ChemComm

Single-wall Carbon Nanotube Separations via Aqueous Two-phase Extraction: New Prospects Enabled by High-throughput Methods

Journal:	*ChemComm*
Manuscript ID	CC-FEA-11-2024-006096.R1
Article Type:	Feature Article

FEATURE ARTICLE

Single-wall Carbon Nanotube Separations *via* Aqueous Two-phase Extraction: New Prospects Enabled by High-throughput Methods

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Christopher M. Sims, Ming Zheng, and Jeffrey A. Fagan

Aqueous two-phase extraction (ATPE) is an effective and scalable liquid-phase processing method for purifying single species of single-wall carbon nanotubes (SWCNTs) from multiple species mixtures. Recent metrological developments have led to advances in the speed of identifying solution parameters leading to more efficient ATPE separations with greater fidelities. In this feature article, we review these developments and discuss their vast potential to further advance SWCNT separations science towards the optimization of production scale processes and the full realization of SWCNT-enabled technologies.

1. Introduction

Single-wall carbon nanotubes (SWCNTs) are a family of cylindrical carbon allotropes with exceptional properties that make them useful for multiple applications, including sensing, digital logic, and single photon sources. Each species (often termed chirality) of SWCNT has a distinct crystalline structure denoted by an (*n,m*) "roll-up" vector on a graphene lattice, which succinctly contains information on the nanotube diameter and lattice orientation. The specific dimensionalities of each individual (*n,m*) SWCNT yield unique mechanical, thermal, optical, and electronic properties. To tailor and select properties for different applications, control over the SWCNT species present in a population is essential. This is desirable to several distinct levels: all nanotubes of the same electronic type, *i.e.,* metallic, or semiconducting; all nanotubes of the same (*n,m*), and most exhaustively, all nanotubes of a single (*n,m*) being of the same enantiomer, *i.e.*, having the same twist direction (left- or right-hand) of their lattice around the long axis. Unfortunately, and despite much progress, commercial SWCNT production methods generate populations containing mixtures of multiple (20+) SWCNT species. As a result, post-production enrichment and isolation of individual SWCNT (*n,m*)s from these polydisperse mixtures, colloquially known as "sorting", has received much research attention over the past two decades.

Many methodologies have been developed to sort SWCNTs, most commonly using liquid-phase processing methods following dispersion into either an aqueous or organic solvent. Our focus is on aqueous sorting methods, although organic-phase conjugated polymer extraction (CPE) and related methods will be described or referenced. Aqueous liquid-phase processing methods include ion exchange chromatography (IEX),[1-4] density gradient ultracentrifugation (DGU),[5-8] electric field fractionation,[9, 10] gel chromatography (GC),[11-14] and aqueous two-phase extraction (ATPE).[15-18] CPE, GC, and ATPE are the techniques currently in wide use. Each of these sorting methods (which have been recently reviewed elsewhere[19-21]) have their own unique capabilities, yet all are principled on the manipulation of non-covalent, physical interactions between the SWCNT surface and dispersing agent. Depending on the dispersing media, these agents include conjugated polymers (*e.g.,* polyfluorene or polycarbazole), surfactants (*e.g.,* sodium dodecyl sulfate (SDS), sodium deoxycholate (DOC), or Triton®), and biomolecules (*e.g.,* DNA, peptides). At the National Institute of Standards and Technology (NIST), our team has focused on the development and optimization of the ATPE method for SWCNT sorting for the past decade. More specifically, we utilize two distinct, but parallel, strategies of controlling the SWCNT selection process *via* ATPE: DNA-wrapping or surfactant competition. In both strategies, ATPE offers several benefits, including a high degree of tailorability, flexibility to target any (*n,m*), a reasonable efficiency of the process, *i.e.,* quick performance, low operating and capital costs, and anticipatable scalability. As such, ATPE will be highlighted through the rest of this feature article.

The ATPE method, first reported by Albertsson,[22] typically involves the mixing of two water-soluble polymers, in water, at concentrations which result in the solution undergoing a thermodynamic phase separation into two separate, but permeable phases. More broadly, self-separating aqueous two-phase systems can also be formed by mixtures of other compounds. Notable examples are two-phase systems formed by mixing a polymer with a salt solution; these are emerging as particularly important for DNA-controlled separations of SWCNTs *(vide infra*), and such systems are also described by the ATPE acronym. For any ATPE sorting process, the functional goal

Materials Science and Engineering Division, National Institute of Standards and Technology, Gaithersburg, MD 21045, USA. E-mails: christopher.sims@nist.gov, ming.zheng@nist.gov, jeffrey.fagan@nist.gov

is for the two phases to have different affinities for any solutes, *i.e.*, particles or nanoparticles, contained within the initial volume of the mixture, such that the microphase separation, physical coalescence, and spatial separation of the polymer phases results in selective fractionation of the dispersed solutes as well. Another key aspect of a well-implemented ATPE system is for the phase separation to occur *via* spinodal decomposition rather than *via* nucleation,[22, 23] as this results in the chemical gradients that enable solute selection of their desired phase to occur (generally) on the approximately micron spatial length scale. Additionally, spinodal decomposition occurs effectively instantaneously and simultaneously throughout the liquid volume due to thermodynamics, whereas nucleation is controlled by random fluctuations and can have a long kinetics-based delay.

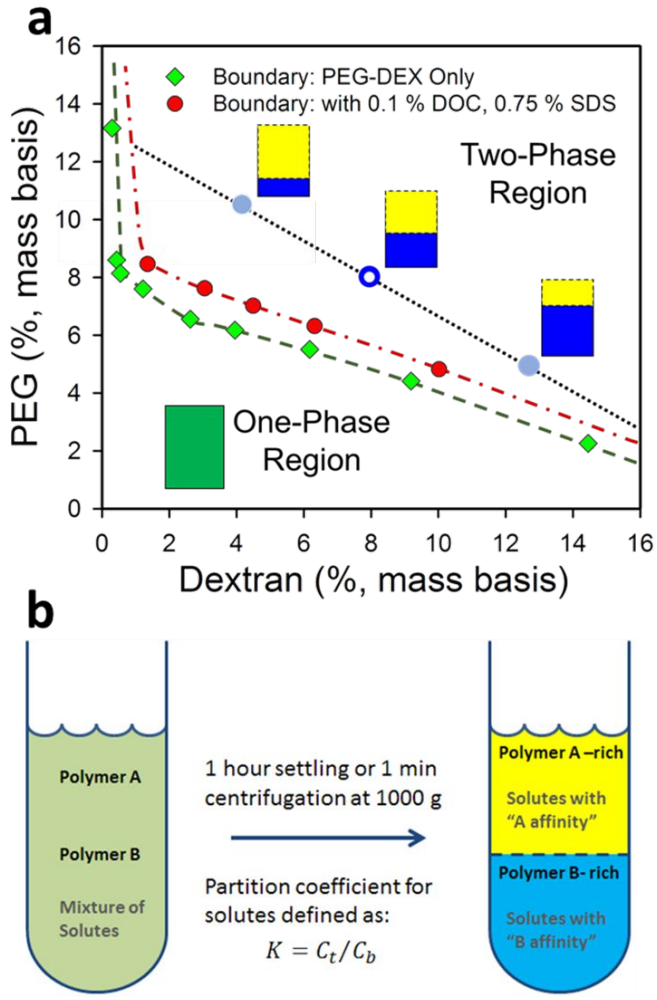

Fig. 1 (a) Phase diagram of an ATPE system composed of water and two polymers, 6 kDa PEG and DEX 70. At equilibrium, all concentrations of PEG and DEX within the two-phase region (such as at the open blue circle) will form two phases at compositions located along and above the coexistence curve (dashed green eyeguide). In the case of PEG:DEX, the upper phase is PEG-rich, while the lower phase is DEX-rich. Any initial composition along a tie-line will produce two phases with identical composition of the two components, but at different volume ratios due to mass conservation (e.g. schematic volumes along the black dotted eyeguide connecting the blue circles). Note that the coexistence curve will shift in the presence of additional non-polymer components such as DOC or SDS (dot-dash red eyeguide) (b) Schematic of an ATPE separation, where the splitting of the mixture into two phases yields separation of solutes with differing affinities for each phase. In most ATPE systems, the spontaneous separation can be sped up with centrifugation. Adapted from Fagan, *Nanoscale Adv.*, 2019, **1**, 3307-3324 with permission from the Royal Society of Chemistry.

Many different two-phase systems can be constructed, each with their own unique phase boundaries and critical concentration points. For two-polymer systems, the location of the two-phase coexistence boundary is dependent on the chemical properties of the polymers, their molecular mass distributions, and often the solution environment (*e.g.*, temperature, concentration of other solution components).[23] At equilibrium, for any set of polymer concentrations beyond the two-phase coexistence boundary, *i.e.*, the critical concentration line, the mixture will form two phases having concentrations at the phase boundaries, with one phase rich in one of the polymers (but containing some of the other) and vice versa. The relative volumes of the two phases are determined by the slope of the "tie-line" connecting the two endpoint concentrations and the position of the initial, global, concentration along its length. Notably, any initial composition of the two polymers along a particular tie-line will separate into phases having identical compositions (dictated by thermodynamics), but of different volumes as set by the mass balance. An example phase diagram of a two-phase system comprised of polyethylene glycol (PEG) 6 kDa and Dextran (DEX) 70 in water is shown in Fig. 1a. Here, three representative examples of the relative volumetric partitioning are shown for three different composition points along the same tie-line; the chemical composition of the blue (yellow) phase is the same for all three example points even though the volume ratio changes.

