2007

Development of ss DNA aptamers for c-Myc:Max by SELEX (Systematic Evolution of Ligands by Exponential Enrichment)

Ying Liu

Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Genetics and Genomics Commons

Recommended Citation

Liu, Ying, "Development of ss DNA aptamers for c-Myc:Max by SELEX (Systematic Evolution of Ligands by Exponential Enrichment)" (2007). Retrospective Theses and Dissertations. 14522.

https://lib.dr.iastate.edu/rtd/14522

This Thesis is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Development of ss DNA aptamers for c-Myc:Max by SELEX
(Systematic Evolution of Ligands by Exponential Enrichment)

by

Ying Liu

A thesis submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Genetics

Program of Study Committee:
Marit Nilsen-Hamilton, Major Professor
David Hannaple
Ted Huiatt

Iowa State University
Ames, Iowa
2007

Copyright © Ying Liu, 2007. All rights reserved.
# TABLE OF CONTENTS

**ABSTRACT** ........................................................................................................................................ iv

**INTRODUCTION** ................................................................................................................................. 1

Thesis organization .............................................................................................................................. 1

The structure and function of c-Myc:Max ......................................................................................... 2

Aptamer .................................................................................................................................................. 12

Comparison of aptamers and antibodies ......................................................................................... 14

Applications of aptamers .................................................................................................................. 15

Diagnostics ........................................................................................................................................... 15

Therapeutics ........................................................................................................................................ 23

**SELEX** ........................................................................................................................................... 27

The establishment of SELEX method ............................................................................................... 27

Development of SELEX ..................................................................................................................... 29

Genomic SELEX .................................................................................................................................. 29

Auto SELEX ......................................................................................................................................... 32

Capillary SELEX ................................................................................................................................. 33

Photo SELEX ........................................................................................................................................ 35

**MATERIALS AND METHODS** ........................................................................................................ 36

Plasmids and oligonucleotides .......................................................................................................... 36

Materials and equipment .................................................................................................................... 37

Checking cDNA sequences of human c-Myc and Max cloned in expression vector ....................... 38

Protein expression and purification .................................................................................................. 38

Expression of recombinant human c-Myc and Max protein in *E. coli* ........................................... 38

Purification of recombinant human Max protein ............................................................................ 38

Purification of recombinant human c-Myc protein ......................................................................... 39

Reconstitution of the Myc:Max complex ......................................................................................... 40

EMSA (Electrophoretic Mobility-Shift Assay) .................................................................................. 40

End labeled E box probe ..................................................................................................................... 40

EMSA .................................................................................................................................................... 41

**RESULTS AND DISCUSSION** ......................................................................................................... 45

Expression and purification of full length of human 6His-tagged c-Myc and Max fusion proteins in *E. coli* ........................................................................................................................................ 45

Refolding and analysis of recombinant full length c-Myc:Max complex ......................................... 46

Choices of temperature and KCl concentration of SELEX ............................................................. 48
Design of original random pool of SELEX ................................. 49
Aptamer development ................................................................. 50
Binding properties of full length and truncated sequences
of selected oligonucleotides ....................................................... 54

CONCLUSIONS AND FUTURE DIRECTIONS ................................. 58
APPENDIX ...................................................................................... 63
FIGURES ....................................................................................... 67
REFERENCES .............................................................................. 104
ABSTRACT

Cancer, a complicated disease, results from various causes including abnormal overexpression of oncogene products. The c-myc gene was found as the cellular homolog of v-myc oncogene. Its deregulated expression can result in a wide range of tumors in human and mouse. c-Myc, as a transcription factor, requires the Max protein to form a heterodimer that binds to its target DNA sequence, CACGTG, termed an E box, to drive some cancer-related downstream gene expression. Here a strategy of developing DNA aptamer(s) using SELEX (Systemic Evolution of ligands by Exponential Enrichment) was performed to isolate single-stranded DNA molecules that could recognize the c-Myc:Max protein and inhibit c-Myc's activity in vitro and in vivo, and therefore to develop a new type of reagent for cancer therapy.

Aptamers are small single-stranded DNA or RNA molecules that can bind to a wide range of targets from small molecules, such as ATP, to proteins and even to whole cells, with high specificity and affinity. In the procedure of SELEX, we attempted to combine the “decoy” approach to develop an anti-c-Myc aptamer. In the design of the original pool, we put a complement of the E box into the 5’ primer followed by a 42 base random sequence. We expected that the selected anti-c-Myc aptamer includes a primary c-Myc binding site that may include a loop-like structure. We also expected that the other complement of E box would be selected from the random region and form a stem structure with the one included in the 5’ primer. This E box could then bind to the E box binding domain of the c-Myc:Max heterodimer. So the expected anti-c-Myc aptamer would bind to the target protein by two mechanisms - primary aptamer binding and decoy. The combination of primary binding site
of the aptamer to c-Myc:Max and the interaction between the decoy component with
c-Myc:Max will make the affinity of the anti-c-Myc aptamer higher than only with one
primary binding site or only with the E box.

The data we obtained suggest that the single-stranded DNA aptamers prepared by this
“decoy” method, including full length and truncated sequences, can bind to the c-Myc:Max
heterodimer with very high affinity (Kd’s from 100-500 nM) and specificity for the target
protein, but can not compete with the E box sequence to bind to c-Myc:Max. It was interesting
to find that none of the single-stranded DNA molecules we selected included E box sequence,
which means that the selected DNA molecules bind to c-Myc:Max at some other sites instead
of an E box binding domain.

There are several directions we could go to optimize the current selected DNA
molecules so as to develop anti-c-Myc aptamers. First, we could gradually truncate the current
aptamers to determine the minimum sequences for each aptamer. Second, an extra E box
sequences could be linked to the DNA sequences by rational design to increase the binding
affinity. Third, some random mutations could be introduced into the previously selected DNA
sequences by doping. Doping means that when the candidate aptamers were synthesized, the
percentage of the four bases, A, T, C and G, is not evenly distributed among each position, so
some random variations will be added to the original aptamer sequences. By introducing the
new mutations, certain members of the degenerate aptamer pool could have higher binding
affinity to target protein than the starting aptamer.
INTRODUCTION

Thesis organization

My work was mainly focused on the development of single-stranded DNA aptamers for c-Myc:Max heterodimer, including preparation of the active protein targets, selection of ssDNA aptamers against c-Myc:Max using SELEX (Systematic Evolution of Ligands by EXponential enrichment), checking the binding ability of selected oligonucleotides to bind c-Myc:Max by EMSA (electrophoretic mobility shift assay). The work that I focused on was part of a collaborative project. Another main participant was Marjan Mokhtarian. Both of us did two SELEX experiments in parallel. Marjan Mokhtarian's work also included testing the binding affinity of selected oligonucleotides using a filter capture assay and estimating the Kd of these molecules by the filter capture assay.

In the “introduction”, I discuss the c-Myc:Max protein heterodimer including its structure and various functions in vivo including the effects on cell cycle progression, genomic stability, apoptosis and cellular metabolism etc. Following that, I introduce some characteristics of the aptamers, single-stranded DNA or RNA molecules that can bind various targets with high affinity and specificity. A comparison of aptamers and antibodies is discussed to demonstrate the advantages of aptamer molecules over antibody. The applications of aptamers in diagnostics and therapeutics are also discussed. The third part of the introduction discusses the establishment and development of SELEX, the method we used to isolate the c-Myc:Max aptamer.

Subsequent section in the thesis are “Materials and Methods” followed by “Results and Discussion”, then “Conclusions and Future directions”. The next section is the
“Appendix”, which contains some early works related to the project including some data of subcloning of the c-Myc and Max cDNAs, protein expression and purification, and EMSA. Finally, the references used in the thesis are listed.

The structure and function of c-Myc:Max

A critical member of the myc proto-oncogenes, c-myc was discovered in the chicken as a homologous gene of the v-myc oncogene, which is from a retrovirus that can induce myeloid leukemia, sarcomas, liver, and kidney tumors in the chicken (Dang 1999). The c-myc gene is widely distributed in organisms, ranging from human to mice, birds, amphibians, fish and Drosophila (Littlewood 1998). Besides c-myc, the myc gene family also includes B-myc, L-myc, N-myc and s-myc. Only c-myc, L-myc and N-myc have oncogenic potential (Dang et al. 1999).

The Myc protein, known as one of the central pillars of carcinogenesis, belongs to bHLH/Z proteins in structure. The structure of this protein family is shown in figure 1. Myc, as a DNA binding transcription factor, contains a helix-loop-helix (HLH) motif at its C-terminus, which is a dimerization domain that can mediate heterodimerization of Myc with Max (Myc association factor X). Max is another HLH protein that can play other important functions in vivo. For c-Myc and Max, there is a short region consisting of basic amino acid residues N terminal to the HLH domain that mediates the specific interaction between the c-Myc:Max protein and its target DNA sequence (CACGTG), which is called an E box. The critical amino acids in the basic region that bind to the E box include glutamic acid (the 9th residue), which contacts the CA of the E box, and an arginine residue that interacts with the internal G of the E box sequence. Besides the HLH and basic region the transcription factor heterodimer contains
another dimerization domain immediately adjacent to C-terminus of the HLH domain, which is the leucine zipper (Littlewood 1998).

The N-terminus of Myc protein also plays important roles in transcriptional regulation of Myc. There exist two conserved sequences among the Myc family protein that are termed MBI and MBII (Myc box I and II). MBI is the glycosylation and phosphorylation site of Myc. The first 143 amino acids at the N-terminus are necessary for Myc:Max binding to its target DNA sequence and transcriptional activation, so it is called the transcriptional activation domain (TAD) (Kato et al. 1990).

Several important factors affect the interaction between the Myc-Max heterodimer with the E box. These include: 1) the flanking sequences immediately before and after CACGTG. Generally G/C dinucleotides predominate over other nucleotides on both sides of the E box sequence; 2) Methylation of CpG of the E box. When the cytosine at the CpG motif is methylated by DNA methyl-transferase, it will inhibit Myc binding to the E box in vitro and in vivo; 3) Other transcription factors that could affect the heterodimer’s binding (Grandori et al. 2000).

When regulating target genes’ transcription, the N-terminal of Myc protein interacts with a variety of proteins involved in transcription. TRRAP is one of earliest found cofactors of Myc:Max, which binds to the TAD of Myc. After binding to the N-terminus of Myc, TRRAP recruits histone acetylase GCN5 and histone acetyl transferase Tip 60 to the promoter of Myc target genes, which may change the structure of chromatin to recruit more other transcription factors including TFIIE or TBP to drive the Myc target genes' transcription (Brough et al. 1995; McMahon et al. 2000).
Besides Myc, the Max protein also can bind to other proteins in HLH family. Max-interacting bHLH proteins include members of the Mad protein family, Mnt (or called Rox) and Mga. Like Myc, these Max protein partners also interact with Max by the HLH domain, but act as repressors and promote cell terminal differentiation instead of cell proliferation by binding the E box sequences. The Max protein itself can form homodimers and repress target gene transcription by binding the same DNA sequences as do other Max heterodimers. So Max, as a stable (t 1/2 ~ 24h), ubiquitously expressed protein can dimerize with various regulated protein partners to form different complexes that either activate or repress specific target gene transcription and control a series of cellular behaviors (Grandori et al. 2000).

Based on the studies of the last two decades, it is known that, in normal cells, the c-Myc gene expression is tightly regulated in response to diverse signals including growth factors, cytokines and mitogens (Dang 1999; Grandori et al. 2000). The expression of c-Myc in normal cells is strongly correlated with cell growth and proliferation. Homologous recombination studies showed that deletion of both alleles of the c-myc gene can cause mouse embryos to die between 9.5 and 10.5 days of gestation (Davis et al. 1993), which proved that normal regulated c-Myc expression is required for embryonic development. Another experiment showed the consequence of a remarkable prolongation of cell doubling time in immortalized rat fibroblasts with c-myc inactivation, which demonstrated that Myc is involved in the mechanism of controlling cell and tissue growth and proliferation (Mateyak et al. 1997).

