UC Riverside

UC Riverside Previously Published Works

Title

Probing and quantifying DNA–protein interactions with asymmetrical flow field-flow fractionation

Permalink

https://escholarship.org/uc/item/2wp28187

Authors

Ashby, Jonathan
Schachermeyer, Samantha
Duan, Yaokai
et al.

Publication Date

2014-09-01

DOI

10.1016/j.chroma.2014.07.002

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Probing and Quantifying DNA-protein Interactions with Asymmetrical Flow Field-Flow Fractionation

Jonathan Ashby,[1†] Samantha Schachermeyer,[1†] Yaokai Duan,[1] Luis A. Jimenez,[2] Wenwan Zhong[1]*

[1] Department of Chemistry, [2]Program in Biomedical Sciences, University of California, Riverside

† These authors contributed equally to the work.

*Corresponding author: Wenwan Zhong

Department of Chemistry

University of California

Riverside, CA 92521

USA

Email: wenwan.zhong@ucr.edu

Tel: 951-827-4925

Fax: 951-827-4713

Abstract

Tools capable of measuring binding affinities as well as amenable to downstream sequencing analysis are needed for study of DNA-protein interaction, particularly in discovery of new DNA sequences with affinity to diverse targets. Asymmetrical flow field-flow fractionation (AF4) is an open-channel separation technique that eliminates interference from column packing to the non-covalently bound complex and could potentially be applied for study of macromolecular interaction. The recovery and elution behaviors of the poly(dA)$_n$ strand and aptamers in AF4 were investigated. Good recovery of ssDNAs was achieved by judicious selection of the channel membrane with consideration of the membrane pore diameter and the radius of gyration (R_g) of the ssDNA, which was obtained with the aid of a Molecular Dynamics tool. The R_g values were also used to assess the folding situation of aptamers based on their migration times in AF4. The interactions between two ssDNA aptamers and their respective protein components were investigated. Using AF4, near-baseline resolution between the free and protein-bound aptamer fractions could be obtained. With this information, dissociation constants of ~16nM and ~57 nM were obtained for an IgE aptamer and a streptavidin aptamer, respectively. In addition, free and protein-bound IgE aptamer was extracted from the AF4 eluate and amplified, illustrating the potential of AF4 in screening ssDNAs with high affinity to targets. Our results demonstrate that AF4 is an effective tool holding several advantages over the existing techniques and should be useful for study of diverse macromolecular interaction systems.

Key words: field-flow fractionation, DNA-protein interaction, aptamer, DNA folding, affinity measurement

1. Introduction

DNA-protein interactions control diverse cellular functions, such as transcription regulation, DNA replication and repair, chromosome maintenance, etc. Interruption of such interactions could lead to numerous diseases, including cancers.[1,2] To achieve a better understanding of the nature of DNA-protein interaction, numerous research efforts have been devoted to measurement of binding kinetics and affinity, and identification of DNA sequences with affinity to transcription factors, polymerases, histones, etc.[3-6] In addition to endogenous interactions, synthetic DNA probes - DNA aptamers - that can interact with proteins and other biological targets strongly and specifically, have been discovered [7,8] with high utility in molecular detection and great potential as therapeutic agents.[9-12] Discovery of additional aptamers with high functionality and obtaining more knowledge on aptamer-target interaction are thus in great demand.

Many techniques have been developed for mapping of DNA-protein interaction.[3] Although high-throughput next generation sequencing techniques have greatly enhanced the discovery rate and reliability of DNA-protein interactions,[5,13] the first step of isolating the DNA-protein complex is limited to immunocapture on a solid support. Similar immobilization is also required for interaction studies using surface plasmon resonance [14] and protein arrays.[6,15] Immobilization could reduce protein stability, change binding epitopes, and introduce interference to the binding from the support surface.[16]

On the other hand, open channel separation methods like capillary electrophoresis (CE) can carry out affinity and kinetic measurements as well as binding screening without protein immobilization, in which the absence of channel packing imposes minimal disturbance to the interacting molecules. CE also provides the advantages of rapid separation and high resolution

for study of various interaction systems, including DNA-protein, protein-nanoparticles, ligand-receptor, etc.[17-20] It may also be used for the selection of aptamers.[21] However, severe Joule heating that destroys separation resolution imposes big problems to CE, since buffers with high salt contents are needed to maintain good interaction,[22] appropriate DNA folding, and intact protein structure.[23-25] CE's low sample capability also makes it difficult to be coupled with downstream discovery of new sequencing with high affinity to targets.

Another open channel separation technique, asymmetric flow field-flow fractionation (AF4), should be useful for study of biological complexes as well.[26,27] Capable of separating analytes based on hydrodynamic size, AF4 employs an asymmetrical trapezoidal channel, the breadth of which decreases from inlet toward outlet. (Supporting Information Figure S1).[26] The carrier solution entering the channel from the inlet continuously escapes through the accumulation wall made of an ultrafiltration membrane, splitting into a cross flow and a decreasing axial flow. All analytes are pressed against the accumulation wall by the cross flow, and simultaneously diffuse away due to Brownian motion. Analytes with larger sizes diffuse slower and thus are located at a position closer to the wall, being eluted by the axial flow at later times than those with smaller sizes.

Compared to CE, AF4 can use carrier solutions with divalent cations and high ionic strength that is beneficial for maintaining complex integrity.[28] Additionally, its high sample capacity allows convenient after-column collection for down-stream analysis.[29] These advantages have been used in determining the affinity of protein-protein interactions[30] and for characterization of protein complexes,[31,32] and should also be useful in the discovery of ssDNAs with high affinity to targets. However, few reports on using AF4 to analyze DNA, in particular short ssDNAs containing motif for protein binding, have been found.[29,33-36] Giddings *et al.*

demonstrated the initial ability of symmetrical F4 to fractionate ssDNA and dsDNA of similar length,[29] while Moon *et al.* demonstrated the ability of miniaturized AF4 to fractionate 50 and 100nt ssDNA, as well as a complex between ssDNA and replication protein A.[33] In addition, Winnik *et al.* utilized AF4 to characterize DNA-chitosan complexes.[34] However, these previous applications did not specifically investigate the affinity of DNA towards specific targets. Herein, to assess AF4's capability in studying ssDNA-protein interaction, the present work explored the recovery and retention behaviors of a series of poly(dA)$_n$ strands and several aptamers; evaluated whether aptamers could maintain certain degrees of folding during separation; measured the binding affinity of two aptamer-protein systems; and investigated the feasibility of using AF4 as an aptamer selection tool through collection and amplification of the eluted protein-bound aptamer.

