

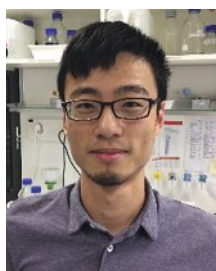
Analytical Approaches for Monitoring DNA–Protein Interactions

Shang-Jung Wu[§] and Ardemis A. Boghossian*

[§]SCS-DSM Award for best poster presentation in Analytical Sciences

Abstract: DNA–protein interactions play a critical role in cellular regulation. We herein review existing analytical methods for investigating these interactions, highlighting methods such as chromatin immunoprecipitation and yeast-one-hybrid that are used to identify undiscovered DNA–protein interactions. We summarize the most common approaches for characterizing known interactions based on DNA–protein structure, thermodynamic and kinetic measurements, and dynamic binding assays. We discuss techniques in optical imaging as well as representative methods, such as electrophoretic mobility shift assay and surface plasmon resonance. The advantages and disadvantages of these techniques are used to assess a proposed optical platform based on single-walled carbon nanotube (SWCNT) fluorescence.

Keywords: Biosensor · Carbon nanotube · DNA–protein interaction · Photoluminescence



Shang-Jung Wu was born in Taiwan. He studied Material Science and Engineering at the National Taiwan University. He obtained his Master degree under the supervision of Prof. Wei-Fang Su, with his thesis focusing on controlling block copolymer morphology. He is now a PhD candidate at École Polytechnique Fédérale de Lausanne (EPFL) in the Laboratory of Nanobiotechnology (LNB) supervised by Prof. Ardemis

Boghossian. His current research focuses on functionalizing single-walled carbon nanotubes for biosensing applications.

DNA–protein interactions lie at the core of essential biological processes such as replication, transcription, and DNA repair. Mechanistic knowledge of these interactions offers insight into cell regulation, which holds the key to unlocking therapeutic targets for diseases such as cancer. As such, several analytical approaches for monitoring DNA–protein interactions have been developed. One common approach for identifying DNA sequence(s) that bind a particular protein-of-interest (POI) is based on a pull-down assay. Pull-down assays, such as chromatin immunoprecipitation (ChIP, Fig. 1a),^[1] can identify the specific region(s) of the genome that associates with a POI. In this assay, DNA-bound proteins are sheared into shorter DNA fragments by either sonication or digestion, and the POI, along with its associated DNA, are selectively immunoprecipitated from the mixture using a protein-specific antibody. The captured DNA sequence(s) is subsequently analyzed in a high-throughput manner by either DNA microarray (ChIP-on-chip) or by next-generation sequencing (ChIP-seq). Alternatively, systematic evolution of ligands by exponential enrichment (SELEX)^[2] screens a DNA sequence library for sequences that preferentially bind with a POI. Bound sequences are eluted and amplified for additional rounds of screening with increasingly stringent elution requirements. After iterative selection, the DNA sequence(s) with the strongest binding affinity is retained and identified.

In contrast to ChIP or SELEX, which aim to identify the DNA target(s) of a POI, alternative strategies focus on identifying the particular protein(s) that binds to a specific DNA sequence of interest. Yeast-one-hybrid (Y1H, Fig. 1b) fuses a protein library (referred to as the ‘prey’ protein) to the activation domain of a transcription factor.^[3] Once a protein specifically binds to the sequence of interest (referred to as the ‘bait’ DNA) that is located upstream of a gene for a reporter protein (*e.g.* luciferase), the binding brings the transcription factor in close proximity to the promoter element of the gene, activating transcription and triggering protein expression. The reporter protein expression can thus be used to signal protein binding to a given DNA sequence. Alternatively, a non-specific assay can be followed by 2D gel electrophoresis or mass spectroscopy (MS) to identify the protein(s) interacting with a specific DNA.^[1]

Once a DNA–protein complex has been identified, it can be further studied using the analytical tools listed in Table 1. In the following sections, we will discuss some of these methods in greater detail, highlighting their advantages as well as limitations. We primarily focus on characterization of the binding event, whereas methods for studying enzymatic interactions between DNA and proteins are discussed elsewhere.^[1]