On the other hand, the c-Myc gene is overexpressed in many cancer cells in an
uncontrolled way and causes various pathological changes, including lung carcinoma, breast carcinoma, cervical carcinoma, ovarian carcinoma and a variety of hematological tumors (Dang et al. 1999). Two main mechanisms have been found that cause abnormal expression and/or regulation of c-Myc protein. One is chromosomal translocation. The c-Myc gene is translocated from chromosome 8 to a position downstream of the promoters of one of three immunoglobulin genes that are on chromosome 2,14 or 22 and is then constitutively activated (Boxer and Dang 2001; Dang et al. 1999). The other reason for Myc activation is due to point mutations that occur in the region around its phosphorylation and glycosylation sites in the transactivation domain at its amino terminus. These mutations affect normal modifications in the N terminus, abolish negative regulation of c-Myc activity and prolong the half-life of the onco protein (Dang et al. 1999).

Deregulated expression of Myc has significant effects on several related cellular behaviors by activating or repressing specific target genes. These cellular activities include cell cycle progression, differentiation, genomic stability, apoptosis and cellular metabolism (Dang et al. 1999). The effects of Myc on these cellular behaviors may be linked to its oncogenic activity.

For cell cycle control, constitutive expression of c-Myc mainly drives cell cycle progression from G1 to S phase (Karn et al. 1989). A series of important molecular events occur in this period. The best identified event is the inactivation of the Retinoblastoma protein (pRb) through phosphorylation by cyclin-dependent kinases (CDKs) (Roussel 1997). The phosphorylated pRb dissociates from transcription factor E2F proteins, following which the released E2F proteins drive their target genes' expression. The target genes include cyclin E
and A, which are required for S phase entry. Myc, as a positive regulator of CDK, can increase cyclin E/CDK2 or cyclin E activity via three distinct pathways. The first is functional inactivation of p27 to prevent p27-mediated growth arrest (Amati et al. 1998; Dang et al. 1999). P27 is one of the main anti-proliferation signals. Constitutively expressed Myc indirectly antagonizes the function of p27 and makes it dissociate from cyclin E/CDK2 (Amati et al. 1998). Studies using Myc-estrogen receptor (MycER) chimeras indicated another way that Myc up-regulates cyclin E expression. The activation of Myc in this study increased cyclin E expression at the transcription level. No other factors were involved, which suggests that cyclin E gene may be a direct target of Myc. But further evidence is needed to support this hypothesis (Amati et al. 1998). The third pathway of cell cycle control by Myc is through phosphatase Cdc25A. The Cdc25A gene is positively regulated by Myc. Cdc25A dephosphorylates and activates CDK2. Active CDK2 drives downstream events of cell cycle progression from G1 to S phase (Amati et al. 1998; Dang et al. 1999). Colony stimulating factor (CSF-1) is required for some cell lines to enter into S phase. Myc is an immediate early response gene of CSF-1. CSF-1 needs to bind its receptor (CSF-1R) to initiate the mitogenic response via different pathways. Roussel et al., (1997) used NIH-3T3 fibroblasts to show that the tyrosine at position 809 in CSF-1R plays key roles to keep cells surviving and proliferating. Engineered human CSF-1R point mutant (CSF-1R Y809F) expressed in NIH-3T3 fibroblasts can produce a considerable reduction in Myc expression, which suggests that Myc is tightly linked to CSF-1-dependent proliferation (Roussel 1997).

Besides transcriptional activation by Myc of specific target genes, Myc also promotes cell proliferation and oncogenesis via transcriptional repression of some cell cycle/growth...
inhibitors, especially repression of transforming growth factor-\(\beta\) (TGF-\(\beta\)) -induced growth arrest. These cell cycle arrest genes include p15, p21, p27, gas1 (growth arrest specific1), Gadd (growth arrest and DNA damaging-inducible) genes (Gartel and Shchors 2003; Wanzel et al. 2003). Basically constitutive expression of Myc represses gene transcription \textit{via} two different mechanisms. One involves the Myc-Max heterodimer and the Inr (transcriptional initiator) element. In this model the Myc-Max heterodimer interacts with Miz-1 (a zinc-finger protein and a partner of Myc) or another transcriptional activator through the c-terminal domain of Myc; the Myc-Max-Miz-1 complex binds to the Inr element of the target genes to suppress their expression. The other mechanism is Sp1-instead of Inr-dependent. In this model Myc interacts with Sp1 only or with active Smad2/3 and Sp1 to form an inactive stable complex of Sp1-Smad-c-Myc through the central domain of Myc; it is not necessary for Myc to bind Max and the DNA element (Gartel and Shchors 2003). The cell cycle inhibitor P15 gene is a good example that is repressed by Myc through both mechanisms. P21 is a CDK inhibitor that is suppressed by c-Myc \textit{via} an Inr-independent mechanism. Microarray analysis showed that P21 is a direct transcriptional repression target of c-Myc. The CDK inhibitor P27 is repressed by c-Myc through the c-Myc-Max heterodimer binding to the Inr element of the p27 promoter (Gartel and Shchors 2003). The c-Myc also represses transcription of the Gas1 gene although the molecular mechanism is not well understood. It has been reported that the Myc box2 (MBII) is required for this repression (Lee et al. 1997). It has also been shown that GADD 45 is suppressed by c-Myc through its binding to the GC-rich binding site of the GADD45 promoter \textit{via} the Sp1 Inr-independent pathway (Wilson 1997). So in conclusion, loss of transcriptional repression of cell cycle/growth arrest genes by Myc is a very important
way by which Myc promotes proliferation and the neoplastic phenotype.

Tumorigenesis induced by the overexpression of Myc is also reflected by genomic instability. Transient excess of Myc can increase the occurrence of the neoplastic phenotype of Rat1A cells by 50-fold and that this phenomenon is accompanied by destabilization of the cellular genome (Dang et al. 1999; Felsher and Bishop 1999; Mai et al. 1996). The genomic instability is reflected by multiple chromosomal abnormalities, such as aneuploidy, double minute chromosomes, dicentric chromosomes and multicentric chromosomes. The authors demonstrated that Myc may induce genomic instability by destroying the G1/S check points for DNA damage thus allowing the cell cycle to continue before the cells can adequately respond to the damage. This can promote the accumulation of DNA damage in the cell (Felsher and Bishop 1999). Genomic instability is also found in NIH 3T3 cells (Felsher and Bishop 1999) and mammary gland tumor cells of transgenic mice when c-Myc is overexpressed (McCormack et al. 1998).

Myc overexpression was also found to link to programmed cell death or apoptosis. The relationship between c-Myc and apoptosis was first found in 32D.3 myeloid progenitor cells, (Askew et al. 1991). Other studies showed that overexpressed c-Myc in Rat-1 fibroblast (Hermeking and Eick 1994), mouse mammary gland cells (McCormack et al. 1998) or activated myc-ER in mouse primary embryo fibroblast (Wagner et al. 1994) could lead to apoptosis. On the other hand c-Myc overexpression induces immortalization of primary rodent fibroblasts (Land et al. 1983; Simm et al. 1994). The main mechanism by which c-Myc induces immortalization is by driving the expression of telomerase hTERT, which maintains the length of telomeres and causes cell immortalization. The c-Myc binding sequence (E box)
was found in the promoter of the hTERT gene, which suggests it could be directly activated by c-Myc (Horikawa et al. 1999; Wu et al. 1999).

Deregulated expression of c-Myc also affects cell metabolism as reflected by changing energy consumption, DNA metabolism and translational regulation. In tumor tissues, cancer cells usually utilize glucose as main source of energy instead of oxygen and they produce lactic acid. This phenomenon is termed the “Warburg effect” (Dang and Semenza 1999). The lactate dehydrogenase-A (LDHA) gene is tightly related to the “Warburg effect”, and its activity is regulated by c-Myc (Tavtigian et al. 1994). Also c-Myc-related transformation requires LDHA overexpression (Shim et al. 1997). Besides LDHA, other c-Myc-responsive genes include the ornithine decarboxylase (ODC) (Packham and Cleveland 1995), dihydrofolate reductase (Mai and Jalava 1994) and thymidine kinase (Pusch et al. 1997) genes. The enzymes encoded by these genes participate in DNA metabolism, which makes it easier to understand c-Myc protein mediated cell cycle progression from G1 to S phase, because DNA synthesis occurs in S phase. The effects of c-Myc on cell metabolism also include the regulation by c-Myc of several translational initiation factor encoding genes including the eIF-2α gene (Rosenwald et al. 1993), Eif-4E (Jones et al. 1996) and ECA39 (Benvenisty et al. 1992).

The deregulated c-Myc gene expression in various cancer cells suggests that this oncogene could be regarded as a potential target of cancer therapy. Traditional approaches of inhibiting oncogene activity have some limitations, such as toxicity due to long time inhibition of associated proto-oncogenes, and the potential of tumor regrowth caused by cessation of the pharmacologic inhibition of target oncogenes. Jain and colleagues used the
tetracycline-regulated system to control the expression of Myc in transgenic mice and used doxycycline as an inhibitor to stop Myc expression in osteogenic sarcoma cells transplanted into syngeneic mice (Jain et al. 2002). The data showed that within twenty-four hours of treating with doxycycline, osteogenic sarcoma cells differentiated into mature osteocytes. What's more, after Myc expression in osteocytes was reactivated by removing the inhibitor, the cells did not restore their malignant properties but went into apoptosis. So even a brief inactivation of c-Myc (24 hours) could result in a sustained regression of tumors in mice.

Strategies of inhibiting c-Myc's activity in cancer cells could focus on the gene, mRNA or protein levels. Triple-helix forming oligonucleotides (TFOs) inhibit gene transcription by blocking the binding of transcription factors to the promoter of target genes. TFOs include GA-rich or GT-rich DNA oligonucleotides and can bind to double-stranded DNA in the major groove. The TFO approach was used to prevent the binding of the transcription factors E2F and MAZ to the c-Myc gene's promoter and effectively decreased the c-Myc gene's transcription (McGuffie et al. 2000).

Quadruplex-forming oligonucleotides are an alternative strategy to inhibit the c-Myc gene's transcription. The sense-strand up-stream of the c-Myc gene promoter contains a G rich region that can form a G-quadruplex structure when separated to a single strand. A Quadruplex-forming oligonucleotide could compete with the G-rich region in the c-Myc gene's promoter for binding to important factors, such as hnRNP K (heterogenous nuclear ribonucleoprotein K) and inhibit recruitment of the pol II complex to the promoter. The effective concentration of the oligonucleotide was in the nanomolar range (Hurley 2001).

Anti-sense DNA and RNAi provided a possibility of inhibiting c-Myc gene's
expression at the mRNA level. Several groups used anti-sense DNA or RNAi with different cell lines to reduce the c-Myc gene's overexpression and block cell proliferation. For example, Leonetti and colleagues utilized phosphorothioate c-myc antisense oligodeoxynucleotides to treat transgenic mouse with c-Myc overexpression and resulted in a prolonged decrease in c-Myc expression, reduced tumor growth and increased survival periods of treated animals (Leonetti et al. 2001). Brummelkamp et al. (2002) used short RNA interference to stably and specifically block the expression of the oncogenic \( K-RAS^{V12} \) expression in human pancreatic carcinoma resulting in the loss of anchorage-independent growth and tumorigenicity (Brummelkamp et al. 2002).