2. Materials and Methods

2.1 Chemicals

Phosphate buffered saline (1× PBS), used as the carrier solution in the recovery and elution studies, contained 137 mM NaCl, 10 mM phosphate sodium salts, 2.7 mM KCl, and 1.0 mM MgCl2. Its pH was adjusted to 7.40 (± 0.03) with 6N HCl. This was also the buffer when preparing the IgE-aptamer binding mixture. Analysis of streptavidin-aptamer binding employed a buffer consisting of 20 mM Tris HCl (pH 7.4), 100 mM NaCl, 5 mM KCl, 2 mM $MgCl_2$, and 1 mM $CaCl_2$. All carrier solutions for AF4 were filtered with a 0.2-μm nitrocellulose filter (Whatman, NJ, USA). Sources and preparation of biochemicals, including proteins and DNAs, were listed in Supporting Information.

2.2 AF4 conditions

An AF2000 system manufactured by Postnova Analytics (UT, USA) was used in this study. The trapezoidal separation channel was 0.350 mm thick (thickness of the spacer), and its tip-to-tip length was 275 mm, with an inlet triangle width of 20 mm and an outlet width of 5 mm. The injection loop volume was 20 µL. The surface area of the accumulation wall was 3160 mm^2, which was made out of the regenerated cellulose ultrafiltration membrane (Postnova Analytics) with the molecular weight cutoff (MWCO) value of 5- or 10-kDa, definition useful for both proteins and nucleic acids. A longer tubing connecting the channel outlet to the detector was used with the 5-kDa membrane to correct for the backpressure change induced by the pore size difference, so that the same axial-outlet flow rates could be used for both membranes.

All separations were done with a channel-inlet flow of 3.30 mL/min, a crossflow of 3.00 mL/min, and an axial-outlet flow of 0.30 mL/min. During the focusing/relaxation step, the channel-inlet flow was 0.30 mL/min and the focusing flow (a flow entering at a position further down from the inlet to focus the analyte into a narrow sample zone) was 3.00 mL/min. The focusing/relaxation step lasted for 8 minutes. A SPD-20A Prominence UV/Vis detector (Shimadzu, MD, USA) at 280 nm was used to monitor the eluent for unlabeled DNAs, and a 474 scanning fluorescence detector with an excitation wavelength of 495 nm and an emission wavelength of 525 nm (Waters, MA, USA) was employed to monitor the fluorescently labeled DNAs. Fluorescence detection was used in the analysis of binding affinity between fluorescently labeled aptamers and their corresponding proteins, while absorbance detection was used in all other measurements.

2.3 Recovery Calculation

Recovery of ssDNAs in AF4 was calculated using the peak areas obtained with or without a global crossflow, i.e. the apparent crossflow rate was zero. We defined the peak area of ssDNA obtained using no global crossflow but both the channel-inlet and axial-outlet flows at 0.30 mL/min as A_{NCF}, which should be proportional to the concentration of ssDNA in the injected sample. The axial-outlet flow passes through the detector. The peak area obtained with the same axial-outlet flow rate at 0.30 mL/min but a channel flow at 3.30 mL/min and a crossflow at 3.00 mL/min, was defined as A. Any reduction from A_{NCF} to A would indicate sample loss inside the AF4 channel. Thus, the ratio of A and A_{NCF} can be used to calculate sample recovery:

$$\% \text{ Recovery} = (A/A_{NCF}) \times 100\% \quad (1)$$

2.4 Calculation of Dissociation Constant

For affinity measurement, a Alexa Fluor 488-labeled aptamer was first denatured in boiling water for 5 minutes and kept on ice until mixed with the target protein. Increasing amounts of protein were added to 2 nM of the Weigand aptamer (IgE) or 10 nM of the streptavidin aptamer. The mixtures were injected into the AF4 and fluorescence fractograms obtained. From the fractograms, the amount of free aptamer and protein-bound aptamer (peak area) was obtained. The bound ratio was determined through division of the area of the protein-bound aptamer peak by the sum of both free and bound aptamer peaks, which was then used to calculate the dissociation constant (K_d) via the Hill equation[34, 35]:

$$\text{Bound ratio} = \text{Bound}_{max} \times [P]^n/(K_d^n + [P]^n) \quad (2)$$

The Hill equation is commonly used to describe ligand-receptor binding. Bound_{max} is the bound ratio when the binding curve reached plateau; [P] represents free protein concentration in the mixture; and n is the cooperativity of binding.

2.5 Molecular Dynamics (MD) Simulation and Folded Percentage Calculation

The initial 3D structures of the poly $(dA)_n$ strands and the unfolded aptamers were in the linear B-form; and those of the folded aptamers were in the forms built using w3DNA (http://w3dna.rutgers.edu)[36] as predicted by Mfold, oligonucleotide secondary structure prediction software (http://www.bioinfo.rpi.edu/applications/mfold),[37] except for the thrombin aptamer. The thrombin aptamer folds into a G-quadruplex not predictable by Mfold; and its initial structure adopted the reported crystal structure (PDB ID: 4DIH), to which three more deoxyribonucleotides (AGT) were added to the 5' terminal. MD simulations were performed on the Stampede supercomputer resource (Texas Advanced Computing Center) provided by XSEDE Science Gateways with the Amber 12 program package.[38, 39] An AMBER 99 force field[40] and TIP3P water model[41] were applied. The counter ions of Na^+, Mg^{2+}, K^+ and Cl^- were used in modeling at the same concentrations as in the AF4 carrier solution to mimic the actual solution environment for the ssDNAs. After energy minimization of the starting molecule structure, 1000 picoseconds (ps) equilibration without restraints was performed with constant pressure and the periodic boundaries maintained by the particle mesh Ewald method,[42] while the volume of the box was allowed to change. The simulated 3D structure allowed the measurement of d_{xyz} of each atom, which was the distance from the atom to the center of mass of all selected atoms. The d_{xyz} values of all atoms were then used for calculation of the radius of gyration (R_g) of the worm-like ssDNA, using the mathematic function (sqrt(SUM[d_{xyz}^2 □ mass] / total mass) available in Amber. R_g measures the root mean square distance of the parts from its center of gravity, and is directly proportional to the hydrodynamic radius of the DNA coil. The calculations were performed for

every coordinate frame obtained for every 2 ps. More details about the MD process can be found in Supporting Information.

