Supporting Information

Expanding the scope of protein-detecting electrochemical DNA "scaffold" sensors

Di Kang[1, //], Claudio Parolo[1, //], Sheng Sun[1], Nathan Ogden,[2] Frederick W. Dahlquist[1,3] and Kevin W. Plaxco[1, 3, *]

[1]Department of Chemistry and Biochemistry

[2]Department of Material Science

[3]Interdepartmental Program in Biomolecular Science and Engineering

University of California, Santa Barbara

Santa Barbara, CA 93106 USA

Materials and Methods

Protein expression and purification.

CdiA-CT specificity domain

The sequence encoding the CdiA-CT^{MHI813} specificity domain was sub-cloned into vector pET24b (Novagen). This plasmid was transformed into Rosetta™ competent cells (Novagen) for expression of a 159 residues protein with a C-terminal hexahistidine tag[1]. Rosetta™ cells (Novagen) carrying the plasmid were grown at 37°C with shaking in lysogeny broth (LB) medium. Expression was induced at OD_{600} ~0.3 by the inclusion of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and the culture was harvested 3 hr after induction. Cell pellets were suspended in 50 mM $NaPO_4$ pH 8.0, 300 mM sodium chloride, 10 mM imidazole, 10 mM β-mercaptoethanol and lysed in a French pressure cell. Cell debris was removed by centrifugation for 1 hour at 13000 g. The supernatant was applied to a nickel column. The column was washed by applying seven column volumes of 50 mM $NaPO_4$ pH 8.0, 300 mM sodium chloride, 35 mM imidazole, 10 mM β-mercaptoethanol. Hexahistidine-tagged CdiA-CT^{MHI813} specificity domain was eluted with 50 mM $NaPO_4$ pH 8.0, 300 mM sodium chloride, 250 mM imidazole, 10 mM β-mercaptoethanol.

CheY

E. coli CheY were expressed in *E. coli* BL21(DE3) (Novagen) strain and purified following published protocols[2]. Specifically, the gene encoding wild-type *E. coli* CheY (residues 1-129) was cloned into pET28a (Novagen) at the NcoI and XhoI sites in frame with the carboxy-terminal hexahistidine tag. Transformants were grown at 37°C in LB medium and induced at OD_{600} = 0.4 by the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The culture was harvested after 3 hr, resuspended in 50 mM sodium phosphate, 10 mM imidazole, 300 mM sodium chloride, pH 8.0 buffer and lysed by French pressure cell press. Cell debris was removed by centrifugation 30,000 g for 30 min before application of the supernatant to an immobilized metal ion (nickel) affinity chromatography column (GE Healthcare). The column was washed with 50 mM sodium phosphate, 300 mM sodium chloride and 35 mM imidazole, pH 8. His-tagged CheY was eluted with 150 mM imidazole in the same buffer. The collected fractions were dialyzed into 50 mM sodium phosphate, 150 mM sodium chloride, pH 7.9 and concentrated to ~1 mM.

CheA P3 and full-length CheA

CheA P3 domain[3] and full-length *E. coli* CheA[4] proteins were expressed in *E. coli* BL21 (DE3) and purified following published protocols. Particularly, genes encoding CheA P3 domain (residues 259-323) and full-length CheA from E coli were subcloned into vector pET22b (Novagen) so as to place a six histidine tail on their carboxyl termini. Transformants were grown at 37°C in lysogeny broth (LB) medium

and induced at OD_{600} = 0.4 by the addition of 1 mM IPTG. The cells were harvested 3 hr after induction. Cells were lysed in a French Press and cell debris was removed by centrifugation. The expressed proteins were purified using Ni-NTA affinity chromatography (GE Healthcare). Purified proteins were dialyzed against the buffer containing 50 mM $NaPO_4$ (pH7.5) and concentrated to a final concentration of 0.5-1.5 mM.

