Soft Matter

Accepted Manuscript

This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

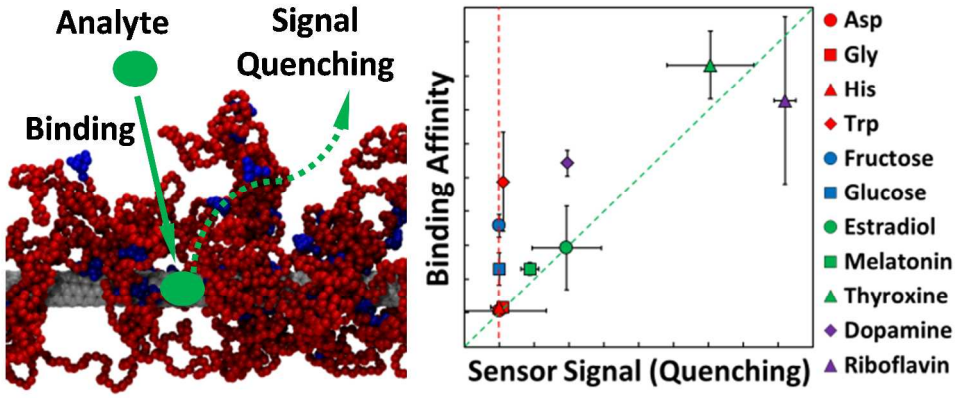
You can find more information about *Accepted Manuscripts* in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/softmatter

Figure Table of Content (TOC)



Soft Matter Accepted Manuscript

Understanding Selective Molecular Recognition in Integrated Carbon Nanotube-Polymer Sensors by Simulating Physical Analyte Binding on Carbon Nanotube-Polymer Scaffolds†

Shangchao Lin,[ab] Jingqing Zhang,[a] Michael S. Strano,[a] and Daniel Blankschtein*[a]

[a]Department of Chemical Engineering and [b]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. E-mail: dblank@mit.edu

†Electronic supplementary information (ESI) available. See DOI:

Soft Matter Accepted Manuscript

ABSTRACT

Macromolecular scaffolds made of polymer-wrapped single-walled carbon nanotubes (SWCNTs) have been explored recently (Zhang et al., *Nature Nanotechnology*, 2013) as a new class of molecular–recognition motifs. However, selective analyte recognition is still challenging and lacks the underlying fundamental understanding needed for its practical implementation in biological sensors. In this report, we combine coarse-grained molecular dynamics (CGMD) simulations, physical adsorption/binding theories, and photoluminescence (PL) experiments to provide molecular insight into the selectivity of such sensors towards a large set of biologically important analytes. We find that the physical binding affinities of the analytes on a bare SWCNT partially correlate with their distribution coefficients in a bulk water/octanol system, suggesting that the analyte hydrophobicity plays a key role in determining the binding affinities of the analytes considered, along with the various specific interactions between the analytes and the polymer anchor groups. Two distinct categories of analytes are identified to demonstrate a

complex picture for the correlation between optical sensor signals and the simulated binding affinities. Specifically, a good correlation was found between the sensor signals and the physical binding affinities of the three hormones (estradiol, melatonin, and thyroxine), the neurotransmitter (dopamine), and the vitamin (riboflavin) to the SWCNT-polymer scaffold. The four amino acids (aspartate, glycine, histidine, and tryptophan) and the two monosaccharides (fructose and glucose) considered were identified as blank analytes which are unable to induce sensor signals. The results indicate great success of our physical adsorption-based model in explaining the ranking in sensor selectivities. The combined framework presented here can be used to screen and select polymers that can potentially be used for creating synthetic molecular recognition motifs.

Keywords: coarse-grained molecular dynamics simulation, physical adsorption and binding, carbon nanotube, biological sensor.

1. INTRODUCTION

Highly-selective molecular recognition (molecular complementarity) of the target species by the host plays an important role in many biological systems involving interactions between receptor-ligand,[1] antigen-antibody,[2] DNA-protein,[3] sugar-lectin,[4] and RNA-ribosome.[5] This selectivity is a key consideration in the pharmaceutical and biomedical industries for the rational design of drugs.[6] It also plays a key role in the functionality of artificial chemical sensors and molecular devices for biomedical (e.g., cancer cell detection),[7, 8] environmental (e.g., toxic agent monitoring),[9, 10] and national security (e.g., anti-biochemical weapon) applications.[11, 12]

Soft Matter Accepted Manuscript

However, sensors that utilize proteins are batch-dependent, fragile, expensive, and difficult to control and manufacture, which limits their use in various applications.[13-15] This has motivated researchers to design and create synthetic constructs that can specifically recognize molecules of interest, including implementing these molecular recognition motifs for sensing applications. Novel chemical sensors based on nanomaterials, including nanoparticles,[16, 17] nanowires,[18, 19] graphene,[20-22] and SWCNTs[7, 23-25] (usually functionalized with polymers as surface receptors), present great challenges in determining the molecular conformation of the new receptor (the nanomaterial-polymer complex), as well as in designing chemical structures of the polymer in order to achieve high selectivities towards certain guest molecules (analytes).

A new class of molecular-recognition motifs for solution-phase chemical sensing has been identified recently using SWCNT-polymer complexes.[26, 27] This new engineering-designed recognition motif relies on the wrapping morphology and solution-phase properties of the polymers, and therefore, differs from the antibody-antigen type mechanism associated with antibody-conjugated SWCNT sensors.[28-30] Very recently, we successfully demonstrated a series of novel synthetic heteropolymers that when constrained onto a SWCNT by chemical adsorption, form a new corona phase composed of polymer "anchors" that exhibit highly selective recognition for specific molecules.[31] We refer to this sensing mechanism as "Corona Phase Molecular Recognition" (CoPhMoRe), suggesting that the polymer corona phase that is exposed to the solution phase indeed imparts selectivity towards certain analytes when they coat the SWCNT surface. Conceptually, SWCNTs are single graphene sheets rolled into a cylindrical geometry that creates quantum confinement of electrons and excitons in a single dimension.[32] Their unique optical and electronic properties make them ideal substrates for fluorescence[24] or field-effect transistor (FET)[25] sensors for molecular recognition. As fluorescence sensors, they

Soft Matter Accepted Manuscript

show photoluminescence (PL) quenching (bleaching), enhancing, or shifting in the near-infrared (nIR) spectrum upon binding of analytes.[33] As FET sensors, they show orders-of-magnitude change in electronic conductance upon physical adsorption of analytes.[25] The attractiveness of SWCNTs as sensor elements, while no doubt enhanced by their excellent sensitivity, is still limited by their poor selectivity towards target analytes. In general, it is believed that the analyte-selective sensor signal results from a combined effect of the analyte binding affinity to the sensor surface and the ability of each adsorbed analyte to introduce the optical or electronic signal. While the optical or electronic signals of these sensors are directly associated with the shift of Fermi levels,[25] exciton quenching by dye molecules,[34] or solvatochromatic shifting from solution environmental perturbations[35], their selectivity is also imparted by the difference in the binding (or physical adsorption) affinities of various analytes on SWCNTs. Such a correlation is quite intuitive, because the analytes need to first reside sufficiently close to the SWCNT surface in order to interfere with electrons and excitons. This notion was recently tested by combining experiments and all-atomistic (AA) molecular dynamics (MD) simulations for SWCNT-FET sensors, but only for a very small set of analytes.[36]

These SWCNT-polymer complexes, which mimic biological antibodies, can serve as scaffolds for biosensors in aqueous environment to detect binding of analytes, even at the single-molecule level.[37-40] In spite of the experimental advances discussed above, a theoretical understanding of the binding selectivity of analytes is still lacking. Although high-throughput screening of many analytes against one type of SWCNT-polymer sensor is already practical, synthesizing many types of polymers without clear design guidelines is still very tedious and time consuming, thereby making the screening of candidate polymers very challenging. Accordingly, developing a theoretical description that provides mechanistic understanding of the

Soft Matter Accepted Manuscript

molecular recognition is critical. To this end, molecular docking is the most common computational method used for high-throughput screening of binding selectivities. Specifically, in this method, one utilizes scoring functions to estimate the ligand-receptor binding free energy, which usually relies on experimentally-determined (X-ray crystallography or solution phase NMR) receptor molecular conformations.[41] On the other hand, all-atomistic (AA) molecular dynamics (MD) simulations are capable of: (*i*) predicting receptor molecular conformations, (*ii*) allowing flexibility in receptor and ligand dynamics, and (*iii*) tracing the trajectory of the entire binding process, but they are only affordable on supercomputers.[42, 43] To make it computationally more efficient, coarse-grained (CG) MD simulations[44] could be utilized. Indeed, CGMD simulations are advantageous to handle large systems (~1,000,000 atoms) and long time scales (~10 μs) with reasonable accuracy and predictive power, when compared to AAMD simulations which are only capable of modeling small adsorbed molecules on SWCNTs.[45-47]

With the above in mind, here, we combine large-scale CGMD simulations and physical adsorption/binding theories to quantify the binding affinities of 11 biologically important analytes on a SWCNT decorated with polyethylene glycol (PEG)-based polymers. The chemical structures of the 11 biologically important analytes, including their corresponding ionic/protonation states under physiological pH (7.4), are shown in **Fig. 1a**. These structures consist of 4 amino acids (aspartate (Asp), glycine (Gly), histidine (His), and tryptophan (Trp)), 2 monosaccharides/sugars (fructose and glucose), 3 hormones (estradiol, melatonin, and thyroxine), a neurotransmitter (dopamine), and vitamin B_2 (riboflavin). These 11 analytes were selected because they exhibit: (*i*) diverse sensing behaviors, and (*ii*) unique selectivity and ranking in our photoluminescence experiments using the high-throughput screening assay. The polymers studied possess a hydrophilic PEG backbone (made of 100 monomers) extended into the aqueous

Soft Matter Accepted Manuscript

solution in order to stabilize the SWCNT. These polymers are also functionalized with two fluorescein isothiocyanate (FITC) groups which act as hydrophobic "anchors" to non-covalently stack onto the SWCNT surface (see **Fig. 1b**). Note that the hydrophilic PEG backbones are not expected to interfere with binding,[31] while the sensing or binding selectivity is imparted by the corona phase composed of the anchors. The predicted ranking of binding affinities was compared with experimental PL quenching results to facilitate understanding of the physical-adsorption mechanism behind the selectivity of SWCNT-polymer sensors.