Using the R_g and the measured retention time of the poly(dA)$_n$ strands, we obtained the calibration plot for Log t vs. Log R_g, which in turn was used to calculate the theoretical elution time of the unfolded or folded form of each aptamer, depicted as t_{UF} and t_F, respectively, based on their corresponding R_g values. The measured elution time (t_M) of each aptamer is in fact the weighted sum of t_{UF} and t_F with the weights being the percentage of each form in the entire population. Therefore, the percentage of aptamer molecules staying in a folded structure during AF4 separation can be calculated by:

$$\%\text{Folded} = (t_{UF} - t_M)/(t_{UF} - t_F) \times 100\% \quad (3)$$

3. Results

3.1 Properties of ssDNAs selected for the study.

For better understanding of the elution behaviors of ssDNAs, we investigated ten strands, the sequences and molecular weights of which are listed in Table 1. Five were poly(dA)$_n$ that should have the lowest possibility of forming intra-molecular H-bonding than other poly(dN)$_n$. Much lower dissociation energy and longer distance were predicted for the H-bond formed from mismatching base-pairing between two A bases than that of T-T, C-C, and G-G, as reported by Otero-Navas *et al.*[37] No folding was predicted by Mfold for all poly(dA)$_n$ used in our study. The other five were capable of forming unique secondary structures as predicted by Mfold (Supporting Information, Figure S2), and referred to as aptamers in the following text. Among them was the anti-IgE aptamer with a stem-loop structure firstly reported by Weigand, *et al.* in 1996[38] and thus called the Weigand aptamer (WA) throughout the text. Its complementary strand, cWA, can form two small stem-loops separated by 4 nucleotides and was also included.

Another anti-IgE aptamer identified by the Liu group,[39] the I9-102 strand, was tested as well, the secondary structure of which has two bulges on its long stem (>13 base-pairs) and a long overhang at the 5'-end of the stem. The 60-mer aptamer against streptavidin (SA) was discovered as Aptamer 31 by Stoltenburg *et al.* in 2005.[40] The free energy of folding calculated by Mfold for these aptamers were listed in Table 1. The last aptamer was the anti-thrombin aptamer (TA) with a G-quadruplex structure, but its folding energy could not be calculated by Mfold.

To better explain their recovery and elution behaviors, we employed MD simulation for calculation of the radius of gyration, R_g, of the ssDNA in 1 mM Mg^{2+} solution and after energy minimization. MD modeling has been found to be a promising tool for studying the structural and dynamic conformation of DNAs.[41-43] To evaluate the ssDNA structure stability during the simulation, the root-mean-squared deviation (RMSD) of all heavy atoms of each ssDNA was obtained. Although we used a shorter simulation time, 1000 ps, than the 50 ns in the cited reports,[41-43] due to limitation of the total service units available to us from XSEDE, the RMSD values in the last 200 ps of the simulation typically had < 5% relative standard deviation (RSD), except for poly$(dA)_{10}$, poly$(dA)_{20}$, and the TA, the RSD of which varied between 8-10% (Supporting Information, Figure S3a). Still, the volume or bulkiness of all strands changed very little, shown by the fairly constant R_g during the simulation duration (Supporting Information, Figure S3b). The R_g values of the ssDNA in the B-form (for both the poly$(A)_n$ and aptamers) and folded structure (for aptamers only) are also shown in Table 1.

3.2 Recovery of ssDNAs in AF4.

Recovery of ssDNAs in AF4 was investigated, because the ssDNAs may escape through the AF4 channel wall made by the porous membrane. The 5- and 10-kDa MWCO regenerated cellulose ultrafiltration membranes were studied. The recovery was calculated by Equation 1 and plotted vs. Mw in Figure 1 for each poly(dA)$_n$ (square) or aptamer (circle) strand. For ssDNAs in both groups, those with Mw close to or larger 10 kDa could yield a recovery > 90% on the 10-kDa membrane (the solid curves in Fig. 1), which agrees with the membrane's MWCO definition. However, on the 5 kDa membrane, only the poly(dA)$_n$ with n ≥ 30 (Mw ≥ 9.33 kDa) and the aptamers with Mw ≥ 11 kDa could be eluted out of the channel at > 95% (the dotted plots in Fig. 1) with negligible adsorption on the membrane. Moreover, the two aptamers, the 5.66 kDa TA and the 8.99 kDa cWA, had much lower recovery than the poly(A)$_n$ with comparable Mw.

