Published in final edited form as:

J Sep Sci. 2009 May ; 32(10): 1654–1664. doi:10.1002/jssc.200900060.

A Genome-Inspired DNA Ligand for Affinity Capture of Insulin and Insulin-like Growth Factor-2

Junfeng Xiao, Jennifer A. Carter, Kimberley A. Frederick, and Linda B. McGown*

Department of Chemistry and Chemical Biology 118 Cogswell Laboratory Rensselaer Polytechnic Institute Troy, NY 12180

Abstract

The insulin-linked polymorphic region (ILPR) of the human insulin gene contains tandem repeats of similar G-rich sequences, some of which form intramolecular G-quadruplex structures *in vitro*. Previous work showed affinity binding of insulin to an intramolecular G-quadruplex formed by ILPR variant **a**. Here we report on interactions of insulin and the highly homologous insulin-like growth factor 2 (IGF-2) with ILPR variants **a**, **h** and **i**. Circular dichroism indicated intramolecular G-quadruplex formation for variants **a** and **h**. Affinity MALDI mass spectrometry and surface plasmon resonance were used to compare protein capture and binding strengths. Insulin and IGF-2 exhibited high binding affinity for variants **a** and **h** but not **i**, indicating the involvement of intramolecular G-quadruplexes. Interaction between insulin and variant **a** was unique in the appearance of two binding interactions with $K_D \sim 10^{-13}$ M and $K_D \sim 10^{-7}$ M, which was not observed for insulin with variant **h** ($K_D \sim 10^{-8}$ M) or IGF-2 with either variant (K_D 's $\sim 10^{-9}$ D M). The results provide a basis for design of DNA binding ligands for insulin and IGF-2 and support a new approach to discovery of DNA affinity binding ligands based on genome-inspired sequences rather than the traditional combinatorial selection route to aptamer discovery.

Keywords

Aptamer; Affinity Capillary Electrophoresis; MALDI-mass spectrometry; Insulin; Surface plasmon resonance

Introduction

The insulin-linked polymorphic region (ILPR) of the human insulin gene promoter region contains tandem repeats of closely related, G-rich sequences that are 14-15 bases in length [1-15]. A dozen or so ILPR variant sequences have been identified, the most common of which is variant **a** (5'-ACAG4TG4-3'). Oligonucleotides containing two or more repeats of some variants including variant **a** have been shown to form intramolecular G-quadruplex structures *in vitro* [11-15].

We previously reported that insulin exhibits high affinity binding *in vitro* through its beta chain to the intramolecular G-quadruplex formed by a 2-repeat variant **a** oligonucleotide [16]. We have since observed that IGF-2 also exhibits high affinity binding to this oligonucleotide. The availability of a DNA binding ligand to human insulin and IGF-2 offers an alternative to antibodies for isolation and purification of these proteins from biological samples. Although the binding ligands are not combinatorially derived aptamers,

*Corresponding author: Dr. Linda B. McGown Department of Chemistry, College of the Holy Cross, 1 College St., Worcester, MA 01610, mcgowl@rpi.edu, fax: +1-518-276-4887 .

they offer the same advantages such as ease of production and manipulation, stability, relatively small size, and re-usability [17-19].

In the present work, the affinity binding interactions of insulin and IGF-2 with oligonucleotides comprising 2-repeat sequences of ILPR variants **a**, **h** and **i** (ILPRa, ILPRh and ILPRi, respectively), were studied using affinity Matrix Assisted Laser Ionization Desorption Mass Spectrometry (affinity MALDI MS) and surface plasmon resonance (SPR). Circular dichroism spectroscopy (CD) was used to examine intramolecular G-quadruplex formation by the oligonucleotides. In the affinity MALDI MS experiments, the protein solution was incubated with DNA oligonucleotides that were immobilized at the MALDI probe surface as previously described [16,20,21]. Affinity capillary electrophoresis (affinity CE) was used to estimate the maximum insulin binding capacity of the ILPRa-modified surface. For SPR measurements, the protein solution flowed across immobilized oligonucleotides on two different types of SPR sensor chips: “C1” chips in which the oligonucleotides are immobilized through electrostatic interactions with a positively charged sensor surface and “SA” chips in which biotinylated oligonucleotides are attached to a streptavidinated surface. In both the affinity MALDI and SPR experiments, the experimental design in which the DNA is immobilized and the protein is unconstrained is a good model for affinity capture, isolation, and purification of proteins.

Experimental

Materials

The amine coupling kit and NSB reducer (10 mg/mL of carboxymethyl dextran sodium salt in 0.15 M NaCl containing 0.02% NaN₃) were from BIAcore (BIAcore Life Science, GE Health Care, NJ, USA). Oligonucleotides that were either thiol-modified or biotinylated at the 5'-end were synthesized by Eurogentec (San Diego, CA). The sequences of oligonucleotides ILPRa, ILPRh and ILPRi are:

ILPRa 5'-(ACA GGGG TGT GGGG)₂-3'

ILPRh 5'-(ACA GGGG TGT GGGC)₂-3'

ILPRi 5'-(ACA GGGT CCT GGGG)₂-3'

A scrambled oligonucleotide (5'-TTTTC CGACC TTCCG GGGGC CCCAG CGTCC TGCAG TG-3') that does not form a well-defined secondary structure [22] was also used as a control in some experiments. Prior to surface attachment in the MALDI and SPR experiments the oligonucleotides were heated to melt any intermolecular and intramolecular structures and then cooled to room temperature. CD spectra were then taken of the solutions to confirm formation of intramolecular G-quadruplexes by ILPRa and ILPRh as evidence by peaks at 280 nm and absence of such intramolecular structures in ILPRi and the scrambled oligonucleotide prior to immobilization. The absence of these intramolecular structures in ILPRi and the scrambled oligonucleotide allow these two oligonucleotides to serve as controls, one with identical length and very similar sequence to ILPRa and ILPRh and the other an unrelated control with different length, base composition and sequence.