The extraction aspect of ATPE relies on selective partitioning, based on differences of affinity, of the SWCNTs into one of the two phases. This is shown schematically for any type of solute with differential affinity in Fig. 1b, in which the phase separation and spatial segregation of the polymer phases representationally fully resolves yellow and blue solutes from an initially homogeneously green-coloured mixture of the solute pair. By extension, starting from an initial mixed population of SWCNT species, if differences in affinity for the two polymer phases can be induced for particular SWCNT populations, then fractionation should be possible. As discussed above, each (*n,m*) SWCNT has distinct material properties, and thus may be tailorable to have a different affinity from all other (*n,m*) species for a particular combination of polymers, dispersants, and other additives by a single or series of ATPE separations. As such, an overarching goal of ATPE research has been to determine which combinations of these controllable parameters best, *i.e.,* optimally, resolve and isolate targeted individual (*n,m*)s from mixed SWCNT populations.

Even from this introduction, in which we have described in general terms the two-phase systems, dispersants, and additives used, one can quickly ascertain that the phase space for potential variations and combinations of dispersant and two-phase chemistries is inordinately large and infeasible to empirically explore in detail. Even with this limitation, substantial progress has been made in identifying many ATPE combinations, with several distinct methodologies of ATPE delivering high yield and high purity (*n,m*) SWCNT populations. Despite the identification of specific methods, and the degree of their optimization over the past decade, there is still an

obvious desire to develop improved separations. This implies a need for efficient means of discovering new ATPE systems and separation conditions.

2. New Methodological Approaches for Facilitating SWCNT Separations

Determining effective separation conditions for specific (*n,m*) SWCNTs, or the means to identify those conditions, has been a driving motivation of our NIST team efforts over the past several years. Much effort has been applied towards developing higher throughput approaches for elucidating effective SWCNT sorting conditions using ATPE. Here, we review the developmental sequence of these advances, with distinct strategies for DNA-wrapping controlled and surfactant competition-based ATPE. To provide context, we start each of the sections with a brief amount of background information specific to the DNA-wrapping or surfactant competition ATPE sorting strategies.

2.1 Sorting SWCNTs *via* Sequence-Dependant DNA-Wrapping

A naturally occurring polymeric biomolecule, DNA is comprised of nucleotide monomers, with each nucleotide consisting of a deoxyribose sugar, phosphate group, and one of the four native nitrogenous bases, guanine (G), adenine (A), cytosine (C) or thymine (T). Single-stranded DNA (ssDNA) results from linking these nucleotides through a hydrophilic sugar-phosphate backbone, leaving the nitrogenous bases unpaired and exposed. It is through these nitrogenous bases that ssDNA is thought to form strong π-π interactions with the SWCNT surface, resulting in stable aqueous dispersions of DNA-wrapped SWCNTs (DNA-SWCNTs), as first reported by Zheng *et al.* (and illustrated in Fig. 3a)[1]

As may be apparent, the potential number of ssDNA sequences is limited only by physical possibility. The number of potential combinations increases as 4^x with *x* here being the number of bases in the sequence (which are usually denoted *x*-mer), with potentially any sequence being useful for dispersion and resolution of the SWCNT (*n,m*)s. Initial efforts investigating DNA-SWCNT interactions began with simple patterns, such as single nucleotide homopolymers [*e.g.*, poly(T)] or alternating purine-pyrimidine sequences [*e.g.*, poly(GT)]. These fundamental explorations suggested that certain ssDNA sequences were capable of resolving different (*n,m*) SWCNT species, presumably by forming highly ordered structures around the SWCNTs *via* an extensive hydrogen bonding network between neighbouring strands.[2] Deeper investigation suggested that these ordered DNA structures also differed based on the electronic type and diameter of the SWCNT they interacted with, which could be leveraged to separate SWCNTs based on their physical characteristics.[1-4] In 2009, it was demonstrated that many single-chirality SWCNTs could be purified *via* IEX using specific short DNA sequences identified through a limited but systematic sequence search.[24] Importantly, these early works confirmed that the DNA-SWCNT interaction was dependant on both the DNA sequence *and* the SWCNT structure, from which we brought forth the concept of finding efficient ways to identify DNA sequences for targeted SWCNT (*n,m*) species purification, which we termed "recognition sequences" at the time.[24-26]

While these initial advances in DNA-SWCNT separations were made using IEX, we have since shifted to using the ATPE method at NIST. Compared to IEX, ATPE has the advantage of having well-defined and highly tunable homogenous solution phases generated from simple and readily purchased chemicals. This allows for greater control and reproducibility, but also independent scaling of separation volume and enabling recollection of all SWCNT materials (*i.e.*, no loss); together all these factors increase the efficiency of the separations. Fortunately, our first DNA-SWCNT studies with ATPE showed that many of the ssDNA sequences previously identified through IEX experiments were transferrable to ATPE, successfully purifying several metallic and semiconducting SWCNT species and their left- and right-handed enantiomers from synthetic mixtures.[16, 27] Another important outcome from moving to the ATPE method was that it allowed us to better

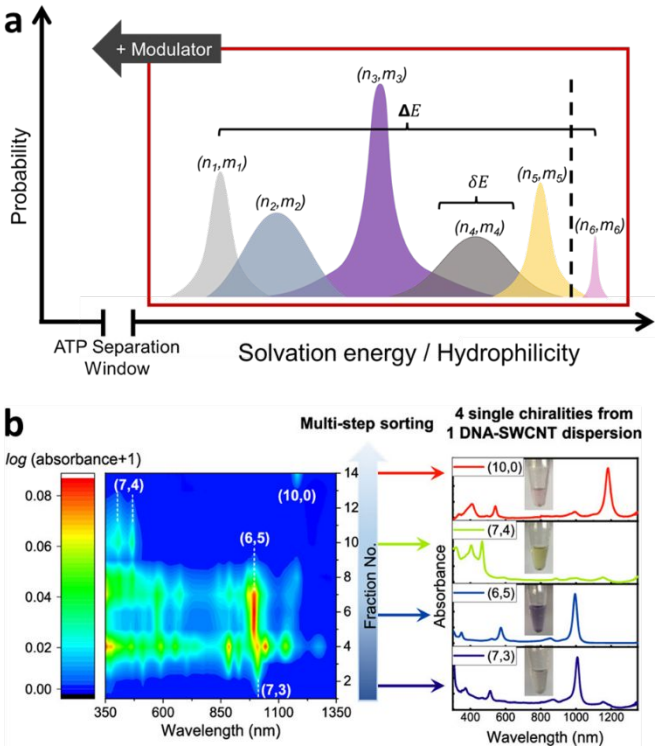

Fig. 2 (a) A schematic of the solvation energy distribution spectrum formed from a given DNA-SWCNT dispersion featuring different (*n,m*) SWCNT species wrapped by the same DNA sequence. Note that each individual DNA-SWCNT hybrid has a solvation energy *distribution* owing to the many ways the DNA could wrap around a SWCNT, with more ordered structures expected to produce narrower distributions. Such ordered structures, such as (n_6,m_6), are less likely to overlap with other species and thus be easier to separate from a mixture (black dashed line). DNA-SWCNTs at either end of the spectrum, such as (n_1,m_1), are also expected to be easier to isolate from the mixture. In this schematic example, the energy spectrum is located completely in the bottom phase of an ATPE system; with the gradual addition of modulators (e.g. salt, polyvinylpyrrolidone (PVP), etc.), the distribution shifts towards the top phase, sequentially extracting the different (*n,m*) SWCNTs. An example of this process is seen in (b), where four different (*n,m*)s were exacted from a single $(TCTCCC)_2TCT$ dispersion through the gradual addition of PVP in a PEG:DEX ATPE system. Adapted from Lyu et al., *J. Am. Chem. Soc.*, 2019, **141**, 20177-20186 with permission from the American Chemical Society.

probe the underlying separation mechanisms, under the hypothesis that better understanding will support design of more optimized DNA-SWCNT purifications. As noted earlier, SWCNT separations in ATPE are dependent on prescribing differing affinities for the many SWCNT species for the two phases, such that some (or one) will segregate into one phase, and the rest into the other. We have previously proposed a framework for thinking about and quantifying affinity differences for any specific SWCNT, in terms of its solvation energy in, and relative difference across, the two phases.[28] In the case of DNA-ATPE, this solvation energy, and hence partitioning behavior, is primarily influenced by the wrapping DNA sequence. We hypothesize that for a given DNA sequence, each (*n,m*) SWCNT will have its own average solvation energy, with different SWCNT species having broader or narrower distributions and more or less distinct values of their solvation energy depending on the structure and specificity of the DNA coverage of the nanotube surface (*i.e.*, wrapping). A schematic of the solvation energy distribution of different (*n,m*) SWCNTs when coated by the same DNA sequence is shown in Fig. 2a.

In practice, there are two modalities of DNA-ATPE separations that differ in how they are performed. In some cases, due to the solvation energy distribution, only a single nanotube species will reside in the top or bottom phase, enabling immediate extraction in a fully purified population. We call the DNA sequences enabling such selective extraction "recognition sequences." An example of this is equivalent to performing an extraction at the conditions of the black dashed line in Fig. 2a. Two notable sequences that fall into this category are the 12-mer $TTA(TAT)_2ATT$, which we term "super sequence (6,5)", and 15-mer $T_3C_3T_3C_6$, which we term "super sequence (8,3)". They are termed "super sequences" because these sequences are respectively able to isolate both the left- and right-handed enantiomers of the designated (*n,m*) species across different ATPE systems. Alternatively, or in addition, solvation modulating molecules can be added to change the location of the solvation energy cut line, through which many other (*n,m*)s can sometimes be extracted in narrow modulator concentration windows into the top phase; these sequences we term "resolving sequences." An example of this is seen in Fig. 2b, where (7,3), (6,5), (7,4), and (10,0) SWCNTs, all coated by a single 15-mer in $(TCTCCC)_2TCT$, were extracted in a 1.5 kDa PEG/DEX 250 ATPE system through the gradual addition of polyvinylpyrrolidone (PVP) as a modulator. It should be emphasized that a "recognition sequence" is a special case of "resolving sequence", corresponding to a DNA sequence that allows only one SWCNT species to be purified.

