_{\text{bound}} \leq 1.2$ nm	$s_{\text{bound}} \leq 1.4$ nm	$s_{\text{bound}} \leq 1.6$ nm
PFOS- $\beta$ -CD	$-35.954 \pm 0.054$	$-36.093 \pm 0.007$	$-36.377 \pm 0.044$	$-36.652 \pm 0.027$	$-36.612 \pm 0.029$
SDS- $\beta$ -CD	$-26.723 \pm 0.057$	$-27.222 \pm 0.014$	$-27.632 \pm 0.029$	$-28.079 \pm 0.032$	$-28.505 \pm 0.098$
TCAA- $\beta$ -CD	$-1.545 \pm 0.056$	$-3.292 \pm 0.006$	$-4.943 \pm 0.007$	$-6.222 \pm 0.012$	$-7.256 \pm 0.008$



**Supplementary Figure 22. Free energy landscapes  $PMF(s)$  projected into the COM separation  $s = \sqrt{r^2 + z^2}$  between the  $\beta$ -CD probe and the three targets, the PFOS analyte and SDS and TCAA interferents.** The data for each system are distinguished by lines of different colors. In all cases, the PMF plots illustrate that the effectively non-interacting regime is reached in the far field where the free energy ceases to be a function of  $s$ . Landscapes are adjusted by the entropic correction accounting for the trivial increase in configurational phase space volume associated with increasing  $s$  and which physically corresponds to subtracting out the  $PMF(s)$  corresponding to non-interacting point particles<sup>13-15</sup>. The two vertical lines delineate the plateau region defined as the range spanning  $1.2 \leq (s = \sqrt{r^2 + z^2}) \leq \sqrt{3.5}$  nm.



**Supplementary Figure 23. Illustration of convergence of the funnel well-tempered parallel bias metadynamics calculations for the  $\beta$ -CD–PFOS system.** **a-c**, CVs  $z$ ,  $r$  and  $\theta$  (from left to right) as a function of simulation time. **d-f**, Bias heights deposited as a function of time for CVs  $z$ ,  $r$  and  $\theta$  (from left to right). **g-i**,  $\Delta G$  and related quantities computed as a function of simulation time. Points are computed at 10 ns intervals of simulation length. **g**, Calculated  $\Delta G$ . **h**, change in  $\Delta G$  computed for each marginal 10 ns of simulation. **i**, standard error in  $\Delta G$  computed from 3-fold block averaging.

## **Supplementary Note 1. All-atom molecular dynamics simulations of analyte interactions with $\beta$ -CD probes.**

We study the interactions of the PFOS analyte and SDS and TCAA interferent molecules with a  $\beta$ -CD probe using all-atom molecular dynamics (MD) simulations conducted in GROMACS 2021.6<sup>13</sup>. The initial structures of PFOS, SDS and TCAA target molecules were generated using Avogadro version 1.2.0<sup>16</sup> in PDB format. The PFOS molecule was generated in anionic state<sup>17, 18</sup> corresponding to the pH 8 conditions explored in this work. The restrained electrostatic potential (RESP) method<sup>19</sup> was used to compute partial charges with Gaussian 16RevA.03<sup>20</sup> software and General Amber Force Field (GAFF2)<sup>21</sup> force field parameters. System topologies were generated in GROMACS<sup>13</sup> format using the acpype<sup>22</sup> library. The GAFF2 compatible q4md-CD force field<sup>23, 24</sup> combining GLYCAM04<sup>25-27</sup> and Amber99SB<sup>28</sup> was used to model the glucose units in the  $\beta$ -CD probe. The structure and q4md-CD force field files in GROMACS format of the  $\beta$ -CD probe was obtained from Zhang et al.<sup>29</sup> (<https://www.virtualchemistry.org/ff.php#refs>). Calculations were performed in cubic boxes of at least  $6 \times 6 \times 6 \text{ nm}^3$  with periodic boundary conditions implemented in all three dimensions. The linear dimensions of the box were sufficiently large for the target-probe pair to reach the effectively non-interacting regime in the far field. The PFOS, SDS, or TCAA molecule being simulated was initialized in the center of the  $\beta$ -CD cavity for each calculation. Each system was solvated with TIP3P water molecules<sup>30</sup> to a density of  $1.0 \text{ g/cm}^3$  of water and PFOS systems were charge neutralized by the addition of a single sodium counterion to balance the formal  $(-1)$  net charge carried by PFOS. The total number of atoms in the simulations lie in the approximate range of 12,000-22,000. Forces arising due to the Lennard-Jones potential were switched smoothly to zero at 1 nm. Electrostatic interactions were treated using the particle mesh Ewald<sup>31</sup> method with a 1 nm real-space cutoffs and a reciprocal grid spacing optimized