The recovery results on the 5-kDa membrane could be explained by the relative size of the ssDNA and the membrane pore. The pore diameter for the 5-kDa regenerated cellulose ultrafiltration membranes has been estimated to be around 2 nm.[44,45] Indeed, both the poly(dA) strands showing low recovery on the 5-kDa membrane are with R_g smaller than 2 nm (Table 1). As for the two aptamers with very low recovery, the TA has the R_g values for both the unfolded B-form and the folded structure below 2 nm; and the cWA could fold into a secondary structure with a R_g around 1.6 nm, although its B-form has an R_g of 2.93 nm. If the folded form is the dominant species of cWA in AF4, its recovery on the 5-kDa membrane would be low. However, the R_g values of the 8.99 kDa cWA could not explain its high recovery on the 10-kDa membrane, the estimated pore size of which is 3 nm. Unlike the case of the 5-kDa membrane, using Mw seemed to be good enough to judge the retention of the ssDNA on the 10-kDa membrane. The discrepancy could be due to the higher membrane pore density of the 5-kDa

membrane. It has been reported by Ramamurthy *et. al* that[46], the pore density for a 3.5 kDa MWCO membrane was about 3 fold higher than that of the 11 kDa membrane. Similar differences may exist between our 5- and 10-kDa membranes. Unlike ultracentrifugation, AF4 lacks a strong centrifugal force that would push most of the ssDNA molecules to go into the membrane pores. While the crossflow presses the ssDNAs against the membrane, increasing their interaction with the pores, the channel flow generates a shear force to strip them away from the pores. Nevertheless, a higher pore density of the 5-kDa membrane could increase the probability of smaller ssDNA passing through the pores. As a result, both the relative size of the pore and its pore density can influence the recovery of ssDNA in AF4. For a larger MWCO membrane (10kDa), the lowered pore density still allows for a smaller degree of small ssDNA loss compared to the smaller yet denser 5kDa membrane.

It should be noted that, recovery difference between the 10-kDa and 5-kDa membrane could be attributed by the channel permeable elements as well. We measured the recovery under a zero global crossflow, which was achieved by balancing the crossflow forth (from channel through the membrane) and the crossflow back (from reservoir to the channel through the membrane). The characteristic length of the channel and the membrane permeability and thickness, could affect these flow rates[47,48] and lead to different levels of analyte adsorption on the membrane. Therefore, these parameters should be taken into account when explaining the recovery difference.

Overall, our results point out that, the R_g or hydrodynamic radius of a DNA and the pore diameter of the membrane are more appropriate than the values of Mw and MWCO, when determining which membrane to use in AF4 for DNA separation. The 10-kDa membrane in general resulted in higher recovery of ssDNA compared to the 5-kDa membrane. Besides, its

larger pore size induces lower back pressure to the channel flow, and allows easy removal of contaminants with Mw much lower than the MWCO value, which may be needed for applications aiming to purify the target complexes. Thus, the 10-kDa membrane was chosen in our following study.

3.3 Elution of ssDNAs in AF4.

Study DNA-protein binding using separation technologies requires sufficient resolution between the free components and the complex. Thus, the Mw difference necessary to obtain adequate resolution needs to be evaluated. AF4 separates analytes based on their diffusion coefficient D. Based on the Stokes-Einstein equation, D is proportional to the reciprocal value of the molecule's hydrodynamic radius, R_H.[49-51] For a free-draining polymer like DNA and protein, the relationship of R_H and M takes the form of $R_H \sim Mw^x$. For globular proteins, x equals to 1/3. For DNA, x ranges from 1 for a rigid rod to 1/2 for a random coil. Thus, ssDNAs would have larger hydrodynamic diameters than globular proteins with the same Mw and be eluted later in AF4.

We compared the elution times of ssDNAs and proteins under identical flow conditions. The fractograms were shown in Figure 2. Cytochrome c (12.4 kDa), streptavidin (60 kDa), transferrin (76.5 kDa), and IgG (150 kDa) were chosen to cover a wide Mw range. Indeed, ssDNAs were eluted at much later times than proteins with comparable Mw. For instance, poly$(dA)_{50}$ and SA with Mw around 15-20 kDa, came out much later than the 12.4-kDa cytochrome c, but at comparable times as the 80-kDa transferrin. The average peak widths at half-peak-height ($w_{0.5h}$) of all ssDNAs and proteins tested were calculated to be 1.46 min. Therefore, to achieve a resolution larger than 1 (Note: Resolution = time difference / 1.7 $W_{0.5h}$ for

a Gaussian peak), the time difference between the target protein and the ssDNA should be larger than 2.48 min under the current flow conditions. This corresponds to proteins larger than 120 kDa if the ssDNA size is around 15-20 kDa, judged from the plots of retention time vs. molecular weight shown in the Supporting Information Figure S4.

Our result highlights that AF4 is more suitable to study ssDNA binding with larger targets, such as bulky proteins like IgG and biological complexes like microvesicles and viral particles, allowing for the ability to do aptamer selection towards these larger targets, such as what is seen in cell-SELEX. Reliable relationship between retention time and molecule hydrodynamic diameter in AF4 make it easy to predict whether sufficient resolution could be obtained for a given pair of ssDNA and protein, using calibration curves from protein and nucleic acid standards. Additionally, adjusting the inlet/crossflow ratio can help to improve resolution for the more challenging ssDNA-protein systems.

3.4 Folding of ssDNA in AF4 and MD simulation

The aptamers need to be folded in order to bind to proteins; and in return, interaction with target proteins could stabilize the folding structures. As discussed above, the diffusion coefficient D of a molecule is proportional to $1/Mw^{x}$. The value of x could indicate the flexibility of the molecule.[49,51,52] For rod-like polymers, in which the monomers are tightly connected, x could be 1; and for polymers behaving like a self-avoiding random walk chain, x could be around 0.5. This means a folded ssDNA should have an x larger than an unfolded ssDNA because of its higher rigidity. Since in AF4, retention time t is proportional to D of the analyte molecule, we evaluated the folding situation of the aptamers in AF4 by plotting the Log (retention time t) against Log (Mw) in Fig. 3. Three linear plots were observed, meaning that

they all followed the scaling law of $D \sim 1/Mw^{x}$. The $poly(dA)_n$ had a x = 0.56, while the aptamers yielded a x of 0.73. The scaling exponent measured for poly(dA) and poly(dT) strands using Small Angle X-ray Scattering (SAXS) was 0.55 in NaCl solutions, to which our result has high agreement.[53] The higher exponent observed for the aptamer molecules supports that they have lower flexibility than the $poly(dA)_n$ strands owing to the formation of secondary structures while being eluted by AF4.