Another potential method to block the c-Myc transformation function \textit{in vivo} is to interfere with the formation of the c-Myc:Max heterodimer. Berg et al. (2002) used the BR-HLH-LZ domain of c-Myc and Max to make the fusion proteins c-Myc-CFP and Max-YFP. When the two fusion protein dimerize, fluorescence energy transfer will occur between CFP and YFP. CFP and YFP are derivatives of the green fluorescent protein GFP. The authors screened around 10,000 peptido-mimetic substances and found four that could interfere with the dimerization of c-Myc and Max as indicated by the difference in fluorescence emission (Berg et al. 2002).

In our project, we proposed to develop a bivalent aptamer that can bind to two sites on the c-Myc:Max protein: 1) the c-Myc:Max E box binding domain and 2) another site on the heterodimer in order to provide additional specificity and higher affinity for the c-Myc:Max target protein.

An obvious advantage of an aptamer compared to other anti-cancer drugs is that it can
be developed into an allosteric aptamer or bivalent aptamer, this means two different aptamers recognizing different targets could be linked together and control each other’s function (allosteric) or function simultaneously. For example the c-Myc:Max aptamer could be linked to another aptamer that can specifically bind to an anti-cancer drug, such as Gleevec. If the aptamer was allosteric and only allowed Gleevec binding when it bound c-Myc:Max, the binding of the c-Myc:Max aptamer to c-Myc:Max would allow the anti-Gleevec aptamer to drive more Gleevec into the cells with c-Myc:Max overexpression, therefore increase the local concentration of the anti-cancer drug.

If the two strategies above could be combined, it will be much more powerful than a single treatment. So the c-Myc oncogene had been regarded as a target of cancer therapy and recent published results have shown that different strategies can be used to inhibit the cellular effects of deregulated Myc gene expression.

**Aptamer**

Aptamers are single stranded DNA or RNA developed *in vitro* from a random pool to execute certain functions. Aptamers can tightly and specifically bind to their target molecules. The targets of aptamers range widely from small molecules, such as ATP, metal ions etc. to macromolecules, such as peptide, protein, even whole cells.

The basic procedure for developing aptamers is termed SELEX (Systemic Evolution of ligands by Exponential Enrichment). SELEX starts with a large random single strand DNA or RNA pool including $10^{14}$ - $10^{15}$ molecules; each molecule has a random region, which is usually from 30-40 nucleotide long (Ellington and Szostak 1990; Tuerk and Gold 1990). On the both sides of the random sequences there are fixed sequences for PCR primer annealing;
for RNA molecules the 5 prime includes the T7 promoter for T7 RNA polymerase binding. The critical step of SELEX is partition from those nonspecific binders of ssDNA or RNA molecules that can specifically bind to targets. Several methods can be used to realize the partition. Examples are nitrocellulose filter binding, EMSA, and affinity chromatography. The selected pool then can be amplified by PCR (polymerase chain reaction). When selecting from an RNA pool, it is necessary to synthesize a complementary DNA strand by reverse transcription and then do PCR to amplify. In order to develop new molecules low fidelity PCR is necessary to introduce some mutations in the selected population. A new single stranded library can be generated by separation of the double stranded DNA PCR product for DNA or by transcription for RNA molecules. The same selected procedure will be repeated until the binding efficiency is up to the expected percent (Jayasena 1999).

One of the critical strategies of SELEX is to determine the ratio of target molecules to selected pool. Another concern during SELEX is to avoid accumulation of those molecules that can bind the support materials (such as filter or column) and related molecules or analogs of the true target. So sometimes negative selections against the support materials or related molecules are necessary. The strategy of negative selection is based on the high specificity of aptamers. It was reported that aptamers can discriminate subtle structural differences between the true target and other molecules that differ by features such as a methyl group (Haller and Sarnow 1997; Jenison et al. 1994) or hydroxyl group (Geiger et al. 1996; Sassanfar M 1993).

The number of rounds of SELEX depends on the target property and selection pressure. Usually 8-15 cycles are required to reach binding efficiency saturation. After the binding efficiency is saturated, the “winner” molecules in the enriched library are cloned and
sequenced. Further analysis is necessary to identify the true aptamers. This includes sequence alignment, affinity analysis and determination of the shortest truncated sequence of aptamers etc. The Kds of aptamers to protein targets is usually range from pM to nM (Gold et al. 1995). Kds for other molecules range from nM to μM. For example, the Kd of theophylline is \(~0.1\mu\text{M}\) (Jenison et al. 1994) and of dopamine is \(~2.8\mu\text{M}\) (Mannironi et al. 1997). Generally the whole procedure including positive selection, negative selection, cloning and sequencing takes 2-3 months.

**Comparison of aptamer and antibody**

The use of antibodies for analysis began in the 1950s and was widely applied in the 1970s. From then on antibodies have shown tremendous power in diagnostic applications based on molecular recognition.

In the 1990s, the first experiment of SELEX (Systematic Evolution of Ligands by Exponential Enrichment) provided a method to develop a different class of molecule-single stranded DNA or RNA termed an aptamer, which could be selected *in vitro* and possess some properties and advantages that antibodies do not have. These include:

A) Antibody identification begins with animals so it is limited to molecules that can be tolerated by animals. This limit can be overcome by the SELEX procedure of aptamer identification. Aptamers are developed in an *in vitro* environment instead of *in vivo* conditions. So aptamer selection does not rely on animals or cultured cells.

B) Because of being prepared using an *in vivo* system antibody preparation is restricted by parameters *in vivo*, such as pH, temperature, salt concentration etc. So the interaction of antibody and target occurs optimally under physiological conditions. While for
aptamer selection, people can change the selection conditions \textit{in vitro} including pH, buffer components, temperature etc. depending on the diagnostic or imaging purposes.

C) The properties of the antibody can vary from batch to batch because a different animal is used each time. Aptamer molecules that are synthesized by a chemical procedure \textit{in vitro} can be reproduced with a high degree of accuracy. It is also very easy to purify DNA or RNA aptamer molecules \textit{in vitro} condition without changing their properties.

D) Stability is a big challenge for protein antibody storage. Proteins are very easily denatured when stored dilute and at low temperatures, or when heated. Aptamers can endure a wide range of temperatures and can very quickly renature (Jayasena 1999).

\textbf{Applications of Aptamers}

The high affinities and specificities of molecular binding and recognition of aptamers as well as the possibility for aptamers to be selected, designed and modified make the single stranded DNA or RNA molecules potential powerful tools in diagnostics and therapeutics.

\textbf{Diagnostics}

In the traditional two-antibody sandwich diagnostics protocol, a monoclonal antibody for the protein of interest is captured on a solid surface; the target protein is then incubated with the monoclonal antibody. In order to improve the signal-to-noise ratio, a second antibody is added to the detection system of the bound protein. The second antibody can be detected by a detection signal of enzyme, biotin, fluorophores or radioactive isotopes (Jayasena 1999).

Since methods were developed for the \textit{in vitro} development of aptamers, the new reagent gradually plays important roles in protein diagnostics assay. The following properties of aptamers suggest that aptamers could be a good substitute for an antibody. First, aptamers,
as single stranded DNA or RNA molecules, can be selected *in vitro* by the SELEX procedure from a random pool, which only requires separation of bound and unbound molecules and PCR amplification. With the construction of a robot that can execute SELEX, the procedure becomes automatic and efficient. Second, the method of photo-SELEX can lead to cross linking between the aptamers and their targets and does not affect the binding between them. So the binding ability and cross-linking ability are doubly advantageous for aptamers that can make them excellent reagents for diagnostics (Brody and Gold 2000; Brody 1999).

A. Aptamer-based sensors

Biosensors have been developed based on recognition between biomolecules, such as proteins and nucleic acids, which allow rapid and selective detection of targets through transducers. Aptamer-based biosensors have obvious advantages compared to protein-based immunosensors that make aptamers powerful tools in diagnostics. First, aptamers can be modified to be attached to a surface support; the interaction between biotin and straptavidin can be utilized to fix aptamers to a proper support or the aptamer can be modified by the addition of one of several chemical groups and then linked chemically to the support. The immobilized aptamer on a sensor surface can be regenerated to be functional for repeated uses. A second advantage of aptamers is their ability to be labeled with a wide range of reporter molecules that can report the binding reaction of aptamer and target (Jayasena 1999).

In one application, an anti-thrombin DNA aptamer was used to selectively detect thrombin in solution by evanescent-wave-induced fluorescence anisotropy. This was the first attempt to develop an aptamer-based biosensor. The author labeled the 5 prime end of the DNA aptamer with FITC (fluorescein isothiocyanate) and covalently immobilized the 3 prime
end to a glass surface via a modified amine. It was found that the fluorescence anisotropy of
the immobilized thrombin aptamer increased as a function of the thrombin concentration and
that the Kd was 1.1μM, which is 10 times higher than that of the unlabeled thrombin aptamer.
This biosensor had a very low detection limit for thrombin of as little as 0.7 amol in 140pL
volume (Potyrailo et al. 1998).

The same strategy was used with a fluorophore-labeled PDGF (Platelet-derived
growth factor) aptamer to detect the PDGF target in real time and in homogeneous solution.
Data showed that this method was selective and sensitive; it could detect PDGF in the
subnanomolar range. The significance of this assay is that it is the first application of aptamers
to detect the growth factor PDGF using fluorescence anisotropy measurements (Fang et al.
2001).

In another application, a RNA aptamer selected to specifically recognize the protein
trans-activator of transcription (Tat) of HIV-1 was utilized as the bio-recognition element of a
biosensor. The RNA\textsuperscript{TAT} aptamer showed 133 times higher binding affinity than the natural
HIV-1 trans-activation response element (TAR) RNA. In this work, the RNA aptamer was
immobilized on the sensor surface through the interaction between biotin molecules attached
to the aptamers and streptavidin fixed on the sensor support. This assay could detect 0.65ppm
(mg/L) of target protein. Results also demonstrated a high specificity of the RNA aptamer for
the specific protein (Tat) compared with other proteins (Minunni et al. 2004).

B. Aptamer in molecular beacons

Molecular beacons are stem-loop structure probes. A fluorophore and a quencher are
linked to each terminus of the stem. The stem structure brings the fluorophine and quencher
close together and the fluorescence is quenched by the fluorophore/quencher pair. When the loop portion of molecular beacon interacts with its complementary sequence, such as part of an aptamer molecule, it will open the hairpin structure of the molecule and drive the fluorophore away from the quencher, so the fluorescence is turned “on” (Li et al. 2002).

In one model of the interaction between aptamer and molecular beacon, part of the aptamer sequences is complementary with the loop structure of molecular beacon only when the aptamer is free from its target protein; so when the target-free aptamer hybridizes with molecular beacon, the conformation of molecular beacon will be changed and the termini of the stem portion will be separated, which results in emission of fluorescence. However when the target protein of aptamer is present in the system it will bind to the aptamer and block the site required to interact with molecular beacon, so the fluorophore/quencher pair will not be separated and there will be no emission of fluorescence signal (Li et al. 2002). This strategy has been used to detect the existence of proteins in buffer and in plasma for which a high-affinity and specificity aptamer had been selected (Jayasena 1999).

Another application related to aptamer and molecular beacon is a construction of the thrombin DNA aptamer beacon to detect the protein thrombin. In this design, the length of the core sequence of the thrombin aptamer was extended from a 15-mer to a 17-mer by adding a thymidine nucleotide at both ends of the aptamer; then a fluorophore and a quencher were separately conjugated to the 5’ and 3’ end of the 17mer to get the fluorophore-quencher-labeled molecular aptamer beacon (MAB). When the thrombin MAB is free from its thrombin target, it exists in equilibrium between a random coil structure and a compact quadruplex (including two G-quartets structure). Because in the absence of thrombin
some parts of the MABs are in the nonstructural random state, the 5' end conjugated fluorophore (6-FAM) will give very high fluorescent signal. When the thrombin target is present in the system, it will bind to MAB and shift the equilibrium to the quadruplex structure. This will drive the fluorophore and quencher at the ends of the MAB closer, resulting in quenching of the fluorescence signal. Similar to the previous sensor, this sensor utilized the signaling characteristics of the molecular beacon and the high specificity of target recognition of the aptamer to provide a powerful analysis method for real time detection of a protein in vitro (Li et al. 2002).

