Use and improvement of the systematic evolution of
ligands by exponential enrichment (SELEX) process to
develop aptamers for mycotoxins and other small
molecules

by

Maureen McKeague
BSc (Hon), Carleton University 2007

A thesis submitted to the Faculty of Graduate and Postdoctoral
Affairs in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Chemistry

Carleton University
Ottawa, Ontario

© 2012

Maureen McKeague

Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 978-0-494-93685-6

Our file Notre référence

ISBN: 978-0-494-93685-6

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

Artificially-selected single-stranded oligonucleotides that bind to targets with high affinity and selectivity are known as aptamers. This relatively new technology is emerging as a viable molecular recognition element for use in biosensor platforms. Similar to antibodies, aptamers have many therapeutic, research and diagnostic applications. Aptamers, however, display several advantages over their antibody counterparts including chemical stability and a lower cost of production. Additionally, aptamer generation is performed *in vitro* using the Systematic Evolution of Ligands by EXponential enrichment (SELEX) process. This procedure screens large combinatorial DNA libraries for desired target binding through a series of iterative selection rounds. Despite their potential for diagnostics and therapeutics, relatively few aptamers exist that bind to small molecules. Small molecules are of interest due to their many roles in biology. Mycotoxins are small molecules of particular interest as they are frequent contaminants of food and feed worldwide. Chronic exposure leads to a variety of toxic endpoints including cancer. This research focuses on the development of aptamers to some mycotoxins and other medically relevant small molecules. Research was conducted to improve the success of SELEX experiments by modifying various steps in the process. For example, increasingly stringent selection conditions have been included in the SELEX procedure. Additionally, computational methods have been used to generate DNA libraries for SELEX that display increased structural diversity. These complex libraries have been demonstrated to lead to the selection of aptamers with improved binding properties. With these modified techniques, we have developed novel aptamers for L-homocysteine and adenosine-5'-triphosphate (ATP) as well as for the mycotoxins

ochratoxin A (OTA) and fumonisin B₁ (FB₁). We are also in the process of selecting aptamers for deoxynivalenol (DON) and have described the procedure for incorporating an OTA aptamer in solid-phase extraction columns for rapid clean-up of OTA from wheat samples. Finally, an aptamer database (Aptamer Base) was developed that captures detailed experimental conditions about SELEX and the resulting aptamers. With the information entered into the database, potential relationships can be evaluated between aptamer affinity and their selection conditions. This information will help streamline SELEX protocols for specific targets of interest and may open up the possibility of rational aptamer design.

Acknowledgements

Firstly, I would like extend my deepest gratitude to my supervisor, Prof. Maria DeRosa. I was fortunate to have met her in my undergraduate degree where I completely fell in love with the research. From the day I met her, she has always been encouraging and enthusiastic. Her passion for science, teaching and research led to me joining her research group and pursuing a PhD. She gave me the freedom and confidence to approach problems independently, yet was always available for advice.

I am very grateful for the guidance of the many of the professors and staff of our Department of Chemistry who have cultivated my interest in science and encouraged me along the way including Elena Munteanu, Prof. Jeff Smith and Prof. Jeff Manthorpe. In particular, I am very thankful for Prof. Miller's assistance and encouragement in achieving my goals. Additionally, thank you to Prof. Bob Burk for supporting my many science extracurricular activities and taking the time to give me career advice. I am also indebted to Prof. Sean Barry for giving me my first research opportunity in my undergraduate degree.

Much of the work presented in this thesis would not have been possible without the contributions of the bioinformatics group and CNR Italy. Thank you for being supportive and patient. Additionally, I had the pleasure of bringing the Aptamer Base to life with Jose Cruz-Toledo. Thank you for the many software/ programming tutorials and our many discussions about future research possibilities and the benefits of open source software.

A special thanks to the CUSERT gang who have kept me in touch with the rest of the world and for not giving me too much grief for missing out on countless events due to

lab work (in particular Andrew Dunbar, Sara Hosseini, Chris Donaldson, David Lubberts, Jackie Merlin, Bradley Nazar, Courtney Paxton, Andrew Whyte and Lindsay Wingham-Smith).

I would like to extend a large thank you to every branch of the DeRosa Lab Family Tree. You have all made day-to-day lab life enjoyable, as well as meetings and classes entertaining. In particular, I thank Erin McConnell, Elyse Bernard, Mike Beking, Tariq Francis, Amanda Giamberardino, Yasir Sultan, Emily Mastronardi and Ryan Walsh. I would also like to thank Amanda Foster, Charlotte Bradley, Yasmine Miguel, Kayla Hill and countless other students that I entrusted with so many of my projects. I would not have been able to manage as many projects without their help and dedication.

I am also very appreciative for the career, research and life advice from Dr. Alex Wahba. Thank you to the Miller lab, my other favourite lab, for many entertaining memories. Last, a special thank you to the honorary member of the DeRosa family, Monica Gill, for several needed late-night discussions and for proof reading my thesis.

Finally, I must single out my parents: Lorraine Hope and Dale McKeague. They have provided me with every possible opportunity for success. Raising me as PhD students themselves, they instilled in me good values and hard work. They have always been overwhelmingly encouraging and supportive in my efforts. They have been excellent role models my whole life, in particular my mom who completed a PhD thesis despite the most challenging of obstacles. Finally, I dedicate this thesis to my brother, Reed, you are the real *boy-who-lived* and for that I will forever be grateful. I am so proud of you and wish you the best with whichever career path you choose.

Table of Contents

Abstract.....	ii
Acknowledgements	iv
List of Tables	xiv
List of Figures.....	xvii
List of Abbreviations	xxiii
Chapter 1. Introduction.....	1
1.1 Aptamers	2
1.1.1 Comparison of aptamers with other affinity probes	3
1.1.2 Analytical applications of aptamers	4
1.2 Systematic evolution of ligands by exponential enrichment (SELEX)	7
1.2.1 Modifications to the SELEX process	10
1.3 SELEX targets	13
1.4 Mycotoxins	16
1.5 Aptamer binding affinity	18
1.5.1 Methods for measuring binding affinity	21
1.6 Thesis objectives	28
Chapter 2. Development of aptamer solid-phase extraction columns for the extraction and detection of ochratoxin A.....	30
2.0 Statement of contributions.....	31
2.1 Introduction	31
2.1.1 Ochratoxin A	31
2.1.2 The first mycotoxin aptamer	35
2.1.3 Aptamers in food safety	36
2.1.4 Chapter objectives	37

2.2	Experimental	37
2.2.1	Materials.....	37
2.2.2	Aptamer synthesis	38
2.2.3	Confirming binding of aptamer 1.12.2.....	41
2.2.4	Standards preparation.....	45
2.2.5	Preparation of the aptamersorbent.....	46
2.2.6	Binding efficiency of SPE columns	47
2.2.7	OTA determination in wheat samples	47
2.2.8	Comparison with IMA columns	48
2.2.9	Recovery experiments	49
2.2.10	Regeneration and re-usability of SPE columns	49
2.2.11	HPLC conditions	49
2.2.12	Statistical analysis	50
2.3	Results and Discussion.....	50
2.3.1	Binding affinity of aptamer 1.12.2	50
2.3.2	Binding efficiency of aptamer-SPE columns	56
2.3.3	Optimization of clean-up time and sample handling.....	61
2.3.4	Statistical overview of the aptamer-SPE clean-up	64
2.3.5	Comparison of aptamer-SPE column clean-up with IMA column clean-up.....	65
2.3.6	Re-usability of aptamer-SPE columns	67
2.4	Conclusion.....	68
Chapter 3.	Selection and binding assessment of new OTA aptamers	71
3.0	Statement of contributions.....	72
3.1	Introduction	72
3.1.1	Multiple aptamers for the same target	72

3.1.2 Nanoparticle binding assay	75
3.1.3 Chapter objectives	78
3.2 Experimental	78
3.2.1 Materials.....	78
3.2.2 Preparation of OTA resin	79
3.2.3 OTA SELEX experiments.....	79
3.2.4 Screening of binding affinity.....	81
3.2.5 Dissociation constant (K_d) determination using gold nanoparticles.....	85
3.3 Results and Discussion.....	86
3.3.1 Selection of OTA aptamers	86
3.3.2 Screening of the binding affinity	90
3.3.3 AuNP dissociation constant (K_d) experiment.....	93
3.4 Conclusion.....	96
Chapter 4. Development of a high affinity aptamer for fumonisin B₁	97
4.0 Statement of contributions.....	98
4.1 Introduction	98
4.1.1 Fumonisin B ₁	98
4.1.2 Current detection methods.....	100
4.1.3 FluMag SELEX.....	101
4.1.4 Chapter objectives	104
4.2 Experimental	105
4.2.1 Materials.....	105
4.2.2 DNA library and primer synthesis.....	106
4.2.3 Derivatization of magnetic beads	107
4.2.4 Fumonisin B ₁ SELEX experiments.....	108

4.2.5	Cloning and sequencing	111
4.2.6	Binding assays.....	112
4.2.7	Dissociation constant (K_d) determination using affinity chromatography	113
4.2.8	Confirmation of aptamer binding in solution	114
4.3	Results and Discussion.....	115
4.3.1	Fumonisin B ₁ immobilization	115
4.3.2	Selection of FB ₁ aptamers	116
4.3.3	Binding assays.....	125
4.3.4	Measuring aptamer affinity in solution	129
4.4	Conclusions	131
Chapter 5.	Development of homocysteine aptamers.....	134
5.0	Statement of contributions.....	135
5.1	Introduction	135
5.1.1	Homocysteine.....	135
5.1.2	Current detection technology for homocysteine.....	139
5.1.3	Chapter objectives	141
5.2	Experimental	141
5.2.1	Materials.....	141
5.2.2	Preparation of L-homocysteine	142
5.2.3	Preparation of homocysteine modified agarose	143
5.2.4	DNA library and primer synthesis.....	144
5.2.5	Homocysteine SELEX experiments	145
5.2.6	Cloning and sequencing	147
5.2.7	Assessing aptamer binding affinity and selectivity	149
5.2.8	Preparation of a homocysteine biosensor	152

5.3 Results and Discussion	153
5.3.1 Preparation of L-homocysteine	153
5.3.2 Agarose column preparation	156
5.3.3 Selection of homocysteine aptamers	157
5.3.4 Dissociation constant (K_d) experiments.....	163
5.3.5 Selectivity.....	167
5.3.6 Nanoparticle sensor	168
5.4 Conclusions	173
Chapter 6. Improvement of SELEX libraries	174
6.0 Statement of contributions.....	175
6.1 Introduction	175
6.1.1 Aptamer secondary structure.....	175
6.1.2 SELEX libraries	178
6.1.3 Genetic algorithms	181
6.1.4 Chapter objectives.....	183
6.2 Experimental	184
6.2.1 Computational method: random filtering.....	184
6.2.2 Computational method: genetic filtering.....	185
6.2.3 Reagents	187
6.2.4 DNA library and primer synthesis.....	187
6.2.5 ATP SELEX experiments	188
6.2.6 Cloning and sequencing	190
6.2.7 Dissociation constant (K_d) experiments.....	191
6.3 Results and Discussion.....	192
6.3.1 Increasing the number of complex structures in SELEX libraries	192

6.3.2	Designing a SELEX library with uniform structural distribution	199
6.3.3	Selection of ATP aptamers using the improved library design	204
6.3.4	Choice of ViennaRNA secondary structure folding package.....	207
6.3.5	Comparison to recent work	208
6.4	Conclusions	210
Chapter 7.	A SELEX database	211
7.0	Statement of contributions.....	212
7.1	Introduction	212
7.1.1	Current aptamer databases.....	212
7.1.2	Freebase.....	213
7.1.3	Chapter objectives	215
7.2	Experimental	216
7.2.1	Data model	216
7.2.2	Curating the aptamer literature.....	220
7.2.3	Querying the database	221
7.3	Results and Discussion.....	222
7.3.1	Summary of data in the Aptamer Base.....	222
7.3.2	Accessing the database.....	223
7.3.3	Mining the database.....	224
7.3.4	Utility of the database.....	224
7.3.5	Comparison with other databases.....	225
7.3.6	SELEX queries	227
7.4	Conclusions	232
Chapter 8.	Contributions to knowledge	233

Appendices.....	237
Appendix A Towards the selection of aptamers for deoxynivalenol.....	237
A.1 Introduction	237
Deoxynivalenol	237
Current detection methods	240
Section objectives.....	241
A.2 Experimental	241
Materials.....	241
Deacetylation of DON.....	242
DON conjugation	243
DNA library and primer synthesis	244
DON SELEX experiments	245
A.3 Results and Discussion.....	248
Preparation of DON derivatized column.....	248
Aptamer selection progress	250
A.4 Conclusions	256
Appendix B Binding isotherms	257
B.1 Ochratoxin A aptamer 1.12.2 (Chapter 2).....	257
B.2 Ochratoxin A aptamer A08 (Chapter 3).....	260
B.3 Fumonisin B ₁ aptamers (Chapter 4)	262
B.4 Homocysteine aptamers (Chapter 5)	268
B.5 ATP aptamers (Chapter 6).....	271
Appendix C Algorithms (Chapter 6)	273
C.1 The <i>CountJunction</i> algorithm for determining the number of junctions in a nucleic acid secondary structure	273

Appendix D Supplemental information for the Aptamer Base (Chapter 7)	274
D.1 Curating the database	274
D.2 Distribution of DNA and RNA K_d values found in the Aptamer Base	283
References	284

List of Tables

Table 1-1: A list of modifications to the SELEX process and their descriptions (adapted from Stoltenberg et al. ⁵²).....	11
Table 1-2: A complete listing of DNA aptamers that have been confirmed to bind to small molecules. The dissociation constant, a measure of binding affinity, is included as well as the year of aptamer development.....	15
Table 1-3: Methods for determining binding affinity of aptamers to their targets.	22
Table 2-1: Examples of maximum tolerated levels of OTA in foodstuffs by country and product (adapted from Duarte et al. ¹⁴⁹).....	33
Table 2-2: Proposed Health Canada maximum limits for the presence of ochratoxin A in various food commodities (adapted from Health Canada 2009 ¹⁶¹).	34
Table 2-3: Dissociation constant (K_d) results of aptamer 1.12.2 with OTA from 7 trials using the one-point equilibrium dialysis assay.	52
Table 2-4: Dissociation constant (K_d) values for aptamer 1.12.2 with OTA estimated using various binding affinity methods.....	56
Table 2-5: OTA recovery from aptamer-SPE columns, using different sorbent volume and column size. Recoveries were determined by running standard solutions of different OTA concentrations through the column. Values in each column correspond to the mean of three measurements $\pm$ 1 standard deviation. n.d.= none detected	58
Table 2-6: Recovery and precision (relative standard deviation, RSD) results of binding experiments (n = 3 measurements) performed by passing different amounts of OTA through aptamer-SPE columns packed with 300 μ L aptamersorbent in the presence of 15% methanol.	60
Table 2-7: Solvent elution flow rate and total clean-up time from 1 mL and 3 mL SPE columns packed with different volumes of aptamersorbent. Each value corresponds to the mean of three measurements $\pm$ 1 standard deviation.	62
Table 2-8 Recovery and precision (relative standard deviation, RSD) results of ochratoxin A (OTA) from spiked wheat (n= 3 measurements) after clean-up on solid-phase-extraction (SPE) columns packed with 300 μ L oligosorbent (24 nmol DNA aptamer)...	65
Table 3-1: Buffer composition used in the equilibrium dialysis binding screen for OTA aptamers.	83

Table 3-2: The putative aptamer sequences (of the N ₄₀ region) cloned after round 15 of the SELEX experiment. The number of clones present is indicated in the column on the right.	87
Table 3-3: Results of the one-point equilibrium dialysis method for screening OTA putative aptamer sequences A08, B08, D07 and E08. The K_d in the various buffer conditions is reported. The average K_d of all 3 trials with the calculated standard deviation is detailed. As the trials using the SB with Triton resulted in significant bubbles, the average is also reported excluding this initial value.....	91
Table 3-4: Results of the one-point equilibrium dialysis method for putative OTA aptamer sequences A08, and D07 compared to aptamer 1.12.2. The K_d in various buffer conditions is reported. The average K_d of all the trials with the calculated standard deviation is detailed.	91
Table 4-1: Summary of fumonisin B ₁ selection round conditions.	109
Table 4-2: Full length sequences obtained after 18 selection rounds for the target fumonisin B ₁	123
Table 4-3: Average K_d values obtained from 3 separate affinity chromatography experiments for FB ₁ sequences obtained from SELEX. Reported error was found using standard deviation.	128
Table 5-1: The various elution methods used to recover putative aptamer sequences from the L-Hcy immobilized sepharose, by selection round.....	146
Table 5-2: Putative aptamer sequences for L-homocysteine obtained after 8 rounds of selection.	163
Table 5-3: Summary of methods attempted to elucidate the binding affinity of L-Hcy aptamers. A brief description explaining why these methods were ineffective is included.	165
Table 5-4: Dissociation constants (K_d) for sequences Hcy8 and Hcy10 obtained using affinity chromatography on L-homocysteine immobilized sepharose or magnetic beads. (All binding isotherms can be found in Appendix B).....	167
Table 6-1: Average structural distribution for a typical SELEX library determined using a different numbers of mutations (variances in brackets).....	193
Table 6-2: <i>In silico</i> and <i>in vitro</i> structural distribution for the <i>RFLibraryA</i> library with primer binding sites GGAAGAGATGGCGAC and AGCTGATCCTGATGG shown in Figure 6-6. The <i>in vitro</i> values were generated by cloning and sequencing of the library following solid-phase synthesis (17 sequences were obtained). The sequences are in Table 6-3.	196

Table 6-3: DNA sequences obtained experimentally from the <i>RFLibraryA</i> and the corresponding number of junctions as predicted by ViennaRNA.	197
Table 6-4: <i>In silico</i> structural distribution of the <i>GFLibrary 1</i> obtained after 571 generations of the genetic filtering method. The structural distributions prior to starting SELEX and after 8 rounds of selection using ATP as the target are also given.....	202
Table 6-5: DNA sequences obtained experimentally from the initial <i>GFLibrary1</i> and the corresponding number of junctions predicted by ViennaRNA.....	203
Table 6-6: DNA sequences obtained experimentally from the <i>GFLibrary1</i> following 8 rounds of SELEX using ATP as the target. The corresponding number of junctions predicted by ViennaRNA are listed.	206
Table 7-1: Comparison of aptamer affinity based on target types (as of March 1, 2012). The mean K_d for each set of target types was compared. The 2-sample t test was used to evaluate whether there was a significant difference between these results (initial results from P Q. Ho ³⁰⁵).	232

List of Figures

Figure 1-1: The thrombin aptamer is able to form a specific binding interaction with its target, thrombin (blue). Image generated using Discovery Studio 2.5 by Accelrys.	2
Figure 1-2: The Systematic Evolution of Ligands by EXponential enrichment (SELEX) process. Beginning with a large library of DNA, iterative cycles of target incubation, library partitioning and amplification are performed to select aptamers.	9
Figure 1-3: Breakdown of the various types of aptamer targets (number of occurrences) that have been selected between 1990 and 2007. This list was generated using aptamer.freebase.com on February 20 th , 2012.	13
Figure 1-4: The equilibrium dialysis apparatus consists of two compartments separated by a semi-permeable membrane. The target (purple) in this case is allowed to move between the chambers, whereas the aptamer (green) is unable to pass through the membrane.	25
Figure 1-5: The process of using ultrafiltration tubes to determine aptamer binding: the target (purple) and aptamer (green) are allowed to incubate. Separation occurs using centrifugation, allowing only unbound target to pass through the filter membrane. The filtrate is measured to determine the concentration of unbound target.	25
Figure 1-6: Using fluorescence polarization to quantify aptamer binding. A) The fluorescent target (purple) experiences rotational diffusion, resulting in the emission of depolarized light. B) When the target is bound by the aptamer (green), the complex experiences less rotational diffusion, leaving the majority of the emitted light polarized.	27
Figure 2-1: Structure of ochratoxin A.	32
Figure 2-2: Total fluorescence (CPS) of OTA measured (ex: 375nm, em: 430nm) before dialysis and in the loading/ receiving chambers following 48 hours of equilibration. Based on fluorescence, no OTA adhered to the membrane and equilibrium was reached.	51
Figure 2-3: Binding isotherm for aptamer 1.12.2 with OTA obtained using equilibrium dialysis. DNA aptamer concentration was varied with a constant concentration of OTA (200 nM). Following 48 hours of dialysis, fluorescence was determined and the fraction of OTA was calculated and plotted against DNA concentration. A theoretical curve was fit using Microsoft Excel Solver. Inset: Plot of the experimentally determined OTA fraction vs. the theoretical calculated OTA fraction.	53
Figure 2-4: Total fluorescence (CPS) of OTA measured (ex: 375nm, em: 430nm) loaded into the 3kDa cut-off Amicon Ultra Centrifuge devices and in the filtrate following a 10	

minute centrifugation at 14,000 g. Based on fluorescence, >90% of the OTA adhered to the filter and did not pass through the filter.	55
Figure 2-5: Linear regression between OTA amount loaded on and recovered from aptamer-SPE column resulting from binding experiments (n = 3 measurements).	61
Figure 2-6: OTA percent recovery from spiked wheat samples (5ppb) obtained using manual extraction or orbital shaking at varying times between 5 and 60 minutes.	64
Figure 2-7: Comparison of ochratoxin A (OTA) content determined by solid-phase-extraction (SPE) column-clean-up versus immunoaffinity (IMA) column clean-up. Each point corresponds to the mean of three measurements. Results are reported without correcting for recovery.....	66
Figure 2-8: Chromatograms of an ochratoxin A (OTA)-free wheat sample (lower line) and a wheat sample naturally contaminated with 1.5ng/g OTA (upper line). Sample extracts were cleaned-up on SPE columns containing 300 μ L oligosorbent (24 nmol DNA aptamer).....	67
Figure 2-9: Normalized OTA recovery for 1 mL SPE columns after 2 to 5 re-uses.	68
Figure 3-1: Structures of adenosyl containing compounds – adenosine monophosphate (AMP), nicotinamide adenine dinucleotide (NAD ⁺) and S-adenosyl methionine (SAM).	73
Figure 3-2: Gold nanoparticle aptamer assay for target detection. A: Sodium citrate stabilized AuNPs undergo salt-induced aggregation, resulting in a blue-purple colour. B: Single-stranded DNA aptamer coils around the AuNP surface, preventing salt-induced aggregation, therefore the colloid solution remains red. C: Upon addition of the aptamer's target, the aptamer folds into a rigid structure around the target, preventing the adsorption of the ssDNA on the AuNP surface. Therefore, salt-induced aggregation returns and a blue-purple colour is observed.....	77
Figure 3-3: Secondary structure prediction for the full length putative OTA aptamer clones. Image generated by Mfold.....	89
Figure 3-4: OTA binding assay of aptamer A08 using AuNPs. Increasing concentrations of OTA resulted in aggregation of the AuNP, and therefore a change in colour from red to blue.....	94
Figure 3-5: UV-Visible absorbance spectra of the OTA binding assay of aptamer A08 using AuNPs. Increasing concentrations of OTA resulted in a decrease in the absorbance at 525nm, and an increase in the absorbance at 705nm. Absorbance at 300 – 350 nm results from high concentrations of OTA.	94
Figure 4-1: Structure of fumonisin B ₁	98

Figure 4-2: The partitioning of the DNA library using target-immobilized magnetic beads.	103
Figure 4-3: PCR performed using FluMag SELEX modified primers. The forward primer is modified with 5' fluorescein (yellow). The reverse primer is modified with a spacer followed by dA ₂₀ . Insert: The end products of PCR product separated via PAGE. Left: DNA absorbance of each product. Right: Fluorescence emission of the fluorescein-modified putative aptamers.	105
Figure 4-4: HPLC chromatographs comparing the unbound FB ₁ remaining in the supernatant following reaction with: A) no EDC present in the reaction B) EDC present as the coupling agent. Excess OPA retention time 4 minutes; OPA derivatized FB ₁ retention time 6 minutes.	117
Figure 4-5: Percent recovery of binding ssDNA to the FB ₁ immobilized magnetic beads after each selection round. With any observable significant increase or plateau in percent recovery, increasingly stringent selection conditions were implemented the following round (see experimental and Table 4-1).	120
Figure 4-6: Secondary structure prediction for the full length FB ₁ aptamer clones. Image generated by Mfold.	124
Figure 4-7: Binding assays of the individual aptamer sequences against FB ₁ beads compared to the unmodified beads (labeled as control beads). The percent of DNA eluted with buffer at 90°C is in blue, the DNA eluted using 7 M urea at 90°C is in grey. The control library (unenriched library) was tested on the FB ₁ modified beads as a comparison.	126
Figure 4-8: 20 % denaturing PAGE separation of products from the DMS/piperidine protection assay. All lanes contained 1 nmol of 6-FAM labeled FB ₁ 39 aptamer. Lane 1 was mixed with buffer, lanes 2, 4-6 were incubated with 1 mM FB ₁ and lane 3 was mixed with 1 mM glutamic acid.	132
Figure 5-1: Structure of L-homocysteine.	135
Figure 5-2: Homocysteine biochemical pathways: The transmethylation pathway and transulfuration pathway of homocysteine is flanked by the folate dependent pathway on the right and the folate independent pathway on the left. S-adenosylmethionine (SAM); methyltransferase (MTs); S-adenosylhomocysteine (SAH); cystathionine β synthase (CBS); γ-cystathionase (CGL); betaine homocysteine methyl transferase (BHMT); dimethylglycine (DMG); 5-methyltetrahydrofolate (5-CH ₃ -THF); methionine synthase (MS); tetrahydrofolate (THF).	136
Figure 5-3: UV-Visible absorbance spectra of L-homocysteine thiolactone HCl (absorbance at 238nm) and of L-homocysteine at 0.05g/L.	154

Figure 5-4: ^1H NMR (300 MHz) of L-homocysteine thiolactone HCl (bottom) and L-homocysteine produced via base hydrolysis (top).	155
Figure 5-5: UV-Visible absorbance spectra of Ellman's reagent (DTNB) and DTNB with L-homocysteine (Hcy). DTNB reacts with the thiol group, producing a yellow coloured thiolate ion that absorbs at 412nm. Inset: the yellow colour observed from mixing DTNB with Hcy compared to the colourless DTNB solution.	157
Figure 5-6: Percent recovery of binding ssDNA to the L-Hcy immobilized sepharose after each selection round. Beginning at round 6, counter-selections with L-cysteine and L-methionine were performed prior to positive selection.	161
Figure 5-7: Secondary structures of putative L-Hcy aptamer sequences generated using Mfold.....	164
Figure 5-8: Binding isotherms using the AuNP assay. Target concentration (either L-homocysteine or L-cysteine) was varied with a constant concentration of aptamer, control ssDNA or minimal aptamer. An apparent K_d was elucidated from the aggregation of the AuNP, measured as the ratio of absorbances at 660 nm and 524nm. A theoretical curve was fit using Microsoft Excel Solver.....	170
Figure 5-9: Derived calibration curve of the adsorption ratio (A_{660}/A_{524}) at the lower concentration range of target (either L-homocysteine or L-cysteine) with a constant concentration of aptamer or control ssDNA.	171
Figure 5-10: Derived calibration curve of the adsorption ratio (A_{660}/A_{524}) for high concentrations of target (either L-homocysteine or L-cysteine) with a constant concentration of aptamer or control ssDNA.	172
Figure 6-1: An example of a 5-way junction where there are five connected helices (S1 to S5).	176
Figure 6-2: Structural distribution (percent of sequences displaying a given junction order: either one-way, two-way, three-way, four-way, or five-way junction structures) of the 2793 known aptamers from the Ellington Aptamer Database. Secondary structures were determined using ViennaRNA and the CountJunctions algorithm was used to determine the number of junctions.	178
Figure 6-3: Average structural distribution (percent of sequences displaying a given junction order: either one-way, two-way, three-way, four-way, or five-way junction structures) of 1 million sequences generated <i>in silico</i> for random SELEX libraries containing 15-nt primer sites at both ends. Secondary structures were determined using ViennaRNA and the CountJunctions algorithm was used to determine the number of junctions.....	179
Figure 6-4: Process of a genetic algorithm.	182

Figure 6-5: Comparison of average structural distribution (percent of sequences displaying a given junction number: either one-way, two-way, three-way, four-way, or five-way junction structures) of 1 million sequences for typical random nucleic acid libraries (light colours) and libraries generated by the random filtering process (dark colours), with (red shades) and without (blue shades) primer binding sites. The random filtering method greatly increases the percentage of high-complex structures such as 5-way junctions by 10% without primers and 6% with primers. 195

Figure 6-6: Sample secondary structure of the *RFLibraryA* DNA sequence including primer sites GGAAGAGATGGCGAC and AGCTGATCCTGATGG designed to generate the highest percentage of 5-way junction sequences (31.13%). The sequence has 19 single-stranded positions, excluding primer sites (noted as N). As a starting library for SELEX experiments, these 19 single-stranded positions are considered as random (N) and the other positions are considered as fixed. Image generated by Mfold. 196

Figure 6-7: Sample sequences and representative secondary structures for *RFLibraryA*, shown in Figure 6-6. A) a 5-way junction structure B) a 4-way junction structure C) a 3-way junction structure and D) a 2-way junction structure. 199

Figure 6-8: The best fitness value observed for each generation of genetic filtering for the 100 nucleotide library containing the primer binding sites ATACCAGCTTATTCAATT and AGATAGTAAGTGCAATCT. Inset: Sample secondary structure of the DNA library sequence (GFLibrary1). This 100nt library, containing 24 random nucleotide positions, was synthesized and used in selection experiments against ATP. Image generated by Mfold. 201

Figure 6-9: Percent recovery of binding ssDNA to the ATP agarose after each selection round. 205

Figure 6-10: Comparison of predicted structural distributions of RNAfold (blue), GTfold (red) and MC-Fold (green) for the same 1000 random sequences of 100 nucleotides each (containing no fixed primer sites). The structural distributions generated by RNAfold and GTfold are almost identical. However, MC-Fold predicted considerably more complex structures. 209

Figure 7-1: Basic type-relationship map used by the Aptamer Base to describe SELEX experiments. The Interaction experiment type has an interaction as the outcome. Each interaction has at least 2 interactors as participants (for SELEX experiments these are aptamers and the aptamer target). Rectangles denote types and arrows represent properties between topics. Overlapping rectangles represent the multiple types associated with a topic. 216

Figure 7-2: Basic type-relationship map used by the Aptamer Base to describe the details of SELEX experiments. Rectangles denote types and arrows represent properties between topics. Overlapping rectangles represent the multiple types associated with a topic. ... 217

Figure 7-3: Basic type-relationship map used by the Aptamer Base to describe the details of affinity experiments. The affinity experiment results in dissociation constant and either 'confirms' or 'disputes' an interaction. Rectangles denote types and arrows represent properties between topics. Overlapping rectangles represent the multiple types associated with a topic..... 218

Figure 7-4: Basic type-relation map describing minimal aptamers. Minimal aptamers are described in the Aptamer Base by creating a new interaction for each individual minimal aptamer..... 219

Figure 7-5: Summary of data (as of March 1, 2012) found in the Aptamer Base. (A) Distribution of the aptamer types (B) Distribution of the aptamer targets. 222

Figure 7-6: Histogram displaying the frequency of RNA aptamers (as of March 1, 2012) of given sequence lengths (within a 5 nucleotide range). The mean RNA aptamer length is 44 nucleotides. Image generated using SigmaXL..... 229

Figure 7-7: Histogram displaying the frequency of DNA aptamers (as of March 1, 2012) of given sequence lengths (within a 5 nucleotide range). The mean DNA aptamer length is 58 nucleotides. Image generated using SigmaXL..... 230

List of Abbreviations

3-ADON – 3-acetylated deoxynivalenol

6-FAM – 6-carboxyfluorescein phosphoramidite

Ado – adenosyl

AMP – adenosine monophosphate

AFM – atomic force microscopy

AOAC – Association of Official Agricultural Chemists

ATP – adenosine triphosphate

API – application programming interface

a.u. – arbitrary units

AuNP – gold nanoparticle

BB – binding buffer

BPA – bisphenol A

CPG – controlled pore glass

CPS – counts per second

CD – circular dichroism

CDI – 1,1'-carbonyldiimidazole

CE – capillary electrophoresis

DADPA – Coupling gel (immobilized diaminodipropylamine)

DDBJ – DNA Database of Japan

DDGS – dried Distillers' grains with solubles

DNA – deoxyribonucleic acid

DMSO – dimethyl sulfoxide

DMS(O)MT – 4,4'-dimethoxy-4'-methylsulfonyltrityl

DMT – dimethoxytrityl

DON – deoxynivalenol

dsDNA – double stranded deoxyribonucleic acid

DTNB – 5,5'-dithiobis-2-nitrobenzoic acid

EDC– 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride

EDTA – ethylenediaminetetraacetic acid

ELISA – enzyme-linked immunosorbent assay

EMBL – European Molecular Biology Laboratory

EMSA – electrophoretic mobility shift assay

ESI-MS – electrospray ionization mass spectrometry

FB₁– fumonisin B₁

FLD – fluorescence detection

FAO – Food and Agriculture Organization

FDA – US Food and Drug Administration

GC/MS – gas chromatography mass spectrometry

HEPES – 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

Hcy – homocysteine

HPLC – high performance liquid chromatography

IARC– International Agency for Research on Cancer

IMA – immunoaffinity

ITC – isothermal titration calorimetry

JECFA– Joint FAO/WHO Expert Committee on Food Additives

JSON – JavaScript object notation

K_a – association constant

K_d – dissociation constant

LC/MS – liquid chromatography mass spectrometry

L-HcyT – L-homocysteine thiolactone hydrochloride

LOD – limit of detection

LOQ – limit of quantification

MIP – molecular imprinted polymer

MWCO – molecular weight cut-off

MQL – meta query language

NAD^+ – nicotinamide adenine dinucleotide

NHS – N-hydroxysuccinimide

NMR – nuclear magnetic resonance

nt – nucleotide

OTA – ochratoxin A

PAGE – poly(acrylamide) gel electrophoresis

PCR – polymerase chain reaction

PBS – phosphate buffered saline

PLA – proximity ligation assay

PMTDI – provisional maximum tolerable daily intake

QCM – quartz crystal microbalance

QD – quantum dot

RCA – rolling circle amplification

REST – representational state transfer

RNA – ribonucleic acid

RP-HPLC – reversed-phase high performance liquid chromatography

RSD – relative standard deviation

SAH – S-adenosyl homocysteine

SAM – S-adenosyl methionine

SB – selection buffer

SPE – solid-phase extraction

SPR – surface plasmon resonance

ssDNA – single-stranded deoxyribonucleic acid

SELEX – systematic evolution of ligands by exponential enrichment

TBE – Tris(hydroxymethyl)aminomethane/borate/ethylenediaminetetraacetic acid buffer

TEMED – N,N,N',N'-tetramethylethylenediamine

TLC – thin layer chromatography

TNB – 5-thio-2-nitrobenzoic acid

Tris – Tris(hydroxymethyl)aminomethane

URL – universal resource locator

UV-Vis – Ultraviolet-visible spectrophotometry

VEGF – vascular endothelial growth factor

X-gal – 5-bromo-4-chloro-indolyl- β -D-galactopyranoside

XML – extensible markup language

Chapter 1. Introduction

1.1 Aptamers

Historically, nucleic acids were associated with the storage and genetic coding of information and have long been thought to be less complex than proteins¹. However, like proteins, nucleic acids are able to fold into intricate tertiary structures that have the potential to perform a variety of functions including gene-regulation, catalytic activity and ligand-binding². Interest in functional nucleic acids was prompted by the number of discoveries of non-coding ribonucleic acids (RNAs) displaying catalytic or binding properties².

Two decades ago, several researchers revolutionized molecular recognition by developing synthetic RNA motifs that bound specifically to molecular ligands³⁻⁵. These RNA structures, called aptamers, were selected using an *in vitro* selection procedure called Systematic Evolution of Ligand by EXponential enrichment (SELEX)³. Like antibodies, these synthetically derived molecular recognition probes were found to be selective and able to bind to their target with high affinity.

Figure 1-1: The thrombin aptamer is able to form a specific binding interaction with its target, thrombin (blue). Image generated using Discovery Studio 2.5 by Accelrys.

1.1.1 Comparison of aptamers with other affinity probes

There is a growing need for rapid, robust and inexpensive methods for sensing and diagnostic purposes⁶. As molecular recognition is the cornerstone of sensing, there has been increased focus on the development of new molecular recognition probes for sensing applications⁷. While antibodies remain the standard in molecular recognition since the 1950's, the relatively new technology of aptamers offers several advantages⁸.

Firstly, the *in vitro* aptamer selection process allows a greater control over aptamer binding conditions. Non-physiological conditions such as high salt concentration or non-physiological temperatures and pH can be used in successful selections⁹. Due to the stability of the oligonucleotide phosphodiester backbone, aptamers exhibit improved stability over their antibody counterparts. In particular, aptamers can be reversibly denatured by changing the conditions that were used for selection. For example, a change in pH, temperature, ionic strength or use of denaturants irreversibly denatures antibodies, while aptamers simply unfold. The aptamer structure can then regain functionality upon return of the original binding conditions⁶. Due to the nucleic acid nature of aptamers, they bind to complementary nucleic acids as well as their targets, which can be exploited in sensing schemes or as “antidotes” *in vivo*. Aptamer generation has been achieved for a wide variety of targets including small molecules¹⁰, proteins¹¹, viruses¹² and whole cells¹³. Unlike antibodies, aptamers can also be generated for targets that are toxic as well as for targets that do not elicit an immune response *in vivo*⁸.

Once selected, aptamers are manufactured using the well-established oligonucleotide automated chemical solid-phase synthesis procedure^{14, 15}. The accuracy

and reproducibility of this procedure allows for a relative ease in producing aptamers at large scales, with very little batch-to-batch variation in activity¹⁶. Additionally, aptamer sequences can be modified with reporter molecules at select nucleotide positions throughout this solid-phase synthesis without affecting the functionality or activity of the aptamer^{17, 18}.

Aptamers offer advantages over other synthetically created molecular recognition systems such as molecular imprinted polymers (MIPs). As artificial recognition systems, MIPs offer several similar advantages as aptamers compared to antibodies¹⁹. For example, MIP synthesis can be simple and cheap. MIPs are also stable for years and are not damaged by changes in heat and pH²⁰. However, MIPs are typically limited to binding small molecules and display high cross-reactivity¹⁹.

1.1.2 Analytical applications of aptamers

Since the development of aptamers in 1990, many applications using aptamers have been reported²¹. The majority of the articles published since 1990 focus on aptamer applications instead of designing new aptamers²¹. One of the most common uses of aptamers is for optical sensors^{22, 23}. Aptamers frequently undergo large conformational changes upon binding to their target ligands²⁴. Introduction of dyes into conformationally labile regions of aptamers allows for a change in optical characteristics such as a shift in colour, change in intensity or polarization upon ligand binding²⁵. Frequently, quantum dots (QDs) or fluorescent – quenching pairs such as fluorescein and dabsyl can be included to observe either signal-on or signal-off behavior in the presence of a target²⁶. Gold nanoparticles (AuNPs) can be conjugated to aptamer systems causing

a colour change upon binding²⁵. Finally, dye displacement assays using fluorophores have also been reported in many papers²⁷. These dyes bind to an aptamer, typically in a non-specific fashion, and can be displaced by an unlabeled target. Reporter molecules such as ethidium bromide and crystal violet are highly fluorescent when they form complexes with the aptamer, however, the fluorescence drastically decreases as the target displaces the reporter. Several of these optical applications have moved from in-solution spectroscopic assays to more practical lateral flow assays²⁸.

Regardless of the optical applications, the analytical applications of aptamers are rapidly moving toward electrochemical strategies. Electrochemical approaches have the potential of high sensitivity, low cost and the ability to readily miniaturize the instrumentation²⁵. With electrochemical detection, a change in aptamer conformation or a binding competition can also be exploited. By covalently conjugating electroactive reporters to either the aptamer or to a complementary strand, binding can be detected using either signal-on or signal-off strategies, depending on the redox reporter. Signal transduction for electrochemical methods has been achieved using a variety of electroactive reporters such as methylene blue^{29, 30}, ferrocene³¹, ruthenium complexes³², and $\text{Fe}(\text{CN})_6^{4-/3-}$ ³³.

Although conformational signal transduction is often exploited in analytical applications, not all aptamers undergo conformational changes required for detection. However, since aptamers can be readily synthesized and conjugated to surfaces, mass-sensitive detection methods performed on or at a surface can be used. For example, aptamer-based surface plasmon resonance (SPR)³⁴ and quartz crystal microbalance (QCM) sensors³⁵ have been developed. While these methods are typically label-free,

analysis is related to a change in mass. These applications are generally only useful for large aptamer targets such as proteins and cells.

Finally, several aptamer applications are reliant on the same techniques used to detect other nucleic acids. The most reliable and accepted techniques make use of a variety of nucleic acid amplification systems. For example, rolling-circle amplification (RCA)³⁶, the polymerase chain reaction (PCR)³⁷ and the proximity ligation assay (PLA)³⁸ have been used for analytical aptamer detection purposes.

While aptamer development in this thesis was primarily for analytical purposes, it is important to note that the utility of aptamers stretches far beyond analytical applications. Several researchers are interested in the therapeutic application of aptamers. Work towards using aptamers as drugs¹⁶, target delivery systems³⁹ and imaging agents⁴⁰ is constantly being investigated. Currently, there are 9 aptamers that have undergone clinical trials and several other aptamers that have been selected for therapeutic purposes¹⁶. The best known aptamer therapeutic is the drug Macugen®. This aptamer is injected intravitreally to improve symptomatic decline in visual acuity typical of age-related macular degeneration. The modified RNA aptamer binds to all isoforms of human vascular endothelial growth factor (VEGF), blocking its ability to interact with the VEGF receptors⁴¹. While this drug is no longer in use due to a competing treatment, it exemplifies the potential of aptamers as therapeutics¹⁶.

Originally, aptamer attention was focused on the use of RNA. Due to its role in natural systems, it was long considered that only RNA could fold into functional motifs. However, only two years following the original *in vitro* selection experiments, several groups reported DNA aptamers that were capable of binding specific molecular targets

with high affinity⁴². The lack of the 2' hydroxyl of the DNA sugar results in an increased chemical and biological stability compared to RNA⁴³, a useful characteristic for almost all applications of aptamers and in particular for *in vitro* sensing applications. Despite this finding, researchers continued to preferentially select RNA aptamers over DNA, resulting in few available DNA aptamers. As a result, DNA aptamers for the targets ATP, thrombin, platelet-derived growth factor and immunoglobulin E have often been used to demonstrate the utility of these analytical systems²⁵. These have been valuable for proof-of-concept purpose. However, these applications can only reach their full potential if they can be used with new, economically relevant aptamer targets. I was particularly interested in developing new aptamers that could be employed for such purposes.

1.2 Systematic evolution of ligands by exponential enrichment (SELEX)

The idea of *in vitro* evolution was first reported in the 1960s when Dr. Sol Spiegelman at the University of Illinois observed that in a cell-free system, the RNA genome of the Q β bacteriophage could be evolved during replication to form RNAs that were more efficiently copied by the viral replicase⁴⁴. Later, Dr. Spiegelman was able to evolve sequences for other traits such as resistance to ethidium bromide⁴⁵. However, the true potential of *in vitro* evolution was not achieved until several decades later. In that time, several biological advances occurred including the invention of polymerase chain reaction (PCR), the isolation of reverse transcriptase and the ability to generate long oligonucleotides containing random nucleotide regions using solid-phase synthesis. Equipped with these modern techniques, *in vitro* selection was able to realize its full potential. In 1990, three separate groups reported *in vitro* selection and evolution of

nucleic acids⁴⁶. Tuerk and Gold³ used the term SELEX for their process of selecting RNA ligands against T4 DNA polymerase; Ellington and Szostak⁴ performed *in vitro* selection to select RNA ligands (for which they coined the term ‘aptamers’) against various organic dyes. Finally, Robertson and Joyce⁵ evolved the *Tetrahymena* self-splicing intron to carry out a DNA cleavage reaction.

Since its invention, several researchers have performed SELEX to isolate nucleic acids with a wide variety of functions. While several modifications have been proposed, the general SELEX process remains the same (Figure 1-2). Typically, SELEX begins with an initial library (often referred to as a “pool”) consisting of random nucleic acid sequences (either RNA or DNA depending on the nature of the research). Typical SELEX libraries consist of 30-80 random nucleotide positions flanked by primer binding sites necessary for PCR steps⁴⁷. The library is then incubated with the target of interest and several washing steps are employed to remove non-functional sequences. For small molecule targets, the target is typically immobilized onto a solid-support matrix such as agarose to permit partitioning⁴⁸. In the selection of aptamers, the sequences that display affinity to the target of interest remain bound to the immobilized target. The next steps in SELEX include the elution of the binding sequences from the target and the polymerase chain reaction (PCR) amplification of those binding sequences to yield an enriched library for subsequent, more stringent, selection rounds¹⁰. As the interactions that lead to molecular recognition between the binding sequences and the target are non-covalent in nature, mild conditions can be used to separate the two species. Elution using high concentrations of the target molecule, high-temperatures or hydrogen bond-disrupting

compounds such as urea can be performed. The strength of the molecular interactions within the target-aptamer complex will dictate the conditions required for elution⁷.

Figure 1-2: The Systematic Evolution of Ligands by EXponential enrichment (SELEX) process. Beginning with a large library of DNA, iterative cycles of target incubation, library partitioning and amplification are performed to select aptamers.

Once separated from the target, the few binding DNA sequences are amplified by PCR to yield a practical amount of sample to continue the process. As PCR generates double-stranded DNA and aptamers are single-stranded, the aptamer sequence is

separated from its complement using one of a number of techniques, such as gel electrophoresis⁴⁹ or using an agarose resin⁵⁰. SELEX progress is monitored by modifying the aptamer strand with a traceable label, to determine when more stringent conditions should be applied⁵¹. The enriched library generated from a round of selection is subjected to further selection rounds that serve to increase the affinity of the library for the target molecule (positive selections). After several rounds, the enriched library is cloned, sequenced and characterized to isolate aptamers with the desired properties. Once these sequences are elucidated, solid-phase chemical synthesis is used to reproducibly synthesize aptamers in large quantities.

1.2.1 Modifications to the SELEX process

An enormous advantage of SELEX is its flexibility with respect to methodology, binding conditions, and library design. The first modifications introduced involved the inclusion of negative or counter-selection steps to eliminate library members displaying affinity for either the solid-support matrix or compounds sharing structural similarity to the target. The majority of modifications involve changing the stringency, the platform on which selection is performed or the type of ligand being targeted^{52, 53}. The library used in the selection can also be modified to include fixed regions of known functionality or additional steps can be included in the selection scheme such as using mutagenic PCR to increase the diversity of the library⁵⁴. Later, the SELEX process was automated by Ellington and others⁵⁵⁻⁵⁷. More recently, as the targets for SELEX have advanced from small molecules to whole cells, several new SELEX techniques including “cell-SELEX”

have been adopted¹³. Table 1-1 below lists several of the modifications to the original SELEX process.

Table 1-1: A list of modifications to the SELEX process and their descriptions (adapted from Stoltenberg et al.⁵²).

Negative selection	An additional step performed typically at the beginning of selection that removes sequences that have an affinity for the selection matrix	58
Counter-selection	Additional rounds added to SELEX that remove sequences that bind to similar target structures	59
Blended SELEX	Incorporation of an extraneous molecule into a nucleic acid library and then selecting the most active of these composite assemblies. Each nucleic acid sequence in the starting library is a variant of the small molecule moiety and allows up to 10^{15} variants of the small molecule to be screened for the most active variant	60
Expression cassette SELEX	A special form of Blended SELEX that involves transcription factors and optimizes aptamer activity for gene therapy application	61
Chimeric SELEX	Uses two or more different oligonucleotide libraries for production of chimeric aptamers with more than one wanted feature or function. Each of the parent libraries will be selected first to a distinct feature; the resulting aptamers are then fused together	62
Multi-stage SELEX	A modification on chimeric SELEX. Here the fused aptamer components then go through an additional selection with all the targets	63
Deconvolution SELEX	Used to generate aptamers for complex targets. Typically selection is performed on mixtures (or a cell). Once aptamers have been generated, a second part of SELEX involves discriminating which aptamers bind to which parts of the complex mixture	64
Covalent/ Crosslinking SELEX	Aptamers contain reactive groups which are capable of covalent linking to a target protein	65
Photo SELEX	Uses aptamers bearing photo-reactive groups that can photo cross-link to a target and/or photo activate a target molecule	66
Spiegelmer technology	Aptamer selection is performed with the natural D-nucleic acids for the mirror-image (enantiomer) target and the final synthesis of the resulting aptamers are L-isomers which bind to the original target	67

Table 1-1 continued

Tailored SELEX	An integrated method to identify aptamers with only 10 fixed nucleotides through ligation and removal of primer binding sites within the SELEX process	68
Genomic SELEX	The SELEX library is constructed from an organism's genome and target proteins and metabolites from the same organism are used to elucidate meaningful interactions	69
Toggle-SELEX	Selection is performed on different targets at alternating rounds	70
Indirect SELEX	The target used in the selection is not the aptamer binder; however, it becomes required for aptamer binding to the new target	71
<i>In vivo</i> SELEX	Uses transient transfection in an iterative procedure in cultured vertebrate cells to select for RNA-processing signals	72
Tissue	Generates aptamers capable of binding to tissue targets	64
FluMag	There are two modifications in this method. First, the library is modified with fluorescein instead of radiolabels for quantification purposes. Secondly, the target is immobilized to magnetic beads instead of agarose	49
CE-SELEX	Separation is performed using capillary electrophoresis	73
Non-SELEX	Involves repetitive steps of partitioning with no amplification steps.	74
EMSA-SELEX	Partitioning occurs through the use of electrophoretic mobility shift assay (EMSA) at every round	75
Atomic force microscopy SELEX (AFM-SELEX)	Makes use of a dynamic atomic force microscopy tip to pick up and visualize aptamer-target complexes. This SELEX requires only one round of selection.	76
Yeast Genetic SELEX	Optimizes <i>in vitro</i> selected aptamers by creating a library of degenerate aptamers and performing a secondary selection <i>in vivo</i> using a yeast three (one)-hybrid system	77
Cell-SELEX	Can generate aptamers that can bind specifically to a cell type of interest. Commonly, a cancer cell line is used as the target to generate aptamers that can differentiate that cell type from other cancers or normal cells	78

The above SELEX modifications include changes to target immobilization, nucleic acid library, selection stringency, amplification or monitoring of the enrichment. The goal of these changes was to either generate improved aptamers for novel applications or to render the currently laborious SELEX procedure, rapid and simple.

1.3 SELEX targets

With the success of the first *in vitro* selection experiments to small organic dyes⁴, much of the original SELEX focus was on developing aptamers for small molecules. Since then, selections using a variety of targets have been performed. Figure 1-3 shows the number of aptamers for various target types selected between 1990 and 2007.

Figure 1-3: Breakdown of the various types of aptamer targets (number of occurrences) that have been selected between 1990 and 2007. This list was generated using aptamer.freebase.com on February 20th, 2012.

However, once it was discovered that aptamers could be easily selected for proteins and cells, new aptamers for small molecules became few and far between. These larger targets, containing more functional groups and structural motifs, give a higher probability of finding sequences that can interact with the target via hydrogen bonds, electrostatic interactions and hydrophobic interactions⁶. Additionally, there are several challenges associated with developing aptamers for small molecules. The primary complication

arises from the need to immobilize the target to a solid support matrix. Several aptamers for small molecules have been selected for targets that can be purchased on pre-made agarose material¹⁰. Should an aptamer researcher choose to take the time to perform conjugation chemistry, several targets do not contain functional groups that can be coupled to agarose columns. For cases where conjugation is possible, the amount of column material available for aptamer binding compared to target is very high, resulting in high non-specific binding of the library. Despite negative selection steps, this problem is difficult to avoid⁷⁹.

In the early 1990's, in an effort to promote the power of SELEX, numerous papers and reviews boasted that *in vitro* selection is easy, inexpensive and fail-safe, resulting in little publication interest for new selections (Dr. Justin Gallivan, Emory University, personal communication). On the contrary, SELEX is very laborious and less than 30% of selections result in aptamers⁵¹. Additionally, patents for virtually every application of aptamers have placed a stranglehold on aptamer innovation⁸⁰. Few research groups have chosen to invest in the development of aptamers for new small molecules and as a result very few aptamers have been made. Table 1-2 lists all 33 small molecules for which DNA aptamers have been selected and characterized at the time of writing. With an average of only three new small-molecule targeted DNA aptamers being developed every two years, and the upcoming expiry of several aptamer patents, there is an exciting drive to develop new aptamers for small molecules.

Table 1-2: A complete listing of DNA aptamers that have been confirmed to bind to small molecules. The dissociation constant, a measure of binding affinity, is included as well as the year of aptamer development.

Target	Binding affinity measured as K_d	Year	Reference
Reactive Green 19	33 μ M	1992	42
AMP and ATP	6 μ M	1995	81
L-arginine	2.5 mM	1995	82
L-argininamide	0.25mM	1995	82
Sulforhodamine B	190 nM	1998	83
Cellobiose	600 nM	1998	84
7,8-dihydro-8-hydroxy-2'-deoxyguanosine	270 nM	1998	85
Cholic acid	5 to 67.5 μ M	2000	86
Hematoporphyrin	1.6 μ M	2000	87
L-tyrosinamide	4.5 μ M	2001	88
Sialyllactose	4.9 μ M	2004	89
Ethanolamine	6–19 nM	2005	90
(R)-thalidomide	1 μ M	2007	92
Hoechst derivative 7e	878 nM	2007	93
Tetracycline	64 nM	2008	94
L and D arginine	590–810 μ M	2008	95
Daunomycin	10 nM	2008	96
Oxytetracycline	10 nM	2008	97
Ochratoxin A (OTA)	200 nM	2008	50
Arsenic (V)	7 nM	2009	98
Arsenic (III)	5 nM	2009	98
Dopamine	700 nM	2009	99
8-hydroxy-2'-deoxyguanosine	100 nM	2009	100
Diclofenac	42.7–166.34 nM	2009	101
Atrazine	2 μ M	2010	102
(S) and (R)-ibuprofen	1.5–5.2 μ M	2010	103
ATP	3.7 μ M	2010	104
Fumonisin B ₁ (FB ₁)	100 nM	2010	105
Acetamiprid	4.98 μ M	2011	106
Kanamycin	78.8 nM	2011	107
L-tryptophan	1.757 μ M	2011	108
Bisphenol A (BPA)	8.3 nM	2011	109

Small molecules (molecular weight of 1000 g/mole) are important in biological processes due to their ability to diffuse across cell membranes¹¹⁰. These compounds may be harmful (e.g. toxins and carcinogens), or they may play key roles in nutrition or treatment, such as drugs or nutrients. In cells, small molecules display a variety of biological functions serving as cell signaling molecules, pigments or as defense

mechanisms¹¹¹. In molecular biology, they can be used as tools, acting as antibiotics or other important drugs¹¹². In food, small molecules are important for energy storage or can act as natural pesticides¹¹³. It is known that aptamers are ideal molecular recognition probes for small molecules^{8, 114}, where the potency of aptamer-target interactions lies in the high degree of selectivity observed with these interactions. Several groups have demonstrated this ability. The first example of this unparalleled selectivity was observed in 1994, when the selected RNA aptamer for theophylline displayed a 10,000 times weaker binding affinity to caffeine, a xanthine that differs by a single methyl group. This selectivity was found to be a 10 fold improvement on the selectivity for the antibodies for these targets⁵⁹.