All other reagents were from Sigma (St. Louis, MO), including polyethylene imine (PEI), tris(2-carboxyethyl)phosphine (TCEP), 3-aminopropyltriethoxysilane (3-APTES), sulfo-succinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate (sulpho-SMCC), sinnapinic acid (SAA), trifluoroacetic acid (TFA), recombinant human insulin, insulin-like growth factor 2 (IGF-2), human alpha-thrombin, neutravidin, streptavidin (SA), and reagents for preparation of HBS-EP buffer (10mM Hepes, 0.15 M NaCl, 3mM EDTA, 0.005% polysorbate 20, pH 7.4) and the phosphate buffer (10 mM potassium phosphate, 150 mM NaCl, pH 7.4). All buffers and solutions were prepared using distilled deionized water.

Circular dichroism measurements

Circular dichroism (CD) was measured using a Jasco Model 715 Spectropolarimeter (Jasco Inc., Easton, MD) with Peltier sample temperature control. Samples were contained in 1 cm pathlength quartz cuvettes for both spectral acquisition and thermal melt experiments. For the thermal melt experiments, temperature was increased at increments of 0.5 °C/min and CD signal was monitored at 294 nm, which corresponds to the intramolecular G-quadruplex peak of interest.

Affinity MALDI MS

Fused silica chips (20mm × 20 mm × 0.75 mm) were made by Valley Design (Westford, MA). 5'-thiol modified DNA oligonucleotides (ILPR variants or scrambled control) were covalently attached to the fused silica chips as previously described [16,20,21]. In brief, the chip surface was first cleaned and activated by rinsing with methanol, water, and sodium hydroxide. The activated chip was then immersed in a 10% solution of 3-APTES at 100 °C for 4 h, followed by addition of the heterobifunctional linker sulpho-SMCC to the 3-APTES-coated fused silica surface to create spots that were approximately 1 mm in diameter. The 5'-thiol modified oligonucleotide was treated with TCEP to give a free sulfhydryl group that was reacted with the linker at the chip surface. Finally, the DNA-modified plates were rinsed with buffer to remove excess reagents including free oligonucleotide, dried with ultra-high purity nitrogen, and stored at room temperature. As an additional precaution to ensure the removal of any non-covalently bound oligonucleotides that could otherwise form intermolecular G-quadruplex structures with the bound oligonucleotides, surfaces were treated with the acidic MALDI matrix to unfold any secondary oligonucleotide structures and rinsed with 50% acetonitrile in water to remove unbound material prior to first use.

For the protein capture experiments, a 1 µL aliquot of protein sample in the phosphate buffer was incubated on each oligonucleotide-coated spot at room temperature for 30 min. The surface was then rinsed with deionized water for 30 s to remove any weakly bound or unbound proteins as well as buffer and salts, and dried under nitrogen gas. This incubation-rinse cycle was repeated two more times, each with a new 1 µL aliquot of sample, to concentrate captured protein at the surface. A 1-1.5 µL aliquot of MALDI matrix consisting of saturated SAA with 0.3% TFA was then applied to each spot and allowed to crystallize. The chips were then mounted on a customized aluminum MALDI target and each spot was analyzed by MALDI-time-of-flight mass spectrometry (MALDI-TOF MS) using a Bruker AutoFlex II instrument (Billerica, MA). After use the spots were rinsed with 50% acetonitrile in water to remove the MALDI matrix, proteins and any concomitants. The phosphate buffer was applied to the surfaces to reconstitute them prior to reuse. Surfaces were periodically tested for protein carryover, contamination and degradation of the oligonucleotide-modified surface by MALDI-MS in the absence of applied protein sample. In no case was any carryover, contamination or degradation indicated.

Affinity Capillary Electrophoresis

Affinity capillary electrophoresis (affinity CE) experiments were performed to determine the insulin binding capacity of fused silica surfaces that were modified with ILPRa. The inner surfaces of bare fused-silica capillaries (Polymicro Technologies, Phoenix, AZ, USA) of 47 cm total length and 100 µm inner diameter (I.D.) were coated with oligonucleotides (ILPRa or a scrambled sequence control) using the same procedure described above for the MALDI plates. Following the attachment process, the capillaries were flushed with Tris buffer (20 mM Tris, 5 mM KCl, pH 7.3), sealed with parafilm and stored at 4 °C.

The experiments were performed using either a Beckman Coulter P/ACE 5000 or P/ACE MDQ CE system according to the following procedure [16]: One capillary volume of protein was loaded onto the capillary under low pressure (0.5 psi) and incubated overnight. This volume was followed by a capillary volume of wash buffer (20 mM Tris, 5 mM KCl, pH 7.3), both of which were collected. This served as the “load” collection. Three more capillary volumes of the wash buffer were passed through the capillary under low pressure to remove loosely associated proteins from the surface and collected. This served as the “wash” collection. A single capillary volume of elute buffer (20 mM Tris, 8 M urea) was then introduced and the capillary temperature, which was maintained at 25 °C up to this point, was raised to 50 °C. After 30 min, the capillary volume and two additional capillary volumes of the elute buffer were passed through the capillary, collected and combined to form the “elute” collection. Each of the collections, load, wash and elute, was diluted to a constant volume of 25 µl using its corresponding wash or elute buffer and stored at 4 °C.