Imagine that the aptamers rapidly change between the unfolded and folded forms.[54-57] The measured retention time (t_M) should then be the weighted sum of the retention times of the aptamer when it is in the unfolded (t_{UF}) and folded (t_F) structures. Since the $poly(dA)_n$ molecules have no secondary structures, they can be used as the standards to obtain the linear relationship between R_g and retention time t (Supporting Information, Figure S5). Using this relationship, the theoretical retention times of the unfolded B-form and the folded structure, t_{UF} and t_F, can be calculated from their corresponding R_g values listed in Table 1. Then, with Equation 3, the percentage of the aptamer molecules in the folded form can be obtained (Table 2). In general, aptamers with larger free energy drop per base when it folds, such as the WA and cWA, exhibited a %Folding > 70%, and those for SA and I9-102 dropped below 40%. The folded structure should be beneficial for maintaining the binding between ssDNAs and their targets.

3.5 Affinity Measurement.

After confirmation that there was negligible loss of WA on the 10-kDa membrane, we went on to measure its affinity to the target protein IgE using AF4. An incubated mixture of Alexa-488 labeled WA and IgE was injected onto the AF4, and the bound and free aptamer fractions separated and fluorescently quantified. Using the bound ratio of WA, the affinity between protein

and aptamer was determined, represented as the dissociation constant K_d. With 2 nM aptamer, steady increase of the protein-bound WA peak with protein concentration changing from 6.7 nM to 107 nM was observed (Figure 4a). Fitting the average peak area ratio and the protein concentration to Equation 2, K_d was found to be 16.3±0.5 nM. This value was about 60% higher than the reported value of 10 nM.[38] The discrepancy between the K_d value obtained by AF4 and the previously reported value can be attributed to the addition of a fluorophore to the short aptamer that may disrupt the binding interaction between aptamer and protein.

We also confirmed that there was no substantial dissociation of the WA-IgE complex during AF4, by measuring the peak area sum of the free WA and the complex peaks in four repeated injections. The representative fractograms for both samples were shown in Figure 5. With the peak area sum of the free and IgE-bound WA peaks being found to be 2565.2±256.7, the WA peak in the aptamer-only sample was 2417.3±250.5. No significant difference was revealed by the two-Sample *t*-test at the confidence interval of 0.95 (n = 4). In addition, the average symmetry of the bound peak was 0.987±0.047, a value very close to 1, indicating the peak was essentially Gaussian in nature. Overall, it was determined that the WA-IgE complex did not have any significant dissociation while traveling through the column, confirming the validity of using AF4 to obtain the K_d value of WA binding to IgE.

The WA used for obtaining results displayed in Fig. 5 carried the primer binding regions on its 3' and 5' ends. Thus, the free WA and the IgE-bound WA were collected (Supporting Information for the collection procedure) and amplified by PCR, the collecting regions being marked by the grey brackets in Fig. 5. The PCR products were analyzed by PAGE (insert in Fig. 5). Successful amplification of the WA in collections from the WA-IgE incubation sample was observed in the collection regions of 1-5, which corresponded to the free WA peak, as well as in

region 7 and 8, for the IgE-bound WA peak. No band was shown in region 6, the valley in between the free and IgE-bound WA peak, again proving no complex dissociation. Additionally, no DNA band was observed in the regions 7-8 in the WA only sample. Our experiment has proved the principle of collecting and amplifying of the protein-bound ssDNAs for downstream analysis by AF4.

The binding of SA to its target, streptavidin, was also tested. With the same buffer as used for discovery of the particular SA, the Tris-buffered saline containing 2 mM $MgCl_2$ and 1 mM $CaCl_2$, the K_d value was found to be 20.9 ± 1.8 nM (Figure 4b), ~63% lower than the reported value of ~ 56.7 nM.[40] The increased affinity between SA and streptavidin in AF4 could be attributed to the nature of the interaction – in the previously reported study, streptavidin was on a solid support. This immobilization could result in conformation difference in streptavidin, resulting in differences in affinity between the two. The low %Folding of SA did not prevent successful observation of the aptamer-protien complex in AF4, probably because the complex was formed before injection to AF4 and the binding stabilized the secondary structure of the SA, making it less unlikely to unfold. On the other hand, if a different solution, the 1□PBS with 1 mM Mg^{2+}, was used as the carrier solution for testing the SA-streptavidin interaction, a much larger K_d was found (Supporting Information Figure S6) which was about 800% different than the reported value. As can be seen, the selection of buffer conditions can greatly affect the affinity between aptamer and target. Ideally, the buffer chosen for aptamer selection should be identical to that used in a subsequent assay. AF4, due to its tolerance to a variety of salts and reagents, is an ideal platform for this selection process.

4 Conclusions

In this study, successful affinity measurement for ssDNA-protein binding was performed by AF4, after examining the recovery and elution behavior of ssDNAs. We also demonstrated that the sizes for the ssDNAs in unfolded and folded structures obtained by MD simulation can help to illustrate the folding situation based on the elution time in AF4. Both are new attempts not reported previously, based on our knowledge.

Additionally, our result demonstrates that AF4 has the freedom to use different carrier solutions with high salt concentrations and divalent cations like Mg^{2+} and Ca^{2+}. Such buffers could be necessary to produce the best binding environment for the interactive system. In particular, the divalent cations can stabilize the folding of ssDNAs and thus be beneficial for maintaining target binding. Although our AF4 measurement yielded K_d values for IgE and streptavidin with ~60% deviation from the reported values, the deviation can be attributed to the difference between the binding strengths towards free or immobilized subjects. Both reported values were from immobilized proteins, and AF4 measurement was performed over freely suspended proteins.

Moreover, our results highlight the additional advantage of AF4 over existing techniques in study of DNA-protein binding: its high sample capacity, open separation channel, and size-based separation mechanism of AF4 allow it to be employed for screening the target-binding ssDNAs from a complex mixture without immobilization of either the target. With its unique capability of analyzing objects with relatively large sizes, AF4 could be highly useful in study of binding systems not suitable for the existing techniques, for example, binding involves proteins larger than hundreds of kDa, pathogenic particles, nanoparticles larger than 50 nm, etc.