2. COMPUTATIONAL METHODS

2.1. Simulation of SWCNT-Polymer Scaffold. The CGMD simulations were carried out using the GROMACS 4.5 software package.[48] The MARTINI force field developed by Marrink et al. was used to model the SWCNT-polymer scaffold in water.[49] Four water molecules were modeled using one P4 particle in the MARTINI force field. The PEG chain in the polymer was modeled according to Lee et al.[50, 51] The SWCNT was modeled using the 3:1 mapping scheme by Wallace and Sansom,[52] and described by the SC3 particle in the MARTINI force field. An (8,6) SWCNT was used as a model SWCNT because it is semiconducting and contributes significantly to the fluorescence intensity among SWCNTs of various chiralities.[26] The SWCNT was fixed in the middle of the simulation box with 3D periodic boundary conditions. The length of the simulation box along the z-axis, L_z, was kept constant in order to mimic an infinitely long nanotube along this axis. The FITC groups in the PEG-based polymer were coarse-grained from their energy minimized all-atomistic (AA) structures predicted from the PRODRG2 program,[53] following the philosophy of the MARTINI force field (see **Sections 2.2. and 2.3.**). The internal

Soft Matter Accepted Manuscript

coordinates of the FITC groups, as well as their connections to the PEG chain, were modeled using an elastic network approach with a spring constant of 17,000 kJ/mol/nm^2.[54]

All simulations involving SWCNTs were conducted under the isothermal-semiisobaric ensemble ($NP_{xy}T$, fixed numbers of molecules, constant pressure of 1 bar along the x and y-axes, and constant temperature of 300 K). All the non-bonded interactions (Lennard-Jones and Coulombic) were treated using the default values in the original MARTINI paper.[49] The equation of motion was integrated at a time step of 0.01 ps for the SWCNT-polymer scaffold. At the beginning, to facilitate polymer adsorption onto the SWCNT surface, which occurs through a time-consuming diffusion process, a non-equilibrium MD simulation was carried out by pulling the FITC groups towards the nanotube surface. This approach is appropriate because only the hydrophobic FITC groups are supposed to stack onto the nanotube surface. Subsequently, the obtained molecular structure was equilibrated without any constraints for 500 ns to allow rearrangements of the PEG backbone and the FITC groups on the nanotube surface. The last 200 ns of the simulation was collected for data analysis. Note that due to the larger diffusivities of the CG particles relative to those of their AA counterparts, the real time spent in a simulation is approximately 4 times that of the simulated time.[49] Therefore, the system was effectively simulated for around 2 μs. Note that the simulated system consists of around 70,000 CG particles, which is equivalent to around 280,000 atomic groups or more than 1 million atoms. The equilibration of the SWCNT-polymer simulation was confirmed by monitoring the time-dependent 2D radius of gyration (the square roots of the squared summations of the x and y components) of the PEG backbone perpendicular to the nanotube central axis (**Fig. S1 in the Supporting Information**). Note that although the simulations were carried out in 3D, our primary interest is understanding the radial (2D) behavior, perpendicular to the cylindrical (z)

Soft Matter Accepted Manuscript

axis, in order to monitor how far the polymer scaffolds stretch out radially. Specifically, we carried out the simulation for 500 ns to make sure that the fluctuation in the 2D radius of gyration profile converges to a tolerance value (~ 5% of the mean value).

2.2. Simulation of Analyte Binding Affinity. The binding affinities of 11 analytes on the SWCNT-polymer scaffold were predicted by simulating the adsorption process of each analyte on the scaffold. The 2 monosaccharides and the 4 amino acids were modeled using the extended MARTINI force field for carbohydrates[55] and proteins,[56] respectively. The remaining 5 analyte molecules were coarse-grained from their energy minimized AA structures from the PRODRG2 program,[53] following the philosophy of the MARTINI force field. Note that due to the coarse-grained nature of the MARTINI force field, some detailed analyte properties, such as the dipole moment of the analyte molecule and the charge transfer process on the SWCNT-polymer surface, are not included in the simulation. In fact, the charge transfer process can only be described using quantum mechanical methods, which are limited to very small systems and short time scales.[57] The dipole moment of the analyte can only be described in all-atomistic simulations, which are also limited to small systems and nanosecond time scales.[36] On the other hand, the net charge of the analyte molecule is well described by the current CG model. For example, dopamine and Asp molecules are modeled to carry net charges of +1 and –1 under physiological pH, respectively. The net charge, in fact, is taken into account in the computation of Henry's constants or hydrophobicity of the analyte (e.g., higher net charge leads to larger analyte hydrophilicity and a lower Henry's constant). Therefore, the overall goal of the current paper is to utilize CG models to understand sensor selectivity from the perspective of classical polymer

Soft Matter Accepted Manuscript

morphology and physical adsorption, rather than charge transfer, dipole moments, and other quantum mechanical aspects.

To create completely new CG force fields for the 5 additional analyte molecules, internal coordinates of each of them were modeled using the same elastic network approach used for the PEG-based polymers.[54] The force-field parameters of FITC and the 5 additional analytes were calibrated (or trained) against their thermodynamic properties, specifically, against the octanol/water distribution (or partition) coefficients,[58] log D_{OW}, at the physiological pH = 7.4. Force-field training is conducted by iteratively modifying force-field parameters such that the simulated log D_{OW} values (see **Section 2.3**) match the experimental log D_{OW} values, which is typically carried out to obtain parameters of the MARTINI force field.[49] The resulting force-field parameters for FITC and the 5 additional analytes are listed in **Table S1**. Note that the counterions associated with aspartate (Na^+) and dopamine (Cl^-) were modeled using the positively-charged Qd and the negatively-charged Qa particles in the MARTINI force field, respectively.[49] Since experimental values of log D_{OW} of FITC and the 5 additional analytes are not available, we used the empirical quantitative structure-activity relationship (QSAR) method to theoretically predict their log D_{OW} values as our benchmark. The QSAR method, based on the extended group-contribution approach,[59, 60] is freely available online at www.chemicalize.org.

In addition, we noticed that both dopamine and riboflavin have multiple closely-positioned hydroxyl groups (catechol or dihydroxyl in dopamine, and tetrahydroxyl in riboflavin), which are capable of forming strong hydrogen bonds between the two hydroxyl groups in the FITC groups (**Table S1**). Such multi-hydroxyl group feature is not observed in the other 9 analytes studied here, and are also not accounted for explicitly in the original MARTINI force field. Therefore, additional interaction energies need to be added to the existing MARTINI force-

Soft Matter Accepted Manuscript

field parameters. Specifically, this includes the Lennard-Jones (LJ) parameters for both hydroxyl-containing CG beads in dopamine (one SP2 bead) and riboflavin (two P4 beads) interacting with the hydroxyl-containing CG beads in FITC groups (two SNda beads). Note that all the other 9 analytes possess hydrogen-bonding groups (**Fig. 1a**) which are not as strong as the multi-hydroxyl groups in dopamine and riboflavin. Therefore, we compensated for the extra hydrogen-bonding energies between the multi-hydroxyl groups (in dopamine and riboflavin) and the hydroxyl groups in FITC by adding 5 kJ/mol to their interactive LJ potential energies, a value consistent with the difference between monohydroxyl and dihydroxyl hydrogen-bonding energies with a water molecule.[61] Although a more precise hydrogen-bonding model beyond the non-bonded LJ potential used here could be developed at the coarse-grained level, such a model would be less valuable considering the already approximate nature of the parameters in the MARTINI force field.

The adsorption process of the various analyte molecules on the SWCNT-polymer scaffold was initiated by randomly positioning 10 analyte molecules of each type around the scaffold. Note that the use of 10 analyte molecules corresponds to a bulk analyte concentration of ~2 mM, which allows statistical accuracy in data sampling while satisfying the assumption of infinite dilution required for the subsequent calculation of Henry's adsorption constants. The initial molecular conformation of the scaffold was taken from the last trajectory of the previous SWCNT-polymer scaffold simulation at 500 ns. The 10 analyte molecules were then allowed to diffuse freely for a very long simulation time of 4 μs in order to reach equilibration, as determined when the radial distribution function (RDF) of each species around the SWCNT was found to remain unchanged. The last 2 μs of the simulation was used for data analysis. We confirmed that the sampling time is sufficient to render long-term RDF data by monitoring the

Soft Matter Accepted Manuscript

dynamics of "on" (adsorption) and "off" (desorption) events during the simulation (**Fig. S2**), including verifying that a large number of such events were captured during the sampling period. Due to the larger diffusivities of the CG particles noted earlier,[49] each system was effectively simulated for around 16 μs. The conformation of the polymer on the SWCNT is not fixed during the MD simulation, but evolves while the binding of analytes occurs. Therefore, contrary to ns AA simulations in which the polymer conformation is locked into a certain local minimum, the μs CG simulations carried out here allow sufficient sampling of many polymer conformations.

2.3. Simulation of Distribution Coefficients. In order to compute log D_{OW} directly, the free energy of solvation of FITC and each of the 5 additional analyte molecules considered was calculated in both the octanol and water phases. Their corresponding primary ionization states at pH = 7.4 were predicted from www.chemicalize.org. These free-energy simulations were carried out under the isothermal-isobaric ensemble (*NPT*, fixed numbers of molecules, constant pressure of 1 bar, and constant temperature of 300 K).[55] The D_{OW} values of the 2 monosaccharides and the 4 amino acids were obtained in Refs. 55 and 56, respectively. Given the free energies of solvation, computation of log D_{OW} is straightforward. The difference between the analyte solvation free energy in the octanol phase, ΔG_O, and the water phase, ΔG_W, is the distribution free energy, $\Delta\Delta G_{OW}$, of the solute between a water-saturated octanol solution (consisting of 25 CG water molecules and 450 CG octanol molecules[62]) and water (consisting of 1,000 CG water molecules because the solubility of octanol in water is minimal). Specifically,[55]

$$\Delta\Delta G_{OW} = \Delta G_W - \Delta G_O = -\ln 10\ kT \log D_{OW} \quad (1)$$