In previous work, we showed that capillaries coated with the G-quadruplex DNA thrombin-binding aptamer (TBA) have a binding capacity of approximately 0.08 thrombin molecules per nm² surface area [23]. Extending these results to the present work, as a first approximation, we could predict a total capacity of 60 pmol protein for the capillaries of length and internal diameter that were used. This corresponds to an insulin concentration of 20 µM in a single capillary volume of 3.7 µl. In order to ensure an excess of protein to maximize protein capture, a solution of ten times this concentration, 200 µM, of protein was incubated in the various capillaries.

Total protein in each collected fraction was determined using fluorescence emission spectroscopy. Fluorescence emission spectra of collections from the capillaries and of human insulin standard solutions were collected using an SLM-AMINCO model 8100 Series 2 fluorescence spectrometer with an excitation wavelength of 280 nm. Samples were contained in 0.1-cm pathlength quartz cuvettes from Starna Cells. Concentrations were determined from calibration curves constructed using the standard solutions.

Determination of dissociation constants using SPR

Preparation of sensor surfaces—SPR experiments were performed using a BIAcore 3000 system with C1 and SA sensor chips (BIAcore, Biacore Life Sciences, GE Health Care). The two chips provide a comparison between different oligonucleotide immobilization processes that could result in different surface coverages and possibly different orientations that might affect the experimental results. For the C1 chip, negatively charged oligonucleotides are immobilized through electrostatic interactions with the positively charged sensor surface formed by treatment of the carboxymethylated sensor surface with PEI [24]. First, phosphate buffer was injected for 10 min to equilibrate the system. A 10-15 µL volume of 1 mg/mL PEI solution was then injected and the amount of bound PEI was manually controlled to be approximately 3000 RU. The oligonucleotide (30 µL of 20 µM oligonucleotide in HBS-EP buffer) was then injected at a flow rate of 5 µL/min. The amount of bound oligonucleotide was manually controlled to achieve a signal of 1000 RU as recommended by the BIAcore manufacturer for DNA oligonucleotide immobilization on the C1 chip.

The SA-modified chip from BIAcore is supplied with the gold surface already modified with streptavidin that is anchored to a carboxymethylated dextran layer. Biotinylated oligonucleotides can then be immobilized through affinity biotin-streptavidin binding. The biotinylated oligonucleotides (50-70 µL of 20 µM oligonucleotide in HBS-EP buffer) were injected into the flow cell at a flow rate of 5 µL/min. The amount of bound oligonucleotide was manually controlled to achieve a signal of 500 RU as recommended by the BIAcore manufacturer for DNA oligonucleotide immobilization on the SA chip.

Monitoring ILPR- protein binding interactions—A 60 μL volume of insulin or IGF-2 in phosphate buffer at concentrations ranging from 5 μM to 10 fM was injected into the flow cell on the sensor chip at a flow rate of 30 $\mu\text{L}/\text{min}$, which was chosen through experimentation to maximize response while minimizing contributions to the kinetic profiles from mass transfer (as opposed to contributions from the affinity interaction) that occur at lower flow rates [25]. The dissociation process was allowed to proceed for 2 min, followed by injection of 0.05% SDS in water for the SA chip or 50 mM NaOH with 1 M NaCl for the C1 chip to re-establish the baseline. This was followed by two consecutive 2 min injections of phosphate buffer to regenerate the surface for protein capture.

In the Biacore experiments control cells are used to correct for bulk refractive index changes and non-specific binding at the sensor surfaces. The sensorgram of the control cell is subtracted from the sensorgram of the cell of interest and data analysis is then performed on the control-corrected sensorgram. For the C1 chip, the sensorgrams for the proteins on ILPRi were indistinguishable from the sensorgrams for the scrambled sequence and so both ILPRi and the scrambled sequence were used as controls and gave similar results. The same was true for the SA chip at high ($> \text{nM}$) protein concentrations but at lower concentrations the sensorgrams for the proteins on the scrambled and ILPRi oligonucleotides were too noisy to be used as controls. We therefore used the unmodified streptavidinated surface as the control for all of the SA experiments. A further complication was the presence of significant non-specific binding to the sensor surface that needed to be addressed. Since carboxymethylated dextran at the SA chip surface is a likely source of non-specific binding, we added 1 mg/mL of dextran to the protein solution to compete with the dextran at the sensor surface. Under this competitive binding condition, we were able to greatly reduce the non-specific binding of insulin to both the control and ILPRa cell surfaces. Dextran was therefore included in the insulin and IGF-2 solutions and any contribution from the added dextran was corrected by subtracting a sensorgram for a solution containing dextran only from the sensorgrams of the protein-dextran solutions in both the ILPR and control cells.

Data analysis—BIAcore evaluation software (BIAeval Version 4.1, Biacore) was used to fit the control-corrected sensorgrams to a 1:1 Langmuir binding model based on chi-squared goodness-of-fit minimization. Binding parameters were derived from the BIAcore experiments based on the following second-order rate equations:

where A is the analyte (insulin or IGF-2), B is the ligand (ILPR oligonucleotide), k_a is the association rate constant and k_d is the dissociation rate constant. The rates of association and dissociation can be expressed as:

$$\text{Association rate: } d[AB]/dt = k_a [A][B]$$

$$\text{Dissociation rate: } -d[AB]/dt = k_d [AB]$$

At equilibrium, the association rate equals the dissociation rate and the equilibrium dissociation constant (K_D) can be calculated as:

$$K_D = [A][B] / [AB] = k_d / k_a$$

We also used Scrubber 2.0 software (BioLogic Software Pty Ltd, Cambell, Australia). The results were not significantly different and so all of the results shown here are from BIAeval. The fitting was performed with and without inclusion of mass transfer in the fitting model, however, inclusion of mass transfer did not improve the fits and in some cases, introduced systematic bias in the residuals. Therefore, only the results for fits without mass transfer are reported. Reliability of the results was based on examination of chi-squared goodness-of-fit (<1 considered acceptable), visual examination of the randomness of the fitting residuals, and consistency of k_a , k_d and K_D for different runs across the concentration ranges examined.