While several parameters beyond the DNA sequence still need to be considered in the DNA-ATPE system, such as the composition of the two phases or the identity and use of modulating agents, the observation that the DNA sequence plays the most important role in the process presents a double-edged sword. There are practically limitless options for wrapping SWCNTs with ssDNA, offering boundless potential for designing and employing unique DNA-SWCNT hybrids, but the sheer number of possibilities presents an equally substantial challenge in identifying said hybrids. Until 2019, all of the identified ssDNA sequences leading to purification of (*n,m*) SWCNT species had been empirically identified through multiple years of systematic searching through the DNA sequence space by physical experiments. While these experiments helped us to develop a better understanding of the DNA-SWCNT interactions that drive the purification of targeted SWCNT species, the process was time-consuming and inefficient. Even with the knowledge gained, we were still unable to fully utilize that information to predict new and potentially more favorable DNA sequences as the known sequences did not sufficiently guide future successful discoveries.

2.2 Finding DNA Resolving Sequences with Machine Learning

To complement pattern identification and heuristic approaches we turned to machine learning (ML) techniques, in collaboration with Anand Jagota, to more efficiently search the DNA sequence space.[29] Our initial aims were to build models to categorize DNA sequences as resolving or non-resolving, where a "resolving" sequence could yield (*n,m*)-pure SWCNTs (pure defined as > 90 % purity in either upper or lower phase), whereas a "non-resolving" sequence could not isolate any pure species. These "non-resolving" sequences typically only produce multi-(*n,m*) SWCNT mixtures (Fig. 3a). To reduce the complexity of our initial ML-study the first training set was composed of experimental data from a previous empirical search of 82 12-mers (T/C bases only) that identified 9 sequences as "resolving" (~10 % success).[27] This training set was transformed into numerical vectors using two different methods: position specific vector (*psv*) and term frequency vector (*tfv*) using overlapping *n*-grams (*n* = 1-3). An *n*-gram is defined as a substring of the length of *n* units from a given sequence; for example, if *n* = 2, the sequence "XYZ" can be translated into "XY" or "YZ" by counting each character as gram. The *psv* method conserves each *n*-gram at each position by utilizing one hot encoding for each *n*-gram at each position, while the *tfv* method counts the *n*-grams in the sequence and uses this count as a feature. During the first round of learning, we trained models by using three different algorithms of increasing complexity: Logistic Regression,[30] Support Vector Machine,[31] and Artificial Neural Network (ANN)[31]

After a training and validation cycle, the highest performing models (namely, LR models with bigram *psv* features and ANN models with trigram *tfv* features) were used to predict new resolving sequences from the full sequence library (2^{12}).[29] Given the small training set size relative to the query space (82 of 4096), we cross-validated sequences for ATPE experimental testing by selecting a subset of 10 sequences from the models' intersecting classification results; each was initially designated as "resolving". Of the 10 predicted "resolving" sequences, we found five to be capable of resolving single (*n,m*)s to the sufficient level of purity, a remarkable first result of a 50 % success rate compared to the ~10 % success rate from the empirical training set. We then retrained the models using this updated data set, producing an additional 10 sequences. Although the retrained models showed improved validation

performance, when tested experimentally the successful prediction rate remained at 50 %, showing that we could still make meaningful improvements to our models. Towards improving the next round of models and experiments, we next searched for structural motifs in the predicted "resolving" sequences using saliency measures within the ANN models.[32] This approach found higher saliencies in the first and last segments of the sequences, which suggested that the end bases of the DNA sequences were more influential in whether that sequence was "resolving" or not, which agreed with previous empirical studies of DNA-SWCNTs.[25, 33] In particular, a three-base "CCC" motif was found with notable frequency, especially in the newly identified "resolving" sequences.[27]

The promising results of this work encouraged us to apply this ML approach to systematically explore shorter DNA sequences.[34] Starting with 5-mers, which are the shortest synthetically viable DNA oligos, there was the feasible possibility of screening the entire sequence library ($4^5 = 1024$). We started by testing 56 5-mer sequences with ATPE experiments to serve as the first ML training set, classifying 22 as "resolving" with the remaining 34 as "non-resolving" (Fig. 3b). We then used a similar approach as our previous study[29] to build, train, and validate the models from the initial training set, resulting in competent, well-trained models. Since 56 sequences could not cover the entire 5-mer library, we added a constraint such that all the sequences chosen for the next cycle of physical experiments would differ from each other by at least two bases to better cover the 5-mer library. For example, the 5-mers CC**A**G**G** and CC**G**G**T** are sufficiently different, while the 5-mers GG**G**CG and GG**C**CG are not different enough as they differ

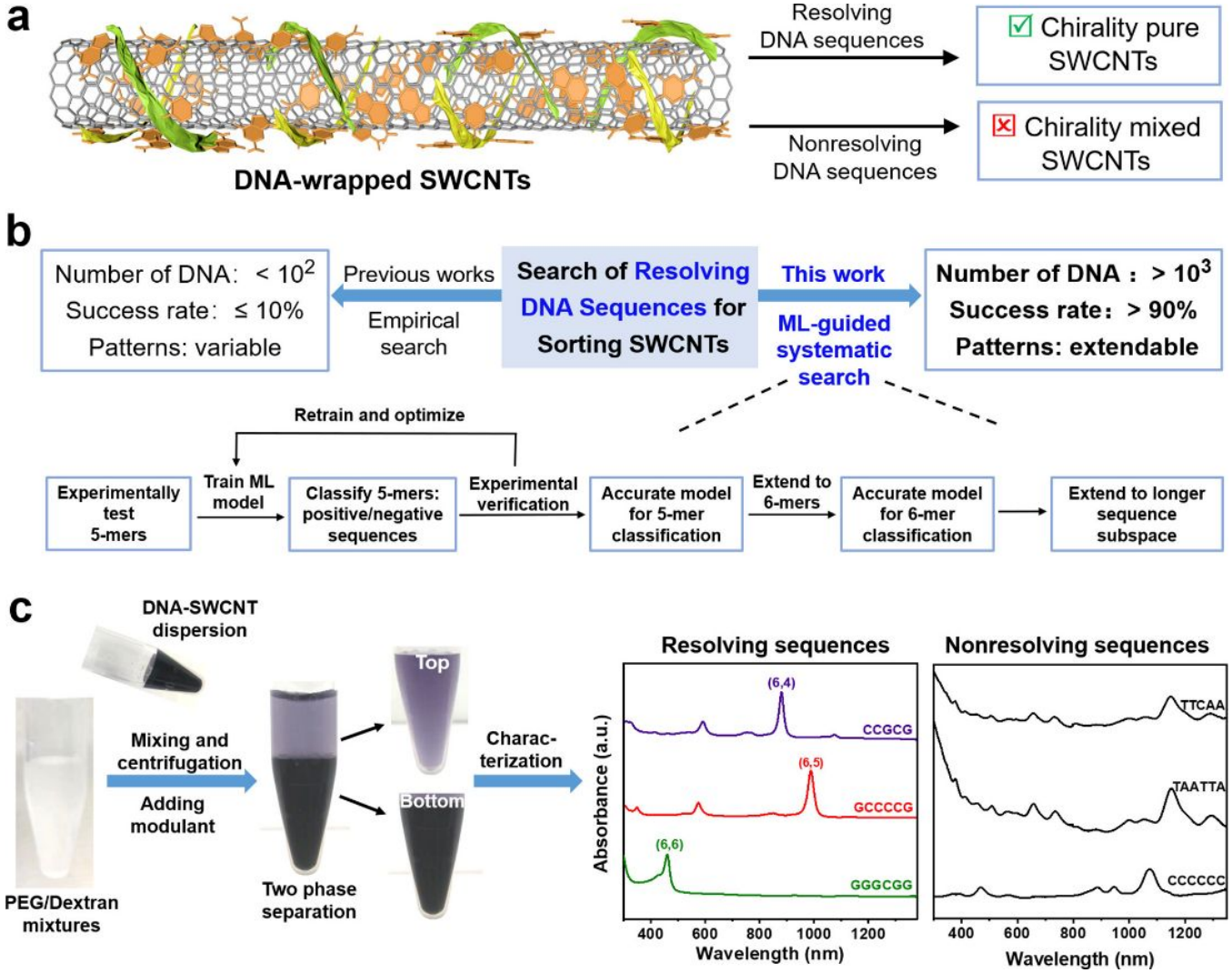

Fig. 3 (a) Machine learning (ML)-guided systematic search of DNA sequences for sorting SWCNTs. (a) Schematic illustration of DNA-wrapped single-wall carbon nanotubes (SWCNTs). Resolving DNA sequences can purify chirality-pure SWCNTs, while non-resolving DNA sequences only give chirality-mixed SWCNTs. (b) Previously, discovery of resolving DNA sequences was generally based on an empirical search. The number of resolving DNA sequences reported was small ($< 10^2$), the success rate was quite low (≤ 10 %), and the resolving sequences have dramatically distinct patterns. Alternatively, the discovery of resolving DNA sequences can be based on a machine learning (ML)-guided systematic search. This approach substantially increases the number ($> 10^3$) and success rate (> 90 %) of finding resolving sequences, which can be extended to sequence search in longer DNA subspaces. (c) Schematic illustration of using the ATPE technique to examine and test resolving/non-resolving DNA sequences. Here, one step separation is demonstrated. The bottom phase can be further extracted multiple times by adding a blank top phase. If the DNA–SWCNT hybrids can give rise to at least one type of single-chirality SWCNTs with a purity of ≥90 % in either the top or bottom phase, that sequence is classified as a resolving sequence; otherwise, it is classified as a non-resolving sequence. The identification relies on the UV–vis-NIR absorbance spectra of sorted fractions. Six typical examples of resolving/non-resolving sequences are illustrated. Reprinted from Lin et al., *ACS Nano*, 2022, **16**, 4075–4713 with permission from the American Chemical Society.

by only one base. From this, we selected 10 sequences predicted to be "resolving" and 10 sequences predicted to be "non-resolving" to test experimentally. Impressively, all 20 sequences performed as predicted, yielding a 100 % success rate (Fig. 3c).