Aptamers are additionally able to select for chirality. Given that the majority of organic compounds and biological molecules are chiral, and their activity relies on this handedness, aptamers are a powerful tool in specific detection and separation of these compounds¹¹⁵. Several groups have reported this stereoselectivity¹¹⁶. Initially, several RNA aptamers displayed some discrimination between various L and D amino acids^{117, 118}. In 1996, Geiger et al.⁵⁸ reported the selection of RNA aptamers that bind to L-arginine with high affinity and enantioselectivity. In 1997, Williams et al.¹¹⁹ produced the first stereoselective DNA aptamer that bound to the D-enantiomer of the peptide hormone vasopressin with high affinity. This aptamer displayed a significantly lower affinity for the L-enantiomer. More recently, enantioselective DNA aptamers have been selected for the small molecule (R)-thalidomide⁹² and separate aptamers have been identified for (S) and (R)-ibuprofen¹⁰³.

1.4 Mycotoxins

Mycotoxins are small-molecules that have to date not received much attention by aptamer researchers. These molecules are fungal secondary metabolites that can contaminate a wide range of agricultural commodities worldwide. Many tens of thousands of analyses are done for mycotoxins each years and are thus high priority targets for the development of new biosensors and detection systems¹²⁰.

Most mycotoxins are secondary metabolites meaning they are typically formed when one or more nutrients has become limiting, after mycelia growth has ceased¹²¹. It is estimated that at least 25% of the maize and small-grains produced worldwide are contaminated with mycotoxins which may be capable of causing disease and death in plants, animals and humans¹²². Several mycotoxins can be cytotoxic, carcinogenic, mutagenic or immunosuppressive. Since the discovery of aflatoxins in 1960, governments have been attempting to limit mycotoxin contamination in foods and feeds. Compared to food contamination with bacteria, mycotoxins are not inactivated by cooking temperatures, therefore careful monitoring of cereals is essential prior to distribution for consumption. Due to the health risks to humans and animals, national and regional food authorities have addressed the mycotoxin problem by adopting and amending regulatory limits¹²³. Although prevention is the best control strategy, mycotoxin contamination cannot be completely avoided¹²⁴. Additionally, post-harvest control, decontamination and detoxification procedures are constantly being investigated; however, no single method has been developed that is effective against a wide variety of mycotoxins¹²⁵. Therefore, rapid and inexpensive monitoring is the best course for reducing mycotoxin exposure to humans and animals.

Current detection methods include thin layer chromatography, enzyme-linked immunosorbent assays, capillary zone electrophoresis, gas chromatography, and high performance liquid chromatography technologies¹²⁶. While these methods are effective, the need for highly trained personnel and expensive equipment make them inaccessible in all but the most well-equipped laboratories. Testing for contamination with these methods is labor-intensive and slow. The use of aptamers in place of antibodies in mycotoxin detection may be particularly useful as mycotoxin extraction involves organic extraction procedures that commonly denature antibodies¹²⁷. The production of mycotoxin binding aptamers is therefore the first step in the realization of inexpensive, robust and innovative detection technology that could offer a solution to the limitations of the existing detection techniques. These tests strategies may be appropriate for crop handlers and agricultural inspectors in the field that are far from laboratories.

1.5 Aptamer binding affinity

The high affinity aptamers exhibit for their targets is quantified using either an association constant (K_a) or dissociation constant (K_d). The majority of aptamers are described using the latter, which represents the target concentration in which half of the aptamers are bound by target. This type of equilibrium constant (K), in molarity, measures the propensity of the aptamer-target complex (AT) to dissociate into its individual components, aptamer (A) and target (T), see the general reaction¹²⁸:

For this general dissociation reaction, the dissociation constant can then be described as:

$$K_d = \frac{[A][T]}{[AT]} = \frac{1}{K_a}$$

From the equation, the following relationships are evident: a lower K_d value represents a higher affinity of the aptamer-target complex and K_a is the inverse of K_d . Aptamers rival antibodies in that they can bind to their targets with dissociation constants in the nanomolar and some cases picomolar range⁸. Sub-picomolar dissociation constants as a result of non-covalent binding interactions between two molecules are rare and have yet to be discovered with aptamers¹²⁹.

To measure the K_d , a constant concentration of either the aptamer or target is titrated with an increasing concentration of the other component. For example, if the initial target concentration is held constant, then the fraction of bound target (f_t) is given by:

$$f_t = \frac{[A]}{K_d + [A]}$$

This equation results in the formation of a binding isotherm in the form of a rectangular hyperbola that saturates at $f_t = 1$, which represents nearly 100% of the target binding to the aptamer. From this binding isotherm, the K_d is found at $f_t = 0.5$ and can be estimated directly using a nonlinear regression analysis. However, this equation requires that the concentration of target is significantly smaller than the K_d so that the free aptamer concentration $[A]$ at equilibrium does not change significantly from the total target concentration upon binding¹²⁸. Depending on the sensitivity of the method being used to measure the K_d , this is not always possible¹²⁸. In the cases where higher concentrations of target are required to obtain a measurable signal, both the target and aptamer

concentrations must be included into the equation. This equation, below, gives a more accurate representation of the dissociation constant.

$$f_t = \frac{([A_0] + [T_0] + K_d) - \sqrt{([A_0] + [T_0] + K_d)^2 - 4[A_0][T_0]}}{2[A_0]}$$

The dissociation equilibrium can also be described as the ratio of the rate constants of the forward (k_{on}) and reverse reactions (k_{off}). Two aptamers may have the same K_d , but one may have much higher k_{on} and k_{off} rates compared to the other. This is related as shown⁶:

$$K_d = \frac{k_{off}}{k_{on}} = \frac{[A][T]}{[AT]}$$

In terms of aptamer-target binding, a few trends have emerged over the past two decades. First, a higher affinity often results in higher selectivity. This is due to the higher number of interactions formed with the target, which may not be present with a different ligand. However, some studies have shown that this is not always the case¹³⁰. Secondly, as mentioned earlier, large molecules bind to aptamers better than smaller molecules due to the increased possible number of non-covalent interactions¹³¹. Sub-nanomolar dissociation constants are rare with small molecules, but are often found with aptamers binding to proteins and cells⁶. Third, targets with heterocyclic functionality, such as theophylline, make better aptamer targets, often binding at much lower dissociation constants. This may be attributed to the increased pi-pi stacking interactions

between the nucleotide bases and the heterocyclic target¹³², though others have postulated that this property is due to a decreased number of rotatable bonds¹³¹.

1.5.1 Methods for measuring binding affinity

Methods of determining binding affinity are constantly being developed and are often the limiting factor in the rapid development and testing of aptamers. Moreover, the method used to measure the K_d may affect the reported apparent K_d . For example, some techniques may not be sensitive enough to effectively measure the interaction¹²⁸.

A list of methods for determining binding affinity can be found in Table 1-3. Binding affinity can be determined by measuring or detecting a change in aptamer conformation upon binding to target. While this has been useful for several aptamers, not all aptamers undergo a measurable conformation change¹³³. Alternatively, fluorescent and absorbance related methods can be used to rapidly determine K_d values, however, these are generally limited to targets that have intrinsic fluorescence or absorbance properties. Surface mass-sensitive detection methods such as quartz crystal microbalance (QCM) and surface plasmon resonance (SPR) are rapid but are typically limited to large targets such as proteins³⁵. Size-exclusion procedures tend to be much more common among several target types. Capillary electrophoresis (CE) and high-performance liquid chromatography (HPLC) methods are limited to laboratories that have access to this equipment. Therefore, small molecule binding aptamers, whose targets lack any intrinsic fluorescent properties, are typically assayed using affinity chromatography, equilibrium dialysis or ultrafiltration.

Table 1-3: Methods for determining binding affinity of aptamers to their targets.

Method	Description	Reference
Binding detected using the conformational change experienced by the aptamer		
Isothermal titration calorimetry (ITC)	Allows simultaneous determination of K_d , stoichiometry and thermodynamic properties. It relies on the fact that formation of the aptamer–target complex is an exothermic process that releases heat.	134
Circular dichroism (CD)	CD refers to the differential absorption of left and right circularly polarized light. DNA displays optically active chirality. Upon aptamer folding around the target, the CD spectra will change.	135
Nuclear magnetic spectroscopy (NMR)	Distinguishes individual amide protons. By comparing the heteronuclear single quantum coherence spectroscopy (HSQC) of the free/ bound aptamer, changes in the chemical shifts of the peaks is possible.	136
Electrophoretic mobility shift assay (EMSA)	The presence of the target increases the molecule weight of the complex, resulting in a gel shift. With small targets, binding may cause a conformational change resulting in a change in migration through the gel matrix.	137
Fluorescence/ absorbance based methods		
Fluorescence intensity	The fluorescence of the aptamer or target may be quenched or increased upon binding.	138
Fluorescence polarization	A fluorophore is excited with polarized light and due to rotational diffusion, the size of the fluorophore will dictate the proportion of polarized light that is emitted.	139
UV-Visible spectrometryabsorption	If there is an intensity change in the maximum absorption of either the aptamer or its target or if another absorption peak appears in the near visible region due to the presence of protein, metal ions or an extended π -electron system, binding can be measured. In some cases, melting studies can be used.	140
Mass-sensitive surface-based measurement		
Surface plasmon resonance (SPR)	Either the target or aptamer can be coupled to a chip. By flowing various concentrations of the non-tethered ligand, changes in refractive index can be measured as the aptamer-target complex forms.	54
Quartz crystal microbalance (QCM)	This method uses piezoelectric crystals to correlate the mass accumulated (target binding) on the surface with a decrease of the resonance frequency of the quartz crystal.	141
Size-exclusion separation methods		
High-performance liquid chromatography	Allows the separation of different interactions with the stationary phase. Zone separations of aptamer, target and complex can be used to assess the equilibrium distribution of these components.	142
Capillary electrophoresis (CE)	This method is similar to HPLC except that it using an electric field to separate the components of the mixture by size and charge.	143
Equilibrium dialysis	Equilibrium dialysis allows the aptamer, target and the complex to equilibrate. However, two compartments are separated by a semi-permeable membrane allowing only the smallest component to pass through.	50
Ultrafiltration/ nitrocellulose filtration	This method is similar to dialysis. The aptamer and target are incubated to allow binding. The fraction of the smallest unbound component is forced through a filter and measured.	144
Affinity chromatography	Either the target or aptamer is covalently immobilized to a solid-phase support. The other component is allowed to bind and non-binders are washed through.	7

Relatively few methods are available for measuring aptamer binding to small molecules, the main methods used in this thesis are described below.

Affinity chromatography is a simple technique that allows separation of biological mixtures based on highly specific interactions. For measuring aptamer-target equilibrium, the target is immobilized to a solid support matrix, typically using the matrix that was used for the aptamer selection. Various concentrations of a fluorescently or radio-labeled aptamer are allowed to incubate with the target-immobilized matrix to permit binding. Washing steps are then performed to remove non-binders⁴⁹. Since the binding of the aptamer is a result of non-covalent interactions, elution of the aptamer from the affinity medium can be performed in the same way as elution in SELEX; therefore, by using denaturants, heat or changing the pH, ionic strength or polarity. The concentration of the aptamer can then be measured via the tracer, using either fluorescence or radioactivity assays. While commonly-used, aptamers that have been directly selected using these solid-support matrices often display non-specific binding to the matrix itself⁷⁹. As a result, K_d values obtained only in this way have large error values or do not always accurately reflect the apparent binding in solution.

Equilibrium dialysis and ultrafiltration techniques (see Figures 1-4 and 1-5 respectively) have the advantage of measuring binding directly in solution. Various concentrations of the aptamer and the target are allowed to incubate and are placed into the loading chamber of the dialysis tube or into the top portion of the ultrafiltration cell. After binding occurs, separation is performed using the semi-permeable membrane in these devices. It is necessary that the smallest component of the aptamer-target complex be specifically measured. With small molecule cases, it is always the target that must be

detected particularly since the smallest reported aptamer (for thrombin) is over 4000 g/mol¹⁴⁵. With equilibrium dialysis, the tubes are allowed to shake at low speeds. Over a period of 24 hours, the small molecule will be allowed to diffuse across the membrane; however, the larger aptamer will remain in the original loading chamber. The more tightly the aptamer is bound to the small molecule, the less likely the small molecule will be able to diffuse into the other chamber⁵⁰. Ultrafiltration can be performed slightly faster, using a brief, centrifugal force to make the unbound small molecular target pass through the filter membrane. With both of these methods, the amount of small molecule present on either side of the membrane is compared, and the fraction can be used to generate a binding curve⁵⁹. If there are no spectroscopic methods available to measure the concentration of the small molecule, liquid chromatography mass-spectrometry (LC/MS) methods may be used. While the dialysis and filtration techniques are low-cost, they can be quite laborious and result in significant error. More importantly, the molecular weight cut-off for these devices cannot always be purchased for the aptamer-target system. Moreover, it has been found that some of the small-molecular targets non-specifically interact with the membranes themselves.

Figure 1-4: The equilibrium dialysis apparatus consists of two compartments separated by a semi-permeable membrane. The target (purple) in this case is allowed to move between the chambers, whereas the aptamer (green) is unable to pass through the membrane.

Figure 1-5: The process of using ultrafiltration tubes to determine aptamer binding: the target (purple) and aptamer (green) are allowed to incubate. Separation occurs using centrifugation, allowing only unbound target to pass through the filter membrane. The filtrate is measured to determine the concentration of unbound target.

A tool that has been frequently used in studying biological interactions is fluorescence polarization. This technique can give information about the size of fluorescent molecules or information about binding. In this method, excitation light is passed through a polarizing filter called a polarizer that causes the light to oscillate in a single orientation. The fluorophore will then absorb the most light energy when it is parallel with the polarization axis of the excitation light. During the nanosecond timescale that a fluorescent molecule is in the excited state, rotation may occur causing a change in orientation of the emission dipole with respect to the excitation dipole. This is also known as depolarization of the fluorescence emission. Due to Brownian motion, small molecules have more rotational diffusion than large molecules and therefore have a greater propensity for depolarization, whereas fluorescence polarization is retained for large molecules displaying limited motion. This phenomenon (see Figure 1-6) can be exploited to measure the fraction of aptamers bound to small fluorescent molecule. Since aptamer binding to the target will result in an increase in size, the rotational diffusion will decrease and a higher polarization of the fluorescent target will be observed¹⁴⁶.

Figure 1-6: Using fluorescence polarization to quantify aptamer binding. A) The fluorescent target (purple) experiences rotational diffusion, resulting in the emission of depolarized light. B) When the target is bound by the aptamer (green), the complex experiences less rotational diffusion, leaving the majority of the emitted light polarized.

1.6 Thesis objectives

While aptamer technology has existed for over two decades, the difficulties on the development of aptamers for small molecules has resulted in very few aptamers that can bind to such compounds. This thesis aims to develop new aptamers for small molecules including medically relevant targets and food-contaminating mycotoxins as well as to improve the procedures used in their selection and testing.

In Chapter 2, the investigation into the feasibility of using aptamers in commercial food-safety testing is presented. Using an existing aptamer for ochratoxin A (OTA), we demonstrate the utility of aptamer-based solid-phase extraction columns for rapid clean-up of OTA from naturally contaminated wheat samples prior to detection.

Apart from this OTA aptamer, no other groups had successfully isolated aptamers that recognized mycotoxins. In Chapter 3, newly selected OTA aptamers with improved binding characteristics are presented. Additionally, the generation and testing of novel aptamers for the mycotoxin FB₁ is presented in Chapter 4, and the progress in the selection of aptamers for DON is presented in Appendix A.

Because homocysteine is a marker for a number of medical conditions, selection of aptamers for homocysteine are investigated. Information on their binding properties and selectivity is presented in Chapter 5. Additionally, the development of new aptamers for ATP (Chapter 6) is discussed, one of which displayed improved binding over pre-existing aptamers for this target.

Chapter 6 explores the use of computational methods to generate SELEX libraries that exhibit higher structural complexity. The value of these libraries was

demonstrated in selection experiments and the development of high affinity aptamers for small molecules.

Finally, Chapter 7 details the development of a new functional database for SELEX and aptamers. The data model that describes the structured information about the experimental conditions under which aptamers were selected and assessed is also described. The utility of this database is investigated by querying the entered data to determine relationships between SELEX conditions and the functionality of the aptamers developed.

Chapter 2. Development of aptamer solid-phase extraction columns for the extraction and detection of ochratoxin A

De Girolamo, A.; McKeague, M.; Miller, J. D.; DeRosa, M. C.; Visconti, A.

Determination of ochratoxin A in wheat after clean-up through a DNA aptamer-based solid-phase extraction column. *Food Chem.* **2011**, *127*, 1378-1384.

2.0 Statement of contributions

Synthesis and binding characterization of the aptamer, and initial preparation of the oligosorbent by M. McKeague. All aptamer solid-phase extraction tests, optimization and wheat extraction were performed at the Istituto di Scienze delle Produzioni Alimentari, Consiglio Nazionale delle Ricerche (CNR) in Bari, Italy by Annalisa De Girolamo and Maria C. DeRosa.

2.1 Introduction

2.1.1 Ochratoxin A

Ochratoxins are a group of mycotoxins produced by several fungi in the genera *Aspergillus* and *Penicillium*¹⁴⁷. Ochratoxin A (OTA), *N*-[(3*R*)-(5-chloro-8-hydroxy-3-methyl-1-oxo-7-isochromanyl)carbonyl]-L-phenylalanine (MW= 403.81 g/mol), see Figure 2-1, is the main compound that accumulates in food. Since its discovery in 1965¹⁴⁸, it has become one of the most studied mycotoxins and has recently attracted regulatory attention in a number of countries¹⁴⁹. This mycotoxin is a structural analogue of the amino acid phenylalanine linked to a chlorinated dihydroisocoumarin moiety via an amide bond¹⁵⁰. It is able to compete with phenylalanine in the reaction catalyzed by phenylalanyl-t-RNA synthetase and thus results in protein synthesis inhibition¹⁵¹. Various studies have since shown that OTA impairs blood coagulation¹⁵², glucose metabolism¹⁵³, is immunosuppressive¹⁵⁴ and teratogenic¹⁵⁵.

Figure 2-1: Structure of ochratoxin A.

In particular, OTA is known for its nephrotoxicity in all animal species tested¹⁵⁶. OTA is also carcinogenic¹⁵⁷ to rodents and has been classified by International Agency for Research on Cancer (IARC) as a possible human carcinogen¹⁵⁰. The occurrence of OTA in cereals such as wheat, barley, corn, oats, and rye is of primary concern due to the importance of the aforementioned cereals in diet. However, several other food products such as beer, wine, nuts, liquorice, coffee, cocoa, meat and spices are also susceptible to OTA contamination. Unlike crops contaminated with other mycotoxins, OTA contamination does not result in any observable physical damage. As a result, careful monitoring and testing is required and several regulatory limits have been established worldwide¹⁵⁸. Table 2-1 lists some of the maximum tolerated levels of OTA in foodstuffs from various countries worldwide. Several countries have also established regulations on OTA in animal feed^{149, 150, 159}. For example, the recommended tolerance levels in Canada for poultry and swine diets is 2 mg/kg¹⁶⁰. The proposed Health Canada maximum limits for OTA in various food commodities can be found in Table 2-2.

Table 2-1: Examples of maximum tolerated levels of OTA in foodstuffs by country and product (adapted from Duarte et al.¹⁴⁹).

Country	Product	Maximum level of OTA (ppb)
Bulgaria	grape juice	3
	dried vine fruit and cereals to be subjected to sorting, or other physical treatment, before human consumption or use as an ingredient in foodstuffs	5
	cereals and processed products thereof intended for direct human consumption or as an ingredient in foodstuffs	3
	Spices	10
	green coffee beans	8
	roasted coffee	4
Cuba	coffee and cereals	5
Czech Republic	child and baby nourishment	1
	flours and cereal products	3
	Foodstuffs	5
Denmark	pig kidney	10
Estonia	pig liver	10
	cereals, cereal flours, cereal groats and flakes, pasta products, ordinary baker's wares, fine baker's wares; isolates, concentrates and hydrolysates of cereals protein	5
Greece	coffee (raw and processed)	20
Hungary	cereals (including the rice and Fagopyrum sp.)	5
	every cereal product including milled products and those cereal products used for direct human consumption	3
	raisin (currant, sultana); roasted coffee and coffee products; other plant originated foods	10
	green/unroasted coffee	15
Indonesia	Coffee	Not detectable
Iran	dates, dried grapes (raisins and sultanas), official figs and all dried fruits	10
	baby food based on cereals without milk and baby instant food (ready to use)	1
	barley and maize	50
	rice and wheat	5
	Legumes	20
Israel	cereals, cereal products and other foods	50
Italy	Coffee	8
	roasted coffee	4
	cocoa and derived products	0.5
	Beer	0.2
	pig meat and derived products	1
	fruit juice	50
	baby food	0.5
Latvia	Cereals	5
Morocco	Cereals	30
Romania	alimentary products	20
	alimentary products for babies of 3 years old	5
Serbia and Montenegro	all foodstuffs	10
Singapore	cereal, raw and roasted coffee beans	2.5
Slovakia	milk, meat, poultry, flour and its products, rice, vegetables, potatoes	5
	infant formulae and food for children	1
	other foodstuffs	10
Sudan	Wheat	15
Switzerland	dried fruit and spices	20
	infant formulae and follow-on formulae and processed cereal-based foods and baby foods for infants and young children	0.5
	all foodstuffs	5
Turkey	raw grain	5
	foodstuffs produced from grain	3
	Raisins	10
Uruguay	rice, barley, beans, coffee, corn	50

Table 2-2: Proposed Health Canada maximum limits for the presence of ochratoxin A in various food commodities (adapted from Health Canada 2009¹⁶¹).

Product	Maximum level of OTA (ppb)
raw cereal grains	5
directly consumer grains (i.e. rice, oats, pearled barley)	3
derived cereal products (flour)	3
derived cereal products (wheat bran)	7
breakfast cereals	3
grape juice (and as ingredients in other beverages) and related products	2
dried vine fruit (currants, raisins, sultanas)	10
baby foods and processed cereal-based foods for infants and young children	0.5
dietary foods for special medicinal purposes intended for infants	0.5

Collected data has shown that OTA levels in Canadian grain are generally below international limits. From 1994-2010, the average contamination of all grain cargo shipments was 0.8 ppb¹⁵⁸. With the technology currently available, 1 ppb is the limit of quantification of OTA in food samples. Typical detection uses chromatographic techniques. Thin-layer chromatography-fluorescence detection (TLC-FD) and high performance liquid chromatography-fluorescence detection (HPLC-FD) have been adopted as official methods¹⁶². However, much work is being done to either improve these techniques or to develop new detection methods to improve the cost, speed, and sensitivity associated with detection. Sensing platforms such as capillary electrophoresis, solid-phase extraction–liquid chromatography–electrospray tandem mass spectrometry and solid-phase microextraction–liquid chromatography–fluorescence detection have been reported. Among a variety of detection methods, toxin-specific antibody is consistently being utilized as the molecular recognition agent¹⁶³.

A major issue with many of these methods is that multiple steps are required prior to detection. These include extraction, extensive sample clean-up, pre-concentration, and even derivatization of the analyte. For compounds regulated in the ng/g level, sample clean-up is performed using one-time use immunoaffinity columns. Additionally, in an attempt to overcome the limitations encountered with chromatographic analysis, many immunoaffinity methods have been developed¹⁶⁴. Additionally, fiber optics¹⁶⁵ and fluorescence polarization¹⁶⁶ have been shown to be useful for rapid mycotoxin analysis.

2.1.2 The first mycotoxin aptamer

The first report of OTA aptamers appears to be from 2006 by Hungarian workers¹¹ (see Chapter 3). However the first that received considerable attention was reported in 2008. This aptamer was selected using conventional SELEX strategies in 2008⁵⁰. A minimal aptamer was found, henceforth referred to as aptamer 1.12.2, that is 36 nucleotides in length and binds to OTA with a reported dissociation constant of 200 nM⁵⁰. While structural studies have not been performed on this aptamer, the high guanine incidence (47 %) suggests it folds into a G-quadruplex¹⁶² like several other DNA aptamers^{81, 145}. Studies with divalent cations show that this aptamer affinity increases 4 fold in the presence of 20 mM Ca^{2+} . This occurrence is presumably because calcium forms a coordination complex between the carboxyl and the 8-hydroxyl groups in OTA⁵⁰, reducing free rotation¹³¹. The aptamer has also been tested for selectivity, showing no binding to compounds with similar structures such as N-acetylphenylalanine and warfarin. Additionally, the aptamer bound with a 100-fold less affinity to ochratoxin B, which lacks the chloro group⁵⁰.

Several groups have integrated this aptamer into typical antibody detection systems and applications. For example, the aptamer has been incorporated into a colorimetric detection system utilizing gold nanoparticles, where accurate measurements can be done in less than 5 minutes¹⁶². Aptamer-based electrochemical and electrochemiluminescent sensors have also been developed using the OTA aptamer^{164, 167, 168}. These systems are sensitive, displaying detection limits in the low ppb range. Other applications include chromatographic and fluorescent test strips^{169, 170}, colorimetric assays with horseradish peroxidase¹⁷¹, fluorescent quenching assays^{163, 172}, rolling circle amplification detection¹⁷³ and fluorescence polarization^{139, 174}.

To demonstrate the feasibility of using this aptamer as a sorbent in place of antibodies, affinity columns for solid-phase extraction (SPE) were prepared and tested for OTA cleanup from wheat extracts⁵⁰. While promising, the study lacked the robustness required to bring the technology to commercial readiness.

2.1.3 Aptamers in food safety

Despite the promise of aptamers as biosensors and the advantages they pose over antibodies, relatively little work has been achieved in aptamer-related food safety testing. Only a handful of aptamers exist for antibiotics, adulterants, toxic metals, pesticides, bacteria and viruses related to food¹⁷⁵. Most of these aptamers have been incorporated into one or two biosensing platforms, but the majority of them have not been used for food-related testing. Additionally, very few of the biosensor platforms have been tested for their ability to function directly in food matrices. Regardless, the few studies that have tested aptamer-biosensors with food matrices have shown great potential. For

example, a screening assay employing solid-phase extraction (SPE) followed by an RNA aptamer-based procedure was employed for rapid detection of malachite green residues in fish tissues¹⁷⁶. Zhang et al. developed a DNA aptamer modified glassy carbon electrode and were able to directly quantify tetracycline in less than 5 minutes in 6 separate milk samples with a coefficient of variance of less than 10%³³. An electrochemical sensor was also developed by Zelada-Guillén that was able to rapidly detect *Escherichia coli* in skim milk¹⁷⁷. An aptamer-magnetic bead and quantum dot sandwich assay was also able to detect *Campylobacter jejuni* in dilute food matrices such as ground beef, 2% milk and chicken juice¹⁷⁸. While all of these studies have demonstrated the promise of aptamer-based biosensors for food safety testing, these studies need to be further optimized and tested for commercial purposes.

2.1.4 Chapter objectives

Despite the promise of aptamers as biosensors, none have progressed to the point of commercial readiness. This chapter investigates the use of a DNA aptamer for preparing solid-phase extraction columns for the rapid extraction of OTA from naturally contaminated wheat samples. The optimization work and comparison with immunoaffinity columns is presented.

2.2 Experimental

2.2.1 Materials

Phosphoramidites, modifiers, activator, deblock, capping, and oxidizing reagents were obtained from Glen Research. Standard support columns and acetonitrile were

purchased from BioAutomation. Ultra High Purity 5.0 argon was purchased from Praxair. All buffers were prepared with Millipore Milli-Q deionized water at 18 MΩcm. Buffering agents, salts, OTA from *Petromyces albertensis* ≥ 98% (TLC) for binding studies and OTA standards were purchased from Sigma-Aldrich. All molecular biology grade electrophoresis chemicals were purchased from BioShop Canada (Burlington, Ontario). All solvents for the solid-phase extraction experiments were HPLC grade (Mallinckrodt Baker, Milan, Italy). Sulfo-NHS acetate, Amicon-Ultra 0.5 mL 3 kDa centrifuge units and Stirred-cell Ultrafiltration Membranes (1 kDa, 25 mm) were purchased from Fisher Scientific Canada. Sample Dispo Equilibrium Dialysis units, 5000MWCO, were purchased from Harvard Apparatus (Holliston, Massachusetts, USA). CarboxyLink Coupling gel (immobilized diaminodipropylamine, DADPA, supplied in a 50% slurry) and 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (EDC) were received from Pierce (Rockford, Illinois, USA). Empty Bond Elut® columns of 1-ml capacity (0.5 mm internal diameter) and 3 ml capacity (0.9 mm internal diameter) were received from Varian (Palo Alto, California, USA). OchraTest™ WB columns were purchased from Vicam L.P. (Watertown, Massachusetts, USA). Unprocessed wheat samples were received from an Italian local milling factory (Bari, Italy). Prior to analysis, the wheat samples were ground with a laboratory mill (Tecator Cyclotec 1093, International PBI, Hoganas, Sweden) to a uniform consistency (particle size < 500 μm).

2.2.2 Aptamer synthesis

The ssDNA aptamer sequence, 1.12.2 (5'-GATCGGGTGTGGGTGGCGTA-AAGGGAGCATCGGACA-3')⁵⁰ was prepared by using standard phosphoramidite

chemistry on a BioAutomation Mermade 6 (Plano, Texas) using 1 μ mol controlled pore glass (CPG) columns (500Å pore size). Synthesis conditions recommended by the manufacturer were followed with the exception of a 60 second increase in the coupling time for any 5' modifier. Unmodified sequence 1.12.2 was synthesized in addition to sequences modified at the 5' end with fluorescein phosphoramidite (6-FAM).

The 6-FAM modified and unmodified DNA columns were removed from the synthesizer. All of the CPG support beads were dried using argon and transferred to microcentrifuge tubes and allowed to incubate in 1 mL of 28% NH_4OH for 16-24 hours at 55°C. The CPG beads were removed using 0.22 μm Spin-X cellulose acetate microcentrifuge filter tubes (Corning) and dried over night by cryopumping on a Savant AES2010 SpeedVac. Cellulose acetate filters were used rather than nylon filters as DNA binds to the nylon membranes, whereas cellulose acetate has very low DNA binding. This DNA was purified using a 12% denaturing polyacrylamide gel procedure adapted from the Molecular Cloning Laboratory Manual¹⁷⁹. A gel solution was prepared using 31.5 g of urea, 23.5 mL of stock acrylamide solution (acrylamide/bis-acrylamide, 40% solution 19:1), 15 mL of 5 x TBE buffer (45 mM Tris, 45 mM boric acid, 1 mM EDTA) and 14 mL of deionized water. The solution was stirred and warmed to 37°C, filtered through Whatman No. 1 filter paper and allowed to cool. The polymerization reaction was initiated by adding 0.45 mL of 10% ammonium persulfate and 35 μL of N,N,N',N'-tetramethylethylenediamine (TEMED). The gel was cast by loading the solution into a glass-plate set up and inserting the one-well comb. After polymerization for 30 minutes, the gel wells were rinsed with deionized water and the gel was pre-run on an SE 600 Chroma Standard Dual cooled gel electrophoresis unit (Hoefer, USA) for 15

minutes using a constant current of 25 mA. The crude DNA sample was dissolved in water and mixed with an equal amount of formamide and heated at 55°C for 5 minutes. One synthesis column of sample was loaded per gel. In place of loading dyes, a single lane of an 18-nt fluorescein modified DNA strand was loaded to track DNA migration. The gels were run at 25 mA for approximately 2 hours, or until the tracer had migrated half-way through the gel. The gel was characterized using absorbance by imaging on a 20 x 20 cm TLC plate (F₂₅₄) using an Alpha Imager Multi Image Light Cabinet (Alpha Innotech) on the epi-UV setting (excitation wavelength of 254nm). Fluorescein-labeled bands were also visualized using emission (excitation at 302nm). The relevant DNA bands were cut from the gel, crushed and incubated in deionized water at 37°C overnight with gentle shaking or rotation.

Various gel extraction techniques were used in aptamer isolation. The above incubation in deionized water proved to be the fastest and most efficient and resulted in the least amount of contamination with other DNA samples. The crush and soak method, adapted from the Molecular Cloning Laboratory Manual¹⁷⁹, follows a similar procedure as above. The relevant DNA bands were cut from the gel, crushed using a pipette tip and incubated in elution buffer containing 0.1% SDS, 0.5 M ammonium acetate and 10 mM magnesium acetate at 37°C overnight with shaking or rotation. The freeze/thaw method¹⁸⁰ was adapted from Chen et al. Again, the relevant DNA bands were cut from the gel. The gel was crushed by passing it through a small-bore syringe. For every 0.5 mL of gel, 3 mL of 10 mM Tris, pH 7.4 buffer was added. The sample was rapidly frozen at -80°C for 30 minutes, followed by a quick 5 minute thaw in a 50°C hot water

bath and a 5 minute soak at 90°C. This was followed by an overnight elution on a rotary shaker at room temperature.

Following extraction, gel was removed via filtration through 0.22 µM cellulose acetate syringe filters. Extracted DNA was purified using filtration with 3kDa cutoff Amicon-Ultra centrifuge units. After concentration of the sample, the filters were rinsed four times with deionized water to remove trace amounts of urea and other remaining gel components. Purified DNA was then quantified using a Cary 300 Bio UV-Visible spectrophotometer (Varian, USA).

2.2.3 Confirming binding of aptamer 1.12.2

Equilibrium dialysis

The same procedure used by Cruz-Aguado et al. was used to verify the binding affinity of the 1.12.2 aptamer for OTA⁵⁰. Single-use Sample Dispo Equilibrium Dialysis units with a 5000Da molecular weight cut-off were used. In each of these units, 75 µL of selection buffer (SB) containing 10 mM HEPES, pH 7.0, 120 mM NaCl, 5 mM KCl, and 5 mM MgCl₂ was placed in the receiving chamber. To ensure equilibrium of OTA could be reached, the loading chamber was filled with 75 µL of 200 nM OTA dissolved in SB. After 48 hours, the fluorescence was measured in each chamber. To determine whether aptamer was able to pass through the semi-permeable membrane, 75 µL of a 20 µM solution of aptamer 1.12.2 in SB was added to the loading chamber. After the 48 hour incubation time, the concentration of DNA in the receiving chamber was measured using UV-Visible spectrometry. Finally, for the single-point binding assays a 75 µL solution of 20 µM aptamer and 200 nM OTA in SB was loaded into the loading chambers of seven

separate units. The units were allowed to gently shake at 60rpm while lying flat for 48 hours to permit equilibration between the chambers. Following the 48 hour agitation period, fluorescence measurements of the loading and receiving chambers were obtained by direct fluorescent measurement (excitation wavelength 375nm; emission wavelength 430nm) using a Fluorolog Fluorescence Spectrophotometer (Horiba Jobin Yvon, USA) with a SpectrAcq controller.

The fraction of bound OTA (f) was determined as:

$$f = \frac{F_l - F_r}{F_l}$$

and the dissociation constant (K_d) was estimated as:

$$K_d = \frac{[A_0]}{f} - [A_0]$$

A full binding curve was produced for aptamer 1.12.2 using the same procedure, but in place of 20 μ M aptamer, 9 samples of aptamer concentrations varying from 0.5 nM to 30 μ M were used. The fraction of OTA bound was measured, calculated as above and plotted against aptamer concentration. The K_d was evaluated by minimizing the residual values between calculated and observed experimental fraction of bound OTA using the solver feature of Microsoft Excel^{181, 182}.

Affinity chromatography

OTA-immobilized column material was prepared as previously reported⁵⁰. A 1 mL aliquot of DADPA slurry was washed ten times with 1 mL of 20 mM phosphate buffer, pH 5.0. 40 μ L of a 5 mM solution of OTA dissolved in dimethylsulfoxide (DMSO) was mixed with a 20 times excess of EDC in 200 μ L cold deionized water and

was added to the DADPA slurry. The mixture was incubated with shaking for 1 hour at room temperature. Following coupling, the DADPA resin was washed 3 times with 0.1 M NaHCO₃, pH 8.5. 100 mg of sulfo-NHS acetate was immediately dissolved in 2 mL of 0.1 M NaHCO₃, pH 8.5 and reacted with the column matrix to quench unreacted amine groups. The mixture was incubated with shaking for another hour, followed by ten 1 mL washes with SB.

An affinity chromatography assay was set up by preparing 12 Spin-X cellulose acetate microcentrifuge filter tubes containing OTA column. 100 μ L of a 1 mg/100 μ L mixture of the OTA-immobilized column in SB buffer was added to each of the filter tubes. Buffer was removed by centrifugation at 14,000 g for 1 minute. Twelve concentrations of fluorescein-modified aptamer were prepared in SB buffer from 5 nM to 50 μ M. Each aptamer sample was heated to 90°C for 5 minutes, followed by 30 minutes of cooling at room temperature. 100 μ L of the aptamer samples were then added to the filter tubes and allowed to incubate with shaking for 30 minutes. Non-binding aptamers were removed by centrifugation followed by 5 washing steps using 100 μ L of SB. The receiving tube was changed and the bound aptamers were eluted from the column by washing twice with 100 μ L of 90°C SB. The DNA elution fractions were pooled and the fluorescence was measured (excitation wavelength 490nm; emission wavelength 520nm). The fluorescence was plotted versus the concentration of aptamer used. The K_d was determined by minimizing the residual values between calculated and observed experimental fraction of OTA.

Fluorescence polarization

Six dilutions of unmodified aptamer in SB were prepared ranging from 1 nM to 20 μ M. This DNA was heated and cooled as described above. 50 μ L of each aptamer dilution was mixed with 50 μ L of a 1 μ M mixture of OTA in SB, and allowed to incubate for 30 minutes. Following incubation, the polarization of each sample was measured using the Fluorolog Fluorescence Spectrophotometer with polarizers. Ten measurements were taken for each sample using a 1 second integration time, slit width of 10 nm and excitation 375nm; emission 430nm. The average polarization values were plotted versus the aptamer concentration and used to calculate the apparent K_d .

Ultrafiltration

To evaluate the feasibility of using centrifuge tubes for ultrafiltration as a means of determining aptamer binding affinity, 200 μ L of a 200 nM mixture of OTA in SB was loaded into a 3kDa molecular weight cut-off Amicon-Ultra centrifuge filter units. Half of this volume was allowed to pass through the membrane by centrifugation at 14,000 g for 10 minutes. Additionally, 200 μ L of a 200 nM mixture of OTA in SB was loaded into a 10kDa molecular weight cut-off Amicon-Ultra centrifuge filter units. To allow half of this volume to pass through the membrane, centrifugation at 14,000 g for 2 minutes was performed. The concentration of OTA in both the retentate and filtrate was determined as before, using fluorescence.

The stirred-cell apparatus (Model 8010, Amicon) was tested as an alternative for determining the aptamer K_d . The 1 mL minimum volume required for the apparatus, consisting of 200 nM OTA in SB was loaded into the chamber. The solution was

allowed to pass through the ultrafiltration membranes (1kDa molecular weight cutoff) using UHP argon at 50 psi of pressure. Again, the concentration of OTA in both the retentate and filtrate was determined using fluorescence.

To measure binding affinity using the stirred-cell apparatus, five 1 mL dilutions of various unmodified aptamer concentrations from 1 nM to 10 μ M were prepared. After heating and cooling the DNA as described above, 1 μ L of a 1 mM OTA solution was added to each DNA dilution. The mixtures were incubated for 30 minutes. Beginning with the lowest DNA concentration, the full 1 mL mixture was added to the stirred-cell apparatus. With the argon-based pressure, the solution was allowed to completely pass through the filter over a 10 minute period. The filtrate was collected and stored while the stirred-cell filter was rinsed 3 times with SB. This procedure was repeated for the following 4 dilutions in ascending order of aptamer concentration. The five filtrates were then measured using fluorescence to determine OTA concentration. The fluorescence of the filtrate was plotted against aptamer concentration and used to calculate K_d .

2.2.4 Standards preparation

An OTA stock solution with a final concentration of 4 μ g/ml was prepared in toluene:acetic acid (99:1; v/v). OTA solutions for spiking purposes were prepared at various concentrations ranging from 50 ng/mL to 2 μ g/mL (i.e. 124 nM to 5 μ M) by drying and reconstituting the stock solution in methanol. Additional spiking solutions between concentrations of 0.4 and 40 ng/mL (i.e. 1-100 nM) were reconstituted in binding buffer (BB). The binding buffer contained 10 mM Tris (pH 8.5), 120 mM NaCl, 5 mM KCl and 20 mM CaCl_2 . The OTA standard solutions necessary for HPLC

calibration were prepared by drying and reconstituting stock in methanol:eluting buffer (20:80, v/v) or into acetonitrile:water:acetic acid (99:99:2, v/v/v) to obtain final concentrations of 0.02 to 40 ng/mL OTA (i.e. 0.05-100 nM). Eluting buffer consisted of 10 mM Tris, pH 9 and 1 mM EDTA.

2.2.5 Preparation of the aptamer sorbent

The unmodified aptamer was conjugated to a slurry of immobilized DADPA through internal phosphate groups as previously described⁵⁰. This coupling gel was used as the sorbent material for the solid-phase extraction (SPE) columns. An aliquot of 5 mL of the DADPA slurry was washed three times with 10 mL of deionized water and once with 5 mL of 0.1 M imidazole (adjusted to a pH of 6.0 using dilute HCl). A 1 mL solution of 400 nmol aptamer in deionized water, 2.5 mL imidazole solution and 5 mL of 156 mM EDC in 0.1 M imidazole (pH 6) were added to the coupling gel and vortexed thoroughly. The gel was allowed to react for 3 hours at room temperature with rotation. To remove the supernatant, the tube was centrifuged for 2 minutes at 1000 g. The gel was then washed with 5 mL of binding buffer 4 times, discarding the washes each time. 1 mL columns were packed with 100 μ L or 200 μ L of the prepared sorbent, using polyethylene diskettes (Varian Inc.) on the bottom. The gel was allowed to settle for 30 minutes and another frit was placed on top. 3 mL columns were packed with either 200 or 300 μ L of the sorbent in the same fashion, always ensuring the sorbent remained wet. The packed SPE columns were then washed with 10 mL of binding buffer. The columns were capped with a small volume of BB and stored at 4°C until the first use.

2.2.6 Binding efficiency of SPE columns

The prepared 1 and 3 mL SPE columns were tested for their ability to bind OTA in standard solution by loading various concentrations of OTA onto the column, followed by elution. Firstly, the columns were conditioned by running 5 mL of BB through each column. After conditioning, 1 mL aliquots (loading fractions) of standard solutions (0.4, 1 and 40 ng/mL OTA in BB) were passed through the columns. Columns were then washed by with 2 mL of binding buffer (washing fraction) and the adsorbed OTA was eluted from the column with 2 mL methanol:eluting buffer (20:80, v/v) (1st eluting fraction) as previously described⁵⁰. Two additional elutions with methanol:eluting buffer were performed (second and third eluting fractions) to completely recover OTA from the column. The loading and washing fractions and the final three eluting fractions were retained at 4°C until HPLC analysis. The injection volume for HPLC analysis was 50 µL. Additional analysis was performed in the same way as described above, with 15% methanol in the final extraction dilutions to be loaded onto the columns.

The flow rates of the loading, washing and eluting steps were measured with both the 1 and 3 mL columns by running 10 mL of diluted extraction solvent through the column (loading step), followed by 6 mL BB (washing step) and 4 mL methanol:eluting buffer (20:80, v/v) (eluting step) to optimize the elution time throughout the clean-up procedure.

2.2.7 OTA determination in wheat samples

Wheat samples were extracted using procedure after Cruz-Aguado et al.⁵⁰. An amount (10 g) of ground sample was weighed and extracted with 40 mL of

methanol:water (60:40, v/v) followed by a one minute shaking period. The extracts were then filtered using Whatman No. 4 filter paper and diluted 4 times in BB. The diluted extract was then filtered through glass microfiber filters (Whatman GF/A) and cleaned-up using either the prepared SPE columns or on immunoaffinity (IMA) columns.

A 10 mL sample of diluted extract was passed through the SPE columns packed with 300 μ L of aptamersorbent. The column was washed with 6 mL of BB followed by OTA elution using 4 mL of methanol:eluting buffer (20:80, v/v). Sample extracts were stored at 4 °C until HPLC analysis. The injection volume for HPLC analysis was 50 μ L and each extraction was measured 3 times.

The performance of the aptamer-SPE columns was evaluated using 10 mL samples of diluted extraction solvent containing 0.4, 0.8, 3.8, 37.5, 75.0, 150.0 and 500.0ng OTA (i.e. 0.001-1.238 nmol OTA). The column was washed and eluted as described above.

2.2.8 Comparison with IMA columns

IMA columns were used for the clean-up of extracts following the AOAC Official Method 2000.03 for the determination of OTA in barley¹⁸³. Briefly, 10 mL of the diluted extract was run through IMA columns at a flow-rate of one drop per second, followed by washing with 2 $\times$ 5 mL wash buffer (2.5% NaCl, 0.5% NaHCO₃, 0.01% Tween 20) and 10 mL distilled water. OTA was eluted with 1.5 mL methanol and collected in a clean vial. The eluted extract was evaporated under nitrogen flow at approximately 50 °C followed by reconstitution in 500 μ L of the HPLC mobile phase. Sample extracts were

stored at 4 °C until HPLC analysis. The injection volume for HPLC analysis was 50 µL. Each extraction sample was measured 3 times.

2.2.9 Recovery experiments

Recovery experiments were performed in triplicate by spiking blank (OTA-free) durum wheat sample with OTA at concentrations of 0.5, 1.0, 5.0, 10.0 and 50.0ng/g. The SPE columns packed with 300 µL of aptamersorbent were used for clean-up as described in 2.2.7.

2.2.10 Regeneration and re-usability of SPE columns

To evaluate the re-usability of SPE columns, five new columns packed with 300 µL of aptamersorbent were re-used up to five times. A 10 mL sample of diluted extract of wheat sample spiked with 5ng/g OTA was cleaned-up using the SPE column procedure described in section 2.2.7. Following each use, the SPE columns were washed with 10 mL of methanol:eluting buffer (20:80, v/v) and 5 mL binding buffer for regeneration purposes.

2.2.11 HPLC conditions

HPLC analyses of OTA in eluates from SPE columns were carried out using the Alliance® High Throughput (HT) System (Waters Corporation, Milford, MA) equipped with a fluorometric detector (Multi-Wavelength Fluorescence Waters 2475) and the Empower™ software systems for instrument control and data processing. The analytical column was a Symmetry C18 (150mm×4.6mm, 5 µm, Waters). The flow rate of the

mobile phase was 1.0 mL/min and the column temperature was maintained at 35°C. The mobile phase consisted of a mixture of acetonitrile:water:acetic acid (99:99:2, v/v/v). OTA was eluted at a retention time of 7.0 minutes and detected with an excitation wavelength of 333 nm and emission wavelength of 460nm. Quantification of OTA was performed by integrating the relevant peak areas and comparing to the appropriate calibration curves.

2.2.12 Statistical analysis

All data was subjected to ANOVA (one-way analysis of variance). Significant differences ($p \leq 0.05$) between mean values were identified by the General Linear Model (GLM) procedure with the Duncan's and Dunnett's tests post-hoc. The data was processed using the statistical software package Statistics for Windows (StatSoft, Tulsa, OK, USA).

2.3 Results and Discussion

2.3.1 Binding affinity of aptamer 1.12.2

To confirm binding of the aptamer 1.12.2, binding studies were performed on the synthesized aptamer. Given the apparent ease of the one-point equilibrium dialysis method previously reported for this aptamer, this method was tested first. To successfully employ this method, it was necessary to first ensure that OTA was able to pass through the semi-permeable membrane; allowing equilibrium to be reached. Figure 2-2 shows the correlation between the amount of OTA (in fluorescence units) loaded into the tubes and recovered after a 48-hour equilibration. Additionally, aptamer was loaded

into the loading chamber of the equilibrium dialysis tube and allowed to equilibrate. As expected, very little aptamer was able to pass through the semi-permeable membrane, in this case 0.65% ($1.3 \times 10^{-7} \text{M}$) was detected in the receiving chamber.

Figure 2-2: Total fluorescence (CPS) of OTA measured (ex: 375nm, em: 430nm) before dialysis and in the loading/ receiving chambers following 48 hours of equilibration. Based on fluorescence, no OTA adhered to the membrane and equilibrium was reached.

The affinity experiment was then performed. However, in addition to several complications that arose using these dialysis tubes, such as bubbles becoming trapped in one of the chambers, the overall results demonstrated that this method is very inconsistent. The K_d was calculated as described in section 2.2.1 and the results of each individual equilibration can be found in Table 2-3. The average K_d ($14 \mu\text{M}$), with a large standard deviation, is 70 times greater than originally reported in 2008 ($0.2 \mu\text{M}$)⁵⁰.

Table 2-3: Dissociation constant (K_d) results of aptamer 1.12.2 with OTA from 7 trials using the one-point equilibrium dialysis assay.

Trial	K_d (μM)
1	40.7
2	15.1
3	15.6
4	14.8
5	0.78
6	10.8
7	4.86
Average	14 ± 13

To avoid the large error associated with this one-point equilibrium dialysis assay, a full binding curve was performed using the equilibrium dialysis units. While fitting this data resulted in a slightly improved K_d ($5.3 \pm 0.7 \mu\text{M}$), the data was relatively noisy (Figure 2-3). This noise was again due to the complication in using these devices such as attempting to avoid trapping bubbles in the chambers. Additionally, binding saturation could not be reached with this method, which could result in an artificially lower estimated K_d .

Since the OTA-immobilized column matrix used in the original selection was relatively facile to prepare, an affinity chromatography binding assay to elucidate the K_d was performed. Again, binding saturation was not reached and more importantly the estimated K_d was much larger than the previous studies ($97 \mu\text{M}$) (see Table 2-4 and binding isotherm in Appendix B). It was hypothesized that the 5 washing steps used to remove non-binding aptamers may have been excessive, removing the majority of the bound aptamers as well. Alternatively, the conjugation reaction of OTA to the DADPA

material may have been unsuccessful. The conjugation reaction was performed with 5 times less OTA than originally described in 2008⁵⁰.

Figure 2-3: Binding isotherm for aptamer 1.12.2 with OTA obtained using equilibrium dialysis. DNA aptamer concentration was varied with a constant concentration of OTA (200 nM). Following 48 hours of dialysis, fluorescence was determined and the fraction of OTA was calculated and plotted against DNA concentration. A theoretical curve was fit using Microsoft Excel Solver. Inset: Plot of the experimentally determined OTA fraction vs. the theoretical calculated OTA fraction.

As OTA is intrinsically fluorescent, this property was exploited for binding studies. Others in our group have had some success using fluorescence polarization. Additionally, fluorescence polarization has been very successful in measuring binding of antibodies to this toxin. Again with this method, binding saturation could not be achieved. This is likely because the aptamer does not display high enough affinity to the

target. Therefore, saturation can only be reached at concentrations that are difficult to synthesize and are outside the detection limit of the individual methods themselves. Regardless, the estimated K_d ($6.1 \pm 0.2 \mu\text{M}$) was in agreement with the estimated K_d obtained using the full curve equilibrium dialysis assay (see Table 2-4 and binding isotherm in Appendix B).

A final attempt was made to measure the binding affinity of this aptamer in the hope that either binding saturation could be achieved or a K_d value closer to the reported K_d could be estimated. Equilibrium filtration (or ultrafiltration) techniques, similar to the principle of equilibrium dialysis, have been used by many to measure the K_d values for small molecule-binding aptamers⁵⁹. The 3kDa cut-off Amicon Ultra Centrifuge devices were tested for their ability to allow OTA to pass through the membrane. Figure 2-4 shows that the majority of the OTA does not go through the filter. While a small percentage (up to 10%) of target sticking to the filter is acceptable, and can typically be corrected for in K_d determinations, these filters trapped more than 90% of the loaded OTA. The 10kDa filters did not display any improvement for this. These filters function by allowing water and small dissolved compounds to pass through the filters placed on the sides of the centrifuge device. Since OTA is only slightly soluble in water, it is possible that the high gravitation speeds prevent OTA from staying in solution and passively moving through the membrane. To test this hypothesis, the Amicon stirred-cell units that use high pressure to help force solution through the filter membranes, were used. For these devices, 90% of the OTA was able to pass through the 1kDa membranes. Using this technique, an equilibrium filtration assay was performed. The K_d for this method was also much higher than the other methods ($30 \pm 1 \mu\text{M}$).

Figure 2-4: Total fluorescence (CPS) of OTA measured (ex: 375nm, em: 430nm) loaded into the 3kDa cut-off Amicon Ultra Centrifuge devices and in the filtrate following a 10 minute centrifugation at 14,000 g. Based on fluorescence, >90% of the OTA adhered to the filter and did not pass through the filter.

We theorize that this is due to the high level of error inherent with this method.

Due to the handling difficulty and long filtration times required, only 5 concentrations were tested. For typical K_d experiments, more points should be used, specifically in the $f = 0.2$ to 0.8 range¹²⁸.

The resulting K_d values for each experiment can be found in Table 2-4 and the binding curves can be found in Appendix B. It is apparent that aptamer 1.12.2 does bind to the target OTA with relatively high affinity for a small molecular target. However, we were unable to reproduce the high affinity originally reported⁵⁰. More importantly, these

results show that binding affinity is affected by the method used to acquire the measurements.

Table 2-4: Dissociation constant (K_d) values for aptamer 1.12.2 with OTA estimated using various binding affinity methods.

Method	K_d (μM)
Reported in 2008 ⁵⁰	0.2
Equilibrium dialysis (1 point)	14 ± 13
Equilibrium dialysis (full curve)	5.3 ± 0.7
Affinity chromatography	97 ± 3
Fluorescence polarization	6.1 ± 0.2
Equilibrium filtration	30 ± 1

Since aptamer 1.12.2 showed binding, and has been previously demonstrated to be highly selective, it was used to prepare the aptamer-solid-phase extraction columns.

2.3.2 Binding efficiency of aptamer-SPE columns

The OTA aptamer 1.12.2⁵⁰ was conjugated to a coupling gel for use as aptamersorbent for the preparation of aptamer-based SPE columns. The columns were used for the clean-up and pre-concentration of OTA prior to HPLC analysis. Beginning from preliminary results reported in 2008 by Cruz-Aguado et al.⁵⁰, we optimized the clean-up procedure by examining the effect of several different parameters.