Results

Investigations of G-quadruplex formation

We previously confirmed G-quadruplex formation by the thrombin-binding aptamer (TBA) that was immobilized on a fused silica MALDI probe surface by investigating its capture of thrombin [20]. Since thrombin exhibits affinity binding only to the intramolecular G-quadruplex conformation of the aptamer, capture of thrombin at the fused silica surface confirmed G-quadruplex formation by the immobilized aptamer. In this work, the affinity MALDI MS probe surfaces were treated prior to first use with the acidic MALDI matrix followed by rinsing with 50% acetonitrile to ensure destruction of any intermolecular G-quadruplexes and removal of any oligonucleotides that were not covalently attached to the surface. Application of the phosphate buffer then facilitated intramolecular G-quadruplex formation by the remaining, individually immobilized single strands at the surface.

Confirmation of intramolecular G-quadruplex formation at the SPR chip surface is more difficult than for the MALDI probe surfaces due to the dynamic flow process that is used to immobilize the oligonucleotides at the sensor surface and remove proteins from the sensor surface between runs. We therefore used CD spectroscopy to investigate formation and stability of the G-quadruplex structures in solution under the various conditions used to prepare and regenerate the SPR chip surfaces. The use of CD to investigate G-quadruplex structures is well established [26-29], although not without limitations since it generally relies on empirical comparisons among spectra [30,31]. Nevertheless, using CD we were able to obtain insight into G-quadruplex formation (although not specific geometric structure) under the conditions of the SPR experiments.

The CD spectra were taken after the oligonucleotide solutions were heated to melt any three-dimensional structures and then cooled back to room temperature, which was performed before immobilization of the oligonucleotides to the MALDI and SPR probe surfaces. The CD spectra of biotinylated ILPRa, ILPRh and ILPRi in the phosphate buffer are shown in Figure 1. Spectra of the thiol-modified oligonucleotides were similar to those shown in Figure 1 for the biotinylated oligonucleotides and the spectra of the oligonucleotides in the HBS-EP buffer were essentially identical to those shown in Figure 1. For both ILPRa, which has been reported to form an intramolecular G-quadruplex [11-13], and ILPRh, the CD spectrum exhibits a peak at approximately 295 nm that is indicative of intramolecular G-quadruplex formation [26-28]. This peak is absent from the spectrum of ILPRi, which exhibits only a single peak around 260 nm. While it is not feasible to deduce the specific geometries of the oligonucleotides from these CD spectra alone, they indicate at the very least that ILPRi lacks the conformation available to ILPRa and ILPRh that results in a peak around 295 nm and is generally indicative of intramolecular G-quadruplex formation.

Thermal melt experiments were performed to compare the stability of the intramolecular G-quadruplexes of ILPRa and ILPRh under the experimental conditions used in this work. The CD signal of the ILPR oligonucleotides in the phosphate buffer was monitored at 294 nm, which corresponds to the CD peak of the intramolecular G-quadruplexes. Analysis of the thermal melting curves (Figure 2) yielded melting temperatures of 75 °C and 55 °C for the ILPRa and ILPRh intramolecular G-quadruplexes, respectively. We conclude that the ILPRa G-quadruplex is more stable than the ILPRh G-quadruplex. Additionally, the slight increase in the ILPRa curve at higher temperatures prior to melting suggests a second, thermally stable structure.

Affinity Capillary Electrophoresis

Affinity CE was used to estimate the insulin capacity of the ILPRa-coated surfaces in the affinity MALDI experiments that were prepared using the same covalent attachment scheme. Replicate capture experiments performed as described in the experimental section on each of several different, identically-prepared ILPRa-coated capillaries consistently yielded a total capture of approximately 20 pmol insulin. A capillary coated with a scrambled DNA sequence yielded at most 3 pmol insulin. After subtraction, the ILPRa protein capacity was calculated to be 0.07 insulin molecule per nm², which is in close agreement with the capacity of 0.08 thrombin molecule per nm² that was previously determined for capillaries coated with the TBA [23]. The implication is that capacity is limited by available oligonucleotide binding sites and that protein size (36 kDa for thrombin vs. 5.7 kDa for insulin) does not impose steric or crowding constraints on the protein capacity at these surface oligonucleotide coverages.

Affinity MALDI MS

Typical results are shown in Figure 3 for affinity MALDI MS capture and detection of insulin and IGF-2 at a bare (unmodified) silica surface and at surfaces coated with ILPRa, ILPRh and ILPRi. The spectra are dominated by peaks that correspond to the +1 ions of the proteins. The occasional shoulders or small peaks are attributed to impurities or matrix interference. For both insulin and IGF-2, the most capture occurs with ILPRh, followed by ILPRa. The proteins are barely detectable at the ILPRi-coated and bare surfaces. No significant peak could be distinguished from the baseline for the scrambled sequence oligonucleotide (not shown). MALDI MS results cannot be compared between different proteins since efficiencies of ionization and desorption may vary among proteins.

Spot-to-spot reproducibility is shown in Table 1 as the mean and standard deviation of the protein peak intensity for replicate runs of the same concentrations used for Figure 3. Additionally, results for a 10-fold lower concentration of insulin are shown. Based on the affinity CE result of 0.07 insulin molecule per nm², which translates to total capacities on the order of 100 fmol insulin per ILPRa-modified spot, and assuming similar surface coverage for ILPRh and ILPRi, application of 500 fmol is greater than the surface capacity of an ILPR-coated spot. Such overloading may increase non-specific binding. Ten-fold reduction of insulin concentration appears to lower non-specific binding as indicated by the disproportionately greater decrease in capture at the ILPRi and bare surfaces relative to the ILPRa and ILPRh surfaces. We can conclude from the affinity MALDI MS capture experiments that for both insulin and IGF-2, capture decreases as ILPRh > ILPRa >> ILPRi.