during runtime. The neighbor list was generated every 10 steps using the Verlet method<sup>32</sup>. The LINCS algorithm<sup>33</sup> was used to update the hydrogen atomic positions relative to their bonded non-hydrogen atoms in the system. High-energy overlaps in the initial system configuration were eliminated by steepest descent energy minimization using the low memory Broyden-Fletcher-Goldfarb-Shanno approach<sup>34</sup> employing a maximum force tolerance of  $10 \text{ kJ mol}^{-1} \text{ nm}^{-1}$  on any single atom. Initial velocities corresponding to  $T = 300 \text{ K}$  were drawn from the Maxwell-Boltzmann distribution. A harmonic position restraint with force constant of  $k = 500 \text{ kJ/mol}\cdot\text{nm}^2$  was applied during these equilibration runs to the positions of the target or interferent-probe complex to prevent drift from their initial positions within the  $\beta$ -CD cavity. The energy minimized systems were equilibrated for 1 ns in a NVT ensemble at temperature  $T = 300 \text{ K}$  with a stochastic velocity rescaling thermostat<sup>35</sup>. Next, a 1 ns NPT ensemble equilibration with a target  $T = 300 \text{ K}$  and a target pressure  $P = 1 \text{ bar}$  was performed employing a stochastic velocity rescaling thermostat<sup>35</sup> and a Berendsen barostat<sup>36</sup>. Equations of motion were integrated with a 2 fs time step using the leapfrog algorithm<sup>37</sup>.

## **Supplementary Note 2. Enhanced sampling using funnel well-tempered parallel bias metadynamics.**

Enhanced sampling techniques are required to accelerate sampling of the interactions between the  $\beta$ -CD probe and the PFOS, SDS or TCAA molecules and assure good exploration of the thermally relevant configurational phase space. We employed parallel bias metadynamics (PBMetaD), a technique in which the molecular dynamics simulation is subjected to biasing potentials along pre-selected collective variables (CVs) to accelerate transitions over free energy barriers in these coordinates and drive good sampling of configurational space<sup>38</sup>. Unbiased free energy landscapes can be extracted from the converged PBMetaD trajectories using thermodynamic reweighting of the configurations according to the applied converged biasing potential to recover the Boltzmann distribution in the thermodynamic ensemble under which the simulation was conducted<sup>39</sup>. We define our PBMetaD CVs within a coordinate system defined by the  $\beta$ -CD probe. Since we are interested in the configurational phase space sampled by the target molecule relative to the probe (or vice versa), it is convenient to use a coordinate system centered on one of the molecules rather than the absolute coordinate system of the simulation box to mod out rigid translational and rotational displacements that alter the absolute position of the two molecules but not relative to one another. We define this coordinate system as illustrated in Supplementary Fig. 20. The origin point is defined as the center of mass (COM) of all carbon and oxygen atoms in the narrow-side plane (primary face) of the D-glucose units of  $\beta$ -CD probes. The  $XY$  plane is defined as this plane constructed from the carbon and oxygen atoms in the narrow-side plane of each 6-sided D-glucose units that constitute the  $\beta$ -CD probes. The  $Z$  axis is defined via the normal vector to this plane passing through the origin of the coordinate system defined at the center of mass (COM) of all carbon and oxygen atoms in the  $\beta$ -CD rings and with the positive sign of the  $z$ -axis pointing from

the broader to the narrower end of the  $\beta$ -CD molecule. Due to the high level of angular symmetry, our PBMetaD CVs do not require definition of the X and Y axes, so it is sufficient to define just the XY plane and the Z axis.