Acknowledgements

This work was supported by the National Science Foundation CAREER Grant # CHE-1111057113. J. Ashby was supported by the National Science Foundation Graduate Research Fellowship under Grant Number DGE-0813967. MD calculation used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number OCI-1053575. All simulations were done on a TACC (Texas Advanced Computing Center) Dell PowerEdge C8220 Cluster with Intel Xeon Phi coprocessors (Stampede). The authors are grateful to Dr. Chia-en Chang for her guidance on the MD work.

References.

[1] F.J. Blanco, G. Montoya, Transient DNA/RNA-protein interactions, Febs J. 278 (2011) 1643-1650.

[2] D. Vuzman, Y. Levy, Intrinsically disordered regions as affinity tuners in protein-DNA interactions, Mol. BioSyst. 8 (2012) 47-57.

[3] B. Dey, S. Thukral, S. Krishnan, M. Chakrobarty, S. Gupta, C. Manghani, V. Rani, DNA-protein interactions: methods for detection and analysis, Mol. Cell. Biochem. 365 (2012) 279-299.

[4] X.-M. Ding, X.-Y. Pan, C. Xu, H.-B. Shen, Computational prediction of DNA-protein interactions: A review, Curr. Comput.-Aided Drug Des. 6 (2010) 197-206.

[5] R. Helwa, J.D. Hoheisel, Analysis of DNA-protein interactions: from nitrocellulose filter binding assays to microarray studies, Anal. Bioanal. Chem. 398 (2010) 2551-2561.

[6] M. Snapyan, V. Sakanyan, DNA interactions with arrayed proteins, Drug Discovery Ser. 8 (2007) 313-332.

[7] J. Liu, M. You, Y. Pu, H. Liu, M. Ye, W. Tan, Recent developments in protein and cell-targeted aptamer selection and applications, Curr. Med. Chem. 18 (2011) 4117-4125.

[8] S.M. Shamah, J.M. Healy, S.T. Cload, Complex Target SELEX, Acc. Chem. Res. 41 (2008) 130-138.

[9] K.M. Bompiani, R.S. Woodruff, R.C. Becker, S.M. Nimjee, B.A. Sullenger, Antidote control of aptamer therapeutics: the road to a safer class of drug agents, Curr. Pharm. Biotechnol. 13 (2012) 1924-1934.

[10] R. Freeman, J. Girsh, I. Willner, Nucleic Acid/Quantum Dots (QDs) Hybrid Systems for Optical and Photoelectrochemical Sensing, ACS Appl. Mater. Interfaces 5 (2013) 2815-2834.

[11] T. Smuc, I.-Y. Ahn, H. Ulrich, Nucleic acid aptamers as high affinity ligands in biotechnology and biosensorics, J. Pharm. Biomed. Anal. 81-82 (2013) 210-217.

[12] P. Sundaram, H. Kurniawan, M.E. Byrne, J. Wower, Therapeutic RNA aptamers in clinical trials, Eur. J. Pharm. Sci. 48 (2013) 259-271.

[13] V. Costa, C. Angelini, I. De Feis, A. Ciccodicola, Uncovering the complexity of transcriptomes with RNA-Seq, J. Biomed. Biotechnol. (2010) doi:10.1155/2010/853916.

[14] J. Majka, C. Speck, Analysis of protein-DNA interactions using surface plasmon resonance, Adv. Biochem. Eng./Biotechnol. 104 (2007) 13-36.

[15] E. Katilius, C. Flores, N.W. Woodbury, Exploring the sequence space of a DNA aptamer using microarrays, Nucleic Acids Research 35 (2007) 7626-7635.

[16] A. Sassolas, L.J. Blum, B.D. Leca-Bouvier, Homogeneous assays using aptamers, Analyst 136 (2011) 257-274.

[17] R.K. Mosing, M.T. Bowser, Isolating aptamers using capillary electrophoresis-SELEX (CE-SELEX), Methods Mol. Biol. 535 (2009) 33-43.

[18] Z. Ronai, Y. Wang, J. Khandurina, P. Budworth, M. Sasvari-Szekely, X. Wang, A. Guttman, Transcription factor binding study by capillary zone electrophoretic mobility shift assay, Electrophoresis 24 (2003) 96-100.

[19] H. Zhang, X.-F. Li, X.C. Le, Tunable Aptamer Capillary Electrophoresis and Its Application to Protein Analysis, J. Am. Chem. Soc. 130 (2008) 34-35.