First, different combinations between the volume (100, 200 or 300 μ L) of packing aptamersorbent and the 1 mL and 3 mL sized cartridges were tested to determine the most suitable analytical clean-up conditions. According to Cruz-Aguado et al., 100 μ L of

aptamersorbent contained approximately 8 nmol of aptamer. To determine the best combination of volume and size, the prepared SPE columns were tested for their binding efficiency of OTA. The amount of OTA recovered from each column in the loading, washing and eluting fractions was measured by HPLC and is reported in Table 2-5.

1 mL Columns

1 mL columns packed with either 100 or 200 μ L of aptamersorbent showed good binding towards OTA. In the first eluting fraction almost all of the OTA (about 78%) was recovered using the methanol:eluting buffer at all the tested levels of OTA (0.1-40 ng) and only low (maximum 5%) or no OTA was detected in the second and third eluting fractions. Up to about 9% OTA was lost in the loading fractions, whereas a very low amount or no OTA was detected in the washing fraction. As the total amount of OTA recovered from columns was lower than 100% (approximately 20%), an attempt to recover this OTA was performed. Washing with up to 5 mL of 50°C eluting buffer was attempted as well as several 2 mL washes with up to 50% methanol in the eluting solvent. The lack of OTA in these washes led us to conclude that some of the OTA is irreversibly binding to the aptamersorbent under these conditions.

3 mL Columns

When 3 mL columns were packed with 200 or 300 μ L aptamersorbent, the mean OTA binding efficiency was also good (about 81%). However, compared to the 1 mL columns where the majority of the OTA was recovered in the first fraction, the 3 mL columns gave only 47% to 63% OTA recovery in the first eluting fraction. The

Table 2-5: OTA recovery from aptamer-SPE columns, using different sorbent volume and column size. Recoveries were determined by running standard solutions of different OTA concentrations through the column. Values in each column correspond to the mean of three measurements $\pm$ 1 standard deviation. n.d.= none detected

OTA standard concentration (ng/ mL)	OTA recovery (%)				
	Loading fraction	Washing fraction	1 st eluting fraction	2 nd eluting fraction	Total eluting recovery
100 μ L aptamersorbent/1 mL column					
0.4	8.7 $\pm$ 1.5	n.d.	76.6 $\pm$ 0.8	n.d.	76.6 $\pm$ 2.8
1.0	7.5 $\pm$ 2.6	n.d.	78.3 $\pm$ 3.9	0.7 $\pm$ 1.1	78.9 $\pm$ 3.2
40.0	5.7 $\pm$ 2.0	0.7 $\pm$ 0.2	78.7 $\pm$ 4.4	0.5 $\pm$ 0.01	79.2 $\pm$ 4.4
200 μ L aptamersorbent/1 mL column					
0.4	4.5 $\pm$ 0.5	n.d.	77.4 $\pm$ 1.9	n.d.	77.4 $\pm$ 1.4
1.0	4.1 $\pm$ 2.5	n.d.	78.4 $\pm$ 4.0	4.0 $\pm$ 0.9	82.4 $\pm$ 3.1
40.0	0.6 $\pm$ 0.4	n.d.	81.4 $\pm$ 1.0	1.9 $\pm$ 1.7	83.3 $\pm$ 1.9
200 μ L aptamersorbent/3 mL column					
0.4	5.7 $\pm$ 2.7	n.d.	63.4 $\pm$ 6.5	21.2 $\pm$ 3.1	84.6 $\pm$ 4.2
1.0	3.3 $\pm$ 3.4	n.d.	62.6 $\pm$ 3.9	20.0 $\pm$ 4.4	82.6 $\pm$ 1.8
40.0	1.8 $\pm$ 0.6	0.7 $\pm$ 0.2	61.6 $\pm$ 6.4	17.3 $\pm$ 5.9	78.9 $\pm$ 0.5
300 μ L aptamersorbent/3 mL column					
0.4	3.3 $\pm$ 2.8	n.d.	47.4 $\pm$ 3.6	31.3 $\pm$ 1.9	78.7 $\pm$ 4.4
1.0	1.8 $\pm$ 1.7	n.d.	60.3 $\pm$ 8.5	24.8 $\pm$ 6.6	85.1 $\pm$ 3.1
40.0	4.1 $\pm$ 1.7	n.d.	55.4 $\pm$ 0.3	24.5 $\pm$ 0.3	80.0 $\pm$ 0.4

remaining 20-25% was recovered in the second elution, but again only a small amount or no OTA was detected in the third eluting fraction. This observation is likely related to the larger number of aptamer sites present in the larger aptamersorbent volumes, resulting in a greater exposure of the OTA to aptamer binding sites. With more aptamers present, some OTA could be partially eluted from some of the aptamersorbent, but bind again to aptamers present in lower portion of the sorbent. Additionally, it is known that binding of OTA to the aptamer depends on the presence of divalent cations. In particular, as noted, OTA forms a coordination complex with calcium present in the binding buffer and this complex enhances binding to the aptamer that is highly negatively charged. The eluting solvent used contained EDTA for chelating the divalent cations, thus aiding the removal of OTA from the aptamersorbent. With the larger 3 mL columns, it is possible that the amount of EDTA in the first eluting step did not completely chelate the calcium present; therefore, an additional eluting step was required.

Despite the fact that approximately 20% of OTA was not recovered, elution with 4 mL methanol:eluting buffer gave a mean recovery higher than 77% which was considered adequate for the purpose of the study.

Binding efficiency in the presence of methanol

Cruz-Aguado et al. reported that the binding affinity between OTA and the aptamer was not significantly affected by concentrations of methanol up to 20%; however, no data on column performances in the presence of methanol was reported⁵⁰. Since methanol is used in the extraction of OTA wheat, the SPE column performance was tested in the presence of methanol. The solvent was spiked with different concentrations of OTA,

diluted 4-fold with binding buffer and cleaned-up on SPE columns. The tested mixture contained 15% methanol. The amount of OTA run through the column was compared to the amount of OTA recovered in the eluting fraction and the results are reported in Table 2-6. Recovery of OTA ranged from 81 to 90% with relative standard deviations (RSD) values < 10% when up to 500ng OTA was loaded onto the column. In particular, no significant difference ($p \geq 0.05$) was observed when comparing OTA recoveries by multivariate analysis with binding buffer (no methanol, Table 2-6) or with diluted extraction solvent (15% methanol) at the spiking level of 0.4, 1.0 and 40 ng (mean recovery values 85%, 81% and 90%, respectively). Binding capacity of the aptamer-SPE columns showed a linear ($r = 0.999$ between loaded and eluted OTA) behaviour range of 0.4-500ng OTA (0.001-1.2 nmol), indicating that the column capacity is higher than the maximum tested level of 4 nmol OTA/ml oligosorbent (Figure 2-5).

Table 2-6: Recovery and precision (relative standard deviation, RSD) results of binding experiments (n = 3 measurements) performed by passing different amounts of OTA through aptamer-SPE columns packed with 300 μ L aptamersorbent in the presence of 15% methanol.

OTA added (ng)	OTA added (pmol)	Recovery (%)	RSD (%)
0.4	1	87	8
0.8	2	81	8
3.8	9.4	90	3
37.5	93	86	4
75.0	186	85	8
150.0	372	82	2
500.0	1238	81	4

Figure 2-5: Linear regression between OTA amount loaded on and recovered from aptamer-SPE column resulting from binding experiments (n = 3 measurements).

2.3.3 Optimization of clean-up time and sample handling

Flowrate

The flow rate and the time elapsed for the entire clean-up procedure using different column sizes and aptamersorbent amount were tested and are reported in Table 2-7. Mean flow rates using the 1 mL columns ranged from 0.25 to 0.43 mL/min, corresponding to a total clean-up time of about 1 hour regardless of the aptamersorbent volume. However, the flow rates using the 3 mL SPE columns was faster, approximately 4 times that of the 1 mL columns. As a result, total clean-up time was less than 20 minutes. These results led us to continue using the 3 mL SPE columns for the clean-up

of wheat samples. For convenience, further studies were performed using the 3 mL columns packed with 300 μ L aptamersorbent. These columns, containing 24 nmol of aptamer, resulted in a 4 fold improvement in clean-up time compared to the conditions reported in 2008⁵⁰. Therefore, a sufficient batch (n = 200) of new columns was prepared and used for remaining experiments.

Table 2-7: Solvent elution flow rate and total clean-up time from 1 mL and 3 mL SPE columns packed with different volumes of aptamersorbent. Each value corresponds to the mean of three measurements $\pm$ 1 standard deviation.

Oligosorbent (μ L)	Column capacity (mL)	Elution flow rate (mL/min)			Total clean-up time (min)
		Loading fraction	Washing fraction	Eluting fraction	
100	1	0.25 $\pm$ 0.01	0.36 $\pm$ 0.03	0.25 $\pm$ 0.02	73 $\pm$ 2
200	1	0.30 $\pm$ 0.06	0.43 $\pm$ 0.03	0.29 $\pm$ 0.01	62 $\pm$ 8
200	3	1.26 $\pm$ 0.16	1.77 $\pm$ 0.25	0.87 $\pm$ 0.13	16 $\pm$ 2
300	3	1.29 $\pm$ 0.21	1.47 $\pm$ 0.06	0.78 $\pm$ 0.13	17 $\pm$ 2

Orbital shaking

Based on the stoichiometry ratio of 1:1 reported by Cruz-Aguado et al., it would be expected that columns containing 24 nmol aptamer can bind up to about 10 μ g of OTA corresponding to about 16 μ g/g OTA in wheat. Experiments at these levels were not performed simply because OTA concentrations in wheat are commonly much lower¹⁸⁴. The 3 mL columns packed with 300 μ L of aptamersorbent were applied to the purification of OTA from wheat extracts prior to determination by HPLC. Preliminary

experiments were carried out with an OTA-free wheat sample (previously analyzed by HPLC) spiked at a level of 5ng/g OTA and extracted. This was performed as previously described⁵⁰ except the 4 mL of methanol:eluting buffer was used for elution and OTA detection was by HPLC/FLD. Although a good OTA recovery (82%, n = 3) was obtained, the extraction by manual shaking is unsuitable for the analysis of a large number of samples. Additionally, the centrifugation required the transfer of samples to appropriate containers, increasing sample handling. Additional recovery experiments were performed to evaluate the efficiency of extraction by orbital shaking (from 5 to 60 min) in combination with filtration of the extracted sample. OTA recovery obtained with manual shaking were similar to that obtain using orbital shaking ($p \geq 0.05$) and ranged from 98 to 105% with relative standard deviations (RSD) ranging from 4 to 11%. The percent recovery obtained using manual extraction compared to orbital shaking between 5 and 60 minutes can be found in Figure 2-6. Extraction by orbital shaking for 5 minutes was found to be suitable for the extraction of OTA from wheat and more efficient than manual shaking.

Figure 2-6: OTA percent recovery from spiked wheat samples (5ppb) obtained using manual extraction or orbital shaking at varying times between 5 and 60 minutes.

2.3.4 Statistical overview of the aptamer-SPE clean-up

Recovery experiments were carried out on OTA-free wheat samples that were spiked at levels ranging from 0.5 to 50 ng/g. Table 2-8 shows the collected recovery and precision results. The mean recovery of OTA from wheat samples spiked at concentrations ranging from 0.5 to 50ng/g ranged from 74 to 90%, with an overall average of 80%. The RSD values for these recoveries ranged from 1 to 7%. These values fulfill the criteria set by the European Commission for the acceptability of OTA analysis (i.e. recoveries from 50 to 120% and $RSD_r \leq 40\%$ for OTA concentrations less than 1 $\mu\text{g/kg}$; recoveries from 70 to 110% and $RSD_r \leq 20\%$ for OTA concentrations in the range of 1-10 $\mu\text{g/kg}$). The limits of detection (LOD), calculated as the OTA content associated with an HPLC signal corresponding to three standard deviations ($n = 10$) from

that of the OTA-free control, was 23pg/g. The limits of quantification (LOQ), calculated as the OTA content associated with an HPLC signal corresponding to ten standard deviations (n=10) from that of the OTA-free control, was 77pg/g.

Table 2-8 Recovery and precision (relative standard deviation, RSD) results of ochratoxin A (OTA) from spiked wheat (n= 3 measurements) after clean-up on solid-phase-extraction (SPE) columns packed with 300 μ L oligosorbent (24 nmol DNA aptamer).

Added (ng/g)	Average detected (ng/g)	Recovery (%)	RSD (%)
0.5	0.45	90	1.1
1.0	0.84	84	1.2
5.0	3.83	77	4.4
10.0	7.24	74	4.5
50.0	36.46	75	5.2

2.3.5 Comparison of aptamer-SPE column clean-up with IMA column clean-up

The applicability of SPE columns to the clean-up of wheat sample extracts was tested on 33 samples of naturally contaminated wheat flour. Since no official method for the determination of OTA in wheat has been reported, the results were compared to those obtained using IMA column clean-up for the AOAC Official Method 2000.03¹⁸³ for the determination of OTA in barley. The results are presented in Figure 2-7. Mean OTA levels ranged from not detected to 9.5ng/g in samples analyzed by SPE/HPLC and from 0.3 to 10.7ng/g in samples analyzed by IMA/HPLC. This difference was likely caused by the different OTA recovery of the two methods. The SPE/HPLC method showed a lower sensitivity as compared to IMA/HPLC as indicated by the slope of the equation (0.891) that was significantly different ($p < 0.05$) from 1. However, a good correlation (r

value, 0.990) was observed between OTA values obtained by using the two different clean-up methods. These findings suggest that the SPE clean-up in combination with HPLC/FLD could be considered a suitable method for the purification and pre-concentration of OTA prior to HPLC analysis in wheat samples at levels that naturally occur in wheat. Figure 2-8 shows the chromatograms of an OTA-free wheat sample (lower line) and of a wheat sample naturally contaminated with 1.5ng/g OTA. The clean-up using aptamer-SPE columns provided a good chromatographic profile due to the specificity of the aptamer contained in the column. Therefore the SPE/HPLC method can be used as an alternative method to HPLC/IMA clean-up.

Figure 2-7: Comparison of ochratoxin A (OTA) content determined by solid-phase-extraction (SPE) column-clean-up versus immunoaffinity (IMA) column clean-up. Each point corresponds to the mean of three measurements. Results are reported without correcting for recovery.

Figure 2-8: Chromatograms of an ochratoxin A (OTA)-free wheat sample (lower line) and a wheat sample naturally contaminated with 1.5ng/g OTA (upper line). Sample extracts were cleaned-up on SPE columns containing 300 μ L oligosorbent (24 nmol DNA aptamer).

2.3.6 Re-usability of aptamer-SPE columns

The reusability of the aptamersorbent for OTA analysis in wheat samples was also investigated. Five concentrations of diluted samples were loaded on five different columns for five consecutive clean-up procedures. There is no change ($p > 0.05$) in extraction recovery between the first and the five subsequent uses. Assuming a 100% OTA recovery after the first use, the mean OTA recovery over the consecutive 4 uses ranged from 92 to 109% (Figure 2-9) with RSD values ranging from 2 to 7%. Based on

these results, it was concluded that SPE columns packed with 300 μ l aptamersorbent could be re-used at least 5 times without affecting the binding affinity of the aptamer towards OTA. Further experiments could be performed to test the performance of 3 mL columns packed with 200 μ L aptamersorbent to reduce the amount of aptamersorbent required.

Figure 2-9: Normalized OTA recovery for 1 mL SPE columns after 2 to 5 re-uses.

2.4 Conclusion

While the high binding affinity of aptamer 1.12.2 as previously reported was not reproducible using the above experiments, binding studies consistently indicated sufficient affinity for OTA. Moreover, we have investigated several methods for K_d

determination, recognizing that some methods are better suited for different target types. More importantly, it has become clear that K_d may vary by affinity method, indicating that full characterization of new aptamers should involve more than one method to confirm high affinity.

Regardless of the measured binding of aptamer 1.12.2 for OTA, SPE columns were prepared with 24 nmol of this aptamer covalently immobilized. With this concentration of aptamer, even aptamers with moderate affinity would be sufficient in OTA extraction of concentrations up to 500ng (using the 300 μ L columns). Assuming a dissociation constant of 5 μ M, 94% of the 500ng of OTA would be bound by this concentration of aptamer. As a result, the aptamer-SPE columns showed good performance in the range 0.4-500 ng of OTA and were then applied to the clean-up of OTA from wheat samples. Recovery and precision data of the aptamer-SPE/HPLC method were acceptable and within the criteria established by the European Union. Comparative analysis of several naturally contaminated wheat samples using IMA clean-up and aptamer-SPE clean-up showed a good correlation between the two clean-up procedures, suggesting that the aptamer-SPE/HPLC method can be applied to the quantitative determination of OTA in unprocessed wheat at levels below the maximum admissible levels proposed by the European Commission for durum wheat and could be potentially useful for the management of OTA in commodities. The successful outcome of this research provides evidence that DNA aptamers can be used to replace antibodies as molecular recognition and binding reagents in diagnostic assays for commercial applications for food testing. Future work will evaluate the applicability of the proposed method for the analysis of OTA in other food matrices. Further research is needed to

verify if other aptamer sequences and aptamersorbent preparations might have improved operational characteristics. In particular, new aptamers for OTA that show an improved affinity for the target should be developed.

Chapter 3. Selection and binding assessment of new OTA aptamers

3.0 Statement of contributions

Selection of the putative aptamer sequences was performed at the Department of Medical Chemistry, Semmelweis University in Budapest, Hungary by Viola Bardóczy et al. The binding affinity experiments were performed by M. McKeague.

3.1 Introduction

3.1.1 Multiple aptamers for the same target

A typical SELEX library contains on average 40 random nucleotides (N_{40}). With this number of random positions, the size of the sequence space (the number of possible sequences) is 4^{40} or 1×10^{24} . However, this number of DNA sequences would correspond to 24 kg of DNA, an expensive and impractical amount to synthesize. Therefore, practical and technical considerations limit the number of random sequences that can actually be used in SELEX to approximately 10 nmoles, which corresponds to approximately 10^{16} different sequences. As a result, the sequence space is substantially unrepresented⁴⁷. Despite this statistical limitation, aptamer selections often succeed. For example, the RNA ATP aptamer binds to the adenosyl (Ado) moiety of ATP and AMP with submicromolar affinity ($K_d = 0.7 \mu\text{M}$ in solution)¹⁸⁵. The same motif was then observed in aptamers selected to other targets such as nicotinamide adenine dinucleotide (NAD^+)¹⁸⁶ and S-adenosyl methionine (SAM)¹⁸⁷. The repeated isolation of this motif for adenosine-containing targets (see Figure 3-1) exemplifies reproducibility and success of SELEX experiments.

Figure 3-1: Structures of adenosyl containing compounds – adenosine monophosphate (AMP), nicotinamide adenine dinucleotide (NAD⁺) and S-adenosyl methionine (SAM).

Regardless, there are several factors that may affect the result of *in vitro* selection experiments. Larry Gold, the co-inventor of SELEX, indicated that less than 30% of selections are successful⁵¹. Many experimental variables must be considered for SELEX, including pH, temperature, choice of buffer, use of an organic co-solvent, metal ion concentrations and incubation time. The choices for all of these variables depend on the specific target being used, and there is no general rule for optimal selection conditions⁴⁷. Therefore, some selections for the same target may be more successful than others. As a result, researchers continue to perform new selections to find better aptamers for targets when aptamers for such targets already exist.

As mentioned, one of the common aptamer targets is thrombin. The first thrombin aptamer was selected by Bock et al. in 1992. This DNA aptamer is 15 nucleotides in length, folds into a G-quadruplex and binds to the fibrinogen-recognition site of thrombin with a K_d value of 100 nM¹⁴⁵. However, a few years later, Macaya et al.

identified a DNA aptamer containing a duplex that bridges the ends of the quadruplex structure. This aptamer improved the binding affinity by an order of magnitude ($K_d = 10\text{-}25\text{ nM}$)¹⁸⁸. In 1997, Tasset et al. then repeated the selection of thrombin aptamers, using filtration as the partitioning step rather than the affinity chromatography originally employed. From this selection, an aptamer displaying a K_d of 0.5 nM was selected that was found to bind to the heparin recognition exosite of thrombin¹⁸⁹. In 2001, White et al. selected RNA aptamers displaying similar affinity (1-4 nM) using a toggle-SELEX approach⁷⁰. More recently, in 2010, Miyachi performed AFM-SELEX and identified a new thrombin aptamer sequence that displayed a 1000 fold improvement in binding over the original selected aptamer ($K_d = 200\text{ pM}$)⁷⁶. The apparent increase in aptamer binding from this succession of selection experiments gives promise that using novel selection techniques can result in the discovery of improved aptamers for the same target.

However, the efforts made to select improved aptamers by new methods had mixed success. For example, several groups have employed new and fairly extensive SELEX methods at an attempt to discover improved aptamers for the target streptavidin. Tahiri-Alaoui et al. first selected streptavidin binding aptamers displaying a K_d of 7 nM¹⁹⁰. Then, in 2005, a selection called FluMag SELEX that involved selection on magnetic beads in place of agarose yielded aptamers displaying affinities of only 56 to 85 nM⁹⁰. Additionally, Bittker et al. performed several selections using non-homologous recombination and error-prone PCR steps in the selection process and yielded aptamers with only a modest improvement of binding (3.3 nM)¹⁹¹. Regardless, due to the selection conditions available, the variable stringency strategies possible and the inability to

sample the entire sequence space, new selections are a worthwhile effort to discover novel sequences for the same target.

3.1.2 Nanoparticle binding assay

As evidenced in Chapter 2, K_d determination is often the bottle-neck in aptamer development, particularly for aptamer-small molecule interactions. Therefore the possibility of using a new method for K_d determination, involving gold nanoparticles (AuNPs), was investigated. Gold nanoparticles are described as a suspension (or colloid) of sub-micrometre sized particles of gold in water or other liquid. When the suspension of gold is on the nanometer scale, the prominent spectroscopic feature of surface plasmon resonance is observed. Because gold nanoparticles have tunable optical properties and some of the highest extinction coefficients reported, they are useful in colorimetric detection¹⁹².

The ability to tune the optical properties of gold is not possible in the bulk material. In colloid form, however, gold nanoparticles are in a dispersed state due to the negatively charged citrate molecules present in the suspension. The negatively charged citrate ions act as capping agents of the nanoparticles and result in relatively large inter-particle distances, which results in the visualization of a red colour. Aggregation of the nanoparticles can result from varying the temperature, pH and ionic strength of the solution. Manipulation of these properties promotes the nanoparticles to come closer together, causing the inter-particle distance to be smaller than the average diameter of the particle. This aggregated state results in a change of the inter-particle plasmon coupling which generates a significant absorption band shift in the visible region of the

electromagnetic spectrum. Therefore a change in colour from red to purple or blue is observed¹⁹³.

Several methods exist for gold nanoparticle synthesis. The Turkevich method, named after John Turkevich for his work on gold nanoparticles in 1951¹⁹⁴, is one of the most simple. Generally, this method produces monodisperse spherical gold nanoparticles of approximately 10–20 nm in diameter. In this method, sodium citrate ($\text{Na}_3\text{C}_6\text{H}_5\text{O}_7$) is reacted with chloroauric acid (HAuCl_4). The citrate ion ($\text{C}_6\text{H}_5\text{O}_7^{3-}$) oxidizes, forming acetone dicarboxylic acid ($\text{C}_5\text{H}_6\text{O}_5$), which can then form a complex with Au(III). Reduction of Au(III) to Au(0) occurs when the complex reaches a critical mass that is greater than its thermodynamic stability. This forms a suspension of nuclei that continues to grow with the addition of more gold particles. Once the gold is consumed, this growth is terminated. The sodium citrate also acts as a capping agent that stabilizes the colloid suspension by electrostatic repulsion¹⁹⁵.

Due to their optical properties and ease of use and synthesis, gold nanoparticles have been incorporated into a number of novel DNA sensor strategies¹⁹⁶. In 2004, Li and Rothberg discovered that single-stranded DNA (ssDNA) adsorbs onto citrate-coated gold nanoparticles but double-stranded DNA (dsDNA) does not¹⁹⁷. This was a remarkable observation that has since been exploited in several biosensor platforms. While the reasons for this observation have yet to be verified, the recent theory is that the electrostatic interaction between ssDNA and gold nanoparticles is of little importance to the mode of interaction and the largest effects are from solvation forces, in particular, the hydrophobic effect¹⁹⁸.

This difference in interaction of ssDNA and dsDNA with nanoparticles has been applied to aptamer-based biosensors. For example, Wei et al. were able to use the thrombin aptamer to adsorb to the surface of AuNPs, preventing their aggregation upon addition of salt. However, with the addition of thrombin, the aptamers preferentially interact with the thrombin and fold into a rigid tertiary structure that is unable to adsorb around the AuNP surface. This leaves the AuNPs susceptible to salt-induced aggregation causing them to turn blue¹⁹³. The same concept was applied to the DNA aptamer for adenosine¹⁹⁹ and more recently for OTA¹⁶². Figure 3-2 describes this biosensing scheme. The ability of these aptamer-nanoparticles sensing schemes to detect small concentrations of their target provide precedent that a similar platform could be used to determine the binding affinity of newly developed aptamers.

Figure 3-2: Gold nanoparticle aptamer assay for target detection. A: Sodium citrate stabilized AuNPs undergo salt-induced aggregation, resulting in a blue-purple colour. **B:** Single-stranded DNA aptamer coils around the AuNP surface, preventing salt-induced aggregation, therefore the colloid solution remains red. **C:** Upon addition of the aptamer's target, the aptamer folds into a rigid structure around the target, preventing the adsorption of the ssDNA on the AuNP surface. Therefore, salt-induced aggregation returns and a blue-purple colour is observed.

3.1.3 Chapter objectives

This chapter investigates the possibility of selecting new aptamers that bind to the target OTA with high affinity. Such aptamers may be useful in developing improved aptamer-based biosensors and extraction columns for this toxin. The selection strategies as well as the binding characterization of the aptamers are described.

3.2 Experimental

3.2.1 Materials

All buffers were prepared with Millipore Milli-Q deionized water at 18 M Ω cm. Buffering agents, salts, reagents for nanoparticles synthesis and OTA from *Petromyces albertensis* $\geq 98\%$ (TLC) for binding studies were purchased from Sigma-Aldrich. All molecular biology grade electrophoresis chemicals were purchased from BioShop Canada (Burlington, Ontario). EAH sepharose 4B was obtained from GE Healthcare. The streptavidin-coated high-binding capacity plate was purchased from Thermo Fisher Scientific (Rockford, IL, USA). The DNA Taq polymerase and PCR components were purchased from Invitrogen. A pGEM-T-easy vector by T/A cloning kit was received from Promega (Madison, WI, USA). Phosphoramidites, modifiers, activator, deblock, capping, and oxidizing reagents were obtained from Glen Research. Standard support columns and acetonitrile were purchased from BioAutomation. Ultra High Purity 5.0 argon was purchased from Praxair. Sulfo-NHS acetate and Amicon-Ultra 0.5 mL 3kDa centrifuge units were purchased from Fisher Scientific Canada. Sample Dispo Equilibrium Dialysis units, 5000MWCO, were purchased from Harvard Apparatus (Holliston, Massachusetts, USA). CarboxyLink Coupling gel (immobilized

diaminodipropylamine, DADPA, supplied in a 50% slurry) and 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (EDC) were received from Pierce (Rockford, Illinois, USA).

3.2.2 Preparation of OTA resin

Ochratoxin A was coupled to EAH Sepharose 4B according to manufacturer's protocol. Briefly, a 5 mM solution of OTA in deionized water was prepared and adjusted to a pH of 4.5. The OTA solution was mixed with sepharose at a 1:1 v/v ratio. 0.1 M of EDC was prepared in water with the pH adjusted to 4.5 and added to the OTA and sepharose mixture. The reaction was incubated at room temperature for 4 hours with gentle shaking. After the first 3 hours, the pH was again adjusted to 4.5. After the incubation period, the gel matrix was washed with three cycles of alternating washes consisting of 0.1 M acetate buffer pH 4.0 containing 0.5 M NaCl and 0.1 M Tris-HCl buffer pH 8 containing 0.5 M NaCl. The OTA resin was stored at 4°C until use.

3.2.3 OTA SELEX experiments

The oligonucleotide library was synthesized on a 1 μ mol scale and purified using RP-HPLC (Metabion, Martinsried, Germany). This library was composed of 40 random nucleotides flanked by two fixed regions at the 5' and 3' ends. 5'-AGCCTCGTCTGT TCTCCC-40N-GGGAAGACAAGCAGACGT-3'. The forward primer, SELF, 5'-AGC-CTCGTCTGTTCTCCC -3', and reverse primer, SELR, 5'-BIOTIN-ACGTCTGCTT-GTCTTCCC-3', annealing to the flanking regions were also synthesized and prepared for the selection procedure.

SELEX was performed as described by Murphy et al. with some modification²⁰⁰. The SELEX library (1 nmol) was dissolved in selection buffer (SB) containing 10 mM Na₂HPO₄, 2 mM KH₂PO₄, 2.7 mM KCl, 137 mM NaCl and 0.05% Triton-X, pH 7.4. This library was incubated with 10 µL of the OTA resin and allowed to shake at room temperature for 1 hour. The resin was washed 3 times with the SB to eliminate non-specifically bound sequences. The putative aptamers were then eluted using 10 µL of 50 mM NaOH and cleaned using a phosphate buffered saline (PBS)-equilibrated desalting column and transferred directly to a PCR reaction.

The 100 µL PCR reaction mixture contained 10 µL of 10 x reaction buffer, 1.5 units of Taq polymerase, 1 M of the two primers (SELF and biotinylated SELR), 0.8 mM of each deoxyribonucleotidetriphosphate (dNTPs) and 1.5 mM MgCl₂. Amplification was performed by heating to 95°C for 2 minutes followed by 18 cycles of 30 seconds at 95°C, 30 seconds at 56°C and 30 seconds at 68°C, followed by a final extension at 68°C for 2 minutes. After checking the PCR product quality on 2.5% agarose gel using 10 µL samples, 23 µL of 5 M NaCl was added to the remaining 90 µL PCR mixture. This solution was divided into two parts and incubated with a Reacti-Bind streptavidin-coated high-binding capacity plate (Thermo Fisher Scientific, Rockford, IL, USA) then washed 3 times using 1 mL of PBS-Tween. The nonbiotinylated strands were separated from the immobilized complementary strand by incubating with a 30 µL/well aliquot of fresh 100 mM NaOH for 5 minutes. The ssDNA was removed and diluted using SB. The pH was adjusted and the ssDNA solution was heated to 95°C for 5 minutes and immediately placed on ice and kept at 4°C until the next round of SELEX.

10 μ L of the OTA resin was used in cycles 1 through 7. To increase the stringency of selections, only 5 μ L of OTA resin was used in cycles 8 through 11 and 2.5 μ L in cycles 12 through 15. Negative selections were performed using unmodified EAH Sephadex 4B at cycles 4, 8 and 12. After cycle 15, the final PCR of the selected ssDNA sequences was performed using the SELF and nonbiotinylated SELR primers. 4 μ L of this PCR mixture was ligated into p-GEM-T-easy vector by T/A cloning and transformed into DH5 α *Escherichia coli* competent cells. Twenty-one colonies were selected and tested for the presence of the correct inserts using colony PCR with M13 forward and reverse primers. The PCR products were purified by polyethylene glycol precipitation and sequenced using the M13 promoter primer²⁰¹.

3.2.4 Screening of binding affinity

Aptamer synthesis

Aptamer sequences A08, B08, D07 and E08 were synthesized, unmodified, as described in section 2.2.2 (sequences are found in Table 3-2). Aptamer A08 was also synthesized with the 5' fluorescein modifier, 6-FAM. All aptamers were purified by PAGE as described in section 2.2.2.

Equilibrium dialysis

The same procedure used by Cruz-Aguado et al. (described in section 2.2.3) was used to screen four of the selected aptamers for binding affinity to OTA⁵⁰. Single-use Sample Dispo Equilibrium Dialysis units with a 10,000Da molecular weight cut-off were used. To ensure the aptamer sequences were unable to pass through the semi-permeable

membrane, 75 μ L of a 20 μ M solution of aptamer A08 in SB was added to the loading chamber. After the 48-hour incubation time, the concentration of DNA in the receiving chamber was measured using UV-Visible spectroscopy.

For the equilibration studies, several variants of the original selection buffer (SB) were used. Firstly, SB was used as listed in section 3.2.3. However, the Triton present in this original buffer resulted in the production of bubbles upon agitation of the dialysis units. Since the presence of these bubbles may affect equilibrium, SB without Triton was also used. Additionally, to ensure that OTA could be solubilized, SB with 2.5% DMSO in place of Triton was also tested. All 4 unmodified aptamer sequences were individually dissolved in these 3 buffers at a concentration of 20 μ M and mixed with 200 nM OTA. 75 μ L of this solution was loaded into the loading chambers of separate dialysis units with 75 μ L of the corresponding buffer solution in the receiving chamber. The units were allowed to shake at 60 rpm while lying flat for 48 hours to permit equilibration between the chambers. Following the 48-hour agitation period, fluorescence measurements of the loading and receiving chambers were obtained by direct fluorescent measurement (excitation wavelength 375nm; emission wavelength 430nm) using a Fluorolog Fluorescence Spectrophotometer (Horiba Jobin Yvon, USA) with a SpectrAcq controller.

The fraction of bound OTA (f) was determined as:

$$f = \frac{F_l - F_r}{F_l}$$

and the dissociation constant (K_d) was estimated as:

$$K_d = \frac{[A_0]}{f} - [A_0]$$

Previous reports indicate that increased concentrations of Ca^{2+} increased the affinity of the reported aptamers for OTA. Investigation was performed to determine whether the newly selected aptamers, A08 and D07, displayed a similar effect. The SB used in these selections contain phosphate that when mixed with calcium salts forms insoluble calcium phosphate compounds { CaHPO_4 solubility = 0.004303 g/100 g H_2O at 20°C and $\text{Ca}_3(\text{PO}_4)_2$ solubility = 0.002 g/100 g H_2O at 20°C }. Therefore the buffer used by Cruz-Aguado et al.⁵⁰ in 2008, containing HEPES, was used so that there would be no precipitation of calcium salts. However, aptamers can be sensitive to the conditions in which they were selected, therefore multiple buffer conditions were investigated. The conditions are summarized in Table 3-1.

Table 3-1: Buffer composition used in the equilibrium dialysis binding screen for OTA aptamers.

Buffer	Composition
SB, no-Triton	10 mM Na_2HPO_4 , 2 mM KH_2PO_4 , 2.7 mM KCl, 137 mM NaCl, pH 7.4
SB, no-Triton DMSO	10 mM Na_2HPO_4 , 2 mM KH_2PO_4 , 2.7 mM KCl, 137 mM NaCl, 2.5% DMSO, pH 7.4
2008 Buffer⁵⁰	10 mM HEPES, 5 mM KCl, 120 mM NaCl, pH 7.0
2008 Buffer DMSO⁵⁰	10 mM HEPES, 5 mM KCl, 120 mM NaCl, 2.5% DMSO, pH 7.0
Combination buffer	10 mM HEPES, 2.7 mM KCl, 137 mM NaCl, pH 7.4
Combination buffer DMSO	10 mM HEPES, 2.7 mM KCl, 137 mM NaCl, 2.5% DMSO, pH 7.4

Using the combination buffer and combination buffer with 2.5% DMSO, the equilibrium dialysis tests were repeated for aptamer A08 and D07 in the presence of 0mM, 10 mM and 20 mM CaCl_2 .

Affinity chromatography

OTA-immobilized column material was prepared as reported by Cruz-Aguado et al.⁵⁰. A 1 mL aliquot of DADPA slurry was washed ten times with 1 mL of 20 mM phosphate buffer, pH 5.0. 40 μ L of a 5 mM solution of OTA dissolved in dimethylsulfoxide (DMSO) was mixed with a 20 times excess of EDC in 200 μ L cold deionized water and was added to the DADPA slurry. The mixture was incubated with shaking for 1 hour at room temperature. Following coupling, the DADPA resin was washed 3 times with 0.1 M NaHCO₃, pH 8.5. 100 mg of sulfo-NHS acetate was immediately dissolved in 2 mL of 0.1 M NaHCO₃, pH 8.5 and reacted with the column matrix to quench unreacted amine groups. The mixture was incubated with shaking for another hour, followed by ten washes with 1 mL SB.

An affinity chromatography assay was set up by preparing 14 Spin-X cellulose acetate microcentrifuge filter tubes containing OTA-modified column. 100 μ L of a 1 mg/100 μ L mixture of the OTA-immobilized column in SB buffer was added to each of the filter tubes. Buffer was removed by centrifugation at 14,000 g for 1 minute. Fourteen concentrations of fluorescein-modified A08 aptamer, from 1 nM to 50 μ M, were prepared in SB buffer. Each aptamer concentration was heated to 95°C for 5 minutes, followed by 30 minutes of cooling on ice. 100 μ L of the aptamer concentrations was then added to the filter tubes and allowed to incubate with shaking for 30 minutes. Non-binding aptamers were removed by centrifugation followed by 3 washing steps using 100 μ L of SB. The receiving tube was changed and the bound aptamers were eluted from the column by washing twice with 95°C SB. The DNA elution fractions were pooled and the fluorescence was measured (excitation wavelength 490nm; emission

wavelength 520nm). The fluorescence was plotted versus the concentration of aptamer. As before, the K_d was determined by minimizing the residual values between calculated and observed experimental fraction of OTA using the solver feature of Microsoft Excel.

3.2.5 Dissociation constant (K_d) determination using gold nanoparticles

Gold nanoparticle synthesis

All glassware used for AuNP synthesis was cleaned by soaking in aqua regia (3:1 mixture of concentrated HCl/HNO₃) for 15 minutes followed by thorough rinsing with deionized water. A 250 mL Erlenmeyer flask was used to mix 98 mL of deionized water and 2 mL of 50 mM HAuCl₄ for a final concentration of 1 mM HAuCl₄. The solution was heated to boiling with magnetic stirring. Upon boiling, 10 mL of 38.8 mM sodium citrate was added. Following a change in suspension colour from pale yellow to dark blue and finally to a wine red, heating was continued for an additional 20 minutes. The flask was removed from the heat and allowed to cool to room temperature with continued stirring. The nanoparticles were characterized by UV-Visible spectrometry.

Binding assay using gold nanoparticles

The assay was performed as described by Zheng et al. with modifications²⁰². Five samples of OTA in water were prepared at various concentrations from 10 nM to 100 μ M. 243 μ L of the OTA dilutions or a blank solution of deionized water was added into microcentrifuge tubes containing 6 μ L of either unmodified aptamer 1.12.2 (10 μ M in 2008 buffer⁵⁰) or unmodified aptamer A08 (10 μ M in SB buffer). After a 30 minute incubation period with shaking, 135 μ L of the AuNP solution (10.5 nM) was added,

vortexed briefly and allowed to react for 5 minutes. 72 μL of 0.25 M NaCl was then added to each microcentrifuge tube and allowed to incubate for a further 5 minutes. Each sample was then diluted 4 times with water and analyzed using UV-Visible spectroscopy. After normalization, the absorbance at 705 nm was plotted against OTA concentration. The K_d was evaluated by minimizing the residual values between calculated and observed absorption using the solver feature of Microsoft Excel^{181, 182}.

Controls were performed to ensure that OTA did not cause aggregation of the AuNP in the absence of NaCl. The procedure was performed as above using a 100 μM dilution of OTA and using deionized water in place of the DNA and salt. To ensure that OTA also did not prevent the salt-induced aggregation of the AuNPs, the procedure was repeated as above but adding the 0.25 M NaCl to observe aggregation.

3.3 Results and Discussion

3.3.1 Selection of OTA aptamers

Fifteen rounds of selection were performed, using relatively high stringency conditions compared to previous OTA selections¹¹. While the same number of counter-selections, 3 overall, were performed in both selections, the selection performed in 2008 employed a total of 13 positive selection rounds (2 less than in this SELEX).

Additionally, Cruz-Aguado et al. performed SELEX in the presence of a large excess of OTA, using 250 μL of OTA-resin each round, compared to the 10 μL or less used throughout this selection. While there is not a specific correlation between the number of selection rounds performed and the binding affinity of the selected aptamers, studies have

shown that using less resin, and therefore less target, drastically increases the stringency of the selection and often results in higher affinity aptamers²⁰³.

The 21 clones that were sequenced after round 15 of the OTA selections are shown in Table 3-2¹¹.

Table 3-2: The putative aptamer sequences (of the N₄₀ region) cloned after round 15 of the SELEX experiment. The number of clones present is indicated in the column on the right.

Name	Sequence (5'-AGCCTCGTCTGTTCTCCC-N₄₀- GGGAAGACAAGCAGACGT-3')	Number of clones present
A08	GGCAGTGTGGGCGAATCTATGCGTACCGTTCGATATCGTG	2
B08	GGCGCATGATCATTCCGGTGGGTAAGGTGGTGGTAACGTTG	2
D07	GGCCAGGAATCCATCTGTCCAGGGACAAGCAGCTAGGGTT	2
E08	GGCGGTGTCTGGCGTATGTCGTGGGTGGGGTGCGAGTGC	2
D09	GGGTGAGTGGGCATTGGGTCCAGGGGTCTGTAAACGGGTCG	1
H08	GGGGTGGGTGTCTCGGGGAGGAGTTCCTGCGGCAGCGGTGG	1
E09	GGCCACGTCTTGGTGCTCGCCGACACGCGTTTGTGGAGTG	1
F07	GCCTGTTCTTCGAGCGCCCGACCCCAAGAATGCGCAGGTAT	1
A09	AGGGGGGATAGGTTATGAGTTGTGGGGGAATTTTGATGTT	1
B09	AATAGGGGAGGAAAGAAGGATGTGTAATAAAAGGGTGAGT	1
A07	AAAAGAGATTGTTTGTGATTTGGTTTGGTGTAATTTGG	1
D08	GAAAGTGATTTATGAAGTTAAAGAGAATAAGGTTTGGATG	1
F08	AAAAAGAAAATTGTTGTTGAGATGTTTTGGAATTGAGTGG	1
F09	AAAGAGGTTGTGAAAGAAGAATGGGTGAAATGTGAGGTTG	1
G09	AAGTAAGGATAAGATATGGAGTTGGTGGGATAAGTAAGTG	1
H07	GAGATGGATGATATGAAAAAATGATTGGGAAAGAAAGGTG	1
H09	GGAGGAAAATTAAGAGTTAATAAGTGATTAAGAGGTAATG	1

Since aptamers A08, B08, D07 and E08 were present multiple times in the sequenced library, these aptamers were synthesized for further analysis. No sequence homology was found between these 17 different sequences. This was not surprising as typically hundreds of individual clones are required to identify consensus aptamer sequences⁴⁷. The cost of sequencing can be quite high (\$20 per sequence) and therefore it is not always practical to sequence so many clones. However, even when dozens of

colonies can be sent for sequencing, it is not uncommon to find a lack in sequence homology, or to find a relatively low sequence incidence of each clone. Regardless, many of these selections have yielded high affinity binders for their target⁵⁴.

Despite the lack of sequence homology observed in the 17 sequences obtained for SELEX, it is interesting that the first 4 putative aptamer sequences contain a relatively high incidence of deoxyguanine (G) residues. Aptamer 1.12.2 contains 47.2% G. Completely random sequences should contain 25% G: however, A08 consists of 35% G, B08 42.5% G, D07 32.5%G and E08 52.5% G. While high contents of G are common in DNA aptamers^{81, 145}, the continued selection of G-rich sequences suggests that the target OTA may have an affinity towards G-rich motifs.

The secondary structure predictions for sequences A08, B08, D07 and E08 were performed using the Mfold software²⁰⁴, using corrected folding parameters of 157 mM Na⁺ concentration and temperature of 25°C²⁰⁵. The structures found to have the lowest minimum free energy can be found in Figure 3-3²⁰⁶. Each predicted structure indicates 2 or 3 stem-loop motifs, very common to aptamers. The high incidence of guanine residues is usually characteristic of G-quadruplexes; however, G-quadruplex structures cannot be predicted using the Mfold software and are therefore not shown in these predicted structures.

Figure 3-3: Secondary structure prediction for the full length putative OTA aptamer clones. Image generated by Mfold.

3.3.2 Screening of the binding affinity

Equilibrium dialysis

Since K_d and structure determination are typically labour-intensive processes, aptamer clones are first screened to determine their relative target-binding properties. From this, the “winning” sequences are then studied in greater detail. The one-point equilibrium dialysis assay, despite the previous trouble experienced, was a good starting point for this screening. As before, 20 μ M aptamer (in this case of sequence A08) was loaded into the loading chamber of the equilibrium dialysis tube and allowed to equilibrate. As expected, very little aptamer was able to pass through the semi-permeable membrane; in this case, 0.36% (7.1×10^{-8} M) was detected in the receiving chamber. This indicated that the amount of DNA lost to the receiving chamber would be negligible and could be excluded from the calculations.

The initial screening was performed in the buffer used in SELEX (SB), however, after a few minutes of incubation, the Triton X-100 had produced significant bubbles in the individual chambers. As bubbles can prevent equilibrium from being reached, the assay was repeated without Triton X-100. However, as OTA is only slightly soluble in water, a third trial in the presence of 2.5% DMSO was performed as well. The results of this screening can be found in Table 3-3.

Despite the large variation for each measurement, sequence A08 and D07 showed the highest binding affinity. Using various buffer conditions and salt concentrations, several dialysis experiments were performed on these two sequences and aptamer 1.12.2, for comparison purposes.

Table 3-3: Results of the one-point equilibrium dialysis method for screening OTA putative aptamer sequences A08, B08, D07 and E08. The K_d in the various buffer conditions is reported. The average K_d of all 3 trials with the calculated standard deviation is detailed. As the trials using the SB with Triton resulted in significant bubbles, the average is also reported excluding this initial value.

	Calculated K_d (μM)			
	A08	B08	D07	E08
SB (with Triton X-100)	12.7	0.69	0.39	22.2
SB (no Triton X-100)	0.769	10.6	0.45	29
SB (with 2.5% DMSO)	1.8	54	1.6	43
Average K_d (all points)	5 ± 6	22 ± 28	0.8 ± 0.7	31 ± 11
Average K_d (excluding Triton buffer)	1.2 ± 0.7	32 ± 30	1.0 ± 0.8	36 ± 10

Table 3-4: Results of the one-point equilibrium dialysis method for putative OTA aptamer sequences A08, and D07 compared to aptamer 1.12.2. The K_d in various buffer conditions is reported. The average K_d of all the trials with the calculated standard deviation is detailed.

	Calculated K_d (μM)		
	1.12.2	A08	D07
SB (with Triton X) Trial 1	10.8	12.7	1360
SB (no Triton X)	4.86	38.1	0.452
SB (2.5% DMSO)	240	0.769	0.391
SB (with Triton X) Trial 2	190	3.07	*
2008 buffer ⁵⁰	210	66.2	1.29
Combination buffer Trial 1	160	0.124	3.45
Combination buffer (2.5% DMSO)	160	50.3	86.5
Combination buffer Trial 2	40.7	0.510	760
Combination buffer (10 mM Ca^{2+})	15.6	0.47	2430
Combination buffer (20 mM Ca^{2+})	14.8	1.75	810
Average K_d (all points)	105 ± 95	17 ± 25	606 ± 844

*Calculated a negative K_d value

The results of the screening were extremely inconsistent and showed a large spread in calculated K_d values. The inconsistency of this method is evident in the large change in calculated K_d values observed in the second trials of repeated dialysis conditions. For example, as shown in Table 3-4, the second dialysis trial using SB with Triton X-100 resulted in a 20x increase in K_d value for aptamer 1.12.2, but resulted in a 4x decrease in K_d value for aptamer A08. In the second trial, the K_d calculation for aptamer D07 was found to be negative. While the negative value was likely caused by bubble formation, it emphasizes the difficulty and inconsistency of this method. Additionally, the effect of buffer and calcium concentration could not be determined from this screen and the calculated standard deviations were in some cases larger than the average K_d values. This method should therefore not be employed for aptamer binding experiments. However, despite the inconsistent results, the overall trend indicated that newly selected aptamer A08 performed well in comparison to the 1.12.2 aptamer reported in 2008⁵⁰ and was therefore chosen for further experimentation.

Affinity chromatography

Using the remaining OTA-column material from the previous OTA affinity chromatography experiments, an affinity chromatography binding assay was performed to elucidate the K_d of the winning aptamer A08. Again, binding saturation was not reached and more importantly the estimated K_d was in the mM range. To ensure that the washing process was not excessive, as was potentially the problem in section 2.3.1, only 3 washes instead of 5 were performed to remove non-binding aptamers. Despite this change, the lack of binding indicates that it is more likely that the conjugation reaction of

OTA to the DADPA resin was unsuccessful. The binding isotherm of the affinity experiment can be found in Appendix B.

3.3.3 AuNP dissociation constant (K_d) experiment

The synthesized nanoparticles were a red-wine colour. UV-Visible spectrometry was used to confirm plasmon absorbance at $\lambda_{\text{max}} = 525\text{nm}$. Given that the $\epsilon_{520} = 2.7 \times 10^8 \text{ M}^{-1}\text{cm}^{-1}$, this extinction coefficient was sufficient for calculating the concentration of AuNP at this λ_{max} ²⁰². The concentration was measured to be 10.5 nM, comparable to the 11 nM used in the method by Zheng et al., and was therefore sufficient for this study. Should the AuNP size be critical for future assays, methods such as dynamic light scattering (DLS), atomic force microscopy (AFM) or scanning electron microscopy (SEM) can be used to determine the size and uniformity of the colloid. However, for these initial studies, characterization with UV-Visible spectrometry was satisfactory²⁰⁷.

The qualitative results of the binding assay with aptamer A08 can be found in Figure 3-4. For both aptamer A08 and 1.12.2, an increasing concentration of OTA resulted in aggregation of the AuNPs and thus a change in colour from red to purple and then to blue. The controls showed that high concentrations (1×10^{-4}) of OTA alone did not induce AuNP aggregation; therefore, AuNP aggregation can only be attributed to the aptamer preferentially binding to the target rather than the AuNP. While this assay provided a quick means of qualitatively assessing aptamer affinity, UV-Visible spectrometry was used to quantify the observed aggregation. Results of the normalized spectra are shown in Figure 3-5. With increased concentrations of OTA, the λ_{max} of the AuNPs shifted from 525 nm (red) to 705 nm (blue). The absorbance at 705 nm was

Figure 3-4: OTA binding assay of aptamer A08 using AuNPs. Increasing concentrations of OTA resulted in aggregation of the AuNP, and therefore a change in colour from red to blue.

Figure 3-5: UV-Visible absorbance spectra of the OTA binding assay of aptamer A08 using AuNPs. Increasing concentrations of OTA resulted in a decrease in the absorbance at 525nm, and an increase in the absorbance at 705nm. Absorbance at 300 – 350 nm results from high concentrations of OTA.

was plotted against the OTA concentration to fit binding isotherms and estimate the K_d values (binding isotherms can be found in Appendix B). Only one trial for each aptamer was performed due to the high concentrations of OTA required for the study. Therefore, K_d error was calculated by determining the error on the slope of the line comparing experimental and theoretical absorbance values. Aptamer A08 was found to have an apparent K_d of $1.4 \pm 0.4 \mu\text{M}$, while aptamer 1.12.2 was estimated to have an apparent K_d of $6 \pm 3 \mu\text{M}$.

This method for determining K_d values was of particular value as it provided evidence that aptamer A08 could be directly incorporated into a AuNP-based OTA sensor. Additionally, the apparent K_d values obtained in this way measure the aptamer binding to OTA in solution. Measuring binding affinities in solution with small molecules often poses a significant problem. As a result, most aptamers that bind to small molecules are verified for binding properties using affinity chromatography on the modified column matrix used in selection²⁰⁸. However, it is often the case that the aptamers display significant affinity to the column matrix, resulting in a K_d value that does not reflect direct binding to the target⁷.

Interestingly, the K_d value obtained here for aptamer 1.12.2 is consistent with the values obtained using the full-curve equilibrium dialysis ($5.3 \pm 0.7 \mu\text{M}$) and fluorescence polarization ($6.1 \pm 0.2 \mu\text{M}$) in section 2.3.1. This validates both this method for K_d determination and provides further evidence that the reported binding affinity of aptamer 1.12.2 could not be reproduced in our hands⁵⁰. Further, the approximate 4 fold improved binding observed with A08 over aptamer 1.12.2 using the AuNP assay is consistent with the approximate 6 fold improvement of binding observed using the equilibrium dialysis

screen. By manual inspection of the data from the biosensor experiment performed with aptamer 1.12.2 by Yang et al., an apparent K_d of 0.3 to 0.6 μM for aptamer 1.12.2 was observed¹⁶². While these values are more closely related to those originally reported in 2008⁵⁰, the higher affinity is likely attributed to the use of 1 mM Mg^{2+} . As described earlier, the use of these divalent cations drastically improved binding affinity. With this concentration, again, aptamer 1.12.2 did not perform as well as reported in the AuNP assay.

3.4 Conclusion

A new selection, using slightly more stringent conditions, was performed for OTA with the hope of discovering improved binding motifs for this target. Several unique sequences were elucidated from this selection. While no consensus sequences emerged, it is clear that OTA shows an affinity towards G-rich motifs. From the screening experiments, A08 emerged as the winning sequence. Further characterization revealed that this sequence binds to OTA in solution with an apparent K_d of 1.4 ± 0.4 μM , showing a 4 to 6 times improvement in binding over the original OTA aptamer 1.12.2. Future work will investigate the selectivity of this aptamer for similar targets such as ochratoxin B and other components frequently found in food-testing systems. Additionally, as a means of reducing the cost of aptamer synthesis, we will test the binding affinity of the aptamer with the original primer sequences removed. Further minimization may be performed by partially fragmenting the full length sequence and then testing the fragments that form well-defined secondary structures as predicted by Mfold.

Chapter 4. Development of a high affinity aptamer for fumonisin B₁

McKeague, M.; Bradley, C. R.; De Girolamo, A.; Visconti, A.; Miller, J. D.; DeRosa, M.

C. Screening and initial binding assessment of fumonisin B₁ aptamers. *Int. J. Mol. Sci.*

2010, *11*, 4864-4881.

4.0 Statement of contributions

Selection of the putative aptamer sequences was performed by M. McKeague, with the assistance of Charlotte R. Bradley.