C. Microarray based on aptamers for protein analysis

It is useful to study gene products at the protein level instead of the DNA or RNA level because regulation during translational and posttranslational modification is not accurately reflected in DNA or RNA levels. Using an aptamer array to understand proteomics has some advantages because of the following properties of aptamers: 1) Aptamers can be quickly obtained by chemical synthesis with high accuracy; 2) Aptamer molecules can be fixed in a specific location on a solid support at the required density to make DNA microarrays; 3) UV cross-linking can generate irreversible links between aptamers and target proteins, which stabilize the specificity provided by affinity; 4) The range of targets for aptamers is wide, ranging from ion, small organism to peptides, proteins, viruses even tissues (Jayasena 1999).

Microarrays based on aptamers can be used as robust tools to analyze protein expression and can provide an important technology for disease diagnostics and development of new therapeutics. Collett and colleagues described a method of producing aptamer-based microarray slides. The authors used this aptamer chip array as a tool to characterize the
binding properties of isolated aptamers to their corresponding target proteins. In this assay twelve selected chicken egg white lysozyme RNA aptamer candidates were biotinylated and printed on streptavidin slides. The slides were treated with decreasing concentrations of a fluorescent target protein. Binding of the aptamers to the target protein at each concentration could be detected by the fluorescence intensity of the spots on the slides when the slides were scanned. The detection limit was 1pg/ml. The aptamer-based microarrays provide a high-throughput strategy to test the aptamer binding affinity (Collett et al. 2005).

The same group also developed an aptamer-based microarray for detecting multiple protein targets. For this assay the authors generated a multiplex aptamer microarray for four different aptamers and their individual target protein including anti-lysozyme RNA aptamer, anti-ricin RNA aptamer, anti-lgE DNA aptamer and anti-thrombin DNA aptamer. The detection limit for each protein target was 5pM for lysozyme, 0.5nM for ricin, 10 pM for lgE and 5nM for thrombin, respectively. Each RNA and DNA aptamer showed a specific response to their target protein and not others (Eun Jeong Cho 2006).

D. Design and application of allosteric nucleic acids based on aptamers

An allosteric effect results when an allosteric regulator binds to a different site than the active site, and thereby induces a structural rearrangement of the molecule that influences the activity of the active site. Based on the concept of allosteric mechanisms several kinds of aptamer-based allosteric nucleic acids were designed. The first one is an allosteric ATP aptamer, which is based on the TRAP (“targeted reversible attenuated probe”) design. In this, the DNA sequence includes an ATP-DNA aptamer followed by an antisense sequence and an attenuator segment. In the absence of regulatory nucleic acid (reg NA), which is
complementary to the antisense sequence, the ATP aptamer part is attenuated by the attenuator therefore unable to bind to its target. When it is present, the reg NA will hybridize with the antisense sequence and separate the attenuator from the aptamer. The released aptamer can bind to its target. So in the TRAP model, the aptamer activity is regulated by hybridization between antisense and reg NA and the antisense sequence and aptamer part play roles of recognizing and signaling (Cong and Nilsen-Hamilton 2005).

A similar design was also used in allosteric ribozyme. Data showed that the activity of hammerhead ribozyme could be increased by around 250 fold in the presence of regulatory sequences (Burke et al. 2002).

CLAMP (cis-linked aptamers for medical and microanalytical procedures) is another design based on an allosteric mechanism. The allosteric CLAMP includes two cis-linked aptamers; the binding of target to one of the aptamers can regulate the activity of the other one. The advantage of the CLAMP structure is allowing the regulator or the target to be any kind of molecule as long as the proper aptamer can be obtained. A good example of CLAMP is currently under development. It consists of an ATP aptamer and a neomycin aptamer in one sequence. The binding of neomycin to neomycin aptamer increases binding of the ATP aptamer to ATP (Stodola 2003).

Another allosteric aptamer design shares some similarity to the CLAMP. In this approach, an allosteric RNA aptamer can bind RVV-X (Russell’s viper venom factor X activator) and VEGF165 (human vascular endothelial growth factor). RVV-X has an enzymatic activity to trigger the reaction of coagulation cascade of blood components that could be used as a detectable signal. When VEGF165 is absent, the RNA aptamer has high
affinity for RVV-X and blocks its enzymatic activity, so no signal can be detected. The inhibitory effect is reversed by binding of the RNA aptamer to the effector molecule-VEGF165. Thus, the VEGF165 functions as a controller for the on/off switch of RVV-X enzymatic activity. The potential mechanism behind the effect is when the domain on the allosteric RNA aptamer which binds to VEGF165 is occupied by the effector; the domain for binding RVV-X on the aptamer is changed somehow, so the binding affinity of RVV-X is decreased. The difference between this allosteric aptamer design and CLAMP is the binding of one molecule to the aptamer decreases instead of increasing the binding of the other molecule (Chelyapov 2006).

Based on the previous studies, RNA and DNA molecules can be designed to make functions as molecular switches in the presence of specific target molecule. Compared to DNA, RNA molecules have more flexibility to form proper structures needed for binding different targets. Breaker et al. (2002) first discovered and discussed the strategy of “riboswitch”. Three general strategies can be used to create engineered riboswitch: rational design, in vitro selection (SELEX) and combined design (rational design and SELEX). One example is the construction of an ATP-dependent RNA riboswitch. The approach was achieved by linking a hammerhead ribozyme domain to an RNA domain that can bind ATP. The ATP binding could create corresponding changes in structure in the catalytic domain of the ribozyme and so it could play roles in catalytic reaction (Breaker 2002). Interestingly, people also found natural versions of “riboswitch” in vivo. In structure, two allosteric glycine-binding RNA subunits are arranged in tandem. The binding of a glycine molecule to subunit I triggered the structure changes of RNA subunit II and helped the second subunit to
bind to glycine. The function of the riboswitch *in vivo* is to regulate expression of a gene related to glycine dependent energy consumption (Famulok 2004).

**Therapeutics**

The therapeutic function of an aptamer is based on the aptamer’s ability to directly inhibit its target molecule's function by folding into a proper 3-dimesntional structure. Thus, the aptamer can specifically and tightly bind to the target with high affinity (Rebekah 2000).

At the beginning of exploration of aptamer function as therapeutic reagent, the Tat protein was regarded as the target molecule to inhibit because it is responsible for HIV-1 RNA transcription through its direct interaction with TAR RNA (HIV trans-activation response element) (Minunni et al. 2004; Rebekah 2000). A “decoy” RNA aptamer was developed that could bind to Tat protein and inhibit HIV RNA transcription. This RNA aptamer was also proved to be functional *in vivo*. This significant study opened the way for an RNA molecule (and later single stranded DNA) to be utilized as a therapeutic reagent by its ability to specifically bind target proteins and inhibit their functions.

Platelet-derived growth factor (PDGF) stimulates mesangial cell proliferation and cell matrix accumulation, which leads to cardiovascular diseases. Floege et al., (1999) utilized a nuclease-resistant and high-affinity DNA aptamer as an antagonist to PDGF and dramatically suppressed PDGF-induced cell proliferation. The high affinity and specificity of the DNA aptamer was reflected in the fact that, at the concentration of 1μg/ml, the anti-PDGF aptamer completely blocked rat mesangial cell growth and proliferation induced by PDGF but not by other growth factors, such as EGF or FGF. The Kd for the PDGF aptamer is around 0.1 nM (Floege et al. 1999).
Receptor tyrosine kinases (RTKs) are transmembrane proteins that initiate several important signaling pathways that regulate cell growth and differentiation in a variety of cancers. So RTKs were regarded as important targets for cancer diagnostics and therapeutics. Nuclease-resistant RNA aptamers that can specifically recognize the human RET (rearranged during transfection) RTK were isolated. The authors utilized RET RTK-expressing cells as the target during SELEX. The isolated RNA aptamer blocked RET-dependent signaling and related down stream molecular events and specifically bound its target. The $K_d$ of the aptamer is 30-70 nM (Cerchia et al. 2005).

The Raf-1 protein is another good therapeutic target for an aptamer. As a serine/threonine kinase in the cytoplasm, Raf plays important roles in transmitting signals of cell proliferation and development from plasma membrane and nucleus. Even though the molecular mechanism of signal transmission involving Raf-1 is not completely clear, it was known that the functional Raf-1 protein activity is tightly regulated by its interaction with Ras, which belongs to GTPases. The critical region of Raf-1 for its association with Ras is in 51-131 amino acid, which is named Ras-binding domain (RBD). The authors isolated and characterized RNA aptamers which can specifically recognize Raf-1 RBD (the $K_d$ is 152±23nM) and can efficiently interfere with the interaction between Raf-1 and Ras, thus the isolated anti-Raf-1 RNA aptamer could be used as a tool to regulate Raf-1 involved signal transmission pathway and as a potential candidate of Raf-1 related cancer therapy (Kimoto et al. 2002).

Several important issues needed to be considered for aptamers to be transformed from specific protein binders into potential therapeutic molecules in vivo. Generally, the size of the
aptamer should be shortened to 40 nucleotides or less from the original length of around 80-100 to get a minimum-sized aptamer with high specificity and affinity to target molecules. Also stability and systemic clearance of aptamer are two critical points for molecules to be effective in vivo. For RNA aptamers, the stability can be increased by selecting modified aptamers using 2'-amino or 2'-fluoro nucleotides to substitute for ribonucleotides. To avoid rapid systemic clearance, the small aptamers with molecular weight from 8,250 to 13,000 kd can be conjugated by polyethylene glycol (PEG) or attached to a liposome to increase molecular size (Rebekah 2000).

VEGF (vascular endothelial growth factor) is a growth factor that is tightly correlated with psoriasis, macular degeneration and tumor proliferation. Development of a reagent that can limit or inhibit VEGF function in vivo could provide a method to effectively cure VEGF-related disease. A 2'-fluoropyrimidine RNA aptamer NX1838 conjugated with PEG was selected in 1999. This RNA aptamer can specifically bind to VEGF165 and block some VEGF 165-mediated cellular events, such as calcium mobilization, signal transduction and cellular proliferation. NX1838 was the first aptamer-based therapeutic molecule approved for use in the human (Bell et al. 1999). This aptamer is now in clinical use under the name of pegaptanib or macugen (Ng and Adamis 2006; Tobin 2006).

A new potential way of aptamer application in therapeutics was generated for controlling gene expression via interaction between RNA aptamer and specific target molecules. In the mechanism of eukaryotic translation, a critical step in translational initiation is ribosome-mRNA interaction and 5'-3' scanning of ribosome from 5'-m7G cap to the start codon. Because an aptamer can be isolated to specifically bind to its target, it was tested that
whether a complex of aptamer-target in the 5' UTR (5' untranslated region) can repress gene translation by blocking ribosome scanning or by the interaction of ribosome and mRNA. This strategy was utilized in Chinese hamster ovary (CHO) host cells. The authors constructed a mammalian β-galactosidase expression vector with a copy of Hoechst dye 33258 aptamer in the gene’s 5’-UTR and a control vector without the aptamer. The constructed aptamer-containing or control plasmids were separately transfected into CHO host cells together with a luciferase reporter gene as an internal control. The results showed that when the target drug hoechst dye 33258 was absent, the expression level of β-galactosidase was similar for the aptamer containing expression vector and control vector, also expression of the luciferase internal control gene was not affected; but when the dye 33258 was present, the expression level of β-galactosidase from the expression plasmids that included either the H10 or H19 Hoechst aptamers in the 5' UTR was dramatically lowered by more than 90% compared with that of control plasmid without H10 or H19 aptamers in 5' UTR. Again luciferase expression was the same for all plasmid constructs (Werstuck and Green 1998). This strategy developed a translational switch by inserting an isolated aptamer sequence in the 5' UTR of a gene of interest to regulate translation by adding the specific target molecules into the in vivo system.