Characterizing Known DNA–Protein Interactions

Common characterization techniques focus on obtaining information on the bound DNA–protein structure, thermodynamic and kinetic parameters, and spatiotemporal mapping of binding events. The structure of the DNA–protein complex can be used to elucidate the nature of the dominant binding force, such as electrostatic and/or hydrophobic interactions. Crystal structures of purified complexes can be obtained using X-ray diffraction (XRD, Fig. 1c), nuclear magnetic resonance (NMR),^[4] or cryogenic electron microscopy (cryo-EM). The DNA-binding domain of the protein can be mutated and crosslinked to covalently attach the DNA to the interacting residues. The crosslinked complex can be analyzed through MS to map the proximity of cer-

*Correspondence: Prof. A. A. Boghossian

E-mail: ardemis.boghossian@epfl.ch

Institute of Chemical Sciences and Engineering (ISIC), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne CH-1015

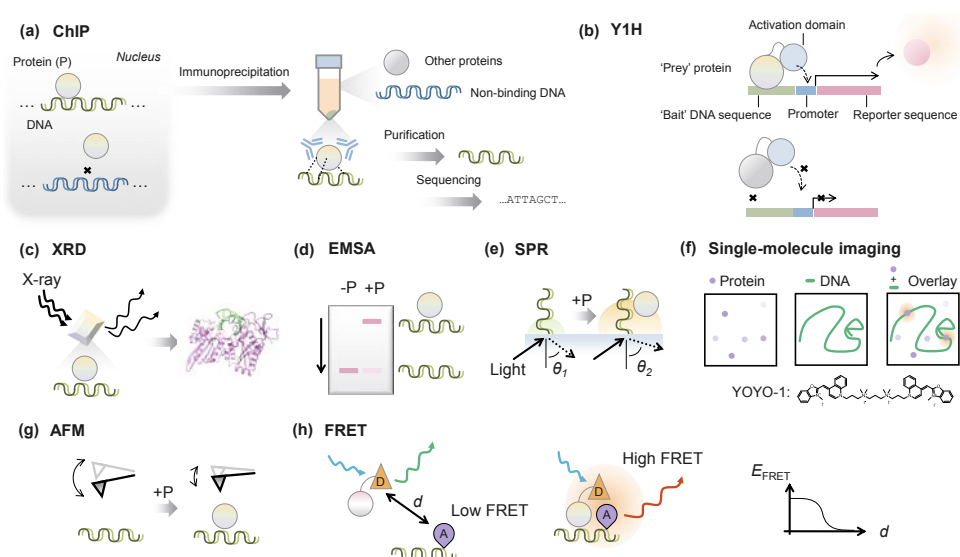


Fig. 1. Working principle of different methods for studying DNA-protein interactions. (a) Chromatin immunoprecipitation (ChIP) is used to identify the sequence(s) that bind to a certain protein (P), while (b) yeast-one-hybrid (Y1H) conversely identifies proteins that bind to a certain 'bait' DNA sequence. (c) X-ray diffraction (XRD) determines the crystal structure of complex. (d) Electrophoretic mobility shift assay (EMSA) can be used to determine thermodynamic parameters such as K_d , while (e) surface plasmon resonance (SPR) can further determine kinetic parameters such as k_{on} and k_{off} . (f) Single-molecule imaging and (g) atomic force microscopy (AFM) investigate binding through imaging. (h) Förster resonance energy transfer (FRET) is used for both sensing and imaging applications *in vitro* and *in vivo*.

Table 1. Summary of common methods for studying DNA-protein interactions.