We next moved on to 6-mers, which have a much larger sequence library (4^6 = 4096). We performed another retraining cycle using the new information from the 5-mer set to prepare our ML models for the 6-mer sequence space. Our models predicted that 744 of the 6-mer sequences would be "resolving", with the remaining 3352 sequences classified as "non-resolving". Similar to the 5-mer exploration, we added a constraint to help better explore the extensive sequence space. In the case of the 6-mers, we selected physical test sequences such that they would differ from each other by at least three bases. Again, choosing 10 sequences predicted to be "resolving" and 10 sequences predicted to be "non-resolving", our experimental results this time found that all 10 "resolving" sequences agreed with prediction, while 2 of the expected 10 "non-resolving" sequences were actually able to resolve a particular (*n,m*). We expanded our physical experiments to include an additional 62 sequences predicted to be "resolving", with 57 of them being able to resolve at least one (*n,m*) to sufficient purity. In total, we tested 82 6-mer sequences using the ATPE method, with 75 of them matching the ML-guided prediction of "resolving" or "non-resolving". This > 90 % success rate is a substantial improvement over the ~10 % previous empirical success rate.

Along with this considerable increase in success rate and anticipated expansion of (*n,m*) resolving sequences, it's also important to note what (*n,m*) SWCNTs we were able to purify and what new information regarding the DNA sequence structure we found that could help inform future designs. In terms of SWCNT species, we found resolving sequences covering the entire spectrum of types, from armchair metallic SWCNTs [(5,5) and (6,6)] and quasi-metallic species [(7,4), (8,5), and (9,6)] to a range of semiconducting species across different chiral angle and diameter ranges, such as the zigzag (10,0) to the near-armchair (8,7). Moreover, we were able to resolve relatively low-abundance SWCNT species such as the (9,5), which was a species we had previously been unable to purify using DNA-ATPE. Additionally, using circular dichroism (CD), which probes the handedness (left- or right-) of (*n,m*) enantiomers, we found that several of our resolved (*n,m*)s were highly enantiomerically enriched, and in many cases, more enriched than the previously reported exemplar [specifically, (6,4), (7,3), (7,5), (8,4), (8,5), (9,4), and (8,6)]. In all, the ML-guided search identified resolving sequences that could purify 22 SWCNT species with a broad range of both (*n,m*) and enantiomerically enriched purities. Finally, we noted that we were able to resolve most of the (*n,m*) species using only (G/C) 6-mers, suggesting a potential structural theme. Taking a subset of four resolving (G/C) 6-mers, we tripled their length to 18-mers, surprisingly finding that all four of the 18-mers could resolve a different (*n,m*) [specifically, (6,5), (7,4), (8,3), and (10,0), a set that includes both a zigzag and a quasi-metallic species]. Remarkably, a further length doubling of the two 18-mers that resolved the (6,5) and (8,3) species to their respective 36-mers found that those two sequences were also able to resolve the (6,5) and (8,3) species. These results suggested that the (G/C) combination might have a "super-resolving" capability that could also be extended to longer sequences. We are already building on these results to investigate 7- and 8-mer sequences for their resolving capabilities, with preliminary findings yielding promising results of new sequences for purifying previously difficult to isolate (*n,m*)s. Based on these successes in sequence identification and the inherent structural insights gleaned, we envision that the ML-guided strategy has immense potential for efficiently discovering ssDNA sequences leading to comprehensive (*n,m*) sorting of single-species and enantiomer SWCNTs. Such a capability will have further distinct and direct applicability for biotechnology applications by way of the unique DNA-SWCNT hybrids that can be produced.[34]

2.3 Surfactant Competition-based ATPE: Improving Determination and Screening of Conditions for (*n,m*) Separations

Surfactants are perhaps the most common dispersants used for liquid-phase sorting of SWCNTs, with a wide variety of both ionic and non-ionic surfactants demonstrated to individually disperse SWCNTs in water with varying levels of effectiveness.[35] Of these, the most common surfactants currently used are detergents such as sodium dodecyl sulfate (SDS) and sodium dodecylbenzenesulfonate (SDBS), or bile salt family surfactants such as sodium cholate (SC) and sodium deoxycholate (DOC). Bile salts, and in particular DOC, are notably effective and selective dispersants for SWCNTs, and are key to common implementations of several of the aforementioned liquid-phase processing methods including DGU, GC, and ATPE. Despite their outward differences, we hypothesize that the many surfactant-controlled SWCNT separation methods mostly share a similar mechanism, explicitly that differential adsorption of competing surfactants on each distinct SWCNT (*n,m*)'s surface amplifies differences in their physicochemical properties and thus enables selective fractionation. In DGU, for example, such differential adsorption induces differences in the solvated density of metallic and semiconducting SWCNT species of the same diameter; this enables spatial segregation of the two types through application of strong acceleration forces to drive the SWCNTs to their different isopycnic locations within a density gradient. For ATPE, while it is instead the density differences between the phase separating layers that drive spatial segregation, the nature of the adsorbed layer on the SWCNT is what controls (*vide infra*) the affinity for the two phases, and thus controls the selection for separation. Careful control of both the type(s) and concentration(s) of surfactant (and other components) competing for adsorption to the SWCNT interface thus results in selective extraction. Explicitly, in the case of surfactant-ATPE, the selective part of the SWCNT partitioning into one of the two polymer-enriched phases occurs as the solution undergoes phase separation, with the best mass transfer efficiency achieved when this occurs *via* spinodal decomposition.

At NIST, we have focused on developing a surfactant-ATPE system using 6 kDa PEG and DEX 70 as the two polymers; this results, as in Fig. 1, in a PEG-rich upper phase and DEX-rich lower phase. Multiple combinations of surfactants enable ATPE selection of SWCNTs in a differential manner. The most common are a competition of SC and SDS to differentiate semiconducting and metallic SWCNT types, and DOC and SDS to fractionate by (*n,m*) species or enantiomeric handedness.[36] In both cases, we hypothesize, and evidence indicates, that the bottom DEX phase contains those SWCNTs covered by the bile salt, with SDS-covered SWCNTs favouring the top PEG phase. Such evidence comes from both conducted experiments, *i.e.*, addition of more bile salt shifts SWCNTs to the DEX phase, and recently published orthogonal analytical ultracentrifugation and fluorescence measurements that strongly support this hypothesis, in particular for SDS and DOC competition.[37] For SDS and DOC, each (*n,m*) SWCNT species (and enantiomer)[18, 38, 39] is observed to have a specific ratio of the two surfactants at which the adsorbed layer flips in composition and the polymer phase of extraction changes. Fig. 4a. shows a schematic of this concept by way of a separation coefficient, which is defined as the ratio of the concentration of the *i*th nanotube species separating into the top (*c*T,*i*) versus bottom (*c*B,*i*) polymer phase; the (7,5) SWCNT is used as an example.

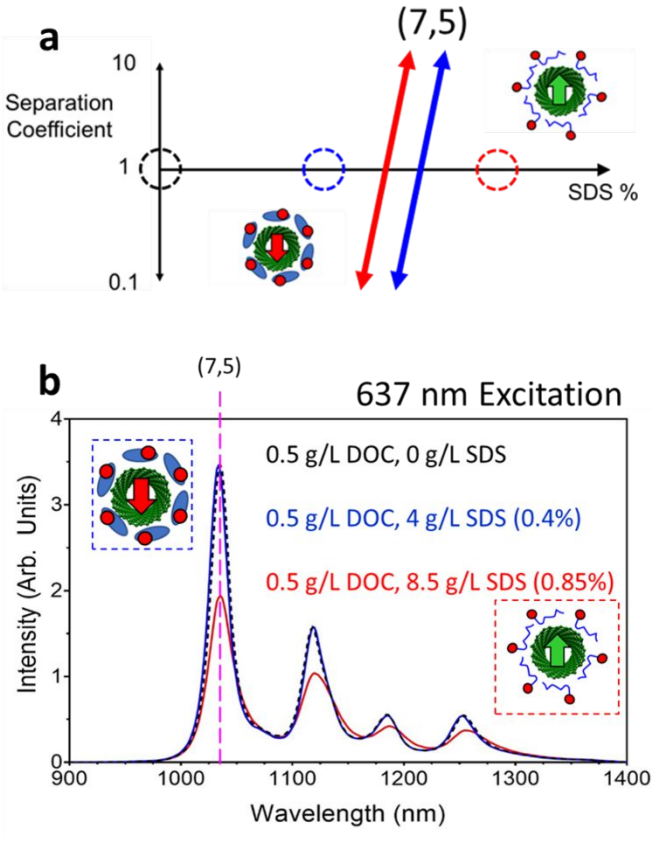

Fig. 4 (a) Diagram of the expected surfactant coating on the (7,5) SWCNT with respect to differing solution conditions. The black circle has no SDS, while the blue circle has an SDS concentration below the empirically known phase transition condition (PCCC value, diagonal lines); both conditions are expected to have DOC coverage as the (7,5) SWCNT would be in the bottom phase of an ATPE mixture under these conditions. The red circle, having an SDS concentration above the empirically known PCCC value, is expected to have SDS coverage as the (7,5) SWCNT is in the ATPE top phase under these conditions. (b) Fluorescence spectra of the SWCNT dispersion at conditions matching panel A. The black curve (no SDS) is overlaid by the blue curve reporting DOC coverage in both scenarios. The strong fluorescence intensity decreases, and mild wavelength shifts seen in the red curve demonstrates a new dielectric environment, where the SWCNTs are now coated by SDS. Adapted from Sims and Fagan, *Carbon*, 2020, **165**, 196-203 with permission from Elsevier Ltd.