Grate et al., (2001) showed the same strategy could work in the yeast strain S. cerevisiae. The authors targeted cyclin B2 (CLB2) gene as the gene of interest. CLB2 controls cell cycle transition from the G2 phase to mitosis in budding yeast cells. A malachite green aptamer sequence was inserted immediately upstream of the start codon of CLB2 gene. It was shown that the expression level of CLB2 was depended on the presence or absence of TMR
(tetramethylrosamine), a malachite green analog that can bind to the malachite green aptamer tightly ($K_d \approx 40\text{nM}$). Western blot analysis showed that $1\mu\text{M}$ TMR lead to a more than 10-fold decrease of CLB2 protein expression, while the normal yeast strains without MG aptamer insertion in clb2 gene showed no difference before and after adding TMR. RT PCR and Western blot showed that the blocking effect of TMR on the expression level of CLB2 by the presence of the MG aptamer in the CLB2 gene did not occur at the transcriptional but at the translational level. 1-D NMR further demonstrated that the TMR ligand binding to the MG aptamer drove the secondary structure of the MG aptamer in the 5' UTR from a less stable state to a more stable state. This stable MG aptamer structure can inhibit movement of the ribosome along the mRNA and inhibit translation of the mRNA (Grate and Wilson 2001). These results showed that the binding of a small ligand to an aptamer in the 5'UTR could induce secondary structure conformational changes that could decrease the translational initiation rate. A strategy based on aptamer-target binding in the 5' UTRs of gene transcripts of interest in vivo could inhibit translation of the message and thereby provide a new way of using aptamers for therapeutic applications to inhibit specific disease-related gene expression.

**SELEX**

**The Establishment of the SELEX Method**

The first SELEX experiments were executed in Dr. Larry Gold’s lab in University of Colorado in 1990 (Ellington and Szostak 1990; Tuerk and Gold 1990). When the SELEX strategy was designed, the researchers in Gold’s lab worked on the translational regulation in T4 bacteriophage-infected *E.coli*. They found that genes 32 and 43 encode DNA binding proteins, and the protein products can negatively regulate their own expression by binding to
the translational initiation region of their corresponding mRNA.

It was known that the mRNA target for T4 DNA polymerase (the gene 43 encoded protein) includes a loop-stem (or hairpin) structure upstream of the AUG. People who were interested in the self translational repression of bacteriophage T4 DNA polymerase tried to figure out the sequences required for the protein-mRNA interactions by generating some mutations in the loop range of the mRNA. They thought a random pool of mRNA sequence could be an ideal source to provide the desired sequences. They started with a library containing $10^{14}$ to $10^{15}$ different mRNA molecules including 8 random sequences of each molecule to incubate with the T4 DNA polymerase target in 3 different ratios of RNA to protein (10:1, 100:1 and 1000:1). After repeated binding, separation, replication (RT PCR and PCR) and in vitro transcription, the aptamer candidate molecules were cloned and sequenced. The data showed that the SELEX experiment yielded two different groups of sequences in the loop region. One was the same as the wild type binding sequence; the other one included four mutations, and was named the “quadruple mutant”. The two groups of aptamers had similar $K_d$ (~4.8x$10^{-9}$ M). Based on these data, the researchers further imagined that single-stranded DNA or RNA aptamers could be developed for many kinds of molecular targets (Gold et al. 1997; Tuerk and Gold 1990).

At the same time when Dr. Gold’s lab isolated the RNA aptamer of T4 DNA polymerase, Dr. Szostak’s lab also began to develop RNA molecules that can bind to small molecules of several dyes via in vitro selection. The starting pool they used included roughly $10^{13}$ different RNA molecules and 100 random oligonucleotides for each molecule. They used affinity column chromatography to separate the specific binders and nonbinders. The $K_d$s of
aptamers specific for different target dyes ranged from 100 μM to 600 μM (Ellington and Szostak 1990).

The isolation of functional RNA molecules that can specifically bind to given targets suggested that a similar method could be used to provide single stranded DNA molecules that can properly fold into certain structures and specifically bind to corresponding targets. Ellington and Szostak (1992) used a single stranded DNA pool from which to select aptamers that bind several dye molecules including Cibacon blue, reactive green and reactive blue. They found that the DNA-target interactions are both sequence and target-specific. The $K_d$s were 30 to 50 μM (Ellington and Szostak 1992). From their experimental data, single-stranded DNA could be regarded as alternative potential diagnostic and pharmaceutical reagents (Ellington and Szostak 1992; Klug 1994)

**Development of SELEX**

*Genomic SELEX*

As people found more and more proteins playing roles in regulating gene expression through binding specific target DNA or RNA sequences, a new method derived from SELEX, which is termed “Genomic SELEX”, was developed. Genomic SELEX utilized the whole genomic sequences of a certain organism as the original pool instead of random sequences to isolate the specific and tight binding oligonucleotides for a target protein. To some extent, genomic SELEX could be thought a “global search” method to predict the potential networks of interaction between protein and nucleic acid *in vivo*.

How to establish a high quality starting genomic library that can represent the whole genome of a certain organism is a critical step for genomic SELEX. Singer et al., (1997)
reported an approach starting with random priming on sheared denaturing genomic DNA to construct the original pool. First the authors designed a 5 prime primer (A ran) and a 3 prime primer (B ran), each including a fixed sequence at 5 prime and a nine random sequence at 3 prime individually. They chose denatured human, yeast and *E.Coli* genome DNA libraries as the templates to be annealed to the 3 prime primer (B ran), which is expected to randomly distribute through the genomic DNA; then Klenow was used to perform the extension reaction. The 5’ primer (A ran) then was added to the system and the extended product above was used as a new template to anneal and extend again. The extended products should include various fragments of different lengths. These products were then resolved through a denaturing gel and purified via electro-elution. The purified eluted products were annealed to new primers, primer A with same fixed sequence as primer “A ran” but also with a T7 promoter at the 5 prime end, and primer B containing the same fixed sequences as primer “B ran” for amplification. So the resulting amplified double stranded products should contain inserts consisting of genomic DNA sequences of different lengths surrounded by 5 prime and 3 prime sequences. The authors also tested the distribution of the end-points (an end-point means the last genomic nucleotide in each fragment in the library) of a certain region of genomic inserts in the library and the sequence accuracy of the library. Data showed that the library “contains overlapping inserts starting at most of the positions within the genome”. In the tested genomic regions, the longest distance with no end-point was only 9 nucleotides long. The sequence fidelity of the tested library inserts was high compared to the published genomic sequence (Gold et al. 1997; Singer et al. 1997).

Besides random priming, several other methods were developed to construct genomic
libraries, such as mechanical fragmentation and blunt-end ligation (Sompayrac and Danna 1990), and restriction digestion and ligation (Kinzler and Vogelstein 1989).

Very similar to SELEX of random sequences that uses oligonucleotides to generate a diverse library, genomic SELEX generates a diverse library of sequences from genomic DNA. Both procedures involve separation of specific and nonspecific binders and non-bound molecules, amplification of the specific binding species, cloning and sequencing of the “winner” aptamer candidates.

In the very early stage of genomic SELEX application, a RNA library derived from \textit{E.coli} by random primer extension was used as the original pool and the MS2 coat protein was chosen as the target molecule because the \textit{in vivo} \textit{E.coli} RNA sequence required to bind MS2 was already known. Results showed that the fixed region in the primer sequences used to amplify the library participated in forming binding site of the selected aptamers. Also, when the library was constructed, the two primers included nine random sequences that may misanneal to the template, thus provided some extra mutations to the library. The mutated region also could be part of the binding sites of the target protein.

In order to avoid these unwanted situations, researchers replaced the fixed and random sequences in the primers with new fixed sequences selected via computer program (STOGEN) design, so the new fixed sequence should form none or very little secondary structure on their own or with genomic sequence region in the library. These changes were made so as to decrease the possibility that the primers would participate in formation of potential binding sites for target protein.
**Auto SELEX**

The normal SELEX procedure is repetitive and time consuming. If an average of 12 rounds of selection is required to select an aptamer, it will take around two months. Sometimes for some steps, such as PCR and PCR product purification of selection pool need to be repeated. Thus the whole procedure of SELEX will take a longer time than generally expect (Cox and Ellington 2001; Cox et al. 1998).

Cox et al., (2002) automated the *in vitro* selection procedure using a specially modified Beckman Biomek 2000 robot. They integrated a thermal cycler for selected pool amplification, a magnetic bead separator for separation of target-nucleic acid complexes and free nucleic acid, an enzyme cooler and pipetting tools into the Biomek system, which are required for auto-SELEX.

In the first attempt of automated selection, oligo (dT)$_{25}$ was attached to Dynabeads as a target to test the feasibility and efficiency of the automated system. A RNA pool including 30 random sequences was used as the starting pool. The time spent to complete one round of auto-SELEX including mixing of pool and target, elution of bound RNA sequences, RT-PCR and *in vitro* transcription was only 212 minutes, which was impressive compared to the days or a week of manual SELEX. The sequence data showed that all cloned sequence included continuous poly A, which was expected for oligonucleotide (dT)$_{25}$ target’s binding sites. This successful attempt provided high possibilities for other targets, such as protein, to be used in the auto-selection procedure (Cox et al. 1998).

The researchers then utilized the automated aptamer selection method to develop an anti-lysozyme aptamer. In this selection procedure, filtration was used to separate the specific
binders and the unbound nucleic acids. Target lysozyme was biotinylated and fixed on streptavidin beads, the mixture of selected pool and target-bound beads was then filtered through a low protein-binding membrane. In this experiment, selection was repeated for 12 rounds; the whole procedure took around 12 hours. Sequence comparison results showed that six binding species were developed from the original pool. One of the binding species constituted 61% of the total chosen colonies (22/33). The dissociation constant of the highest affinity aptamer was 31nM. The automated workstation can simultaneously deal with eight selections and can finish 12 rounds of selection in two days. The same auto-SELEX strategy were utilized to isolate aptamers of some other protein targets including human U1A (a component of the nuclear splicesome), MEK1 (a human MAP kinase) etc. Generally the K_d value is from pico molar to mid-nanomolar (Cox et al. 2002).

**Capillary SELEX**

The procedures of SELEX have been gradually improved with time. For affinity column chromatography, the conventional method of capturing potential aptamers in SELEX, a certain target molecule needed to be linked to a matrix support, such as agarose; some potential binding sites of the target may be blocked by linking to the matrix, which could decrease the binding efficiency of SELEX. Another commonly used SELEX method, the filter assay could avoid the problem above because the binding between oligonucleotides and target occurs in free solution without linking to a support material. But sometimes it could provide a high binding background of oligonucleotides. Considering the shortcomings of conventional SELEX methods, Mendonsa and Bowser (2004) at the University of Minnesota utilized capillary electrophoresis (CE) and showed it could work well in SELEX for different targets.
CE showed several important advantages compared to affinity column and filtration as a method of capturing the target-aptamer complex. First, with CE, single-stranded DNA or RNA molecules bind to their corresponding targets in solution instead of the target being attached to a support matrix. This avoids the necessity to block binding sites. Second, oligonucleotide molecules specifically bound to targets can be separated from those nonbinding oligonucleotides according to size and charge differences. Thus washing steps are unnecessary. Third, the whole procedure of CE SELEX only take 2-4 days instead of the weeks to months that it takes for other SELEX methods.