Method	Purpose				
	Structure	Binding affinity	Time-resolved	<i>in vitro</i> sensing	<i>in vitro</i> imaging
X-ray diffraction	X				
Nuclear magnetic resonance	X				
Cryogenic electron microscope	X				
Cross-linked mass spectrometry	(X)	X		X	
Filter-binding assay		X	Limited	X	
Electrophoretic mobility shift assay		X	Limited	X	
DNA footprint		X		X	
Southwestern blotting		X		X	
Fluorescence anisotropy		X		X	
Microscale thermophoresis		X		X	
Isothermal calorimetry		X		X	
Quartz crystal microbalance		X		X	
Analytical ultracentrifuge		X		X	
Field effect transistor		X	X	X	
Electrochemical method		X	X	X	
Surface plasmon resonance		X	X	X	
Protein induced fluorescence enhancement		X	X	X	X
Single-molecule imaging		X	X	X	X
Atomic force microscope		X	X		X
FRET-based		X	X	X	X

tain residues^[5] and identify contributions from post-translational modifications.

The binding interaction is typically quantified by the thermodynamic (*e.g.* dissociation constant, K_d) and kinetic (*e.g.* binding on/off rate, k_{on} and k_{off}) parameters of the system. The K_d can be determined using an electrophoretic gel-based method.^[6] For example, in the electrophoretic mobility shift assay (EMSA, Fig. 1d), electrophoretic movement of DNA is compared between free DNA (in the absence of protein) and DNA in the presence of the binding protein. Since DNA movement is hindered when bound to a protein(s), the bound DNA shows a distinct mobility band compared to the free DNA, and the ratio of the two bands can be used to calculate the apparent affinity of the DNA–protein complex when the stoichiometry and starting concentrations of the DNA and protein(s) are known.^[6a] In addition, the DNA–protein complex can be separated from free DNA and visualized with radiolabelled reagents, and the K_d can be determined through titration. Surface plasmon resonance (SPR, Fig. 1e) is an alternative, label-free method for quantifying affinity using optical measurements.^[7] In SPR, a metallic surface is modified with DNA, and the refractive index of the surface becomes altered when protein binds to the DNA. This binding modifies the surface plasmon and evanescent field, which can be optically probed using a total-reflection configuration. Unlike gel-based methods, SPR can not only be used to determine K_d , but it can also be used for real-time analysis (k_{on}

and k_{off}) in complex fluids. Other methods like fluorescence polarization (FP)^[8] and microscale thermoelectrophoresis (MST)^[9] can also optically detect DNA–protein interactions, though these methods require additional fluorophore labelling (see Table 1 for comparison). More information on other methods, such as isothermal calorimetry (ITC)^[1] and electrochemical techniques,^[10] which are summarized in Table 1, are discussed elsewhere.

Though the thermodynamic and kinetic parameters offer a glimpse into the overall behavior of the system, the binding dynamics of the DNA–protein complex are critical for determining the detailed mechanism. Single-molecule imaging in particular is capable of tracking individual proteins and/or DNA targets as they are translocated or diffuse towards one another (Fig. 1f). In a DNA-curtain assay,^[11] DNA molecules stained by YOYO1 (structure is shown in Fig. 1f) were fixed in a lipid bilayer on a cover slip, and RNA polymerase (RNAP) was fused to green fluorescence protein (GFP). Through total internal reflection fluorescence microscopy (TIRFM), the diffusion of RNAP along the DNA was monitored at the single-molecule level. In addition to optical techniques, high-speed atomic force microscope (AFM, Fig. 1g) offers an alternative imaging approach for studying dynamics.^[12] Though the superior spatial resolution of this technique can offer a preliminary look into the detailed DNA–protein conformations, the throughput and temporal resolution are limited.