We term the combination of surfactant concentrations (or their ratio) that drive bottom-to-top change in phase selection as the "partition coefficient change condition" (PCCC). The PCCC values for the (7,5) SWCNT and its enantiomer are respectfully represented by the value at which the red and blue lines cross the separation coefficient equals one axis in Fig. 4a. While previously identified PCCC values represent a large step towards the overall goal of sorting individual (*n,m*)s from an initial mixture, there are challenges hindering the practical application of species targeting. First, is that although each (*n,m*) and its enantiomer have a specific SDS:DOC partitioning ratio, there are many different SWCNT species (≈ 50 in the sub-1 nm regime) and for multiple combinations of (*n,m*)s these ratios can effectively overlap for a reasonable number of ATPE separation steps. Secondly, identifying the PCCC points by extraction experiments is extremely tedious; it takes weeks to months to measure a set of perhaps 15 values and requires a relatively large amount of SWCNTs for precise measurements. Thirdly, and relating to the first challenge, the surfactant ratios of the PCCC values are also dependant on the presence/concentration of other components making up the two-phase system, which leads to the extensive parameter space for ATPE separations mentioned earlier.

This was our metrology challenge: how to identify PCCC values rapidly and accurately for specific (*n,m*) SWCNT species (and enantiomers) without performing tedious series of extractions at each condition. Feasibly addressing both challenges, and a preference for higher precision resolution of PCCC values than extraction, pointed towards use of a property of the SWCNT-surfactant hybrid that would report on the composition of the adsorbed surfactant layer. The optical properties of SWCNTs are one such potential measurand, as the peak locations and intensity of optical transitions are strongly affected by their surrounding environment due to the size of an exciton exceeding the physical diameter of an individual SWCNT. Changing the composition and packing arrangement of the adsorbed surfactant molecules on the SWCNT surface results in a perturbation of the dielectric environment around the SWCNT;[40] in the case of fluorescing semiconducting SWCNTs, as will be discussed through the remainder of this section, this perturbation yields differences in the emission peak intensity, wavelength, and linewidth of the fluorescence for the same (*n,m*) SWCNT species. We proposed that this modulation could be leveraged to more quickly evaluate the partitioning conditions of (*n,m*) SWCNT species in ATPE by measuring the fluorescence of SWCNT species as a function of surfactant concentrations.[36, 41]

2.4 Fluorescence Detection of ATPE Separation Conditions

To test this hypothesis, we first replicated previous literature results showing that SWCNTs covered by DOC exhibited greater fluorescence intensities than their counterparts covered with SDS.[42, 43] We next conducted test experiments with DOC-SDS mixed surfactant systems with various concentrations (0.5 g/L DOC, variable SDS), simulating those of an actual ATPE preparation. The results of these test experiments are shown in Fig. 4b. The fluorescence spectra of the example (7,5) SWCNT species at a relatively low SDS concentration (4 g/L SDS, Fig. 4b, blue) matches that of the SDS-free sample (Fig. 4b, black). Under these conditions, the SWCNT would partition to the bottom phase of an ATPE separation. Conversely, at a higher SDS concentration (8.5 g/L SDS, Fig. 4b, red) the fluorescence is greatly reduced (≈ 40 % intensity decrease) with slight wavelength shifts, behaving similarly to a DOC-free, SDS-covered sample. At these conditions, the SWCNTs would partition to the top phase of an ATPE separation. Conceptually then, there should be a specific SDS concentration corresponding to the point at which the fluorescence intensity decreases, and we proposed that this would also coincide with the aforementioned PCCC point.

Constructing a series of individual mixed surfactant SWCNT solutions of constant DOC and SWCNT concentration, but increasing SDS concentration, we measured the fluorescence intensity of several (*n,m*)s as functions of their environments, the results of which are shown in Fig. 5a. The fluorescence intensity curve for each (*n,m*) species generally features roughly five distinct regions of behavior with increasing SDS concentrations: 1) an initial region of high fluorescence intensity, consistent with DOC-coverage, followed by 2) a region where the intensity sharply decreases to 3) a small region of relatively consistent intensities of intermediate value, followed by 4) a 2nd region with sharp decreases in intensity ending in 5) a final region of consistently low intensity, consistent with an SDS-dominated surface. The sharp transition between the intensity regions suggests that the dominant adsorbed surfactant undergoes a rapid and discrete exchange (from DOC to SDS) at those values, and thus should also be the PCCC value for that (*n,m*) SWCNT. From Fig. 5a, it is readily apparent that these anticipated PCCC values are quite different for each of the noted (*n,m*)s.

To confirm our initial proposition, we conducted physical separations demonstrating the co-localization of the fluorescence change and extraction phase change. An example of species sorting is shown in Fig. 5b, where an ATPE separation at 1.0 % SDS (Fig. 5a. black line, intermediate of the two species' PCCCs) yields the blue (7,6) in the top phase, leaving the purple (6,5) in the bottom phase. The fluorescence-based method is even sensitive enough to resolve the PCCCs of two enantiomers of a single (*n,m*) species. Enantiomer enrichment is demonstrated in Fig. 5c, where an ATPE separation at ≈ 0.73 % SDS (Fig. 5a red line, intermediate of the (7,5) curve's two sharp transitions) results in a top phase enriched in the negative CD signal (7,5) enantiomer, leaving the positive CD signal (7,5) enantiomer in the bottom phase. This is consistent with the diagram shown in Fig. 4a, where this cut would be between the red and blue diagonal lines.

Having validated the fluorescence methodology for identifying (*n,m*)-specific mixture conditions leading to ATPE phase transitions (PCCC values), we next sought to use the method to evaluate the dependence of various parameters on the ATPE sorting process.[41, 44] These parameters included temperature, polymer concentrations, and surfactant ratios for different variants of the bile salt and alkyl chain surfactant (*e.g.*, tail length). These investigations could be done faster and more quantitatively than before given the ability to precisely control the solution environment of the SWCNTs and that physical separations were no longer required. Many important insights into the governing principles underlying the ATPE method were gleaned from these experiments: 1) quantitatively determining

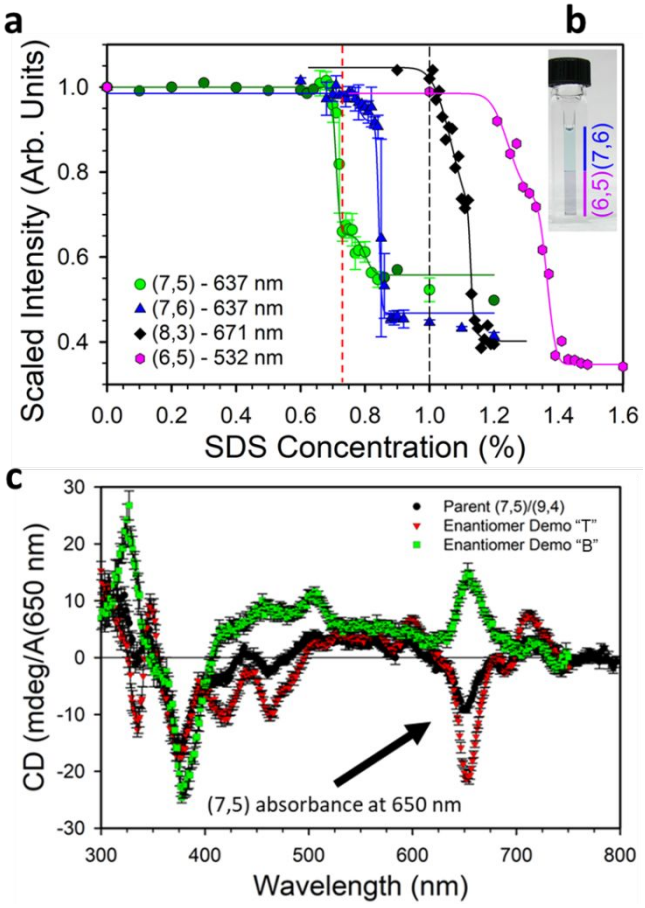

Fig. 5 (a) Scaled fluorescence intensity versus SDS concentration data for (7,5), (7,6), (8,3), and (6,5) SWCNT populations at 20.0 °C, 25 g/L PEG, and a constant DOC concentration of 0.5 g/L. Note the differing curve positions with respect to SDS concentration as well as the behavior of each curve having a two-step decrease in fluorescence intensity. (b) Conducting an ATPE separation at conditions matching the black line in panel A (1.0 % SDS), results in discrimination of the blue (7,6) and purple (6,5) SWCNT species. (c) CD measurements on the parent SWCNT sample and its two daughter fractions separated by an ATPE system at conditions matching the red line in panel A (0.73 % SDS). The parent sample (black circles) mildly enriched with the negative signal (7,5) enantiomer is further enriched in the top-phase daughter fraction ("T", red triangles), leaving the bottom-phase daughter fraction enriched with the positive signal (7,5) enantiomer ("B", green squares). This shows that the two-step decrease behavior in the fluorescence intensity plots are indicative of differential signal from the two enantiomers of the same SWCNT species. Adapted from Sims and Fagan, *Carbon*, 2020, **165**, 196-203 with permission from Elsevier Ltd.

the effect of temperature on ATPE partitioning, supporting empirical observations of (*n,m*) separation challenges at

elevated temperatures; 2) showing that for the common PEG-DEX two-phase system, the concentration of DEX in the bottom phase has little to no effect on the separation, while the PEG concentration in the bottom phase plays a large role and can also be quantified; 3) determination of specific surfactant ratios (SDS:DOC) linked to each (*n,m*) SWCNT enantiomer, which not only guides sorting, but further supports the simple competitive adsorption mechanism in surfactant-ATPE (at minimum for the SDS:DOC pair and supported by our recent work).[37]

We estimate that utilization of fluorescence to identify PCCC values for these experiments resulted in an ≈ 20-fold throughput increase relative to conducting physical separations. However, and despite the significant improvements gained from the fluorescence method, the sheer phase space of potentially viable ATPE systems was still too large to feasibly explore through serial measurements of individually constructed solution populations.