We now define the three CVs used to drive our PBMetaD calculations (i)  $z$ : the Z coordinate of the COM of the target molecule, (ii)  $r$ : the radial distance of the COM of the target molecule away from the Z axis in the XY plane, and (iii)  $\theta$ : the planar angle between a vector drawn between the head and tail of the target molecule and the Z axis. Physically,  $z$  controls the vertical separation between the  $\beta$ -CD probe and the target molecule, while  $r$  controls the orthogonal distance of the target molecule from the axis of the  $\beta$ -CD ring. In selecting as our CV the radial distance  $r$  – as opposed to its decomposition into the  $x$  and  $y$  components, where  $r^2 = x^2 + y^2$  – we have implicitly moded out the approximate cylindrical symmetry of the cyclodextrin ring in the XY plane, and chosen  $r$  to drive sampling in a single lumped CV that just controls the radial component of the intermolecular COM separation. Finally,  $\theta$  characterizes whether the target molecule approaches the cyclodextrin ring head-first or tail-first, and driving sampling in this CV helps sample both possibilities as the target molecule approaches the narrow ( $z > 0$ ) and broad ( $z < 0$ ) sides of the cyclodextrin ring. Since the target molecule can exist as a guest within the core of the  $\beta$ -CD host, this CV is particularly useful in driving sampling of the different docked orientations.

Having defined the PBMetaD CVs, we then conduct the enhanced sampling calculations employing a well-tempered metadynamics procedure<sup>40</sup> in which we use a bias factor of  $\gamma = 20$ , an initial Gaussian height of  $W = 1.2$  kJ/mol, and a Gaussian width of  $\sigma = 0.2$  nm for  $z$  and  $r$  and 0.2 radians for  $\theta$ . We deposit Gaussians with a pace corresponding to once every 500 time steps. Furthermore, to focus PBMetaD sampling on the region of configurational space of most interest

where the target molecule is close to the  $\beta$ -CD probe, we utilized funnel metadynamics<sup>41</sup> following a procedure similar to that detailed in Rizzi et al.<sup>42</sup>. The principle of the approach is that within the funnel region defined within the cyclodextrin-centric coordinate system, the target molecule does not feel any additional restraining forces and the metadynamics simulation proceeds as normal, whereas outside the funnel region, the target molecule experiences restraining forces that drive it back into the funnel. This focuses sampling on the region of space within the funnel to improve computational efficiency and the bias introduced by the restraining potentials can be analytically removed using reweighting<sup>39</sup>. In this work, we elect to use a spherical “funnel” of radius 2 nm centered on the origin of the cyclodextrin-centric coordinate system (Supplementary Fig. 20) and applied a harmonic restraining potential at the funnel walls with a force constant of  $k = 5,000,000$  kJ/mol·nm<sup>2</sup>. As demonstrated in Supplementary Fig. 22, the volume of the funnel was sufficiently small to achieve good sampling of the full configurational phase space within the funnel using PBMetaD, and sufficiently large that the target molecule was adequately far from the probe at the funnel wall that it entered the effectively non-interacting regime such that the free energy landscape plateaued and no longer changed as a function of COM distance between the molecules.

We conducted the funnel well-tempered PBMetaD simulations in the NPT ensemble at  $T = 300$  K and  $P = 1$  bar using a Nosé-Hoover thermostat<sup>43</sup> and Parrinello-Rahman barostat<sup>44</sup> using GROMACS 2021.6 patched with PLUMED-2.8.1<sup>45</sup>. The simulation parameters were identical to those detailed above for system initialization, with the exception that integration of the equations of motion was performed with a 2 fs time step for both SDS and PFOS and a 1 fs time step for TCAA. Empirically, we found the TCAA system required a shorter timestep to maintain numerical stability of the numerical integration. Snapshots were saved for analysis at a period of 2 ps. Following best metadynamics practices<sup>39</sup>, simulation convergence was validated by (i) monitoring