4.1 Introduction

4.1.1 Fumonisin B₁

The class of mycotoxins known as fumonisins is produced mainly by the fungi *Fusarium verticillioides* and *Fusarium proliferatum*²⁰⁹. In 1970, *F. verticillioides* was isolated from infected corn that was thought to cause an outbreak of leukoencephalomalacia in horses in South Africa. As a result, the fungus was studied extensively and led to the isolation of fumonisins in 1988²¹⁰. These molecules consist of a hydroxylated hydrocarbon chain to which tricarballic acid, methyl and other groups are bonded. There are currently 15 different reported fumonisins grouped into four main categories (A, B, C, and P). In particular members of group B contain a free amino group, group A contain amides and group C lack the C-1 terminal methyl group^{210, 211}. Among the many different fumonisins identified, fumonisin B₁ (FB₁), also known as macrofusine, is of particular concern as it is the most abundant and toxic (Figure 4-1)²¹².

Figure 4-1: Structure of fumonisin B₁.

Fusarium is found as a natural contaminant in corn (maize) and corn-based foods worldwide. However, they have also been found in several other foodstuffs including rice^{213, 214}. Insect damage and some other environmental conditions, such as high temperatures and drought, result in favorable conditions for fungal infection and therefore the accumulation of fumonisins^{209, 215}. FB₁ comprises approximately 70% to 80% of the total fumonisins found in maize grain samples and has been detected in corn in all parts of the world. The concentrations found are highly variable, but levels as high as 300 mg/kg have been observed in feed in Hungary and United States. The mean daily intake of FB₁ is estimated at 12 to 140 µg per person, depending on geographical factors²¹⁰.

At high concentrations, FB₁ causes a variety of acute toxicological effects in domestic and laboratory animals, depending on the species. For example, it causes leukoencephalomalacia in horses, pulmonary edema in pigs and nephrotoxicity in rats, rabbits and lambs. It also has been associated with a reduction in body weight in chickens and causes hepatotoxicity in all species tested²¹⁶. It has been shown to be a powerful cancer promoter and a mild complete carcinogen in rodents²¹⁷. Esophageal cancer has also been associated with FB₁ in countries where consumption is high, such as China and South Africa²¹¹. As a result, FB₁ has been declared by the IARC to be a group 2B carcinogen²¹³.

Though FB₁ toxicity results in wide variety of clinical symptoms dependent on species and exposure, its toxicity stems from its clear structural similarity to the long-chain backbones of sphingolipids bases, co-substrates of the enzyme ceramide synthase. Due to the structural similarity with sphingoids bases, FB₁ inhibits ceramide synthase, disrupting sphingolipid biosynthesis. This results in an accumulation of bioactive

intermediates of sphingolipid metabolism such as sphinganine and sphingoid bases as well as depletion of complex sphingolipids²¹⁸. Sphingolipids have important roles in membrane structures, inter-cell communication, cell growth, differentiation, cellular death and immune response²¹⁹. Therefore, FB₁ causes inhibition of cell-growth as well as interference with apoptosis. Additionally, interference with the function of membrane proteins, including the folate-binding protein, resulting in inhibition of the uptake of folate, can occur²¹¹.

The Joint FAO/WHO Expert Committee on Food Additives and Contaminants established a provisional maximum tolerable daily intake of 2 µg/kg of body weight per day for total FB₁, FB₂ and FB₃^{184, 212}. Due to the concern of potential human exposure to fumonisins, the European Commission established regulatory limits for the total level of FB₁ and FB₂ in corn and corn-based foods. Depending on intended use and type of finished food product, the limits range from 200 to 4000 mg/kg. In the United States, recommended maximum levels for total fumonisins (FB₁, FB₂ and FB₃) exist for human food and animal feed. These levels ranging from 2 to 4 ppm for fumonisins in human food and from 5 to 10 ppm for fumonisins in animal feed²²⁰. Regulatory levels have not been established in Canada.

4.1.2 Current detection methods

Given its toxicity and prevalence, several analytical methods have been reported for the detection of FB₁. For regulatory purposes, the fully validated AOAC Official Method for FB₁, FB₂ and FB₃ in corn uses a strong anion-exchange column cleanup, pre-column derivatization with *ortho*-phthaldialdehyde followed by HPLC analysis with

fluorometric detection^{220, 221}. However, several other detection methods exist including gas-chromatography mass spectrometry (GC/MS) analysis and capillary zone electrophoresis²¹¹. These methods are sensitive and selective but require expensive equipment²¹¹. Immunoaffinity-based methods, such as fluorescence polarization assays, have also been developed¹²⁰. These techniques are often tedious and require time-consuming cleanup and pre-concentration steps following extraction. A new liquid chromatography, LC/MS/MS, approach has been developed for the fast sample screening of up to 87 mycotoxins simultaneously²²². However, more rapid approaches for FB₁ detection are increasingly being requested by the food and feed industry. Immunochromatographic lateral flow tests allow rapid on-site detection of mycotoxins within a few minutes. The strip tests are easy to perform, require only a sample extraction step before use and provide results within a few minutes. These tests are gaining acceptance and are being integrated into routine quality-monitoring procedures. While these lateral flow assays are easy to perform and do not require extensive equipment, they can only be used for qualitative screening of FB₁. Typically, these assays result in higher estimates of FB₁ concentrations than those determined using GC/MS and HPLC methods. The over-estimation is thought to occur due to the cross reactivity of the antibodies used in these assays²¹³.

4.1.3 FluMag SELEX

FluMag SELEX⁴⁹ utilizes fluorescent labels for DNA quantification and magnetic beads for target immobilization. These modifications have proven useful for aptamer

selection of small molecules, particularly those that are difficult to isolate and are thus valuable and expensive.

During the SELEX procedure, separation of the binding DNA from the sequences lacking affinity is required. Typically, this is achieved by immobilizing the target of interest to a column matrix, usually agarose or sepharose. The sequences displaying affinity for the target will bind to the immobilized target, allowing easy partitioning of unwanted sequences through multiple washes. FluMag SELEX makes use of magnetic beads as the solid matrix⁴⁹. These beads can be surface-functionalized with almost any chemical group to allow for easy covalent attachment of a target. The uniformity of the beads provides optimal and reproducible binding of target molecules and prevents fusion between beads⁹⁰. Additionally, only a relatively small amount of target (1 mg) is required for efficient conjugation to the beads compared to conventional column materials (10 mg or more). Use of the magnetic beads also provides a means for facile separation of the bound sequences from the non-binding sequences. In the presence of a magnetic field, the bound target is attracted to the side of the tube and can be isolated. This allows easy removal of unwanted material left in the supernatant. Therefore, the DNA library can be introduced into a solution of the magnetic beads with bound target. Sequences displaying affinity for the target will bind. When brought near the supermagnet, the target-immobilized beads along with any bound DNA sequences will be brought towards the magnet, allowing easy separation of unbound sequences in solution (see Figure 4-2). This method is much more rapid than typical agarose column separation by gravity and milder than performing separation using a high centrifugal force.

Figure 4-2: The partitioning of the DNA library using target-immobilized magnetic beads.

Following library partitioning, radioactive labels are usually used to quantify the binding and non-specific fractions of oligonucleotides. Radioactivity assays are sensitive, enabling detection of the low concentrations of DNA isolated in a round of SELEX. However, this technique is expensive and environmentally incompatible and requires that the entirety of the SELEX process be managed within an isotope laboratory. To avoid these complications, the FluMag SELEX process makes use of fluorescent labels⁴⁹. Fluorescein can be purchased as 5'-fluorescein phosphoramidite and be directly incorporated into the synthesis of the forward primer. In this way, the 5' end of the

putative aptamer sequences become labeled with fluorescein during PCR amplification. The reverse primer is modified with a hexaethylene glycol spacer followed by an additional 20 deoxyadenine residues to allow separation of the PCR double stranded product. As a result, DNA concentrations can then be easily quantified using fluorescence detection. In Chapter 3, isolation of the aptamer sequence was performed using a biotin-labeled reverse primer in PCR amplification. Following PCR, separation of the dsDNA PCR was achieved by removing the complement strand with streptavidin-coated beads. However, research shows that this method is ineffective and that the undesired complement strand often contaminates the eluted library of putative aptamer sequences²²³. The FluMag method alleviates the possibility of isolating complementary aptamer sequences by separation of the dsDNA PCR product using PAGE and further allows the fluorescent visualization of the single-stranded product of interest (see Figure 4-3).

4.1.4 Chapter objectives

This chapter describes the selection of an aptamer for the mycotoxin FB₁. An aptamer for FB₁ could be the needed molecular recognition agent for the development of inexpensive, robust and innovative biosensor platforms that could offer solutions to the limitations of the existing detection technology. In turn, this testing process may be simple enough to be performed by crop handlers and agricultural inspectors in the field, ultimately reducing its exposure and millions of dollars in waste. The binding characterization of the selected aptamers is also presented.

Figure 4-3: PCR performed using FluMag SELEX modified primers. The forward primer is modified with 5' fluorescein (yellow). The reverse primer is modified with a spacer followed by dA₂₀. Insert: The end products of PCR product separated via PAGE. Left: DNA absorbance of each product. Right: Fluorescence emission of the fluorescein-modified putative aptamers.

4.2 Experimental

4.2.1 Materials

Phosphoramidites, modifiers, activator, deblock, capping, and oxidizing reagents were obtained from Glen Research. Standard support columns and acetonitrile were purchased from BioAutomation. Ultra High Purity 5.0 argon was purchased from Praxair. ProMag 3 Series Carboxylic Acid Surfactant Free magnetic beads were

purchased from Bangs Laboratories Inc (Fishers, IN, USA). Fumonisin B₁ was a gift from Prof. Miller (Carleton University). All PCR and electrophoresis components were purchased from BioShop Canada. 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (EDC), chemicals for buffering agents and all other purposes were purchased from Sigma-Aldrich. A StrataClone PCR Cloning Kit was received from Agilent Technologies. A TempliPhi Amplification kit was purchased from GE Healthcare. Amicon-Ultra 0.5 mL 3kDa centrifuge units were purchased from Fisher Scientific Canada.

4.2.2 DNA library and primer synthesis

The starting DNA library was prepared by using standard phosphoramidite chemistry on a BioAutomation Mermade 6 (Plano, Texas, USA) as described in section 2.2.2 using 200 nmol controlled pore glass (CPG) columns with the larger, 1000Å pore size. The sequence is 96 nucleotides in length, containing a central region of 60 random nucleotides flanked by two primer binding sites necessary for PCR and cloning: 5'-ATACCAGCTTATTCAATT-N₆₀-AGATAGTAAGTGCAATCT-3'. The following primers used for amplification and cloning of the selected oligonucleotides were synthesized: Primer1: 5'- ATACCAGCTTATTCAATT-3' and Primer2: 5'-AGATTGCACTTACTATCT-3'. Modified primers were also synthesized to allow for fluorescent labeling of the binding DNA and separation of strands on a denaturing polyacrylamide gel: ModPrimer1: 5'-6FAM-ATACCAGCTTATTCAATT-3' and ModPrimer2 5'-poly-dA₂₀-HEG-AGATTGCACTTACTATCT -3'. HEG represents the hexaethylene glycol Spacer Phosphoramidate 18 (Glen Research). Library DNA and

modified primers were purified using denaturing polyacrylamide gel electrophoresis (12%) followed by clean-up with Amicon YM-3 Centrifugal Filter Devices as described in section 2.2.2. Unmodified primers were purified using reverse phase-HPLC (RP-HPLC) using an Agilent 1200 Series HPLC unit (Agilent Technologies, USA). A Zorbax 300-Ext C 18, 5 μ M, 9.4 x 250 mm column was used for DNA collection. Samples were run using a gradient elution starting at 95% 50 mM ammonium acetate and 5% acetonitrile. Over a 30 minute period, the gradient was changed to 60% 50 mM ammonium acetate and 40% acetonitrile. Purified sample was identified by UV-Visible spectrometry monitoring at 260 nm and 330nm. The purified sample was collected, lyophilized, resuspended in deionized water and quantified using UV-Visible spectrometry.

4.2.3 Derivatization of magnetic beads

ProMag 3 Series Carboxylic acid Surfactant Free Bangs magnetic beads were covalently modified through amino groups of various target compounds according to the instructions from the beads' manufacturer. Aliquots of 500 μ L (approximately 5×10^8 beads) were washed 5 times with Coupling Buffer (0.1 M K_2HPO_4 , 0.15 M NaCl, pH 5.5) and magnetically separated from the supernatant using a Dynal MPC-S, 6 x 1.5 mL tube supermagnet (Invitrogen). To test the coupling efficiency, 500 μ L of a 5×10^{-5} M solution of FB₁ was reacted at room temperature with the beads both in the presence and absence of excess 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) hydrochloride (~5 mM) for 60 minutes, followed by 5 washes with 10 mM Tris-HCl, pH 7.4 to remove excess FB₁ and EDC. Coupling efficiency was measured by determining the concentration of

unbound FB₁ remaining in the supernatant following the coupling reaction. The concentration of FB₁ was determined using HPLC with fluorescence detection (FLD) as described by Visconti et al²²⁴. 20 µL of a sample of FB₁ was derivatized with 50 µL of OPA solution (20 mg *o*-phthaldialdehyde (OPA), 2.5 mL 0.1 M NaB₄O₇, 0.5 mL methanol and 20 µL 2-mercaptoethanol) and mixed for 30 seconds at room temperature three minutes prior to injection into the HPLC. Separation of OPA from OPA derivatized FB₁ was achieved with a C-18 Reverse Phase HPLC column (5 µM, 250 x 4.6 mm) using a 77:23 v/v methanol: 0.1 M NaH₂PO₄ (pH 3.35) solvent system. Quantification was done using fluorescence detection, excitation at 336 nm and emission at 440nm. Bound FB₁ was calculated as the difference between the initial and unbound concentrations. Separate aliquots of beads were derivatized with one of L-homocysteine, L-cysteine, L-methionine or L-glutamic acid for counter-selection steps. Coupling efficiency for these targets was not verified.

4.2.4 Fumonisin B₁ SELEX experiments

Fresh aliquots of FB₁-derivatized beads and unmodified beads or beads modified with a counter-selection target were washed 5 times with 0.5 mL of Column Buffer (CB), 100 mM NaCl, 20 mM Tris, 2 mM MgCl₂, 5 mM KCl, 1 mM CaCl₂, pH 7.6, before each SELEX round. Prior to the first selection, the library (2 nmol) was suspended in 0.5 mL of CB and heated for 10 minutes at 90°C then cooled at 4°C for 15 minutes, followed by incubation at room temperature for at least 7 minutes. In subsequent rounds, 100pmol of library was used.

The library was first incubated with either unmodified beads or beads derivatized with counter-selection target molecules with mild shaking for 60 minutes. Only DNA lacking affinity for the counter-selection targets was collected by removal of the bead supernatant. This library, presumably devoid of sequences displaying non-specific affinity, was subjected to the same heating and cooling process described above, followed by incubation with a varying amount of FB₁-derivatized beads for a specific amount of time, depending on the stringency of binding selection desired (see Table 4-1).

Table 4-1: Summary of fumonisin B₁ selection round conditions.

Selection Round	Derivatization of beads used for counter-selection	Number of FB₁ - derivatized beads used	Incubation time with FB₁-derivatized beads (min)
1	Unmodified	5×10^8	60
2	Unmodified	5×10^8	60
3	Unmodified	5×10^8	60
4	Unmodified	5×10^8	60
5	L-homocysteine	1×10^8	60
6	L-homocysteine	1×10^8	60
7	L-homocysteine	1×10^8	45
8	L-homocysteine	1×10^8	45
9	L-homocysteine	1×10^8	45
10	L-homocysteine	5×10^7	45
11	L-cysteine and L-methionine	5×10^7	45
12	L-cysteine and L-methionine	5×10^7	30
13	L-cysteine and L-methionine	5×10^7	30
14	L-cysteine and L-methionine	5×10^7	30
15	L-cysteine and L-methionine	1×10^7	30
16	L-glutamic acid	1×10^7	30
17	L-glutamic acid	1×10^7	30
18	L-glutamic acid	1×10^7	20

Following incubation, sequences with little to no affinity for FB₁ were removed by washing 3 to 8 times with 0.1 mL of Column Buffer. DNA bound to the FB₁ was then removed with 7 M urea at 90°C. Five elution fractions of 0.1 mL each, were collected, purified using Amicon Ultra 3kDa 0.5 mL centrifugal filters and quantified using UV-visible spectroscopy and/or fluorescence. In the final two selection rounds, binding DNA was removed with five 30 minute elutions using 0.2 mM FB₁.

The entire selected oligonucleotide library was amplified in 15-30 parallel PCR reactions for the first round; in subsequent rounds, only 90% of the selected library was amplified with the remaining 10% set aside for analysis and in case of PCR failure. Each reaction consisted of 0.1 M Tris HCl pH 9, 50 mM KCl, 1% Triton X-100, 1.9 mM MgCl₂, 0.3 mM dNTP mix, 1 µM each modified primer and 5 units of Taq DNA polymerase. The DNA was initially melted for 10 minutes at 94.0°C, followed by 25 cycles of 94.0°C (1 minute), 47.0°C (1 minute) and 72.0°C (1 minute). Final extension occurred at 72.0°C for 10 minutes after the last cycle. In rounds 10-18, asymmetric PCR (1:50 ratio of primers, 35 cycles) was performed to facilitate the isolation of the aptamer strand from the complement. PCR products were dried down, heated at 55°C for 5 minutes in the presence of formamide and run on 12% denaturing PAGE to separate the double stranded product. Details of the PAGE set-up can be found in section 2.2.2. The 5' fluorescein-labeled DNA strands (the selected sequences) were identified using an Alpha Imager UV-illuminator. The corresponding DNA bands were cut from the gel and extracted using the freeze/rapid thaw method described by Chen & Ruffner (described in section 2.2.2) in 10 mM Tris HCl buffer, pH 7.4. The DNA was then purified using the

Amicon centrifugal filters, quantified and re-suspended in CB and could be used in the next selection round.

Enrichment of the DNA library was assessed throughout the rounds of selection by monitoring the percent of DNA library binding to the FB₁-derivatized beads by measuring absorbance at 260 nm using a Cary 300 BioUV-visible spectrophotometer and measuring fluorescence (excitation 490nm, emission 520nm) with a Fluorolog Fluorescence Spectrophotometer (Horiba Jobin Yvon, USA) with a SpectrAcq controller. After round 16, no additional enrichment was observed; therefore a total of eighteen selection rounds were performed.

4.2.5 Cloning and sequencing

The enriched DNA library obtained from SELEX round 18 was PCR amplified using the unmodified primers (Primer1 and Primer2) and cloned using the StrataClone PCR Cloning Kit according to the manufacturer's protocol. Briefly, the ligation reaction mixture was prepared by combining 3 μ L of StrataClone Cloning Buffer, 2 μ L of the PCR product and 1 μ L of the StrataClone Vector Mix in a microcentrifuge tube and mixed by repeated pipetting. The ligation reaction was incubated at room temperature for 5 minutes and then placed on ice. 1 μ L of this ligation reaction was added to the StrataClone Solopack competent cells that were thawed on ice. The competent cells were incubated on ice for 20 minutes, followed by a 45 second heat-shock at 42°C and a 2 minute incubation on ice. 250 μ L of LB broth at 37°C was added to the transformation reaction mixture, which was then allowed to recover for 1 hour at 37°C with gentle agitation. LB-ampicillin plates for blue-white screening were prepared by spreading 40

μL of 2% X-gal on each LB-ampicillin agar plates (1.0% NaCl, 1.0 % tryptone, 0.5 % yeast extract, 2.0% agar, and 100 μg/mL ampicillin). 20 to 100 μL of the competent cells were spread onto the plates. The colonies were left to grow overnight at 37°C. Positive colonies (white) were picked. DNA was prepared for sequencing by rolling-circle amplification with the TempliPhi Amplification kit using the protocol provided. A small portion of each colony plaque was transferred, using a pipette tip, into a 0.2 mL PCR tube containing 5 μL of TempliPhi sample buffer. The samples were denatured by heating at 95°C for 3 minutes and cooled to room temperature. The TempliPhi premix was prepared by combining 5 μL of reaction buffer and 0.2 μL of enzyme mix in a separate tube. 5 μL of this mix was added to each of the cooled, denatured samples. Amplification was allowed to proceed for 18 hours at 30°C followed by inactivating the enzyme at 65°C for 10 minutes. The samples were cooled to 4°C and sent for full service sequencing at the University of Calgary University Core DNA Services.

4.2.6 Binding assays

6 unique sequences obtained from the cloning experiment were tested for binding to FB₁. These sequences, as well as unenriched library, were synthesized labeled with 6-FAM on a 200 nmol scale using standard phosphoramidite chemistry. Sequence synthesis was verified through molecular weight verification using electrospray ionization mass spectrometry, ESI-MS (Novatia LLC).

Each sequence was heated to 90°C for 10 minutes, cooled to 4°C and incubated at room temperature for 7 minutes. Aliquots of 2x10⁸ FB₁-derivatized beads were washed with CB and incubated for 30 minutes at room temperature with 15pmol of each DNA

sequence separately. Non-binding DNA was removed with 3 washing steps using column buffer. Bound DNA was eluted first with two 0.1 mL fractions of CB at 90°C for 10 minutes each. To ensure complete removal of bound DNA, elution using two fractions of 0.1 mL 7 M urea at 90°C for 10 minutes was also performed. The percent of DNA eluted with column buffer and urea were determined using fluorescence (excitation at 490nm, emission at 518nm) and calculated using calibration curves. The same experiment was performed on unmodified beads to ensure the observed binding of sequences was specific to the target.

4.2.7 Dissociation constant (K_d) determination using affinity chromatography

The dissociation constant for each aptamer sequence was determined by performing binding assays as above. Varying concentrations of aptamer (1 nM to 1 μ M) were incubated as above with 2×10^8 FB₁-derivatized beads in column buffer. Non-binding DNA was removed with five washes. Bound DNA was eluted with two 0.1 mL fractions of column buffer at 90°C for 10 minutes each. The fluorescence of the eluted aptamers was recorded and plotted against aptamer concentration. The dissociation constants were evaluated as described earlier, by minimizing the residual values between calculated and observed experimental fluorescence data using the solver feature of Microsoft Excel. This experiment was performed three times to obtain the reported dissociation constants.

4.2.8 Confirmation of aptamer binding in solution

Ultrafiltration

To evaluate the feasibility of using ultrafiltration as a means of determining aptamer binding affinity, 200 μL of a 2 μM mixture of FB_1 was loaded into 10kDa molecular weight cut-off Amicon-Ultra centrifuge filter units. Half of this volume was allowed to pass through the membrane by centrifugation at 14,000 g for 2 minutes. The concentration of FB_1 in both the retentate and filtrate was determined using HPLC with FLD as described in section 4.2.3 following derivatization with OPA.

The stirred-cell apparatus (Model 8010, Amicon) using ultrafiltration membranes (1kDa and 5kDa molecular weight cutoff) was tested as an alternative for determining the aptamer K_d . 1 mL of 2 μM mixture of FB_1 was loaded into the chamber. The solution was allowed to pass through the membrane using UHP argon at 50psi of pressure. Again, the concentration of OTA in both the retentate and filtrate was determined using HPLC with FLD as described in section 4.2.3.

DMS-piperidine protection assay

6-FAM labeled aptamer FB_{139} was diluted to a concentration of 20 μM in CB and heated to 90°C for 10 minutes, cooled to 4°C and incubated at room temperature for 7 minutes. 50 μL of this DNA was incubated for 30 minutes with 50 μL of 1 mM FB_1 in CB, 50 μL of 1 mM glutamic acid in buffer or 50 μL of CB alone. A fresh solution of 10% v/v of dimethyl sulfate (DMS) was prepared and 1 μL was added to each DNA mixture to give a final concentration of 0.1% (v/v) DMS. This was allowed to incubate at room temperature for 10 minutes. Methylation was stopped by adding 5 μL stop solution

(1.5 M sodium acetate pH 7.0, 1 M 2-mercaptoethanol and 100 µg tRNA/ml). The reactants were ethanol precipitated with 50 µL 3 M NaCl and 1 mL anhydrous ethanol. The DNA was allowed to precipitate on dry ice for 30 minutes followed by centrifugation at -4°C for 20 minutes at 14,000 g. The supernatant was removed and the pellet was washed with 80% chilled ethanol, followed by centrifugation for an additional 10 minutes. The majority of the supernatant was removed using a pipette, the remaining few microliters were removed *in vacuo*. 100 µL of 10% (w/v) piperidine was added to each DNA sample, vortexed briefly and incubated at 90°C for 30 minutes. Following piperidine cleavage, the sample was dried, redissolved in 20 µL water and 20 µL formamide followed by heating at 55°C for 5 minutes. The entire reaction was separated on a 7 M urea, 19% polyacrylamide gel at 20 mA. After 2 hours, fluorescein fluorescence imaging was performed using the AlphaImager.

4.3 Results and Discussion

4.3.1 Fumonisin B₁ immobilization

To effectively separate binding ssDNA from non-binding DNA, immobilization of the target to a solid matrix was necessary. As magnetic separation technology is fast, simple and effective, the use of surface modified magnetic beads is a superior alternative to conventional immobilization to sepharose or agarose. In particular, substantially less target is required for efficient immobilization compared to the amounts required for column immobilization⁹⁰. This is particularly important in the case of expensive or rare targets such as FB₁. Carboxylic acid modified beads were chosen due to the simple coupling chemistry possible with the sole primary amine of FB₁. Fumonisin B₂, B₃ and

B₄ all have this amine which could allow for a similar aptamer selection protocol for these other members of the fumonisin family in the future. In addition, some evidence has shown that the primary amine is a critical part of the anti-FB₁ antibody binding epitope. Conjugation via the amine would force the aptamer to exclude the amine in its binding epitope which could allow for the development of antibody/aptamer sandwich assays²²⁵. Coupling via the carboxylic acid groups to an amine modified bead was also possible. However, this coupling strategy would require an additional reaction step to block remaining amines; otherwise DNA would non-specifically bind to these amines via electrostatic forces. Successful conjugation of the target, FB₁, was confirmed using HPLC. *O*-phthaldialdehyde (OPA) derivatized FB₁ was fluorescently detected at a retention time of approximately 6 minutes. Any remaining derivatizing agent (OPA) was detected at 4 minutes. Quantification of unbound FB₁ using a standard curve, measured as the total peak area, showed 1-5 nmol of target successfully conjugated to aliquots of 5×10^8 beads. Comparison of the immobilized FB₁ in the presence and absence of the required coupling agent (EDC) showed that a small amount of FB₁ (less than 0.1 nmol) can non-specifically interact with the carboxylic acid surface activated beads (see Figure 4-4).

4.3.2 Selection of FB₁ aptamers

Careful consideration was made when choosing the selection buffer for this experiment. Since the DNA backbone is polyanionic, the negative-negative charge repulsion may inhibit formation of complex structures in the absence of counterions such as Na⁺. This may ultimately impede binding with the target. Divalent cations in

Figure 4-4: HPLC chromatographs comparing the unbound FB₁ remaining in the supernatant following reaction with: A) no EDC present in the reaction B) EDC present as the coupling agent. Excess OPA retention time 4 minutes; OPA derivatized FB₁ retention time 6 minutes.

particular, such as Mg^{2+} , block the negative charge repulsion and allow the backbone to form tighter folding. Interestingly, aptamers with higher affinity show less magnesium dependence than weaker binding aptamers¹³¹. Therefore, performing SELEX under physiological conditions (1–2.5 mM Mg^{2+}) may increase the stringency of selection conditions, leading to the production of high-affinity binding aptamers. While the only other mycotoxin aptamer, the OTA aptamer, was selected under buffer conditions lacking calcium, an increase in binding affinity was observed upon the addition of calcium ions⁵⁰. It is suggested that the OTA forms a coordination complex with divalent ions such as calcium with the aid of the carboxyl and hydroxyl groups. Calcium ion was included in the buffer for this experiment to allow for a similar coordination complex that may form between neighboring carboxyl groups in FB₁. As mentioned, many published aptamers fold into G-quadruplex structures such as the thrombin-binding aptamer²²⁶ and the HIV-1

integrase aptamer²²⁷. Potassium was included in the selection buffer since monovalent cations, especially K^+ , are required to help stabilize G-rich sequences into these guanosine tetrads.

While aptamers have been selected using a variety of random region lengths within the initial library, a 60 nucleotide length was chosen for these experiments. A minimum of 25 nucleotides should be employed to achieve a diversity of 10^{15} different sequences. Generally, as the length of the random region increases, the structural complexity of the library increases. This increase in complexity allows for the presence of complex structures which are often high-affinity binders¹⁰⁴. For this reason, lengths over 200 nucleotides have been used. However, an increased library length is associated with a lower synthesis yield and higher cost of production. An initial library with 60 random nucleotides was chosen to strike a balance between these two competing factors.

When performing SELEX on a solid matrix for target immobilization, it is extremely important to ensure that any sequences displaying binding properties towards the matrix and/or the linker arm are eliminated. Once aptamers are generated, they are typically used to measure free target in solution, therefore any aptamer affinity derived from any partial binding to the matrix will reduce the functionality of the aptamer for sensing applications. For example, the published rhodamine aptamer displays a weaker binding to the target rhodamine when in solution compared to when it is immobilized on the matrix used in the selection⁸³. To mitigate this, it is necessary to remove all sequences that display even weak binding or partial recognition to the matrix. Typical SELEX experiments do this by performing a pre-selection step prior to the first selection round to remove any initial sequences displaying this unwanted binding⁴⁷. However, as

mutation and evolution of the library occurs with each selection and amplification step, this negative selection step was performed prior to each round to ensure as best as possible that the library was only binding to the target. To further enhance the selectivity of the aptamers, negative selection was performed using various targets beginning at round 4. These negative selection steps, often called counter-selections, remove any sequences from the library that bind to structurally similar molecules, or molecules that may be found in a sample containing the target of interest. Counter-selections using L-homocysteine (rounds 5-10), L-cysteine, L-methionine (rounds 11-15) and L-glutamic acid (rounds 16-18) immobilized to the magnetic beads were used since these amino acids could be easily coupled to the beads via the amino group. Library sequences that would bind to any exposed carboxylic acid functional group would be removed by these negative selections. A negative selection using FB₂ was not employed in the selections for several reasons. The difference between these two structures is very subtle: FB₂ lacks the 11-hydroxy group. While aptamers have been selected that can recognize as small a difference as this, the counter-selection against FB₂ would likely reduce the DNA library and eliminate many high affinity binders. The loss of high affinity binders in order to gain selectivity may not be worthwhile in this case as FB₂ is often found in samples containing FB₁. Both the USA recommended levels and the European Commission maximum tolerable intakes include the total of FB₁ and FB₂ intake; therefore, total fumonisin is often detected²²⁸. Note that the absence of counter-selections against FB₂ does not guarantee that the selected aptamers will be able to bind both targets.

To assess the enrichment of the library for target binding, as well as to determine when increased stringency could be applied to the selection, the percent binding of the

DNA library was monitored after each round using fluorescence and absorbance measurements (see Figure 4-5). Typical selection matrices include μmol amounts of target. However, due to the cost and availability of the target FB₁, as well as the possibility of using small concentrations of target with the magnetic beads, the selection matrix contained a maximum of 5 nmol of target. In the later rounds, the amount of target available for binding could be as low as 20pmol (1:5 ratio of FB₁:DNA). This accounts for the seemingly low percent recovery of binding sequences with each round. In the initial few rounds, the maximum amount of target immobilized beads (5×10^8 beads) was

Figure 4-5: Percent recovery of binding ssDNA to the FB₁ immobilized magnetic beads after each selection round. With any observable significant increase or plateau in percent recovery, increasingly stringent selection conditions were implemented the following round (see experimental and Table 4-1).

used for selections. This was to ensure complete capture of binding sequences. However, every 4-5 rounds, the number of magnetic beads used for positive selection, and therefore the amount of target available for binding, was reduced 2 to 5 fold (see Table 4-1). This decrease in the concentration of the target increased the stringency of the selection round, allowing for competition between sequences within the selection, similar to the SELEX strategy employed in Chapter 3. Competition for the fewer binding sites should ultimately yield tighter-binding aptamers. Tok et al.²⁰³ described a novel method to generate DNA aptamers by using a single target-immobilized microbead. They found that the weak binding aptamers were displaced by the higher affinity aptamers. After only 2 selection cycles, aptamers displaying dissociation constants in the low nanomolar range were obtained.

Two other conditions were manipulated to increase the stringency of selections (see Table 4-1). First, non-binding aptamers were washed away with column buffer (prior to elution) only 3 times in the initial rounds. As the number of rounds increased, this was increased to 8 washing steps to wash away not only the non-binding sequences but also any weaker-binding sequences. Second, the binding incubation time was slowly decreased as the selections progressed. In the initial rounds, a one hour period was used to allow for slow folding or refolding in the presence of the target. However, in the last rounds, 20 minute incubations were used since higher affinity binding species can be readily selected within a short equilibration time⁴⁷.

Several methods are possible for the elution of target-binding sequences. Direct cleavage of the target-bound aptamer from the bead was not possible in this case. While

elution by affinity with free ligand is typically used⁴⁷, the amount required for successful elution for each round (often up to 50 mM) of 98% purity would be prohibitively expensive. Typically, the target used for elution should be as pure as possible: this would cost closer to \$12,000 if obtained through a standard chemical supply house. For this reason, the elution methods available included changing the ionic strength or use of denaturants. It was determined that simply changing the ionic strength by addition of metal chelators, such as EDTA, or using deionized water only allowed removal of weak binders. This is not surprising as some aptamers have been shown to maintain their binding in the absence of Na^+ and K^+ cations⁵⁰. To fully remove the high affinity binders, several elutions with 7 M urea at 90°C over several minutes were required. However, to ensure the final aptamers would bind to FB_1 free in solution, elution with pure FB_1 was performed in the last 2 selection rounds. As high affinity binders in these late rounds of SELEX should have very long residence times, a few long elution times were performed to completely remove all bound species. Following elution in these last 2 rounds, extensive clean-up to remove the FB_1 from the binding sequences was required as the FB_1 inhibited PCR and quenched fluorescence measurements.

Six unique sequences were obtained after cloning from round 18 (see Table 4-2). Surprisingly, sequences having the highest incidence (FB_1 31 and FB_1 39) have the lowest G content. As discussed previously, high G content is typical of several DNA aptamers.

Table 4-2: Full length sequences obtained after 18 selection rounds for the target fumonisin B₁.

Sequence	5'- ATACCAGCTTATTCAATT-N₆₀- AGATAGTAAGTGCAATCT-3'	G Content (%)	Sequence incidence (% of Total)
FB₁ 14	CTATACGGAGTGGATATCGATCTGTAACGT GAGTGAGATAATGTGATGCATAGTCGTGG	32	12.5
FB₁ 16	CATCCAGTAACAAACACATAAGTAACGGC GATATGTCAAAGCGGTATCGGCTACAGATG	22	12.5
FB₁ 23	GCGGATGCGTAAATGACGATAAACATAGA TGGGGTATATCGCGATGCGACAGGGTGT	35	12.5
FB₁ 31	CGGGGACGTGTATACCAGCTTATTCAATTC ACAGTTATGTCCTATACCAGCTTATTCAATT	17	25
FB₁ 32	AATGTACGATGTGTGGGCAACATGAGTATG TCGTGTGATATCTAGATGAGGTAGCGGTGG	37	12.5
FB₁ 39	AATCGCATTACCTTATACCAGCTTATTCAA TTACGTCTGCACATACCAGCTTATTCAATT	8	25

Secondary structures of the sequences were predicted using Mfold²⁰⁴ (see Figure 4-6) with a folding temperature of 25°C and Na⁺ and Mg²⁺ concentrations as in the CB²⁰⁵. The resulting predicted secondary structures are all of very low complexity. All sequences consist of simple hairpin-loop motifs. This is very typical of selected aptamers¹⁰⁴. At this early stage, however, it is unclear if the large, initially unstructured regions of these sequences are involved in target binding, nor is it known whether these stretches fold in the presence of the target molecule as a result of new molecular interactions. These questions will be addressed in our future work.

Figure 4-6: Secondary structure prediction for the full length FB_1 aptamer clones. Image generated by Mfold.

4.3.3 Binding assays

Each aptamer clone was screened for its relative ability to bind to FB₁-derivatized beads using a chromatography binding assay. This allowed for comparative determination of the relative binding ability of each aptamer generated in the selection experiment. The selection conditions used in the SELEX experiment were used for these binding assays, using aliquots of FB₁-derivatized beads as the target. To remove any non-binding or weakly-bound DNA, many washes were performed. Bound DNA was eluted first with column buffer at 90°C. Elution using high temperatures is sufficient to denature most target-DNA interactions; however, the addition of the denaturant urea, in our experience, is able to displace even the most tight, non-covalent, interaction. The comparison of binding DNA eluted only with 90°C buffer compared to 90°C urea can be found in Figure 4-7. The specific binding of the aptamers to the target was compared to binding of the unenriched initial library (Control Library). Binding assays were also performed on unmodified (control) magnetic beads. This was done to ensure that the individual clones were binding specifically to FB₁ rather than interacting with the surface of the magnetic beads.

Figure 4-7: Binding assays of the individual aptamer sequences against FB₁ beads compared to the unmodified beads (labeled as control beads). The percent of DNA eluted with buffer at 90°C is in blue, the DNA eluted using 7 M urea at 90°C is in grey. The control library (unenriched library) was tested on the FB₁ modified beads as a comparison.

The control library, which did not undergo enrichment, has a small degree of non-specific affinity for the beads as evidenced by the low percentage of DNA bound. This amount is comparable to the non-specific binding seen between each of the sequences and the unmodified beads (Control Beads). For most sequences, 50% or more of the DNA non-specifically bound to the unmodified beads can be removed using the 90°C buffer, indicative that these interactions are relatively weak. The opposite is true in the case of the aptamer sequences and their affinity for FB₁ beads. Across all sequences, a 2 to 5 fold increase in affinity is seen in the presence of the target. Additionally all of the FB₁ tests required high concentrations of 90°C urea in order to remove the majority of the bound sequences. The observed percent binding of the aptamer sequences is sizeable for this type of assay. A similar binding assay performed using streptavidin conjugated magnetic beads by Stoltenburg et al. showed a maximum percent binding of 21% for their aptamer clones. Despite the lower binding percents observed, their aptamers were found to bind to the target with dissociation constants in the low nanomolar range⁴⁹.

Interestingly, FB₁ 39 which yielded the highest binding to the target contains the lowest guanosine content (only 8%). As mentioned earlier, this is atypical of most aptamers. This sequence is slightly T-rich (35%) compared to C (27%) or A (30%). While less common, there is some precedent for T-rich aptamer sequences. For example, the DNA aptamer for tubulin is a T-rich aptamer displaying binding in the micromolar range. The random library used in their selection was also able to generate G-rich aptamers for other targets such as chitin, indicating that the T-rich aptamer property is specific for the target type and not simply a chance event due to the nature of the original library²²⁹.

As the screen revealed that all sequences showed a reasonably high affinity for the FB₁ immobilized beads, all sequences were further analyzed to determine their K_d values. The average K_d values from 3 trials, with the corresponding standard deviation values, can be found in Table 4-3. The binding isotherms can be found in Appendix B. As expected, sequence FB₁ 39 displayed the highest affinity for the target with a K_d value of 100 nM. For a small molecular target, this is considered a relatively low K_d value (high affinity). A few DNA aptamers have recently been isolated for small molecules and small ions, displaying K_d values in the low nanomolar range such as the aptamer for BPA¹⁰⁹ and the aptamer for arsenic⁹⁸. However, the majority of aptamers that bind to small molecules display affinities in the low to mid micromolar range. Our stringent selection procedure may have resulted in the isolation of this high affinity aptamer. Additionally, all other selected sequences displayed good to moderate binding to the target, with all K_d values lower than 1 μ M.

Table 4-3: Average K_d values obtained from 3 separate affinity chromatography experiments for FB₁ sequences obtained from SELEX. Reported error was found using standard deviation.

Sequence	K_d (nM)
FB ₁ 14	920 $\pm$ 40
FB ₁ 16	200 $\pm$ 100
FB ₁ 23	1000 $\pm$ 20
FB ₁ 31	470 $\pm$ 60
FB ₁ 32	400 $\pm$ 100
FB ₁ 39	100 $\pm$ 30

4.3.4 Measuring aptamer affinity in solution

As previously discussed, aptamers selected using a solid support matrix may display affinity towards this matrix and/or the linker arm. Despite the efforts to eliminate any of this non-specific interaction, it is clear in Figure 4-7, that each sequence displays some affinity (less than 10%) to the matrix. As a result, binding affinity of the best aptamer, FB₁ 39, was tested in solution.

Equilibrium filtration

The 10kDa cut-off Amicon Ultra Centrifuge devices were tested for their ability to allow FB₁ to pass through the membrane. Similar to the problem encountered with OTA, the majority of the FB₁ does not pass through the filter. While a small percentage (up to 10%) of target sticking to the filter is acceptable, and can typically be corrected for in K_d determinations, these filters trapped more than 90% of the loaded target. As an improvement had been observed in filter binding with OTA using the Amicon Stirred-cell units, these were tested with FB₁ as well. Using the 1kDa MWCO filters, 60% of the FB₁ was able to pass through the filter. Typically these filter pore sizes are $\pm 50\%$. Therefore, given that FB₁ has a molecular weight of 721.83g/mol, it is not surprising that larger quantities of FB₁ were retained by the filter compared to the tests with OTA (MW=403.81g/mol). However, using the 5kDa filters, 90% of the FB₁ was recovered following filtration. This method could have been exploited to determine the K_d value of our aptamer in solution in a similar manner to the OTA experiment performed in section 2.3.1. However, detection of FB₁ could only be performed using the OPA HPLC FLD method. Using our fluorescence detector, the lowest detection limit was 1 μ M, which

was insufficient for producing a binding isotherm. While access to a fluorescence detector with much better sensitivity may be useful for measuring the K_d value, alternate methods for determining binding in solution should be investigated. This is of particular importance since the results for the OTA aptamer 1.12.2 (in section 2.3.1) did not agree with all other reported data.

DMS/ piperidine protection assay

The binding affinity of many RNA aptamers can be determined using assays such as in-line probing. This is owing to the presence of the 2' hydroxyl of RNA that can act as a nucleophile and non-enzymatically cleave the phosphodiester backbone on the adjacent phosphorus centre. The 2' oxygen, phosphorus and adjacent 5' oxygen must be in line to perform this cleavage²³⁰. As a result, this process is dependent on the flexibility at each RNA position and therefore the overall tertiary structure of the sequence. Since aptamer conformation often changes with target binding, a change in the backbone cleavage can be detected at increasing concentrations of the target. While the lack of the 2' hydroxyl results in benefits in terms of stability for applications of DNA aptamers, it also causes more difficulties in measuring DNA aptamer target equilibria.

However, protection assays such as the DMS piperidine assay has been useful for elucidating the structure of aptamers²³¹ and DNazymes²³² and could be used to measure binding affinity. Based on Maxam-Gilbert sequencing, DMS methylates purine residues, specifically the N7 position of G and N3 position of A. This results in purine ring opening. Next, a break in the glycosidic bond between the ribose sugar and the nucleobase occurs, displacing the base and leaving the phosphodiester backbone

susceptible to cleavage by piperidine. In general, DMS preferentially methylates G residues over A. When a labeled DNA strand, cleaved by DMS/piperidine, is separated on a denaturing gel, an intense band is visualized at every single G residue and a weak band can be seen at each A residue. This phenomenon provided the basis for one of the first DNA sequencing techniques in the 1970s²³³. However, the same principle can be used for measuring the binding affinity of an aptamer. If the target interacts with a G or A residue on an aptamer strand, such that methylation is either more or less favoured, a change in the banding pattern can be seen compared to the aptamer alone.

The results of the DMS/piperidine protection assay can be seen in Figure 4-8. Unfortunately, no consistent change in banding pattern can be observed between the aptamer incubated with target and the aptamer with controls (no target or glutamic acid). While unhelpful, this does not imply that the aptamer lacks affinity for the target FB₁. In fact the results are not surprising. Most researchers look at the guanine residues (as they are the more intense bands), however, FB₁ 39 has a very low incidence of G residues (only 8%). Therefore it is possible that the aptamer does not interact with these few bases.

4.4 Conclusions

Eighteen rounds of increasingly stringent selection conditions were performed in order to obtain DNA sequences displaying high binding affinity to FB₁. Of the six individual sequences obtained from the SELEX experiment, several candidates displayed superior binding to the target compared to controls. Aptamer clone FB₁ 39, displaying low complexity and an unexpectedly low percent of G, displayed the highest affinity for

Figure 4-8: 20 % denaturing PAGE separation of products from the DMS/piperidine protection assay. All lanes contained 1 nmol of 6-FAM labeled FB₁ 39 aptamer. Lane 1 was mixed with buffer, lanes 2, 4-6 were incubated with 1 mM FB₁ and lane 3 was mixed with 1 mM glutamic acid.

the target. Continued efforts will be made to confirm affinity of this aptamer in solution.

Due to the difficulty with most methods, the nanoparticle binding assay described in

Chapter 3 will be used. Further analysis will attempt to identify the minimal target-

binding sequence in order to shorten the aptamer and improve binding while lowering its

production cost. Dissociation constants for the final minimal aptamer will be determined

in solution and cross-reactivity tests will be run against other mycotoxins. Should this

aptamer not exhibit similar binding with other members of the fumonisin family, some of the partially enriched library from an earlier SELEX round will be used to select aptamers for those fumonisins. Stability of the aptamer under typical fumonisin extraction conditions and the feasibility of incorporating the aptamer into solid-phase extraction columns and biosensor systems will also be examined.

Chapter 5. Development of homocysteine aptamers

5.0 Statement of contributions

The selection experiments and aptamer characterization was performed by M. McKeague. The L-homocysteine aptasensor experiments were performed by Amanda Giamberardino.

5.1 Introduction

5.1.1 Homocysteine

Homocysteine (Hcy), 2-amino-4-sulfanylbutoic acid (see Figure 5-1), is a non-proteogenic amino acid that was first isolated by Butz and du Vigneaud in 1932²³⁴. This amino acid serves as an intermediate in methionine metabolism and is thus produced at the crossroads of two major biochemical pathways, where it is irreversibly transformed via the transsulfuration pathway into cysteine or is remethylated back to methionine²³⁵ (see Figure 5-2).

Figure 5-1: Structure of L-homocysteine

The primary source of homocysteine is from the S-adenosylmethionine (SAM)-dependent transmethylation pathway. Here, methionine (Met) is converted to SAM by the ATP-dependent action of methionine adenosyltransferase. SAM is demethylated to S-adenosylhomocysteine (SAH) by methyltransferase (MTs), which is then hydrolyzed by SAH hydrolase to Hcy and adenosine²³⁶.

Figure 5-2: Homocysteine biochemical pathways: The transmethylation pathway and transulfuration pathway of homocysteine is flanked by the folate dependent pathway on the right and the folate independent pathway on the left. S-adenosylmethionine (SAM); methyltransferase (MTs); S-adenosylhomocysteine (SAH); cystathionine β synthase (CBS); γ -cystathionase (CGL); betaine homocysteine methyl transferase (BHMT); dimethylglycine (DMG); 5-methyltetrahydrofolate (5-CH₃-THF); methionine synthase (MS); tetrahydrofolate (THF).

At this point in the pathway, homocysteine can be processed in two different ways. It can be remethylated when methionine levels are low or it can undergo

transulfuration. There are two known remethylation pathways for methionine salvage: the folate dependent pathway and the folate independent pathway. In the folate dependent pathway, the vitamin B₁₂²³⁷ dependent enzyme, methionine synthase (MS), catalyzes remethylation using 5-methyltetrahydrofolate (5-methylTHF) as a substrate²³⁸. The folate independent pathway makes use of the oxidation product of choline, betaine, as the methyl donor. With this substrate, betaine-homocysteine S-methyltransferase (BHMT) catalyzes the production of methionine. BHMT and MS are considered to contribute equally in the process of homocysteine remethylation in the liver²³⁹.

Transulfuration is the primary means of homocysteine catabolism. The enzyme cystathionine β synthase (CBS) along with pyridoxal phosphate (the active form of vitamin B₆) condenses homocysteine and serine to produce cystathionine. Cystathionine is then cleaved into cysteine and α -oxobutyrate by the vitamin B₆ dependent γ -cystathionase (CGL). In addition to protein synthesis, cysteine is used for glutathione synthesis and other biologically important molecules²³⁹.

Reduced expression of any protein along the remethylation or transulfuration pathway will result in a disruption in homocysteine metabolism, leading to hyperhomocysteinemia. The proposed upper reference limit for plasma homocysteine levels in adults leading a healthy lifestyle or taking adequate vitamins is 12 μ M and 15-20 μ M for those without proper vitamin supplements or a healthy lifestyle²⁴⁰. Cases of elevated homocysteine levels are considered mild when the plasma levels are between 15–30 μ M. Typically, slight increases in levels are due to moderate renal disease, drug use (such as antiepileptic drugs, methotrexate, theophylline, niacin, immunosuppressive drugs), hypothyroidism, certain cancers, psoriasis, mild folate or vitamin B₁₂ deficiency,

increasing age, high protein intake, low intake of vegetables or fruits and sickle-cell anemia. Homocysteine plasma levels are considered moderate between 30–100 μM and are typically a result of end-stage renal disease, moderate vitamin B₁₂ deficiency and severe folate deficiency. Finally, homocysteine plasma levels more than 100 μM are considered severe and are a result of severe vitamin B₁₂ deficiency and almost always cystathionine beta-synthase deficiency²⁴¹.

It was only in 1964 that Gibson et al. reported that patients with homocystinuria had vascular anomalies and arterial thrombosis. Then in 1969, McCully proposed the “homocysteine hypothesis” of vascular disease, making the connection between elevated homocysteine levels and the risk of atherosclerosis²⁴². Several possible mechanisms for the association between homocysteine and atherosclerosis have been demonstrated in experimental models including stimulation of smooth muscle growth, reduction in endothelial cell growth, weakened endothelial cell relaxation and reduced synthesis of high-density lipoprotein²⁴¹.

In addition to cardiovascular disease, increased homocysteine levels have been shown to be associated with renal failure, Alzheimer’s disease, dementia, psychiatric disorders, pregnancy complications and birth defects (such as congenital heart disease, cleft lip and palate), late pregnancy complications, osteoporosis and cancer²³⁵. In particular, homocystinuria is a rare autosomal recessive condition where homocysteine plasma concentrations are considered severe ($> 100 \mu\text{M}$). Homocystinuria results in a high risk of vascular disease early-on if left untreated. Individuals that are homozygous for homocystinuria also experience ocular, skeletal, and neurologic complications²⁴².

In patients with homocystinuria, it was discovered that treatment with folic acid lowered homocysteine. As a result, researchers hypothesized that treatment with folic acid, vitamin B₆ and vitamin B₁₂ would decrease the risk of cardiovascular disease²⁴¹. However, it is unclear whether such supplementation actually prevents cardiovascular events. As a result, there is no consensus about whose homocysteine levels should be measured and if homocysteine-lowering therapies should be used. Therefore, the United States Preventive Services Task Force currently does not recommend measuring the plasma homocysteine levels in the evaluation of either low risk or intermediate-risk populations. Regardless, Humphrey et al. performed a systematic review and meta-analysis and found that for every 5 µM increase in serum homocysteine levels, the risk of coronary heart disease (CHD) events increased by approximately 20 %, independent of traditional CHD risk factors²⁴³. While experts suggest that homocysteine reduction therapies do not lower the risk of cardiovascular disease, most of the clinical trials have not specifically involved patients with elevated homocysteine levels and have instead included all patients regardless of homocysteine level. Future studies must clarify whether patients with elevated homocysteine would benefit from such treatment and at which level of homocysteine would it be useful to start treatment²⁴¹.

5.1.2 Current detection technology for homocysteine

As previously noted, a multitude of factors have been implicated in disturbing homocysteine metabolism including nutrient deficiency, use of pharmacological agents, genetic defects and certain disease states. While hyperhomocysteinemia is a hallmark of the aforementioned conditions, the impact on other related metabolites is condition

specific. As such, analytical techniques for homocysteine detection and related metabolites in biological samples are important to gain specific knowledge about the metabolic disturbance²⁴⁴.

In 2003, Nekrassova et al. reviewed 24 detection methods for homocysteine. The comparison included 15 high performance liquid chromatography (HPLC) separation-based methods. The remaining 9 methods included gas chromatography/ mass spectrometry (GC/MS), capillary electrophoresis (CE), electrochemical detection and capillary zone electrophoresis (CZE). The majority of the HPLC separation techniques were coupled with fluorescence detection based on various derivatization methods and had a limit of detection (LOD) ranging from 2.5 nM to 120 nM. The remaining separation techniques involved UV detection or electrochemical detection and had LOD in the low micromolar range. The LOD for the GC/MS methods ranged from 0.17 – 5 μ M. The three capillary electrophoresis methods used were coupled to various detection methods, namely UV, electrochemical and laser induced fluorescence detection. These three methods had a LOD range of 0.5 – 3.0 μ M²⁴⁵. These methods are sensitive and allow analysis of a wide range of analytes; therefore, simultaneous detection of cysteine and glutathione is possible alongside homocysteine. However, the difficulty in resolution of compound peaks often compromises the specificity of these methods²⁴⁴. Additionally, these tedious analytical procedures require expensive reagents and instrumentation²⁴⁶. In general, immunoassays such as chemiluminescence immunoassays, fluorescence polarization immunoassays, ELISA and enzyme assays were found to be rapid and simple but are costly²⁴⁷.

Despite the many analytical methods, none have emerged that are rapid, sensitive, selective and inexpensive. Alternative approaches are therefore required and aptamers may be valuable for mitigating these detection issues. Several aptamer-based biosensors for detection of analytes in serum have been successfully developed and tested^{51, 248}. Additionally, an RNA aptamer for SAH²⁴⁹ with an affinity of approximately 0.1 μ M has already been selected. However, competition studies of this aptamer revealed that the adenine is the main contributor to binding affinity, with only a minor aptamer-target interaction at the thioether of homocysteine. As a result, new DNA aptamers selective for homocysteine should be developed for detection purposes.

5.1.3 Chapter objectives

This chapter describes the selection of novel high affinity aptamers that are selective for the target, L-homocysteine. The binding characterization and selectivity are described. One of the selected aptamers, Hcy8, is used in the development of a sensitive and selective gold-nanoparticle biosensor that may be used for the rapid detection of L-homocysteine in serum samples.

5.2 Experimental

5.2.1 Materials

Phosphoramidites, modifiers, activator, deblock, capping, and oxidizing reagents were obtained from Glen Research. Standard support columns and acetonitrile were purchased from BioAutomation. Ultra High Purity 5.0 argon was purchased from Praxair. All PCR and electrophoresis components were purchased from BioShop

Canada. L-homocysteine thiolactone hydrochloride (L-HcyT), DL-homocysteine, L-homocysteine (L-Hcy), L-cysteine, L-methionine, NHS Sepharose, 5,5'-dithiobis-(2-nitrobenzoic acid (DTNB) and chemicals for buffering agents and all other purposes were purchased from Sigma-Aldrich and prepared with Millipore Milli-Q deionized water at 18 MΩcm. Amino-agarose was purchased from MP Biomedicals (Solon, OH, USA). 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (EDC) and sulfo-*N*-hydroxysulfosuccinimide (sulfo-NHS) were received from Pierce (Rockford, Illinois, USA). Spin-X cellulose acetate microcentrifuge filter tubes, Amicon-Ultra 0.5 mL 3kDa centrifuge units and sulfo-NHS acetate were purchased from Fisher Scientific Canada. ProMag 3 Series Carboxylic Acid Surfactant Free magnetic beads were purchased from Bangs Laboratories Inc (Fishers, IN, USA).