GlnRS

E. coli GlnRS was expressed and purified following published protocols[5]. Expression of the protein was achieved by addition of 1 mM IPTG and 10 mM β-mercaptoethanol to the culture at the time of induction (OD_{600} ~ 0.4). The expressed protein was purified using Ni-NTA affinity chromatography (GE Healthcare).

Green fluorescent protein (GFP)

His-tagged emerald green fluorescent protein (GFP) was expressed from pRSET-EmGFP expression plasmid (Invitrogen) in *E. coli* BL21 (DE3) and purified via nickel ion affinity chromatography. The resin was washed with 8 column volumes of 50 mM sodium phosphate, 300 mM sodium chloride, and 20 mM imidazole, pH 8.0. The his-tagged GFP was then eluted with 1 column volume of 250 mM imidazole in the same buffer. GFP containing fractions were dialyzed and concentrated. Purity was >95% as determined by SDS-PAGE electrophoresis.

Electrode fabrication and measurement

Sensors were prepared by analogy to previously described E-DNA sensors[6,7]. In brief, prior to sensor fabrication, gold disk electrodes (2.0 mm diameter, CH Instruments, Austin, TX) were cleaned both mechanically (by successively polishing with 1 μm diamond and 0.05 μm aluminum oxide slurries) and electrochemically (through successive scans in 0.5 M sulfuric acid and 0.1 M sulfuric acid 0.01M KCl). Surface area was determined by cycle voltammetry scan in 0.05 M sulfuric acid. Methylene blue (MB) and thiol modified anchor strand DNA (5'-C6-Thiol-G CAG TAA CAA GAA TAA AAC GCC ACT G-MB-3') (Biosearch Technologies, U.S.A) was reduced for 1 hr at room temperature in the dark in 10 mM tris(2-carboxyethyl)phosphine hydrochloride (Molecular Probes, Carlsbad, U.S.A). The reduced solution was diluted to a final concentration of 30 nM DNA in 10 mM phosphate/100 mM NaCl buffer, pH 7.5. Cleaned electrodes were incubated in the DNA solution for 30 min at room temperature, rinsed with deionized water, and then all electrodes were "backfilled" with 3 mM 6-mercapto-1-hexanol (Sigma, U.S.A) overnight (~16 hr) at 4°C.

Following fabrication all sensors were rinsed in deionized water and tested with square wave voltammetry (SWV) with a 25 mV amplitude signal at a frequency of 20 Hz in 50 mM sodium phosphate, 100 mM NaCl buffer to confirm deposition of the anchor strand. The electrodes were then incubated with 100 nM of complementary

DNA modified with either NTA (5'- NTA - AGT GGC GTT TTA TTC TTG TTA CTG -3'; Gene Link, U.S.A), Dig (5' Dig - AGT GGC GTT TTA TTC TTG TTA CTG -3'; Biosearch, U.S.A); or biotin (5'- biotin - AGT GGC GTT TTA TTC TTG TTA CTG -3'; Biosearch, U.S.A) for 30 min. Following this we measured their SWV peak current to define the signal produced by the unmodified scaffold.

Recognition elements were grafted onto the NTA-modified scaffold using Cu/NTA/His-tag complexation. We employed copper rather than nickel as histidine-copper complexes are less kinetically labile[8]. To do so we incubated the NTA-scaffold-modified electrodes with 100 μM copper sulfate (Sigma, U.S.A) in 50 mM sodium phosphate, 100 mM NaCl for 20 min. These electrodes were then incubated with 10 μM of the purified recognition element for 30 min at room temperature followed by washing with 150 mM imidazole in 50 mM sodium phosphate, 100 mM sodium chloride, pH 7.5. We then rinsed the resulting sensor with 50 mM sodium phosphate, 100 mM sodium chloride buffer, and immediately tested with square wave voltammetry (SWV).

Signal response data were collected using square wave voltammetry (SWV) with a 25 mV amplitude signal at a frequency of 20 Hz in either 50 mM sodium phosphate, 100 mM NaCl. Error bars represent the standard deviations of measurements performed on at least three independently fabricated sensors.