<i>in vivo</i> sensing/Imaging	Requirement			Sensitivity	ref.
	Purified sample (P) Crude extract (C)	Immobilization chemistry	Labelling		
	P			–	–
	P		Isotope	0.1–1 mM	[4]
	P			-	-
	P/C			10–100 nM	[5]
	P/C		Radiolabel/Fluorophore	0.1 nM	[1]
	P/C		Radiolabel	0.1 nM	[6a]
	P/C		Radiolabel	0.1 nM	[16]
	P/C		Radiolabel	0.1 nM	[1]
	P		Fluorophore	1-10 nM	[8]
	P/C		Fluorophore	1-10 nM	[9]
	P			1-10 uM	[1]
	P	X		100 uM	[22]
	P/C		Fluorophore	~1 nM	[23]
	P	X		Single molecule	[24]
	P/C	X		~Single molecule	[10]
	P/C	X		1 nM	[7]
	P	(X)	Fluorophore	Single molecule	[25]
X	P/C	X	Fluorophore	Single molecule	[11]
	P	X		Single molecule	[12]
X	P/C	(X)	Fluorophore	Single molecule	[14]

Since the majority of the aforementioned methods require sample purification, real-time detection of DNA–protein interactions *in vivo* remains challenging. Though methods like SPR can use crude extract for measurements, *in vivo* measurements often result in noisy background and low signal-to-noise ratios. A simple way to overcome this limitation is to label DNA and/or protein with fluorophores and use the colocalization of the different fluorescence emissions to monitor the interaction *in vivo* (Fig. 1f). Advancements in microscopy have also pushed the limits of feasibility in the field. For example, super resolution microscopy has allowed researchers to overcome the imaging diffraction limit. One such setup based on photoactivated localization microscopy (PALM) uses a photoactivable fluorescent protein to label a POI for tracking.^[13] DNA polymerase and ligase are fused to photoactivable mCherry to track the mechanism of DNA repair, and the delayed protein diffusion is reflective of the site searching mechanism.

Techniques based on Förster resonance energy transfer (FRET, Fig. 1h) are flexible with the type and purity of samples that can be used for measurement.^[1] In this system, one fluorophore behaves as a donor that absorbs a photon to achieve an excited state. Before it radiatively relaxes to the ground state, the excited donor non-radiatively transfers the energy to a neighboring fluorophore that behaves as an acceptor. This acceptor emits radiatively as it relaxes to its ground state. The FRET efficiency correlates with the distance between two fluorophores, particularly when the fluorophores are in the range of 2–10 nm. This distance-dependent efficiency has been exploited in studies that seek information on DNA–protein proximity. In one example, GFP that was fused to a POI served as a donor to nucleic acids that were stained with Sytox orange, which behaved as the acceptor.^[14] When single-molecule FRET is coupled with a fluorescence lifetime imaging microscope (FLIM), the DNA–protein FRET interaction can be probed independent of local concentration.

Optical Monitoring of DNA–Protein Interactions Using Single-walled Carbon Nanotubes (SWCNTs)

We recently developed a novel optical platform for investigating DNA–protein interactions using SWCNTs. SWCNTs can be conceptualized as rolled-up graphene sheets (Fig. 2a). The wrapping direction determines the (n,m) chirality indices of a particular nanotube, and the chirality ultimately governs the optoelectronic properties of a particular SWCNT, which can be either metallic or semiconducting.^[15] The semiconducting SWCNTs with diameters of 0.7–1.0 nm (Fig. 2b) can be excited by light between 400 nm and 800 nm (*i.e.* E_{22} van Hove transition). The emitted fluorescence occurs at near-infrared (nIR) wavelengths between 900 nm and 1400 nm (*i.e.* relaxation at E_{11} van Hove transition). The characteristic excitation and emission wavelengths of a particular SWCNT is specific to the (n,m) chirality, which determines the bandgap. For example, a (6,5) SWCNT is excited around 570 nm, and the peak fluorescence emission is around 990 nm.