[20] H. Zhang, X.-F. Li, X.C. Le, Differentiation and detection of PDGF isomers and their receptors by tunable aptamer capillary electrophoresis, Anal. Chem. 81 (2009) 7795-7800.
[21] S.D. Mendonsa, M.T. Bowser, In Vitro Evolution of Functional DNA Using Capillary Electrophoresis, J. Am. Chem. Soc. 126 (2003) 20-21.
[22] M. Dvorak, J. Svobodova, M. Benes, B. Gas, Applicability and limitations of affinity capillary electrophoresis and vacancy affinity capillary electrophoresis methods for determination of complexation constants, Electrophoresis 34 (2013) 761-767.
[23] T. Iwaki, T. Saito, K. Yoshikawa, How are small ions involved in the compaction of DNA molecules?, Colloids Surf., B 56 (2007) 126-133.
[24] A.N. Lane, The stability of intramolecular DNA G-quadruplexes compared with other macromolecules, Biochimie 94 (2012) 277-286.
[25] A.A. Lubin, K.W. Plaxco, Folding-Based Electrochemical Biosensors: The Case for Responsive Nucleic Acid Architectures, Acc. Chem. Res. 43 (2010) 496-505.
[26] K.-G. Wahlund, L. Nilsson, Flow FFF -- Basics and Key Applications, in: S.K.R. Williams, K.D. Caldwall (Eds.), Field-Flow Fractionation in Biopolymer Analysis, Springer-Verlag/Wien, 2012, p. 1-21.
[27] S.K.R. Williams, D. Lee, Field-flow fractionation of proteins, polysaccharides, synthetic polymers, and supramolecular assemblies, J. Sep. Sci. 29 (2006) 1720-1732.
[28] S. Schachermeyer, J. Ashby, W. Zhong, Aptamer-protein binding detected by asymmetric flow field flow fractionation, J. Chromatogr. A 1295 (2013) 107-113.
[29] M.-K. Liu, C.J. Giddings, Separation and Measurement of Diffusion Coefficients of Linear and Circular DNAs by Flow Field-Flow Fractionation, Macromolecules 26 (1993) 3576-3588.
[30] J. Pollastrini, T.M. Dillon, P. Bondarenko, R.Y.T. Chou, Field flow fractionation for assessing neonatal Fc receptor and Fc gamma receptor binding to monoclonal antibodies in solution, Analytical Biochemistry 414 (2011) 88-98.
[31] J.Y. Lee, D. Choi, C. Johan, M.H. Moon, Improvement of lipoprotein separation with a guard channel prior to asymmetrical flow field-flow fractionation using fluorescence detection, Journal of Chromatography, A. 1218 (2011) 4144-4148.
[32] A. Hawe, M. Wiggenhorn, M. Van De Weert, J.H.O. Garbe, H.-C. Mahler, W. Jiskoot, Forced degradation of therapeutic proteins, Journal of Pharmaceutical Sciences 101 (2012) 895-913.
[33] D. Kang, M.H. Moon, Miniaturization of frit inlet asymmetrical flow field-flow fractionation, Anal. Chem. 76 (2004) 3851-3855.
[34] P.L. Ma, M.D. Buschmann, F.M. Winnik, Complete Physicochemical Characterization of DNA/Chitosan Complexes by Multiple Detection Using Asymmetrical Flow Field-Flow Fractionation, Anal. Chem. (Washington, DC, U. S.) 82 (2010) 9636-9643.
[35] K.G. Wahlund, A. Litzen, Application of an Asymmetrical Flow Field-Flow Fractionation Channel to the Separation and Characterization of Proteins, Plasmids, Plasmid Fragments, Polysaccharides and Unicellular Algae, Journal of Chromatography 461 (1989) 73-87.
[36] A. Litzen, K.G. Wahlund, Improved Separation Speed and Efficiency for Proteins, Nucleic-Acids and Viruses in Asymmetrical Flow Field Flow Fractionation, Journal of Chromatography 476 (1989) 413-421.

507[37] I. Otero-Navas, J.M. Seminario, Molecular electrostatic potentials of DNA base–base
508 pairing and mispairing, J. Mol. Model. (2011).
509[38] T.W. Wiegand, P.B. Williams, S.C. Dreskin, M.H. Jouvin, J.P. Kinet, D. Tasset, High-
510 affinity oligonucleotide ligands to human IgE inhibit binding to Fc epsilon receptor I, J.
511 Immunol. 157 (1996) 221-230.
512[39] K.M. Ruff, T.M. Snyder, D.R. Liu, Enhanced Functional Potential of Nucleic Acid
513 Aptamer Libraries Patterned to Increase Secondary Structure, J. Am. Chem. Soc. 132
514 (2010) 9453–9464.
515[40] R. Stoltenburg, C. Reinemann, B. Strehlitz, FluMag-SELEX as an advantageous method
516 for DNA aptamer selection, Anal. Bioanal. Chem. 383 (2005) 83-91.
517[41] S. Sen, L. Nilsson, MD simulations of homomorphous PNA, DNA, and RNA single
518 strands: characterization and comparison of conformations and dynamics, J. Am. Chem.
519 Soc. 123 (2001) 7414-7422.
520[42] J.M. Martinez, S.K. Elmroth, L. Kloo, Influence of sodium ions on the dynamics and
521 structure of single-stranded DNAoligomers: a molecular dynamics study, J. Am. Chem.
522 Soc. 123 (2001) 12279-12289.
523[43] M. Biyani, K. Nishigaki, Single-Strand Conformation Polymorphism (SSCP) of
524 Oligodeoxyribonucleotides: An Insight into Solution Structural Dynamics of DNAs
525 Provided by Gel Electrophoresis and Molecular Dynamics Simulations, J. Biochem. 138
526 (2005) 363-373.
527[44] D.A. Ladner, M. Steele, M.A. Weir, K. Hristovski, P. Westerhoff, Functionalized
528 nanoparticle interactions with polymeric membranes, Journal of hazardous materials 211-
529 212 (2012) 288-295.
530[45] J. Ren, Z. Li, F.-S. Wong, A new method for the prediction of pore size distribution and
531 MWCO of ultrafiltration membranes, Journal of Membrane Science 279 (2006) 558–569.
532[46] S. Singh, K.C. Khulbe, T. Matsuura, P. Ramamurthy, Membrane characterization by
533 solute transport and atomic force microscopy, Journal of Membrane Science 142 (1998)
534 111-127.
535[47] P. Déjardin, Determination of the functioning parameters in asymmetrical flow field-flow
536 fractionation with an exponential channel, Journal of Chromatography A 1305 (2013)
537 213-220.
538[48] M. Martin, M. Hoyos, On the no-field method for void time determination in flow field-
539 flow fractionation, Journal of Chromatography, A. 1218 (2011) 4117-4125.
540[49] A.V. Dobrynin, R.H. Colby, M. Rubinstein, Scaling Theory of Polyelectrolyte Solutions,
541 Macromolecule 28 (1996) 1859-1871.
542[50] L. He, B. Niemeyer, A Novel Correlation for Protein Diffusion Coefficients Based on
543 Molecular Weight and Radius of Gyration, Biotechnol. Prog. 19 (2003) 544-548.
544[51] A.E. Nkodo, J.M. Garnier, B. Tinland, H. Ren, C. Desruisseaux, L.C. McCormick, G.
545 Drouin, G.W. Slater, Diffusion coefficient of DNA molecules during free solution
546 electrophoresis, Electrophoresis 22 (2001) 2424-2432.
547[52] E. Stellwagen, N.C. Stellwagen, Determining the electrophoretic mobility and
548 translational diffusion coefficients of DNA molecules in free solution, Electrophoresis 23
549 (2002) 2794-2803.
550[53] Y.L. Sim Adelene, J. Lipfert, D. Herschlag, S. Doniach, Salt dependence of the radius of
551 gyration and flexibility of single-stranded DNA in solution probed by small-angle x-ray
552 scattering, Phys Rev E Stat Nonlin Soft Matter Phys 86 (2012) 021901.