Simulations

To model our system, we developed a simple Monte Carlo simulation using Matlab. For each conformation in our ensembles we rebuilt the DNA scaffold to give it some degree of intrinsic flexibility. To build each new scaffold conformation the position of the first base pair was defined by randomly choosing a rotational angle, ρ, from the evenly distributed range 0 and 2π and an angle, θ, the angle between the axis and the surface, which was randomly selected from a Gaussian distribution centered on 0; the width of the Gaussian we employ defines the flexibility of the linker connecting the scaffold to the surface (Figure 4a). The center of the base pair was then placed 6 Å from the origin along this vector. The position of the DNA backbone was then determined via appropriate coordinate transfers. For each subsequent base, new angles, ρ' and θ', (relative to the coordinates of the previous base) were randomly selected. The rotational angle, ρ', was selected from a Gaussian distribution centered on the average rotational angle between DNA base pairs (0.6178 radians)[9] with a standard deviation of 0.0125, while the bend angle, θ', was selected from a distribution centered on 0 with a standard deviation of $\frac{\pi}{60}$. A sphere representing the center of the next DNA base pair was placed 3.38 Å (the per-base translation of DNA) away along this vector and the position of the backbone again determined. This process was reiterated 27 times to create our scaffold. Upon reaching the final base pair the position of the methylene blue reporter was defined by placing a sphere 7 Å away (to account for the size of the methylene blue and its linker) normal to one of

the backbone chains. The "protein" was then simulated by attaching a variable radius sphere 3 Å away (to account for the his-tag) from the opposite backbone. To simulate a flexible attachment from the scaffold to the protein we defined an angle, ϕ, defined as the angle of the linker relative to the vector of the last base pair, which was randomly selected from a Gaussian distribution of varying standard deviation. For comparision with our experimental data we converted these radii into molecular weights according to the following formula[10]:

$$R_g = 7.78(Mw)^{0.37}$$

To build our conformational ensembles, we next determined the energy of each conformation, which is comprised of the internal energy of the DNA and any interaction with the surface. The internal energy is given by:

$$U_{bend} = \sum_{k=2}^{n} \frac{\frac{k_b T}{2} L_p}{3.38} (1 - u_k \cdot u_{k-1})$$

$$U_{twist} = \sum_{k=2}^{n} \frac{k_b T}{2} \kappa (\rho_k - \bar{\rho})^2$$

Here L_p is the persistence length of double-stranded DNA (approximated to be 53.5 nm), u_k is the unit vector defining the kth base relative to the previous base, κ is the twist force constant for double-stranded DNA (203.49 radians^{-2})[9], and $(\rho_k - \bar{\rho})$ is the difference between the rotation angle of the kth base and the average rotational angle of double-stranded DNA. Interactions with the surface were defined by a hard-wall approach in which the energy of these interactions are zero for conformations that do

not overlap with the monolayer and infinity if there is any overlap between the protein and the monolayer. The monolayer was simulated as an exclusion layer 9.25 Å from the surface. Using this potential, we used Monte Carlo approaches to simulate 1,000,000 conformations for each set of parameters (i.e., for each discrete value of θ, ϕ, and protein radius) to generate our ensembles.

Per Uzawa et al.[11] we calculated a weighted mean effective electron transfer rate (k_{eff}) for each ensemble according to:

$$k_{eff} = \frac{1}{n}\sum_{i=1}^{n} P_i \cdot k_i(z)$$

Here n is the number of scaffolds with non-zero probability of occurring, P_i is the probability that the i[th] scaffold is in its current conformation as determined by Maxwell Boltzmann statistics, and $k_i(z)$ is a distance dependent electron transfer rate. The latter is given by a 1-D tunneling equation with a length constant of 1.2 $Å^{-1}$ [12].