Soft Matter Accepted Manuscript

Note that ΔG_O and ΔG_W were calculated as the free-energy difference, ΔG_{BA}, of the solute in vacuum (state A) and in the condensed phase (octanol or water) (state B) using the thermodynamic-integration (TI) approach:[63]

$$\Delta G_{BA} = G_B - G_A = \int_{\lambda_A}^{\lambda_B} \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda \tag{2}$$

$$\Delta G_W = \Delta G_{BA}(B = W) \quad and \quad \Delta G_O = \Delta G_{BA}(B = O) \tag{3}$$

where $U(\lambda)$ denotes the potential energy function describing the total solute-solvent interaction, the average $\langle ... \rangle$ is taken over the MD trajectory, and λ is a coupling parameter that regulates the strength of U and varies linearly from zero interaction (λ_A = 0, in vacuum) to full interaction (λ_B = 1, in the condensed phase). All the bonded interactions were interpolated linearly. To remove the singularities in the potentials for the non-bonded interactions, a soft-core approach was used.[64] Calculations were performed at 19 intermediate λ values with a constant spacing of $\Delta\lambda$ = 0.05 until a smooth curve for the free-energy derivative $\langle \partial U / \partial \lambda \rangle$ was obtained, which was then integrated numerically using a trapezoidal scheme (**Fig. S3a**). For each individual λ point, 100 ns of CGMD simulation using a stochastic-dynamics (SD) approach was carried out, and the last 50 ns of simulation was used for data analysis. The CG simulated log D_{OW} values agree very well (within an error of ~0.4) with the QSAR predicted values (**Fig. S3b**). Note that a positive log D_{OW} value indicates preferential partitioning into the octanol phase (hydrophobic), while a negative value indicates preferential partitioning into the water phase (hydrophilic). Therefore, the higher the log D_{OW} value, the more hydrophobic the analyte molecule is. Note also that an error of ~0.4 in log D_{OW} represents an error of ~2 kT in the calculations of $\Delta\Delta G_{OW}$. Most importantly, the CGMD simulated log D_{OW} values match the trend in the QSAR predicted values.

Soft Matter Accepted Manuscript

3. EXPERIMENTAL METHOD

SWCNTs were purchased from *Nano-C*. Amine difunctionalized poly(ethylene glycol) (NH_2-PEG-NH_2) was purchased from *Creative PEGworks*. Sodium cholate (SC), reaction reagents (fluorescein isothiocyanate (FITC), dichloromethane (CH_2Cl_2), dimethylformamide (DMF), N,N-diisopropylethylamine (DIEA), and dimethyl sulfoxide (DMSO)), and the 11 analytes considered (Asp, Gly, His, Trp, fructose, glucose, estradiol, melatonin, thyroxine, dopamine, and riboflavin) were purchased from *Sigma Aldrich* and used as received, unless indicated otherwise. The synthesis scheme of FITC-PEG-FITC is shown in **Scheme S1**. Briefly, NH_2-PEG-NH_2 (5kDa, 0.1 mol/l) and FITC (0.22 mol/l) were dissolved in a mixture of CH_2Cl_2 and DMF (1:1), followed by the addition of 0.2 mol/l of DIEA to the reaction mixture. After three hours, the reaction product was flocculated by ether (10x in volume) and re-dissolved in DMF, and this cycle was repeated twice. Finally, the product was flocculated by ether again, and collected using a vacuum filtration system.

SWCNTs were first suspended in a 2 wt % SC aqueous solution using previously published methods.[24, 33] Briefly, 1 mg/mL SWCNTs were added to 40 mL 2 wt% SC in *NanoPure* H_2O, and were sonicated with a 6 mm probe tip at 40% amplitude (~12W) for 1 hr in an ice bath. The resulting dark black solution was ultracentrifuged in an SW32 Ti rotor (Beckman Coulter) at 153,700 RCF (max) for 4 hrs to remove unsuspended SWCNT aggregates and catalyst particles. FITC-PEG-FITC was then dissolved at 2 wt % in the SC-SWCNT suspension, and the mixture was placed in a 3.5 kDa MWCO dialysis bag and dialyzed against 2L 1X PBS buffer (50 mM, pH 7.4) for 24 hours to remove free SC and allow the polymer to self-assemble on the nanotube surface.

Soft Matter Accepted Manuscript

Polymer-wrapped SWCNT solutions were diluted to a final SWCNT concentration of 2 mg/l. The three analytes, Asp, Gly, and His, were initially dissolved in DMSO. The other 8 analytes were dissolved in 50 mM PBS (pH 7.4). The resulting analyte solutions were added to the SWCNTs at the various bulk concentrations listed in **Table S2**, such that the final DMSO concentration was 1 vol%. The mixture was then incubated for 1 hr, and the resulting SWCNT PL was measured and repeated 3 times with a home-built near infrared (nIR) fluorescence microscope. Briefly, a Zeiss AxioVision inverted microscope was coupled to a Princeton Instruments InGaAs OMA V array detector through a PI Acton SP2500 spectrometer. The software package Winspec32 was used to collect the spectra. The sample excitation was from a 785 nm photodiode laser, 450 mW at the source and 150 mW at the sample. A custom-written Visual Basic program communicates between the microscope stage and Winspec32, automating the process of taking the SWCNT spectra of the SWNT in a 96 well-plate. For more detailed discussions on the experimental method, please refer to our recently published work.[31]

4. RESULTS AND DISCUSSION

4.1. Morphology of the SWCNT-Polymer Scaffold. The molecular conformations of the polymer on the SWCNT are shown in **Fig. 2a**. Two polymer surface coverages were investigated: 1.44 (low) and 2.41 (high) polymers per nm of the SWCNT. Note that the two surface coverages selected are close to those corresponding to the surfactant cholate ion adsorption case (≤ 3.60 molecules/nm), as reported in recent MD simulations.[45] In addition, the molecular weights of the FITC group (which is the only portion of the polymer that adsorbs onto the SWCNT surface) and the cholate ion are both ~ 400 g/mol, which make them similar in terms of their excluded volumes. In general, the flexible polymeric PEG backbone (red beads in **Fig. 2a**) forms a

Soft Matter Accepted Manuscript

polymer coating or "cloud" around the SWCNT. Most of the hydrophobic FITC groups (blue beads in **Fig. 2a**) bind closely to the nanotube surface to form tangled and twisted PEG "loops". A closer examination of the distributions of FITC groups, polymeric PEG backbones, and water molecules around the SWCNT, through their CG particle number densities, $\rho(r)$, in units of number of molecules per nm^3, are shown in **Fig. 2b**. As shown in **Fig. 2b** (top), the sharp, primary peaks at r = 0.9 nm in the $\rho(r)$ profile of FITC groups corroborate their close binding to the nanotube surface to form the new corona phase which is responsible for the sensing selectivity, while the nonzero $\rho(r)$ regions beyond r = 1.3 nm indicates their minor detachment from the nanotube surface. As shown in **Fig. 2b** (middle), most of the PEG backbones do not bind to the nanotube surface (tiny peak at r = 0.9 nm), but instead extend far into the water phase (to around r = 6 nm) due to their hydrophilic nature. As shown in **Fig. 2b** (middle and bottom), the multiple peaks in the $\rho(r)$ profile of PEG correlate with the ordering of the water molecules around the hydrophobic nanotube surface, which is well known in the literature,[65] and is similar to previous simulation reports on water ordering around a graphene surface.[66] As expected, the $\rho(r)$ profiles of the FITC groups and the PEG backbones at low surface coverages are lower than those at high surface coverages.

Based on the simulated distribution profiles, the simulation cell was divided into three distinct phases, as shown in **Figs. 2b and 2c** using three different colors: I – FITC phase ($0 \le r \le R_1$), which is equivalent to the SWCNT surface; II – polymeric PEG phase ($R_1 \le r \le R_2$), and III – bulk analyte solution phase ($R_2 \le r \le R_3$). Note that: (*i*) R_1 = 1.3 nm was determined based on the position of the local minima located to the right of the primary peaks in the $\rho(r)$ profiles of the FITC groups (see **Fig. 2b** (top)), (*ii*) R_2 = 6.0 nm was determined based on the fact that the $\rho(r)$ profiles of the PEG backbones approach zero beyond r = 6.0 nm (see **Fig. 2b** (middle)), and

Soft Matter Accepted Manuscript

(*iii*) R_3 = 9.7 nm simply reflects the size of the simulation cell used, which is sufficiently large to prevent interactions between the SWCNT-polymer scaffolds across the periodic image.

The number of analyte molecules in phase *i*, N_i (*i* = 1, 2, and 3), was determined using their cumulative radial distribution functions (RDFs), $N(r)$, where r is the distance to the nanotube central axis. Note that $N(r)$, which is equivalent to the number of analyte molecules located within the cylindrical space of radius r and length L_z = 20.8 nm (corresponding to the length of the SWCNT along the *z*-axis of the simulation cell) was computed from the normalized RDFs, $g(r)$, through integrations over the cylindrical space. Specifically,

$$N(r) = \int_0^r g(r) 2\pi r L_z dr \tag{4}$$

$$N_1 = N(R_1) \tag{5}$$

$$N_2 = N(R_2) - N(R_1) \tag{6}$$

$$N_3 = N(R_3) - N(R_2) \tag{7}$$

where R_i (*i* = 1, 2, and 3) is the boundary of phase *i*. Note that the $N(r)$ profile was determined for each analyte type on the SWCNT-polymer scaffold under the bare SWCNT solution case (no polymer), in addition to the low-polymer and the high-polymer coverage cases.

4.2. Binding Affinities of the Analytes. The binding affinities of the 11 analytes considered onto the SWCNT-polymer scaffold can be quantified in terms of the Henry's constants. This corresponds to their physical adsorption at infinite dilution, reflecting the fact that their concentrations are generally very low in our experiments (see **Table S2**). Two Henry's constants

Soft Matter Accepted Manuscript

K_1 and K_2 (both independent of the analyte concentration), in units of nm^2, can be defined in phases I and II, respectively, with respect to bulk phase III. Specifically (see **Fig. 2c**),

$$K_1 = \frac{\theta_1}{C_0} = \frac{N_1 / L_z}{N_3 / \pi\left(R_3^2 - R_2^2\right)L_z} = \pi\left(R_3^2 - R_2^2\right)\frac{N_1}{N_3} \tag{8}$$

$$K_2 = \frac{\theta_2}{C_0} = \frac{N_2 / L_z}{N_3 / \pi\left(R_3^2 - R_2^2\right)L_z} = \pi\left(R_3^2 - R_2^2\right)\frac{N_2}{N_3} \tag{9}$$

where θ_1 and θ_2 are the numbers of adsorbed analyte molecules per nm of the SWCNT-polymer scaffold (or linear adsorption surface coverages) in phases I and II, respectively, N_i (i = 1, 2, and 3) is the number of analyte molecules in phase i determined using their cumulative RDFs from the MD simulations, C_0 is the analyte bulk concentration determined in phase III, and L_z = 20.8 nm is the length of the SWCNT along the z-axis of the simulation cell. Note that the two Henry's constants are, as defined, intrinsic properties of the analytes, and are therefore independent of the analyte bulk concentrations C_0. This follows because the concentration effects on K_1 and K_2 are cancelled out when θ_1 and θ_2 are divided (or normalized) by the bulk concentration C_0, or from a simulations perspective, because N_1 and N_2 are normalized by N_3 to yield concentration-independent K_1 and K_2.