2.5 Automation of Fluorescence-based Condition Detection by Implementation of Gradient Titration

The desire to further increase the rate of solution condition screening, and to additionally reduce the required SWCNT mass, led to the development of an even faster approach by incorporating an automated titration system into our existing fluorescence method (Fig. 6a).[45] In doing so, we were able to increase the measurement throughput and data resolution multi-fold, while also greatly reducing the amount of SWCNT sample needed for experimentation. For example, the black diamond data set in Fig. 6b represents the resolution (roughly 0.01 % SDS) of a typical individual solution experiment set, with the data shown being generated in ≈ 1 h using 20 individually prepared SWCNT samples. Automated titration experiments, such as the blue and red data sets in Fig. 6b, can cover the same concentration range in less than 10 min with over 5 times the data density, each using only the same amount of SWCNT solution as a single individually constructed data point. For what we call a ”forward” experiment, the sample fluorescence (initially DOC covered) starts out high, gradually diminishes with dilution by the added titrant, then suddenly steps to a lower intensity value across a short range of concentration centered at the PCCC value. A “reverse” experiment, in contrast, starts at a greater SDS percentage above the PCCC value (low fluorescence intensity) and by dilution of the titrant reduces the SDS concentration below the PCCC value, inducing a fluorescence intensity increase by DOC coverage. Note that since the initial volume and addition rate are known, the dilution follows a simple algebraic equation, which we correct for before data presentation (*e.g.*, in Fig 6b). For robustness, in our current implementation we dilute an initial volume of all components (*e.g.*, DOC, SWCNT, SDS, PEG) by an equal volume of liquid containing only one component at a different concentration. Successful continuous mixing and temperature control are key features of the setup, and concentration gradients (excepting the SWCNTs which are always diluted by a known degree as a function of time) can be either positive or negative. A key observation using this functionality is also shown in Fig. 6b, in which identical, within resolution, PCCC values are observed for both (7,5) enantiomers and fully reversible.

This gradient titration method brings vastly (and arbitrarily) increased resolution (through control of the titration rate), enabling us to more accurately and precisely quantify the SWCNTs’ PCCC values compared to the original method. Encouragingly, we also found that slow experiments were not necessary for accurate and precise measurements. We were able to increase the measurement and titration rates and decrease the time of a single sweep to as little as 8 min while still collecting hundreds of data points and maintaining consistent results. Such a finding showcases the robustness of the method and the potential for extension to accurately investigate fast kinetic effects. Even at normal titration rates, the increased resolution in the surfactant ratios is exposing previously unseen subtle fluorescence intensity and wavelength changes that occur during the surfactant exchange process within the PCCC slope. Although not yet fully understood, these findings open the door to improved probing and understanding of the physicochemical interactions and mechanisms that

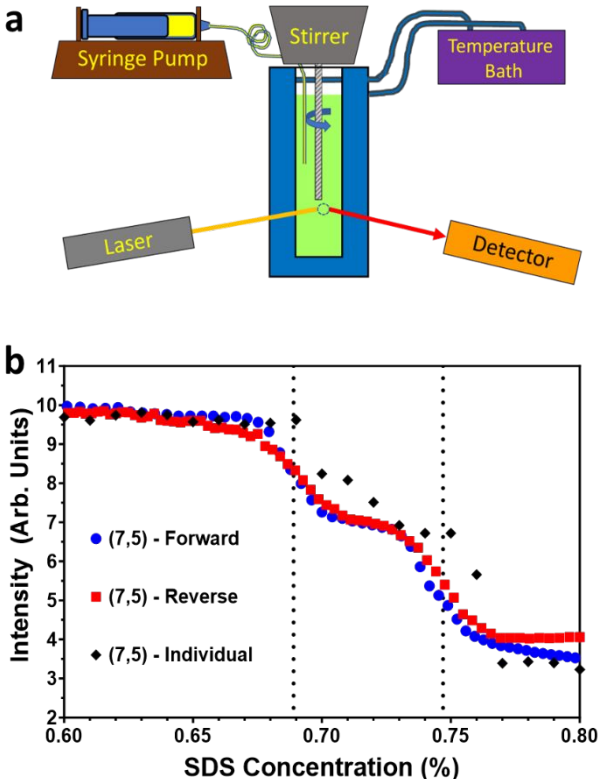

Fig. 6 (a) Schematic of the automated gradient titration setup housed within a fluorescence spectrophotometer. A stirrer provides rapid mixing of the SWCNT solution during titrant addition, while a temperature bath sets and maintains a fixed solution temperature. Fluorescence is measured in a front face collection geometry. (b) Fluorescence intensity of the (7,5) SWCNT for the two experimental method variations as a function of the SDS concentration. Comparing the results of the individually prepared solutions (black diamonds) to the automated gradient titration approach (blue circles or red squares), the vast increase in data density and hence, resolution is clear. Of note is that identical PCCC values (black dashed lines) are obtained for both (7,5) enantiomers from both forward (blue circles) and reverse (red squares) titrations, showing that the surfactant exchange process is thermodynamically and kinetically independent of direction. Adapted from Sims and Fagan, *Carbon*, 2024, **219**, 118813 with permission from Elsevier Ltd.

underlie the ATPE separation method. Altogether, the measurement improvements advanced by the automated gradient titration fluorescence methodology have massive potential for accelerating SWCNT separations design and facilitating SWCNT technologies development.[45]

3. Method Deployment Opportunities for Enhancing SWCNT Technology Development

The developed methods described above illustrate the promise of high-throughput approaches for determining SWCNT sorting conditions, rapidly exploring the exceptionally large parameter space of ATPE. We now offer our perspective of further advancements enabled by these methodologies and the potential future impact on SWCNT technology development.

3.1 Future Development Needs for DNA-based SWCNT Sorting

3.1.1 Developing New Tools to Study DNA-SWCNT Hybrids There are two fundamental issues in the study of DNA-SWCNT hybrids, the first being sequence selection. As described earlier, ordered DNA-SWCNT structures and corresponding resolving sequences had previously been identified *via* a time-consuming trial-and-error SWCNT sorting process,[2, 24, 27] but emerging new methods are changing this situation. As reviewed in section 2, our recent work shows that ML can be a useful tool to identify resolving DNA sequences.[29, 34] In parallel, multiple studies have shown that helicity- and handedness-dependent DNA coating structures can be probed by small molecules using fluorescence spectroscopy.[46-48] Further instrumentation and methodology development in this direction is expected to solidify the mechanistic connection between sorting and sensing, and establish more efficient experimental methods to identify resolving sequences.

The second issue is DNA-SWCNT structure determination, especially for specific DNA-(*n,m*) enantiomer combinations that are known to be resolving in ATPE. There has been a lack of methods to observe the coating on a nanotube at atomic resolution. In this case, the rapidly developing and increasingly accessible single-particle cryo-electron microscopy (cryo-EM) may address this need.[49] This method appears to be especially suited for quasi-1D noncovalent assemblies that may lack long-range order. The stunning structure of a conductive protein nanowire recently determined by cryo-EM is an inspiring example.[50] It is quite likely that atomic details of certain DNA-SWCNTs will soon be revealed.[51] Taken together, solutions to these fundamental issues will bring us closer to the future goal of *de novo* design of DNA-SWCNTs.

3.1.2 What Lies Ahead for DNA-ATPE Sorting? Despite what has been demonstrated thus far for DNA-based SWCNT sorting, there is still enormous room for further improvement. Listed below are two specific issues that can be addressed in the near future:

1. Investigators should be able to identify (a) more "resolving" sequences that resemble super sequence (6,5) or super sequence (8,3), which allow robust sorting across several two-phase systems, and (b) more sequences that can resolve more than one (hopefully many) chiral species in a synthetic mixture.
2. We need to understand why only certain two-phase systems work for DNA-SWCNT sorting. Unlike surfactant-coated SWCNTs, DNA-SWCNTs are stable in many two-phase systems *(i.e.*, comprised of mixtures of different salts and polymers), only a subset of which afford sorting. For instance, there are many PEG/salt systems, but very few can support DNA-SWCNT sorting. It is possible that some salts make the water phase more structured, which results in a better-defined solvation energy for a DNA-SWCNT hybrid, and a better sorting outcome. Understanding of the underlying mechanism of DNA-based ATPE sorting will surely lead to better outcomes in separation.

We envision that someday we will be able to resolve each and every single (*n,m*) SWCNT species in any given synthetic mixture through a single separation run on an automated machine [*e.g.*, counter current chromatography (CCC)[52]]. The DNA sequence to be used for separation could be predicted by a ML algorithm that is built on solved structures of DNA-SWCNTs, and quantitative understanding of solvation energies of DNA-SWCNTs in polymer- and salt-modified aqueous phases. Finally, commercialization of this technology will make SWCNTs with defined handedness, helicity, and length a readily accessible material for SWCNT technology developments.

3.2 Automated Exploration of Surfactant-SWCNT Interactions

As noted earlier, there are many potential directions for changing or modifying the surfactant-ATPE system to improve separations. Variables also include details of the polymers and temperature, but we are particularly interested in different combinations (3+) of surfactant(s) and inclusion of new components such as electrolytes or oxidants.

3.2.1 3+ Surfactant Competition Both published works and our ongoing efforts have identified variations in the surfactant-ATPE method that can be used to separate (*n,m*) species with similar SDS:DOC PCCC values by adding an additional component, such as SC, and performing another ATPE separation with different selectivity. One example of this identified by gradient titration is the addition of 0.9 % SC to a PEG:DEX ATPE system with 0.05 % DOC and 0.65 % SDS as a means to separate (8,4) from (7,5).[45] The gradient titration fluorescence methodology is efficient enough to enable phase space searches for such differentiated conditions as a function of the 3rd component concentration, particularly with a shift to performing titration on samples containing many SWCNT species and deconvolution of the obtained spectra.[45] Given identification of sufficient numbers of PCCC points for different (*n,m*) and surfactant combinations, identification of new combinations towards optimization of a multistage cascade may even be a feasible target for ML.