The researchers first used human IgE as the binding target. After two rounds of selection using CE, they had around 95% binding but they didn’t do negative selection. After cloning and sequencing the winning sequences, they showed that the average $K_d$ value was less than 100 $\mu$M, and the best one was 27±8 nM (Mendonsa and Bowser 2004).

The same group also determined the minimum size of target molecules that could be used in CE SELEX. They chose NPY (neuropeptide Y), a 36-amino acid peptide, which is smaller than a 80-mer single stranded DNA. Results showed that the $K_d$ of the single stranded DNA aptamer of NPY isolated via CE-SELEX was in the high nanomolar range after four rounds of selection. The specificity of the single stranded DNA aptamer to NPY was also tested and it showed 42-fold selectivity for NPY compared to the human pancreatic polypeptide. So the data showed that even a molecule that is as small as NPY could be used as a target of CE-SELEX. The strategy of SELEX using CE could be considered for selection of other single stranded DNA or RNA aptamers (Mendonsa and Bowser 2005).
photo selex

Photochemical SELEX (PhotoSELEX) is the procedure through which the aptamers are isolated through a covalent, *in vitro* selection method (Golden et al. 2000). The establishment of photo SELEX is based on the development of photoaptamers, which are derived from aptamers by replacing the thymidine (T) with a brominated deoxyuridine (BrdU). Photoaptamers possess the ability to crosslink to specific sites on target proteins by forming covalent bonds when electronically excited by long wave UV light (Dom Zichi 2002). The binding ability and crosslinkability is based on recognizing the shape and charge distribution of the targets and make photoaptamers an excellent reagent for diagnostics (Brody and Gold 2000). The development of multiplexed photoaptamer-based arrays allow multiple proteins of interest to be simultaneously measured and the relative data showed that a 17-plex photoaptamer array could achieve the detection of target proteins including interleukin-16, vascular endothelial growth factor at the 10 fM level (Bock et al. 2004).
MATERIALS AND METHODS

Plasmids and oligonucleotides

Express vector pRSET-6His-c-Myc (human) and pET15b(+)-6His-Max (human) were donated by Dr. Ernest Martinez; the following oligonucleotides were synthesized by Integrated DNA Technologies (Coralville, IA): Oligonucleotide pool 2001: 5’-GAG AGA GGT AGG AAT GGC ACG TGN (NNN)13 NNG AAC ACA GCC TTG CCT GGA AT-3’; oligo 1201: 5’-Bio/ATT CCA GGC AAG GCT GTG TTC-3’; oligo 1931 5’-GAG AGA GGT ACG TG-3’; oligo W-1218: 5’-GAG AGA GGT AGG AAT GGC ACG TGA TCG TGG TGG GGG AGG AGG GGA GCC AAG TTC GGT GAA TGA GGA ACA CAG CCT TGC GTG GAA T-3’; oligo W-1223: 5’-GAG AGA GGT AGG AAT GGC ACG TGT GGG GGG ATA CCC CAG TCG TGC AGC ACT AGG TAC GTT GAG TCG AAC ACA GCC TTG CCT GGA AT-3’; oligo W-1234: 5’-GAG AGA GGT AGG AAT GGC ACG TGT ACG TTG GGT GTC AGT GGT CTC AGT GC C GTG GTT CTG GGG TGA ACA CAG CCT TGC CTG GAA T-3’; oligo T1-1218: 5’-GGG GAG GAG GGG AGC CAA GTT CGG TGA ATG AGG AAC ACA GCC TTG CC-3’; oligo T1-1223: 5’-GCA CGT GTG GGG GGA TAC CCC ATG CGT GC-3’; oligo T2-1218: 5’-CAC GTG ATC GTG GTG GGG GAG GGG AGC CAA GTT CGG TGA ATG AGG AAC ACA GCC TTG CC-3’; oligo T1-1223: 5’-GCA CGT GTG GGG GAG GGG AGC CAA GTT CGG TGA ATG AGG AAC ACA GCC TTG CC-3’; oligo T2-1223: 5’-CAC GTG TGG GGG GAT ACC CCA GTC GTG CAG CAC TAG GTA CGT TGA GTC-3’; oligo T2-1234: 5’-CAC GTG TAC GTT GGG TGT CAG TGG TCT CAG TGC GGT GGT TCT GGG GT-3’; the following oligonucleotides were synthesized from DNA facility of Iowa State University oligo E box: 5’-GCT CAG GGA CCA CGT GTG CGG GGA TC-3’; oligo 154: 5’-CGG GGA GAC AAC GAC GGC GGT GG-3’; oligo 155: 5’-CGG ATT CTC TGC TCT CCT CG-3’; oligo 156 5’-
GGA GCG CCA GAG GAG GAA CG-3'; oligo 180: 5'-CTG CGA CGA GGA GGA GAA C-3'; oligo 191: 5'-CCA CAC ACA CCA GCA AGA-3'; oligo T7-1: 5'-AAT ACG ACT CAC TAT AG-3'; oligo PETREV: 5'-GCT AGT TAT TGC TCA GCG G-3'.

Materials and Equipment

BL21-codonPlus (DE3)-RP competent cells for expression human c-Myc and BL21 (DE3) pLysS Competent cells for expression human Max were purchased from Stratagene (La Jolla, CA); BD TALON Metal Affinity Resins for protein purification was purchased from BD Biosciences; the SELEX buffer was 13.5 mM NaCl, 100 mM KCl, 0.22 mM Na2HPO4, 5 mM MgSO4, 120 mM CaCl2, 5 mM MgCl2, 0.05% NP40, 20 mM HEPES, pH 7.25. The buffers for dialysis and storage of protein included BC500 and BC100. BC500 contains 20 mM Tris-HCl pH 7.9, 20% glycerol, 500 mM KCl, 0.05% NP40, and 10 mM 2-mercaptoethanol; BC100 contains the same components as BC500 except 100 mM KCl. The Coomassie (Bradford) Protein Assay Kit was ordered from PIERCE (Rockford, IL). The ECL Western blotting detection reagents and analysis system were purchased from Amersham Biosciences (Piscataway, NJ). The mixed cellulose nitrate and cellulose acetate membrane filters 0.22um/25mm GSTF 025 00 for SELEX were from Millipore (Bedford, MA). The Dynabeads-280 Streptavidin-linked beads were from Dynal Biotech, the TOPO TA cloning kit was from Invitrogen (Carlsbad, CA), the γ32P-ATP was from ICN (Costa Mesa, CA), the Micro Bio-Spin 30 chromatography columns for purifying oligonucleotides were from BD-BioRad (Hercules, CA), and the phosphorimager screen for exposure of EMSA gels was from Amersham Biosciences (Piscataway, NJ). Equipment used included a Typhoon 8600 variable mode imager from Amersham Pharmacia Biotech (Piscataway, NJ), an air-driven
ultracentrifuge from Beckman (Fullerton, CA) and a minicycler PCR machine from MJ Research (Waltham, MA). The software used for predicting ssDNA structure was Mfold (Zuker 2003).

**Checking cDNA sequences of human c-Myc and Max cloned in expression vectors**

Primers T7-1, 154, 155, 156 and 180 were used for sequencing the human c-Myc cDNA and primers T7-1, PETREV, and 191 were used for sequencing the human Max cDNA by the DNA facility of Iowa State University.

**Protein Expression and Purification**

*Expression of recombinant human c-Myc and Max protein in E.coli*

The pET15b(+)-6His-Max plasmid was transfected into BL21 (DE3) pLysS competent cells (Stratagene). The pRSET-6His-Myc was transfected into BL21-codonPlus (DE3)-RP competent cells (Stratagene). Transformed bacteria were cultured in 500 ml LB culture media including 100 μg/ml ampicillin at 37°C for 16 hours. Five ml of the above culture media was transferred to 1L fresh LB media including 100μg/ml ampicillin and the culture shaken at 225 rpm at 30°C until the cell density had reached 0.3-0.4 OD₆₀₀ₙₙₚ. IPTG (0.5mM) was added and shaking was continued under the same conditions for another 3 hours. Bacterial pellets were collected by centrifugation. The bacteria pellets were stored at -80°C or protein purification was continued.

*Purification of recombinant human Max protein*

The bacterial pellet was resuspended in 15ml cold lyses buffer (20 mM Heps, pH7.9, 500 mM NaCl, 10% glycerol, 0.1% NP 40, 10 mM 2-mercaptoethanol, 1 mM PMSF) and
sonicated on ice. The lysate was centrifuged and the supernatant containing soluble 6His-Max protein was collected and incubated at 4°C for 4 hours with rotation with 1ml TALON Metal Affinity Resins containing 5 mM imidazole. The resin was washed 3 times with 5ml cold lysed buffer containing 5 mM imidazole then 3 times with 1ml buffer BC500 containing with 5mM imidazole, followed by once with BC100 containing 15 mM imidazole then once with BC100 containing 30mM imidazole. Recombinant 6His-Max protein was eluted from the TALON resin with 300 mM imidazole in BC100. All the steps were performed at 4°C. The concentration of purified protein was determined using the Coomassie (Bradford) Protein Assay Kit. The size of the protein was determined by SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) with Coomassie blue staining as well as Western blotting.

**Purification of recombinant human c-Myc protein**

The bacterial pellet was suspended with 15 ml cold lysed buffer (20 mM Hepes, pH7.9, 500 mM NaCl, 10% glycerol, 0.1% NP 40, 10 mM 2-mercaptoethanol, 1 mM PMSF) and sonicated on ice. The lysate was centrifuged, and then the insoluble 6His-cMyc in the pellet was resuspended in 10 ml E-buffer (50 mM Hepes, pH7.9, 5% glycerol, 0.5 mM 2-mercaptoethanol, 0.05% Na-deoxycholate, and 1% NP40) and homogenized. The inclusion bodies were separated by centrifugation; the pellet that contained the insoluble 6-His-cMyc was suspended with S-buffer (10 mM Hepes, pH7.9, 6 M guanidine-HCl, and 5 mM 2-mercaptoethanol) and shaken overnight at 4°C. The mixture was separated by centrifugation to remove aggregates then the supernatant containing denatured 6His-cMyc was brought to 5 mM imidazole and incubated with 1ml TALON Metal Affinity Resin for 2
hours at 4°C. The resin was washed 3 times with S-buffer containing 5 mM imidazole, 3 times
with 1ml BC500 containing 7M urea and 5mM imidazole, once with BC100 containing 7M
urea and 15mM imidazole then once with BC100 containing 7 M urea and 30 mM imidazole.
The recombinant 6-His-cMyc was eluted with 1ml BC100 containing 300 mM imidazole and
7 M urea. The concentration of purified protein was determined using the Coomassie
(Bradford) Protein Assay. The size of the protein was determined by SDS-PAGE and
Coomassie blue staining as well as by Western blot.

**Reconstitution of the c-Myc:Max complex**

To obtain the c-Myc-Max heterodimer, c-Myc and Max were refolded together by
step-wise dialysis starting with a molar ratio of 3:1. One mg c-Myc and 100 μg Max were
mixed in 1ml BC100 containing 7 M urea and incubated at 37°C for 30 min. The protein
mixture was dialyzed in “Mini Dialysis Tubing” (MW cutoff: 3.5kDa) in 8 successive steps of
2 hour each against BC500 containing 4 M urea, BC500 containing 2 M urea, BC500
containing 1 M urea, BC500 containing 0.5 M urea, BC500, BC100 (for twice) and SELEX
buffer. To remove aggregates formed during refolding, the preparation was then spun for 2h at
103,000 x g (80,000 rpm) using an air-driven ultracentrifuge. The reconstituted heterodimer
was stored at 4°C.