Supporting Figures

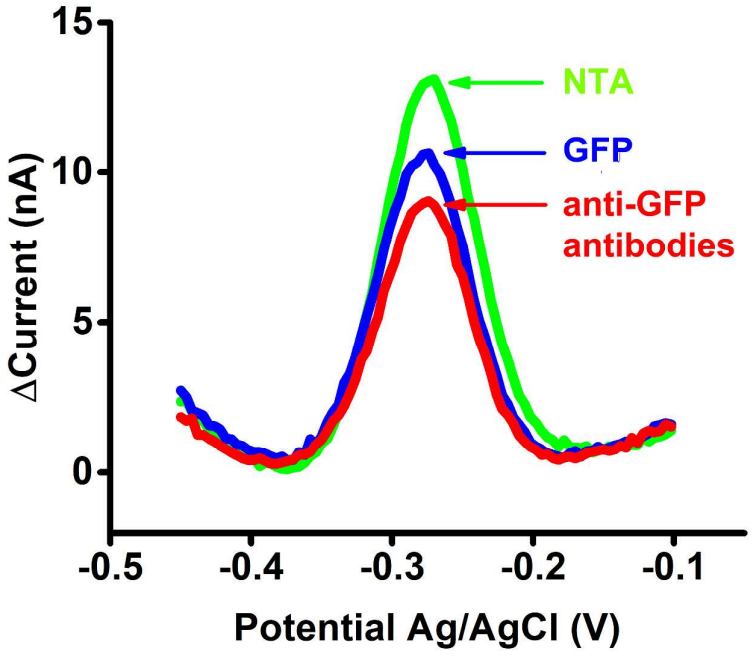

Figure S1. Square-wave voltammagrams collected for an unmodified scaffold, the scaffold after the attachment of GPF (as a recognition element), and after the binding of polyclonal anti-GFP antibodies. In the absence of any recognition element the methylene blue is free to approach the electrode, producing a large current (Green). Upon addition of the recognition element collisions, and thus the current, is reduced (Blue). This current reduced more upon the binding of anti-GFP antibodies (Red).

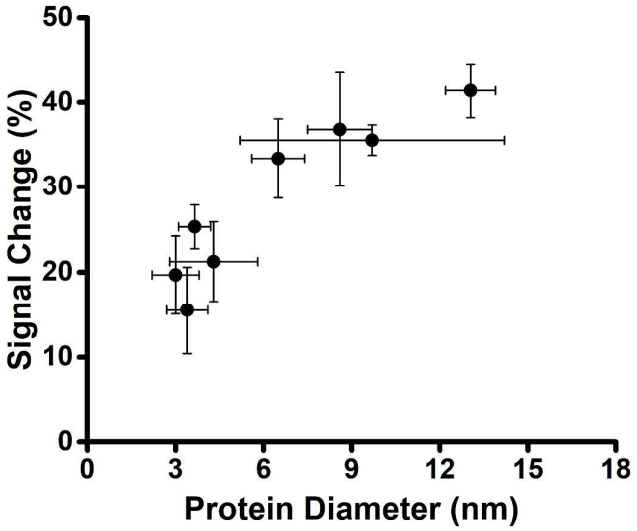

Figure S2. As the recognition element diameter becomes larger, the signaling current is reduced until ultimately plateauing at a cutoff that corresponds to a protein diameter of ~7 nm. The 3.5 nm radius of such a protein corresponds to the diameter of the double helical scaffold plus the length of the methylene blue and its linker, suggesting that the observed signal change arises due to simple steric blocking by the attached protein. Y-axis error bars reflect the standard deviations of replicates performed using independently fabricated sensors. X-axis error bars reflect the maximum and minimum diameters of the protein structure, the circles representing the average of the two.