5.2.2 Preparation of L-homocysteine

Originally, the thiolactone form of homocysteine was the only form commercially available as the pure L isomer. The acyclic homocysteine was only available as a racemic mixture. Therefore, the L-homocysteine thiolactone ring was opened with a base hydrolysis procedure adapted from Duerre and Miller²⁵⁰. 35 mg of L-HcyT was reacted with 250 μL of 5M NaOH for 15 minutes at room temperature. The reaction was then heated to 100°C for 25 minutes and 1.25 mL of 1 M NaH₂PO₄ was added to the reaction followed by storage at 4°C for 3-4 days to allow precipitation of excess NaCl. Following centrifugation at 10,000 g for 10 minutes, the supernatant was removed and filtered through 0.22 μm Spin-X filter tubes. Solvent was removed *in vacuo* and the product was stored under argon at 4°C until further use.

To verify reaction completion, absorption of L-HcyT at $\lambda=238$ nm was compared prior to and following base hydrolysis using UV-Visible spectrometryspectrophotometry.

NMR spectra

on a Bruker 300 MHz Spectrometer were run in D₂O for both L-HcyT and L-Hcy to assess the reaction success and purity. L-Hcy: ¹H NMR (300 MHz D₂O): δ 3.80 (m, 1H, CHNH₂), 2.55 (m, 2H, CH₂SH), 2.10 (m, 2H, CH₂CH₂S-) and L-HcyT: ¹H NMR (300 MHz D₂O): δ 4.25 (dd, 1H, CHNH₂), 3.45 (m, 2H, CHS), 2.78 (m, 1H, CH₂CH₂S-), 2.20 (m, 1H, CH₂CH₂S-).

5.2.3 Preparation of homocysteine modified agarose

Coupling via the primary amine

Both L-Hcy and L-HcyT were coupled to a pre-activated carboxyl-agarose NHS-Sepharose. 10 mg of L-Hcy was dissolved in 0.5 mL coupling buffer containing 0.1 M 2-(*N*-morpholino)ethanesulfonic acid (MES) and 0.5 M NaCl. Drained NHS-Sepharose (1 mL) was mixed and allowed to react for 2 hours with mild shaking. The supernatant was removed following a 5 minute centrifugation at 10,000 g. The sepharose was washed 5 times by incubation with 1 mL of 50 mM Tris, pH 10.4 for 30 minutes to quench unreacted amines. The newly modified sepharose was stored at 4°C until further use. Sepharose modified with L-cysteine and L-methionine were prepared in the same way.

Coupling via the carboxylic acid

As L-Hcy could be immobilized to agarose via either the primary amine or carboxylic acid, coupling using both chemistries were performed and compared.

Coupling of the carboxylic acid to amino-agarose was performed as described by Graberek et al. with moderate changes. 10 mg (82 μ moles) of L-Hcy was dissolved in 0.5 mL of coupling buffer, 82 μ moles of EDC and 210 μ moles of sulfo-NHS were added and allowed to react at room temperature for 15 minutes. This solution was added to 1 mL of drained amino agarose and allowed to react for 2 hours at room temperature. Five washes with 1 mL of 50 mM Tris, pH 10.4 were performed to quench unreacted amines as described above and stored at 4°C overnight.

Assessing coupling efficiency

Coupling efficiency of L-Hcy was determined using the Ellman's test. A 40 μ M DTNB solution was prepared in 10 mM phosphate buffer, pH 8.0. 15 mg of drained L-Hcy bound agarose (NHS sepharose or amino agarose) was reacted for 5 minutes with 0.5 mL of the prepared DTNB solution. The supernatant was removed by filtration through Spin-X filter tubes. Coupling efficiency was assessed by measuring the concentration of the produced thiolate ion in solution at 412nm²⁵¹.

5.2.4 DNA library and primer synthesis

The starting DNA library was prepared using standard phosphoramidite chemistry on a BioAutomation Mermade 6 (Plano, Texas) as described in section 2.2.2 using 200 nmol controlled pore glass (CPG) columns with the larger, 1000Å pore size. The sequence is 96 nucleotides in length, containing a central region of 60 random nucleotides flanked by two primer binding sites necessary for PCR and cloning: 5'-ATACCAGCTTATTCAATT-N₆₀-AGATAGTAAGTGCAATCT-3'. The following

primers used for amplification and cloning of the selected oligonucleotides were synthesized on a 1 μ mol scale: Primer1: 5'- ATACCAGCTTATTCAATT-3' and Primer2: 5'-AGATTGCACTTACTATCT-3'. Modified primers were also synthesized to allow for fluorescent labeling of the binding DNA and separation of strands on a denaturing polyacrylamide gel: ModPrimer1: 5'-6FAM-ATACCAGCTTATTCAATT-3' and ModPrimer2 5'-poly-dA₂₀-HEG-AGATTGCACTTACTATCT -3'. HEG represents the hexaethylene glycol Spacer Phosphoramidite 18 (Glen Research). DNA was purified using denaturing polyacrylamide gel electrophoresis (12%) followed by clean-up with Amicon YM-3 Centrifugal Filter Devices as described in section 2.2.2.

5.2.5 Homocysteine SELEX experiments

SELEX was performed as described by Mannironi et al.²⁵² with several changes. Fresh aliquots of L-Hcy-immobilized sepharose, unmodified sepharose or sepharose modified with a counter-selection target were pre-equilibrated by washing 10 times with 0.5 mL of homocysteine binding buffer (HBB), 50 mM Tris-HCl, 5 mM MgCl₂, 0.5 M NaCl (pH 7.4), before each SELEX round. Prior to the first selection, the library (5 nmol) was suspended in 0.5 mL of HBB and heated for 5 minutes at 65°C then cooled to room temperature. In subsequent rounds, 200pmol of library was used.

The library was first incubated with 250 μ L of either unmodified column material (rounds 1– 5) or sepharose derivatized with counter-selection target molecules (round 6 – 8) with mild shaking for 30 minutes. Only DNA lacking affinity for the counter-selection targets was collected by removal of supernatant following a 1 minute centrifugation at 5000 g. This library, presumably devoid of sequences displaying non-specific affinity,

was subjected to the same heating and cooling process described above, followed by incubation with 100 μ L of L-Hcy derivatized column for 30 minutes. Following incubation, sequences with little to no affinity for L-Hcy were removed with 10 washing steps using 200 μ L of HBB. DNA bound to L-Hcy was then removed using a variety of recovery methods depending on the round, as described in Table 5-1. Recovery of putative aptamers was performed until the fluorescence of the elution fractions was negligible. The elution fractions were collected, purified using Amicon Ultra 3kDa 0.5 mL centrifugal filters, and quantified using UV-visible spectroscopy and/or fluorescence.

Table 5-1: The various elution methods used to recover putative aptamer sequences from the L-Hcy immobilized sepharose, by selection round.

SELEX round	Recovery of putative aptamers (elution)
1	5 x 250 μ L washes with 0.1 M L-Hcy
2	10 x 250 μ L washes with 0.1 M L-Hcy
3	3 x 250 μ L washes with 10 mM L-Hcy; 5 x 250 μ L 3.5 M urea in Tris at 90°C
4	5 x 250 μ L 3.5 M urea in Tris at 90°C; 5 x 250 μ L 7 M urea at 90°C
5	5 x 250 μ L dI H ₂ O at 95°C; 5 x 250 μ L 7 M urea at 90°C
6	5 x 250 μ L 0.15 M NaOH; 5 x 250 μ L 7 M urea at 90°C
7 & 8	5 x 250 μ L 7 M urea at 90°C

The entire selected oligonucleotide library was amplified in 15-30 parallel PCR reactions for the first round, in subsequent rounds only 90% of the selected library was amplified with the remaining 10% set aside for analysis and in case of PCR failure. Each reaction consisted of 0.1 M Tris HCl pH 9, 50 mM KCl, 1% Triton X-100, 1.9 mM MgCl₂, 0.3 mM dNTP mix, 1 μ M each modified primer and 5 units of Taq DNA polymerase. The DNA was initially melted for 10 minutes at 94.0°C, followed by 25

cycles of 94.0°C (1 minute), 47.0°C (1 minute) and 72.0°C (1 minute). The final extension occurred at 72.0°C for 10 minutes after the last cycle. In rounds 10-18, asymmetric PCR (1:50 ratio of primers, 35 cycles) was performed to facilitate the isolation of the aptamer strand from the complement. PCR products were dried down, heated at 55°C for 5 minutes in the presence of formamide and run on a 12% denaturing PAGE to separate the double stranded product. Details of the PAGE set-up can be found in section 2.2.2. The 5' fluorescein-labeled DNA strands (the selected sequences) could be identified using an Alpha Imager UV-illuminator. The corresponding DNA bands were cut from the gel and extracted using the freeze/rapid thaw method described by Chen & Ruffner (described in section 2.2.2) in 10 mM Tris HCl buffer, pH 7.4. The DNA was then purified using the Amicon centrifugal filters, quantified and re-suspended in HBB and could be used in the next selection round.

Enrichment of the DNA library was assessed throughout the rounds of selection by monitoring the percent of DNA library binding to the J.-Hcy column each round by measuring DNA absorbance at 260 nm using a Cary 300 BioUV-visible spectrophotometer and measuring fluorescence (excitation 490nm, emission 520nm) with a Fluorolog Fluorescence Spectrophotometer (Horiba Jobin Yvon, USA) with a SpectrAcq controller. A total of 8 rounds were performed.

5.2.6 Cloning and sequencing

The enriched DNA library obtained from SELEX round 8 was PCR amplified using the unmodified primers (Primer1 and Primer2) and cloned using the StrataClone PCR Cloning Kit according to the manufacturer's protocol. Briefly, the ligation reaction

mixture was prepared by combining 3 μL of StrataClone Cloning Buffer, 2 μL of the PCR product and 1 μL of the StrataClone Vector Mix in a microcentrifuge tube and mixed by repeated pipetting. The ligation reaction was incubated at room temperature for 5 minutes and then placed on ice. 1 μL of this ligation reaction was added to the StrataClone Solopack competent cells that were thawed on ice. The competent cells were incubated on ice for 20 minutes, followed by a 45 second heat-shock at 42°C and a 2 minute incubation on ice. 250 μL of LB broth at 37°C was added to the transformation reaction mixture which was then allowed to recover for 1 hour at 37°C with gentle agitation. LB-ampicillin plates for blue-white screening were prepared by spreading 40 μL of 2% X-gal on each LB-ampicillin agar plates (1.0% NaCl, 1.0 % tryptone, 0.5 % yeast extract, 2.0% agar, and 100 $\mu\text{g/mL}$ ampicillin). 20 to 100 μL of the competent cells were spread onto the plates. The colonies were left to grow overnight at 37°C . Positive colonies (white) were picked. DNA was prepared for sequencing by rolling-circle amplification with the TempliPhi Amplification kit using the protocol provided. A small portion of each colony plaque was transferred, using a pipette tip, into a 0.2 mL PCR tube containing 5 μL of TempliPhi sample buffer. The samples were denatured by heating at 95°C for 3 minutes and cooled to room temperature. The TempliPhi premix was prepared by combining 5 μL of reaction buffer and 0.2 μL of enzyme mix in a separate tube. 5 μL of this mix was added to each of the cooled, denatured samples. Amplification was allowed to proceed for 18 hours at 30°C followed by inactivating the enzyme at 65°C for 10 minutes. The samples were cooled to 4°C and sent for full service sequencing at the University of Calgary University Core DNA Services.

5.2.7 Assessing aptamer binding affinity and selectivity

Aptamer synthesis

Aptamer sequences Hcy8 and Hcy10 were synthesized, unmodified, as described in section 2.2.2 (sequences are found in Table 5-2). Both putative aptamers were also synthesized with the 5' fluorescein modifier, 6-FAM. All aptamers were purified by PAGE as described in 2.2.2. A truncated version of aptamer Hcy8, Hcy8.1, was also synthesized with the primer regions removed. An unmodified control sequence, Control ssDNA, 5'-TGCTCCTACAAATGCCATCATTAGTCCGTGGTAGGGCAGGT TGGGGTGACTGCTGCAGCGAGCTTACG-3' was also synthesized. Fluorescein-modified control sequence FB₁ 32 (from section 4.3.2) was obtained from previous experiments.

Equilibrium filtration

200 μ L of a 2 μ M mixture of L-Hcy in HBB was loaded into a 3kDa molecular weight cut-off Amicon-Ultra centrifuge filter units. Half of this volume was allowed to pass through the membrane by centrifugation at 14,000 g for 10 minutes. The concentration of L-Hcy in both the retentate and filtrate was determined as described in section 6.2.3, using the Ellman's test.

For the K_d determinations, serial dilutions of Hcy8 were made ranging from 10 μ M to 10 nM in HBB. These dilutions were heated to 65°C for 5 minutes then cooled to room temperature. 100 μ L of each DNA dilution was mixed with 100 μ L of homocysteine in HBB (1 μ M) and incubated with mild shaking for 30 minutes. Following the incubation, the samples were transferred to a 3kDa molecular weight cut-

off Amicon-Ultra centrifuge filter units and centrifuged for 6 minutes at 14,000 g. The Detect X® ThiolStar fluorescence detection kit (Arbor Assays) was used to quantitatively measure thiol groups (in this case homocysteine) as described by the manufacturer. 10 µL of the filtrates and retentates from each dilution were added to 90 µL of assay buffer (two times concentrated phosphate buffer containing detergents and stabilizers) along with 25 µL of thioStar® thiol detection reagent. The samples were incubated for a period of 30 minutes and the fluorescence of the samples was measured at excitation wavelength of 390 nm and emission wavelength of 570nm.

Quartz crystal microbalance

A quartz crystal microbalance (QCM) from Stanford Research Systems (QCM200, 5 MHz AT-cut quartz crystal oscillator with 0.1 sec gate time) was employed in an attempt to measure homocysteine-aptamer interactions. The gold electrodes were cleaned by treatment with a solution of H₂O₂/NH₄OH/H₂O (1:1:5) for 10 min at 70°C, followed by several washes with deionized water. To assess the sensitivity of the QCM at detecting homocysteine, a 1 mM solution of L-Hcy in water was prepared and placed on the gold surface for 30 minutes, followed by several washes to remove unbound species. After drying under argon, the change in frequency was noted, ensuring that any capacitance was nulled by manual control of the bias voltage.

Affinity chromatography: Sepharose

L-Hcy-immobilized sepharose was prepared as described in section 6.2.3 using DL homocysteine purchased from Sigma-Aldrich. An affinity chromatography assay

was set up by preparing 14 Spin-X cellulose acetate microcentrifuge filter tubes containing L-Hcy-modified column. 100 μ L of a 1 mg/100 μ L mixture of the L-Hcy-immobilized column in HBB buffer was added to each of the filter tubes. Buffer was removed by centrifugation at 14,000 g for 1 minute. Fourteen concentrations of fluorescein-modified Hcy aptamer (Hcy8 or Hcy10) or control sequence FB₁ 32 from section 4.3.2 were prepared in HBB buffer at concentrations ranging from 1 nM to 50 μ M. Each aptamer dilution was heated to 65°C for 5 minutes, followed by cooling to room temperature. 100 μ L of each aptamer sample was then added to the filter tubes and allowed to incubate with shaking for 30 minutes. Non-binding aptamers were removed by centrifugation followed by 2 washing steps using 100 μ L of HBB. The receiving tube was changed and the bound aptamers were eluted from the column by washing twice with 90°C HBB. The DNA elution fractions were pooled and the fluorescence was measured (excitation wavelength 490nm; emission wavelength 520nm). The fluorescence was plotted versus the concentration of aptamer. As before, the K_d was determined by minimizing the residual values between calculated and observed experimental fluorescence using the solver feature of Microsoft Excel.

Affinity chromatography: Magnetic beads

ProMag 3 Series Carboxylic acid Surfactant Free Bangs magnetic beads were covalently modified through amino groups of L-homocysteine, L-methionine and L-cysteine according to the instructions from the beads manufacturer. In this case, $\geq 98\%$ L-homocysteine from Sigma-Aldrich was used for these tests. Aliquots of 500 μ L (approximately 5×10^8 beads) were washed 5 times with Coupling Buffer (0.1 M K₂HPO₄,

0.15 M NaCl, pH 5.5) and magnetically separated from the supernatant using a Dynal MPC-S, 6 x 1.5 mL tube supermagnet (Invitrogen). 500 μ L of a 5×10^{-5} M solution of target and 5 mM EDC was reacted with the beads for 60 minutes, followed by 5 washes with 10 mM Tris-HCl, pH 7.4 to remove excess target and EDC. Coupling efficiency of L-homocysteine and L-cysteine was verified using the Ellman's test as described in section 6.2.3 using 500 μ L of derivatized beads.

The dissociation constant for each fluorescein modified aptamer sequence or control (FB₁ 32 from section 4.2.3) was determined by performing binding assays as described in section 4.2.6. Each sequence was heated to 95°C for 5 minutes and cooled to room temperature. Varying concentrations of aptamer (5 nM to 10 μ M) were incubated as above with 2×10^8 L-Hcy-derivatized beads in HBB for 30 minutes with mild shaking. Non-binding DNA was removed with 2 washes of HBB buffer. Bound DNA was eluted with two 0.1 mL fractions of HBB at 90°C for 10 minutes each. The fluorescence of the eluted aptamers was recorded and plotted against aptamer concentration. The dissociation constants were evaluated as described earlier, by minimizing the residual values between calculated and observed experimental change in fluorescence data using the solver feature of Microsoft Excel. For selectivity tests, the aptamer sequences were incubated with the L-cysteine or L-methionine modified beads in the same way.

5.2.8 Preparation of a homocysteine biosensor

Nanoparticles were synthesized as in section 3.2.5 and displayed a $\lambda_{\text{max}}=520\text{nm}$. The assay was performed as described by Zheng et al. with modifications²⁰². Seven

samples of L-Hcy in water were prepared at various concentrations ranging from 200 nM to 36 μ M. 243 μ L of the L-Hcy dilutions or a blank solution of deionized water was added into microcentrifuges tubes containing 6 μ L of either unmodified aptamer Hcy8 (10 μ M in HBB), unmodified truncated aptamer Hcy8.1 (10 μ M in HBB buffer) or control sequence. After a 30 minute incubation period with shaking, 135 μ L of the AuNP solution (10.5 nM) was added, vortexed briefly and allowed to react for 5 minutes. 72 μ L of 0.25 M NaCl was then added to each microcentrifuge tube and allowed to incubate for a further 5 minutes. Each sample was then analyzed using UV-Visible spectrometry. After normalization, the relative absorption ratio between 660 nm and 524 nm $[(A_{660}/A_{524})_n - (A_{660}/A_{524})_0]$ was plotted against the corrected L-Hcy concentration. The K_d was evaluated by minimizing the residual values between calculated and observed absorption ratio using the solver feature of Microsoft Excel^{181, 182}. This experiment was replicated an additional 3 times as described. Furthermore, the experiment was repeated using unmodified aptamer Hcy8, replacing L-homocysteine with L-cysteine. Additional one-point controls were performed using 3.6 μ M L-methionine, as well as L-Hcy in diluted serum (diluted between 2 and 10 times) as described above.

5.3 Results and Discussion

5.3.1 Preparation of L-homocysteine

Given that the pure L isomer of homocysteine could only be purchased in the thiolactone form, it had to be hydrolyzed with base to L-Hcy²⁵⁰. As thiolactone absorbs in the UV region, between 235-238nm, and the open L-Hcy form displays no absorption properties, UV-Visible spectrometry was used to assess the reaction completion. Small

amounts of the reacted L-HcyT were dissolved in deionized water to a concentration of 0.05g/L. At this concentration, L-HcyT results in an absorbance greater than 1.0 between the wavelengths of 235-238nm. The complete disappearance of this peak was indicative of reaction completion (see Figure 5-3). ^1H NMR was used for

Figure 5-3: UV-Visible absorbance spectra of L-homocysteine thiolactone HCl (absorbance at 238nm) and of L-homocysteine at 0.05g/L.

comparison purposes with L-HcyT and for verification of the L-Hcy product. As shown in Figure 5-4, all the chemical shifts in the thiolactone ^1H NMR spectrum were shifted downfield relative to the thiol. This is due to the presence of the significantly more electron withdrawing thiolactone group. The protons on the chiral carbon 1 of the thiolactone produce a doublet of doublets. The protons on the carbon adjacent to the chiral center are diastereotopic; thus each will couple separately to the proton on the chiral center. Well-resolved peaks for the diastereotopic protons are not observed in the

open form of homocysteine. Although the protons adjacent to the chiral center are still diastereotopic, the effect is not as pronounced due to the free rotation about the C-C bond between the chiral center and the adjacent carbon. In the thiolactone spectrum, each proton of carbon 2 is well resolved due to the locked ring structure and results in two separate peaks at chemical shifts of 2.78ppm and 2.20ppm. The amine and hydroxyl protons in both L-Hcy and L-HcyT did not produce any peaks as these protons are readily exchanging with the deuterium from D₂O. Some residual L-HcyT peaks were observed in the L-Hcy spectrum. Integration of these peaks indicated that L-Hcy was contaminated with approximately 2% thiolactone. Due to the slightly alkaline buffer used in SELEX, it is expected that this remaining thiolactone would likely be converted to L-Hcy.

Figure 5-4: ¹H NMR (300 MHz) of L-homocysteine thiolactone HCl (bottom) and L-homocysteine produced via base hydrolysis (top).

5.3.2 Agarose column preparation

Immobilization of L-Hcy to a solid support matrix could be achieved using the primary amine, the carboxylic acid or even the thiol group. Given that the thiol group distinguishes L-Hcy from other amino acids (except cysteine), it was important that this group remained free for aptamer binding. Therefore, immobilization was tested using the amino group to a carboxylic acid modified support matrix as well as immobilization via the carboxylic acids groups to an amino modified column material. Ellman's test for the determination of thiols was used to determine the coupling efficiency for each column preparation. Ellman's reagent (DTNB)²⁵¹ absorbs light in the UV range at 325nm. When this disulfide is mixed with a free thiol group, the mixed disulfide is preferentially produced²⁵³, releasing free thiolate ion that is yellow in colour and absorbs light at 412nm. The thiolate ion is produced at a 1:1 ratio with the free thiol, therefore, by determining the concentration of thiolate ion, the amount of free thiol (in this case L-Hcy) can be indirectly quantified (see Figure 5-5).

Utilizing the amino-modified agarose would ensure that only the open form of homocysteine could be immobilized as the thiolactone does not have a free carboxylic acid group. This would eliminate the potential of immobilizing a fraction of the 2% L-HcyT that was present in the prepared L-Hcy. However, the NHS sepharose allows for a shorter functionalization time and an improved coupling efficiency since only one nucleophilic substitution step is required. As expected, the carboxyl-agarose preactivated with N-hydroxysuccinimide (NHS) had the best coupling efficiency determined as 4.3

Figure 5-5: UV-Visible absorbance spectra of Ellman's reagent (DTNB) and DTNB with L-homocysteine (Hcy). DTNB reacts with the thiol group, producing a yellow coloured thiolate ion that absorbs at 412nm. Inset: the yellow colour observed from mixing DTNB with Hcy compared to the colourless DTNB solution.

nmol of L-Hcy per mg of dry material. The amino agarose had a coupling efficiency of 1.2 nmol L-Hcy/mg column. As a result, the carboxyl-agarose was used for the L-Hcy coupling and for the SELEX experiments. This also eliminated the need to block remaining amino groups to prevent non-specific interaction of the negatively charged backbone of DNA with the positively charge amino groups.

5.3.3 Selection of homocysteine aptamers

As with each selection, careful consideration was made when choosing the selection buffer for this experiment. Since the DNA backbone is polyanionic, the

negative–negative charge repulsion may inhibit formation of complex structures in the absence of counterions such as Na^+ . Divalent cations in particular, such as Mg^{2+} , block the negative charge repulsion and allow the backbone to form tighter folding. Due to the very small molecular weight of the target (L-homocysteine = 135.18g/mol), there are less sites for DNA interaction via intermolecular forces. However, Carothers et al. found that using higher concentrations of Mg^{2+} increased the likelihood of finding sequences with binding properties¹³¹. Therefore, compared to other selections, a higher concentration of Mg^{2+} (5mM) was included, though this concentration was comparable to other small molecule aptamer selections²⁵².

Due to the success of the fumonisin B₁ selection (Chapter 4), the same library template and primers were used for the homocysteine selections. In particular, these primers had been previously optimized by Crameri et al. for SELEX experiments²⁵⁴. The design prevented the accumulation of higher molecular weight products from PCR due to the 3' ends priming onto another product. Additionally, the primers were designed to minimize the formation of smaller products, such as primer dimers, through the reduction of possible intramolecular secondary structure formation²⁵⁴. The same random region length of 60 nucleotides was chosen to strike a balance between two competing factors: shorter sequences are easier to synthesize and handle, while longer sequences allow for a greater structural complexity and possibly new functional motifs¹⁰⁴.

As described in Chapters 4 and 5, a negative selection step was employed prior to each selection round. Once aptamers are generated, they are typically used to measure free target in solution, therefore any aptamer affinity derived from any partial binding to the matrix will reduce the functionality of the aptamer for sensing applications. Given

the relatively small surface area of L-Hcy compared to the column matrix, the DNA sequences are much more likely to interact with the column matrix. Therefore, a negative selection step was performed with each SELEX round to ensure that sequences that display even weak binding or partial recognition to the matrix are removed from the selection library.

To further enhance the selectivity of the aptamers, counter-selections were performed to remove any sequences from the library that bind to structurally similar molecules. Counter-selections using L-cysteine and L-methionine were performed simultaneously in rounds 6 through 8 once significant enrichment of the library was observed. These amino acids were chosen as they are the only other amino acids with side chains containing a sulfur atom; methionine consisting of a thioether and cysteine containing a free sulfhydryl group. There are several other forms of homocysteine found in the body. For example, in its most common form, homocysteine is bound to the protein albumin by a disulfide linkage. Additionally, homocystine, which consists of two homocysteine molecules bound by a disulfide bridge and the homocysteine-cysteine disulfide (Hcy-Cys) occur. The extra counter-selection steps to gain selectivity against these additional targets were not considered to be worthwhile since total homocysteine is typically measured, which requires a general reduction step to convert all these forms to homocysteine²⁴⁰.

It is well known that several aptamers have been selected that are able to distinguish between various enantiomers of targets. For example, several examples of RNA aptamers exist that display some discrimination between various L and D amino acids^{58, 117, 118}. DNA aptamers have also been selected that are enantioselective; for

example, aptamers have been identified that differentiate (S) and (R)-ibuprofen¹⁰³.

However, in this case, counter-selections with D-homocysteine were not possible due to the unavailability of either D-homocysteine or D-homocysteine thiolactone.

To assess the enrichment of the library for target binding, as well as to determine when increased stringency could be applied to the selection, the percent binding of the DNA library was monitored after each round using fluorescence and absorbance measurements (see Figure 5-6). Compared to the fumonisin selection in Chapter 4, where only nanomoles or picomoles were available for sequence binding, the coupling efficiency of the NHS sepharose allowed the selections to be performed in the presence of micromole amounts of target. This accounts for the much higher percent recovery of binding sequences with each round in this selection compared to Chapter 4. The purpose of large concentrations of target was to ensure complete capture of binding sequences. Beginning with round 6, the counter-selections with L-cysteine and L-methionine resulted in a drastic reduction in the percentage of DNA eluted from the L-Hcy immobilized column. This was not surprising as L-cysteine in particular, is structurally very similar to L-Hcy. With the drastic reduction in library binding, it is expected that the resulting aptamers will have a significant preference in binding L-Hcy over the counter targets L-cysteine and L-methionine.

Figure 5-6: Percent recovery of binding ssDNA to the L-Hcy immobilized sepharose after each selection round. Beginning at round 6, counter-selections with L-cysteine and L-methionine were performed prior to positive selection.

Several methods are possible for the elution of target-binding sequences. Direct cleavage of the target-bound aptamer from the bead was not possible in this case. Elution by affinity with free ligand is often used⁴⁷. This method was employed in the first two rounds of selection, with 10 mM L-Hcy. While originally successful, this method required several washing steps with target to fully remove all the bound sequences. The several washes necessary may have been a result of very slow k_{off} values for some of the high affinity sequences. To mitigate this, longer incubation times of the library with the free target were attempted. However, unlike some SELEX methods that allow monitoring of the k_{off} and k_{on} rates⁷⁹, the column based selection does not provide this

information²⁵⁵. Therefore, despite several elution steps, some high affinity aptamers may not have been recovered from the column. Eluting with free L-Hcy was additionally problematic as the proton NMR spectrum indicated the presence of a 2% impurity of L-HcyT. As a result, sequences with a binding preference to the thiolactone form of homocysteine may have been selected.

Several other elution methods were employed to ensure the recovery of all high affinity binders. Other than direct cleavage and elution with the free target, use of heat, change in ionic strength or the use of denaturants can be used. It was found that simply using heat and changing the ionic strength by washing with 95°C deionized water only allowed removal of weak binders. This is not surprising as some aptamers have been shown to maintain their binding in the absence of Na⁺ cations⁵⁰. Elution with moderate concentrations of urea (3.5M) or with NaOH were also performed. In every case, further elution with 7 M urea at 90°C over several minutes was able to remove additional DNA from the column. Therefore, it was determined that elution with 7 M urea at high temperatures was required to ensure full recovery of putative aptamers. Later, it was determined that 6 M urea at high temperatures was as effective as 7 M urea and resulted in a less rigorous DNA clean-up prior to PCR.

Four unique sequences were obtained after cloning from round 8 (see Table 5-2). Secondary structures of these sequences were predicted using Mfold²⁰⁴ (see Figure 5-7) with a folding temperature of 25°C and Na⁺ and Mg²⁺ concentrations as found in the HBB (0.5 M Na⁺, 5 mM Mg²⁺)²⁰⁵. At this early stage, it is unclear if the large, initially unstructured regions of these sequences are involved in target binding, nor is it known

whether these stretches fold in the presence of the target molecule as a result of new molecular interactions. These questions will be addressed in our future work.

Table 5-2: Putative aptamer sequences for L-homocysteine obtained after 8 rounds of selection.

Sequence	5'- ATACCAGCTTATTCAATT-N₆₀ - AGATAGTAAGTGCAATCT- 3'	Sequence incidence (%)
Hcy1	ATGATCGTCCATATCCGGCACAACACCATATCTAT GTACCACACGCCAAGCATGTTGTGT	12.5
Hcy3	CCAATACTATGACGCAGCAATGAGACCACTAGAA TATCCAGGCACCCATACAACGGTTAC	12.5
Hcy8	ACCAGCACATTTCGATTATACCAGCTTATTCAATTC ACAGCTATGTCCTATACCAGCTTATTCAATT	50
Hcy10	GTGGAAAGCCGAATGTGATTAGGGACCAGTGGA GAAGTAGTACGGACTGACCTCGCGTGTA	25

5.3.4 Dissociation constant (K_d) experiments

As L-homocysteine is one of the smallest targets used in aptamer selection, determining the affinity of the selected aptamers was exceedingly difficult⁹⁰. Furthermore, homocysteine does not absorb light, therefore, the majority of the methods for measuring target and aptamer equilibria, described in Table 1-3, are ineffective for this target. However, the presence of the thiol group provided new chemistry to exploit at an attempt to measure the selected aptamer K_d values. Unfortunately, these attempts were unsuccessful for a variety of reasons described in Table 5-3.

Figure 5-7: Secondary structures of putative L-Hcy aptamer sequences generated using Mfold.

As a result, affinity chromatography was again employed to measure the affinity for the selected aptamers. Since this technique is laborious, only sequences Hcy8 and Hcy10 were tested, due to their high incidence in sequencing results. For this analysis, sepharose was prepared with the commercially available DL-homocysteine for time purposes and to ensure that no other compounds such as L-HcyT were present. The measured K_d values of sequence Hcy8 and Hcy10 compared to the control sequence are listed in Table 5-4 (binding isotherms can be found in Appendix B).

Table 5-3: Summary of methods attempted to elucidate the binding affinity of L-Hcy aptamers. A brief description explaining why these methods were ineffective is included.

Method	Reason for ineffectiveness in determining K_d
Equilibrium filtration with L-Hcy detection using DTNB	DTNB not sensitive enough: 7 μ M LOD
Equilibrium filtration with L-Hcy detection using ThiolStar Kit	ThiolStar detection kit results in background noise; not sensitive enough: 200 nM LOD
Quartz crystal microbalance	Low sensitivity and extremely time-consuming
AuNP assay	Interaction of thiol with the gold surface is non-specific (further discussion in section 6.3.6)

As previously discussed, aptamers selected using a solid support matrix may display affinity towards this matrix and/or the linker arm. Despite the efforts to eliminate any of this non-specific interaction by performing multiple negative selections, aptamers from the selection in section 4.3.3 displayed some affinity for the matrix. Therefore, an attempt to test the binding affinity of the aptamer sequences using L-Hcy derivatized magnetic beads was performed. L-Hcy was commercially available at this point and was therefore used for coupling purposes. While there is a possibility for some non-specific interaction with the beads, it is far less likely since the beads were not used in the SELEX

experiments. The measured K_d values of sequence Hcy8 and Hcy10 compared to the control sequence are listed in Table 5-4 (binding isotherms can be found in Appendix B).

Interestingly, the K_d values obtained using the magnetic beads are half the value of those obtained using the sepharose. This may be partially attributed to the fact that DL-homocysteine was used for coupling to the sepharose. If the aptamer sequences are enantioselective, only half of the total concentration of target would be available for the sepharose based affinity chromatography experiment compared to the beads based experiment. Even without performing counter-selection steps, it is possible that the selected aptamers would display such an enantioselectivity. However, when comparing the K_d values for the control sequence, the affinity decreases by half as well with the beads based method. This result could be attributed to the fact that the control sequence, FB_I 32, which resulted from the beads based selection of an aptamer for fumonisin B₁ (Chapter 4), was shown to have some affinity for the beads. Alternatively, the consistently smaller K_d values obtained with the beads based method may provide further proof (see Chapter 2) that the affinity method used to elucidate aptamer target equilibria has an effect on the measured K_d value. Regardless, these results suggest that both sequence Hcy8 and Hcy10 bind to homocysteine with high affinity. Hcy8 consistently resulted in the best K_d value, in the low micromolar range. However, both of the sequences' measured affinities are suitable for detecting homocysteine within clinically relevant ranges.

Table 5-4: Dissociation constants (K_d) for sequences Hcy8 and Hcy10 obtained using affinity chromatography on L-homocysteine immobilized sepharose or magnetic beads. (All binding isotherms can be found in Appendix B).

Sequence	Sepharose (μM)	Beads (μM)
Control	28 ± 3	14.7 ± 0.7
Hcy 8	0.8 ± 0.3	0.21 ± 0.03
Hcy 10	3.0 ± 0.3	1.8 ± 0.1

5.3.5 Selectivity

Hcy8 displayed the best measured K_d value; therefore, further testing was performed to determine its selectivity for similar amino acid targets. Using the magnetic beads, Hcy8 was found to bind to L-methionine with a K_d of $44.3 \pm 0.6 \mu\text{M}$, over 200 times weaker than to its selected target (binding isotherm can be found in Appendix B). The actual value is likely to be much higher as binding saturation was not observed. Even for small molecules, aptamers with K_d values in this range are considered very weak. Especially considering the lack of binding saturation, it is possible that sequence Hcy8 does not bind to methionine at all. This indicates that the aptamer is merely interacting with the amino acid backbone, characteristic of all amino acids. Instead the thiol group on the amino acid side chain must be critical in binding recognition. Using the magnetic beads, Hcy8 was found to bind to L-cysteine with a K_d of $22.6 \pm 0.8 \mu\text{M}$, over 100 times weaker than to its selected target (binding isotherm can be found in Appendix B). Again, no binding saturation was achieved so it is possible that there is no binding recognition with this target. However, the lower K_d value for cysteine compared to methionine is not unexpected. L-cysteine and L-homocysteine vary by a single extra carbon position in the side group carbon chain. Regardless, the significant reduction in

binding indicates that the aptamer is interacting with the entire side group of the amino acid and not simply the sulhydryl group. Further selectivity studies with other biologically relevant thiols, such as glutathione will be tested in the future.

5.3.6 Nanoparticle sensor

Several cysteine sensors using citrate-capped AuNPs have been developed. All of these sensors have relatively high limits of detection and lack specificity^{256, 257}. As a result, Wang et al. developed a cysteine sensor using citrate-capped AuNPs and ssDNA²⁵⁸. As described earlier, ssDNA coils around AuNP resulting in the nitrogen atoms of the bases adsorbing to the gold surface, stabilizing the AuNPs against salt-induced aggregation. However, due to the strong gold-thiol interaction, cysteine displaces ssDNA molecules adsorbed on AuNPs, making them susceptible to salt-induced aggregation. With this system, Wang et al. were able to develop a AuNP biosensor that was selective for thiol-based compounds.

The AuNP aptamer assay in Chapter 3 was useful for measuring the binding affinity of the developed OTA aptamers. In that case, the target-binding to the aptamer prevented the aptamer from adsorbing to the AuNP, resulting in salt-induced aggregation. Unfortunately, with homocysteine and cysteine, the competing interaction of the thiol group with the AuNP as well as with thiol group with the the aptamer prevented accurate K_d determination using this AuNP assay. However, given that the sensor described by Wang et al. was not selective for homocysteine over cysteine, we tested whether using the aptamers selected for L-Hcy, rather than using random DNA, would either improve the sensitivity of the assay, or impart further selectivity to this sensor.

Figure 5-8 shows the results of the nanoparticle binding assays. Despite the likelihood of a non-specific gold-thiol interaction, the preliminary comparison of the binding of the aptamer shows a 3 times improved response with homocysteine ($K_{d \text{ apparent}} = 1.9 \mu\text{M}$) compared to cysteine ($K_{d \text{ apparent}} = 5.9 \mu\text{M}$). While these binding response curves do reflect a preferential aptamer response with homocysteine over cysteine, none of the binding models accurately reflect the experimental data. This suggests that several competing interactions are involved, including the target-aptamer interaction, the aptamer-AuNP interaction and the target-AuNP interaction. This is further emphasized by the difference of the observed K_d value for the aptamer obtained here ($1.9 \mu\text{M}$) and in section 5.3.4 ($0.8 \mu\text{M}$). Furthermore, the selectivity tests in section 5.3.5 suggest that L-cysteine binds to the aptamer with 100 times weaker affinity. With the AuNP system, cysteine displayed on a three times reduction in affinity while high concentrations (3.6M) of L-methionine did not result in any aggregation of the AuNPs. L-methionine will not interact with the AuNP due to the lack of thiol. These AuNP responses with cysteine and methionine underscore the importance of the gold-thiol interaction in this assay.

Despite the interference of these competing factors, it is clear from the binding isotherms that the aptamer imparts some selectivity in this system as a weaker binding using control ssDNA results in a $K_{d \text{ apparent}} = 4.3 \mu\text{M}$.

The preliminary truncation experiments, using the aptamer sequence with the primer regions removed, resulted in a $K_{d \text{ apparent}} = 3.6 \mu\text{M}$. However, the binding model again does not accurately reflect the data points which are not dissimilar from the full length aptamer data points. With the shorter sequence length, this minimal aptamer may be easier to remove from the AuNP surface. Therefore, at this time it is unclear whether

this apparent reduction in binding compared to the full-length aptamer indicates that the primer regions are partially involved in L-Hcy binding or aptamer secondary structure. Further minimal aptamers will be tested using the beads based affinity chromatography method.

Figure 5-8: Binding isotherms using the AuNP assay. Target concentration (either L-homocysteine or L-cysteine) was varied with a constant concentration of aptamer, control ssDNA or minimal aptamer. An apparent K_d was elucidated from the aggregation of the AuNP, measured as the ratio of absorbances at 660 nm and 524nm. A theoretical curve was fit using Microsoft Excel Solver.

While the competing interactions may not be ideal for homocysteine detection at high concentrations, the sensitivity and selectivity of this system may be useful for detecting low concentrations of homocysteine. When the lower linear range of the biosensor is plotted (see Figure 5-9), it is clear that the aptamer provides a 2-3 times

signal enhancement compared to controls (non-apramer with homocysteine or aptamer with cysteine). From the R^2 values, it is clear that the linear relationships for the control samples are not ideal. Therefore, at these concentrations it is possible that the aptamer-target interaction

Figure 5-9: Derived calibration curve of the adsorption ratio (A_{660}/A_{524}) at the lower concentration range of target (either L-homocysteine or L-cysteine) with a constant concentration of aptamer or control ssDNA.

dominates, resulting in an observed aggregation of the AuNP. Within the higher linear range (see Figure 5-10) that more accurately reflect the control data, the signal enhancement is not as pronounced. At these concentrations, the gold-thiol interaction likely dominates and the sensor loses its selectivity for homocysteine.

The AuNP assay was carried out in serum at several dilutions. Even at a 1:10v/v dilution of serum in water, using high concentrations of L-Hcy (3.6 μ M), salt-induced aggregation did not occur. This is likely due to the high concentration of protein and other macromolecules in the serum. These may coat the nanoparticles, preventing aggregation. While optimization of the salt conditions may help restore aggregation, using the current conditions, only highly diluted serum samples can be tested. While the typical gold-thiol interaction may not result in the sensitivity needed to detect L-homocysteine at these diluted concentrations, it is clear that our aptamer-based system that favours these conditions. At these concentrations it is expected that only the L-homocysteine will be detected. Therefore, the aptamer imparts both sensitivity and selectivity for homocysteine in this AuNP assay.

Figure 5-10: Derived calibration curve of the adsorption ratio (A_{660}/A_{524}) for high concentrations of target (either L-homocysteine or L-cysteine) with a constant concentration of aptamer or control ssDNA.

5.4 Conclusions

Eight rounds of SELEX were performed to obtain DNA sequences displaying high binding affinity to the target, L-homocysteine. Two of the 4 sequences isolated from selections displayed high affinity for the target. Sequence Hcy8 bound to L-Hcy with a K_d in the low micromolar range. Given that this would represent the second smallest DNA aptamer target ever reported, and given the challenge of finding aptamers for such a small molecule, this is an excellent binding affinity. This sequence was also found to be selective, binding to the similar target L-methionine and L-cysteine with 200 and 100 times less affinity, respectively. There will be continued efforts to confirm affinity of this aptamer in solution. Further analysis will attempt to identify the minimal target-binding sequence in order to shorten the aptamer and improve binding while lowering its production cost. The stability and selectivity of the aptamer under physiological conditions and in serum will also be further examined. With this full length sequence, we have developed a AuNP based biosensor able to detect low concentrations of L-homocysteine. Results suggest that using this aptamer, the biosensor is selective for homocysteine.

Chapter 6. Improvement of SELEX libraries

Luo, X.; McKeague, M.; Pitre, S.; Dumontier, M.; Green, J.; Golshani, A.; DeRosa, M. C.; Dehne, F. Computational approaches toward the design of pools for the in vitro selection of complex aptamers. *RNA* **2010**, *16*, 2252-2262.

6.0 Statement of contributions

The computational experiments were run by Xuemei Luo under the supervision of Prof. Frank Dehne from the Carleton University School of Computer Science with library design constraints suggested by M. McKeague. Additional design conditions were suggested by the Carleton Bioinformatics group. All experimental work including the aptamer selection, library synthesis, cloning and binding characterization was performed by M. McKeague.

6.1 Introduction

6.1.1 Aptamer secondary structure

Structural analysis of many aptamers shows that they form intricate binding pockets for their ligands. In the free form of the aptamer, the ligand binding region is often completely unstructured and binding of the target proceeds by an induced fit mechanism²⁵⁹. However, a better understanding of the binding mechanism requires further knowledge of the structure. In particular, there is much interest in the three-dimensional arrangement of the many secondary structural elements of aptamers, including double-stranded helices, single-stranded loops, bulges and hairpins²⁶⁰.

When examining the secondary structure of nucleic acids, a common structural element is the junction, which connects different helical segments²⁶¹. For example, in Figure 6-1 there are 5 connected helical domains; thus, the overall structure is a 5-way junction. It has been shown that junctions are involved in a variety of different functional and structural roles in nucleic acids in general, including the self-cleaving catalytic properties of the hammerhead ribozyme²⁶², promotion of functional folded states of the

hairpin ribozyme²⁶³ and translation initiation of the hepatitis C virus at the internal ribosome entry site²⁶⁴.

Figure 6-1: An example of a 5-way junction where there are five connected helices (S1 to S5).

Much research has been performed on the structure and function of riboswitches, which are considered to be natural aptamers^{265, 266}. More specifically, riboswitches are RNA devices that consist of an aptamer domain coupled to an expression platform that regulates a gene²⁶⁷. Binding of a ligand to the aptamer alters its structure and, subsequently, the structure of the expression platform. Riboswitches are commonly found to fold into high-order junction secondary structures (such as 4-way and 5-way junction structures). For example, the lysine riboswitch involves three-helical and two-helical bundles joined by a five-way junction²⁶⁸. Almost all other riboswitches can be classified into one of two types based on their high-junction number. Type I riboswitches

consist of a fundamental three-way junction architecture. For Type II riboswitches, such as the SAM-I binding riboswitch, the aptamer domain is arranged in a four-way junction structure²⁵⁹. This higher-order structure characteristic of riboswitches is critical to target binding, for example in recognition of the binding pocket domain by purine riboswitches²⁶⁹.

However, analysis of existing synthetic aptamers has revealed that the majority of oligonucleotide sequences generated from SELEX experiments have simple structures with low degrees of complexity and that highly complex structures are far less abundant^{270, 271}. For example, aptamers that bind to ATP, chloramphenicol, neomycin B and streptomycin all have linear stem-loop or slightly branched structures²⁷². In fact, of the 2793 known aptamers from the Ellington Lab aptamer database²⁷³, less than 1% of the aptamer predicted secondary structures are 4-way and 5-way junctions (see Figure 6-2).

This low degree of complexity is concerning as it has been demonstrated that there is a general relationship between structure and function in aptamers. For example, Carothers et al. specifically noted a relationship in complexity and affinity for aptamers generated for the target GTP: an increased number of stems of the aptamer structure resulted in an improved affinity to the target²⁷⁴.

Figure 6-2: Structural distribution (percent of sequences displaying a given junction order: either one-way, two-way, three-way, four-way, or five-way junction structures) of the 2793 known aptamers from the Ellington Aptamer Database. Secondary structures were determined using ViennaRNA and the CountJunctions algorithm was used to determine the number of junctions.

6.1.2 SELEX libraries

As aptamer binding is related to its structure, and examples from nature and preliminary research show that increased complexity is critical for improved function, the question remains: why do synthetic aptamers form simple secondary motifs? In a recent computational analysis of *in vitro* random RNA libraries²⁷⁵, it was demonstrated that random libraries are themselves not structurally diverse and heavily favor simple topological structures. The structural distribution for random libraries shown in Figure 6-3, which is consistent with the analytical results from Sabeti et al., illustrates that 99.80%

of structures in a random library are expected to be 1-way, 2-way, or 3-way junctions²⁷⁶.

This lack of structural diversity in random libraries suggests that complex structure motifs with high-order junctions would be rare in selected aptamers.

Figure 6-3: Average structural distribution (percent of sequences displaying a given junction order: either one-way, two-way, three-way, four-way, or five-way junction structures) of 1 million sequences generated *in silico* for random SELEX libraries containing 15-nt primer sites at both ends. Secondary structures were determined using ViennaRNA and the CountJunctions algorithm was used to determine the number of junctions.

It is known that structural complexity is largely dependent on sequence length.

As the length increases, there is a greater probability of accessing more complex structural motifs²⁷⁶. However, oligonucleotide libraries for SELEX are usually restricted to sequences of at most 100nt in length due to the decrease in coupling efficiency for

their synthesis²⁷⁷. Additionally, primer binding regions of at least 15 nucleotides at each of the 3' and 5' ends must be included within this length for PCR purposes. While these primer regions may participate in the formation of secondary structure motifs, it is less likely. Therefore, the number of variable positions available for the formation of new complex motifs is reduced from 100 to 70 nucleotides.

However, Carothers et al. demonstrated that increasing the structural diversity of the starting oligonucleotide library could enhance the possibility of finding novel aptamers with improved activity^{130, 274, 278}. As a result, novel approaches have been employed that aim at increasing the structural complexity of starting libraries. One approach to increase the structural complexity is to keep the structure constant and introduce random segments in the proximity of the existing structure²⁷⁹⁻²⁸². Other approaches involve changing sequence length and composition^{283, 284}. Davis et al. performed SELEX using a typical random library as well as a partially structured RNA library consisting of a designed stem-loop within the random region. Subsequent analysis of the obtained aptamers revealed that most of those with highest affinity originated from the partially structured portion of their library⁵³. However, the above observations are based on adhoc library designs for individual experiments and targets. Furthermore, these adhoc libraries with more complex structures were generated by inserting particular stem-loops that were known a priori to be relevant for the given target. Such prior knowledge is not available in many cases. Hence, it is necessary to develop more general library design approaches that can be used for targets without such prior knowledge. Such designs should make use of primer binding regions as part of their strategy for predicting high complexity structures for use in SELEX. Additionally,

to increase the number of functional sequences available in the starting library for SELEX, the starting library should be engineered to have a uniform distribution over all possible structures²⁷⁵.

6.1.3 Genetic algorithms

Computational methods, in particular genetic algorithms, may be useful in designing a structured SELEX library. Genetic algorithms are search techniques used to find solutions to optimization problems and often applied in bioinformatics, computational science, engineering, economics, chemistry, manufacturing, mathematics and physics. It is a very powerful process as it mimics the natural process of evolution, using techniques inspired from nature including mutation and recombination. While several variations exist, the overall process for genetic algorithms consists of initialization, selection, reproduction and termination (see Figure 6-4)²⁸⁵.

Initialization

The method begins with an initial generation of individual solutions forming the current population. This population is usually randomly generated to allow a greater range of possible solutions (i.e. the search space). While the population size depends on the particular problem, there are usually several hundreds or thousands of possibilities.

Selection

Each individual in this initial population is evaluated and assigned a fitness score based on its ability to solve the problem. A fraction of the existing population is then

stochastically selected to breed a new generation. The individuals with the best fitness scores are more likely to be selected²⁸⁵.

Figure 6-4: Process of a genetic algorithm.

Reproduction

The selection members of the first population are then modified using three operations: copy, crossover and mutation for optimization purposes. These new solutions are then assigned a fitness score and form the new population. Generally, the average fitness will have increased by this procedure for the new population, as typically the best individuals from the previous generation were selected for breeding.

Termination

This overall process is repeated hundreds or thousands of times until a termination position is reached. Ideally, termination occurs if a solution is found that satisfies minimum criteria. However, termination may also result if a plateau occurs where successive rounds of selection and reproduction do not produced improved results, or until a maximum number of generations is reached due to time constraints.

6.1.4 Chapter objectives

The majority of selected aptamers fold into simple structural motifs, such as hairpins. The lack of structural complexity in random DNA libraries may account for this lack of complex structural motifs in existing aptamers. This chapter investigates whether initial SELEX DNA libraries containing sufficient pre-determined structural complexity would be useful in finding new functional motifs. These initial SELEX libraries should be sufficiently diverse such that they contain new complex structural motifs without losing the simple motifs that are known to result in high affinity aptamers.

As such, a computational method called *Genetic Filtering* is used to design a DNA library recipe with uniform structural distribution (i.e. having equal proportion of simple and complex structural motifs). This library is tested in a new selection for the target ATP. The initial ATP aptamer selection by Huizenga et al. resulted in an aptamer with moderate affinity and low complexity that has been well-studied by several groups. This target therefore serves as an excellent comparison for the structured library, to investigate whether, if given the opportunity, selection of new functional motifs for the same target is possible.

6.2 Experimental

6.2.1 Computational method: random filtering

A random filtering method was used to selectively increase the number of 5-way junctions in a SELEX DNA library. The initial population consisted of 1 million 100 nt random sequences with and without primer binding regions. From these sequences, those that formed 5-way junction secondary structures, as determined by ViennaRNA folding software, were isolated. These 5-way junctions were used as the starting point for potential SELEX libraries and the theoretical structural distribution of each 5-way junction was determined by generating 1 million mutated sequences *in silico*. These 1 million sequences were generated by mutating every non-primer, single-stranded position with a randomly chosen DNA base (either A,C,G or T) to substitute for the original. ViennaRNA was used to generate all secondary structures. The ViennaRNA 1.6.5²⁸⁶ software package includes RNAfold which generates a single RNA or DNA secondary structure prediction through energy minimization based on a dynamic programming

algorithm originally developed by Zuker et al²⁸⁷. ViennaRNA provides a nucleotide pairing list indicating for each nucleotide whether or not it is paired with another nucleotide. From this information, the number of junctions for each sequence was computed using the CountJunctions algorithm (see Appendix C).

6.2.2 Computational method: genetic filtering

Genetic filtering was used to design a DNA SELEX library with a uniform structural distribution, i.e. 20 % of each 1-way, 2-way, 3-way, 4-way and 5-way junctions. The pseudo-code is provided in Appendix C. The initial generation of library designs was based on a set of 5-way junctions which contained a given set of primer binding regions. Each library design was assigned a fitness score indicative of its proximity to the desired uniform structural distribution. The fitness was assessed as follows: For a sequence S , let jct_i be the percentage of i -way junctions counted after mutating its single-stranded positions 10^4 times, and let $djct_i$ be the desired percentage of i -way junctions ($i = 1, \dots, 5$). The fitness score for S was calculated as $\sum_{i=1}^5 |jct_i - djct_i|$. The desired percentage of i -way junctions is the uniform structural distribution and each $djct_i$ was set to 20 %. The libraries were selected such that there would be 24 single-stranded random positions, excluding primer binding sites, to ensure complete sequence space coverage ($4^{24} \sim 10^{14}$). Therefore, when each library design was assigned a fitness score, sequences that contained a number of single-stranded positions (num_ss) not equal to the desired number of single-stranded positions ($dnum_ss$) were penalized. The details of this Algorithm, *Fitness_function_with_penalty*, is in the Appendix C.

New generations of library designs were then obtained by selecting designs from previous generations with better (i.e. smaller) fitness scores. Here, a roulette wheel selection method was used to select library designs from the current population where each design was chosen with a probability that corresponded to its relative fitness. By repeatedly spinning the roulette wheel, the new intermediate population was selected, where fitter individuals had a greater chance to be selected than weaker ones. Finally, crossover, mutation and copy operations were applied to the intermediate population to create the next generation. For the crossover method, two sequences S1 and S2 from the intermediate population of library designs were chosen. A random number, *randomPosition*, was generated. To create the next generation, the sub-sequences of S1 and S2 from a given position were switched to allow potential switching of sub-structures between two sequences.