for diffusive sampling in the three CVs, (ii) measuring the fall-off in the heights of the deposited Gaussian biases in each CV deposited over time, and (iii) assessing convergence of the calculated  $\Delta G$ s of interaction for each simulated system using a block average estimate of uncertainty. Convergence was defined to have been reached when (i) transitions in CV-time plot are diffusive (Supplementary Fig. 23a-c) (ii) deposited bias height falls off to less than  $1/100^{\text{th}}$  of the initial height (Supplementary Fig. 23d-f), and (iii)  $\Delta G$ , uncertainties in  $\Delta G$ , and change in  $\Delta G$  as a function of simulation length remain static with increased simulation length (Supplementary Fig. 23g-i). In practice, we ran calculations for each of the three systems consisting of a  $\beta$ -CD probe with one of PFOS, SDS or TCAA targets for a total of  $1\ \mu\text{s}$  (500,000,000 steps for PFOS and SDS or 1,000,000,000 steps for TCAA). We observed convergence in all three criteria within approximately 500 ns. The production MD simulation with enhanced sampling simulation took approximately 4-11 days (between 90-250 ns/day) per system, depending on the system size and integration time step, on a shared supercomputing node utilizing either  $4 \times$  NVIDIA Quadro RTX 5000 GPUs and  $2 \times$  Intel Xeon E5-2620 v4 Broadwell CPUs or  $1 \times$  V100 GPU and  $20 \times$  CPU cores of Intel Skylake processor.

### Supplementary Note 3. Estimation of equilibrium binding free energies $\Delta G$ and dimensionless standard state binding constants $K_b$ .

We extract from the funnel well-tempered PBMetaD simulations for each of the three systems estimates of the equilibrium binding free energies and standard state binding constants. To do so, we compile biased histograms  $\tilde{p}(r, z)$  in the radial and vertical CVs used to drive our PBMetaD calculations over the converged portion of our simulation trajectories. Since the bias continues to (slowly) evolve over the converged portion, to obtain the unbiased histograms  $p(r, z)$ , up to an unknown normalizing constant, we are obliged to reweight each frame according to the instantaneous bias applied in the CVs and funnel restraints using the reweighting method detailed in Zhao et al.<sup>46</sup>. The unknown normalizing constant reflects our ignorance of the absolute free energy scale or, equivalently, the partition function, although the probability ratio of various  $(r, z)$  configurations can be quantitatively compared in a relative sense. As an aside, although the 2D  $r$ - $z$  projection turns out to be a useful projection of the free energy surface for this system, we note that we are at liberty to accumulate unbiased histograms in this manner in any CVs of our choosing provided we properly reweight as a function of the PBMetaD CVs (and funnel restraints) in which the accelerating biases were applied<sup>39</sup>. In this particular case, we conducted enhanced sampling in  $\{r, z, \theta\}$  but chose to project our free energy landscapes into  $\{r, z\}$  and marginalize out  $\theta$  such that we integrate over the various different probe orientations sampled by our PBMetaD calculations at each  $(r, z)$  configuration.

Having obtained our estimates of the unbiased, but unnormalized, probability distributions through our histogram approximation, we then convert this into a potential of mean force (PMF) using the statistical mechanical relation  $PMF(r, z) = -k_B T \log(p(r, z)) + C$ , where  $C$  is an arbitrary additive constant reflecting our ignorance of the absolute free energy scale. Since we will

only appeal to relative free energies (i.e., free energy differences or probability ratios) we are at liberty to specify  $C$  for numerical convenience, although the specific choice does not affect any downstream numerical results, including our calculation of  $K_b$ . In **Fig. 6a-c** in the main text, we present the  $PMF(r, z)$  for the three systems to resolve the preferred binding location of each target molecule with each of the three  $\beta$ -CD probes and choose to set  $C$  such that the mean PMF of the non-interacting region has a free energy of zero for visualization convenience. For the purposes of the binding constant calculations described below, it is also convenient to adopt a convention wherein the PMF takes on a value of zero in the far field where the target and probe are distantly separated at large  $r$  and/or  $z$  corresponding to the effectively non-interacting state (i.e.,  $PMF(r \rightarrow \infty, z \rightarrow \infty)=0$ ). In practice, we set  $C$  for each system by projecting the PMF into the COM separation  $s = \sqrt{r^2 + z^2}$  between the target and the probe, subtracting out the entropic correction<sup>13-15</sup>, and computing the mean value of  $PMF(s)$  in the plateau region  $1.2 \leq (s = \sqrt{r^2 + z^2}) \leq \sqrt{3.5}$ . In performing this calculation, we exclude the region near the funnel wall to avoid edge effects associated with the large restraining potentials in these regions of configurational space. We present in Supplementary Fig. 22 the resulting  $PMF(s)$  free energy landscapes for each of the three target-probe systems. These plots confirm that the box size was sufficiently large in each case that the non-interacting far field region was sampled in the calculations corresponding to the plateau region where the free energy no longer changes with increasing separation. As described below, this observation is important for the purposes of reliable estimation of binding constants.