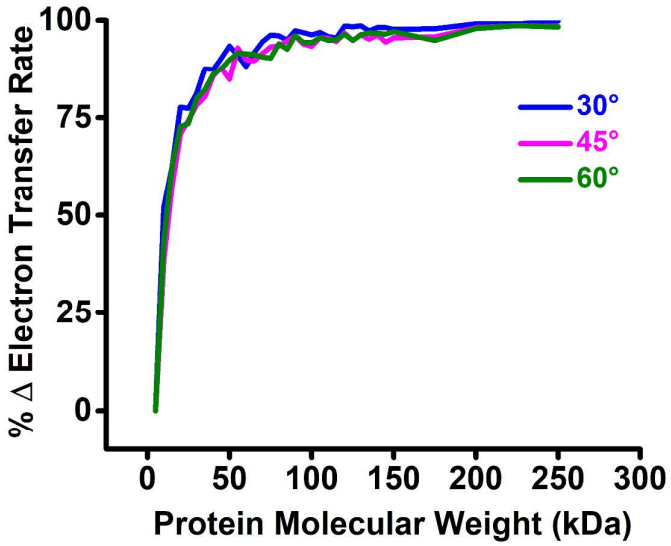

Figure S3. The simulated, relative electron transfer rate is effectively independent of the flexibility of the linker connecting the scaffold to the surface (the distribution of the angle ϕ in Figure 4A). After normalizing each of the curves against the transfer rate of the unmodified scaffold, we see that increasing the surface-to-scaffold linker flexibility has no effect on the shape of the gain curve or the value at which it plateaus. Here we employed a fixed 15˚ Gaussian width defining the flexibility of the linker connecting the recognition element to the scaffold.

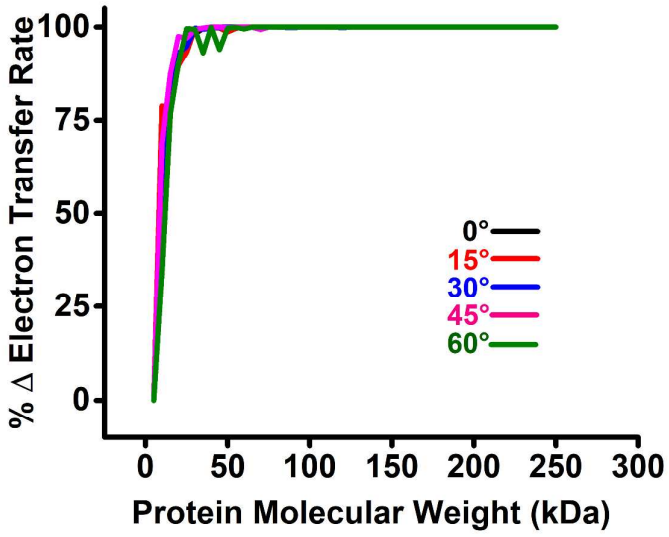

Figure S4. The simulated, relative electron transfer rate is effectively independent of the flexibility of the linker connecting the methylene blue redox reporter to the DNA scaffold. Here the recognition element to scaffold linker angle was fixed at 0˚ and not allowed to vary.