Fig. 3a shows the simulated K_1 values for each of the 11 analytes considered on the SWCNT-polymer scaffold for the bare SWCNT solution case (no polymers, red), and for the low-polymer (green) and high-polymer (blue) surface coverage cases. The information in **Fig. 3a** was re-plotted in **Fig. 3b** to illustrate the partial correlations between the K_1 values for the no-polymer case and those for the low-surface coverage and high-surface coverage cases, indicating that the interactions between the SWCNT and the analytes contribute significantly to the simulated K_1 values, in addition to the contribution from the polymer phase (indeed, the FITC

Soft Matter Accepted Manuscript

groups). The only two exceptions are dopamine and riboflavin, in which the low-polymer and high-polymer surface coverage cases have greatly enhanced K_1 values compared to the case without polymers. This significant contribution to K_1 from the polymer phase is due to the very strong hydrogen bonding between the hydroxyl groups in FITC and the multi-hydroxyl groups in dopamine and riboflavin, as discussed in **Section 2.2**. In general, the values of K_1 increase from the no-polymer case to the low-polymer surface coverage case, as a result of additional analytes binding to the FITC groups stacked on the SWCNT surface (see the blue diamonds in **Fig. 3b**). On the other hand, the changes in K_1 between the low-polymer and the high-polymer surface coverage cases do not exhibit a consistent trend, primarily due to the competing effects of the increased binding sites resulting from the FITC groups and the reduced binding sites available on the bare SWCNT surface (see the red squares in **Fig. 3b**).

On the other hand, an examination of **Fig. 3c** shows that the simulated values of K_2 correlate inversely with the values of K_1, except for fructose and glucose which both exhibit high K_1 and K_2 values (possibly due to their high affinity for both the PEG phase and the FITC groups). As expected, this inverse correlation occurs because the SWCNT surface and the PEG phase compete for analyte binding. Since the PEG phase is located away from the SWCNT surface, analytes adsorbed to this phase are not likely to contribute to sensor signals. Accordingly, unlike K_1, K_2 is not an appropriate measure of the binding affinity of an analyte, which is also clear from the strong correlation between K_1 and the experimental results (see below for details).

Since the value of K_1 is analyte-specific and varies with respect to the polymer surface coverage, the averaged K_1 value over the low-polymer and the high-polymer surface coverage cases, K_1^{avg}, would be a good representative value to quantify the binding affinity of each analyte. The K_1^{avg} values and the associated standard errors in averaging the low-polymer and the high-

Soft Matter Accepted Manuscript

polymer surface coverage cases are shown in **Fig. 4a**. Specifically, thyroxine exhibits the highest average binding affinity K_1^{avg}, followed by riboflavin, dopamine, tryptophan, estradiol, fructose, melatonin, and glucose. Three out of the four amino acids, except for tryptophan which is the only hydrophobic amino acid considered here, have almost no binding affinities to the SWCNT surface (all with $K_1 < 4$ nm^2).

The fact that the values of K_1 for the 11 analytes, particularly in the no polymer case, partially correlate with the ranking of their octanol/water distribution (partition) coefficients, D_{OW} (see **Fig. 4b**) is consistent with the picture that the binding of analytes to the SWCNT-polymer scaffold is analogous to the partition/distribution of the analytes between a hydrophobic phase and a hydrophilic phase. Recall that D_{OW} is a quantitative measure of the intrinsic hydrophobicity of an analyte, and was used to obtain the CG force-field parameters of the analyte. Therefore, the hydrophobicity/hydrophilicity of the analyte molecules, as reflected in their distribution coefficients, D_{OW}, is a key factor in determining their binding affinities. The specific interactions between the FITC groups and the analytes, combined with the hydrophobicity of the analytes for binding onto the SWCNT surface, control the overall binding affinities of the analytes on the SWCNT-polymer scaffold. Although there are no simple theories to quantify the specific interactions between the FITC groups and the analytes considered here (or many other types of analytes not considered), the molecular simulation method presented here can serve as a screening tool for case-by-case studies.

4.3. Specific Binding due to the Polymer Anchors. As shown experimentally, the hydrophilic PEG phase of the polymer coating is merely responsible for the dispersion stability of the SWCNTs.[31] Indeed, the PEG phase is distributed too far away from the SWCNT surface

Soft Matter Accepted Manuscript

so that it cannot promote any specific binding effects to attract analytes that lead to PL quenching. Therefore, specific binding towards certain analytes can be promoted only by the polymer anchors (the FITC groups here). The RDF profiles of highly-bound analyte molecules, including thyroxine, riboflavin, dopamine, estradiol, melatonin, tryptophan, and fructose, around the CG particles in the SWCNT (under no-polymer, and low-polymer and high-polymer surface coverage cases) and FITC groups (under low-polymer and high-polymer surface coverage cases) are shown in **Fig. 5** (with corresponding simulation snapshots shown on the right column) to help clarify the specific binding effects. In general, the peak heights in the RDF profiles around the SWCNT correlate directly with the simulated analyte binding affinities K_1 in **Fig. 3a**.

More interestingly, the peak heights in the RDF profiles around the FITC groups show various analyte-specific features. For example, thyroxine, tryptophan, fructose, dopamine, and riboflavin (**Fig. 5a, d-g**) do exhibit higher binding affinities towards the FITC groups (higher RDF peak heights) compared to estradiol and melatonin (**Fig. 5b-c**). For dopamine and riboflavin, it is simply due to the highly strong hydrogen bonding energies between the hydroxyl groups in FITC and the multi-hydroxyl groups in dopamine and riboflavin. For thyroxine and tryptophan, this may be due to their amphiphilic chemical structures with hydrophobic moieties that are modeled by three SC4 beads adjacent to only one hydrophilic moiety, that is, modeled by the Qda or SP1 bead, as shown in **Table S1** for thyroxine and in Ref. 56 for tryptophan. This structure can match the similar amphiphilic chemical structure of FITC (with three SC4 beads adjacent to one Nd bead, as shown in Table S1) and bind with each other closely like two aggregated surfactants in a small micelle or dimer. In the case of fructose (modeled by three hydrophilic beads, P1-P3-P4, in Ref. 55), the higher binding affinity towards the FITC groups may result from the strong binding affinity to the PEG phase (adjacent to the FITC groups),

Soft Matter Accepted Manuscript

which is reflected by the shifted peaks in the RDF profile of fructose away from the SWCNT (or towards the PEG phase). Similar scenarios also apply in the case of glucose. However, since both fructose and glucose act as reference/blank analytes (see **Section 4.4** below) and are not able to quench the fluorescent response of SWCNTs, such specific binding effect will not affect the fluorescent signals of the sensor.

The weaker binding affinities of estradiol and melatonin towards the FITC groups also helps explain why the presence of the polymer coating on the SWCNT, even at low-polymer surface coverage, is not able to enhance binding affinities of the no-polymer cases (**Fig. 3b**). On the other hand, tryptophan, fructose, and dopamine (**Fig. 5d-e**) exhibit enhanced binding affinities towards the FITC groups as the polymer surface coverage increases, due to the increased total number of FITC groups present. However, thyroxine, estradiol, and riboflavin (**Fig. 5a-b**) show, counter-intuitively, decreased binding affinities towards the FITC groups as the polymer surface coverage increases, suggesting that an increase in the number of FITC groups reduces the available binding sites for analytes with very large molecular sizes, as reflected by the number of CG beads used to model them. Specifically, thyroxine, estradiol, and riboflavin are modeled by 8, 6, and 8 beads, respectively, (**Table S1**), while tryptophan, fructose, and dopamine are modeled by only 5, 3, and 3 beads,[55, 56] respectively.

Competitive adsorption between the polymer FITC groups and the analytes are taken into account automatically in the simulations by allowing dynamic adsorption and desorption of the FITC groups or the analytes from the SWCNT surface. The FITC groups are more hydrophobic (or attractive to the SWCNT surface) than most of the analytes studied here, except for thyroxine and estradiol (as ranked by the partition coefficients (log D_{ow}) reported in **Fig. S3b,** and the Henry's constants K_1 reported in **Fig. 3a** for the no-polymer case). Therefore, most analytes are

Soft Matter Accepted Manuscript

not able to compete with the polymers for adsorption. Even in the case of thyroxine and estradiol, due to the low analyte concentration studied here, they can replace only a very small number of the polymer FITC groups on the SWCNT surface (a total of 10 analytes in the system vs. 60 ~ 100 FITC groups for low-polymer ~ high-polymer surface coverages, respectively), and hence, will not affect the overall findings reported here.

4.4. Comparison with PL Quenching Experiments. As we indicated earlier, the experimentally-observed sensor signal (the degree of optical quenching in the PL spectra) upon the addition of analytes corresponds to a combined effect of the analyte binding affinity, K_1, as well as of the ability of each adsorbed analyte molecule to quench the PL spectra, which we denote here by the unknown parameter, Q. Specifically, $Q = 0$ corresponds to the complete inability of an analyte to quench the PL spectra and generate a sensor signal, while $Q > 0$ corresponds to the finite ability of an analyte to quench the PL spectra and generate a sensor signal. Based on the intuitive correlation between the experimental SWCNT sensor signals (the directly-observed degree of PL quenching upon analyte binding), F, and the analyte adsorption surface coverage on the SWCNT surface, θ_1, one can propose a simple equation that relates F to θ_1, including accounting for the factor of Q introduced above. Because θ_1 and Q are independent of each other, a separation-of-variables type product rule can be used to correlate F with θ_1 and Q. Specifically,

$$F = \theta_1 \cdot Q = \frac{I_0 - I_F}{I_0} \tag{10}$$

where I_0 and I_F are the initial (or reference) and the final fluorescence intensities measured in the PL experiments, respectively. Using **Eq. 8** for Henry's adsorption isotherm in **Eq. 10** yields:

Soft Matter Accepted Manuscript

$$F = \theta_1 \cdot Q = K_1 C_0 \cdot Q \rightarrow \frac{F}{C_0} \equiv f = K_1 \cdot Q \quad (11)$$

where C_0 is the experimental bulk concentration of the corresponding analyte as listed in **Table S2**, and f is the analyte concentration-normalized degree of quenching for each analyte, in units of mM^{-1}. Note that, for sensor applications, f is a better quantitative measure of the intrinsic selectivity of each analyte than F, because it does not depend on C_0. Note also that C_0 used in **Eq. 8** is computed from the CGMD simulations, and is conceptually the same as C_0 used in **Eq. 11**.