Determination of Dissociation Constants using SPR Measurements

Studies of experimental conditions and binding specificity for ILPRa—Binding strength measurements in SPR are sensitive to experimental conditions such as buffer system, pH, temperature and flow rate. We chose the phosphate buffer as a reasonable approximation of biological conditions that would influence protein conformation, stability

and electrostatic interactions with the DNA oligonucleotides. Initial experiments were performed at both 25 °C and 37 °C, however, a substantial increase in non-specific binding at the higher temperature led us to choose 25 °C for this work. Flow rate was evaluated and 30 $\mu\text{L}/\text{min}$ was chosen based on minimization of mass transfer contributions to the binding, as determined by the shapes of the sensorgrams and the agreement between BIAeval analysis with and without inclusion of mass transfer in the model. Protein concentration in the buffer solution was varied over a wide range (pM- μM) in order to ensure that measurements would include concentrations at or below the K_D of the binding interaction [32]. We used both SA and C1 chips but only the SA chips yielded useful sensorgrams for sub-nM concentrations of the proteins.

Figure 4 shows sensorgrams and fitting results for insulin and IGF-2 (each 250 nM) on an ILPRa flow cell in a C1 chip and for insulin at low (1-100 pM) concentrations on an ILPRa flow cell in a SA chip. The sensorgrams are much noisier at the low concentrations as would be expected due to the extremely low concentrations of the protein solutions as well as the small masses (< 10 kDa) of both protein and DNA with respect to SPR measurements. Nevertheless, the K_D s obtained from the fits to sensorgrams for very low protein concentrations showed excellent consistency, as described in a later section.

The protein binding specificity for ILPRa was studied using neutravidin and alpha-thrombin (both at 10 nM) in the C1 chip and insulin-like growth factor 1 (IGF-1) at concentrations of 850 nM-10 μM in the SA sensor chip. Neutravidin serves as a nonspecific protein binding control. Thrombin has a high affinity for the intramolecular G-quadruplex structure of the TBA and its sequence (5'-GGTTGGTGTGGTTGG-3') [33] is similar to the ILPRa sequence in that they both contain a -T-G-T- loop sequence, although the CD spectrum of TBA is unlike that of ILPRa, indicating significant conformational differences. The sequence similarities and conformational differences between TBA and ILPRa present an opportunity to examine the specificity of the ILPRa G-quadruplex conformation towards insulin and IGF-2. IGF-1 shares a high degree of structural homology with insulin and IGF-2 and is also a good test of specificity of the ILPRa G-quadruplex towards insulin and IGF-2.

The sensorgrams yielded K_D values in the μM range for neutravidin, 10 μM for thrombin, and no difference between the IGF-1 and control sensorgrams. These results confirm that the ILPRa G-quadruplex binding is highly specific for insulin and IGF-2, rather than for insulin-like proteins in general (as demonstrated by the lack of response to IGF-1) or for G-quadruplex-binding proteins in general (as demonstrated by the lack of response to thrombin), and that there is no response to other, unrelated proteins such as neutravidin.

Determination of Dissociation Constants for insulin and IGF-2 with ILPRa and ILPRh—Table 2 shows the results for binding strengths, expressed as the mean binding dissociation constant $K_D \pm$ standard deviation, for insulin and IGF-2 with ILPRa and ILPRh for the SA and C1 sensors. The K_D values for individual runs combined for both sensor types are shown as a function of protein concentration in Figure 5. Additionally, the association rate constant k_a was examined as a function of protein concentration (Table 3). The k_a values were generally independent of protein concentration, with a few exceptions when compiled over extremely broad (e.g., pM- μM) concentration ranges. This absence of systematic concentration dependence supports the use of the second-order (1:1) fitting model for the protein-DNA binding interactions.

For insulin, the K_D values were independent of sensor type. This is supported by the residuals of the fits, which were generally randomly distributed for insulin on both sensors, indicating that the C1 and SA chips both yielded accurate results and supporting the use of

the 1:1 fitting model. For insulin with ILPRa, the K_D was consistently around 10^{-7} M at high (>nM) insulin but showed a surprising, dramatic decrease to 10^{-13} M at sub-nM insulin. For insulin with ILPRh, the K_D values were consistently around $2\text{--}3 \times 10^{-8}$ M across the entire concentration range. The observation of the second, high affinity interaction of insulin with ILPRa but not ILPRh under identical experimental conditions indicates that the result is real rather than artifactual.

For IGF-2 with ILPRa, the K_D values were consistently in the $1\text{--}3 \times 10^{-9}$ M range across the entire concentration range when determined on the SA sensor. The values were, however, an order of magnitude higher when determined on the C1 sensor surface. IGF-2 with ILPRh showed a similar dependence on sensor type. The K_D on the SA sensor was 3×10^{-9} M, which is similar to the value for IGF-2 with ILPRa. The K_D on the C1 chip was an order of magnitude lower. These differences may be due to the concentration ranges of IGF-2 relative to the K_D values. SPR measurements are best performed with the concentration of binding species in solution at or below the K_D of the interaction [32]. Since the K_D for IGF-2 with both ILPR variants appears to be in the range of 10^{-9} M, it makes sense that higher IGF-2 concentrations might be more sensitive to sensor type, surface coverage and heterogeneity, etc. In these experiments, it appears that SA provided the most robust determination of K_D for IGF-2. This is supported by comparison of the fitting residuals (not shown), which showed systematic variation for IGF-2 on C1 but not on SA, indicating that SA is the preferred sensor for IGF-2.