**EMSA (Electrophoretic Mobility-Shift Assay)**

**End labeled E box probe**

One hundred pmol double-stranded E box was end labeled with γ\(^{32}\)P using 20 μCi
γ\(^{32}\)P-ATP at 37°C for 2 hours. The labeled DNA probe was separated from the ATP and
phosphate using a micro Bio-Spin 30 chromatography column. The specific activity of the probe was determined by TCA precipitation. The concentration of DNA was determined using a Nano Drop NP-1000 spectrophotometer and stored at -20°C.

**EMSA**

The interaction between the c-Myc:Max heterodimer and the E box probe was performed in a 20 μl final volume containing 25 mM Tris-HCl pH 7.9, 15% glycerol, 100 mM KCl, 0.15 mM EDTA, 0.075% NP40, 7.5 mM 2-mercaptoethanol, 160 ng poly(dIdC):(dIdC), 375 μg/ml BSA. The incubation period was 30 min at room temperature (~24°C). The binding activities were analyzed by running the samples through 6% native PAGE in 20 mM Tris-acetate, 0.5 mM EDTA, pH 7.9 at 150 V and 2 hours at 4°C.

**SELEX (Systematic Evolution of Ligands by Exponential Enrichment)**

**Positive Selection against c-Myc:Max**

A 100-pmol single stranded DNA pool, including 42 random sequences, was end labeled as described previously. Certain amounts of c-Myc:Max protein and labeled ssDNA pool were incubated at ~24°C with constant rotation under SELEX buffer for 1 hour. The DNA-protein complexes were collected on a filter by loading the mixture of target protein and single-stranded DNA molecules on a membrane filter GSTF 02500 and filtering at a pressure of 125 mm Hg. Then the filter was washed 3 times with 1 ml of SELEX buffer. Protein bound DNAs were eluted with hot buffer containing 7M urea, 100 mM sodium citrate, 3 mM EDTA pH 9. The total cpm in the protein-bound pool was determined using Cerenkov radiation measured with a scintillation counter. The selected ssDNA pool was concentrated by
precipitation with 100% ethanol then amplified by PCR in which one primer was biotinylated.

The PCR product and streptavidin beads were incubated at ~24°C for 30 min in B&W buffer (5 mM Tris-HCl pH7.5, 0.5 mM EDTA, 1 M NaCl) then the buffer was removed by removing the magnetic beads using magnetic separation. The beads were incubated in 1X SSC (0.15 M NaCl, 0.015 M NaCitrate pH 7.0) for 5 min at 95°C then the ssDNA eluted in this buffer was retrieved.

**Negative Selection against the membrane filter**

The ssDNA pool was end-labeled with γ-32P as described previously. The 32P-labeled selected ssDNA pool was loaded on a membrane filter GSTF 02500 and pulled through under a pressure of 125 mm Hg. The cpm (counts per minute) of the flow-through was determined by a scintillation counter and used to calculate the binding efficiency of the pool. The flow-through was collected and concentrated by precipitation with 100% ethanol.

**Negative Selection against human Max and the membrane filter**

The ssDNA pool from a positive selection was end labeled and incubated with a chosen ratio of human Max and ssDNA pool at ~24°C for 1 hour. The protein-ssDNA mixture was loaded onto membrane filter GSTF 025 00 and pulled through at the pressure of 125 mm Hg. The cpm of the flow-through was determined with a scintillation counter and used to calculate the binding efficiency and the flow-through was concentrated by precipitation with 100% ethanol.

The positive and negative selections were repeated for 13 rounds until the binding efficiency was up to ~80% of the total cpm.
Cloning of candidate aptamer molecules for sequencing

After the binding efficiency reached 50% the PCR products were cloned into PCR 2.1 TOPO TA plasmid and transformed into One Shot Top10 competent cells to produce clones with the plasmids that included the PCR product inserts. Fifty colonies were chosen for preparing plasmids and sequencing.

Filter Assay to check the binding abilities of the cloned oligonucleotides

The 50 sequenced oligonucleotides (including the E box and the random region but not the majority of the primer sequences) were synthesized and end labeled with $\gamma^{32}\text{P}$. The oligonucleotides were each incubated with c-Myc:Max at a molar ratio of 1:1 in SELEX buffer at ~24°C for 1 hour, and then the binding efficiency of each oligonucleotide for c-Myc:Max was determined by filter assay using the same procedure described in “Positive Selection against c-Myc:Max”.

EMSA to check the binding ability of the putative aptamers from SELEX to c-Myc:Max

Five oligonucleotides were chosen from the results of the filter assay for which 15-23% of the total cpm bound c-Myc-Max. The complete sequences of these oligonucleotides were determined and synthesized, and then their 2-dimensional structures were determined using Mfold. This analysis was used to establish truncated oligonucleotides that are likely to contain the aptamer sequence. EMSA was performed with each oligonucleotide including the complete sequences (including primer sequences) and truncated sequences. For each oligonucleotide BSA was substituted for c-Myc:Max as a control. In some studies, the $^{32}\text{P}$-labeled oligonucleotide was mixed with an unlabeled oligonucleotide
(such as the E-box) at ratios of labeled to unlabeled oligonucleotide of 1:50 and 1:100.
RESULTS AND DISCUSSION

Expression and Purification of full length of human 6His-tagged c-Myc and Max fusion proteins in E.coli

The Max protein was expressed in the host cell line BL21(DE3) pLysS (Stratagene). The Myc protein includes 8% proline residues and its c-terminal contains around 14% arginine residues. A strong bias of codon usage for these two amino acids is known to exist between humans and bacteria. In order to avoid truncated or limited expression of the c-Myc protein in E.coli we chose BL21-CodonPlus (DE3)-RP as the host cells to express the Myc protein. This E.coli strain includes extra copies of genes encoding tRNAs that carry arg and pro and recognize the arginine codons “AGA/AGG” and the proline codon “CCC”, which are commonly used in human but not bacterial cells (Stratagene).

Recombinant Max was soluble and could be accumulated without degradation in nondenaturing conditions by collecting several elution fractions from cobalt-affinity chromatography on “TALON” resin (Clontech). The c-Myc protein was insoluble and needed to be purified under denaturing conditions. The left panel of figure 2 shows the Coomassie blue staining of the c-Myc; the result showed that the majority of the elution fraction was the c-Myc. The right panel of figure 2 shows the Western blotting result using the primary mouse anti-his antibody and the secondary goat anti-mouse IgG-HRP antibody; the data showed that the c-Myc was successfully expressed and purified from the host cells. The apparent molecular weight of c-Myc was around 64Kd in 15% PAGE, which is as same as the published literature (Farina et al. 2004), but theoretically it should be 49 Kd. The inconsistence here could be explained that under SDS and heating conditions, the c-Myc
protein may not bind as much SDS as other proteins of the same length.

Figure 3 showed that the Max was also successfully expressed and purified from host cells. The left panel in figure 3 shows the Coomassie blue staining of Max and the right one shows the Western blot result. The Western blot shows that there was some contamination in the elution fraction of the Max, but the major component was Max as shown by Coomassie blue staining. The contamination did not seem to affect the refolding of c-Myc:Max heterodimer as described below and shown in figure 4.

**Refolding and analysis of recombinant full length c-Myc:Max complex**

Previously, others have tried to prepare the Myc:Max heterodimer by separately isolating and refolding denatured recombinant Myc and Max, mixing them together, and heating at 37 °C or 42 °C to allow the Max homodimer to dissociate and the Myc:Max heterodimer to form (Kato et al. 1992). Using this protocol resulted in a very unstable Myc:Max heterodimer and required several micrograms of Myc protein to form the heterodimer (Kato et al. 1992; Reddy et al. 1992). In our experiment, we used a protocol of co-refolding full length of c-Myc and Max protein by slow stepwise dialysis as described previously (Farina et al. 2004). The molar ratio of c-Myc to Max used for refolding was 3:1. To prove that the c-Myc:Max complex properly refolded and was able to bind its target DNA sequence, EMSA (electrophoretic mobility-shift assay) was performed using a γ-32P labeled E box-containing double stranded DNA probe. It could be seen on 6% native polyacrylamide gel that the c-Myc:Max-E box complex formed two shifted bands (figure 4). The relative mobility of the upper band was 0.22. It was previously shown that the c-Myc protein is unstable and can not form a homodimer or bind to the E box (Dang 1999). Thus the shifted band was
identified as the c-Myc:Max heterodimer. The relative mobility of this band is similar to that shown by others for c-Myc:Max (Farina et al. 2004). The relative mobility of the lower band was 0.36, and it was identified as the Max:Max homodimer. This was demonstrated using purified Max, which gave only one band on the gel at a mobility of around 0.36. BSA was used as a negative control for binding activity and showed no binding to the probe (Figure 4).

It was found that a certain proportion of the Max or c-Myc:Max preparation was in the form of big aggregates that did not travel out of the wells during EMSA. These aggregates could affect the SELEX procedure if not removed. The Max aggregates could be dissociated by diluting 10-fold, but the aggregates of c-Myc:Max did not dissociate after dilution. Ultracentrifugation was performed after dialysis to remove the aggregates of c-Myc:Max from the preparation. The c-Myc:Max heterodimer itself was not very stable at 4°C. The Max:Max homodimer replaced the c-Myc:Max heterodimer with time. Therefore it was necessary to make a new protein prep every two to three rounds of SELEX to make sure the correct target was available for selection. One possible reason for this dimer substitution is that the Myc:Max complex dissociates slowly with time at 4°C. Max forms homodimers much more readily than it forms a heterodimer with Myc because the relative concentration of Myc is not high enough. This is consistent with previous studies that several micrograms of c-Myc and a large molar excess of c-Myc over Max are required to form the heterodimer of c-Myc:Max when the two proteins are separately refolded and mixed together. Another explanation of these results could be that when Myc and Max dissociate, the Myc protein is not stable as a monomer and denatures to form aggregates. Thus the concentration of Myc monomer available to form dimers with Max is not high enough, and Max instead will form homodimer.
Choices of temperature and KCl concentration of SELEX

In the original design of the SELEX experiments, 37°C was chosen as the proper temperature for selection because it is close to the physiological conditions in vivo. In order to make sure the refolded c-Myc:Max complex was active through the SELEX procedure, 2μM c-Myc:Max was incubated in SELEX buffer at 37°C for 2 hours in a total volume of 100μl. After two hours the sample was diluted into the EMSA binding buffer including 100mM KCl and the activity of the c-Myc:Max heterodimer was tested by EMSA using a 0.05μM γ-32P labeled E box probe as shown in figure 5. The majority of the protein was present as the Max:Max homodimer instead of the c-Myc:Max heterodimer. Room temperature (~24°C) was then chosen to replace 37°C in the same incubation procedure of c-Myc:Max in SELEX buffer. When tested at this temperature c-Myc:Max activity was observed on the EMSA gel (Figure 5). Therefore we chose 24°C for SELEX. We didn’t try other temperatures. One explanation for why 37°C is not a good choice for optimal c-Myc:Max activity may be that at 37°C or higher temperatures, Max:Max and c-Myc:Max dissociate. The c-Myc monomer is not stable and denatures. In the absence of c-Myc, Max will form homodimers (Farina et al. 2004).