In **Fig. 4a**, we compare f with the simulated values of K_1^{avg}, averaged over the low-polymer and the high-polymer surface coverage cases. Two distinct categories of analytes can be identified in the plot: (i) along the diagonal green dashed line, the K_1^{avg} values of the three hormones (thyroxine, estradiol, and melatonin), dopamine, and riboflavin correlate well with the normalized degree of quenching, f, with a constant $Q > 0$, and (ii) along the red dashed line, the normalized degrees of quenching are zero, independent of the K_1^{avg} values, for the four amino acids (Asp, Gly, His, and Trp) and the two monosaccharides (fructose and glucose), which implies that $Q = 0$.

Category (i) demonstrates the success of the current physical adsorption model to explain the experimental sensor signals for analytes such as the three hormones (thyroxine, estradiol, and melatonin), dopamine, and riboflavin. The hormones studied also exhibit the same normalized degree of quenching under the same analyte adsorption surface coverage (linear number density) on the SWCNT-polymer scaffold among different hormone types based on the good linear correlation shown in **Fig. 4a**. Linear regression to the hormone data in this category gives rise to $Q \sim 0.04\ mM^{-1} \cdot nm^{-2}$, based on the definition of $Q = f/K_1$ in **Eq. 11**. Note that 1 mM of analyte is equivalent to $\sim 6.02 \times 10^{-4}$ of analyte molecules per nm^3. Therefore, based on the more

Soft Matter Accepted Manuscript

comprehensive definition of $Q = F/\theta_1$ (i.e., the degree of PL quenching per analyte surface coverage on the SWCNT-polymer scaffold) in **Eq. 11**, we converted it to $Q \sim 0.04 / (6.02 \times 10^{-4})$ nm ~ 70 nm/molecule, or $Q \sim 0.7$ (molecule/100 nm)$^{-1}$ (i.e., 0.7 degree of PL quenching for each hormone molecule adsorbed on every 100 nm of the SWCNT-polymer scaffold). Note that the experimentally reported exciton diffusion length is ~100 nm for bare SWCNTs,[39, 40] which gives a better physical picture that ~70% of the excitons will be absorbed by the analytes before they radiatively decay and emit signals.

Category (ii) indicates that amino acids and monosaccharides cannot by recognized by the SWCNT sensor coated with the PEG-based polymers studied here, although some of them possess high binding affinities to the SWCNT surface (e.g., tryptophan and fructose). Therefore, these analytes can be categorized as reference/blank analytes without sensitivities ($Q = 0$), unless different surface coatings are introduced to selectively bind to these analytes and help interfere with surface excitons in the SWCNT, such as those demonstrated in a glucose sensor using boronic acid[67] or glucose oxidase enzyme.[33]

In addition to effects resulting from physical binding, it is possible that analytes, such as dopamine and riboflavin, can introduce different exciton diffusion pathways where excitons can travel far from the SWCNT surface and move along the polymer chains to reach the analytes. This could possibly explain their slight deviations from the linear diagonal region in **Fig. 4a**, as compared to the three hormones which exhibit good linear correlations in K_1 vs. f. A recent exciton diffusion model based on experimental data suggests that the characteristic diffusion length (or mean-free path) of excitons along a (7,5) SWCNT is 3.9 nm,[68] which is larger than the R_1 = 1.3 nm value considered here for the FITC phase. In addition, note that dopamine and riboflavin are the only two self-fluorescent analytes among the 11 analytes considered here,[69, 70]

Soft Matter Accepted Manuscript

which could complicate the sensor signals, and make them special and unpredictable using just the physical adsorption mechanism. Understanding the exact exciton diffusion and decaying mechanism is still an active area of investigation, which is beyond the scope of the current work. Nevertheless, the combined CGMD simulation and adsorption theory framework presented here allows qualitative predictions of the distribution coefficient, D_{OW}, and the resulting binding affinity, K_1, of the analyte on the SWCNT-polymer scaffold. Even without quantifying the ability of each adsorbed analyte molecule to quench the PL spectra, Q, using quantum mechanics (QM) calculations, our framework can be used to screen out polymer candidates when designing sensors by identifying those analytes that cannot bind to the SWCNT-polymer scaffold. This could lead to the first step in fast sensor-selectivity screening across a large library of analytes of diverse chemical and electronic structures by simply comparing their distribution coefficients, D_{OW}.

5. CONCLUSIONS

We combined large-scale CGMD simulations, physical adsorption/binding theories, and PL experiments to understand the physical-adsorption mechanism behind the selectivity of SWCNT-polymer sensors towards a large set of biologically important analytes. The analytes considered here include: three hormones (estradiol, melatonin, and thyroxine), four amino acids (Asp, Gly, His, and Trp), two monosaccharides (fructose and glucose), a neurotransmitter (dopamine), and a vitamin (riboflavin). The identification of two distinct categories of analytes demonstrates a very rich and complex picture for the correlation between the experimental sensor signals, f, and the simulated physical binding/adsorption affinities of the analytes on the SWCNT surface, K_1. Specific binding affinities between the analytes and the FITC groups were

Soft Matter Accepted Manuscript

shown to be extremely important for the resulting K_1 values under different polymer surface coverages, being very complex given the diversity in analyte chemical structures. While the FITC groups can enhance the binding affinity of some analytes due to specific interactions, they can also reduce the available binding sites of some analytes with large molecular sizes (e.g., estradiol, thyroxine, and riboflavin) when their surface coverage is increased.

The specific interactions involve: (i) the strong binding of amphiphilic chemical structures within FITC towards similar amphiphilic structures within thyroxine, tryptophan, and fructose, in analogy to surfactant aggregations that form a micelle or dimer, and (ii) the highly strong hydrogen-bonding between the hydroxyl groups within FITC and the multi-hydroxyl groups within dopamine and riboflavin. This complex picture reflects the diversity in analyte types and the complexity in exciton diffusion pathways, which can be modeled by introducing an unknown parameter, Q, reflecting the ability of each analyte to quench the PL spectra. The simulation-modeling framework presented here predicts the experimental sensor selectivities well for all the analytes considered. Specifically, we found a good linear correlation between the sensor signals and the binding affinities for the three hormones considered, as well as consistent correlation, although not perfectly linear, for dopamine and riboflavin. In addition, we found that amino acids and monosaccharides are not sensitive to the sensor studied here, probably due to the inability of these analytes to interact with the surface excitons of the SWCNT via the PEG-based polymer coating.

Beyond specific binding to the FITC groups, the key factor determining the binding affinity of the analyte is its degree of hydrophobicity, as reflected in its distribution coefficient, D_{OW}, in the bulk water/octanol system. Specific interactions between the FITC groups and the analytes can only be quantified using molecular simulations on a case-by-case manner. In spite

Soft Matter Accepted Manuscript

of the diversity in analyte chemical and electronic structures, the combined CGMD simulation and adsorption theory framework presented here can serve as a valuable baseline for screening out polymer candidates when designing sensors by identifying which analytes cannot bind to the SWCNT-polymer scaffold. This could serve as the first step in fast sensor-selectivity screening across a large library of analytes by simply comparing their distribution coefficients, D_{OW}. Furthermore, in general, this framework can advance the understanding of physical adsorption at any nanomaterial-biological interfaces to complement experimental observations.[71, 72]

Acknowledgement. S. Lin and D. Blankschtein are grateful for the financial support from the National Science Foundation (NSF) under contract number CBET-1133813. Computational resources were partially supported by the XSEDE program under contract number TG-DMR120003. This work is also supported by grants from NSF, the Arnold and Mable Beckman Foundation, the DuPont-MIT Alliance (DMA), MIT Deshpande Center, Novartis, and a Sloan Fellowship awarded to M. S. Strano. We thank Zachary Ulissi and Vishnu Sresht for helpful discussions.