When a sequence in the intermediate population was selected for mutation, a random number was generated between 0 and 1 for each position of the sequence excluding primer sites. If the number was less than a parameter $P_{position}$, the position was mutated by choosing one of the other three nucleotides, with equal probability. The mutated sequence was then used in the next generation.

When a sequence in the intermediate population was selected for a copy operation, the sequence was simply copied unchanged to the next generation. The above crossover, mutation and copy methods have parameters $P_{crossover}$, $P_{mutation}$ and P_{copy} , respectively, which represent the likelihood of applying the respective method to a given sequence. Furthermore, for every 10 generations, the best sequence of all the generations was re-inserted into the new generation of library designs.

6.2.3 Reagents

Phosphoramidites, modifiers, activator, deblock, capping, and oxidizing reagents were obtained from Glen Research. Standard support columns and acetonitrile were purchased from BioAutomation. Ultra High Purity 5.0 argon was purchased from Praxair. All PCR and electrophoresis components were purchased from BioShop Canada. Adenosine 5'-triphosphate, immobilized on Agarose 4B, and chemicals for buffering agents and all other purposes were purchased from Sigma-Aldrich and prepared with Millipore Milli-Q deionized water at 18 MΩcm. Spin-X cellulose acetate microcentrifuge filter tubes, Amicon-Ultra 0.5 mL 3kDa centrifuge units and stirred-cell ultrafiltration membranes (1kDa, 25mm) were purchased from Fisher Scientific Canada.

6.2.4 DNA library and primer synthesis

The starting DNA library was prepared using standard phosphoramidite chemistry on a BioAutomation Mermade 6 (Plano, Texas) as described in section 2.2.2 using 200 nmol controlled pore glass (CPG) columns with the larger, 1000 Å pore size. The sequence *GFLibrary1* was obtained from the genetic filtering procedure. This library contained 24 random nucleotide positions and two primer binding sites necessary for PCR (shown in bold): 5'-

ATACCAGCTTATTCAATTGCNNNNGCAATTNNNGTCNGGACNNNNGTTCNG
ACNNTCGGCGNNNCGCCGANCTATCTNNNNN**AGATAGTAAGTGCAATCT**-3'.

A fraction of this library was set aside prior to SELEX for cloning and sequencing experiments. The following primers used for amplification and cloning of the selected

oligonucleotides were synthesized on a 1 μ mol scale: Primer1: 5'- ATACCAGCT-TATTCAATT-3' and Primer2: 5'-AGATTGCACTTACTATCT-3'. Modified primers were also synthesized to allow for fluorescent labeling of the binding DNA and separation of strands on a denaturing polyacrylamide gel: ModPrimer1: 5'-6FAM-ATACCAGCTTATTCAATT-3' and ModPrimer2 5'-poly-dA₂₀-HEG-AGATTGCACTTACTATCT -3'. HEG represents the hexaethylene glycol Spacer Phosphoramidite 18 (Glen Research). DNA was purified using denaturing polyacrylamide gel electrophoresis (12%) followed by clean-up with Amicon YM-3 centrifugal filter devices as described in section 2.2.2. Similarly, the *RFLibraryA* from the random filtering experiments and the corresponding primers required for PCR amplification were synthesized with the following sequences:

RFLibraryA 5'-GGAAGAGATGGCGACGCTACTCCCATCNGGTACCGTNNNNA-CGGTATTNNGACGCNNNNNGCGTCNNNACACNNNNGTGTGGATCAGCTGATCCTGATGG-3'

RFPrimer1 5'-GGAAGAGATGGCGACGCT-3' and *RFPrimer2*: 5'- CCATCAGGAT-CAGCTGAT-3'

6.2.5 ATP SELEX experiments

ATP-agarose and unmodified agarose column, 0.25 mL in volume, were washed with 10 mL of column buffer (300 mM NaCl, 5 mM MgCl₂, 20 mM Tris HCl, pH 7.6). The DNA library (5 nmoles), suspended in 1 mL of column buffer, was heated for 5 minutes at 75°C then cooled to room temperature over a 20 minute period, as performed by Huizenga et al⁸¹. The library was then immediately loaded into the unmodified

agarose column and left to incubate with mild shaking at room temperature for 30 minutes. The column was washed with 4 mL of column buffer, collecting the DNA that did not bind to the unmodified agarose material to be used for the selection process. After this pre-selection, 8 rounds of positive selection were performed as follows: The DNA library was subjected to 5 minutes of heating and a 20 minute incubation at room temperature followed by immediate incubation with the ATP-modified agarose column, as described above. After 30 minutes, sequences that did not have an affinity for ATP were washed away with 5 mL of column buffer. To obtain the sequences bound to the ATP, the column was incubated for 10 minutes with 0.5 mL elution buffer (40 mM Tris HCl, 3.5 M urea, pH 8) at 80°C and removed using centrifugation at 10,000 g for 1 minute. Five elution fractions were collected and the DNA was purified using an ultrafiltration stirred cell (Millipore).

The entire selected oligonucleotide library was amplified in 5-15 parallel PCR reactions. Each reaction consisted of 0.1 M Tris HCl pH 9.0, 50 mM KCl, 1% Triton X-100, 1.9 mM MgCl₂, 0.3 mM dNTP mix, 1 µM each primer and 5 units of Taq DNA polymerase. The DNA was initially melted for 10 minutes at 94.0°C, followed by 25 cycles of 94.0°C (1 min), 47.0°C (1 min) and 72.0°C (1 min). Final extension occurred at 72.0°C for 10 minutes after the last cycle. PCR products were dried down, heated at 55°C for 5 minutes in the presence of formamide and run on a 12% denaturing PAGE to separate the double stranded product. The fluorescein-labeled DNA strand (the selected sequences) was identified using an Alpha Imager UV-illuminator. The corresponding DNA bands were cut from the gel and extracted using the freeze/rapid thaw method described by Chen & Ruffner¹⁸⁰ (described in section 2.2.2) in 10 mM Tris HCl buffer,

pH 7.4. After purifying the DNA on the stirred cell and re-suspending it in column buffer, the DNA was used for the next selection round.

Enrichment of the DNA library was assessed throughout the rounds of selection by monitoring the percent of DNA library binding to the ATP column each round by measuring DNA absorbance at 260 nm using a Cary 300 BioUV-visible spectrophotometer and by measuring fluorescence (excitation 490nm, emission 520nm) with a Fluorolog Fluorescence Spectrophotometer (Horiba Jobin Yvon, USA) with a SpectrAcq controller.

6.2.6 Cloning and sequencing

The selected oligonucleotides from SELEX round 8 and a fraction of the starting library were amplified using the unmodified primers (Primer 1 and 2) and cloned using a StrataClone PCR Cloning Kit (Agilent Technologies) according to the manufacturer's protocol. The *RFLibraryA* was amplified using *RFPrimer1* and *RFPrimer2* and subjected to the same cloning procedure. Briefly, the ligation reaction mixture was prepared by combining 3 μ L of StrataClone Cloning Buffer, 2 μ L of the PCR product and 1 μ L of the StrataClone Vector Mix in a microcentrifuge tube and mixed by repeated pipetting. The ligation reaction was incubated at room temperature for 5 minutes and then placed on ice. 1 μ L of this ligation reaction was added to the StrataClone Solopack competent cells that were thawed on ice. The competent cells were incubated on ice for 20 minutes, followed by a 45 second heat-shock at 42°C and a 2 minute incubation on ice. 250 μ L of LB broth at 37°C was added to the transformation reaction mixture, which was then allowed to recover for 1 hour at 37°C with gentle agitation. LB-ampicillin

plates for blue-white screening were prepared by spreading 40 μ L of 2% X-gal on each LB-ampicillin agar plates (1.0% NaCl, 1.0 % tryptone, 0.5 % yeast extract, 2.0% agar, and 100 μ g/mL ampicillin). 20 to 100 μ L of the competent cells were spread onto the plates. The colonies were left to grow overnight at 37°C. Positive colonies (white) were picked and removed carefully from the agar and vortexed in 50 μ L of deionized water. These samples were sent for direct sequencing at the University of Calgary University Core DNA Services using the T3 promoter primer and T7 promoter primer. Sequencing results were analyzed and vector sequences were deleted from the total sequence to retain only the data that represent the selected random sequences. The secondary structures of these sequences were predicted using ViennaRNA.

6.2.7 Dissociation constant (K_d) experiments

Aptamer synthesis

Aptamer sequences ATPF4 and ATPF8 were synthesized with the 5' fluorescein modifier 6-FAM (sequences are found in Table 6-6) and purified by PAGE as described in section 2.2.2.

Affinity chromatography method

An affinity chromatography assay was set up by weighing samples of 1 mg of ATP agarose (approximately 23 nanomoles of ATP) into Spin-X cellulose acetate microcentrifuge filter tubes. These were then washed with column buffer. Dilutions of fluorescein-modified ATP aptamer were prepared in column buffer at concentrations ranging from 1 nM to 10 μ M. After incubation for 30 minutes, the tubes were centrifuged at 10,000 g for 10 minutes. 100 μ L of 90°C column buffer was added and the

tubes were vortexed briefly, followed again by centrifugation at 10,000 g for 10 minutes. The fluorescence of the DNA fraction was measured (excitation wavelength 490nm; emission wavelength 520nm). The fluorescence was plotted versus the concentration of aptamer. As before, the K_d was determined by minimizing the residual values between calculated and observed experimental fluorescence using the solver feature of Microsoft Excel^{181, 182}.

6.3 Results and Discussion

6.3.1 Increasing the number of complex structures in SELEX libraries

The first goal was to increase the percentage of complex structures in initial libraries for SELEX experiments. Random filtering was used to selectively increase the number of 5-way junctions in such libraries. It was initially unclear whether using a million rounds of mutations was sufficient to determine the structural distribution of a random library with a reasonable degree of accuracy. Table 6-1 shows a typical example for the structural distribution of a given library determined with a different numbers of mutations. Each experiment was repeated 100 times, and the numbers shown represent the average and variance for each value. The averages are very stable for different numbers of mutations. However, the variance is clearly reduced with additional experiments, reaching very close to zero when the number of mutations reaches 1 million. Hence, using 1 million mutations was determined to be sufficient to reliably estimate the structural distribution.

Table 6-1: Average structural distribution for a typical SELEX library determined using a different numbers of mutations (variances in brackets).

Number of mutations	Percentage of sequences forming 1 to 5-way junction structures (%) and the corresponding variances				
	1	2	3	4	5
10^2	10.04 (9.72)	49.45 (30.23)	27.12 (20.19)	8.78 (9.71)	4.61 (4.54)
10^3	10.26 (0.70)	49.99 (3.05)	26.45 (2.04)	8.44 (0.65)	4.86 (0.41)
10^4	10.31 (0.09)	49.87 (0.29)	26.69 (0.19)	8.34 (0.07)	4.80 (0.04)
10^5	10.36 (0.01)	49.93 (0.02)	26.57 (0.02)	8.36 (0.01)	4.78 (0.00)
10^6	10.35 (0.00)	49.93 (0.00)	26.58 (0.00)	8.35 (0.00)	4.79 (0.00)

Using ViennaRNA to predict the secondary structures of 10^6 100nt random sequences, 76 5-way junction sequences were found. These sequences were subjected to Random filtering to generate an enhanced library, consisting of a larger proportion of 5-way junction structures. Figure 6-5 shows an increase in the structural complexity (using junction order) of a library generated from random filtering (using the 76 5-way junctions as a starting point) compared to a library of 1 million 100nt random sequences. The most frequent structures in the random filtering libraries are 3-way junctions (48.65%) while the most frequent structures in the random libraries are 2-way junctions (64.35%). In the random filtering libraries, 10.23% of the sequences are 5-way junctions and 15.45% compared with 0.01% in a random library.

As the primer binding sites are generally used in SELEX experiments, the effect of including them on structure diversity was determined. The 15 nucleotide primer binding regions GGAAGAGATGGCGAC at the 5' end and AGCTGATCCTGATGG at the 3' end were chosen, thereby occupying 30 fixed positions in the starting sequence.

With only 70nt left for design modification compared to the original 100nt, finding complex structures was made considerably more difficult²⁷⁶. From a 100nt library of 20 million random sequences containing these primer binding regions, only 35 sequences were found exhibiting 5-way junctions. These sequences were used to generate a library using random filtering. Compared with 76 5-way junctions found in a 1 million 100nt random library without primer sites, 5-way junctions are 40 fold less abundant in the same library containing fixed primer binding sites. According to Sabeti et al.²⁷⁶, a 5-way junction motif in a 70nt random library is approximately 100 to 200 fold less abundant than a 5-way junction motif in a 100nt random library. However, primer binding sites are able to participate in the formation of junctions even though the effective random library length has been decreased. Therefore, the 5-way junction abundances in a 100nt library containing 30nt primer sites should be higher than a simple 70nt random library, explaining the 40 fold difference between these two cases.

Similar to the case without primer binding sites, random filtering produced a more structurally complex library (Figure 6-5). The most frequent structures in the random filtering libraries are 3-way junction sequences (45.76%) while the most frequent structures in random libraries are again 2-way junction sequences (71.72%). In the libraries containing primer binding sites, random filtering produces an average of 6.65% 5-way junctions compared to 0.00% for completely random libraries. Therefore, random filtering increases the percentage of highly complex structures available for selection in the SELEX experiments. The sequence shown in Figure 6-6 called *RFLibraryA* has 19 single-stranded positions (indicated as N) excluding primer sites and resulted in the largest number of 5-way junctions (31.13%) see Table 6-2.

Figure 6-5: Comparison of average structural distribution (percent of sequences displaying a given junction number: either one-way, two-way, three-way, four-way, or five-way junction structures) of 1 million sequences for typical random nucleic acid libraries (light colours) and libraries generated by the random filtering process (dark colours), with (red shades) and without (blue shades) primer binding sites. The random filtering method greatly increases the percentage of high-complex structures such as 5-way junctions by 10% without primers and 6% with primers.

The *RFLibraryA* was synthesized, amplified, cloned and sequenced and sample sequences were folded using ViennaRNA. From the 17 sequences recovered, 11.76% of them formed 5-way junctions (see Table 6-2); sequences are in Table 6-3.

Figure 6-6: Sample secondary structure of the *RFLibraryA* DNA sequence including primer sites GGAAGAGATGGCGAC and AGCTGATCCTGATGG designed to generate the highest percentage of 5-way junction sequences (31.13%). The sequence has 19 single-stranded positions, excluding primer sites (noted as N). As a starting library for SELEX experiments, these 19 single-stranded positions are considered as random (N) and the other positions are considered as fixed. Image generated by Mfold.

Table 6-2: *In silico* and *in vitro* structural distribution for the *RFLibraryA* library with primer binding sites GGAAGAGATGGCGAC and AGCTGATCCTGATGG shown in Figure 6-6. The *in vitro* values were generated by cloning and sequencing of the library following solid-phase synthesis (17 sequences were obtained). The sequences are in Table 6-3.

	Percent of sequences (%)				
	1-way junctions	2-way junctions	3-way junctions	4-way junctions	5-way junctions
<i>In silico</i>	3.28	25.68	24.68	15.22	31.13
<i>In vitro</i>	0	47.06	29.41	11.76	11.76

Table 6-3: DNA sequences obtained experimentally from the *RFLibraryA* and the corresponding number of junctions as predicted by ViennaRNA.

Name	Sequence (5' to 3')	Number of junctions
RF1	GGAAGAGATGGCGACGCTACTCCATCGGTCCGTACACGGTTTAAGA CGCTTTTGCCTCTTAACACGGGCGTGTGATCAGCTGATCCTGATGG	2
RF2	GGAAGAGATGGCGACGCTACTTCCATCCGGTCCGTCCGAACGGTAT TGAGACGCTTTGAGCGTCTATACACGCACGTGTGGATCAGCTGATC CTGATGG	2
RF5	GGAAGAGATGGCGACGCTATCCACCGGTACCGTAGATACGGTATTC GGACGCACGCCGCGTCGATACACATTAGTGTGGATCAGCTGATCCT GATGG	3
RF6	GGAAGAGATGGCGACGCTACTCCCATCTGGTACCGTAGTTACGGTA TTTTGACGCGCTGCGTCGTAAACACTGTATAGATCAGCTGATCCTGA TGG	2
RF7	GGAGAGATGGCGACGCTACTCCCATCTGGTACCGTTATAACGGTAT TAAGACGCTTAATGCGTCATTACACTTCCATAAGGATCAGCTGATCC TGATGG	2
RF8	GGAAGAGATGGCGACGCTACTCCCATCCGGTACCGTCGGACGGTAT TGGGACGCGCCACGCGTCGGGACACAGTAGTGTGGATCAGCTGATC CTGATGG	3
RF9	GGAAGAGATGGCGACGCTATCCCATCGGGTACCTTAATAACGGTAT TAAGACGCAAAGGGCGTCGTTACACGCACGTGTGGATCAGCTGATC CTGATGG	5
RF10	GGAAGAGATGGCGACGCTACTCCCATCTGGTACCGTACTCACGGTA TTTTGACGCATTTAGCGTCTATACACGAATGTGTCGATCAGCTGACC TGATGG	3
RF11	GGAAGAGATGGCGACGCTATCCATCTGGTACCGTGAAGGGCGAATT CCACAGGGACGGTATTTGGACGCTTCGGCGTCTCGACACAAGTGTG TGGATCAGCTGATCCTGATGG	2
RF12	GGAAGAGATGGCGACGCTACTCCCATCTGGTACCGTAGTTACGGTA TTTTGACGCGCTGCGTCGTAAACACTGTATAGATCAGCTGATCCTGA TGG	2
RF13	GGAAGAGATGGCGACGCTCCCCACCGGCCTGAAACGGTATTTGACG CTAATAGCGTCTGGACACCCATAAGGATCAGCTGATCCTGATGG	2
RF14	GGAAGAGATGGCGACGCTACTCCCATCGGGTACCGTATTTACGGTA TTCAGACGCCGGACGCGTCTCAACACATTAGTGTGGATCAGCTGAT CCTGATGG	5
RF15	GGAAGAGATGGCGACGCTACTCCCATCAGGTACCGTATGACGGTAT TTAAGACGCTCAAAGCGTCGGCACACATTTGTGTGGATCAGCTGAT CCTGATGG	4
RF16	GGAAGAGATGGCGACGCTACTCCCATCAGGTACCGTATGACGGTAT TCTGACGCCAGTTGCGTCATAACACAGCAGTGTGGATCAGCTGATC CTGATGG	3
RF18	GGAAGAGATGGCGACGCTCTCCCATCCGGTACCGTATGTACGGTAT TGTGACGCTGCGTGCCTCACAACACATACGTCCAGATCAGCTGATC CTGATGG	2
RF19	GGAAGAGATGGCGACGCTACTCCCATCCGGTACCGTGTTACGGTAT TCAGACGCTAGATGCGTCCCTACACCTTAGTGTGGATCAGCTGACC TGATGG	4
RF20	GGAAGAGATGGCGACGCTACTTCCATCTGGTACCGTAAGCACGGTA TTTCGACGCAGTGGGCGTCGACACACTCGTGTGTGGATCAGCTGAT CCTGATGG	3

It may initially appear that the structural diversity of the libraries designed using random filtering will be overly constrained given the stretches of fixed positions in the stems. However, an examination of a number of structures resulting from the same design demonstrates that while the sequence in these regions is constrained, the corresponding structure is not. Examples of different junction structures formed from the *RFlibraryA* in Figure 6-6 are shown in Figure 6-7. These sample structures illustrate that the same fixed sequence region may participate in a stem in one structure while forming a loop in another. For example, for the 5-way junction structure shown in Figure 6-7A, the sequence GCGT starting at base 61 participates in a stem while the same fixed sequence forms a loop in the corresponding 4-way junction structure illustrated in Figure 6-7B.

Overall, compared to nearly zero five-way junctions and only 0.2% four-way junctions in a typical random DNA library with 10^6 structures of 100 nucleotides each, random filtering was able to generate libraries of the same size where more than 10% of the sequences are 5-way junctions and more than 15% of the sequences are 4-way junctions. Very similar results were found for 100nt DNA libraries with fixed primer binding sites. Compared to a completely random library, the random filtering method was able to dramatically increase the likelihood of obtaining high-order junctions experimentally.

Figure 6-7: Sample sequences and representative secondary structures for *RFLibraryA*, shown in Figure 6-6. A) a 5-way junction structure B) a 4-way junction structure C) a 3-way junction structure and D) a 2-way junction structure.

6.3.2 Designing a SELEX library with uniform structural distribution

The genetic filtering method was initiated using a set of 5-way junctions for the initial library design. The reason for this is that it is possible to obtain low complexity structures through the mutation of high complexity structures; however, the opposite is highly unlikely. For example, by mutating 50 different 3-way junction sequences 1 million times each, for a total of 50 million sequences, only 62 5-way junction structures were obtained. This observation was also reported in a recent investigation where 5-way

junctions were derived from the pre-existing 3-way junctions through evolutionary pathways, but it was observed that the probability was very small²⁸⁸.

For the purposes of the genetic filtering, only 10^4 instead of 10^6 mutations of single-stranded positions were performed for the fitness score calculations. This number was reduced for fitness score estimation simply to save computation time as the method takes hours on a 200 processor PC cluster. As shown in Table 6-1 above, the structural distribution estimated with 10^4 mutations is close to that estimated with 10^6 mutations. Furthermore, all final results were confirmed through a re-calculation of the the fitness score using 10^6 mutations.

A uniform structural distribution was chosen such that the highest possible ratio of high-order junction structures could be sampled without losing the ability to select for simple motifs such as 2-way junction structures that are known to make high affinity aptamers.

The reason for selecting 24 single-stranded positions lies in the SELEX experiment. Most SELEX experiments begin with a starting library consisting of nanomoles of sequences, as this is a reasonable amount of DNA that can be synthesized. This equates to approximately 10^{15} sequences⁴⁷. Given that 4 possible bases can be inserted at any random sequence position and $4^{24} \approx 10^{14}$, this implies that 24 random positions allows complete sequence coverage of the initial library (ie. allowing several expected copies of every different sequence), as well as sufficient diversity (10^{14} different sequences).

The genetic filtering method was used to create a SELEX library of 100nt in length, containing the primer binding sites ATACCAGCTTATTCAATT and

AGATAGTAAGTGCAATCT. With each generation, the fitness score decreased steadily (see Figure 6-8). After 571 generations with a runtime of 2hrs 10min on a 200 processor cluster, the library with the best fitness score was obtained.

Figure 6-8: The best fitness value observed for each generation of genetic filtering for the 100 nucleotide library containing the primer binding sites ATACCAGCTTATTCAATT and AGATAGTAAGTGCAATCT. Inset: Sample secondary structure of the DNA library sequence (GFLibrary1). This 100nt library, containing 24 random nucleotide positions, was synthesized and used in selection experiments against ATP. Image generated by Mfold.

This design, *GFLibrary1* shown in Figure 6-8, had a structural distribution close to a uniform distribution with a distance from the optimal of 2.95 (see Table 6-4). This library was synthesized, cloned and sequenced (see Table 6-5). In the subset of only 19 sequences, 26.3% were found to be 5-way junctions and 21.1% were 4-way junctions.

Compared to a completely random library, the random filtering method was able to dramatically increase the likelihood of obtaining high-order junctions experimentally. More importantly, even within this small sample set, the percentage of each junction structures is fairly uniform (nearly 20%), with a distance from the optimal of only 27.4.

Table 6-4: *In silico* structural distribution of the *GFLibrary 1* obtained after 571 generations of the genetic filtering method. The structural distributions prior to starting SELEX and after 8 rounds of selection using ATP as the target are also given.

	Percent of sequences (%)					Distance from 20%
	1-way junctions	2-way junctions	3-way junctions	4-way junctions	5-way junctions	
<i>In silico</i>	19.6	20.4	18.9	20.4	20.7	2.95
<i>In vitro</i> (initial)	15.8	26.3	10.5	21.1	26.3	27.4
<i>In vitro</i> (after 8 rounds of SELEX)	0	50	30	0	20	80

This second method, genetic filtering, was able to generate DNA library designs of 100 nucleotides each, where 30nt are fixed primer sites, with distributions *in silico* that are very close to uniform. Analysis of a small set of the sequences obtained from the *GFLibrary1 in vitro* showed significant structural diversity, with each of the junction structures present at levels of at least 10%. It is expected that with a larger sample of sequences, the distribution of junction structures would be near the optimal uniform distribution. However, for cost reasons, only a small number of sequences were obtained.

Table 6-5: DNA sequences obtained experimentally from the initial *GFLibrary1* and the corresponding number of junctions predicted by ViennaRNA.

Name	Sequence (5' to 3')	Number of junctions
ATPS11	ATACCAGCTTATTCAATTGCTGCCGCAATTTGCGTCGGGACTAGTGTT CAGACTGTCTGGCGGGCCGCCGATCTATCTTAAGAAGATAGTAAGTGCAATCT	1
ATPS12	ATACCAGCTTATTCAATTGCTTTAGCAATTGGGGTCCGGACTGATGTT CCGACCCCTCGGCGCTGCGCCGATCTATCTAACAGAGATAGTAAGTGCAATCT	2
ATPS14	ATACCAGCTTATTCAATTGCCTGGGCAATTCATGTCAGGACTAAAGTT CCGACATTCGGCGAATCGCCGAATCTATCTGTCGAGATAGTAAGTGCAATCT	2
ATPS15	ATACCAGCTTATTCAATTGCGCGTGCAATTCGGGTCAGGACAACCTGTT CAGACTATCGGCGTTACGCCGAATCTATCTGCTAGATAGTAAGTGCAATCT	2
ATPS16	ATACCAGCTTATTCAATTGCTGATGCAATTGCGGTCAGGACCCTAGTT CAGACGTTCTGGCGCCACGACGATCTATCTTGCTAAGATAGTAAGTGCAATCT	1
ATPS17	ATACCAGCTTATTCAATTGCTATCGCAATTACCGTCAGGACTTCAGTT CGGACGTTTCAGCGCCACGCCGAATCTATCTCTCAGATAGTAAGTGCAATCT	2
ATPS18	ATACCAGCTTATTCAATTGCAGCAGCAATTAACGTCAGGACGGGAGTT CCGACTTTCTGGCGTGGCGCCGATCTATCTACTGTAGATAGTAAGTGCAATCT	5
ATPS19	ATACCAGCTTATTCAATTGCTTGAGGCAATTACTGTCAGGACTTGTTC GACTGTCTGGCGCCTCGCAGACCTATCTATGTAGATAGTAAGTGCAATCT	3
ATPS20	ATACCAGCTTATTCAATTGCGGGTGCAATTAGAGTCGGGACATIGGTT CTGACTCTCGGCGATCCGCCGAATCTATCTGCCCTAGATAGTAAGTGCAATCT	2
ATPS21	ATACCAGCTTATTCAATTGCCTCGGCAATTGTTGTCCGGACTATAGTT CCGACTGTCTGGCCTCTCGCCGAATCTATCTGCTCAGATAGTAAGTGCAATCT	1
ATPS22	ATACCAGCTTATTCAATTGCTTGCGCAATTCGTGTCAGGACTAATGTT CTGACCCTCGGCGAGACGCCGATCTATCTTTAGTAGATAGTAAGTGCAATCT	4
ATPS23	ATACCAGCTTATTCAATTGCGGCTGCAATTATAGTCGGGACTAATGTT CAGACTTTCTGGCGTTACGCCGATCTATCTAGATGAGATAGTAAGTGCAATCT	5
ATPS24	ATACCAGCTTATTCAATTGCAACCGCAATTTTGTGTCAGGACGGATGTT CTGACGATCGGCGCAACGCCGATCTATCTTCTAGATAGTAAGTGCAATCT	4
ATPS25	AGATTGCACTTACTATCTGGTTAAGATAGATCAGCGAGGCGCCGATAG TCTGAAGTGTCTCCAGATTAAATTGCCTACGCAATTGAATAAGCTGGTAT	3
ATPS26	ATACCAGCTTATTCAATTGCCTCAGCAATTTACGTCTGGACTCAAGTTC CGATCTCGGCGCCTCGCCGAGCTATCTCAAGGAGATAGTAAGTGCAATCT	5
ATPS27	ATACCAGCTTATTCAATTGCCTTGGAATTTTCGTCGGGACTTCCGTCC GACAATCGGCGATCCGCCGACCTATCTATACTAGATAGTAAGTGCAATCT	5
ATPS28	ATACCAGCTTATTCAATTGCAAGTGCAATTAAGGTCAGGACGCCAGTT CAGACATTCGGCGATGCGCCGACCTATCTAGTAAAGATAGTAAGTGCAATCT	5
ATPS29	ATACCAGCTTATTCAATTGCGACTGCAATTCAGTCTGGACTTTTGT CGGACACTCGGCGACTCGCCGAATGTCTACTTAAGATAGTAAGTGCAATCT	4
ATPS30	ATACCAGCTTATTCAATTGCGCTTGCAATTCGGGTCCGGACTATCGTT CAGACTCTCGGCGTGCCGCCGATCTATCTTCGAGAGATAGTAAGTGCAATCT	4

6.3.3 Selection of ATP aptamers using the improved library design

SELEX experiments were performed using the uniform structure library, *GFLibrary1*, generated by the genetic filtering method, in an effort to confirm that the design approach yielded a library that allowed for sampling of an increased range of complex structures during the selection process. Selections were performed to find DNA sequences with affinity for ATP⁸¹. In the original ATP aptamer SELEX experiments, the random DNA library used to screen for ATP binders consisted of $\sim 2 \times 10^{14}$ different sequences using a 72 nucleotide random region flanked by defined primer binding sites. After 8 rounds of selection, the random region of the sequence with the best binding affinity was determined to be 5'-

GACTGGGCTTGTGCTTGGGGGAGTATTGCGGAGGAAAGCGGCCCTGCTGAAG
TGGGATACATGTGGATACCC -3' (personal communication, J.W. Szostak). This sequence, with its flanking primer sites, is of low complexity, predicted to form a 2-way junction. Given the extensive study of this system, ATP was chosen as the target for testing the uniform structure library, *GFLibrary1*. A total of eight rounds of SELEX were performed, equivalent to the original procedure by Szostak. The recovery of DNA after each round, corresponding to the amount of DNA binding to the ATP target, was monitored using UV-Visible spectrometry absorbance and fluorescence measurements. With each round, a low but increasing percentage of the library DNA bound to the target was observed (see Figure 6-9).

Figure 6-9: Percent recovery of binding ssDNA to the ATP agarose after each selection round.

DNA from *GFLibrary1* after 8 rounds of enrichment was cloned and sequences were obtained. The secondary structures were analyzed using ViennaRNA. From the 10 sequences, 2-way junctions, 3-way junctions and 5-way junctions were found in the enriched library (see Table 6-4), confirming not only the structural diversity of this designed library, but also that complex structures were retained after several rounds of enrichment. Sequences following the 8 rounds of selection are in Table 6-6.

Table 6-6: DNA sequences obtained experimentally from the *GFLibrary1* following 8 rounds of SELEX using ATP as the target. The corresponding number of junctions predicted by ViennaRNA are listed.

Name	Sequence (5' to 3')	Number of junctions
ATPF1	ATACCAGCTTATTCAATTGCGACAGCAATTAGTGCTGGACGTGTGTCC GACTCTCGGCGAATCGCCGAACATCTCCGTCAGATAGTAAGTGCAATCT	2
ATPF2	ATACCAGCTTATTCAATTGCTTCCGCAATTTGGGTCCGACTATGGTTCC GACTTCGGCGATGCGCCGATCTATCTTAGACAGATAGTAAGTGCAATCT	2
ATPF4	ATACCAGCTTATTCAATTGCCACCGCAATTTATGTCAGGACGCCAGTTCA GACGATCGGCGTACGCCGACCTATCTTAATTAGATAGTAAGTGCAATCT	5
ATPF5	ATACCAGCTTATTCAATTGCCATTGCAATTCGGGTCTGGACACAGGTTCC GACGTTCGGCGCTCCGCCGATCTATCTCGTAGAGATAGTAAGTGCAATCT	3
ATPF6	ATACCAGCTTATTCAATTGCCGTAGCAATTCGGGTCGGGACGGGGGTTCC GACTATCGGCGAAACGCCGGCTATCTTGTAGATAGTAAGTGCAATCT	2
ATPF7	ATACCAGCTTATTCAATTGCCCAAGCAATTTGGGTCTAGGACATGTGTTCC GACTATCGGCGTCCGCCGAACATCTTAAATAGATAGTAAGTGCAATCT	2
ATPF8	ATACCAGCTTATTCAATTGCTCCGGCAATTTCCGTCAGGACCGGGGTTCT GACACTCGGCGTGACGCCGACCTATCTATTCAAGATAGTAAGTGCAATCT	5
ATPF9	ATACCAGCTTATTCAATTGCCGGTGCAATTAAGTCAGGACTGCGTTCCGG ACTTTCGGCGCGCCGCCGATCTATCTTCTTAGATAGTAAGTGCAATCT	3
ATPF10	ATCCCAGCTTATTCAATTGCCAGGCCAATCTTGCTGGACCACAGTCCGA CTTTCGGCGCCTCGCCGAGCTATCTCACGTAGATAGTAAGTGCAATCT	2
ATPF11	ATACCAGCTTATTCAATTGCGTTAGCAATTGCAGTCTGGACTCCCGTTGA CCCTCGGCCAACGCCGATCTATCTTGACTAGATAGTAAGTGCAATCT	3

After the *GFLibrary1* was subjected to 8 rounds of selection against the target ATP, complex 5-way junction structures accounted for 20% of the sampled sequences. However, in typical random libraries out of 20 million sequences, only 35 fold into 5-way junction structures. Therefore, the likelihood of observing this concentration (20%) of 5-way junction structures in a random library is approximately 10^{-10} . These results confirm that the suggested approach not only leads to better access to more complex structures for in vitro selection experiments but also that SELEX is capable of preserving highly complex structures if they are present in the initial DNA library.

Two of the 5-way junction sequences, ATPF4 and ATPF8 were tested for binding affinity to ATP using affinity chromatography, yielding dissociation constant (K_d) values of $24.4 \pm 0.8 \mu\text{M}$ and $3.7 \pm 0.1 \mu\text{M}$, respectively (binding isotherms can be found in

Appendix B). In comparison, a K_d of $6 \pm 3 \mu\text{M}$ was reported for the published ATP aptamer sequence⁸¹. It is important to note that the published ATP aptamer sequence underwent extensive optimization and minimization prior to binding affinity experiments. One 5-way junction generated from the SELEX experiments shows equivalent of slightly improved binding over the optimized system; however, the other 5-way junction does not. This suggests that complexity alone does not guarantee better target binding, but higher complexity structures do have the potential to yield better aptamers. This clearly underscores the need for structural diversity, not just structural complexity, in the initial starting library.

6.3.4 Choice of ViennaRNA secondary structure folding package

Both approaches for SELEX library design are based on nucleic acid secondary structure folding, using RNAfold from the ViennaRNA package. As a result, the computational methods are subjected to the limits of the folding algorithms. In particular, RNA sequences that adopt pseudoknot structures are not sampled in our approach since RNAfold in the ViennaRNA package cannot fold a sequence into a pseudoknot. Mfold is another nucleic acid folding package using thermodynamic methods²⁰⁴. Hofacker et al. compared the performance of RNAfold and Mfold, and found that RNAfold is considerably faster than Mfold²⁸⁶. Recently, other folding programs such as GTfold²⁸⁹ and MC-Fold²⁹⁰ have been developed. Figure 6-10 shows the comparison of the structural distributions of the same 1000 random sequences predicted by RNAfold, GTfold and MC-Fold. The structural distributions generated by RNAfold and GTfold are almost identical. However, MC-Fold predicted considerably

more complex structures. Upon closer inspection of MC-Fold, it was determined that the secondary structure prediction begins by enumerating all possible junction configurations and therefore has a tendency to overestimate the number of complex structures. The runtime for each of these methods was also compared. On a single machine, the run time required for a single sequence by RNAfold and GTfold are very similar²⁸⁹. However, MC-Fold takes several orders of magnitude longer. On a PC using an Intel Core 2 Quad Q6600 processor (2.4GHz and 8GB DDR2 RAM), the run time of RNAfold for folding one sequence was about 0.008029 seconds, while it took about 845 seconds for MC-Fold, which is approximately 100,000 times slower than RNAfold. Thus, it would take 27 years for MC-Fold to predict the secondary structure for one million sequences. It is therefore not feasible to use MC-Fold to predict secondary structures in these computational methods despite its capacity to predict pseudoknot structures. Although there are more recent methods available for secondary structure prediction, ViennaRNA is faster and sufficiently accurate for the purpose of computing structural distributions for a large number of sequences.

6.3.5 Comparison to recent work

Recently, Schlick and coworkers developed a computational approach for designing structured RNA libraries by modeling the library generation process using mixing matrices²⁹¹⁻²⁹³. However, synthesis of several individual libraries is required for this approach, compared to the typical one-column synthesis required for the random and genetic filtering approaches. Therefore, these methods are less costly and less time

Figure 6-10: Comparison of predicted structural distributions of RNAfold (blue), GTfold (red) and MC-Fold (green) for the same 1000 random sequences of 100 nucleotides each (containing no fixed primer sites). The structural distributions generated by RNAfold and GTfold are almost identical. However, MC-Fold predicted considerably more complex structures.

consuming. Additionally, random filtering and genetic filtering are able to add more complex structures to the library. For example, the most complex structures created via mixing matrices are 4-way junction structures whereas random filtering and genetic filtering can increase and control, respectively, the number of 5-way junction structures in the library. The mixing matrices approach also does not include primer binding sites. However, primer-binding regions are often required for SELEX experiments. Since the primer regions are considered in the random and genetic filtering approaches, the resulting libraries can be directly implemented into SELEX experiments. While we

demonstrated that genetic filtering generates libraries with uniform structural distributions, the method can be used to design any given structural distribution, and not only for junction structures. This would not only be possible for different junction structures, but also for other defined secondary structures.

6.4 Conclusions

We have developed two methods for designing improved starting libraries for functional nucleic acid selection experiments: random filtering and genetic filtering. Random filtering yields a DNA library with significant enrichment in the number of highly complex structures present. Genetic filtering results in DNA libraries with a nearly uniform distribution over all possible structures from 1-way junctions to 5-way junctions. A uniform library design created with genetic filtering was synthesized and subjected to a SELEX experiment to isolate ATP binding aptamers. It was found that after 8 rounds of selection, complex 5-way junction structures accounted for a sizeable percentage of the library, confirming that our methods greatly improved generation of high complexity structures and that these structures were preserved during the selection process. The disparate binding affinity of two 5-way junction sequences from this enriched library confirmed that higher complexity sequences have the potential to lead to better aptamers, but certainly do not guarantee improved binding. This suggests that SELEX library designs should aim to increase structural complexity while still preserving a diversity of both low and high complexity structures. Future work will examine in depth whether these designed libraries more generally lead to aptamers with improved binding characteristics.

Chapter 7. A SELEX database

Cruz-Toledo, J.; McKeague, M.; Zhang, X.; Giamberardino, A.; McConnell, E.; Francis, T.; DeRosa, M. C.; Dumontier, M. Aptamer base: a collaborative knowledge base to describe aptamers and SELEX experiments. *Database (Oxford)* **2012**, 2012, bas006.

7.0 Statement of contributions

The database model was conceptualized by M. McKeague. The Aptamer Base and the aptamerbase.semanticscience.org web page were developed by Jose Cruz-Toledo from the Department of Biology. Curation was performed by Jose Cruz-Toledo, Tariq Francis, Amanda Giamberardino, Erin McConnell, M. McKeague and Xueru Zhang under the supervision of M. McKeague. The query programs were developed by Jose Cruz-Toledo with initial analysis by Phuong Que Ho and M. McKeague.

7.1 Introduction

7.1.1 Current aptamer databases

Over the past two decades, hundreds of new aptamers have been selected and incorporated into many biosensor platforms and therapeutic systems: a recent search on pubmed for ‘aptamer’ results in almost 3000 total publications. Additionally, the amount of sequence data generated by these experiments has been accumulating exponentially²⁹⁴. With the growing interest in development and application of aptamers, this trend is expected to continue. Despite this, most of the information describing the SELEX methods, resulting aptamers and binding capabilities remain in the unstructured and textual format of peer reviewed publications.

Efforts for storing aptamer information in a database format have been made by a few groups. For example, Ponomarenko et al. were interested in the applicability of *in vitro* selected aptamers to the analysis of natural genes. As a result, they developed the first aptamer database, SELEX_DB, in 2002 that focused on sequences from SELEX that helped define natural DNA and RNA recognition sites for proteins²⁹⁵. In 2004, the

Ellington group created a more complete database, the Aptamer Database, which encompassed information from all *in vitro* experiments. This database was intended for identifying aptamers and unnatural ribozymes that already existed as well as for collecting information about SELEX experiments to develop a better understanding of the distribution of functional nucleic acids in the sequence space²⁷³. In 2006, a separate group created RiboaptDB that contained the same information as the Aptamer Database with a greater emphasis on artificial ribozymes²⁹⁴.

Unfortunately, RiboaptDB is no longer accessible to the public. Additionally, none of these databases have been updated since 2006. More importantly, these databases cannot be updated with information about new aptamers and SELEX experiments by the aptamer community, outside of the groups that developed the databases. Furthermore, the collected information is restricted to constricted HyperText Markup Language (HTML) forms, limiting the type and number of queries that can be posed against these datasets²⁰⁸.

Several nucleotide sequence data providers exist such as GenBank²⁹⁶, the European Molecular Biology Laboratory Nucleotide Sequence Database (EMBL-Bank)²⁹⁷ and DNA Database of Japan (DDBJ)²⁹⁸. These databases do not maintain a list of artificially created sequences. Additionally, journals do not require entry of these non-natural sequences into any of these databases. Consequently, the community lacks an updated data resource in which information about aptamers, their sequences and the experimental conditions used in their selection can be collected and searched.

7.1.2 Freebase

The vast quantity of information that one has access to in 2012 is unprecedented in human history. With this information overload there is a risk of creating more noise than value, thereby limiting its effective use. Innovations to organize and process this information are therefore critical to maintain its utility²⁹⁹. The semantic web is one such innovation.

The term Semantic Web was coined by the inventor of the World Wide Web, Tim Berners-Lee. The Semantic Web is a component of the Web 3.0 which would enable users to find, share, and combine information more easily. While humans are currently capable of using the Web to carry out general search tasks, machines cannot accomplish this without human direction. The semantic web, however, would be a system of information that could be readily interpreted by computers to perform the work involved in finding, combining, and acting upon information found in the system. This system would require that the relevant information sources be semantically structured³⁰⁰.

Freebase is a semantically enriched database platform that aims to store information in a structured manner. Freebase (<http://www.freebase.com>) is a free, openly licensed community-built resource for structured data that currently provides information about more than 22 million topics. Over 360 million facts are organized into bases, which are collections of thematically related topics.

Data in Freebase is stored as a structured graph in which entities or topics represent single concepts or individuals. This is based on the entity relationship model described which is a conceptual representation of data first described by Peter Chen³⁰¹. Every topic in Freebase is identified by a unique identifier and is accessible by web browser through a Uniform Resource Locator (URL). Topics can in turn be categorized

or "typed", representing a relationship between topics and their respective types (termed an "IS A" relationship). For example, typing a given topic with both the RNA and aptamer types indicates that the particular topic is a ribonucleic acid and is also an aptamer. Every type in Freebase may hold any number of properties with any given value. The topic therefore is described as having a "HAS A" relationship with the given property. For example, a ribonucleic acid has a sequence property, which may in one case be GATTACA. This results in an intricately connect graph of topics and data that span across virtually any combination of information tables. Connecting the information in this way allows unlimited and detailed queries to be posed of the stored data.

7.1.3 Chapter objectives

This chapter details the development of a new aptamer database that contains detailed information about aptamers and the SELEX experiments used for their selection. Freebase was chosen for the database platform to allow other members of the aptamer community to contribute, to help ensure that the database remains current. The data model allows easy searches for aptamers for targets of interest. Additionally, the information stored in the database provides other aptamer researchers the necessary information to repeat any given SELEX experiment. Using the information in the database, we investigate whether experimental design conditions, such as library length, are important for the development of high affinity aptamers.

7.2 Experimental

7.2.1 Data model

The Aptamer Base can be found at <http://aptamer.freebase.com/>. The data model is built upon a central *interaction experiment* which describes the results and experimental procedures used to elucidate biomolecular interactions. These experiments may report one or more resulting *interactions*. Each interaction is related to the participants of the reported interaction, through the “has participants” property (Figure 7-1). In the case of aptamers derived from SELEX, this interaction experiment is more specifically described as a *SELEX experiment* and results in *aptamers* and an *aptamer target* as participants in an interaction.

Figure 7-1: Basic type-relationship map used by the Aptamer Base to describe SELEX experiments. The Interaction experiment type has an interaction as the outcome. Each interaction has at least 2 interactors as participants (for SELEX experiments these are aptamers and the aptamer target). Rectangles denote types and arrows represent properties between topics. Overlapping rectangles represent the multiple types associated with a topic.

The SELEX experiment provides details about the SELEX methods used to generate the aptamer sequences. This includes the *partitioning method* used to isolate putative aptamers from sequences lacking binding properties as well as the *recovery method* often required to isolate the putative aptamers from the aptamer-target complex. Within the SELEX experiment entry, additional details required to reproduce a given experiment are entered. For example, the particular type of SELEX method (from Table 1-1 in Chapter 1), the number of SELEX rounds performed, the initial SELEX library template sequence as well as the selection buffer conditions (termed *selection solution*). All fields within the data model describing the details of the SELEX experiment can be found in the data map, Figure 7-2.

Figure 7-2: Basic type-relationship map used by the Aptamer Base to describe the details of SELEX experiments. Rectangles denote types and arrows represent properties between topics. Overlapping rectangles represent the multiple types associated with a topic.

Details are included about the aptamers such as the nucleic acid type (DNA or RNA), sequence, length and secondary structure. The aptamer target may have several properties depending on its type (i.e. protein, small molecule, cell, etc), including peptide sequence, molecular weight, melting temperature, etc. The resulting interactions can be either “confirmed” or “disputed” with the *affinity experiment* under a set of experimental conditions, thus enabling searching for sequences displaying or lacking binding properties. The affinity experiment results in a *dissociation constant* value describing the affinity of the interaction (see Figure 7-3).

Figure 7-3: Basic type-relationship map used by the Aptamer Base to describe the details of affinity experiments. The affinity experiment results in dissociation constant and either ‘confirms’ or ‘disputes’ an interaction. Rectangles denote types and arrows represent properties between topics. Overlapping rectangles represent the multiple types associated with a topic.

This data model also provides details about minimal aptamers, sub-sequences or variants of larger sequences that exhibit binding to a target ligand (Figure 7-4). While sequences obtained from SELEX are typically greater than 30 nucleotides in length, minimization often occurs to reduce the cost of synthesis or to improve binding characteristics. The reported minimal aptamers participate in interactions that are distinct from those of their parent aptamers, allowing these aptamers to be described with a *dissociation constant* topic that is distinct from that of its parent.

Figure 7-4: Basic type-relation map describing minimal aptamers. Minimal aptamers are described in the Aptamer Base by creating a new interaction for each individual minimal aptamer.

For every aptamer reported in the Aptamer Base the *predicted secondary structure* using RNAfold is included²⁸⁶. The resulting predictions associated to aptamer topics that have been typed as a nucleic acid (either RNA or RNA). The dot bracket notation, minimum free energy and version of RNAfold are provided for each prediction. Given that the nucleic acid property "has predicted secondary structure" is a one-to-many relation, the Aptamer Base could accommodate the incorporation of alternative secondary structure predictions or even tertiary structures made by other prediction programs^{204, 290}.

7.2.2 Curating the aptamer literature

Approximately 3000 articles from the primary scientific literature, from 1990 until December 2011, were identified as potential candidates for inclusion in the Aptamer Base. This collection was created by performing a general search on PubMed for any articles containing the terms "SELEX" or "aptamer" in either the title or abstract. The resulting collection was manually inspected to ensure that only relevant articles, describing a SELEX experiment used to generate aptamers, were assembled for the database entry process. Therefore, review papers, articles describing applications of previously selected aptamers, articles characterizing or modifying existing aptamers and articles detailing SELEX of *in vitro* ribozymes were excluded from the final aptamer database collection. The final set of 157 articles published between 1990 and 2006 was curated by six expert graduate students, all of whom have performed experimental research on aptamers or SELEX. Additionally, only aptamers that had been verified using an *affinity experiment* and had an experimentally determined dissociation constant

were considered for entry. 200 articles published after 2006 were identified for entry into the database and are currently being curated on a daily basis.

To ensure data curation was consistent and accurate, each curator was given the same training in extracting relevant details from aptamer publications (see Appendix D for the documentation describing the data entry process). Any curator questions were directed to a central SELEX expert for consistency purposes. Finally, each entry was verified for completion and accuracy by at least one other curator, and proper typing of the generated Aptamer Base topics was monitored using a correction program.

At times, information required for the data model was not present in the primary literature. In some cases this was due to the particular experimental design chosen by the author and in other cases due to incomplete information presented in the research article. To distinguish between these cases and an incomplete entry on behalf of a curator, the terms “not described” and “not needed” could be selected for each entry.

7.2.3 Querying the database

Data in the Aptamer Base was programmatically accessed through Freebase's application programming interface (API) using the Metaweb Query Language. A suite of Java 1.6 programs were developed which implement Google's latest Java API clients (<http://code.google.com/p/google-api-java-client/>) to query the Aptamer Base.

7.3 Results and Discussion

7.3.1 Summary of data in the Aptamer Base

The data model was carefully constructed based on the methods and results from typical SELEX publications and as suggested by experts in the field. In particular, the required entries were chosen such that the SELEX experiments can be independently repeated²⁵. The Aptamer Base currently contains information about 1068 interactions involving 1239 aptamers binding to 221 different targets totaling approximately 45,000 facts (updated statistics can be found <http://aptamerbase.semanticscience.org/?q=node/5>). These facts were added on a weekly basis between September 2010 and March 2012. More specifically, the Aptamer Base describes interactions involving 497 DNA aptamers and 660 RNA aptamers, with proteins being the most frequently occurring target type (Figure 7-5).

Figure 7-5: Summary of data (as of March 1, 2012) found in the Aptamer Base. (A) Distribution of the aptamer types (B) Distribution of the aptamer targets.

7.3.2 Accessing the database

Searching and browsing data in the Aptamer Base is possible in a number of ways. For researchers simply interested in searching for a list of aptamers, there is a newly created custom website (<http://aptamerbase.semanticscience.org/?q=base/search>) which includes a search interface that directly queries the contents of the Aptamer Base through a simple user-friendly search form. The search can be limited by selecting either RNA or DNA aptamers as well as by selecting a particular target type of interest (see list in Figure 7-5). Users can also search for a particular target name by inputting all or part of the target name. The user can choose from a list of available options which are auto-completed by Freebase Suggest (<http://www.freebase.com/docs/suggest>). PubMed identifiers can also be specified to limit the search to specific publications. This custom website also provides users with documentation on using and accessing the data in the Aptamer Base.

Secondly, users have programmatic access to the Aptamer Base through the Freebase application programming interface (API) (<http://www.freebase.com/docs/data>). The Freebase API allows users to query the Aptamer Base using the Metaweb Query Language (MQL), a powerful tool that allows users to formulate complex queries that are unlimited unlike the searches possible through our simple search interface. Using the API, queries are serialized as JavaScript Object Notation (JSON) through a Representational State Transfer (REST) interface. While this is a powerful query tool, the average aptamer researcher is likely to make use of the custom built search form.

7.3.3 Mining the database

Data in Freebase is also available as part of the Linking Open Data (LOD) project³⁰², a community effort that collects data from various open data sources on the Web and publishes and interlinks the different sources using the Resource Description Framework (RDF) in a dereferenceable manner³⁰³. As such, any topic in Freebase and specifically in the Aptamer Base can be retrieved as RDF/ (Extensible Markup Language) XML by using the following URL: [http://rdf.freebase.com/rdf/\[TOPIC_ID\]](http://rdf.freebase.com/rdf/[TOPIC_ID]), where TOPIC_ID is the unique identifier given by Freebase to a topic.

Lastly, full data dumps of the Aptamer Base are available as tab-delimited files from <http://download.freebase.com/datadumps/latest/browse/base/aptamer/>.

7.3.4 Utility of the database

The Aptamer Base is a new collaboratively developed knowledge base motivated by the information needs of aptamer researchers. The Aptamer Base provides detailed information about the experimental conditions under which aptamers were selected and their binding affinity determined. The data model provides the required types and properties to capture experimental details of both the partitioning method that eliminates sequences lacking affinity from the SELEX library and the recovery method of the putative aptamer sequences from aptamer-target complexes. The Aptamer Base also details the pH, temperature, metal cation concentration and the buffering agent for all SELEX experiments. Providing access to these experimental conditions is crucial in ensuring the reproducibility of the reported experiments²⁵. This information can be leveraged by aptamer researchers wishing to either reproduce or select new aptamers.

For example, researchers attempting to develop aptamers for a specific target can query the Aptamer Base for the conditions of successful SELEX experiments for identical or similar targets. Additionally, researchers wishing to develop new diagnostic and therapeutic applications for aptamers may find all the necessary conditions required to reproduce the aptamer binding.

7.3.5 Comparison with other databases

While other aptamer databases have been documented^{273, 294, 295}, the Aptamer Base provides significant benefits. As the Aptamer Base is part of the larger resource of Freebase, it reuses knowledge that is already in Freebase. For example, the Aptamer Base reuses information from Chemistry Commons (<http://freebase.com/view/chemistry>) to provide links to PubChem and Wikipedia for the majority of aptamer target topics. As a result, several properties about the aptamer targets, such as molecular weight are automatically included.

As stated previously, the RiboaptDB is no longer accessible to the public. Furthermore the other aptamer databases, the Ellington Aptamer Database and SELEX_db, have not been updated since 2006. Aptamer Base is currently in the process of being updated to include all characterized aptamers until 2011. The risk of Aptamer Base following the fate of past databases and being left un-updated is mitigated by the fact that data entry can be performed by any registered member of Freebase. This unique feature enables researchers, educators and students within and outside of the aptamer community to contribute and make use of the knowledge in the Aptamer Base. There is

currently a stable community of over 60,000 users that contribute to Freebase on a regular basis.