These optical properties are influenced by the polarizability in the vicinity of SWCNT sidewall.^[15] In other words, molecules that adsorb or interact with the surface of the SWCNT can change its nIR fluorescence spectrum (Fig. 2c). These interactions can change the fluorescence intensity and/or shift the peak position. In the past, this sensitivity has been exploited to detect different bio-analytes. For example, small molecules such as neurotransmitters^[16] and nitric oxide^[17] have been detected by monitoring their effects on SWCNT fluorescence. In addition to small molecules, SWCNTs have also been used to detect aptamer-binding proteins that interact with DNA–aptamers immobilized onto the SWCNT surface.^[18]

Analogous approaches can be developed for optically detecting specific DNA–protein interactions. For example, DNA can be non-covalently immobilized onto the nanotube surface (Fig. 2d).^[17–19] As demonstrated in our previous work, we anchored double-stranded DNA (dsDNA) onto SWCNTs with the help of a short single-stranded DNA (ssDNA) overhang.^[19] Even when

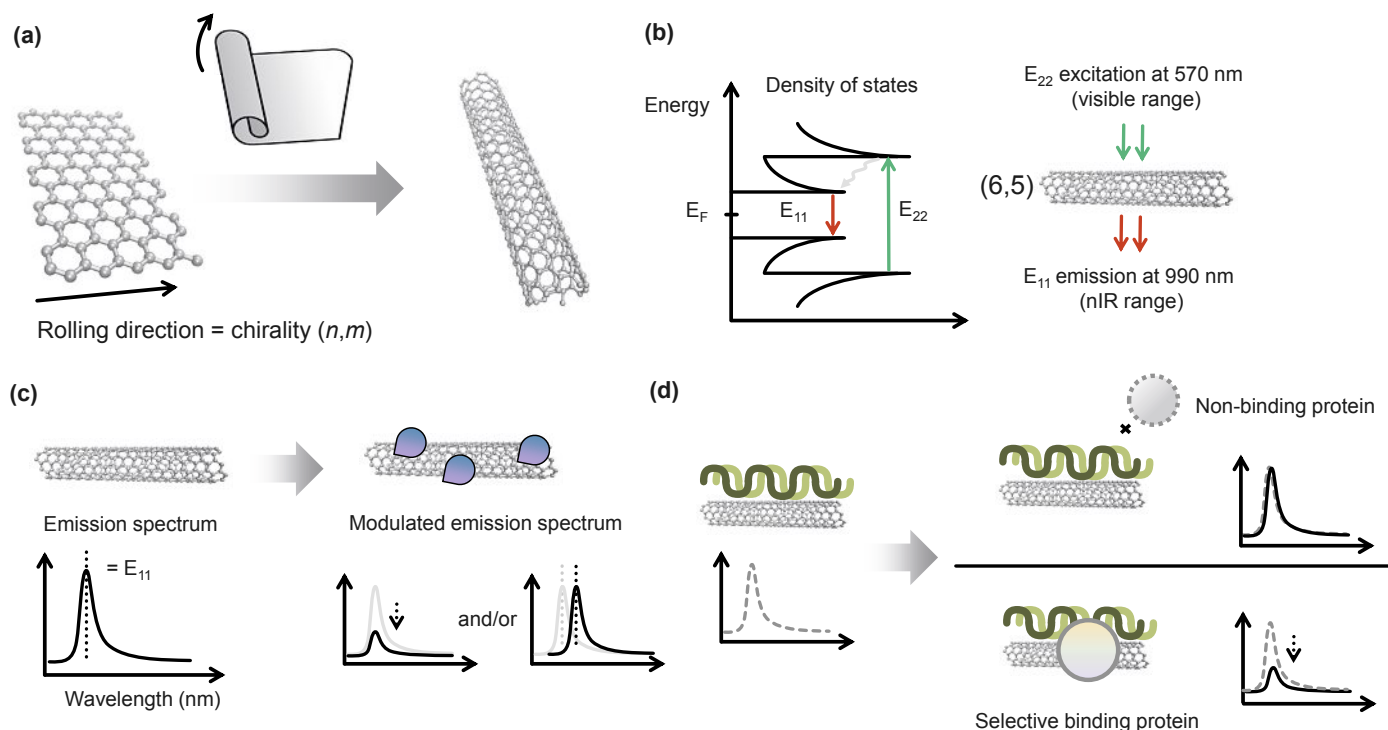


Fig. 2. Single-walled carbon nanotubes (SWCNTs) as optical DNA–protein biosensors. (a) A SWCNT is conceptualized as a rolled-up graphene sheet, and the rolling direction determines the (n,m) chirality. (b) For semiconducting SWCNTs like the (6,5) nanotube, the energy diagram and optical properties are characteristic of the van Hove singularities (E_{22} and E_{11} transitions). (c) The SWCNT fluorescence spectrum is modulated by the surface chemistry. (d) The surface of the SWCNT is functionalized with dsDNA, which can interact with selective binding proteins that yield a change in SWCNT fluorescence (bottom). SWCNTs that lack the target DNA sequence or solutions that lack the binding protein (top) contribute to minimal interaction and negligible fluorescence change.

anchored to the SWCNT surface, the dsDNA remained accessible to restriction enzyme binding and cutting, as confirmed through gel electrophoresis. The sustained targeted enzyme activity on the immobilized surface confirmed (1) that the dsDNA retained its native conformation (B-DNA) in the vicinity of the SWCNT and (2) the specificity of the restriction enzyme towards its target sequence. This technique was applied to different DNA sequences with distinct enzyme recognition sites, demonstrating the versatility of this approach for various DNA targets.