Soft Matter Accepted Manuscript

References and Notes

1. R. E. Babine and S. L. Bender, *Chemical Reviews*, 1997, **97**, 1359-1472.
2. D. R. Davies and G. H. Cohen, *P Natl Acad Sci USA*, 1996, **93**, 7-12.
3. C. O. Pabo and R. T. Sauer, *Annu Rev Biochem*, 1984, **53**, 293-321.
4. D. Neumann, C. M. Lehr, H. P. Lenhof and O. Kohlbacher, *Adv Drug Deliver Rev*, 2004, **56**, 437-457.
5. P. Mangeol, T. Bizebard, C. Chiaruttini, M. Dreyfus, M. Springer and U. Bockelmann, *P Natl Acad Sci USA*, 2011, **108**, 18272-18276.
6. S. H. Gellman, *Chemical Reviews*, 1997, **97**, 1231-1232.
7. D. A. Heller, H. Jin, B. M. Martinez, D. Patel, B. M. Miller, T. K. Yeung, P. V. Jena, C. Hobartner, T. Ha, S. K. Silverman and M. S. Strano, *Nat Nanotechnol*, 2009, **4**, 114-120.
8. S. Andreescu and O. A. Sadik, *Methods*, 2005, **37**, 84-93.
9. G. Aragay, J. Pons and A. Merkoci, *Chemical Reviews*, 2011, **111**, 3433-3458.
10. D. Diamond, *Principles of chemical and biological sensors*, Wiley, New York, 1998.
11. O. A. Sadik, W. H. Land and J. Wang, *Electroanal*, 2003, **15**, 1149-1159.
12. K. E. Sapsford, C. Bradburne, J. B. Detehanty and I. L. Medintz, *Mater Today*, 2008, **11**, 38-49.
13. D. Saerens, L. Huang, K. Bonroy and S. Muyldermans, *Sensors-Basel*, 2008, **8**, 4669-4686.
14. P. D. Skottrup, M. Nicolaisen and A. F. Justesen, *Biosens Bioelectron*, 2008, **24**, 339-348.
15. P. Peluso, D. S. Wilson, D. Do, H. Tran, M. Venkatasubbaiah, D. Quincy, B. Heidecker, K. Poindexter, N. Tolani, M. Phelan, K. Witte, L. S. Jung, P. Wagner and S. Nock, *Anal Biochem*, 2003, **312**, 113-124.
16. K. Saha, S. S. Agasti, C. Kim, X. N. Li and V. M. Rotello, *Chemical Reviews*, 2012, **112**, 2739-2779.
17. G. Raschke, S. Kowarik, T. Franzl, C. Sonnichsen, T. A. Klar, J. Feldmann, A. Nichtl and K. Kurzinger, *Nano Lett*, 2003, **3**, 935-938.
18. G. F. Zheng, F. Patolsky, Y. Cui, W. U. Wang and C. M. Lieber, *Nat Biotechnol*, 2005, **23**, 1294-1301.
19. Y. Cui, Q. Q. Wei, H. K. Park and C. M. Lieber, *Science*, 2001, **293**, 1289-1292.
20. F. Yavari and N. Koratkar, *The Journal of Physical Chemistry Letters*, 2012, **3**, 1746-1753.
21. Y. Lu, B. R. Goldsmith, N. J. Kybert and A. T. C. Johnson, *Appl Phys Lett*, 2010, **97**.
22. F. Schedin, A. K. Geim, S. V. Morozov, E. W. Hill, P. Blake, M. I. Katsnelson and K. S. Novoselov, *Nat Mater*, 2007, **6**, 652-655.
23. C. Staii, A. T. Johnson, M. Chen and A. Gelperin, *Nano Lett*, 2005, **5**, 1774-1778.
24. M. J. O'Connell, S. M. Bachilo, C. B. Huffman, V. C. Moore, M. S. Strano, E. H. Haroz, K. L. Rialon, P. J. Boul, W. H. Noon, C. Kittrell, J. P. Ma, R. H. Hauge, R. B. Weisman and R. E. Smalley, *Science*, 2002, **297**, 593-596.
25. J. Kong, N. R. Franklin, C. W. Zhou, M. G. Chapline, S. Peng, K. J. Cho and H. J. Dai, *Science*, 2000, **287**, 622-625.
26. J. Q. Zhang, A. A. Boghossian, P. W. Barone, A. Rwei, J. H. Kim, D. H. Lin, D. A. Heller, A. J. Hilmer, N. Nair, N. F. Reuel and M. S. Strano, *J Am Chem Soc*, 2011, **133**, 567-581.

Soft Matter Accepted Manuscript

27. D. A. Heller, G. W. Pratt, J. Q. Zhang, N. Nair, A. J. Hansborough, A. A. Boghossian, N. F. Reuel, P. W. Barone and M. S. Strano, *P Natl Acad Sci USA*, 2011, **108**, 8544-8549.
28. S. Wang, E. S. Humphreys, S.-Y. Chung, D. F. Delduco, S. R. Lustig, H. Wang, K. N. Parker, N. W. Rizzo, S. Subramoney, Y.-M. Chiang and A. Jagota, *Nat Mater*, 2003, **2**, 196-200.
29. R. R. Johnson, B. J. Rego, A. T. C. Johnson and M. L. Klein, *The Journal of Physical Chemistry B*, 2009, **113**, 11589-11593.
30. M. Shim, N. W. Shi Kam, R. J. Chen, Y. Li and H. Dai, *Nano Lett*, 2002, **2**, 285-288.
31. J. Zhang, M. P. Landry, P. W. Barone, J.-H. Kim, S. Lin, Z. W. Ulissi, D. Lin, B. Mu, A. A. Boghossian, A. J. Hilmer, A. Rwei, A. C. Hinckley, S. Kruss, M. A. Shandell, N. Nair, S. Blake, F. Sen, S. Sen, R. G. Croy, D. Li, K. Yum, J.-H. Ahn, H. Jin, D. A. Heller, J. M. Essigmann, D. Blankschtein and M. S. Strano, *Nat Nanotechnol*, 2013, **8**, 959-968.
32. M. Endo, S. Iijima and M. S. Dresselhaus, *Carbon nanotubes*, Pergamon, Oxford ; Tarrytown, N.Y., 1996.
33. P. W. Barone, S. Baik, D. A. Heller and M. S. Strano, *Nat Mater*, 2005, **4**, 86-92.
34. J. K. Holt, H. G. Park, Y. M. Wang, M. Stadermann, A. B. Artyukhin, C. P. Grigoropoulos, A. Noy and O. Bakajin, *Science*, 2006, **312**, 1034-1037.
35. B. A. Larsen, P. Deria, J. M. Holt, I. N. Stanton, M. J. Heben, M. J. Therien and J. L. Blackburn, *J Am Chem Soc*, 2012, **134**, 12485-12491.
36. Z. F. Kuang, S. N. Kim, W. J. Crookes-Goodson, B. L. Farmer and R. R. Naik, *Acs Nano*, 2010, **4**, 452-458.
37. H. Jin, D. A. Heller, J. H. Kim and M. S. Strano, *Nano letters*, 2008, **8**, 4299-4304.
38. J. Zhang, A. A. Boghossian, P. W. Barone, A. Rwei, J.-H. Kim, D. Lin, D. A. Heller, A. J. Hilmer, N. Nair, N. F. Reuel and M. S. Strano, *Journal of the American Chemical Society*, 2010, **133**, 567-581.
39. A. J. Siitonen, D. A. Tsyboulski, S. M. Bachilo and R. B. Weisman, *Nano Letters*, 2010, **10**, 1595-1599.
40. L. Cognet, D. A. Tsyboulski, J. D. R. Rocha, C. D. Doyle, J. M. Tour and R. B. Weisman, *Science*, 2007, **316**, 1465-1468.
41. T. Lengauer and M. Rarey, *Curr Opin Struc Biol*, 1996, **6**, 402-406.
42. Y. Shan, E. T. Kim, M. P. Eastwood, R. O. Dror, M. A. Seeliger and D. E. Shaw, *J Am Chem Soc*, 2011, **133**, 9181-9183.
43. R. O. Dror, A. C. Pan, D. H. Arlow, D. W. Borhani, P. Maragakis, Y. B. Shan, H. F. Xu and D. E. Shaw, *P Natl Acad Sci USA*, 2011, **108**, 13118-13123.
44. S. C. L. Kamerlin, S. Vicatos, A. Dryga and A. Warshel, *Annu Rev Phys Chem*, 2011, **62**, 41-64.
45. S. Lin and D. Blankschtein, *The Journal of Physical Chemistry B*, 2010, **114**, 15616–15625.
46. S. Lin, A. J. Hilmer, J. D. Mendenhall, M. S. Strano and D. Blankschtein, *J Am Chem Soc*, 2012, **134**, 8194-8204.
47. A. J. Hilmer, T. P. McNicholas, S. Lin, J. Zhang, Q. H. Wang, J. D. Mendenhall, C. Song, D. A. Heller, P. W. Barone, D. Blankschtein and M. S. Strano, *Langmuir*, 2012, **28**, 1309-1321.
48. B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *Journal of Chemical Theory and Computation*, 2008, **4**, 435-447.

Soft Matter Accepted Manuscript

49. S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman and A. H. de Vries, *The Journal of Physical Chemistry B*, 2007, **111**, 7812-7824.
50. H. Lee, A. H. de Vries, S.-J. Marrink and R. W. Pastor, *The Journal of Physical Chemistry B*, 2009, **113**, 13186-13194.
51. H. Lee and R. G. Larson, *The Journal of Physical Chemistry B*, 2009, **113**, 13202-13207.
52. E. J. Wallace and M. S. P. Sansom, *Nano Lett*, 2007, **7**, 1923-1928.
53. A. W. Schuttelkopf and D. M. F. van Aalten, *Acta Crystallographica Section D*, 2004, **60**, 1355-1363.
54. X. Periole, M. Cavalli, S.-J. Marrink and M. A. Ceruso, *Journal of Chemical Theory and Computation*, 2009, **5**, 2531-2543.
55. C. A. Lopez, A. J. Rzepiela, A. H. de Vries, L. Dijkhuizen, P. H. Hunenberger and S. J. Marrink, *Journal of Chemical Theory and Computation*, 2009, **5**, 3195-3210.
56. L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman and S. J. Marrink, *Journal of Chemical Theory and Computation*, 2008, **4**, 819-834.
57. S. Peng and K. J. Cho, *Nano Lett*, 2003, **3**, 513-517.
58. A. Leo, C. Hansch and D. Elkins, *Chemical Reviews*, 1971, **71**, 525-616.
59. V. N. Viswanadhan, A. K. Ghose, G. R. Revankar and R. K. Robins, *Journal of Chemical Information and Computer Sciences*, 1989, **29**, 163-172.
60. G. Klopman, J.-Y. Li, S. Wang and M. Dimayuga, *Journal of Chemical Information and Computer Sciences*, 1994, **34**, 752-781.
61. B. Gomez-Zaleta, R. Gomez-Balderas and J. Hernandez-Trujillo, *Phys Chem Chem Phys*, 2010, **12**, 4783-4790.
62. S. A. Best, K. M. Merz and C. H. Reynolds, *J Phys Chem B*, 1999, **103**, 714-726.
63. W. F. van Gunsteren and H. J. C. Berendsen, *Journal of Computer-Aided Molecular Design*, 1987, **1**, 171-176.
64. T. C. Beutler, A. E. Mark, R. C. Vanschaik, P. R. Gerber and W. F. Vangunsteren, *Chem Phys Lett*, 1994, **222**, 529-539.
65. D. M. Huang and D. Chandler, *J Phys Chem B*, 2002, **106**, 2047-2053.
66. S. Lin, C.-J. Shih, M. S. Strano and D. Blankschtein, *J Am Chem Soc*, 2011, **133**, 12810-12823.
67. K. Yum, J. H. Ahn, T. P. McNicholas, P. W. Barone, B. Mu, J. H. Kim, R. M. Jain and M. S. Strano, *Acs Nano*, 2012, **6**, 819-830.
68. A. Rajan, M. S. Strano, D. A. Heller, T. Hertel and K. Schulten, *The Journal of Physical Chemistry B*, 2008, **112**, 6211-6213.
69. M. Mabuchi, J. Shimada, K. Okamoto, Y. Kawakami, S. Fujita and K. Matsushige, *P Soc Photo-Opt Ins*, 2001, **2**, 140-148.
70. B. Koziol, M. Markowicz, J. Kruk and B. Plytycz, *Photochem Photobiol*, 2006, **82**, 570-573.
71. X. R. Xia, N. A. Monteiro-Riviere and J. E. Riviere, *Nat Nanotechnol*, 2010, **5**, 671-675.
72. X. R. Xia, N. A. Monteiro-Riviere, S. Mathur, X. F. Song, L. S. Xiao, S. J. Oldenberg, B. Fadeel and J. E. Riviere, *Acs Nano*, 2011, **5**, 9074-9081.