Though providing open access to the Aptamer Base poses a risk of vandalism, Freebase relies on community collaboration to maintain a complete and accurate dataset, similar to other successful open data projects such as Wikipedia. Users can correct or augment facts in the Aptamer Base, but may be suspended if in violation of basic Freebase content creation guidelines. Furthermore, administrators of the Aptamer Base can at any point modify, remove or undo changes to any entry to ensure data quality and consistency. For this purposes, Freebase provides the community with the ability to revert any changes made to the Aptamer Base through the use of Freebase datadumps (<http://download.freebase.com/datadumps/>), which have complete backups of the data, updated weekly, in Freebase for the past two months. Access to older backups is also available by contacting the Freebase staff.

Unlike the Ellington Aptamer Database and RiboaptDB, the Aptamer Base currently only provides information about aptamers. The data model used by the Aptamer Base has been tailored to provide a core set of the necessary details required to generically reproduce current aptamer *in vitro* experiments. However, our data model can be used as a common set of requirements for reporting future *in vitro* aptamer experiments and subsequently expanded to accommodate results of other *in vitro* experiments that involve biomolecular interactions such ribozymes. Additional links describing the large field of applications of aptamers may also be easily incorporated into the current data model.

SELEX often generates several putative aptamer sequences. Often due to time constraints, the binding abilities of all sequences are not tested. Additionally, many sequences derived from SELEX do not bind to the target and are more likely a result of mutation or contamination²⁵². The Ellington Aptamer Database contains almost three times more aptamer sequences than the Aptamer Base. As the majority of these sequences have not been verified to be aptamers using an affinity experiment, there is always a possibility that these sequences actually have little or no affinity for the target. To ensure the relevance of the aptamer data in the database, the sequences found in the Aptamer Base have been verified using an affinity experiment to either bind (“confirm”) or not to bind (“dispute”) to the target in question. In almost all cases, the corresponding dissociation constant is included.

7.3.6 SELEX queries

As discussed earlier, the majority of initial *in vitro* experiments involved selecting RNA aptamers as it was considered that only RNA could fold into functional motifs. However, several groups reported that DNA aptamers were capable of binding specific molecular targets with high affinity. Additionally, the lack of the 2' hydroxyl of the DNA sugar results in an increased chemical and biological stability compared to RNA, a useful characteristic for almost all applications of aptamers. Despite this finding, researchers frequently select RNA aptamers due to the belief that RNA result in higher affinity binders. As a result, more RNA aptamers exist (see Figure 7-5). Using the information in the database we sought to investigate whether this was the case. From our queries, the average dissociation constant of RNA aptamers is 4.7×10^{-6} M, which is 50% better than

that of DNA aptamers ($1 \times 10^{-5} \text{M}$). Additionally, the highest affinity aptamer reported, binding to its target with a K_d of 2 pM (for the target VEGF), is in fact RNA. However, there is also a more significant spread between various binding capabilities of RNA. As a result, the two-sample t-test for equal means indicates that there is no significant difference between these observed K_d averages.

The effect of aptamer length was then further investigated. Firstly, there is a wide range of various aptamer lengths for both RNA and DNA (see Figure 7-6 and 7-7). Additionally, RNA aptamers tend to be shorter in length. From Figure 7-6, 95% of RNA aptamers are from 20 to 68 nucleotides in length with an average length of 44 nucleotides. This range is extended to between 26 and 90 nucleotides in length for DNA with an average of 58nt (Figure 7-7). Here, the two-sample t-test for equal means rejects the null hypothesis, indicating that there is a difference between these 2 values. The data therefore does indicate that RNA aptamers are shorter in length on average. While the average binding affinity of RNA aptamers is not significantly different than that of DNA, the difference in size may be sufficient reason for choosing to perform RNA selections. However, this observed difference may be a result in the difficulty in RNA synthesis. Due to the extra protecting steps, RNA-based selections may begin with much smaller random regions of SELEX libraries. Alternatively, more effort for the minimization of final aptamers may occur due to this increased cost and effort associated with RNA synthesis. Alternatively, the hydroxyl group at the 2' sugar position, which can participate in additional intermolecular interactions with the target, may result in the formation of tighter binding motifs within a smaller sequence length. Comparison of the length of the random regions in the initial library as well as the length of original aptamer

clones must be performed to provide further insight into this difference. Furthermore, the relationship between aptamer length and K_d was investigated. Initial inspection of both the RNA and DNA aptamer datasets results in no discernable trend (see Appendix D). As these datasets include all aptamers lengths binding to every target, there is likely several other factors involved that need to be inspected individually. From the work in Chapter 6, there is precedent that likely secondary structure plays a more defined role in aptamer binding affinity¹⁰⁴.

Figure 7-6: Histogram displaying the frequency of RNA aptamers (as of March 1, 2012) of given sequence lengths (within a 5 nucleotide range). The mean RNA aptamer length is 44 nucleotides. Image generated using SigmaXL.

Figure 7-7: Histogram displaying the frequency of DNA aptamers (as of March 1, 2012) of given sequence lengths (within a 5 nucleotide range). The mean DNA aptamer length is 58 nucleotides. Image generated using SigmaXL.

With the infinite variables available in the database, the effect of particular small molecular target type on the ability to select high affinity aptamers was investigated. While the majority of aptamer targets in the database are proteins, focus was on small molecules due to our interest in selecting new aptamers for these targets. Work has been performed by Carothers et al. at an attempt to determine the effect of target structure and size with binding affinity¹³¹. Using aptamers for only 6 small molecules from the literature, as well as aptamers obtained from his own selections for two other small molecules, Carothers determined that the target molecular weight was proportional to the resulting aptamer affinity (larger targets resulted in lower K_d values). This finding was consistent with the findings of other studies between affinity and target mass³⁰⁴. However, the target theophylline which has a very small mass (180 g/mol) did not follow

this general trend. It was therefore concluded that targets with fewer rotatable bonds, and therefore a reduced degrees of freedom, can result in improved aptamer affinity. Using a larger dataset, several affinity comparisons based on target types were performed. The average K_d values for a set of target types was determined. The significance of the difference in these values was evaluated using the two-Sample t-Test for equal means. Each of the 4 comparisons was consistent with the findings of Carothers et al.(see Table 7-1). For example, aptamers bind to nucleotides with improved affinity compared to aptamers to amino acids. This effect may be a result of the molecular weight argument (nucleotides have molecular weights of 300 g/mol or greater whereas amino acids are 204g/mol or less). However, this effect may be a result of the smaller fraction of rotatable bonds present in nucleotide structures compared to amino acids. Alternatively, the planar structure of the nucleotides as well as the high proportion of hydrogen bonding sites may contribute to the improved affinity compared to amino acids. To further investigate this effect, the average affinities of nucleotides vs. all other small molecules were compared. Again, nucleotides result in improved affinities comparatively. Here the molecular weight of all other small molecules spans 61g/mol to 1000 g/mol. Therefore, it is more likely that the reduced number of rotatable bonds as well as the ideal planar structure and hydrogen bonding sites result in the high affinity of these targets. The molecular weight argument however, was supported in the comparison of amino acids and oligosaccharides, in which the larger oligosaccharides had significantly improved binding affinities. Finally, cyclic compounds having less rotatable bonds resulted in improved affinities compared to acyclic compounds.

Table 7-1: Comparison of aptamer affinity based on target types (as of March 1, 2012). The mean K_d for each set of target types was compared. The 2-sample t test was used to evaluate whether there was a significant difference between these results (initial results from P Q. Ho³⁰⁵).

Target comparison	Mean K_d (μ M)		p value*
	Target 1	Target 2	
Nucleotides and amino acids	58	700	0.0068
Nucleotides and other small molecules	58	280	0.0197
Amino acids and oligosaccharides	700	0.5	0.0033
Cyclic and acyclic small molecules	34	430	0.0071

*2-sided p value, assume unequal variance (if less than 0.05 = true)

7.4 Conclusions

We have created a comprehensive database that stores structured information about aptamers and the conditions under which they were selected and tested. Currently this database contains additional up-to-date information compared to other aptamer databases and we are in the process of including all selected aptamers to January 2012. The platform on which this database was created, Freebase, will help ensure that up-to-date information is available in the future. Aptamer researchers are currently able to use our simple search form to search for aptamers for specific targets or target types. Additionally, more comprehensive queries can be made through Freebase. The database has been queried to extract the sequence length, dissociation constant and target information for aptamers stored in the database. This data was statistically manipulated and revealed several trends between aptamer binding affinity and target type. Further queries will be performed to establish trends in aptamer binding and the experimental conditions and methods used in their selection with the goal of streamlining aptamer selection protocols in the future.

Chapter 8. Contributions to knowledge

While aptamer technology has existed for over two decades, the challenges imposed on the development of aptamers for small molecules has resulted in very few novel aptamers that can bind to practical small molecules. While aptamer detection systems are often superior to other biomolecular-based sensing schemes, relatively few have emerged for practical use. This is mainly due to two reasons: relatively few studies have been performed at examining the robustness of such systems in practical scenarios and in relevant food matrices. Secondly, as there is a lack of produced practical aptamers, the majority of aptamer detection systems use targets for proof-of-concept.

To evaluate whether aptamers display the robustness needed for commercial applicability, we examined the feasibility of using ochratoxin A (OTA) aptamers in commercial food-safety testing. The OTA aptamer 1.12.2 used in the study was able to perform well in aptamer-based solid-phase extraction columns (SPE) for rapid clean-up of OTA from naturally contaminated wheat samples. These aptamer-based SPE columns showed good performance, with recovery and precision data within the criteria established by the European Union. Comparative analysis with the well-used immunoaffinity (IMA) columns showed good correlation; however, the aptamer-based SPE columns displayed improved reusability properties and a lower cost of production compared to the IMA columns.

Selections were performed on several useful small molecule targets relevant for diagnostics and therapeutics. From our SELEX experiments, strategies were evolved that can help improve aptamer selection for expensive or difficult targets. For example, using magnetic beads greatly reduced the amount of target, and therefore cost, required for a SELEX experiment. Conditions which increased selection stringency may lead to better

aptamers. However, the implementation of the stringent conditions must be chosen carefully and correspond to the resulting enrichment of the library. The research found that computationally designed SELEX libraries with an increased structural diversity result in an improved sampling of different secondary motifs during selection experiments. This can lead to an improved likelihood of selecting aptamers with desired binding properties.

These strategies were applied for the selection of improved aptamers for the targets OTA and ATP. These aptamers, displaying higher affinity for their targets, may be useful in the development of sensitive biosensing platforms for detection of these targets. Furthermore, we were the first to select novel aptamers for the targets fumonisin B₁ (FB₁) and L-homocysteine (L-Hcy). Given the importance of FB₁ in food-safety and the clinical relevancy of L-Hcy, these new aptamers may be used in rapid, inexpensive detection systems.

With partners that are experts in computer science, a new aptamer database called Aptamer Base was developed. The database included all confirmed aptamers up to 2007. This database allows aptamer researchers to search for aptamers that bind to either a group of targets or a specific target. The availability of a user-friendly and complete aptamer database will help researchers find practical targets for the development of new therapeutic and diagnostic designs. Additionally, SELEX labs can easily compare their selected aptamers with all of those in the literature. The hope is that by having all the aptamer targets located in one collection, researchers will chose to select aptamers for new targets that are currently not available.

The Aptamer Base captures detailed experimental conditions about SELEX and the properties of the resulting aptamers. With this information we have determined that RNA aptamers do not necessarily result in improved binding characteristics compared to their more stable counterpart, DNA. Additionally, we have confirmed previous reports that aptamer affinity is related to the molecular weight of the target, with the exception of targets with less rotational freedom. Additional links comparing the conditions used in SELEX and aptamer affinity will be made with the hope of streamline SELEX protocols for specific targets of interest and work towards the possibility of intelligent aptamer design. Ultimately, we hope that this work will result the rapid and facile development of novel aptamers.

Appendices

Appendix A Towards the selection of aptamers for deoxynivalenol

A.1 Introduction

Deoxynivalenol

Deoxynivalenol (DON) is a trichothecene mycotoxin produced by several fungal species of the genera *Fusarium*, mainly *F. graminearum* and *F. culmorum*³⁰⁶. These species are responsible for *Fusarium* head blight of wheat and barley and *Fusarium* ear rot in maize³⁰⁷. In addition to the financial loss from these plant diseases, the resulting contamination with deoxynivalenol and other trichothecenes results in significant waste. For example, direct loss of more than \$100 million resulted from a *Fusarium* epidemic in Ontario winter wheat alone in 1996³⁰⁸.

There are approximately 150 structurally related sesquiterpenoid compounds that are members of the trichothecene group. This group is characterized by a tetracyclic 12,13 epoxytrichothec-9-en backbone and has been classified according to their chemical characteristics into four groups: type A (T-2 toxin group), type B (nivalenol group), type C (crotocin group) and type D (macrocyclic). DON is a member of group B, and is a polar organic compound with the chemical structure 12,13-epoxy-3 α ,7 α ,15-trihydroxytrichothec-9-en-8-on (see Figure A.1-1). It consists of a carbonyl group at the C8 position that distinguishes it from type A trichothecenes³⁰⁹. The other members of type B trichothecenes include nivalenol and DON's acetylated derivatives, 3-

acetyldeoxynivalenol, 15-acetyldeoxynivalenol and fusarenon X³¹⁰. DON is of particular importance due to its frequent occurrence of high levels of contamination³⁰⁹.

Figure A.1- 1: Structure of deoxynivalenol

Humans are chronically exposed to DON through ingestion of food grains including corn, barley, rye, oats and wheat that are contaminated in the field and during storage³¹¹. As a result, DON has been found worldwide in foods such as bread, noodles, beer and malt^{306, 312}.

Deoxynivalenol was first characterized in 1972 following its isolation from *Fusarium*-infected barley in Japan³¹³. It was rediscovered in 1973 from *Fusarium*-infected corn in the United States and given its colloquial name “vomitoxin” due to its emetic effects in swine³¹⁴. DON possesses a relatively low toxicity compared to other trichothecenes such as T-2 toxin. Though acute toxicity may rarely be an issue, the widespread contamination of DON is a problem of concern in public health. DON is responsible for a wide range of disorders in animals and humans. As with most mycotoxins, the dangers persist throughout the food-processing system as the toxin is not affected by cooking. The effects of DON include anorexic symptoms (such as nausea, abdominal pain, food refuse, vomiting diarrhea, weight loss)³¹⁵, as well as dizziness,

fever, headache³¹⁶, reduced immunity, anemia, growth retardation and reproductive disorders³¹⁷.

The toxic effects of DON stem mainly from its ability to inhibit the synthesis of DNA, RNA and proteins. This toxin binds to the large ribosomal subunit in eukaryotes therefore inhibiting the peptidyltransferase mediated translation process. This also triggers a ribotoxic stress response³⁰⁹ and activates intracellular protein kinases that both mediate selective gene expression and apoptosis³¹⁸. Additionally, DON is readily transported to the brain, where it interferes with dopamine receptors. This is thought to cause the anorexic-like symptoms³⁰⁶.

DON is one of the most common grain contaminants, contributing more than 90% of total trichothecene contamination³⁰⁶. As a result, levels of DON contamination are tightly regulated in several regions. Recently, JECFA updated its evaluation of DON. In particular, due to *in vivo* deacetylation of the acetylated derivatives of DON, which therefore contribute to total toxicity, the committee decided to convert the provisional maximum tolerable daily intake (PMTDI) to a group PMTDI of 1g/kg of body weight per day for DON and its acetyl derivatives³¹⁹.

To reduce the risk of exposure to DON, Canada developed standards for the maximum permissible levels in foods. Currently, the Canadian guideline for DON is 2 mg/kg in wheat for human consumption and 5 mg/kg for grain products used for cattle and poultry diets. Additionally, the Canadian National Feed Inspection Program analyzes approximately 300 samples annually for DON¹⁶⁰. The European Commission has set more stringent maximum permitted levels for DON in a variety of cereal-based foods. For example, 1750 and 1250 µg/kg were set as maximum levels for unprocessed durum

and common wheat; 750 µg/kg for wheat flour, bran, germ and pasta; 500 µg/kg for wheat products for direct human consumption and 200 µg/kg for wheat-based infant and baby food³¹⁵. Additionally, the U.S. Food and Drug Administration has set limits for DON in food and feed products. Food for human consumption such as bran, flour, and germ has a limit of 1 mg/kg DON³²⁰, grains and grain products intended for swine have a limit of 5 mg/kg and other grain feed has a limit of 10 mg/kg³²¹.

Current detection methods

DON is of particular concern as it is formed in the field prior to harvest and can continue to be produced during storage. Despite several production minimizing strategies, DON occurrence is unavoidable due to the major impact of weather conditions³⁰⁹. For example, the US Dried Distillers' Grains with Solubles (DDGS) survey that took place between August 2009 and January 2011 confirmed that two of the eight plants sampled contained DON levels that surpassed the 5ppm US Food and Drug Administration (FDA) advisory level. As a result, careful regular monitoring of this toxin is required.

There are two AOAC official methods for DON analysis: a TLC method and a GC method. The GC based method that uses electron capture detection is the most widely used. However, the procedure is complicated and requires the investment of both expensive equipment and a skilled technician³²². Therefore, several other techniques for DON detection are being developed and tested. For example HPLC, LC/MS or tandem mass spectrometry (LC/MS/MS) have been reported for DON analysis. These methods

are very sensitive and accurate but require sample clean-up procedures and technical skills. Other promising detection techniques that are being developed include Fourier-transform near infrared (FT-NIR) spectroscopy³¹⁵ and SPR³¹². As described earlier, immunoassays are ideal for routine analysis due to their simple clean-up procedure. Immunoassays that have been developed for DON include ELISA, lateral-flow immunoassays, radioimmunoassays and fluorescence polarization immunoassays³¹⁵. While effective, rapid and simple, these assays often display strong cross-reactivity with other trichothecenes, require long incubation times for complete antigen reaction and are expensive. As previously discussed, aptamer-based biosensors may function as useful alternatives in DON detection.

Section objectives

This chapter describes the progress towards the selection of an aptamer for deoxynivalenol. To-date, no aptamers have been selected for this target. The observed enrichment of the DNA library observed within the first 10 rounds of selection is presented.

A.2 Experimental

Materials

Phosphoramidites, modifiers, activator, deblock, capping, and oxidizing reagents were obtained from Glen Research. Standard support columns and acetonitrile were purchased from BioAutomation. Ultra High Purity 5.0 argon was purchased from

Praxair. 3-acetyldeoxynivalenol was a gift from Prof. Miller. All HPLC grade solvents were obtained from Caledon Labs. All PCR and electrophoresis components were purchased from BioShop Canada. 1,1'-carbonyldiimidazole (CDI) and chemicals for buffering agents and all other purposes were purchased from Sigma-Aldrich and prepared with Millipore Milli-Q deionized water at 18 MΩcm. CarboxyLink Coupling gel (immobilized diaminodipropylamine, DADPA, supplied as a 50% slurry) was purchased from Pierce (Illinois, USA). Spin-X cellulose acetate microcentrifuge filter tubes, Amicon-Ultra 0.5 mL 3kDa centrifuge units and sulfo-NHS acetate were purchased from Fisher Scientific Canada.

Deacetylation of DON

250.0 mg of 3-acetyldeoxynivalenol (3-ADON) was deacetylated to DON as described by Bily et al³²³. The 3-ADON was dissolved in 5 mL of HPLC grade methanol and mixed with 1 mL of 1 M NaOH in a round bottom flask for 2 hours with magnetic stirring. The methanol was removed *in vacuo*. The residue was dissolved in deionized water and added to a 60 mL separatory funnel and extracted with ethyl acetate (3 x 12 mL). The combined extracts were dried *in vacuo*.

TLC confirmed the production of DON (R_f = 0.2, EtOAc) from 3-ADON (R_f = 0.6, EtOAc). The entire product was redissolved in 2 mL of a 1:1v/v acetonitrile: water solution and purified using 80 µL injection volumes through preparative RP-HPLC on a Zorbax 300-Ext C18, 5 µM, 9.4 x 250mm, with the mobile phase program described in Table A.1-1 and a flow rate of 3 mL/minute.

Table A.1- 1: Reverse phase HPLC program for isolation of DON.

Time (minutes)	Acetonitrile (%)	Deionized water (%)
0	5	95
10	5	95
18	50	50
20	50	50
20.10	5	95
30	5	95

DON was detected by monitoring UV absorbance at 220nm. The collected peak was verified using a Biopure deoxynivalenol standard, purity $99.4 \pm 0.6\%$, from Romer labs.

DON conjugation

DON was conjugated to DADPA via the primary alcohol using CDI chemistry as described by Vatele³²⁴. 0.8 mg of CDI was mixed with 2.0 mg of DON in a 1 mL glass vial containing 20 μ L of reagent grade acetone in order to convert the primary alcohol into a reactive carbamate. The reaction was left to shake for 7 hours. 5 mL of DADPA slurry was washed with a mixture of 0.2 M NaHCO₃, pH 10 and acetone (v/v) beginning with 0% acetone, each time increasing the fraction of acetone by 5%, until the column was suspended in 5 mL of 70% 0.2 M NaHCO₃ and 30% acetone. The reactive carbamate was then added to the DADPA suspension and allowed to react for 3 days at room temperature. The slurry was then rinsed 5 times with 0.1 M NaHCO₃, pH 8.5 and reacted with 100 mg of sulfo-NHS acetate dissolved in 3 mL of 0.1 M NaHCO₃, pH 8.5 to quench unreacted amine groups. The mixture was incubated with shaking for another hour, followed by ten washes with DON SELEX buffer (DSB): 20 mM HEPES, 125

mM NaCl, 5 mM KCl, 2 mM MgCl₂, 1 mM CaCl₂, pH 7.6 . 5 mL of unmodified DADPA was also prepared for negative selection purposes by blocking amino groups with 100 mg sulfo-NHS as described above.

300 µL of both the DON-immobilized column and negative column were tested for non-specific DNA binding. 200pmol of fluorescein modified DNA (sequence FB₁ 32 from section 4.3.2) was incubated with each column material for 30 minutes and recovered by centrifugation at 14,000 g. The recovered DNA was quantified using UV-Visible spectrometry.

DNA library and primer synthesis

The starting DNA library was prepared by using standard phosphoramidite chemistry on a BioAutomation Mermade 6 (Plano, Texas) as described in section 2.2.2 using 200 nmol controlled pore glass (CPG) columns with the larger, 1000Å pore size. The sequence is 81 nucleotides in length, containing a central region of 45 random nucleotides flanked by two primer binding sites necessary for PCR and cloning: 5'-ATCCAGAGTGACGCAGCA-N45-TGGACACGGTGGCTTAGT -3'. The following primers used for amplification and cloning of the selected oligonucleotides were synthesized: Primer1: 5'- ATCCAGAGTGACGCAGCA -3' and Primer2: 5'-ACTAAGCCACCGTGTCCA -3'. Modified primers were also synthesized to allow for fluorescent labeling of the binding DNA and separation of strands on a denaturing polyacrylamide gel: ModPrimer1: 5'-6FAM- ATCCAGAGTGACGCAGCA -3' and ModPrimer2 5'-poly-dA₂₀-HEG- ACTAAGCCACCGTGTCCA -3'. HEG represents the

hexaethylene glycol Spacer Phosphoramidite 18 (Glen Research). Library DNA and primers were purified using denaturing polyacrylamide gel electrophoresis (12%) followed by clean-up with Amicon YM-3 Centrifugal Filter Devices as described in section 2.2.2.

DON SELEX experiments

0.5 mL of DON derivatized column matrix and negative column matrix were placed in new Spin-X cellulose acetate filter tubes and washed 5 times with 0.2 mL of DSB before each SELEX round. Prior to the first selection, the library (10 nmol) was suspended in 0.2 mL of DSB and heated for 5 minutes at 92°C then cooled at 4°C for 10 minutes, followed by incubation at room temperature for at 10 minutes. In round 2, the amount of library used for the selection was 1 nmol while 200pmol of library was used in subsequent rounds.

The library was first incubated with the negative column with mild shaking for 30 minutes. The DNA lacking affinity for the negative column material was collected by removing the filtrate following centrifugation at 1000 g for 3 minutes. This library, devoid of sequences displaying non-specific affinity, was subjected to the same heating and cooling process described above, followed by incubation with the DON column material for 30 minutes. Following incubation, sequences with little to no affinity for DON were removed with five washes with 200 µL of DSB. On the fourth wash, the column was transferred to a new Spin-X cellulose acetate filter tube. This was to ensure

that any sequences non-specifically interacting with the cellulose acetate filter or tube would not be eluted with the DON-specific DNA in the next steps.

In odd numbered rounds (i.e. 1, 3, 5, 7, 9), bound sequences were eluted by incubating the column for 10 minutes using 200 μ L of 6 M urea at 90°C. This was repeated an additional 3 times to ensure all bound species were removed from the column. The eluted fractions were pooled and purified using Amicon Ultra 3kDa 0.5 mL centrifugal filters, quantified using UV-visible spectroscopy and/or fluorescence and amplified using PCR. The same procedure was performed for any DNA bound to the negative column. For even numbered rounds (i.e. 2, 4, 6, 8, 10), bound sequences were eluted from the column with 3 washes of 200 μ L 2 mM DON in DSB with 2.5% DMSO. These fractions were directly quantified using UV-visible spectroscopy and/or fluorescence and amplified using PCR. In these even numbered rounds, no elution of DNA from the negative column was performed.

The entire selected oligonucleotide library was amplified in 15-30 parallel PCR reactions for the first round; in subsequent rounds only 90% of the selected library was amplified, with the remaining 10% set aside for analysis and in case of PCR failure. Each reaction consisted of 0.1 M Tris HCl pH 9, 50 mM KCl, 1% Triton X-100, 1.9 mM MgCl₂, 0.3 mM dNTP mix, 1 μ M each modified primer and 5 units of Taq DNA polymerase. The DNA was initially melted for 10 minutes at 94.0°C, followed by 25 cycles of 94.0°C (1 minute), 47.0°C (1 minute) and 72.0°C (1 minute). Final extension occurred at 72.0°C for 10 minutes after the last cycle. PCR products were dried down, heated at 55°C for 5 minutes in the presence of formamide and run on a 12% denaturing PAGE to separate the double stranded product. Details of the PAGE set-up can be found

in section 2.2.2. The 5' fluorescein-labeled DNA strands (the selected sequences) could be identified using an Alpha Imager UV-illuminator. The relevant DNA bands were cut from the gel, crushed and incubated in deionized water at 37°C overnight with gentle shaking or rotation. The DNA was then purified using the Amicon centrifugal filters, quantified and re-suspended in DSB and could be used in the next selection round.

Enrichment of the DNA library was assessed throughout the rounds of selection by monitoring the percent of DNA library binding to the DON derivatized column at each round by measuring DNA absorbance at 260 nm using a Cary 300 BioUV-visible spectrophotometer and measuring fluorescence (excitation 490nm, emission 520nm) with a Fluorolog Fluorescence Spectrophotometer (Horiba Jobin Yvon, USA) with a SpectrAcq controller. Where appropriate, the percent of library binding to the DON column was compared to the percent of library bound to the negative column. As one of the other mycotoxins, FB₁, displayed significant quenching of fluorescein, the effect of DON on fluorescein was tested by mixing a 100 µL solution of 2 mM DON with 200pmol and 2pmol of fluorescein modified DNA (sequence FB₁ 32 from section 4.3.2). The fluorescence intensity of these samples without 2 mM DON and in the presence of 2 mM DON were compared (excitation: 490nm, emission: 520nm). As no significant quenching effects could be detected, fluorescence was used for monitoring the library enrichment.

Table A.1- 2: Changes to the SELEX procedure and the round in which these changes were incorporated.

Selection Round	Change to reported procedure
3	Increased number of washes to 7, from 5, to remove non-specific sequences
6	Increased number of washes to 8, from 7, to additionally remove weak-binding sequences
7	Eluted 5 times with 6 M urea to ensure full removal of high affinity sequences
8	Reduce the volume of DON column from 500 μ L to 250 μ L to select only the highest affinity sequences
11	Eluted using 5 mM DON, instead of 2mM, to ensure removal of bound sequences from immobilized DON

A.3 Results and Discussion

Preparation of DON derivatized column

Aptamers are able to discriminate between small structure differences^{59, 325}. There have been cases where aptamers have been selected for an undesired target⁹⁰. As a result, when performing SELEX, it is critical that the target for selection be as pure as possible. Analytical standards of DON are considered the purest samples available commercially. However, a complete purity assessment revealed that the products from both Biopure Romer Labs and Sigma-Aldrich are only of > 98% and >96% purity respectively³²⁶. However, pure crystalline 3-ADON (fine colourless needles), isolated from *Fusarium culmorum*³²⁷ and purified by chromatography and recrystallization³²⁸, were received as a gift from Prof. Miller. As DON is much more prevalent than its acetylated derivatives, an aptamer for DON would be the most useful. As a result, 3-ADON was deacetylated using a strong base, NaOH. Deacetylation was confirmed using

TLC, by slower migration in ethyl acetate. Further characterization of the desired product was performed using RP-HPLC with UV detection. By comparison with the Biopure analytical standard, the product was eluted at a retention time of 14.5 minutes in the mobile phase described in Table A.1-1. Complete purification of the deacetylated product was performed over 25 runs of preparative HPLC, collecting from 13.5 to 15.5 minutes, resulting in a high purity sample. Due to the rigorous purification, the final yield from 3-ADON was 14.6%. Regardless, the resulting 32 mg was sufficient for preparing the column matrix and performing several elution steps with high concentrations of DON.

DON was immobilized to DAPDA column matrix using standard carbamate chemistry. In the typical SELEX experiments, a significant amount of effort is focused on confirming successful coupling of target to the column matrix. Unfortunately, due to the small amount of DON available, this was not possible. However, it is likely that DON was able to couple to the column material. All the reagents were newly purchased and the coupling chemistry used is efficient. Additionally, as the coupling chemistry used for column derivatization is generally well established, it is rare for researchers to invest time to verify successful coupling^{54, 90, 102}.

Our initial experiments were performed on unblocked DON modified column. Within the first round, more than 90% of the DNA was non-specifically interacting with the column. Several elution steps using high concentrations of urea were required to remove the non-specific binding sequences. As a consequence of using DADPA column, the unreacted amines bear a positive charge in neutral pH conditions. Therefore, the negatively charged backbone of DNA binds to the column through electrostatic

interactions. To mitigate this, a selection buffer with higher salt concentrations (to block the positive charge) and higher pH conditions (to deprotonate the amine) was used. While this improved the non-specific interaction, buffers containing 5 M NaCl or with a pH of 12 were required to reduce the non-specific binding to less than 20%. Instead, organically reacting the remaining amine groups using NHS acetate resulted in less than 16% non-specific binding, which was considered acceptable for selection purposes¹⁰.

Aptamer selection progress

As with the SELEX experiments described in Chapter 4 and 5, careful consideration was made when choosing the selection buffer for this experiment. Since the DNA backbone is polyanionic, the negative–negative charge repulsion may inhibit formation of complex structures in the absence of counterions such as Na^+ . Divalent cations in particular, such as Mg^{2+} , block the negative charge repulsion and allow the backbone to form tighter folding. As aptamers displaying less magnesium dependence are of typically higher affinity¹³¹, only a moderate concentration of Mg^{2+} was included. Calcium was also included due to the increase in binding observed of the OTA aptamer upon the addition of calcium ions⁵⁰. Finally, as many published aptamers fold into G-quadruplex structures²²⁷, potassium was included to help stabilize G-rich sequences into G-quadruplex structures.

While aptamers have been selected using a variety of random region lengths within the initial library, a 45 nucleotide length was chosen for this selection. As discussed in Chapter 4 and 5 this was chosen to strike a balance between two competing

factors: shorter sequences are easier to synthesize and handle, while longer sequences allow for a greater structural complexity and possibly new functional motifs¹⁰⁴. Given the number of selections being performed in our research laboratory, a 45nt random region was also chosen to give a significant size difference (10 or more nucleotides) between other SELEX libraries. As a result, should cross-contamination between SELEX experiments occur, isolation of the library could be performed easily using electrophoretic gel separation.

As evidenced in Chapter 4 and 5, it is extremely important to ensure that any sequences displaying binding properties towards the selection matrix are eliminated if SELEX is being performed using a solid support matrix for target immobilization. As aptamers are used to measure free target in solution, affinity to the matrix would reduce the functionality of the aptamer in many applications. To mitigate this, it is necessary to remove all sequences that display even weak binding or partial recognition to the matrix. Typical SELEX experiments alleviate this using a pre-selection step prior to the first selection round to remove any initial sequences displaying this unwanted binding⁴⁷. However, due to mutation and evolution of the library, as well as the large ratio of available matrix compared to target for aptamer binding, performing only one pre-selection step may be insufficient. Elution with high concentrations of target, in this case DON, will however, help ensure that the aptamers function well in solution. As the coupling efficiency could not be determined, there is at least some assurance that the sequence eluted DON will have affinity for the target. However, for aptamers that have very slow k_{off} values, long incubation times are required with free target for complete recovery. While there are some expensive SELEX methods that have been developed

that allow monitoring of the k_{off} and k_{on} rates such as the SPR based flow-cell SELEX by Dr. Larry Gold⁷⁹, the column based selection does not provide this information²⁵⁵. As a result, it is unknown whether the highest affinity aptamers have been eluted and recovered or not. This, as well as the small amount of pure DON available resulted in the strategy of performing alternative elution techniques. In odd numbered rounds, rigorous elutions with urea at 90°C were performed. These selection rounds would remove all sequences displaying high affinity for the target or the column matrix. To qualify that the sequences have a preferred affinity for the DON over the column matrix, the percent of eluted sequences from the positive column was monitored and compared to the percent of eluted sequences from the negative column. Additionally, even numbered rounds involved eluting with a high concentration of DON. Here, only sequences that bind to DON in solution would be removed.

The results of the first 11 selection rounds can be found in Figure A.1-2. By monitoring the enrichment throughout each round, changes in the selection protocol were made to increase the stringency. For example, the percent of bound DNA enhanced by a factor of 17 in round 2 compared to the first round. Therefore, in round 3, the number of washes was increased from 5 to 7 to attempt to ensure all non-specifically bound sequences were removed. While the subsequent rounds did not show an increased enrichment compared to round 2, the comparison with the percent of DNA bound to the negative column in round 3 (4:1) and 5 (7:1) indicated that the DNA was preferentially binding to DON (see Figure A.1-2). Consequently, the number of washes was increased to 8 in round 6, with an additional urea elution step in round 7 to ensure even the highest affinity aptamers were collected. In round 8, the column volume was reduced by 2 to

Figure A.1-2: Percent recovery of binding ssDNA to DON immobilized column after each selection round. With any observable significant increase in percent recovery, more stringent selection conditions were implemented the following round (see experimental and Table A.1-2).

select the aptamers with the highest residency times. Unfortunately, the percent of DNA eluted was significantly lower in this round (0.1%). This drop in percent binding could not be solely attributed to the reduction in target available for binding. It is possible that the increased stringency imposed was too vigorous, causing a loss of putative aptamer sequences. Alternatively, it is possible that the 2 mM DON was insufficient to elute the majority of the enriched aptamer library. To alleviate this problem, the stringency remained unchanged in the following 3 rounds. Unfortunately, the percent of DNA that

bound to the DON column in round 9 was comparable to the concentration of DNA that bound to the negative column. This further indicated that high affinity binders had been removed from the current selection library. As a result, the round 11 elution was performed with DON rather than urea to ensure the sequences displayed binding to the target. A higher concentration of DON, 5 mM was used for elution, over longer incubation times, at an attempt to ensure complete removal of aptamer sequences. The lack of enrichment observed with this round provided the final evidence that the selection conditions were too stringent.

It is possible that the current library may have several high affinity aptamers for the target. To test for this, cloning and sequencing as well as laborious testing would be required. However, the time and effort for this without some indication that the results may be promising, is not worthwhile. Typically, cloning and sequencing steps are employed once the library is significantly enriched and more than 50% of the library binds to the target¹⁰². Instead, selections will be re-started beginning after round 3. A completely new selection is not necessary as it is clear that the round 3 selected library contains DON binding sequences. With the new selection, the stringency will be incorporated much more carefully. In particular, the approach used by Dr. Larry Gold in his selection of almost 1000 protein-binding aptamers will be employed. For consistency purposes, Gold strategically reduced the target concentration based on the ratio of oligonucleotides interacting with the target-immobilized support matrix (T) compared to the negative support matrix (N) (ratio = T/N) according to this equation⁵¹.

$$\text{If } \frac{T}{N} < 10, [T]_{(i+1)} = [T]_i$$

$$\text{If } 10 < \frac{T}{N} < 100, [T]_{(i+1)} = \frac{[T]_i}{3.2}$$

$$\text{If } \frac{T}{N} > 100, [T]_{(i+1)} = \frac{[T]_i}{10}$$

Where, [T] = target concentration and i = current round number. In other words, if the fraction of DNA binding to the target compared to the negative column is less than 10, the target concentration will remain constant in the next round. Should the fraction of DNA binding to the target compared to the negative column be between 10 and 100 times, the target concentration will be reduced by 3.2 in the subsequent round. Finally, if there is a 100 times difference in the amount of DNA binding to the target compared to the negative column, the target concentration will be reduced by 10 in the following round.

In addition to this strategic stringency approach, counter-selections with similar mycotoxins may be employed. However, these negative selection steps will not be performed with the acetylated derivatives of DON, for several reasons. First, the difference between these structures is very subtle (addition of an acetyl functional group on the 3-OH or 15-OH). While aptamers have been selected that can recognize such minute differences, these counter-selection would likely severely reduce the DNA library and eliminate many high affinity binders. The loss of high affinity binders in order to gain selectivity is not worthwhile in this case as JECFA recommends a group provisional maximum tolerable daily intake (PMTDI) for DON and its acetyl derivatives³¹⁹.

A.4 Conclusions

We have developed a strategy to perform selections on the target deoxynivalenol. While there is interest in developing an aptamer for this target and several groups have worked towards its selection, none have been reported to-date³²⁹. The monitoring of our selection progress has demonstrated that enrichment of the initial library was achieved. However, over-zealous stringency steps may have resulted in loss of the potentially high affinity aptamers for DON. As a result, new selections will be performed using the partially enriched library from round 3. As noted, the structure of DON is has made it a challenging target for the isolation of high affinity aptamers, but the progress made is promising. By employing more carefully planned selection strategies, we expect to isolate novel binding sequences for DON.

Appendix B Binding isotherms

B.1 Ochratoxin A aptamer 1.12.2 (Chapter 2)

Figure B.1-1: Binding isotherm for aptamer 1.12.2 with OTA obtained using affinity chromatography. DNA aptamer concentration was varied with a constant amount of OTA immobilized column (1 mg). Bound aptamer was eluted by washing with 95°C buffer and measured using total fluorescence. A theoretical curve was fit using Microsoft Excel Solver. K_d was determined to be $97 \pm 3 \mu\text{M}$. Inset: Plot of the experimentally determined DNA fluorescence vs. the theoretical DNA fluorescence.

Figure B.1-2: Binding isotherm for aptamer 1.12.2 with OTA obtained using fluorescence polarization. DNA aptamer concentration was varied with a constant concentration of OTA (200 nM). A theoretical curve was fit using Microsoft Excel Solver. K_d was determined to be $6.1 \pm 0.2 \mu\text{M}$. Inset: Plot of the experimentally determined fluorescence polarization vs. the theoretical fluorescence polarization.

Figure B.1-3: Binding isotherm for aptamer 1.12.2 with OTA obtained using equilibrium filtration. DNA aptamer concentration was varied with a constant amount of OTA (1 μ M). Retained OTA was measured using total fluorescence. A theoretical curve was fit using Microsoft Excel Solver. K_d was determined to be 30 ± 1 μ M. Inset: Plot of the experimentally determined OTA fluorescence vs. the theoretical OTA fluorescence.

B.2 Ochratoxin A aptamer A08 (Chapter 3)

Figure B.2-1: Binding isotherm for aptamer A08 with OTA obtained using affinity chromatography. DNA aptamer concentration was varied with a constant amount of OTA immobilized column (1 mg). Bound aptamer was eluted by washing with 95°C buffer and measured using total fluorescence. A theoretical curve was fit using Microsoft Excel Solver. K_d was in the mM range. Inset: Plot of the experimentally determined DNA fluorescence vs. the theoretical DNA fluorescence.

Figure B.2-2: Binding isotherm for aptamer A08 and 1.12.2 with OTA obtained using the AuNP assay. OTA concentration was varied with a constant concentration of aptamer. An apparent K_d was elucidated from the aggregation of the AuNP, measured as the increased absorbance at 705nm. A theoretical curve was fit using Microsoft Excel Solver. K_d was determined to be $1.4 \pm 0.4 \mu\text{M}$ for aptamer A08 and $6 \pm 3 \mu\text{M}$ for aptamer 1.12.2. Insets: Plots of the experimentally determined absorbance vs. the theoretical absorbance.

B.3 Fumonisin B₁ aptamers (Chapter 4)

Figure B.3-1: Binding isotherm for aptamer FB₁ 14 with FB₁ obtained using affinity chromatography. Aptamer concentration was varied using a constant number of FB₁ immobilized magnetic beads (5×10^6). Bound aptamer was eluted by washing with 90°C column buffer and measured using fluorescence. The theoretical curve was fit using Microsoft Excel Solver. K_d was determined to be 920 ± 40 nM.

Figure B.3-2: Binding isotherm for aptamer FB₁ 16 with FB1 obtained using affinity chromatography. Aptamer concentration was varied using a constant number of FB₁ immobilized magnetic beads (5×10^8). Bound aptamer was eluted by washing with 90°C column buffer and measured using fluorescence. The theoretical curve was fit using Microsoft Excel Solver. K_d was determined to be 200 ± 100 nM.

Figure B.3-3: Binding isotherm for aptamer FB₁ 23 with FB1 obtained using affinity chromatography. Aptamer concentration was varied using a constant number of FB₁ immobilized magnetic beads (5×10^8). Bound aptamer was eluted by washing with 90°C column buffer and measured using fluorescence. The theoretical curve was fit using Microsoft Excel Solver. K_d was determined to be 1000 ± 20 nM.

Figure B.3-4: Binding isotherm for aptamer FB₁ 31 with FB1 obtained using affinity chromatography. Aptamer concentration was varied using a constant number of FB₁ immobilized magnetic beads (5×10^8). Bound aptamer was eluted by washing with 90°C column buffer and measured using fluorescence. The theoretical curve was fit using Microsoft Excel Solver. K_d was determined to be 470 ± 60 nM.

Figure B.3-5: Binding isotherm for aptamer FB₁ 32 with FB1 obtained using affinity chromatography. Aptamer concentration was varied using a constant number of FB₁ immobilized magnetic beads (5×10^8). Bound aptamer was eluted by washing with 90°C column buffer and measured using fluorescence. The theoretical curve was fit using Microsoft Excel Solver. K_d was determined to be 400 ± 100 nM.

Figure B.3-6: Binding isotherm for aptamer FB₁ 39 with FB1 obtained using affinity chromatography. Aptamer concentration was varied using a constant number of FB₁ immobilized magnetic beads (5×10^8). Bound aptamer was eluted by washing with 90°C column buffer and measured using fluorescence. The theoretical curve was fit using Microsoft Excel Solver. K_d was determined to be 100 ± 30 nM.

B.4 Homocysteine aptamers (Chapter 5)

Figure B.4-1: Binding isotherms for L-Hcy putative aptamers sequences (and control sequence) obtained using sepharose based affinity chromatography. Aptamer concentration was varied using 1 mg of L-Hcy immobilized column. Bound aptamer was eluted by washing with 90°C column buffer and measured using fluorescence. The theoretical curve was fit using Microsoft Excel Solver.

Figure B.4-2: Binding isotherms for L-Hcy putative aptamers sequences (and control sequence) obtained using magnetic beads based affinity chromatography. Aptamer concentration was varied using a constant number of L-Hcy immobilized magnetic beads (5×10^5). Bound aptamer was eluted by washing with 90°C column buffer and measured using fluorescence. The theoretical curve was fit using Microsoft Excel Solver.

Figure B.4-3: Selectivity tests: Binding isotherms for Hcy8 aptamer sequence obtained using magnetic beads based affinity chromatography. Aptamer concentration was varied using a constant number of cysteine or methionine immobilized magnetic beads (5×10^8). Bound aptamer was eluted by washing with 90°C column buffer and measured using fluorescence. The theoretical curve was fit using Microsoft Excel Solver.

B.5 ATP aptamers (Chapter 6)

Figure B.5-1: Binding isotherm for aptamer ATPF4 with ATP agarose obtained affinity chromatography. DNA aptamer concentration was varied and incubated with 1 mg of column. A theoretical curve was fit using Microsoft Excel Solver. K_d was determined to be $24.4 \pm 0.8 \mu\text{M}$. Inset: Plot of the experimentally determined fluorescence vs. the theoretical fluorescence.

Figure B.5-2: Binding isotherm for aptamer ATPF8 with ATP agarose obtained affinity chromatography. DNA aptamer concentration was varied and incubated with 1 mg of column. A theoretical curve was fit using Microsoft Excel Solver. K_d was determined to be $3.8 \pm 0.1 \mu\text{M}$. Inset: Plot of the experimentally determined fluorescence vs. the theoretical fluorescence.

Appendix C Algorithms (Chapter 6)

C.1 The *CountJunction* algorithm for determining the number of junctions in a nucleic acid secondary structure

From the ViennaRNA secondary structure, the *CountJunctions* algorithm determines the number of junctions. The *CountJunctions* algorithm traverses the sequence from 5' to 3'. For each stem (at least four contiguous base pairs), there are two strands. Starting from one strand, if there is a stem along the path to another strand, the junction number is increased by 1. For example, in Figure 6-1, stem S_1 has two strands, S_1' and S_1'' . Along the path from S_1' to S_1'' in the 5' to 3' direction, each stem S_2 , S_3 , S_4 and S_5 forms one junction, so the number of junctions for stem S_1 is 5. For each stem, the number of junctions is at least 1. The number of junctions for the RNA secondary structure is the maximum number of junctions of all stems. For example, if a secondary structure contains a one-way and a three-way junction, the secondary structure is defined as a three-way junction structure.

Algorithm *CountJunctions*:

1. Scan the nucleotide pairing list created by ViennaRNA and determine the list $[S_1, \dots, S_n]$ of all *stems*. A stem is defined as a maximal list of at least four contiguous pairs.
2. Consider the list L of nucleotides from the 5' end to the 3' end. Each stem S consists of two contiguous subsequences S' and S'' of L . Sort the list $L' = [S_1', S_1'', S_2', S_2'', \dots, S_n', S_n'']$ according to their location in L .
3. For each stem, calculate its number of junctions. A k -way junction consists of a sequence $S_1', S_2', S_2'', S_3', S_3'', \dots, S_k', S_k'', S_1''$.
4. Find the maximum number of junctions.

In Step 1 of Algorithm *CountJunctions*, a stem is defined as a maximal list of at least four contiguous pairs. This minimum stem length was chosen in accordance with Bindewald

et al.³³⁰ Selecting a smaller minimum stem length results in a larger number of predicted high complexity structures.

Appendix D Supplemental information for the Aptamer Base (Chapter 7)

D.1 Curating the database

Create an account

Go to freebase.com. On the top right corner, click the “Sign In or Sign Up” button. Once you have created a Freebase account, join the Aptamer Base at <http://aptamer.freebase.com/> and click the “Join Base” button. In the future, simply sign in anytime you wish to contribute to the database.

Collect information

Choose one publication that describes a SELEX experiment that results in aptamers that have been verified to bind to a target with a measured dissociation constant (K_d). Only articles that describe SELEX experiments, which result in aptamers with corresponding sequences and K_d values will be entered.

The following information must be extracted from the publication:

Data		Description
General		
Document identifier		The PubMed ID (PMID) or digital object identifier (DOI) if PMID is not available
SELEX		
Method		Named method used for selection (likely “Generic SELEX if not specified)
Partitioning method		Method used to remove the non-specific sequences from the initial library
Recovery method		Method used to recover the putative aptamers from the target (if required)
Number of rounds		Number of selection rounds required to obtain the aptamers
Template sequence		The random library used for selection including the 5' and 3' primer binding sites
Selection solution	Buffering agent	Weak acid or base used to maintain the pH of the solution in which SELEX proceeds (generally Tris, phosphate or HEPES)
	pH	Measure of the acidity or basicity of the SELEX solution (typically neutral pH)
	Temperature	Temperature of the selection (usually 37°C or room temperature = 25°C)
	Metal cation concentration	Concentration of any positively charged metals in the SELEX solution (mainly Na, K, Ca, Mg)
Aptamers		
Aptamer type		Either DNA or RNA. Aptamers may additionally be modified at the base, sugar or 5' and 3' ends.
Target		Full name of the target for which the aptamers were selected
Number of aptamers		Only included the total number of aptamers that have measured K_d values (this includes any minimal aptamers of the original full-length clone).
Aptamer sequences		Sequence 5' to 3'; note the total length and any observed patterns
Experimentally determined structure		If any experiments were performed to confirm a predicted structure, indicate the structure. (Using predicting tools such as Mfold are insufficient)
Mutational analysis		Was the aptamer sequence mutated to determine which positions were critical for binding?
Application		Were any of the aptamers applied in a therapeutic or

		detection system?
Affinity experiment		
Method		The method used to measuring the binding affinity and eventually elucidate a K_d value
Binding solution	Buffering agent	Weak acid or base used to maintain the pH of the solution in which the affinity experiment was performed (generally Tris, phosphate or HEPES)
	pH	Measure of the acidity or basicity of the binding solution (typically neutral pH)
	Temperature	Temperature of the binding solution (usually 37°C or room temperature = 25°C)
	Metal cation concentration	Concentration of any positively charged metals in the binding solution (mainly Na, K, Ca, Mg)
Dissociation constants		The values obtained from the affinity experiment for each aptamer. Include the range or error if described.

Create topic names

Each entry into Freebase has a unique name. To create these names, make use of the topic creator found here: <http://134.117.108.159/aptamer/>

Aptamer Information

Document Identifier *:

Aptamer Molecule Type *: DNA

Target Name *:

Number of Interactions *:

Number of Aptamers *:

Number of Minimal Aptamers:

Number of Dissociation Constants:

Number of Affinity Experiments:

Add/ Remove Selection Solutions

Concentration

Cation

[* indicates a required field]

In this form, enter the document identifier and the aptamer type. For the target name, a short-form is preferred. The number of interactions represents the total number of aptamers that have unique K_d values. The number of aptamers should correspond to the total number of *full-length* aptamer clones that have any corresponding K_d value. This distinction will be important later. The number of minimal aptamers is the total number of aptamers that are derived from a full-length aptamer clone. This may be a truncated sequence, an extended sequence, a mutated sequence or a modified aptamer. The number of dissociation constants should correspond to the number of different K_d values reported in the publication. The number of affinity experiments corresponds to the number of affinity experiments performed in which either the conditions or the method changes. Therefore, if one method is performed 7 times to measure the K_d value of 7 different aptamers, but the conditions remain the same, only 1 affinity experiment exists. If the binding solution conditions are changed or the method used is changed, indicate the total number of different affinity experiments that have been performed (even if for the same aptamer). Finally, include the total concentration (in mol/L) and corresponding symbol of each metal cation present in the buffer used in the SELEX and affinity experiments. (For example if 100 mM NaCl and 40 mM NaOH was used, indicated 0.14 and Na). More rows can be added to the form to include the many different metal cations present in the buffer.

Entering information into the database – general

There are several different pieces of information that must be entered into the database.

Topics: First, for any piece of data, a topic must be first created. In general, this is done by using an “Add it!” link or by simply entering one of the created topic names from above. If the topic (or similar topic) exists, a list of the currently available topics will be presented by Freebase. If the intent is to create a completely new topic, do not choose any of these suggestions, simply select “Create new.” Once a topic has been created it can be edited by clicking the “edit this topic” button.

Types: Once a topic has been created, certain types (categories) may be automatically associated with it. Other times, we must specifically indicate that a topic is of a given type (category). For example, if we create an aptamer, we must indicate that it is either of the type DNA or RNA. To categorize a given topic, use the “+Add a type” button. Fill out the box with the type you would like to add (ensure it is the correct type as there are often several similar possibilities).

Properties: Once a topic has been typed, Freebase links several properties to this topic. Typically, we are required to add these properties. Depending on the property, you may be required to enter text, enter a value or choose from a drop-down menu. To ensure that the curator has not forgot to complete an entry, if a publication has not provided a given detail the following will be used: n/a for text; if -1 for a value; if not-described or not required if a drop-down menu.

Entering a SELEX article into the database

The following steps are a suggestion. Any order can be used to fill the database with information extracted from a publication as long as the entry is complete

- 1) Go to http://www.freebase.com/view/base/aptamer/aptamer_experiment and at the bottom of the page, click on “Add it!”. Enter the SELEX experiment, partitioning method and recovery method from the topic creator. Ensure that you click on “Create New” each time to add these new topics. Additionally, using the drop-down menu, chose the particular SELEX method.
- 2) Scroll back to the top of the SELEX experiment page and select the SELEX experiment that you have just created. Once on your SELEX experiment page, chose “Edit this topic” on the right.
- 3) Using the “+Add type” button, add the type “SELEX experiment”, this should generate a list of properties that you now must fill out.
- 4) Add the PMID by clicking on “edit” and entering the appropriate number.
- 5) Add all the interactions from the topic creator into the “has outcome” box. Again, each time ensure that you select “Create new” to add these interactions.
- 6) Add the SELEX conditions topic from the topic creator to the “has experimental conditions.”
- 7) Select the partitioning method for this experiment and edit the topic. Using the drop-down menu, select the appropriate method. Return to the SELEX experiment.
- 8) Select the recovery method for this experiment and edit the topic. Using the drop-down menu, select the appropriate method. Return to the SELEX experiment.
- 9) Select the SELEX conditions for this experiment and edit the topic. Using the, “+Add type” button, type this topic as “SELEX conditions”. You will see several new properties appear that we now must fill out.

- 10) Enter the number of selection rounds required to obtain the aptamers (number).
- 11) Enter the library template from 5' to 3' (do not include those numbers in the sequence). For the random region, use the symbol N followed by the total number of random nucleotide positions, surrounded with dashes. For example GAATTCAGTCGGACAGCG-N₄₀-GATGGACGAATATCGTCTCCC.
- 12) Enter the template variable region summation (number). This is the total number of random positions (N) throughout the entire template sequence. For example, if the template is GAATT-N10-TCGGACAGCG-N15-GGA-N5-CGAATATCGTCTCCC, the template variable region summation is 30. If two separate pools are used in the selection, include the total number (for example GGGAGGACGATGCGG-N30-CAGACGACTCGCCCGA and GGGAGGACGATGCGG-N₄₀-CAGACGACTCGCCCGA, the total would be 70).
- 13) Indicate if there is template bias (yes or no in the drop-down menu). This means that instead of using an equal proportion of A,C,G,T bases for N, the researchers may have used excess of A and T to increase the likelihood of secondary structures, or perhaps a higher concentration of G to increase the likelihood of obtaining G-quadruplex structures.
- 14) Add the selection solution topic from the topic creator to the “has selection solution” box. Click on this entry and edit is as usual.
- 15) For the selection solution, select the buffer (from a drop-down menu), add the pH and temperature (values) and add the metal cation concentrations as outputted in the topic creator.
- 16) Return to the SELEX experiment.