We estimate the free energy difference  $\Delta G$  between the bound and unbound states using a statistical mechanical formalism based on the ratio of the partition functions in each of the two states<sup>47-49</sup>. Specifically, we define,

$$\Delta G = -k_B T \ln \left( \frac{Z_b}{Z_u} \right), \quad (\text{S1})$$

where the bound and unbound state partition functions are,

$$Z_b = \int_{\Gamma_b} \exp(-\beta U(\mathbf{x})) d\mathbf{x} \quad (\text{S2})$$

$$Z_u = \int_{\Gamma_u} \exp(-\beta U(\mathbf{x})) d\mathbf{x}, \quad (\text{S3})$$

where  $U(\mathbf{x})$  is the potential energy of the system as a function of the full dimensional configurational phase space coordinates  $\mathbf{x}$ ,  $\Gamma_b$  and  $\Gamma_u$  respectively delimit the configurational phase space volume occupied of the bound and unbound states, and  $\beta = 1/(k_B T)$ .

We choose to specify the extent of the bound state as a sphere of radius 1.2 nm centered around the COM of the  $\beta$ -CD probe. The limits of this bounded region were motivated by the  $PMF(s)$  flattening as we enter the non interacting region, shown in Supplementary Fig. 22. Alternative definitions for the bound state can result in slightly different values for  $\Delta G$ , and we quantify these differences below. Importantly, this sphere is contained fully within the spherical funnel region used in the funnel well-tempered PBMetaD calculations and so we need not be concerned about any issues of poor sampling outside of the funnel region. The bound state partition function then becomes,

$$Z_b = 2\pi \int_0^{r_{\text{bound}}} \int_{-z_{\text{bound}}}^{z_{\text{bound}}} r \exp(-\beta(PMF(r, z) - PMF(r \rightarrow \infty, z \rightarrow \infty))) dr dz \quad (\text{S4})=$$

$$2\pi \int_0^{r_{\text{bound}}} \int_{-z_{\text{bound}}}^{z_{\text{bound}}} r \exp(-\beta PMF(r, z)) dr dz, \quad (\text{S5})$$

We perform the integration numerically with 2D Simpson's rule by setting  $r_{\text{bound}}$  and  $z_{\text{bound}}$  such that  $s_{\text{bound}} = \sqrt{r_{\text{bound}}^2 + z_{\text{bound}}^2} = 1.2$  nm, using a form consistent with the bounded region and limits informed by  $PMF(s)$  (Supplementary Fig. 22). In the first line, we note that the

potential energy in the full configurational phase space coordinates  $U(\mathbf{x})$  has been replaced by the potential of mean force  $PMF(r, z)$  in the Boltzmann factor due to marginalization over all of the other degrees of freedom in constructing the cylindrical integral, and we have performed the angular integral since there is no dependence on  $\theta$  in  $PMF(r, z)$ . In the second line, we note that under our specification of the arbitrary additive constant in the potential of mean force, we have specified that  $PMF(r \rightarrow \infty, z \rightarrow \infty) = 0$ . We note that asserting this limit is numerically convenient but the final results of our  $\Delta G$  calculation do not depend on this choice due to cancelation in the numerator and denominator of the ratio in Eqn. S1.

The extent of the unbound state is naturally defined as the volume of the simulation box not occupied by the bound state. For a sufficiently large simulation box and rapid decay of  $PMF(r, z)$  to the  $PMF(r \rightarrow \infty, z \rightarrow \infty)$  limit, it is a reasonable approximation to take  $PMF(r, z) = PMF(r \rightarrow \infty, z \rightarrow \infty) = 0$  and perform the integral over the full volume of the simulation box. In the present case, assuming the restraining potentials in our metadynamics funnel to be sufficiently strong, we may approximate them as hard walls and conduct the integration over the volume of the spherical funnel,