In summary, the K_D values for insulin and IGF-2 with ILPRa and ILPRh obtained using the SA chip are: 7×10^{-7} M for insulin with ILPRa at insulin concentrations above nM, 4×10^{-13} M for insulin with ILPRa at insulin concentrations below nM, 3×10^{-8} M for insulin with ILPRh, 2×10^{-9} M for IGF-2 with ILPRa and 3×10^{-9} M for IGF-2 with ILPRh.

Discussion

Several important considerations impact the interpretation of the experimental results. First, the oligonucleotides were heated to melt any secondary structures prior to immobilization and care was taken to remove any non-immobilized oligonucleotides that might otherwise form intermolecular G-quadruplexes with the immobilized strands. Intermolecular G-quadruplex formation between adjacent immobilized strands is unlikely due to the rapid kinetics of intramolecular folding compared to intermolecular association [31] combined with the large spacing between strands that is suggested by the protein capacity of 0.07 insulin per nm^2 . This result agrees closely with the previously determined surface capacity of 0.08 thrombin per nm^2 for capillaries coated with TBA using identical attachment methodology [23], which indicates that the capacity is independent of both protein and sequence/length of the G-quadruplex oligonucleotide, further supporting the independence of the individual, immobilized strands. Therefore, it is reasonable to conclude that the only structures on the surface are those formed by intramolecular folding of the single, immobilized oligonucleotide. A second consideration is that it is unlikely that one immobilized oligonucleotide could interact with more than one protein given the small size of the proteins (5.7 kDa for insulin and 7.4 kDa for IGF-2), the highly specific nature of the interaction, and the low protein concentrations used in the SPR experiments. Together, these considerations support the validity of the 1:1 binding assumption, which is further supported by the consistency of the K_D values across wide protein concentration ranges.

The anomalously high affinity binding of insulin with ILPRa that is observed in the SPR experiments at very low insulin concentrations appears to be real rather than artifactual since it was not observed for insulin with ILPRh under identical conditions. One possible explanation is that ILPRa can assume more than one intramolecular conformation, which

would be consistent with the melting curve in Figure 2. In any case, the anomalously high affinity of insulin toward the G-quadruplex formed by the most common ILPR variant, ILPRa, suggests a unique mode of interaction that is not available to insulin with ILPRh or to IGF-2 with either variant.

Based on the SPR results using the SA chip, the relative affinities are insulin- ILPRa (interaction mode 1) >> IGF-2-ILPRh = IGF-2-ILPRa > insulin-ILPRh > insulin-ILPRa (interaction mode 2). The only difference between ILPRa and ILPRh is the single substitution of C for G, yet this sufficient to significantly lower the melting temperature of the intramolecular G-quadruplex. The lower thermal stability of the ILPRh G-quadruplex may increase its structural flexibility, allowing it to better adapt to the IGF-2 and insulin molecules. This would be consistent with results of other *in vitro* studies, in which it was found that increasing temperature promoted interactions between proteins and intramolecular G-quadruplex DNA, presumably by increasing the flexibility of the G-quadruplex to accommodate the protein [34,35]; this is also seen in ability of aptamers to bolster “less than perfect fits” to proteins through electrostatic, hydrogen bonding and hydrophobic interactions [19]. The anomalously high affinity interaction between insulin monomer and ILPRa, on the other hand, may arise from a more “perfect” fit that benefits from greater rigidity of a second, highly stable ILPRa G-quadruplex conformation suggested by the increase in the thermal melting curve for ILPRa at high temperatures. Further studies are needed to determine the G-quadruplex structures and identify the specific sites of interactions and structures of their resulting complexes with insulin and IGF-2.

Conclusion

The studies of the ILPR sequences provide a basis for design of DNA binding ligands for insulin and IGF-2. Indeed, the discovery that G-quadruplex-forming sequences from the ILPR bind with insulin [16] was recently used as the basis for further selection of insulin-binding aptamers [36]. More generally, the genome-inspired oligonucleotide sequences offer an alternative to the combinatorial route to aptamer selection to protein targets that actually reverses the selection process: in aptamer selection, an artificial, combinatorial pool of oligonucleotides is selected for binding to a specific protein target, whereas the present work suggests the use of a single oligonucleotide with a genomic sequence to capture potential target proteins from natural candidate pools such as serum or cell lysates. Discovery of new DNA binding ligands will augment existing aptamer ligands that have been successfully applied in protein analysis [37-39]. The K_D s reported here are at least as strong as those reported for combinatorially selected aptamers (generally in the nM- μ M range [19]), and even stronger in the case of the insulin monomer with ILPRa. Furthermore, the use of naturally occurring genomic sequences may increase the likelihood of discovering binding interactions that are both analytically and biologically significant.

Acknowledgments

Funding National Institutes of Health (1 R21 DK 70762).