Conclusions and Outlook

In summary, we briefly reviewed analytical methods for investigating DNA–protein interactions, discussing approaches for identifying unknown interactions to characterizing thermodynamic and kinetic parameters for known interactions. We also discussed imaging techniques for single-molecule tracking and *in vivo* applications. Finally, we proposed a novel approach based on the single-molecule sensitivity of nIR SWCNT fluorescence. Though, to the best of our knowledge, this approach has yet to be used to analyze protein activity on DNA, it may offer several advantages for optical DNA–protein imaging. First, the platform benefits from a label-free approach with a facile immobilization procedure based on autonomous DNA adsorption. Also, unlike EMSA and most pull-down assays, this platform uses a non-destructive analytical method that can be extended to real-time monitoring of single-molecule events.^[17,18] Thus, both transient and equilibrium interactions can be studied, and the binding affinity can also be determined after appropriate calibration.^[20] The indefinite photostability of the SWCNT fluorescence further allows continuous, long-term monitoring. Additionally, the different nanotube chiralities offer distinct fluorescence signals for detecting multiple targets simultaneously, without the need for orthogonal bioconjugative procedures. This platform can therefore be extended for use in high throughput, single-molecule measurements of multiple DNA–protein interactions.

Furthermore, the nIR emissions of the SWCNT is also conducive to *in vivo* applications. These emissions are minimally absorbed by biological fluids, tissue, and other biomolecules that typically absorb and/or fluoresce in the visible region of the optical spectrum. Cellular uptake of SWCNTs have already been shown in living cells through endocytosis with minimum cytotoxicity.^[21] Combined with the aforementioned advantages, this technology lays the foundation in achieving the coveted single-molecule, long-term, continuous, and non-destructive monitoring of DNA–protein interactions.

Acknowledgements

The authors are thankful for support from the Swiss National Science Foundation (SNSF) Assistant Professor (AP) Energy Grant.