$$Z_u \approx \int_0^{r_{\text{funnel}}} 4\pi r^2 \exp(-\beta PMF(\infty)) dr \quad (\text{S6})$$

$$= \int_0^{r_{\text{funnel}}} 4\pi r^2 dr \quad (\text{S7})$$

$$= \frac{4}{3} \pi r_{\text{funnel}}^3 \quad (\text{S8})$$

$$= V_{\text{funnel}}, \quad (\text{S9})$$

where  $r_{\text{funnel}} = 2 \text{ nm}$  is the radius of the spherical funnel and  $V_{\text{funnel}} = \frac{32}{3} \pi \text{ nm}^3$  is the volume of the funnel.

Combining Eqns. S1, S5 and S9, we arrive at our final expression for  $\Delta G$ ,

$$\Delta G = -k_B T \ln \left[ \frac{2\pi}{V_{\text{funnel}}} \int_0^{r_{\text{bound}}} \int_{-z_{\text{bound}}}^{z_{\text{bound}}} r \exp(-\beta PMF(r, z)) dr dz \right] \quad (\text{S10})$$

We recognize that the concentration of each species in the simulation box is  $C = C_{\text{probe}} = C_{\text{target}} = C_{\text{probe-target}} = 1/V_{\text{funnel}}$  and that, in general,  $C$  differs from the standard concentration of  $C^\theta = 1 \text{ M}$  corresponding to a standard volume of  $V^\theta = 1/C^\theta = 1661 \text{ \AA}^3$ <sup>50</sup>. We are obliged, therefore, to apply a concentration correction to estimate the binding free energy at the standard concentration<sup>47</sup>,

$$\Delta G^\theta = \Delta G + k_B T \ln \left( \frac{C}{C^\theta} \right) \quad (\text{S11})$$

$$= -k_B T \ln \left[ C^\theta 2\pi \int_0^{r_{\text{bound}}} \int_{-z_{\text{bound}}}^{z_{\text{bound}}} r \exp(-\beta PMF(r, z)) dr dz \right] \quad (\text{S12})$$

Comparing this expression to the relation<sup>47</sup>,

$$\Delta G^\theta = -k_B T \ln K_b, \quad (\text{S13})$$

we identify the binding constant  $K_b$  as,

$$K_b = C^\theta 2\pi \int_0^{r_{\text{bound}}} \int_{-z_{\text{bound}}}^{z_{\text{bound}}} r \exp(-\beta PMF(r, z)) dr dz \quad (\text{S14})$$

We report in Supplementary Fig. 21 the calculated values of  $\Delta G^\theta$  and  $K_b$  from our funnel well-tempered PBMetaD calculations for the binding of the PFOS, SDS or TCAA molecule with the  $\beta$ -CD probes. Numerical quadrature of the 2D integral was performed using Simpson's rule. Uncertainties were estimated by three-fold block averaging over the converged portion of the simulation trajectories.

As mentioned above, the numerical values of  $\Delta G^\theta$  and  $K_b$  can depend on the definition of the bound state. In Supplementary Tables 4 and 5, we report the calculated values of  $\Delta G^\theta$  and  $K_b$  under different choices for the definition of the limits of the spherically defined bound state  $s \leq s_{\text{bound}} = \sqrt{r_{\text{bound}}^2 + z_{\text{bound}}^2} = 1.2 \text{ nm}$ . Above, we selected  $s_{\text{bound}} = 1.2 \text{ nm}$ , and our calculations for different values of  $s_{\text{bound}}$  reveal only minor sensitivity to the choice of bound state phase space volume, and that the trends in these quantities are preserved under all choices for the bound state.