References

1. Owerbach D, Nerup J. Diabetes. 1982; 31:275–277. [PubMed: 6295855]
2. Owerbach D, Gabbay KH. Diabetes. 1996; 45:544–551. [PubMed: 8621001]
3. Owerbach D, Gabbay KH. Diabetes. 1993; 42:1708–1714. [PubMed: 8243816]
4. Bell GI, Horita S, Karam JH. Diabetes. 1984; 33:176–183. [PubMed: 6363172]
5. Bennett ST, Lucassen AM, Gough SC, Powell EE, Undlien DE, Pritchard LE, Merriman ME, Kawaguchi Y, Dronsfield MJ, Pociot F, et al. Nat Genet. 1995; 9:284–292. [PubMed: 7773291]

6. Kennedy GC, German MS, Rutter WJ. *Nature Genetics*. 1995; 9:293–298. [PubMed: 7773292]
7. Pugliese A, Miceli D. *Diabetes Metab Res Rev*. 2002; 18:13–25. [PubMed: 11921414]
8. Stead JD, Buard J, Todd JA, Jeffreys AJ. *Hum Mol Genet*. 2000; 9:2929–2935. [PubMed: 11115836]
9. Bell GI, Karam JH, Rutter WJ. *Proc Natl Acad Sci USA*. 1981; 78:5759–5763. [PubMed: 6272317]
10. Ullrich A, Dull TJ, Gray A, Philips JA, Peter S. *Nucl Acids Res*. 1982; 10:2225–2240. [PubMed: 6283472]
11. Hammond-Kossack MC, Docherty K. *FEBS Lett*. 1992; 301:79–82. [PubMed: 1451791]
12. Hammond-Kossack MC, Dobrinski B, Lurz R, Docherty K, Kilpatrick MW. *Nucl Acids Res*. 1992; 20:231–236. [PubMed: 1741248]
13. Hammond-Kossack MC, Kilpatrick MW, Docherty K. *J Mol Endocrinol*. 1992; 9:221–225. [PubMed: 1476609]
14. Castasti P, Chen X, Moyzis RK, Bradbury EM, Gupta G. *J Mol Biol*. 1996; 264:534–545. [PubMed: 8969303]
15. Lew A, Rutter WJ, Kennedy GC. *Proc Natl Acad Sci USA*. 2000; 97:12508–12512. [PubMed: 11070077]
16. Connor AC, Frederick KA, Morgan EJ, McGown LB. *J Am Chem Soc*. 2006; 128:4986–4991. [PubMed: 16608332]
17. McGown LB, Joseph MJ, Pitner JB, Vonk GP, Linn CP. *Anal Chem*. 1995; 67:663A–668A.
18. Mukhopadhyay R. *Anal Chem*. 2005; 77:115A–118A.
19. Hermann T, Patel DJ. *Science*. 2000; 287:820–825. [PubMed: 10657289]
20. Dick LW Jr, McGown LB. *Anal Chem*. 2004; 76:3037–3041. [PubMed: 15167780]
21. Cole JR, Dick LW Jr, Morgan EJ, McGown LB. *Anal Chem*. 2007; 79:273–279. [PubMed: 17194151]
22. Mendonsa SD, Bowser MT. *Anal Chem*. 2004; 76:5387–92. [PubMed: 15362896]
23. Connor AC, McGown LB. *J Chromatogr A*. 2006; 1111:115–119. [PubMed: 16569569]
24. Cui X, Yang F, Li A, Yang X. *Anal Biochem*. 2005; 342:173–175. [PubMed: 15958197]
25. Karlsson R, Falt A. *J Immunol Methods*. 1997; 200:121–133. [PubMed: 9005951]
26. Burge S, Parkinson GN, Hazel P, Todd AK, Neidle S. *Nucl Acids Res*. 2006; 34:5402–5415. [PubMed: 17012276]
27. Paramasivan S, Rujan I, Bolton PH. *Methods*. 2007; 43:324–331. [PubMed: 17967702]
28. Kypr J, Vorlickova M. *Biopolymers (Biospectroscopy)*. 2002; 67:275–277. [PubMed: 12012446]
29. Hardin CC, Perry AG, White K. *Biopolymers*. 2001; 56:147–194. [PubMed: 11745110]
30. Gray DM, Wen J-D, Gray CW, Repges R, Repges C, Raabe G, Fleischhauer J. *Chirality*. 2008; 20:431–440. [PubMed: 17853398]
31. Lane AN, Chaires JB, Gray RD, Trent JO. *Nucl Acids Res*. 2008 Advance access published August 21, 2008.
32. Blake RC 2nd, Pavlov AR, Blake DA. *Anal Biochem*. 1999; 272:123–134. [PubMed: 10415080]
33. Macaya RF, Schultze P, Smith FW, Roe JA, Feigon J. *Proc Natl Acad Sci USA*. 1993; 90:3745–3749. [PubMed: 8475124]
34. Vo TU, McGown LB. *Electrophoresis*. 2004; 25:1230–1236. [PubMed: 15174042]
35. Vo TU, McGown LB. *Electrophoresis*. 2006; 27:749–756. [PubMed: 16421942]
36. Yoshida W, Mochizuki E, Takase M, Hasegawa H, Morita Y, Yamazaki H, Sode K, Ikebukuro K. *Biosens Bioelectron*. Jun 18.2008 [Epub ahead of print].
37. Clark SL, Remcho VT. *Electrophoresis*. 2002; 23:1335–40. [PubMed: 12007135]
38. Mascini, M., editor. *Aptamers in Bioanalysis*. John Wiley & Sons; 2008.
39. Berezovski MV, Lechmann M, Musheev MU, Mak TW, Krylov SN. *J Am Chem Soc*. 2008; 130:9137–43. [PubMed: 18558676]

Figure 1. CD spectra of 10 μ M ILPRh (top), ILPRa (middle) and ILPRi (bottom) in phosphate buffer after heating to melt secondary structures and cooling to 25 $^{\circ}$ C.

Figure 2. Thermal melting curves for 10 μ M ILPRa (top) and 10 μ M ILPRh (bottom) in phosphate buffer (10 mM potassium phosphate, 150 mM NaCl, pH 7.4). The CD signal was monitored at 294 nm.

Figure 3.