Received: January 29, 2019

- [1] ‘DNA-Protein Interactions, Methods and Protocols’, Eds. T. Moss, B. Leblanc, Vol. 543, Springer, **2009**.
- [2] a) ‘Plant Transcription Factors, Methods and Protocols’, Eds. L. Yuan, S. E. Perry, Vol. 754, Springer, **2011**; b) J. Zhou, J. Rossi, *Nature Rev. Drug Discov.* **2016**, *16*, 181; c) M. R. Gotrik, T. A. Feagin, A. T. Csordas, M. A. Nakamoto, H. T. Soh, *Acc. Chem. Res.* **2016**, *49*, 1903.
- [3] P. B. F. Ouwkerk, A. H. Meijer, *Curr. Prot. Mol. Biol.* **2001**, *55*, 12.12.1.
- [4] S. Campagne, V. Gervais, A. Milon, *J. Royal Soc. Interf.* **2011**, *8*, 1065.
- [5] M. M. Makowski, C. Grawe, B. M. Foster, N. V. Nguyen, T. Bartke, M. Vermeulen, *Nature Comm.* **2018**, *9*, 1653.
- [6] a) L. M. Hellman, M. G. Fried, *Nature Prot.* **2007**, *2*, 1849; b) M. Brenowitz, D. F. Seneor, R. E. Kingston, *Curr. Prot. Mol. Biol.* **1989**, *7*, 12.4.1.
- [7] M. Ritzefeld, N. Sewald, *J. Amino Acids* **2012**, 816032.
- [8] B. J. Anderson, C. Larkin, K. Guja, J. F. Schildbach, *Methods Enzymol.* **2008**, *450*, 253.
- [9] M. Jerabek-Willemsen, T. André, R. Wanner, H. M. Roth, S. Duhr, P. Baaske, D. Breitsprecher, *J. Mol. Struct.* **2014**, *1077*, 101.
- [10] Y. Nasiri Khonsari, S. Sun, *Chem. Commun.* **2017**, *53*, 9042.
- [11] B. E. Collins, L. F. Ye, D. Duzdevich, E. C. Greene, *Meth. Cell Biol.* **2014**, *123*, 217.
- [12] Y. L. Lyubchenko, L. S. Shlyakhtenko, *Crit. Rev. Eukaryotic Gene Exp.* **2016**, *26*, 63.
- [13] S. Uphoff, R. Reyes-Lamothe, F. Garza de Leon, D. J. Sherratt, A. N. Kapanidis, *Proc Natl Acad Sci USA* **2013**, *110*, 8063.
- [14] L. Camborde, A. Jauneau, C. Briere, L. Deslandes, B. Dumas, E. Gaulin, *Nature Prot.* **2017**, *12*, 1933.
- [15] J.-C. Charlier, X. Blase, S. Roche, *Rev. Modern Phys.* **2007**, *79*, 677.
- [16] S. Kruss, M. P. Landry, E. Vander Ende, B. M. A. Lima, N. F. Reuel, J. Zhang, J. Nelson, B. Mu, A. Hilmer, M. Strano, *J. Am. Chem. Soc.* **2014**, *136*, 713.
- [17] 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, M. S. Strano, *J. Am. Chem. Soc.* **2011**, *133*, 567.
- [18] M. P. Landry, H. Ando, A. Y. Chen, J. Cao, V. I. Kottadiel, L. Chio, D. Yang, J. Dong, T. K. Lu, M. S. Strano, *Nat. Nanotechnol.* **2017**, *12*, 368.
- [19] S.-J. Wu, N. Schuergers, K.-H. Lin, A. J. Gillen, C. Corminboeuf, A. A. Boghossian, *ACS Appl. Mater. Interf.* **2018**, *10*, 37386.
- [20] G. Bisker, J. Dong, H. D. Park, N. M. Iverson, J. Ahn, J. T. Nelson, M. P. Landry, S. Kruss, M. S. Strano, *Nature Comm.* **2016**, *7*, 10241.
- [21] P. N. Yaron, B. D. Holt, P. A. Short, M. Lösche, M. F. Islam, K. N. Dahl, *J. Nanobiotechnol.* **2011**, *9*, 45.
- [22] J. Xu, K. W. Liu, K. S. Matthews, S. L. Biswal, *Langmuir* **2011**, *27*, 4900.
- [23] T. C. Yang, C. E. Catalano, N. K. Maluf, *Methods Enzymol.* **2015**, *562*, 305.
- [24] S. Liu, X. Zhang, W. Luo, Z. Wang, X. Guo, M. L. Steigerwald, X. Fang, *Angew. Chem. Int. Ed.* **2011**, *50*, 2496.
- [25] H. Hwang, S. Myong, *Chem. Soc. Rev.* **2014**, *43*, 1221.