3.2.2 Other Additives Modulation of surfactant-ATPE can alternatively be achieved by adding new chemical components, rather than just additional surfactant types, to the initial mixture.[53] A straightforward example from our work is the addition of NaCl at low molar concentrations, which we found to reduce the ratio of SDS to DOC needed to reach PCCC values for monitored (*n,m*)s, but maintained or increased the difference between PCCC values of different SWCNT species.[45]

The ubiquity of NaCl as a potential contaminant to stock polymers or surfactants may even be the source of some observed variability in PCCC values across different labs, *i.e.,* due to such (unintentional) additives. However, even simple salts have significantly different effects as cations or anions to surfactants (for instance, potassium deoxycholate is much less soluble than DOC) or Hofmeister series-style effects on their own. Exploring such effects may improve repeatability, adjust separations effectively, or simply be economical by reducing surfactant use in favour of a less expensive salt.

3.2.3 Measuring Thermodynamics of Bile Salt Adsorption Another way in which titration experiments hold promise is in the measurement of polymer free dispersions of SWCNTs during the competition of surfactants for adsorption to the nanotube interface. As mentioned above, to accurately determine the PCCC values for ATPE extractions by titration experiments, an appropriate concentration of PEG (for PEG:DEX ATPE) consistent with the concentration that would occur in the DEX phase of a separation was required. Addition of this amount of PEG increases the SDS:DOC ratio at which the PCCC points occur, and widens differences in absolute concentration between (*n,m*)s and enantiomers of the same (*n,m*). However, measurements without PEG present should yield adsorption information, as interpreted from fluorescence data, that is simpler to relate to thermodynamic models, either in manners addressable as single surfactant binding determination[54] or multi-dispersant competition. Gradient titration is well suited for such efforts because it can achieve and probe arbitrarily small increments of concentration variation simply by the setting of the concentration(s) with the titrant volume. The precise data being obtained is anticipated to be sufficient to enable modelling of such systems.[55, 56]

3.2.4 What Lies Ahead for Surfactant-ATPE? Looking to the future of surfactant-ATPE metrology, it is likely that a significant number of new and precise data sets of PCCC values for different surfactant competitions and additives will be published in the near future. An obvious path, similar to the DNA-ATPE method, would be to apply ML to estimating PCCC values for non-measured (*n,m*)s at a given set of surfactant compositions, or to project effects of new additives from sparse(r) initial data sets. As the phase space of additive chemistries is large, this could be a particularly effective, and *in silico*, route for effective ATPE optimization towards key (*n,m*) populations.

Beyond the level of single (*n,m*) separation, biosensing and single photon emission applications also are likely to require single enantiomer SWCNT populations. This is due to the differential interaction of the left- and right-hand twist SWCNTs with handed environments. It cannot be assumed that one hand of an (*n,m*) will always extract first, as reversal of extraction order under different surfactant competitions has been observed, and parameters such as relative fluorescence efficiencies for surfactant-(*n,m*) hand pairs can be different.[57] Currently we can only learn which enantiomer of the (*n,m*) a PCCC value represents by offline performance of CD, either by preparing and measuring a sorted population, or by measuring a known single enantiomer sample. A measurement solution that may overcome this restriction is to perform fluorescence-detected CD on titrated or directly constructed samples. Initial efforts towards this goal are underway.

One metrology development that we have not yet significantly advanced is in screening the PCCC values for metallic SWCNTs. Determining the partition of these (*n,m*) species is still limited to conducting ATPE separations and determining the concentration of a species *via* absorbance or Raman spectroscopy.[18, 38] Progress in separation-free determination of values should enable improved technology development as has occurred for semiconducting species.

3.3 How Metrology Development Will Impact Scale-Up and Commercialization

Given the desire for single species SWCNT populations for commercial development, scale-up is an obvious area for technological improvement and demonstration. Such scale-up could be based on either efficiency improvements or direct implementation of larger volume processing equipment, but would also benefit by benchmarks for cost per mg of (*n,m*) produced and identification of potential points for improvement in robustness and circularity. Any improvements in SWCNT starting material quality such as improvement in tailoring synthesized distributions for desired (*n,m*)s, increased fractions of carbon as high crystallinity SWCNT, and reduced or easily removed catalyst contents, would also offer significant value.

From the metrology side we believe that several pathways are promising for improving the separation efficiency of ATPE. In batch separations, one potential path to improvement is the identification of more optimal series of ATPE conditions to rapidly isolate specific SWCNT species. Such identification, as noted above, is a primary driver of the gradient titration metrology. Implementation of CCC, a technique that enables application of hundreds of theoretical plates of ATPE separations to a mixture of SWCNTs automatically, is another promising route. Although previously applied at small scale for surfactant-ATPE, the application of this technique has been limited by equipment difficulties in handling the common 6 kDa PEG-DEX 70 ATPE system in terms of column stationary phase retention, temperature regulation, and sub-optimal observed colloidal stability. Implementation and extension of CCC for other ATPE systems, however, holds promise. Reduced viscosity, higher temperature stability, and faster phase-separation kinetics of many polymer-salt systems are particularly attractive, but PCCC values will need to be re-determined or translated from current data sets.

Another way metrology should directly impact development and scale-up is in determining causes of remaining variability in ATPE separations, and whether this variability arises from differences in synthesis methods, synthesis implementations by manufacturers, lot number variations, or other physical characteristics. For actual separations we have observed that limiting ATPE to purer, or at least more homogeneous, SWCNT populations improves the yield and purity of resulting populations. Examples include rate-zonal ultracentrifugation

processed populations,[58] alkane-filled populations,[59] and polymer depletion-based length sorting (PDLS) populations. The PDLS approach, previously shown to be compatible with and improve ATPE results,[60] was recently expanded in demonstration to include a wider set of SWCNT source materials, and to be useable to remove undesired SWCNT (and some non-SWCNT) subpopulations *via* a low energy and simple process. Integrating an impurity removing (and concentrating) procedure in advance directly reduces the ATPE volume needed and improves selection efficiency by removing non-desired possible contaminates.[61] Measurement of PCCC values for even a single (*n,m*) systematically from a broad range of parent sources or of varied characteristics has not been significantly reported. A limitation of such measurements is that they would address the unknown inter-batch and source polydispersity, but leave polydispersity within a batch to be inferred from PCCC widths. Conducting comparative fluorescence measurements for PCCC values by counting the response of single nanotubes in a hyperspectral microscope is a possible route for determining within batch polydispersity,[62] as well as potentially the range of differences in fluorescence efficiencies for interactions with handed surfactants and SWCNT enantiomers.

Perhaps most importantly, the use of surfactant-based and DNA-based separations in concert may reduce cost and improve (*n,m*) selection by enabling specific targeting of desired populations *via* DNA-ATPE on pre-processed surfactant method populations. Metrology is useful here because of the identification of more optimal cascades and DNA/surfactant exchange conditions through the large sets of PCCC values and DNA sequences.[63, 64] Such a route, enabled by surfactant-to-DNA dispersion exchange, is enabled by the same technology utilized to control the adsorbed DNA dispersant layer for SWCNT biosensors and controlled specific placement *via* DNA-hybridization for building electronic devices.

Lastly, several of the above metrology advances are needed towards solving the measurement challenge of assessing very high purity SWCNT (*n,m*) populations desired for electronics and photonic applications. Potential methods such as photothermal probing[65] and hyperspectral microscopy[62] need development and validation supported by well-defined separated populations, preferably with tracked production characteristics and universal availability to support direct interlaboratory comparison of results. Such a need argues for a new generation of SWCNT reference materials (RM); production and specifications for such RM are under active consideration.

4. Conclusion

ATPE, whether surfactant-based or DNA-directed, is an increasingly powerful tool for separating specific SWCNT (*n,m*) populations from polydisperse synthetic sources. Used in various workflows, the ability to isolate extremely high purity populations to the level of a single (*n,m*) enantiomer, and in combination with other tools to isolate specific length distributions, attach particular functional groups, and direct assembly have all been demonstrated. In this contribution we described how high-throughput metrology advances and the generation of data sets on ATPE selectivity have developed, and hypothesize about how these advances will be utilized for further progress. We anticipate that future work will expand the number of specific ATPE systems available for tailored separations, provide foundations for insight and control of adsorbed surfactant or DNA layers on the SWCNT surface, and facilitate the translation from a bench scale to commercialized technology.

Author Contributions

Christopher M. Sims: Writing – Original Draft; Writing – Review & Editing. **Ming Zheng:** Writing – Original Draft; Writing – Review & Editing. **Jeffrey A. Fagan:** Writing – Original Draft; Writing – Review & Editing

Data Availability

No primary research results, software or code have been included and no new data were generated or analysed as part of this feature article.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

All authors are funded through internal National Institute of Science and Technology funds.

Notes and references

‡ Certain equipment, instruments, software, or materials, commercial or noncommercial, may be identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement of any product or service by NIST, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

§ Although no new experimental data was generated in this feature article, uncertainties are presented as reported as in their original source publications.

1. M. Zheng, A. Jagota, E. D. Semke, B. A. Diner, R. S. McLean, S. R. Lustig, R. E. Richardson and N. G. Tassi, *Nat. Mater.*, 2003, **2**, 338-342.
2. M. Zheng, A. Jagota, M. S. Strano, A. P. Santos, P. Barone, S. G. Chou, B. A. Diner, M. S. Dresselhaus, R. S. McLean, G. B. Onoa, G. G. Samsonidze, E. D. Semke, M. Usrey and D. J. Walls, *Science*, 2003, **302**, 1545-1548.
3. S. R. Lustig, A. Jagota, C. Khripin and M. Zheng, *J. Phys. Chem. B*, 2005, **109**, 2559-2566.