- 17) Select one of the interactions for this experiment and edit the topic.
- 18) This interaction must be typed as either a “collective interaction” or “pairwise interaction”. A collective interaction is one in which several different aptamers have the same dissociation constant. All other cases are “pairwise interactions”.
- 19) Add the aptamer topics from the topic creator to the “has participants” box.
- 20) Add the full target name to the “has participants” box. If an appropriate entry already exists, use this. (The entry may already include useful information about the target).
- 21) Add the affinity experiment topic from the topic creator to the “in confirmed by” box.
- 22) Add the dissociation constant topic from the topic creator to the “has dissociation constant” box.
- 23) Select the target topic that you created in step 19, edit the topic and type it as an “aptamer target”. Under the “aptamer target” property section, add all the categories to which the aptamer target belongs (for example small molecule and toxin). If the target is a small molecule, add the CHEBI ID from <http://www.ebi.ac.uk/chebi/>.
- 24) Return to the interaction (step 16). Select one of the aptamers and edit the topic and type this as an “aptamer” and “RNA or “DNA” as appropriate.
- 25) Fill out the boxes as follows
 - Application: text
 - Has minimal aptamer: add the topic from the topic creator if appropriate
 - Has mutational analysis: yes or no drop-down menu

Has experimentally determined structure: yes or no drop-down menu

Secondary structure: select from the drop-down menu (only if the answer above is yes)

Sequence: complete the full sequence from 5' to 3' (using letters)

Sequence length: calculate the total number of positions (number)

Sequence pattern: describe any pattern using text

- 26) Return to the interaction or SELEX experiment and select another one of the aptamers and perform steps 23-24.
- 27) Return to the interaction (step 16), select the affinity experiment, edit the topic.
Add the PMID and select the affinity method (from the drop-down menu).
- 28) From the topic creator, add all the dissociation constants that were determined with this affinity experiment. Add this to the “has outcome” box. Since we already added these topics to Freebase in step 21, Freebase should recognize that they exist and provide them as examples. Ensure that you select these and DO NOT “create new” as we have previously done.
- 29) Add the affinity conditions topic from the topic creator to the “has experimental conditions” box. Select this, edit it, and add the type “affinity conditions”.
- 30) Add the binding solution topic from the topic creator to the “has binding solution” box. Select and edit this. Add the appropriate information, similar to the method used in step 14.
- 31) Return to the affinity experiment in step 26, select one of the dissociation constants and edit the topic. Here, include the dissociation constant value in mol/L. If a range is given, include the range. If error is provided, include those

values as well. If Freebase recognizes the value as 0, use the “temporary string value” to enter the number. Freebase recognizes very small numbers, having only 1 significant figure as being equal to zero.

32) Continue step 30 for each dissociation constant.

33)

D.2 Distribution of DNA and RNA K_d values found in the Aptamer Base

Figure D.2-1: Scatter of aptamer dissociation constant vs. aptamer length for DNA aptamers (blue) and RNA aptamers (red).

References

- (1) Deigan, K. E.; Ferre-D'Amare, A. R. *Acc. Chem. Res.* **2011**, *44*, 1329-1338.
- (2) Weigand, J. E.; Suess, B. *Appl. Microbiol. Biotechnol.* **2009**, *85*, 229-236.
- (3) Tuerk, C.; Gold, L. *Science* **1990**, *249*, 505-510.
- (4) Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818-822.
- (5) Robertson, D. L.; Joyce, G. F. *Nature* **1990**, *344*, 467-468.
- (6) Mascini, M. In *Aptamers in bioanalysis*; John Wiley & Sons, Inc: Hoboken, New Jersey, USA, 2009; .
- (7) McKeague, M.; Bradley, C. R.; De Girolamo, A.; Visconti, A.; Miller, J. D.; DeRosa, M. C. *Int. J. Mol. Sci.* **2010**, *11*, 4864-4881.
- (8) Jayasena, S. D. *Clin. Chem.* **1999**, *45*, 1628-1650.
- (9) Nimjee, S. M.; Rusconi, C. P.; Sullenger, B. A. *Annu. Rev. Med.* **2005**, *56*, 555-583.
- (10) Jhaveri, S.; Ellington, A. *Curr. Protoc. Nucleic Acid Chem.* **2002**, *Chapter 9*, Unit 9.5.
- (11) Bardoczy, V.; Meszaros, T. *Periodica Polytechnica Abstracts of PhD Conference* **2006**, .
- (12) Balogh, Z.; Lautner, G.; Bardoczy, V.; Komorowska, B.; Gyurcsanyi, R. E.; Meszaros, T. *FASEB J.* **2010**, *24*, 4187-4195.
- (13) Sefah, K.; Shanguan, D.; Xiong, X.; O'Donoghue, M. B.; Tan, W. *Nat. Protoc.* **2010**, *5*, 1169-1185.
- (14) Beaucage, S. L.; Caruthers, M. H. *Tetrahedron Lett.* **1981**, *22*, 1859-1862.
- (15) Beaucage, S. L.; Iyer, R. P. *Tetrahedron* **1992**, *48*, 2223-2311.
- (16) Keefe, A. D.; Pai, S.; Ellington, A. *Nat. Rev. Drug Discov.* **2010**, *9*, 537-550.
- (17) Beaucage, S. L.; Iyer, R. P. *Tetrahedron* **1993**, *49*, 6123-6194.
- (18) Beaucage, S. L.; Iyer, R. P. *Tetrahedron* **1993**, *49*, 1925-1963.

- (19) Xu, Z. X.; Gao, H. J.; Zhang, L. M.; Chen, X. Q.; Qiao, X. G. *J. Food Sci.* **2011**, *76*, R69-75.
- (20) Tse Sum Bui, B.; Haupt, K. *Anal. Bioanal Chem.* **2010**, *398*, 2481-2492.
- (21) Baird, G. S. *Am. J. Clin. Pathol.* **2010**, *134*, 529-531.
- (22) Sassolas, A.; Blum, L. J.; Leca-Bouvier, B. D. *Analyst* **2011**, *136*, 257-274.
- (23) Deisingh, A. K. *Handb. Exp. Pharmacol.* **2006**, (173), 341-357.
- (24) Hermann, T.; Patel, D. J. *Science* **2000**, *287*, 820-825.
- (25) Cho, E. J.; Lee, J. W.; Ellington, A. D. *Annu. Rev. Anal. Chem. (Palo Alto Calif)* **2009**, *2*, 241-264.
- (26) Johansson, M. K. *Methods Mol. Biol.* **2006**, *335*, 17-29.
- (27) Jin, Y.; Bai, J.; Li, H. *Analyst* **2010**, *135*, 1731-1735.
- (28) Xu, H.; Mao, X.; Zeng, Q.; Wang, S.; Kawde, A. N.; Liu, G. *Anal. Chem.* **2009**, *81*, 669-675.
- (29) Xiao, Y.; Lai, R. Y.; Plaxco, K. W. *Nat. Protoc.* **2007**, *2*, 2875-2880.
- (30) Xiao, Y.; Piorek, B. D.; Plaxco, K. W.; Heeger, A. J. *J. Am. Chem. Soc.* **2005**, *127*, 17990-17991.
- (31) Radi, A. E.; O'Sullivan, C. K. *Chem. Commun. (Camb)* **2006**, (32), 3432-3434.
- (32) Wang, X.; Zhou, J.; Yun, W.; Xiao, S.; Chang, Z.; He, P.; Fang, Y. *Anal. Chim. Acta* **2007**, *598*, 242-248.
- (33) Zhang, J.; Zhang, B.; Wu, Y.; Jia, S.; Fan, T.; Zhang, Z.; Zhang, C. *Analyst* **2010**, *135*, 2706-2710.
- (34) Lautner, G.; Balogh, Z.; Bardoczy, V.; Meszaros, T.; Gyurcsanyi, R. E. *Analyst* **2010**, *135*, 918-926.
- (35) Yao, C.; Qi, Y.; Zhao, Y.; Xiang, Y.; Chen, Q.; Fu, W. *Biosensors and Bioelectronics* **2009**, *24*, 2499-2503.
- (36) Ma, C.; Wang, W.; Yang, Q.; Shi, C.; Cao, L. *Biosensors and Bioelectronics* **2011**, *26*, 3309-3312.

- (37) Jyoti, A.; Vajpayee, P.; Singh, G.; Patel, C. B.; Gupta, K. C.; Shanker, R. *Environ. Sci. Technol.* **2011**, *45*, 8996-9002.
- (38) Pai, S. S.; Ellington, A. D. *Methods Mol. Biol.* **2009**, *504*, 385-398.
- (39) Sultan, Y.; DeRosa, M. C. *Small* **2011**, *7*, 1219-1226.
- (40) Xu, W.; Lu, Y. *Chem. Commun. (Camb)* **2011**, *47*, 4998-5000.
- (41) Kourlas, H.; Schiller, D. S. *Clin. Ther.* **2006**, *28*, 36-44.
- (42) Ellington, A. D.; Szostak, J. W. *Nature* **1992**, *355*, 850-852.
- (43) Wang, S.; Kool, E. T. *Biochemistry* **1995**, *34*, 4125-4132.
- (44) Mills, D. R.; Peterson, R. L.; Spiegelman, S. *Proc. Natl. Acad. Sci. U. S. A.* **1967**, *58*, 217-224.
- (45) Saffhill, R.; Schneider-Bernloehr, H.; Orgel, L. E.; Spiegelman, S. *J. Mol. Biol.* **1970**, *51*, 531-539.
- (46) Wilson, D. S.; Szostak, J. W. *Annu. Rev. Biochem.* **1999**, *68*, 611-647.
- (47) Silverman, S., K. *Functional Nucleic Acids for Analytical Applications* **2009**, *1*, 47-108.
- (48) Vant-Hull, B.; Gold, L.; Zichi, D. A. *Curr. Protoc. Nucleic Acid Chem.* **2000**, Chapter 9, Unit 9.1.
- (49) Stoltenburg, R.; Reinemann, C.; Strehlitz, B. *Anal. Bioanal. Chem.* **2005**, *383*, 83-91.
- (50) Cruz-Aguado, J. A.; Penner, G. *J. Agric. Food Chem.* **2008**, *56*, 10456-10461.
- (51) Gold, L., et al *PLoS One* **2010**, *5*, e15004.
- (52) Stoltenburg, R.; Reinemann, C.; Strehlitz, B. *Biomol. Eng.* **2007**, *24*, 381-403.
- (53) Davis, J. H.; Szostak, J. W. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 11616-11621.
- (54) Win, M. N.; Klein, J. S.; Smolke, C. D. *Nucleic Acids Res.* **2006**, *34*, 5670-5682.
- (55) Cox, J. C.; Rudolph, P.; Ellington, A. D. *Biotechnol. Prog.* **1998**, *14*, 845-850.

- (56) Goertz, P. W.; Colin Cox, J.; Ellington, A. D. *Journal of the Association for Laboratory Automation* **2004**, *9*, 150-154.
- (57) Wochner, A.; Cech, B.; Menger, M.; Erdmann, V. A.; Glokler, J. *BioTechniques* **2007**, *43*, 344, 346, 348 passim.
- (58) Geiger, A.; Burgstaller, P.; von der Eltz, H.; Roeder, A.; Famulok, M. *Nucleic Acids Res.* **1996**, *24*, 1029-1036.
- (59) Jenison, R. D.; Gill, S. C.; Pardi, A.; Polisky, B. *Science* **1994**, *263*, 1425-1429.
- (60) Smith, D.; Kirschenheuter, G. P.; Charlton, J.; Guidot, D. M.; Repine, J. E. *Chem. Biol.* **1995**, *2*, 741-750.
- (61) Martell, R. E.; Nevins, J. R.; Sullenger, B. A. *Mol. Ther.* **2002**, *6*, 30-34.
- (62) Burke, D. H.; Willis, J. H. *RNA* **1998**, *4*, 1165-1175.
- (63) Wu, L.; Curran, J. F. *Nucleic Acids Res.* **1999**, *27*, 1512-1516.
- (64) Morris, K. N.; Jensen, K. B.; Julin, C. M.; Weil, M.; Gold, L. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 2902-2907.
- (65) Jensen, K. B.; Atkinson, B. L.; Willis, M. C.; Koch, T. H.; Gold, L. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 12220-12224.
- (66) Brody, E. N.; Willis, M. C.; Smith, J. D.; Jayasena, S.; Zichi, D.; Gold, L. *Mol. Diagn.* **1999**, *4*, 381-388.
- (67) Klussmann, S.; Nolte, A.; Bald, R.; Erdmann, V. A.; Furste, J. P. *Nat. Biotechnol.* **1996**, *14*, 1112-1115.
- (68) Vater, A.; Jarosch, F.; Buchner, K.; Klussmann, S. *Nucleic Acids Res.* **2003**, *31*, e130.
- (69) Dobbelsstein, M.; Shenk, T. J. *J. Virol.* **1995**, *69*, 8027-8034.
- (70) White, R.; Rusconi, C.; Scardino, E.; Wolberg, A.; Lawson, J.; Hoffman, M.; Sullenger, B. *Mol. Ther.* **2001**, *4*, 567-573.
- (71) Kawakami, J.; Imanaka, H.; Yokota, Y.; Sugimoto, N. *J. Inorg. Biochem.* **2000**, *82*, 197-206.
- (72) Coulter, L. R.; Landree, M. A.; Cooper, T. A. *Mol. Cell. Biol.* **1997**, *17*, 2143-2150.

- (73) Mendonsa, S. D.; Bowser, M. T. *Anal. Chem.* **2004**, *76*, 5387-5392.
- (74) Berezovski, M.; Musheev, M.; Drabovich, A.; Krylov, S. N. *J. Am. Chem. Soc.* **2006**, *128*, 1410-1411.
- (75) Tsai, R. Y.; Reed, R. R. *Mol. Cell. Biol.* **1998**, *18*, 6447-6456.
- (76) Miyachi, Y.; Shimizu, N.; Ogino, C.; Kondo, A. *Nucleic Acids Res.* **2010**, *38*, e21.
- (77) Cassiday, L. A.; Maher, L. J., 3rd *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3930-3935.
- (78) Kim, Y.; Liu, C.; Tan, W. *Biomark Med.* **2009**, *3*, 193-202.
- (79) Schneider, D. J.; Vanderslice, R.; Gold, L. (. 5,861,254, 1999.
- (80) Missailidis, S.; Hardy, A. *Expert Opin. Ther. Pat.* **2009**, *19*, 1073-1082.
- (81) Huizenga, D. E.; Szostak, J. W. *Biochemistry (N. Y.)* **1995**, *34*, 656-665.
- (82) Harada, K.; Frankel, A. D. *EMBO J.* **1995**, *14*, 5798-5811.
- (83) Wilson, C.; Szostak, J. W. *Chem. Biol.* **1998**, *5*, 609-617.
- (84) Yang, Q.; Goldstein, I. J.; Mei, H. Y.; Engelke, D. R. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 5462-5467.
- (85) Rink, S. M.; Shen, J. C.; Loeb, L. A. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 11619-11624.
- (86) Kato, T.; Takemura, T.; Yano, K.; Ikebukuro, K.; Karube, I. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **2000**, *1493*, 12-18.
- (87) Okazawa, A.; Maeda, H.; Fukusaki, E.; Katakura, Y.; Kobayashi, A. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 2653-2656.
- (88) Vianini, E.; Palumbo, M.; Gatto, B. *Bioorg. Med. Chem.* **2001**, *9*, 2543-2548.
- (89) Mehedi Masud, M.; Kuwahara, M.; Ozaki, H.; Sawai, H. *Bioorg. Med. Chem.* **2004**, *12*, 1111-1120.
- (90) Mann, D.; Reinemann, C.; Stoltenburg, R.; Strehlitz, B. *Biochem. Biophys. Res. Commun.* **2005**, *338*, 1928-1934.
- (91) Tang, J.; Xie, J.; Shao, N.; Yan, Y. *Electrophoresis* **2006**, *27*, 1303-1311.

- (92) Shoji, A.; Kuwahara, M.; Ozaki, H.; Sawai, H. *J. Am. Chem. Soc.* **2007**, *129*, 1456-1464.
- (93) Sando, S.; Narita, A.; Aoyama, Y. *Chembiochem* **2007**, *8*, 1795-1803.
- (94) Niazi, J. H.; Lee, S. J.; Gu, M. B. *Bioorg. Med. Chem.* **2008**, *16*, 7245-7253.
- (95) Ohsawa, K.; Kasamatsu, T.; Nagashima, J.; Hanawa, K.; Kuwahara, M.; Ozaki, H.; Sawai, H. *Anal. Sci.* **2008**, *24*, 167-172.
- (96) Wochner, A.; Menger, M.; Orgel, D.; Cech, B.; Rimmele, M.; Erdmann, V. A.; Glöckler, J. *Anal. Biochem.* **2008**, *373*, 34-42.
- (97) Niazi, J. H.; Lee, S. J.; Kim, Y. S.; Gu, M. B. *Bioorg. Med. Chem.* **2008**, *16*, 1254-1261.
- (98) Kim, M.; Um, H. J.; Bang, S.; Lee, S. H.; Oh, S. J.; Han, J. H.; Kim, K. W.; Min, J.; Kim, Y. H. *Environ. Sci. Technol.* **2009**, *43*, 9335-9340.
- (99) Walsh, R.; DeRosa, M. C. *Biochem. Biophys. Res. Commun.* **2009**, *388*, 732-735.
- (100) Miyachi, Y.; Shimizu, N.; Ogino, C.; Fukuda, H.; Kondo, A. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 3619-3622.
- (101) Joeng, C. B.; Niazi, J. H.; Lee, S. J.; Gu, M. B. *Bioorg. Med. Chem.* **2009**, *17*, 5380-5387.
- (102) Sinha, J.; Reyes, S. J.; Gallivan, J. P. *Nat. Chem. Biol.* **2010**, *6*, 464-470.
- (103) Kim, Y. S.; Hyun, C. J.; Kim, I. A.; Gu, M. B. *Bioorg. Med. Chem.* **2010**, *18*, 3467-3473.
- (104) Luo, X.; McKeague, M.; Pitre, S.; Dumontier, M.; Green, J.; Golshani, A.; Derosa, M. C.; Dehne, F. *RNA* **2010**, *16*, 2252-2262.
- (105) McKeague, M.; Bradley, C. R.; De Girolamo, A.; Visconti, A.; Miller, J. D.; Derosa, M. C. *Int. J. Mol. Sci.* **2010**, *11*, 4864-4881.
- (106) He, J.; Liu, Y.; Fan, M.; Liu, X. *J. Agric. Food Chem.* **2011**, *59*, 1582-1586.
- (107) Song, K.; Cho, M.; Jo, H.; Min, K.; Jeon, S. H.; Kim, T.; Han, M. S.; Ku, J. K.; Ban, C. *Anal. Biochem.* **2011**, *415*, 175-181.
- (108) Yang, X.; Bing, T.; Mei, H.; Fang, C.; Cao, Z.; Shangguan, D. *Analyst* **2011**, *136*, 577-585.

- (109) Jo, M.; Ahn, J. Y.; Lee, J.; Lee, S.; Hong, S. W.; Yoo, J. W.; Kang, J.; Dua, P.; Lee, D. K.; Hong, S.; Kim, S. *Oligonucleotides* **2011**, *21*, 85-91.
- (110) Cho, M. J.; Juliano, R. *Trends Biotechnol.* **1996**, *14*, 153-158.
- (111) Ashour, M. L.; Wink, M. *J. Pharm. Pharmacol.* **2011**, *63*, 305-321.
- (112) Roemer, T.; Davies, J.; Giaever, G.; Nislow, C. *Nat. Chem. Biol.* **2011**, *8*, 46-56.
- (113) Walsh, T. A. *Pest Manag. Sci.* **2007**, *63*, 1165-1171.
- (114) Michael, F. *Curr. Opin. Struct. Biol.* **1999**, *9*, 324-329.
- (115) Lin, P. H.; Tong, S. J.; Louis, S. R.; Chang, Y.; Chen, W. Y. *Phys. Chem. Chem. Phys.* **2009**, *11*, 9744-9750.
- (116) Michaud, M.; Jourdan, E.; Villet, A.; Ravel, A.; Grosset, C.; Peyrin, E. *J. Am. Chem. Soc.* **2003**, *125*, 8672-8679.
- (117) Famulok, M.; Szostak, J. W. *J. Am. Chem. Soc.* **1992**, *114*, 3990-3991.
- (118) Famulok, M. *J. Am. Chem. Soc.* **1994**, *116*, 1698-1706.
- (119) Williams, K. P.; Liu, X. H.; Schumacher, T. N.; Lin, H. Y.; Ausiello, D. A.; Kim, P. S.; Bartel, D. P. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 11285-11290.
- (120) Maragos, C. M.; Busman, M. *Food Addit Contam. Part A. Chem. Anal. Control. Expo. Risk Assess.* **2010**, *27*, 688-700.
- (121) Miller, J. D. In *Mycotoxins*; Francis, F. J., Ed.; Encyclopedia of food science and technology. John Wiley: New York, 1999; pp 1698-1698-1706.
- (122) Reverberi, M.; Ricelli, A.; Zjalic, S.; Fabbri, A. A.; Fanelli, C. *Appl. Microbiol. Biotechnol.* **2010**, *87*, 899-911.
- (123) Krska, R.; Molinelli, A. E. *Analytical and Bioanalytical Chemistry* **2007**, *387*, 145-148.
- (124) Lopez-Garcia, R.; Park, D. L.; Phillips, T. D. *FAO: Food, nutrition, agriculture* **1999**, .
- (125) Varga, J.; Kocsube, S.; Peteri, Z.; Vagvolgyi, C.; Toth, B. *Toxins (Basel)* **2010**, *2*, 1718-1750.

- (126) Turner, N. W.; Subrahmanyam, S.; Piletsky, S. A. *Anal. Chim. Acta* **2009**, *632*, 168-180.
- (127) Maragos, C. M. *Anal. Bioanal Chem.* **2009**, *395*, 1205-1213.
- (128) Jing, M.; Bowser, M. T. *Anal. Chim. Acta* **2011**, *686*, 9-18.
- (129) Wilchek, M.; Bayer, E. A.; Livnah, O. *Immunol. Lett.* **2006**, *103*, 27-32.
- (130) Carothers, J. M.; Oestreich, S. C.; Szostak, J. W. *J. Am. Chem. Soc.* **2006**, *128*, 7929-7937.
- (131) Carothers, J. M.; Goler, J. A.; Kapoor, Y.; Lara, L.; Keasling, J. D. *Nucleic Acids Res.* **2010**, *38*, 2736-2747.
- (132) Ellington, A. D. *Curr. Biol.* **1994**, *4*, 427-429.
- (133) White, R. J.; Rowe, A. A.; Plaxco, K. W. *Analyst* **2010**, *135*, 589-594.
- (134) Potty, A. S.; Kourentzi, K.; Fang, H.; Jackson, G. W.; Zhang, X.; Legge, G. B.; Willson, R. C. *Biopolymers* **2009**, *91*, 145-156.
- (135) Lin, P.; Chen, R.; Lee, C.; Chang, Y.; Chen, C.; Chen, W. *Colloids and Surfaces B: Biointerfaces* **2011**, *88*, 552-558.
- (136) Lee, J. H.; Canny, M. D.; De Erkenez, A.; Krilleke, D.; Ng, Y. S.; Shima, D. T.; Pardi, A.; Jucker, F. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 18902-18907.
- (137) Gaillard, C.; Strauss, F. *BMC Biochem.* **2000**, *1*, 1.
- (138) Flinders, J.; DeFina, S. C.; Brackett, D. M.; Baugh, C.; Wilson, C.; Dieckmann, T. *Chembiochem* **2004**, *5*, 62-72.
- (139) Cruz-Aguado, J. A.; Penner, G. *Anal. Chem.* **2008**, *80*, 8853-8855.
- (140) Guedin, A.; Lacroix, L.; Mergny, J. L. *Methods Mol. Biol.* **2010**, *613*, 25-35.
- (141) Sultan, Y.; Walsh, R.; Monreal, C.; DeRosa, M. C. *Biomacromolecules* **2009**, *10*, 1149-1154.
- (142) Deng, Q.; German, I.; Buchanan, D.; Kennedy, R. T. *Anal. Chem.* **2001**, *73*, 5415-5421.
- (143) Drabovich, A. P.; Berezovski, M.; Okhonin, V.; Krylov, S. N. *Anal. Chem.* **2006**, *78*, 3171-3178.

- (144) Hall, B.; Arshad, S.; Seo, K.; Bowman, C.; Corley, M.; Jhaveri, S. D.; Ellington, A. D. *Curr. Protoc. Mol. Biol.* **2009**, Chapter 24, Unit 24.3.
- (145) Bock, L. C.; Griffin, L. C.; Latham, J. A.; Vermaas, E. H.; Toole, J. J. *Nature* **1992**, 355, 564-566.
- (146) Juskowiak, B. *Anal. Bioanal Chem.* **2011**, 399, 3157-3176.
- (147) Turesky, R. J. *Chem. Res. Toxicol.* **2005**, 18, 1082-1090.
- (148) van der Merwe, K. J.; Steyn, P. S.; Fourie, L.; Scott, D. B.; Theron, J. J. *Nature* **1965**, 205, 1112-1113.
- (149) Duarte, S. C.; Lino, C. M.; Pena, A. *Food Addit Contam. Part A. Chem. Anal. Control. Expo. Risk Assess.* **2010**, 27, 1440-1450.
- (150) Tittlemier, S. A.; Varga, E.; Scott, P. M.; Krska, R. *Food Addit Contam. Part A. Chem. Anal. Control. Expo. Risk Assess.* **2011**, 28, 775-785.
- (151) Creppy, E. -.; Lugnier, A. A. J.; Fasiolo, F.; Heller, K.; Rösenthaller, R.; Dirheimer, G. *Chem. Biol. Interact.* **1979**, 24, 257-261.
- (152) Galtier, P.; Boneu, B.; Charpentreau, J. L.; Bodin, G.; Alvinerie, M.; More, J. *Food Cosmet. Toxicol.* **1979**, 17, 49-53.
- (153) Pitout, M. J. *Toxicol. Appl. Pharmacol.* **1968**, 13, 299-306.
- (154) Creppy, E. E.; Stormer, F. C.; Rosenthaller, R.; Dirheimer, G. *Infect. Immun.* **1983**, 39, 1015-1018.
- (155) Mayura, K.; Parker, R.; Berndt, W. O.; Phillips, T. D. *Appl. Environ. Microbiol.* **1984**, 48, 1186-1188.
- (156) Creppy, E. E.; Chiarappa, P.; Baudrimont, I.; Borracci, P.; Moukha, S.; Carratù, M. R. *Toxicology* **2004**, 201, 115-123.
- (157) Pfohl-Leszkowicz, A.; Manderville, R. A. *Mol. Nutr. Food Res.* **2007**, 51, 61-99.
- (158) Canadian, G. C. **Prevent ochratoxin A in stored grain.**
<http://www.grainscanada.gc.ca/storage-entrepot/ota/ota-eng.htm> (accessed March/8, 2012).
- (159) De Girolamo, A.; McKeague, M.; Miller, J. D.; DeRosa, M. C.; Visconti, A. *Food Chem.* **2011**, 127, 1378-1384.

- (160) Charmley, L. L.; Trenholm, H. L. *Canadian Food Inspection Agency, Feed Program* **2009**.
- (161) Bureau of Chemical Safety Food Directorate Health Products and Food Branch **February 2009**.
- (162) Yang, C.; Wang, Y.; Marty, J.; Yang, X. *Biosensors and Bioelectronics* **2010**, *In Press, Accepted Manuscript*.
- (163) Sheng, L.; Ren, J.; Miao, Y.; Wang, J.; Wang, E. *Biosensors and Bioelectronics* **2011**, *26*, 3494-3499.
- (164) Wang, Z.; Duan, N.; Hun, X.; Wu, S. *Anal. Bioanal Chem.* **2010**, *398*, 2125-2132.
- (165) Maragos, C. M.; Thompson, V. S. *Nat. Toxins* **1999**, *7*, 371-376.
- (166) Zezza, F.; Longobardi, F.; Pascale, M.; Eremin, S. A.; Visconti, A. *Anal. Bioanal Chem.* **2009**, *395*, 1317-1323.
- (167) Prabhakar, N.; Matharu, Z.; Malhotra, B. D. *Biosensors and Bioelectronics* **2011**, *26*, 4006-4011.
- (168) Kuang, H.; Chen, W.; Xu, D.; Xu, L.; Zhu, Y.; Liu, L.; Chu, H.; Peng, C.; Xu, C.; Zhu, S. *Biosensors and Bioelectronics* **2010**, *26*, 710-716.
- (169) Wang, L.; Ma, W.; Chen, W.; Liu, L.; Ma, W.; Zhu, Y.; Xu, L.; Kuang, H.; Xu, C. *Biosensors and Bioelectronics* **2011**, *26*, 3059-3062.
- (170) Wang, L.; Chen, W.; Ma, W.; Liu, L.; Ma, W.; Zhao, Y.; Zhu, Y.; Xu, L.; Kuang, H.; Xu, C. *Chem. Commun. (Camb)* **2011**, *47*, 1574-1576.
- (171) Yang, C.; Lates, V.; Prieto-Simón, B.; Marty, J.; Yang, X. *Biosensors and Bioelectronics* **2012**, *32*, 208-212.
- (172) Guo, Z.; Ren, J.; Wang, J.; Wang, E. *Talanta* **2011**, *85*, 2517-2521.
- (173) Tong, P.; Zhao, W.; Zhang, L.; Xu, J.; Chen, H. *Biosensors and Bioelectronics* **2012**, *33*, 146-151.
- (174) Kidd, A.; Guieu, V.; Perrier, S.; Ravelet, C.; Peyrin, E. *Anal. Bioanal Chem.* **2011**, *401*, 3229-3234.
- (175) McKeague, M.; Giamberardino, A.; DeRosa, M. C. In *Advances in Aptamer-Based Biosensors for Food Safety*; Environmental Biosensors; Vernon Somerset, InTech: SBN: 9789533074863, 2011; pp 17-18-42.

- (176) Stead, S. L.; Ashwin, H.; Johnston, B.; Dallas, A.; Kazakov, S. A.; Tarbin, J. A.; Sharman, M.; Kay, J.; Keely, B. J. *Anal. Chem.* **2010**, *82*, 2652-2660.
- (177) Zelada-Guillen, G. A.; Bhosale, S. V.; Riu, J.; Rius, F. X. *Anal. Chem.* **2010**, *82*, 9254-9260.
- (178) Bruno, J. G.; Phillips, T.; Carrillo, M. P.; Crowell, R. J. *J. Fluoresc.* **2009**, *19*, 427-435.
- (179) Sambrook, J. R. D. In *Molecular cloning : A Laboratory Manual*. 2001; .
- (180) Chen, Z.; Ruffner, D. E. *BioTechniques* **1996**, *21*, 820-822.
- (181) Fylstra, D.; Lasdon, L.; Watson, J.; Waren, A. *INTERFACES* **1998**, *28*, 29-55.
- (182) Nenov, I. P.; Fylstra, D. H. *Reliable Comput.* **2003**, *9*, 143-159.
- (183) Entwisle, A. C.; Williams, A. C.; Mann, P. J.; Slack, P. T.; Gilbert, J. J. *AOAC Int.* **2000**, *83*, 1377-1383.
- (184) Joint FAO/WHO Expert Committee on Food Additives, (JECFA) **2001**, .
- (185) Sassanfar, M.; Szostak, J. W. *Nature* **1993**, *364*, 550-553.
- (186) BURGSTALLER, P.; FAMULOK, M. *Angew. Chem. -Int. Edit. Engl.* **1994**, *33*, 1084-1087.
- (187) Burke, D. H.; Gold, L. *Nucleic Acids Res.* **1997**, *25*, 2020-2024.
- (188) Macaya, R. F.; Waldron, J. A.; Beutel, B. A.; Gao, H.; Joesten, M. E.; Yang, M.; Patel, R.; Bertelsen, A. H.; Cook, A. F. *Biochemistry* **1995**, *34*, 4478-4492.
- (189) Tasset, D. M.; Kubik, M. F.; Steiner, W. *J. Mol. Biol.* **1997**, *272*, 688-698.
- (190) Tahiri-Alaoui, A.; Frigotto, L.; Manville, N.; Ibrahim, J.; Romby, P.; James, W. *Nucleic Acids Res.* **2002**, *30*, e45.
- (191) Bittker, J. A.; Le, B. V.; Liu, D. R. *Nat. Biotechnol.* **2002**, *20*, 1024-1029.
- (192) Song, Y.; Wei, W.; Qu, X. *Adv Mater* **2011**, *23*, 4215-4236.
- (193) Wei, H.; Li, B.; Li, J.; Wang, E.; Dong, S. *Chem. Commun. (Camb)* **2007**, (36), 3735-3737.
- (194) Turkevich, J.; Stevenson, P. C.; Hillier, J. *Discuss. Faraday Soc.* **1951**, *11*.

- (195) Frens, G. *Colloid & Polymer Science* **1972**, *250*, 736-741.
- (196) Azzazy, H. M. E.; Mansour, M. M. H. *Clinica Chimica Acta* **2009**, *403*, 1-8.
- (197) Li, H.; Rothberg, L. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 14036-14039.
- (198) Nelson, E. M.; Rothberg, L. J. *Langmuir* **2011**, *27*, 1770-1777.
- (199) Liu Xueping; Zhou Zhenhua; Zhang Liangliang; Tan Zhongyang; Shen Guoli; Yu Ruqin *Chin. J. Chem.* **2009**, *27*, 1855-1859.
- (200) Murphy, M. B.; Fuller, S. T.; Richardson, P. M.; Doyle, S. A. *Nucleic Acids Res.* **2003**, *31*, e110.
- (201) Rosenthal, A.; Coutelle, O.; Craxton, M. *Nucleic Acids Res.* **1993**, *21*, 173-174.
- (202) Zheng, Y.; Wang, Y.; Yang, X. *Sensors Actuators B: Chem.* **2011**, *156*, 95-99.
- (203) Tok, J. B.; Fischer, N. O. *Chem. Commun. (Camb)* **2008**, (16), 1883-1885.
- (204) Zuker, M. *Nucleic Acids Res.* **2003**, *31*, 3406-3415.
- (205) Peyret, N. Prediction of Nucleic Acid Hybridization: Parameters and Algorithms
PhD dissertation, Wayne State University, Department of Chemistry, Detroit, MI,
2000.
- (206) SantaLucia, J., Jr *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 1460-1465.
- (207) Khlebtsov, N. G. *Anal. Chem.* **2008**, *80*, 6620-6625.
- (208) Cruz-Toledo, J.; McKeague, M.; Zhang, X.; Giamberardino, A.; McConnell, E.;
Francis, T.; Derosa, M. C.; Dumontier, M. *Database (Oxford)* **2012**, *2012*, bas006.
- (209) de la Campa, R.; Hooker, D. C.; Miller, J. D.; Schaafsma, A. W.; Hammond, B. G.
Mycopathologia **2005**, *159*, 539-552.
- (210) Stockmann-Juvala, H.; Savolainen, K. *Hum. Exp. Toxicol.* **2008**, *27*, 799-809.
- (211) Arranz, I.; Baeyens, W. R.; Van der Weken, G.; De Saeger, S.; Van Peteghem, C.
Crit. Rev. Food Sci. Nutr. **2004**, *44*, 195-203.
- (212) SCF Updated opinion on fumonisin B1, B2 and B3. (accessed October/20, 2010).
- (213) Molinelli, A.; Grossalber, K.; Krska, R. *Anal. Bioanal Chem.* **2009**, *395*, 1309-1316.

- (214) Bansal, J.; Pantazopoulos, P.; Tam, J.; Cavlovic, P.; Kwong, K.; Turcotte, A. M.; Lau, B. P.; Scott, P. M. *Food Addit Contam. Part A. Chem. Anal. Control. Expo. Risk Assess.* **2011**, *28*, 767-774.
- (215) Miller, J. D. *Environ. Health Perspect.* **2001**, *109 Suppl 2*, 321-324.
- (216) Taranu, I.; Marin, D. E.; Bouhet, S.; Pascale, F.; Bailly, J. D.; Miller, J. D.; Pinton, P.; Oswald, I. P. *Toxicol. Sci.* **2005**, *84*, 301-307.
- (217) Bolger, M.; Coker, R. D.; DiNovi, M.; Gaylor, D.; Gelderblom, W.; Olsen, M.; Paster, N.; Riley, R. T.; Shephard, G.; Speijers, G. J. A. *WHO Food Additives Series* **2001**, *47*, 557-680.
- (218) Marasas, W. F., et al *J. Nutr.* **2004**, *134*, 711-716.
- (219) Loiseau, N.; Debrauwer, L.; Sambou, T.; Bouhet, S.; Miller, J. D.; Martin, P. G.; Viadère, J.; Pinton, P.; Puel, O.; Pineau, T.; Tulliez, J.; Galtier, P.; Oswald, I. P. *Biochem. Pharmacol.* **2007**, *74*, 144-152.
- (220) Senyuva, H. Z., et al *J. AOAC Int.* **2010**, *93*, 611-621.
- (221) De Girolamo, A.; Pascale, M.; Visconti, A. *Food Addit Contam. Part A. Chem. Anal. Control. Expo. Risk Assess.* **2011**, *28*, 667-675.
- (222) Sulyok, M.; Krska, R.; Schuhmacher, R. *Food Chem.* **2010**, *119*, 408-416.
- (223) Paul, A.; Avci-Adali, M.; Ziemer, G.; Wendel, H. P. *Oligonucleotides* **2009**, *19*, 243-254.
- (224) Visconti, A.; Solfrizzo, M.; De Girolamo, A. *J. AOAC Int.* **2001**, *84*, 1828-1837.
- (225) Anderson, G. P.; Kowtha, V. A.; Taitt, C. R. *Toxins* **2010**, *2*, 297-309.
- (226) Paborsky, L. R.; McCurdy, S. N.; Griffin, L. C.; Toole, J. J.; Leung, L. L. *J. Biol. Chem.* **1993**, *268*, 20808-20811.
- (227) de Soultrait, V. R.; Lozach, P.; Altmeyer, R.; Tarrago-Litvak, L.; Litvak, S.; Andréola, M. L. *J. Mol. Biol.* **2002**, *324*, 195-203.
- (228) Hopmans, E. C.; Murphy, P. A. *J. Agric. Food Chem.* **1993**, *41*, 1655-1658.
- (229) Fukusaki, E.; Hasunuma, T.; Kajiyama, S.; Okazawa, A.; Itoh, T. J.; Kobayashi, A. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2927-2930.
- (230) Regulski, E. E.; Breaker, R. R. *Methods Mol. Biol.* **2008**, *419*, 53-67.

- (231) Yoshida, W.; Sode, K.; Ikebukuro, K. *Biochem. Biophys. Res. Commun.* **2006**, *348*, 245-252.
- (232) McManus, S. A.; Li, Y. *Biochemistry* **2007**, *46*, 2198-2204.
- (233) Maxam, A. M.; Gilbert, W. *Proc. Natl. Acad. Sci. U. S. A.* **1977**, *74*, 560-564.
- (234) Butz, L.; du Vigneaud, V. *J Biol Chem* **1932**, *99*, 135-136,137,138,139,140,141,121.
- (235) Blom, H. J.; Smulders, Y. *J. Inherit. Metab. Dis.* **2011**, *34*, 75-81.
- (236) Perla-Kajan, J.; Twardowski, T.; Jakubowski, H. *Amino Acids* **2007**, *32*, 561-572.
- (237) Gruber, K.; Puffer, B.; Krautler, B. *Chem. Soc. Rev.* **2011**, *40*, 4346-4363.
- (238) Brustolin, S.; Giugliani, R.; Felix, T. M. *Braz. J. Med. Biol. Res.* **2010**, *43*, 1-7.
- (239) Williams, K. T.; Schalinske, K. L. *Biofactors* **2010**, *36*, 19-24.
- (240) Refsum, H.; Smith, A. D.; Ueland, P. M.; Nexø, E.; Clarke, R.; McPartlin, J.; Johnston, C.; Engbaek, F.; Schneede, J.; McPartlin, C.; Scott, J. M. *Clin. Chem.* **2004**, *50*, 3-32.
- (241) Abraham, J. M.; Cho, L. *Cleve. Clin. J. Med.* **2010**, *77*, 911-918.
- (242) Clarke, R.; Halsey, J.; Bennett, D.; Lewington, S. *J. Inherit. Metab. Dis.* **2011**, *34*, 83-91.
- (243) Humphrey, L. L.; Fu, R.; Rogers, K.; Freeman, M.; Helfand, M. *Mayo Clin. Proc.* **2008**, *83*, 1203-1212.
- (244) Rafii, M.; Elango, R.; House, J. D.; Courtney-Martin, G.; Darling, P.; Fisher, L.; Pencharz, P. B. *Journal of Chromatography B* **2009**, *877*, 3282-3291.
- (245) Nekrassova, O.; Lawrence, N. S.; Compton, R. G. *Talanta* **2003**, *60*, 1085-1095.
- (246) Mu, X.; Zhao, S.; Huang, Y.; Ye, F. *J. Sep. Sci.* **2011**.
- (247) Ducros, V.; Demuth, K.; Sauvant, M.; Quillard, M.; Causse, E.; Candito, M.; Read, M.; Drai, J.; Garcia, I.; Gerhardt, M. *Journal of Chromatography B* **2002**, *781*, 207-226.
- (248) Shevchenko, Y.; Francis, T. J.; Blair, D. A.; Walsh, R.; DeRosa, M. C.; Albert, J. *Anal. Chem.* **2011**, *83*, 7027-7034.

- (249) Gebhardt, K.; Shokraei, A.; Babaie, E.; Lindqvist, B. H. *Biochemistry* **2000**, *39*, 7255-7265.
- (250) Duerre, J. A.; Miller, C. H. *Anal. Biochem.* **1966**, *17*, 310-315.
- (251) Sedlak, J.; Lindsay, R. H. *Anal. Biochem.* **1968**, *25*, 192-205.
- (252) Mannironi, C.; Di Nardo, A.; Fruscoloni, P.; Tocchini-Valentini, G. P. *Biochemistry* **1997**, *36*, 9726-9734.
- (253) Gilbert, H. F. *Methods Enzymol.* **1995**, *251*, 8-28.
- (254) Cramer, A.; Stemmer, W. P. *Nucleic Acids Res.* **1993**, *21*, 4410.
- (255) Gopinath, S. C. *Anal. Bioanal Chem.* **2007**, *387*, 171-182.
- (256) Wu, H.; Huang, C.; Cheng, T.; Tseng, W. *Talanta* **2008**, *76*, 347-352.
- (257) Zhong, Z.; Patskovskyy, S.; Bouvrette, P.; Luong, J. H. T.; Gedanken, A. *J Phys Chem B* **2004**, *108*, 4046-4052.
- (258) Wang, Y.; Wang, J.; Yang, F.; Yang, X. *Anal. Sci.* **2010**, *26*, 545-549.
- (259) Schwalbe, H.; Buck, J.; Furtig, B.; Noeske, J.; Wohner, J. *Angew. Chem. Int. Ed Engl.* **2007**, *46*, 1212-1219.
- (260) Laing, C.; Schlick, T. *Curr. Opin. Struct. Biol.* **2011**, *21*, 306-318.
- (261) Lilley, D. M.; Clegg, R. M.; Diekmann, S.; Seeman, N. C.; Von Kitzing, E.; Hagerman, P. J. *Nucleic Acids Res.* **1995**, *23*, 3363-3364.
- (262) Scott, W. G.; Murray, J. B.; Arnold, J. R.; Stoddard, B. L.; Klug, A. *Science* **1996**, *274*, 2065-2069.
- (263) Wilson, T. J.; Nahas, M.; Ha, T.; Lilley, D. M. *Biochem. Soc. Trans.* **2005**, *33*, 461-465.
- (264) Kieft, J. S.; Zhou, K.; Grech, A.; Jubin, R.; Doudna, J. A. *Nat. Struct. Biol.* **2002**, *9*, 370-374.
- (265) Montange, R. K.; Batey, R. T. *Annu. Rev. Biophys.* **2008**, *37*, 117-133.
- (266) Gilbert, S. D.; Batey, R. T. *Cell Mol. Life Sci.* **2005**, *62*, 2401-2404.

- (267) Tremblay, R.; Mulhbachter, J.; Blouin, S.; Penedo, J. C.; Lafontaine, D. A. In *Natural Functional Nucleic Acids: Ribozymes and Riboswitches*; Yingfu, L., Yi, L., Eds.; Functional Nucleic Acids for Analytical Applications; Springer New York: 2009; pp 11-46.
- (268) Serganov, A.; Huang, L.; Patel, D. J. *Nature* **2008**, *455*, 1263-U76.
- (269) Batey, R. T.; Gilbert, S. D.; Montange, R. K. *Nature* **2004**, *432*, 411-415.
- (270) Zinnen, S. P.; Domenico, K.; Wilson, M.; Dickinson, B. A.; Beaudry, A.; Mokler, V.; Daniher, A. T.; Burgin, A.; Beigelman, L. *RNA* **2002**, *8*, 214-228.
- (271) Khoo, D.; Perez, C.; Mohr, I. *J. Virol.* **2002**, *76*, 11971-11981.
- (272) Laserson, U.; Gan, H. H.; Schlick, T. *Proceedings of the 12th Annual Symposium on Computational geometry* **2004**, , 373-374,375,376,377.
- (273) Lee, J. F.; Hesselberth, J. R.; Meyers, L. A.; Ellington, A. D. *Nucleic Acids Res.* **2004**, *32*, D95-100.
- (274) Carothers, J. M.; Davis, J. H.; Chou, J. J.; Szostak, J. W. *RNA* **2006**, *12*, 567-579.
- (275) Gevertz, J.; Gan, H. H.; Schlick, T. *RNA* **2005**, *11*, 853-863.
- (276) Sabeti, P. C.; Unrau, P. J.; Bartel, D. P. *Chem. Biol.* **1997**, *4*, 767-774.
- (277) Beattie, K. L.; Logsdon, N. J.; Anderson, R. S.; Espinosa-Lara, J. M.; Maldonado-Rodriguez, R.; Frost, J. D.,3rd *Biotechnol. Appl. Biochem.* **1988**, *10*, 510-521.
- (278) Carothers, J. M.; Oestreich, S. C.; Davis, J. H.; Szostak, J. W. *J. Am. Chem. Soc.* **2004**, *126*, 5130-5137.
- (279) Jaeger, L.; Wright, M. C.; Joyce, G. F. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 14712-14717.
- (280) Ohuchi, S. J.; Ikawa, Y.; Shiraishi, H.; Inoue, T. *Nucleic Acids Res.* **2002**, *30*, 3473-3480.
- (281) Lau, M. W.; Cadieux, K. E.; Unrau, P. J. *J. Am. Chem. Soc.* **2004**, *126*, 15686-15693.
- (282) Yoshioka, W.; Ikawa, Y.; Jaeger, L.; Shiraishi, H.; Inoue, T. *RNA* **2004**, *10*, 1900-1906.

- (283) Knight, R.; De Sterck, H.; Markel, R.; Smit, S.; Oshmyansky, A.; Yarus, M. *Nucl. Acids Res.* **2005**, *33*, 5924-5935.
- (284) Legiewicz, M.; Lozupone, C.; Knight, R.; Yarus, M. *RNA* **2005**, *11*, 1701-1709.
- (285) Fernandez, M.; Caballero, J.; Fernandez, L.; Sarai, A. *Mol. Divers.* **2011**, *15*, 269-289.
- (286) Hofacker, I. L. *Nucleic Acids Res.* **2003**, *31*, 3429-3431.
- (287) Zuker, M.; Stiegler, P. *Nucleic Acids Res.* **1981**, *9*, 133-148.
- (288) Chiuman, W.; Li, Y. *J. Mol. Biol.* **2006**, *357*, 748-754.
- (289) Amrita Mathuriya* David A. Bader† Christine E. Heitsch‡ Stephen C. Harvey§ , .
- (290) Parisien, M.; Major, F. *Nature* **2008**, *452*, 51-55.
- (291) Kim, N.; Shin, J. S.; Elmetwaly, S.; Gan, H. H.; Schlick, T. *Bioinformatics* **2007**, *23*, 2959-2960.
- (292) Kim, N.; Gan, H. H.; Schlick, T. *RNA-Publ. RNA Soc.* **2007**, *13*, 478-492.
- (293) Kim, N.; Izzo, J. A.; Elmetwaly, S.; Gan, H. H.; Schlick, T. *Nucleic Acids Res.* **2010**, *38*, e139.
- (294) Thodima, V.; Pirooznia, M.; Deng, Y. *BMC Bioinformatics* **2006**, *7 Suppl 2*, S6.
- (295) Ponomarenko, J. V.; Orlova, G. V.; Frolov, A. S.; Gelfand, M. S.; Ponomarenko, M. P. *Nucleic Acids Res.* **2002**, *30*, 195-199.
- (296) Benson, D. A.; Karsch-Mizrachi, I.; Lipman, D. J.; Ostell, J.; Sayers, E. W. *Nucleic Acids Res.* **2010**, *38*, D46-51.
- (297) Cochrane, G., et al *Nucleic Acids Res.* **2009**, *37*, D19-25.
- (298) Kaminuma, E.; Mashima, J.; Kodama, Y.; Gojobori, T.; Ogasawara, O.; Okubo, K.; Takagi, T.; Nakamura, Y. *Nucleic Acids Res.* **2010**, *38*, D33-8.
- (299) World Economic Forum. *Global agenda council on emerging technologies*, **2011**.
- (300) Pattal, M. M. I.; Li Yuan; Zeng Jianqiu In *Web 3.0: A real personal Web! More opportunities & more threats*; AlBegain, K Balakrishna, C Casado, AC, Ed.; IEEE COMPUTER SOC: LOS ALAMITOS; 10662 LOS VAQUEROS CIRCLE, PO BOX 3014, LOS ALAMITOS, CA 90720-1264 USA, 2009; , pp 128.

- (301) Chen, P. P. *ACM TODS* **1976**, *1*, 0-36.
- (302) W3C SweoIG/TaskForces/CommunityProjects/LinkingOpenData.
<http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData>
 (accessed March/2012, 2012).
- (303) W3C W3C Recommendation 10 February 2004. <http://www.w3.org/TR/rdf-primer/>
 (accessed March/2012, 2012).
- (304) Pfeffer, P.; Gohlke, H. *J. Chem. Inf. Model.* **2007**, *47*, 1868-1876.
- (305) Ho, P. Q. A statistical analysis of aptamer target affinity, Carleton University, Ottawa, 2012.
- (306) Sobrova, P.; Adam, V.; Vasatkova, A.; Beklova, M.; Zeman, L.; Kizek, R. *Interdiscip. Toxicol.* **2010**, *3*, 94-99.
- (307) Awad, W. A.; Ghareeb, K.; Bohm, J.; Zentek, J. *Food Addit Contam. Part A. Chem. Anal. Control. Expo. Risk Assess.* **2010**, *27*, 510-520.
- (308) Schaafsma, A. W. *Adv. Exp. Med. Biol.* **2002**, *504*, 271-276.
- (309) Döll, S.; Dänicke, S. *Prev. Vet. Med.* **2011**, *102*, 132-145.
- (310) Klotzel, M.; Gutsche, B.; Lauber, U.; Humpf, H. U. *J. Agric. Food Chem.* **2005**, *53*, 8904-8910.
- (311) Warth, B.; Sulyok, M.; Berthiller, F.; Schuhmacher, R.; Fruhmman, P.; Hametner, C.; Adam, G.; Frohlich, J.; Krska, R. *Anal. Bioanal Chem.* **2011**, *401*, 195-200.
- (312) Kadota, T.; Takezawa, Y.; Hirano, S.; Tajima, O.; Maragos, C. M.; Nakajima, T.; Tanaka, T.; Kamata, Y.; Sugita-Konishi, Y. *Anal. Chim. Acta* **2010**, *673*, 173-178.
- (313) Clifford, L. J.; Jia, Q.; Pestka, J. J. *J. Agric. Food Chem.* **2003**, *51*, 521-523.
- (314) Vesonder, R. F.; Ciegler, A.; Jensen, A. H. *Appl. Microbiol.* **1973**, *26*, 1008-1010.
- (315) De Girolamo, A.; Lippolis, V.; Nordkvist, E.; Visconti, A. *Food Addit Contam. Part A. Chem. Anal. Control. Expo. Risk Assess.* **2009**, *26*, 907-917.
- (316) Lattanzio, V. M. T.; Solfrizzo, M.; De Girolamo, A.; Chulze, S. N.; Torres, A. M.; Visconti, A. *Journal of Chromatography B* **2011**, *879*, 707-715.
- (317) Karlovsky, P. *Appl. Microbiol. Biotechnol.* **2011**, *91*, 491-504.

- (318) Pestka, J. J. *Arch. Toxicol.* **2010**, *84*, 663-679.
- (319) Solfrizzo, M.; Gambacorta, L.; Lattanzio, V. M.; Powers, S.; Visconti, A. *Anal. Bioanal Chem.* **2011**, *401*, 2831-2841.
- (320) Pestka, J. J.; Smolinski, A. T. *J. Toxicol. Environ. Health B Crit. Rev.* **2005**, *8*, 39-69.
- (321) Food and Drug Administration, USA. **2010**.
- (322) Lupo, A.; Roebuck, C.; Settimo, K.; Quain, A.; Kennedy, J.; Abouzied, M. *J. AOAC Int.* **2010**, *93*, 600-610.
- (323) Bily, A. C.; Reid, L. M.; Savard, M. E.; Reddy, R.; Blackwell, B. A.; Campbell, C. M.; Krantis, A.; Durst, T.; Philogene, B. J.; Arnason, J. T.; Regnault-Roger, C. *Mycopathologia* **2004**, *157*, 117-126.
- (324) Jean-Michel, V. *Tetrahedron* **2004**, *60*, 4251-4260.
- (325) Wallace, S. T.; Schroeder, R. *RNA* **1998**, *4*, 112-123.
- (326) Krska, R.; Szenté, E.; Freudenschuss, M.; Hametner, C.; Zollner, P. *J. AOAC Int.* **2004**, *87*, 909-919.
- (327) Miller, J. D.; Blackwell, B. A. *Can J Botany* **1986**, *64*, 1-2,3,4,5.
- (328) Greenhalgh, R.; Levandier, D.; Adams, W.; Miller, J. D.; Blackwell, B. A.; McAlees, A. J.; Taylor, A. *J. Agric. Food Chem.* **1986**, *34*, 98-102.
- (329) Yang, X. *Dissertation Thesis Microbiology Unit of Nanchang University* **2007**.
- (330) Bindewald, E.; Hayes, R.; Yingling, Y. G.; Kasprzak, W.; Shapiro, B. A. *Nucleic Acids Res.* **2008**, *36*, D392-7.