[54] G.R. Bishop, J. Ren, B.C. Polander, B.D. Jeanfreau, J.O. Trent, J.B. Chaires, Energetic basis of molecular recognition in a DNA aptamer, Biophys. Chem. 126 (2007) 165-175.

[55] E. Kim, C. Yang, Y. Pak, Free-Energy Landscape of a Thrombin-Binding DNA Aptamer in Aqueous Environment, J. Chem. Theory Comput. 8 (2012) 4845-4851.

[56] M. Biyani, K. Nishigaki, Single-strand conformation polymorphism (SSCP) of oligodeoxyribonucleotides: An insight into solution structural dynamics of DNAs provided by gel electrophoresis and molecular dynamics simulations, J. Biochem. 138 (2005) 363-373.

[57] J. Zhou, S.K. Gregurick, S. Krueger, F.P. Schwarz, Conformational changes in single-strand DNA as a function of temperature by SANS, Biophys. J. 90 (2006) 544-551.

Tables

Table 1. Molecular weight, sequence, and unfolded/folded radius of gyration (R_g) of ssDNA strands investigated. "A#" described the poly(dA)n with # for the base number. Abbreviation for each aptamer was explained in the text.

Name	MW	Sequence (5'-3')	R_g (nm) (B-form)	R_g (nm) (folded structure)	ΔG (kcal/J) of folding
Poly(dA)$_{10}$	3.07 kDa	5'-(A)$_{10}$-3'	1.20	N/A	N/A
Poly(dA)$_{20}$	6.20 kDa	5'-(A)$_{20}$-3'	1.94	N/A	N/A
Poly(dA)$_{30}$	9.33 kDa	5'-(A)$_{30}$-3'	3.03	N/A	N/A
Poly(dA)$_{40}$	12.47 kDa	5'-(A)$_{40}$-3'	3.69	N/A	N/A
Poly(dA)$_{50}$	15.60 kDa	5'-(A)$_{50}$-3'	5.05	N/A	N/A
TA	5.66 kDa	5'-ATAGGTTGGTGTGGTTGG-3'	1.93	1.22	N/A
cWA	8.99 kDa	5'-CACGCCACTAGGAGGGACGGATAAACGTG-3'	2.93	1.64	-4.01
WA	11.29 kDa	5'-GGGGCACGTTTATCCGTCCCTCCTAGTGGCGTGCCCC-3'	3.64	2.21	-7.63
SA	18.31	5'-CTATACTCCACTTTGCTATTTCTCGGTTCCTTCACGCGCC GATCGCAGGCTGATGAATTG-3'	6.43	4.04	-2.49
I9-102	20.62	5'-CAGCTGACGTACGTGCATGGCAAACACACTTCATCCGTACCTTCTAGTGGGTGTGTAGCAAGCGCGC-3'	5.57	3.61	-5.18

Table 2. Determination of percent folding of aptamers. T_M, T_{UF}, and T_F were the measured retention time and the predicted times for the fully extended and completely folded structures, respectively.

Name	t_{UF} (min)	t_F (min)	t_M (min)	%Folded	ΔG (kcal/J) of folding per base
cWA	6.26	4.48	5.00	71%	-0.14
WA	7.10	5.32	5.61	84%	-0.21
SA	9.85	7.06	8.60	38%	-0.04
I9-102	9.08	7.54	8.79	31%	-0.08

Figure Captions

Figure 1. Percent recovery of poly(dA)$_n$ and aptamers under AF4 conditions (3.3 mL/min inlet flow, 3 mL/min crossflow, 0.3 mL/min axial-outlet flow). Recovery is in comparison to cross-flow free conditions (both the outlet and inlet flow rates at 0.3 mL/min), and based off of relative peak areas at 280 nm.

Figure 2. Stacked AF4 fractograms of poly(dA)$_n$ strands (top), DNA aptamers (middle), and proteins (bottom). The last 1 or 2 peaks were labeled with the Mw of the corresponding ssDNA or protein to show that the ssDNAs would be eluted much later than proteins with comparable molecular weights. Poly(dA)$_{10}$ was not visible in figure 2a due to its low recovery. AF4 separations were conducted with 3.3/3,0/0.3 mL/min inlet/cross/outlet flow rates and absorbance detection at 280 nm. The theoretical system void time was calculated to be 0.86 min with the channel dimensions and the cross/output flow rates used in the study.[28]

Figure 3. Plotting of log(retention time) as a function of log(molecular weight) for the poly(dA)$_n$ and aptamer strands, demonstrating the differences in rigidity between the unfolded and folded ssDNA.

Figure 4. Hill plots determining the affinity between IgE and the Weigand aptamer (a), or streptavidin and the streptavidin aptamer (b). Each Hill plot point is the average of three to five runs. Inset: Representative fractograms of the aptamer at a fixed concentration incubated with increasing concentrations of the target protein. The arrow indicated the change direction of the peak area along with increase in protein concentration: the area of the free aptamer peak decreased and that of the protein-bound aptamer peak increased. All fractograms were detected at 495 nm excitation and 520 nm emission wavelengths.

Figure 5. Successful collection of IgE-bound WA after AF4 separation. Collections were done on both the WA-only (dotted curve) and the WA-IgE incubated sample (solid curve). Fractions were collected every 2 minutes starting from the elution time of 6 min. Inset: PAGE images of AF4 eluates of WA only and the WA incubated with IgE. The gel band seen corresponds with WA isolated and amplified from each AF4 fraction. The gel was stained with SYBR Gold and imaged using a UV trans-illuminator.