### Supplementary References:

1. Kwong Hong Tsang, D. et al. Chemically functionalised graphene FET biosensor for the label-free sensing of exosomes. *Sci Rep* **9**, 13946 (2019).
2. Jang, H.J. et al. Remote floating-gate field-effect transistor with 2-dimensional reduced graphene oxide sensing layer for reliable detection of SARS-CoV-2 spike proteins. *ACS Appl Mater Interfaces* **14**, 24187-24196 (2022).
3. López-Díaz, D., López Holgado, M., García-Fierro, J.L. & Velázquez, M.M. Evolution of the Raman spectrum with the chemical composition of graphene oxide. *The Journal of Physical Chemistry C* **121**, 20489-20497 (2017).
4. Barbey, R., Laporte, V., Alnabulsi, S. & Klok, H.-A. Postpolymerization modification of poly(glycidyl methacrylate) brushes: an XPS depth-profiling study. *Macromolecules* **46**, 6151-6158 (2013).
5. Wengel, M., Kothe, E., Schmidt, C.M., Heide, K. & Gleixner, G. Degradation of organic matter from black shales and charcoal by the wood-rotting fungus *Schizophyllum commune* and release of DOC and heavy metals in the aqueous phase. *Sci Total Environ* **367**, 383-393 (2006).
6. Cui, Y., Wen, S., Stegen, J.C., Hu, A. & Wang, J. Chemodiversity of riverine dissolved organic matter: Effects of local environments and watershed characteristics. *Water Res* **250**, 121054 (2023).
7. Leenheer, J.A.a.C., J.P. Peer reviewed: characterizing aquatic dissolved organic matter. *Environmental Science & Technology* **37**, 18A-26A (2003).
8. Lowe, B.M., Skylaris, C.K., Green, N.G., Shibuta, Y. & Sakata, T. Molecular dynamics simulation of potentiometric sensor response: the effect of biomolecules, surface morphology and surface charge. *Nanoscale* **10**, 8650-8666 (2018).
9. Zenobio, J.E., Salawu, O.A., Han, Z. & Adeleye, A.S. Adsorption of per- and polyfluoroalkyl substances (PFAS) to containers. *Journal of Hazardous Materials Advances* **7** (2022).
10. Woudneh, M.B., Chandramouli, B., Hamilton, C. & Grace, R. Effect of sample storage on the quantitative determination of 29 PFAS: observation of analyte interconversions during storage. *Environ Sci Technol* **53**, 12576-12585 (2019).
11. Torres, J., Zhu, Y., Liu, P., Lim, S.C. & Yun, M. Adhesion energies of 2D graphene and MoS<sub>2</sub> to silicon and metal substrates. *physica status solidi (a)* **215**, 1700512 (2018).
12. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **14**, 33-38 (1996).
13. Abraham, M.J. et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19-25 (2015).
14. Páll, S., Abraham, M.J., Kutzner, C., Hess, B. & Lindahl, E. in Solving software challenges for exascale. (ed. S.a.L. Markidis, E. ) 3-27 (Springer International Publishing, Cham; 2015).
15. Neumann, R.M. Entropic approach to Brownian movement. *American Journal of Physics* **48**, 354-357 (1980).
16. Hanwell, M.D. et al. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics* **4**, 17 (2012).
17. Meegoda, J.N., Kewalramani, J.A., Li, B. & Marsh, R.W. A review of the applications, environmental release, and remediation technologies of per- and polyfluoroalkyl