Capture and detection of insulin (170 nM, 500 fmol total applied) and IGF-2 (100 nM, 300 fmol total applied) at different MALDI probe surfaces including untreated (bare) fused silica and fused silica coated with ILPRa, ILPRh and ILPRi. Peaks in all cases are insulin⁺¹ peak (5805 m/z) or IGF-2⁺¹ peak (7400 m/z). For each spot, three successive incubation/rinse cycles were performed using 1 μ L aliquot of protein solution per 30 min incubation cycle followed by application of MALDI matrix and MALDI TOF MS analysis.

Figure 4. SPR sensorgrams after reference subtraction (noisy line) and fitting results to 1:1 model (smooth solid line). Top: insulin and IGF-2 (250 nM) on an ILPRa flow cell in a C1 sensor chip. Bottom: insulin at low concentrations on an ILPRa flow cell in a SA sensor chip.

Figure 5. K_D vs. protein concentration for individual determinations of insulin and IGF-2 with ILPRa and ILPRh surfaces. Results from both SA and C1 sensor chips are included. K_D s were obtained using 1:1 fitting models.

Table 1

Spot-to-spot reproducibility of Insulin and IGF-2 captured from increasingly dilute solutions at different MALDI probe surfaces, expressed as mean MALDI-TOF MS peak intensity $\pm$ one standard deviation (number of spots “n” in parentheses).^a

	Mean peak intensity $\pm$ 1 standard deviation for (n) spots		
	Insulin		IGF-2
Surface			
	170 nM (500 fmol)	17 nM (50 fmol)	100 nM (300 fmol)
h	878 $\pm$ 422 (6)	512 $\pm$ 197 (4)	591 $\pm$ 337 (3)
a	348 $\pm$ 187 (6)	292 $\pm$ 32 (3)	156 $\pm$ 209 (4)
i	168 $\pm$ 127 (6)	36 $\pm$ 10 (3)	Not detected
bare	67 $\pm$ 23 (4)	Not detected	Not detected

^aFor each spot, 3 repeats of the incubation/rinse cycle using 1 μ L of the protein of indicated concentration per cycle, for a total application of the amount shown in parentheses, were performed before application of MALDI matrix and MALDI-MS analysis.

Dissociation constants K_D determined using SA and C1 sensor chips for insulin and IGF-2 with ILPRa and ILPRh recovered from a 1:1 fitting model.

Table 2

Variant	Protein	Concentration Range	Sensor Type	Mean $K_D \pm$ standard deviation (M)	n ^a
ILPRa	Insulin	850-5000 nM	SA	$(7 \pm 3) \times 10^{-7}$	4
		0.1-100 pM	SA	$(4 \pm 3) \times 10^{-13}$	5
		250-1000 nM	C1	$(2.1 \pm 0.9) \times 10^{-7}$	4
	IGF-2	45-900 nM	SA	$(1.3 \pm 0.5) \times 10^{-9}$	6
		1-100 pM	SA	$(2 \pm 1) \times 10^{-9}$	5
ILPRh	Insulin	100-750 nM	C1	$(3 \pm 2) \times 10^{-8}$	7
		1-1000 pM	SA	$(3 \pm 3) \times 10^{-8}$	4
		500 nM	C1	7.0×10^{-8}	2
	IGF-2	0.1-100 pM	SA	$(3.0 \pm 0.8) \times 10^{-9}$	4
		0.5-50 nM	C1	$(2.1 \pm 0.6) \times 10^{-10}$	5

^a n = Number of replicate runs.

Association rate constants k_a ($M^{-1}s^{-1}$) for individual determinations of insulin and IGF-2 at different concentrations with ILPRa and ILPRh surfaces of SA and C1 sensor chips. Results were obtained using 1:1 fitting models to reference-subtracted sensorgrams.

Table 3

Insulin-ILPRa SA Insulin concentration	k_a	Insulin-ILPRa C1 Insulin concentration	k_a	Insulin-ILPRh SA Insulin concentration	k_a	Insulin-ILPRh C1 Insulin concentration	k_a
100 pM	4.55×10^5	1000 nM	2.18×10^5	1 nM	6.48×10^4	500 nM	9.46×10^4
100 pM	3.52×10^5	1000 nM	1.78×10^5	1 nM	4.69×10^4	500 nM	4.37×10^4
10 pM	8.21×10^5	500 nM	2.32×10^5	100 pM	1.87×10^5		
1 pM	3.13×10^5	250 nM	1.46×10^5	100 pM	2.23×10^5		
				100 pM	1.77×10^5		
				10 pM	5.24×10^5		
				1 pM	2.33×10^6		
				1 pM	8.47×10^5		

IGF-2-ILPRa C1 IGF-2 concentration	k_a	IGF-2-ILPRa SA IGF-2 concentration	k_a	IGF-2-ILPRh C1 IGF-2 concentration	k_a	IGF-2-ILPRh SA IGF-2 concentration	k_a
750 nM	2.42×10^4	900 nM	1.74×10^4	50 nM	1.35×10^4	100 pM	1.65×10^5
500 nM	4.87×10^4	630 nM	1.06×10^5	25 nM	2.25×10^4	10 pM	7.85×10^5
500 nM	5.76×10^4	450 nM	9.27×10^4	12.5 nM	2.00×10^4	1 pM	1.57×10^6
250 nM	3.05×10^5	360 nM	8.81×10^4	5 nM	3.53×10^4	100 fM	1.63×10^5
100 nM	1.84×10^5	180 nM	1.61×10^5				
100 nM	1.72×10^5	100 pM	1.84×10^5				
100 nM	1.83×10^5	100 pM	2.27×10^5				
25 nM	2.71×10^4	10 pM	1.09×10^6				
		10 pM	6.87×10^5				
		1 pM	2.42×10^5				