Soft Matter Accepted Manuscript

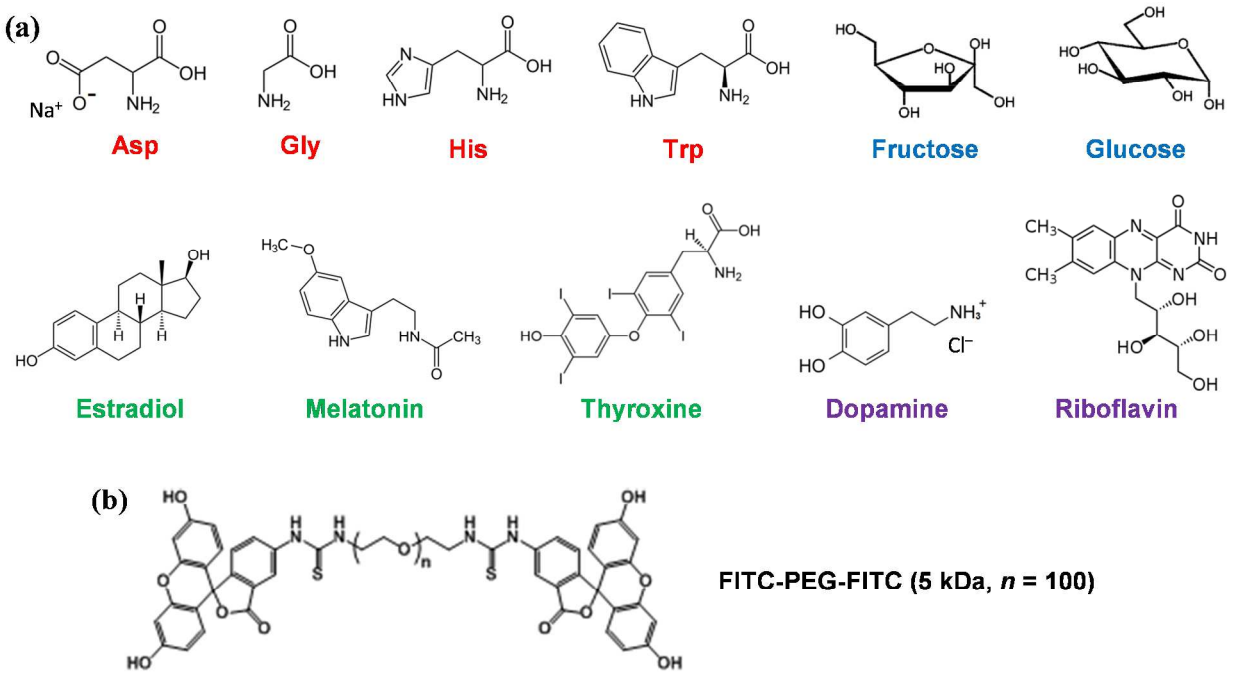
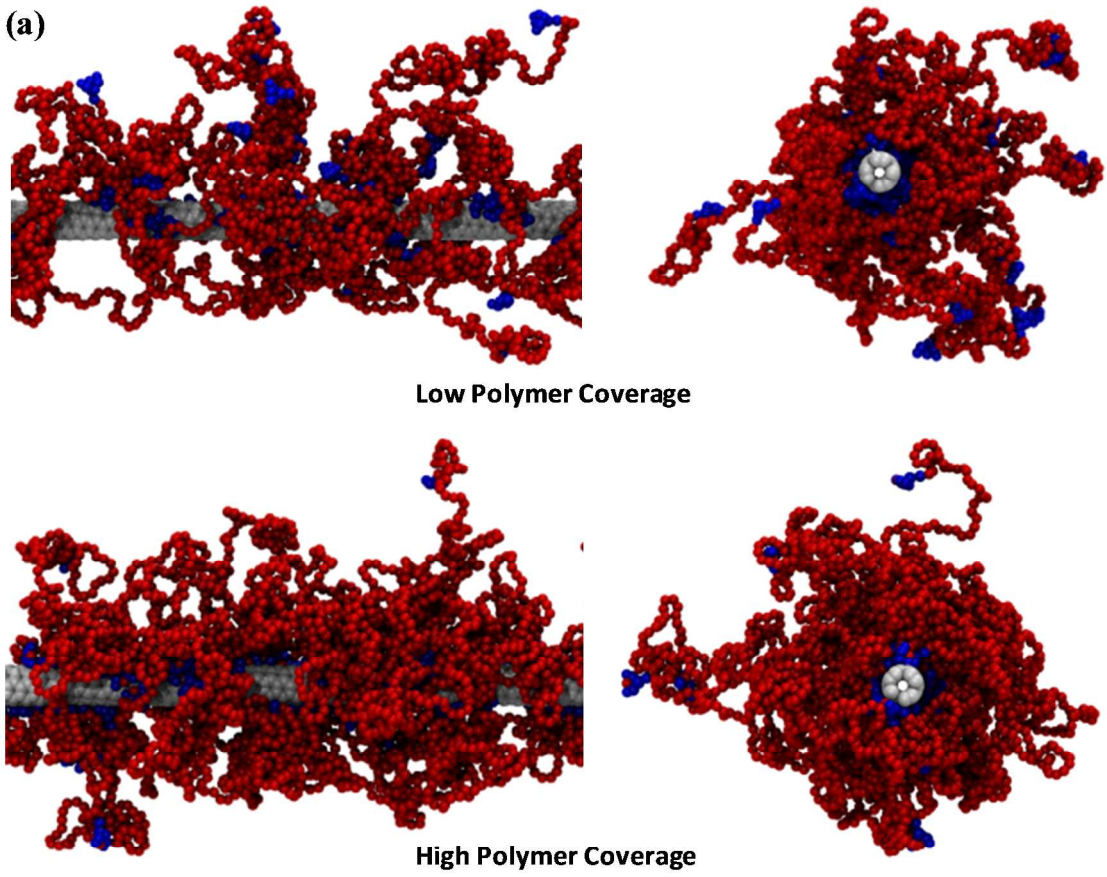


Figure 1. (a) Chemical structures of the 11 biologically-important analytes considered. Note that the colors used for the names of the analytes correspond to the colors of lines in Fig. 4 below for the same analytes. **(b)** Chemical structure of the PEG-based polymer considered.

Soft Matter Accepted Manuscript



Soft Matter Accepted Manuscript

Figure 2. (a) Representative post-equilibrium simulation snapshots of the SWCNT-polymer scaffold studied here (left panel – side view, and right panel – front view). Color codes: white – carbon nanotube, red – PEG phase, and blue – FITC groups. **(b)** CG particle number densities, $\rho(r)$, of FITC groups (top), PEG backbones (middle), and water molecules (bottom) around the SWCNT (at a distance, r, from the nanotube central axis) under both low (dashed lines) and high (solid lines) surface coverages. The lighter blue, red, and green areas represent the three phases in the simulated system, which are illustrated in detail in (c). **(c)** Illustration of the three phases in the simulated system using the simulation snapshot of the high coverage case in (a) as an example.

Soft Matter Accepted Manuscript

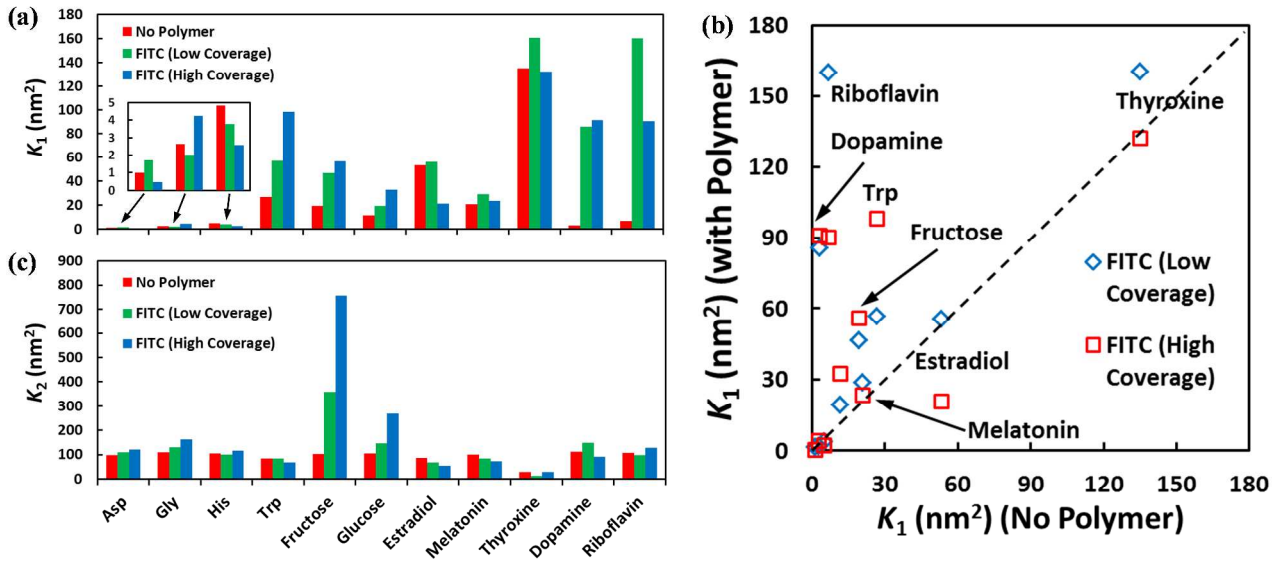


Figure 3. (a) Simulated Henry's constants, K_1, for each of the 11 analytes considered on the SWCNT-polymer scaffold for the bare SWCNT solution case (no polymer, red), the low polymer surface-coverage case (green), and the high polymer surface-coverage case (blue). The inset is zoom-in plot for the K_1 values of Asp, Gly, and His. **(b)** Correlations between the simulated Henry's constant, K_1, for the no-polymer case and those for the low-surface coverage (blue diamonds) and the high-surface coverage cases (red squares). The black dashed line denotes the region of perfect correlation. **(c)** Simulated Henry's constants, K_2, for each of the 11 analytes considered on the SWCNT-polymer scaffold.