4. M. Zheng and E. D. Semke, *J. Am. Chem. Soc.*, 2007, **129**, 6084-6085.
5. M. S. Arnold, S. I. Stupp and M. C. Hersam, *Nano Lett.*, 2005, **5**, 713-718.
6. M. S. Arnold, A. A. Green, J. F. Hulvat, S. I. Stupp and M. C. Hersam, *Nat. Nanotechnol.*, 2006, **1**, 60-65.
7. S. Ghosh, S. M. Bachilo and R. B. Weisman, *Nat. Nanotechnol.*, 2010, **5**, 443-450.
8. A. A. Green and M. C. Hersam, *Adv. Mater.*, 2011, **23**, 2185-2190.
9. K. Ihara, H. Endoh, T. Saito and F. Nihey, *J. Phys. Chem. C*, 2011, **115**, 22827-22832.
10. Y. Kuwahara, F. Sasaki and T. Saito, *J. Phys. Chem. C*, 2019, **123**, 3829-3835.
11. T. Tanaka, H. Jin, Y. Miyata, S. Fujii, H. Suga, Y. Naitoh, T. Minari, T. Miyadera, K. Tsukagoshi and H. Kataura, *Nano Lett.*, 2009, **9**, 1497-1500.
12. H. Liu, D. Nishide, T. Tanaka and H. Kataura, *Nat. Commun.*, 2011, **2**, 309.
13. H. Liu, T. Tanaka, Y. Urabe and H. Kataura, *Nano Lett.*, 2013, **13**, 1996-2003.
14. H. Liu, T. Tanaka and H. Kataura, *Nano Lett.*, 2014, **14**, 6237-6243.
15. C. Y. Khripin, J. A. Fagan and M. Zheng, *J. Am. Chem. Soc.*, 2013, **135**, 6822-6825.
16. G. Ao, C. Y. Khripin and M. Zheng, *J. Am. Chem. Soc.*, 2014, **136**, 10383-10392.
17. J. A. Fagan, C. Y. Khripin, C. A. Silvera Batista, J. R. Simpson, E. H. Haroz, A. R. Hight Walker and M. Zheng, *Adv. Mater.*, 2014, **26**, 2800-2804.
18. J. A. Fagan, E. H. Haroz, R. Ihly, H. Gui, J. L. Blackburn, J. R. Simpson, S. Lam, A. R. Hight Walker, S. K. Doorn and M. Zheng, *ACS Nano*, 2015, **9**, 5377-5390.
19. D. Janas, *Mater. Chem. Front.*, 2018, **2**, 36-63.
20. F. Yang, M. Wang, D. Zhang, J. Yang, M. Zheng and Y. Li, *Chem. Rev.*, 2020, **120**, 2693-2758.
21. X. Wei, S. Li, W. Wang, X. Zhang, W. Zhou, S. Xie and H. Liu, *Adv. Sci.*, 2022, **9**, 2200054.
22. P.-Å. Albertsson, in *Adv. Protein Chem.*, eds. C. B. Anfinsen, J. T. Edsall and F. M. Richards, Academic Press, 1970, vol. 24, pp. 309-341.
23. B. Y. Zaslavsky, *Aqueous Two-Phase Partitioning: Physical Chemistry and Bioanalytical Applications*, Marcel Dekker, Inc., New York, Basel, Oxford, 1995.
24. X. Tu, S. Manohar, A. Jagota and M. Zheng, *Nature*, 2009, **460**, 250-253.
25. D. Roxbury, X. Tu, M. Zheng and A. Jagota, *Langmuir*, 2011, **27**, 8282-8293.
26. X. Tu, A. R. Hight Walker, C. Y. Khripin and M. Zheng, *Journal of the American Chemical Society*, 2011, **133**, 12998-13001.
27. G. Ao, J. K. Streit, J. A. Fagan and M. Zheng, *J. Am. Chem. Soc.*, 2016, **138**, 16677-16685.
28. M. Zheng, *Top. Curr. Chem.*, 2017, **375**, 13.
29. Y. Yang, M. Zheng and A. Jagota, *npj Comput. Mater.*, 2019, **5**, 3.
30. D. R. Cox, *J. R. Stat. Soc. Series B Stat. Methodol.*, 1958, **20**, 215-242.
31. F. Murtagh, *Neurocomputing*, 1991, **2**, 183-197.
32. C. Vens, M.-N. Rosso and E. G. J. Danchin, *Bioinformatics*, 2011, **27**, 1231-1238.
33. A. Shankar, J. Mittal and A. Jagota, *Langmuir*, 2014, **30**, 3176-3183.
34. Z. Lin, Y. Yang, A. Jagota and M. Zheng, *ACS Nano*, 2022, **16**, 4705-4713.
35. W. Wenseleers, I. I. Vlasov, E. Goovaerts, E. D. Obraztsova, A. S. Lobach and A. Bouwen, *Adv. Funct. Mater.*, 2004, **14**, 1105-1112.
36. J. A. Fagan, *Nanoscale Adv.*, 2019, **1**, 3307-3324.
37. C. M. Sims and J. A. Fagan, *J. Phys. Chem. C*, 2024, **128**, 13064-13073.
38. J. Defillet, M. Avramenko, M. Martinati, M. Á. López Carrillo, D. Van der Elst, W. Wenseleers and S. Cambré, *Carbon*, 2022, **195**, 349-363.
39. M. Avramenko, J. Defillet, M. A. Lopez Carrillo, M. Martinati, W. Wenseleers and S. Cambre, *Nanoscale*, 2022, **14**, 15484-15497.
40. J. G. Duque, L. Oudjedi, J. J. Crochet, S. Tretiak, B. Lounis, S. K. Doorn and L. Cognet, *J. Am. Chem. Soc.*, 2013, **135**, 3379-3382.
41. C. M. Sims and J. A. Fagan, *Carbon*, 2020, **165**, 196-203.
42. R. Haggenmueller, S. S. Rahatekar, J. A. Fagan, J. Chun, M. L. Becker, R. R. Naik, T. Krauss, L. Carlson, J. F. Kadla, P. C. Trulove, D. F. Fox, H. C. Delong, Z. Fang, S. O. Kelley and J. W. Gilman, *Langmuir*, 2008, **24**, 5070-5078.
43. J. G. Duque, M. Pasquali, L. Cognet and B. Lounis, *ACS Nano*, 2009, **3**, 2153-2156.
44. C. M. Sims and J. A. Fagan, *Carbon*, 2022, **191**, 215-226.
45. C. M. Sims and J. A. Fagan, *Carbon*, 2024, **219**, 118813.
46. Y. Zheng, A. A. Alizadehmojarad, S. M. Bachilo, A. B. Kolomeisky and R. B. Weisman, *ACS Nano*, 2020, **14**, 12148-12158.
47. Y. Zheng, S. M. Bachilo and R. B. Weisman, *J. Phys. Chem. Lett.*, 2018, **9**, 3793-3797.
48. Y. Zheng, S. M. Bachilo and R. B. Weisman, *J. Phys. Chem. Lett.*, 2017, **8**, 1952-1955.
49. Y. Cheng, *Science*, 2018, **361**, 876-880.
50. F. Wang, Y. Gu, J. P. O'Brien, S. M. Yi, S. E. Yalcin, V. Srikanth, C. Shen, D. Vu, N. L. Ing, A. I. Hochbaum, E. H. Egelman and N. S. Malvankar, *Cell*, 2019, **177**, 361-369.e310.
51. Z. Lin, L. C. Beltrán, Z. A. De los Santos, Y. Li, T. Adel, J. A. Fagan, A. R. Hight Walker, E. H. Egelman and M. Zheng, *Science*, 2022, **377**, 535-539.
52. M. Zhang, C. Y. Khripin, J. A. Fagan, P. McPhie, Y. Ito and M. Zheng, *Anal. Chem.*, 2014, **86**, 3980-3984.
53. B. Podlesny, T. Shiraki and D. Janas, *Sci. Rep.*, 2020, **10**, 9250.
54. F. F. Bergler, S. Stahl, A. Goy, F. Schoppler and T. Hertel, *Langmuir*, 2016, **32**, 9598-9603.
55. M. Park, J. Park, J. Lee and S. Y. Ju, *Carbon*, 2018, **139**, 427-436.
56. S. Hwang, S. Son, M. Park, I.-S. Choi and S.-Y. Ju, *Small Science*, 2024, **4**, 2400011.
57. H. Li, C. M. Sims, R. Kang, F. Biedermann, J. A. Fagan and B. S. Flavel, *Carbon*, 2023, **204**, 475-483.
58. J. A. Fagan, J. Y. Huh, J. R. Simpson, J. L. Blackburn, J. M. Holt, B. A. Larsen and A. R. Walker, *ACS Nano*, 2011, **5**, 3943-3953.
59. H. Li, G. Gordeev, O. Garrity, N. A. Peyyety, P. B. Selvasundaram, S. Dehm, R. Krupke, S. Cambre, W. Wenseleers, S. Reich, M. Zheng, J. A. Fagan and B. S. Flavel, *ACS Nano*, 2020, **14**, 948-963.

60. H. Gui, H. Chen, C. Y. Khripin, B. Liu, J. A. Fagan, C. Zhou and M. Zheng, *Nanoscale*, 2016, **8**, 3467-3473.
61. P. Shapturenka, B. K. Barnes, E. Mansfield, M. M. Noor and J. A. Fagan, *RSC Adv.*, 2024, **14**, 25490-25506.
62. M. Erkens, W. Wenseleers, M. Á. López Carrillo, B. Botka, Z. Zahiri, J. G. Duque and S. Cambré, *ACS Nano*, 2024, **18**, 14532-14545.
63. J. K. Streit, J. A. Fagan and M. Zheng, *Anal. Chem.*, 2017, **89**, 10496-10503.
64. Y. Yang, A. Sharma, G. Noetinger, M. Zheng and A. Jagota, *J. Phys. Chem. C*, 2020, **124**, 9045-9055.
65. J. Miyazaki and H. Maeda, *Jpn. J. Appl. Phys.*, 2023, **62**, 102003.

Data Availability Statement for

Single-wall Carbon Nanotube Separations via Aqueous Two-phase Extraction: New Prospects Enabled by High-throughput Methods

Christopher M. Sims, Ming Zheng, and Jeffrey A. Fagan

Materials Science and Engineering Division
National Institute of Standards and Technology, Gaithersburg, MD USA 20899

No primary research results, software or code have been included and no new data were generated or analysed as part of this feature article.