- substances. *International Journal of Environmental Research and Public Health* **17**, 8117 (2020).
18. Wang, Y., Darling, S.B. & Chen, J. Selectivity of per- and polyfluoroalkyl substance sensors and sorbents in water. *ACS Appl Mater Interfaces* **13**, 60789-60814 (2021).
  19. Bayly, C.I., Cieplak, P., Cornell, W. & Kollman, P.A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *The Journal of Physical Chemistry* **97**, 10269-10280 (1993).
  20. Frisch, M.J. et al. (Wallingford, CT; 2016).
  21. He, X., Man, V.H., Yang, W., Lee, T.-S. & Wang, J. A fast and high-quality charge model for the next generation general AMBER force field. *The Journal of Chemical Physics* **153**, 114502 (2020).
  22. Sousa da Silva, A.W. & Vranken, W.F. ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Research Notes* **5**, 367 (2012).
  23. Cézard, C., Trivelli, X., Aubry, F., Djedaïni-Pilard, F. & Dupradeau, F.-Y. Molecular dynamics studies of native and substituted cyclodextrins in different media: 1. Charge derivation and force field performances. *Physical Chemistry Chemical Physics* **13**, 15103-15121 (2011).
  24. Dupradeau, F.-Y. et al. The R.E.D. tools: advances in RESP and ESP charge derivation and force field library building. *Physical Chemistry Chemical Physics* **12**, 7821-7839 (2010).
  25. Kirschner, K.N. & Woods, R.J. Solvent interactions determine carbohydrate conformation. *Proceedings of the National Academy of Sciences* **98**, 10541-10545 (2001).
  26. Basma, M., Sundara, S., Çalgan, D., Vernali, T. & Woods, R.J. Solvated ensemble averaging in the calculation of partial atomic charges. *Journal of Computational Chemistry* **22**, 1125-1137 (2001).
  27. Kirschner, K.N. & Woods, R.J. Quantum mechanical study of the nonbonded forces in water-methanol complexes. *The Journal of Physical Chemistry A* **105**, 4150-4155 (2001).
  28. Cornell, W.D. et al. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society* **117**, 5179-5197 (1995).
  29. Zhang, H., Tan, T., Feng, W. & van der Spoel, D. Molecular recognition in different environments:  $\beta$ -cyclodextrin dimer formation in organic solvents. *The Journal of Physical Chemistry B* **116**, 12684-12693 (2012).
  30. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. & Klein, M.L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **79**, 926-935 (1983).
  31. Darden, T.A., York, D.M. & Pedersen, L.G. Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. *Journal of Chemical Physics* **98**, 10089-10092 (1993).
  32. Verlet, L. Computer "experiments" on classical fluids. I. thermodynamical properties of Lennard-Jones molecules. *Physical Review* **159**, 98-103 (1967).
  33. Hess, B., Bekker, H., Berendsen, H.J.C. & Fraaije, J.G.E.M. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **18**, 1463-1472 (1997).
  34. Saputro, D.R.S. & Widyaningsih, P. Limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method for the parameter estimation on geographically weighted ordinal logistic regression model (GWOLR). *AIP Conference Proceedings* **1868** (2017).

35. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics* **126** (2007).
36. Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A. & Haak, J.R. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* **81**, 3684-3690 (1984).
37. Hockney, R.W., Goel, S.P. & Eastwood, J.W. Quiet high-resolution computer models of a plasma. *Journal of Computational Physics* **14**, 148-158 (1974).
38. Pfaendtner, J. & Bonomi, M. Efficient sampling of high-dimensional free-energy landscapes with parallel bias metadynamics. *Journal of Chemical Theory and Computation* **11**, 5062-5067 (2015).
39. Bussi, G. & Laio, A. Using metadynamics to explore complex free-energy landscapes. *Nature Reviews Physics* **2**, 200-212 (2020).
40. Barducci, A., Bussi, G. & Parrinello, M. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Physical Review Letters* **100**, 020603 (2008).
41. Limongelli, V., Bonomi, M. & Parrinello, M. Funnel metadynamics as accurate binding free-energy method. *Proceedings of the National Academy of Sciences* **110**, 6358-6363 (2013).
42. Rizzi, V., Bonati, L., Ansari, N. & Parrinello, M. The role of water in host-guest interaction. *Nature Communications* **12**, 93 (2021).
43. Evans, D.J. & Holian, B.L. The Nose–Hoover thermostat. *The Journal of Chemical Physics* **83**, 4069-4074 (1985).
44. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* **52**, 7182-7190 (1981).
45. Bonomi, M. et al. PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Computer Physics Communications* **180**, 1961-1972 (2009).
46. Zhao, M. et al. MARTINI-compatible coarse-grained model for the mesoscale simulation of peptoids. *The Journal of Physical Chemistry B* **124**, 7745-7764 (2020).
47. Dasetty, S. et al. Data-driven discovery of linear molecular probes with optimal selective affinity for PFAS in water. *Journal of Chemical & Engineering Data* **68**, 3148-3161 (2023).
48. Aldeghi, M., Bluck, J.P. & Biggin, P.C. in *Computational drug discovery and design*. (eds. M. Gore & U.B. Jagtap) 199-232 (Springer New York, New York, NY; 2018).
49. General, I.J. A note on the standard state's binding free energy. *Journal of Chemical Theory and Computation* **6**, 2520-2524 (2010).
50. Hermans, J. & Wang, L. Inclusion of loss of translational and rotational freedom in theoretical estimates of free energies of binding. application to a complex of benzene and mutant T4 lysozyme. *Journal of the American Chemical Society* **119**, 2707-2714 (1997).