Soft Matter Accepted Manuscript

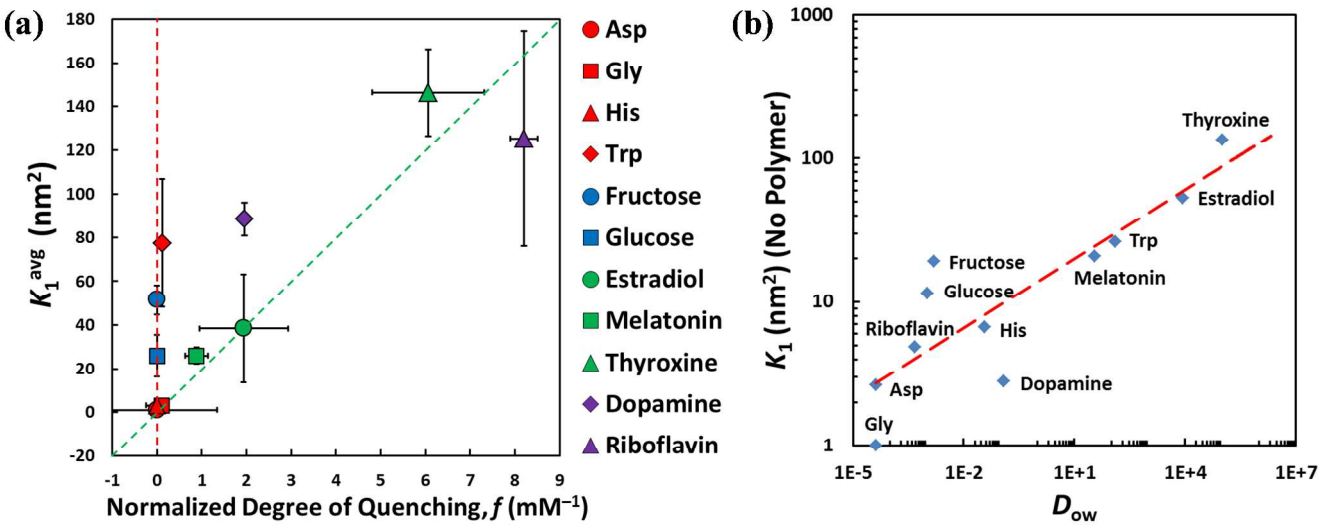
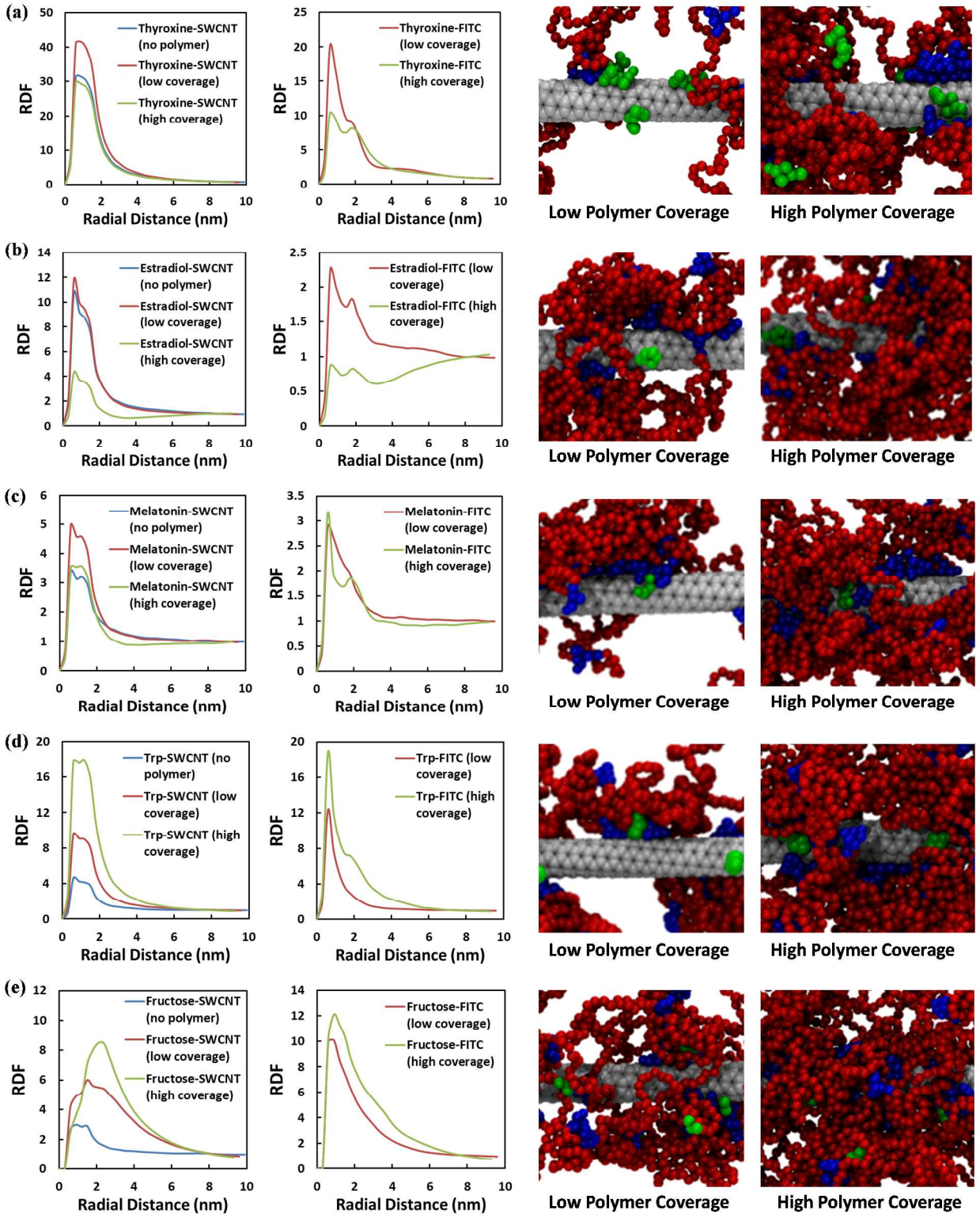


Figure 4. (a) Comparison between the binding affinities of the analytes on the SWCNT surface, K_1^{avg}, averaged over the low-polymer and the high-polymer surface coverage cases, with the experimental PL quenching results, quantified by the normalized degree of PL quenching, f (see **Eq. 11** for the definition). Two clear categories of analytes can be identified along the green (diagonal) and the red (vertical) dashed lines as guides to the eyes. The standard deviations in the experimental PL quenching results (repeated 3 times for each analyte) and in the K_1^{avg} values are shown with error bars. **(b)** Correlation between the simulated K_1 values (for the no-polymer case) and the octanol/water distribution coefficients, D_{OW}, for the 11 analytes considered. The values of D_{OW} for the four amino acids (Asp, Gly, His, and Trp) and the two monosaccharides (glucose and fructose) were obtained in Refs. 56 and 55, respectively. The values of D_{OW} for the other 5 analytes were obtained using the QSAR method, as discussed in **Section 2.2**. The red dashed line is a guide to the eye for the overall trend.

Soft Matter Accepted Manuscript



Soft Matter Accepted Manuscript

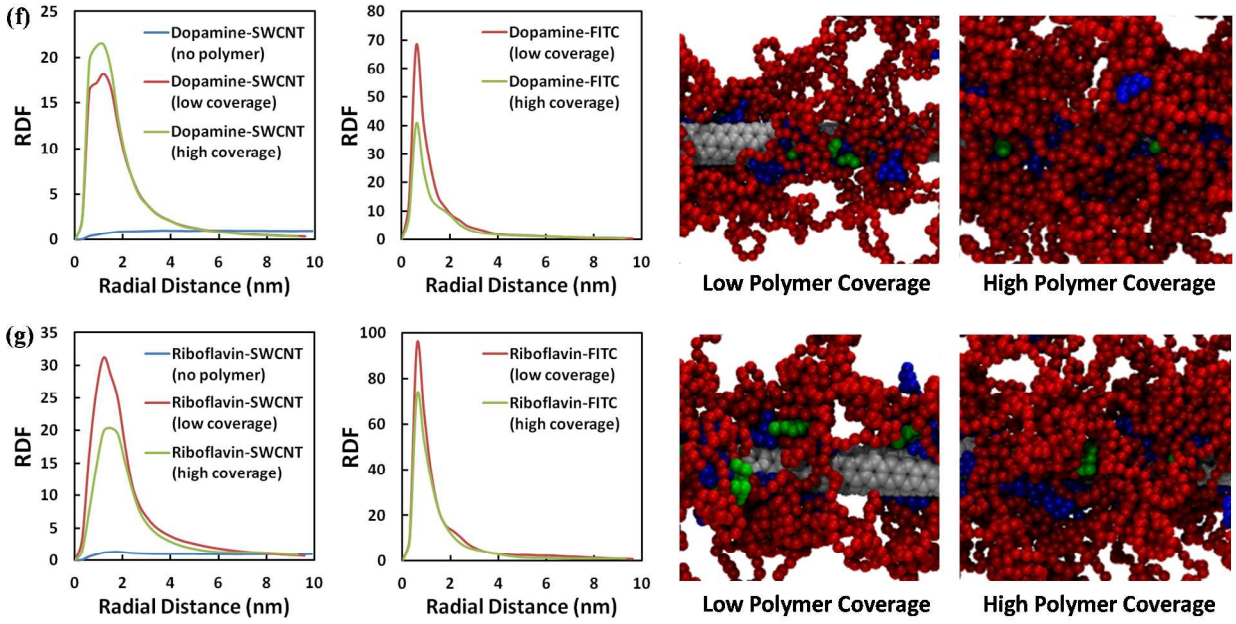


Figure 5. Radial distribution functions (RDFs) of highly-bound analyte molecules around the CG particles of the SWCNT (left column) and the FITC groups (middle column) at various polymer surface coverages: **(a)** thyroxine, **(b)** estradiol, **(c)** melatonin, **(d)** tryptophan, **(e)** fructose, **(f)** dopamine, and **(g)** riboflavin. Corresponding simulation snapshots of the binding events of the above analytes are shown in the right column. The color code is the same as in **Fig. 2,** with green color corresponding to analyte molecules.

Soft Matter Accepted Manuscript