Protein-DNA Affinity Binding Studies Using Mass Spectrometry and Optical Spectroscopy

by

JunFeng Xiao

An Abstract of a Thesis Submitted to the Graduate

Faculty of Rensselaer Polytechnic Institute

in Partial Fulfillment of the

Requirements for the degree of

DOCTOR OF PHILOSOPHY

Major Subject: Chemistry

The original of the complete thesis is on file
In the Rensselaer Polytechnic Institute Library

Examining Committee:

Linda B. McGown, Thesis Adviser

Robert J. Linhardt, Member

Lee Ligon, Member

Mark D. Platt, Member

Rensselaer Polytechnic Institute
Troy, New York

July, 2009
(For Graduation August, 2009)

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, which is the predominant sequence. The present work expands the scope of the research to include interactions of insulin and the highly homologous insulin-like growth factor-2 (IGF-2) with 2-repeat sequences of ILPR variants a, h, and i (ILPRa, ILPRh and ILPRi, respectively). Circular dichroism spectroscopy (CD) indicates that ILPRa and ILPRh form intramolecular G-quadruplexes that are not observed for ILPRi. Affinity MALDI MS was used to compare capture of the proteins by the variants that were immobilized at MALDI probe surfaces and Surface Plasmon Resonance (SPR) was used to determine binding constants and kinetics of the interactions. Insulin and IGF-2 exhibited high binding affinity for ILPRa and ILPRh but not ILPRi, indicating the involvement of intramolecular G-quadruplexes. A particularly interesting result was the observation of a dramatic increase in binding strength, but only for insulin with ILPRa, when insulin concentration was decreased below nM. This result suggests that insulin may have more than one mode of interaction with ILPRa.

Studies were then performed to see if complexes formed between ILPR oligonucleotides ILPRh and ILPRi and the proteins could be detected by MALDI MS. The use of covalent cross-linking to preserve the complexes was investigated. Although the cross-linking experiments were not successful, it was possible to detect significant complexation of both insulin and IGF-2 with ILPRh but not ILPRi. The intramolecular

G-quadruplex structure of ILPRh is the likely source of the interactions since it is absent in ILPRi.

Two different methods were used to determine the location of the binding sites in insulin and IGF-2 for the G-quadruplexes formed by ILPRa and ILPRh. Both methods employed affinity capture of peptides produced by enzymatic digestion of insulin and IGF-2. The sequence VCG(N)RGF was found to be common to all of the insulin and IGF-2 peptides captured by ILPRa and ILPRh, with the exception of those peptides that contained disulfide bonds and those peptides captured by both ILPRh and ILPRi in the affinity MALDI MS experiments. The common sequence is believed to play a key role in the affinity binding of insulin and IGF-2 with the G-quadruplexes formed by ILPRa and ILPRh. Results also showed the importance of arginine in the common sequence in the binding interactions. Knowledge of the affinity binding site in insulin and IGF-2 for G-quadruplexes formed by the ILPR oligonucleotides offers new insight into these interactions that could be used to design affinity binding reagents as well as increasing our understanding of the specificity of G-quadruplex-protein interactions.

The last part of this dissertation describes the investigation of Indium Tin Oxide (ITO)-coated glass slides instead of fused silica plates that were used up to this point in the research as the substrate for immobilization of ILPR oligonucleotides for affinity MALDI MS of proteins. ITO-coated glass slides were shown to provide better signal reproducibility, peak resolution, detectability and hydrophobicity compared to the fused silica plates. The results of the protein capture studies showed that protein detectability was improved by two orders of magnitude for insulin and one order of magnitude for IGF-2. Capture of the insulin β-chain from DTT-treated commercial human serum that

was diluted 10^4-fold in order to reduce interference from non-specific protein interactions in affinity MALDI MS was demonstrated using the ITO-coated plates that were modified with the ILPR oligonucleotides.

This research provides a basis for design of DNA binding ligands for insulin and IGF-2 and supports 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. It has potential applications for the capture and identification of low abundance proteins from biological samples such as serum in applications such as disease biomarker discovery and detection. The results provide fundamental insight into the interactions of insulin and IGF-2 with G-quadruplex-forming sequences from the human ILPR that may lead to new discoveries about regulation of insulin and IGF-2 gene expression and genetic susceptibility to Type 1 Diabetes. In a broader context, these results have important implications for potential roles of genomic G-quadruplexes and their interactions with proteins in biological systems.