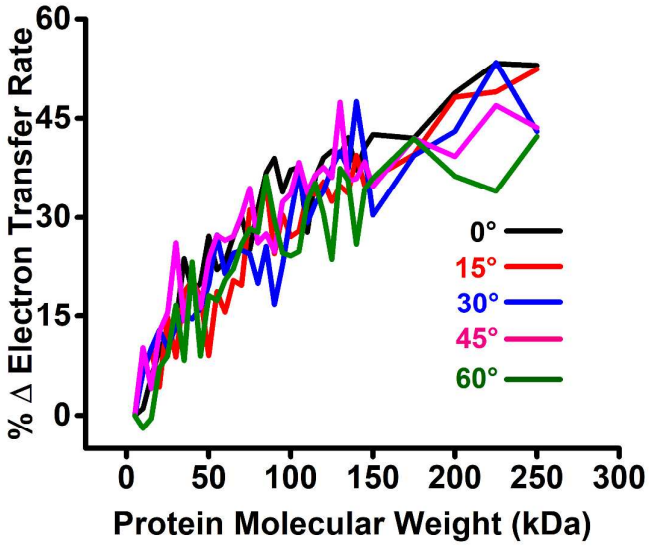

Figure S5. Simulations suggest that moving the methylene blue closer to the electrode (here it is on the 11th base pair of the 27 base pair scaffold) causes the platform becomes less sensitive to changes in the size of the recognition element. Given that binding-induced changes in the size of the complex (between the recognition element and the target) drive signaling in this sensor, this suggests such architectures will exhibit poorer gain for small recognition elements. They may, in contrast, allow sensors to be fabricated using larger recognition elements than is possible with a terminally-placed reporter. As shown (colored curves), this effect is independent of the flexibility of the linker connecting the recognition element to the scaffold.

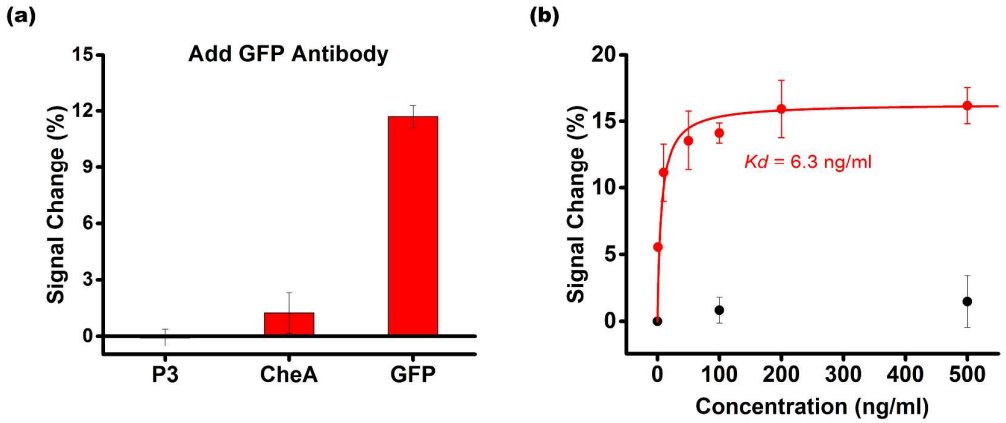

Figure S6. (a) Using green fluorescent protein (GFP) – NTA complex as the recognition element, we easily detect anti-GFP antibodies with a total signal change of about ~12%. In contrast, the specific antibodies did not produce a significant signal change when exposed to E-DNA scaffold sensors incorporating the wrong recognition element (the CheA P3 domain and full-length CheA). (b) Employing green fluorescent protein (GFP) and polyclonal anti-GFP antibodies as our target we observe Langmuir isotherm binding with a dissociation constant 6.3 ng/ml. The detection limit is ~2 ng/ml (the lowest concentration shown on this plot) and the signal gain is about ~15%. In contrast, the sensor does not produce response to non-specific antibodies (black data points). The error bars reflect standard deviation of measurements performed using multiple independently fabricated sensors.

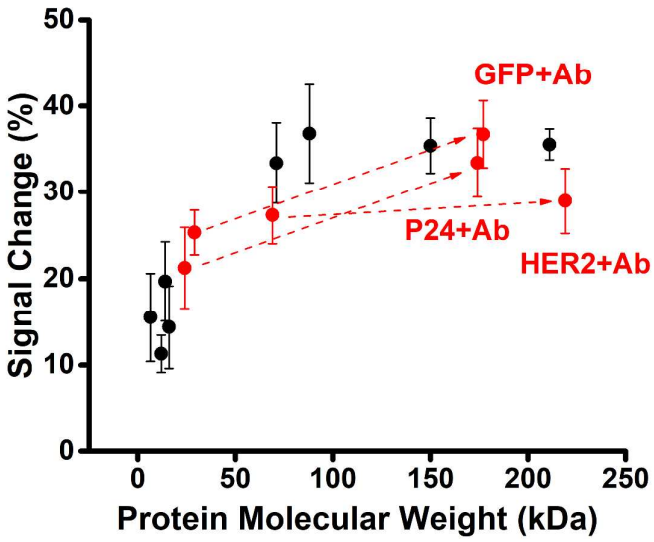

Figure S7. Using the full-length proteins p24 (24 kDa), GFP (29 kDa) and HER2 (69 kDa) as recognition elements. Sensors employing p24 and GFP exhibit 12% and 13% signal changes, respectively, upon binding their target antibodies. A sensor employing the higher molecular weight HER2 protein, in contrast, produces no significant signal change upon antibody binding. The error bars reflect standard deviation of measurements performed using multiple independently fabricated sensors.

References

(1) Willett, J. L.E.; Gucinski, G.C.; Fatherree, J. P.; Low D. A.; Hayes, C. S. Contact-dependent growth inhibition toxins exploit multiple independent cell-entry pathways. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 11341-11346.

(2) Mo, G.; Zhou, H.; Kawamura, T.; Dahlquist, F. W. Solution structure of a complex of the histidine autokinase CheA with its substrate CheY. *Biochemistry* **2012**, *51*, 3786-3798.

(3) Wang, X.; Vallurupalli, P.; Vu, A.; Lee, K.; Sun, S.; Bai, W.; Wu, C.; Zhou, H.; Shea, J. E.; Kay, L. E.; Dahlquist, F. W. The linker between the dimerization and catalytic domains of the CheA histidine kinase propagates changes in structure and dynamics that are important for enzymatic activity. *Biochemistry* **2014**, *53*, 855-861.

(4) Stock, A.; Chen, T.; Welsh, D.; Stock, J. CheA protein, a central regulator of bacterial chemotaxis, belongs to a family of proteins that control gene expression in response to changing environmental conditions. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 1403-1407.

(5) Rodriguez-Hernandez, A.; Bhaskaran, H.; Hadd, A.; Perona, J. J. Synthesis of Glu-tRNA(Gln) by engineered and natural aminoacyl-tRNA synthetases. *Biochemistry* **2010**, *49*, 6727-6736.

(6) Xiao, Y.; Rowe, A. A.; Plaxco, K. W. Electrochemical detection of parts-per-billion lead via an electrode-bound DNAzyme assembly. *J. Am. Chem. Soc.* **2007**, *129*, 262-263.

(7) Rowe, A. A.; White, R. J.; Bonham, A. J.; Plaxco, K. W. Fabrication of electrochemical-DNA biosensors for the reagentless detection of nucleic acids, proteins and small molecules. *J. Vis. Exp.* **2011**, *52*, 2922.

(8) Kang, D.; Sun, S; Kurnik, M.; Morales, D.; Dahlquist, F. W.; Plaxco, K. W. New architecture for reagentless, protein-based electrochemical biosensors. *J. Am. Chem. Soc.* **2017**, *139*, 12113-12116.

(9) Olson, W. K.; Gorin, A. A.; Lu, X. J.; Hock, L. M.; and Zhurkin, V. B. DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 11163-11168.

(10) Smilgies, D. M.; Folta-Stogniew, E. Molecular weight-gyration radius relation of globular proteins: a comparison of light scattering, small-angle X-ray scattering and structure-based data. *J. Appl. Crystallogr.* **2015**, *48*, 1604-1606.

(11) Uzawa, T.; Cheng, R. R.; White, R. J.; Makarov, D. E.; Plaxco, K. W. A mechanistic study of electron transfer from the distal termini of electrode-bound, single-stranded DNAs. *J. Am. Chem. Soc.* **2010**, *132*, 16120-16126.

(12) Bard, A. J.; Faulkner, L. R. *Electrochemical Methods: Fundamentals and Applications*, Wiley, **2000**.