To further test whether the reconstructed c-Myc:Max was able to bind an E box-containing dsDNA oligonucleotide in SELEX buffer, the c-Myc:Max complex was incubated with 32P-end-labeled E box probe at a molar ratio of probe to protein of 17:1 in SELEX buffer at ~24°C in a total volume of 200 μl for 3 hours. Around 30 μl of the reaction solution was then loaded on a 6% native acrylamide gel for EMSA and the complexes were resolved by electrophoresis. The results showed smears below the wells instead of a sharp
band of c-Myc:Max bound to E box probe. The main difference between the SELEX and EMSA binding buffers was the KCl concentration, which was 300mM in SELEX, but 100mM in EMSA binding buffer. Because ideal binding activity had been observed in the EMSA buffer and not in the SELEX buffer, we changed the KCl concentration from 300mM to 100mM in the SELEX buffer and did the same incubation and the EMSA experiment. Under these conditions, the c-Myc:Max heterodimer showed ideal binding to E box probe (Figure 6). These results show that the KCl concentration affects the specific binding of c-Myc:Max to E box, which is consistent with previous data (Farina et al. 2004).

**Design of original random pool of SELEX**

The original pool used in the SELEX procedure contained around $1.4 \times 10^{10}$ different sequences. The pool was amplified by low fidelity PCR from a random pool to increase the copy number of each sequence and to introduce more variation into the pool. Each sequence of the pool includes 42 random nucleotides surrounded by fixed 5’ and 3’ sequences that could be annealed by corresponding primers for amplification as shown in figure 7.

In theory, (an) aptamer sequence(s) could be selected from the random pool. A “decoy” strategy was involved in this design, in which the 5 prime primer contained one complement of the E box sequence (CACGTG); the other complement of the E box was expected to be present within the random region of certain oligonucleotides in the pool and form a stem structure with its complement in the 5’ fixed sequence. This stem structure could bind to the c-Myc:Max DNA binding domain. Other parts in the random region were expected to evolve into an aptamer, maybe a loop-like structure. Thus we were hoping to select an aptamer with two different components that could bind with the c-Myc target protein. The
presence of the two linked components with affinity for c-Myc:Max, the primary aptamer binding region and the E box containing stem structure, would provide the aptamer with a higher affinity for its target than an aptamer with only one binding site.

The “decoy” strategy has been previously utilized for several targets. Mann et al., (1999) targeted E2F, a transcription factor that up-regulates multiple genes' expression related to cell cycle and DNA synthesis, with a double-stranded decoy oligodeoxynucleotide. The E2F decoy contained a consensus E2F DNA binding site and could bind to E2F to efficiently block the activities of some cell-cycle and DNA synthesis related genes (Mann et al. 1999). The oncogenic beta-catenin/T-cell factor (TCF), which regulates some cancer-related genes, such as Cyclin D1 and c-Myc, is another reported target of the decoy strategy. An 18-mer double-stranded DNA decoy of TCF was designed to contain the TCF binding site found in target genes' promoter. The TCF decoy efficiently blocked TCF activity and inhibited downstream target gene expression (Seki et al. 2006). Generally decoy sequences are the same as the recognition sequences of transcription factors in vivo, and a large excess of decoy is required to effectively block the target molecules' function. Here we designed a strategy using a combination of aptamer and decoy in order to improve the binding affinity of the resulting bivalent molecule to the c-Myc:Max target, and to therefore efficiently inhibit the transcription factor's function.

**Aptamer Development**

The whole procedure of SELEX is described in figure 8. The method used for aptamer selection was filter capture. The nitrocellulose membrane used for filtration has very low background binding (less than 2%) of free single stranded DNA molecules, but protein bound
oligonucleotides are retained on the filter. For each round of selection, the proper amounts of c-Myc:Max and single stranded DNA library, end-labeled with $\gamma$-P$^{32}$, were incubated in SELEX buffer at room temperature (around 24°C) for 30 minutes. The mixture was then filtered through a nitrocellulose membrane using vacuum (125 mm Hg). The protein-bound DNA molecules were retained on the filter and nonbound DNAs went through. Then the protein bound DNAs were eluted from the filter with hot urea after washing for three times with SELEX buffer and then were amplified by PCR. The size of the PCR product was 86 bp on 12% PAGE (Figure 9).

For PCR, the 5 prime end of the 3 prime primer was biotinylated to allow the PCR products to be captured by Dynabead-280 Streptavidin (Dynal Biotech) after incubating in B&W buffer (see materials and methods) at 24°C for 30 minutes. After the incubation, the PCR product bound beads were resuspended in 1X SSC (see materials and methods) and incubated at 95°C for 5 minutes to separate the biotinylated and nonbiotinylated DNA strands. It was determined by measuring the cpm values that most of the nonbiotinylated strands, the “sense” strand, could be separated from the biotinylated strand. The “sense” strand used for the next round of SELEX. During the SELEX procedure, negative selections against the filter and Max were included as well as positive selections against c-Myc:Max.

Two parallel SELEX experiments were performed by Ying Liu and Marjan Mokhtarian. Results for Ying Liu’s SELEX are shown in figure 10. Nine rounds of positive and seven rounds of negative selection were done, and the highest binding efficiency for positive selection was 18% at the 11th round of selection. In positive selections, the ratio of target protein to ssDNA pool was from 1:1 to 2.5:1. For negative selections, the first 6 rounds
were against filter membrane and in the last round, Max homodimer was used in the SELEX to remove the filter binders and Max binders. As we expected, the binding percentage of the negative selections was always lower than 3% (Figure 10). This SELEX was not continued for further studies.

For Marjan Mokhtarian's SELEX, seven rounds of positive selections and six rounds of negative selections (three against filter, three against Max protein) were performed. The starting pool was from 12 cycles of low fidelity PCR using the original random pool; it included around $10^{10}$ DNA molecules. In the fourth round of positive selection, double-stranded E box containing oligonucleotide was included to compete with those oligonucleotides that bound to Myc:Max based only on the E box binding domain. During the whole procedure of SELEX, the selection pressure was not constant. For the positive selection in round 1, 2 and 4, the ratio of oligonucleotides to c-Myc:Max protein was 1:1, the ratio in round 3 was 1:2.5; in the fifth round of positive selection, the ratio was 3:1 and then 5:1 and 10:1 in the sixth and seventh round respectively. For the negative selection in round 5, 7 and 8, the Max protein was used as the target; the ratio of oligonucleotides to protein was 1:1, 1:4 and 10:1 respectively.

As shown in figure 11, the binding efficiency after the last round of positive selection was around 27%. For the negative selections against the filter alone, the binding percentage of the pool gradually decreased from 20% to 10% then to 3%; for the negative selections against Max protein, the percentage also decreased from 30% to 22. After the 7th round of positive selection, the binding efficiency was up to around 27% and the selected single stranded DNA molecules were amplified by PCR and cloned into a plasmid using T7 TOPO cloning kit (see
material and methods). Fifty colonies were cloned and forty-six were sequenced.

The sequences from the forty-six colonies then were chemically synthesized. The synthesized sequences started with the complement of the E box sequence (CACGTG) in the 5' prime fixed region followed by the 42 random sequences without the 3' prime fixed region as shown in Figure 12. The 46 sequences were tested for their ability to bind c-Myc:Max by a filter binding assay. Each of the synthesized oligonucleotides was end-labeled by $\gamma^{32}\text{P}$ and incubated with c-Myc:Max (the molar ratio was 1:1), and then was tested the binding efficiency using the filter assay. The data from filtration showed that for five of the synthesized sequences binding efficiency (bound/unbound) was above 15%. These five oligonucleotides were T2-2111, T2-2118, T2-2123, T2-2126 and T2-2134 (refer to Table 1 for the sequences).

EMSA was performed to test the binding ability of the five selected oligonucleotides. The five oligosnucleotides were end-labeled with $\gamma^{32}\text{P}$. Results (Figure 13) showed that three of these five sequences, T2-2118, T2-2123 and T2-2134, could bind to the c-Myc:Max complex as shown by the presence of a band at the expected position on the 6% native PAGE.

Manual alignment of the three oligonucleotides showed that there were some matches, especially G, among these sequences with some gaps (Figure 14). The full length sequences of the three identified oligonucleotides (including the complete 3' and 5' fixed regions) from above then were analyzed using the Mfold program (Zuker 2003) to predict the DNA structures in two dimensions (Figure 15). It was found that some concentrations of the matches among these sequences existed in a loop or stem-loop region in these molecules, which are highlighted with red circles in figure 15. Also, the loop or stem-loop region of each
oligo contained several Gs, which indicated that G-quartet structures could be formed in these areas.

A G-quartet is a common DNA binding structure for protein targets (Jing et al. 2004) and consists of four guanine bases arranged on the four corners of a square planar (Shafer and Smirnov 2000). Thus, it is highly likely that the conserved region in each oligo could play some roles in target binding. For example, the thrombin binding aptamer (TBA) was one of the earliest studied DNA aptamers that can form G-quartet structure. TBA contains 15 core oligonucleotides and forms a unimolecular quadruplex consisting of two G-quartets linked by two TT loops and a TGT loop (Macaya et al. 1993).

We also analyzed the structures of two different truncated versions of W-2118 (T1-2118, T2-2118), W-2123 (T1-2123, T2-2123), and one truncated version of W-2134 (T2-2134) using Mfold. The “T1” truncated version of each putative aptamer includes the stem-loop portion of the corresponding whole sequence shown in figure 16, and “T2” truncated version consists of the complement of the E box sequence following by the random region of each corresponding whole sequence (Figure 12 and 16). Data showed that each truncated version of the three oligonucleotides contains the same conserved region as the corresponding whole sequences (emphasized in red circle) (Figure 16), which indicated that these conserved regions could be involved in binding to the target c-Myc protein.

**Binding properties of full length and truncated sequences of selected oligonucleotides**

After predicting the two dimensional structure of the oligonucleotides, we tested the binding ability of each oligo to c-Myc:Max including the full length sequences and truncated regions determined according to the refolded structure shown in figure 16. The name and
sequence of each oligo are shown in Table 1.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>E box</td>
<td>GCT CAG GGA CCA CGT GGT CGG GGA TC</td>
<td></td>
</tr>
<tr>
<td>1931</td>
<td>GAG AGA GGT AGG AAT GGC ACG TG</td>
<td></td>
</tr>
<tr>
<td>1201</td>
<td>Bio/ATT CCA GGC AAG GCT GTG TTC</td>
<td></td>
</tr>
<tr>
<td>T2-2111</td>
<td>CAC GTG ACT TCC CGC CCA CTG ACA CCA CGG CAT CCA TTC TCT GTT CGT</td>
<td></td>
</tr>
<tr>
<td>T2-2118</td>
<td>CAC GTG ATC GTG GTG GGG GAG GAG GGG AGC CAA GTT CGG TGA ATG AG</td>
<td></td>
</tr>
<tr>
<td>T2-2123</td>
<td>CAC GTG TGG GGG GAT ACC CCA GTC GTG CAG CAC TAG GTA CGT TGA GTC</td>
<td></td>
</tr>
<tr>
<td>T2-2126</td>
<td>CAC GTG TAC CAG TAC GGC TTC CCG ACT TTC CTA AAC GCA CCC TGA TCG</td>
<td></td>
</tr>
<tr>
<td>T2-2134</td>
<td>CAC GTG TAC GTT GGG TGT CAG TGG TCT CAG TGC GGT GGT TCT GGG GT</td>
<td></td>
</tr>
<tr>
<td>T1-2118</td>
<td>GGG GAG GAG GGG AGC CAA GTT CGG TGA ATG AGG AAC ACA GCC TTG CC</td>
<td></td>
</tr>
<tr>
<td>T1-2123</td>
<td>GCA CGT GTG GGG GGA TAC CCC AGT CGT GC</td>
<td></td>
</tr>
<tr>
<td>W-2118</td>
<td>GAG AGA GGT AGG AAT GGC ACG TGA TCG TGG TGG GGG AGG AGG GGA GCC AAG TTC GTT GGT GAA TGA GGA ACA CAG CCT TGC CTG GAA T</td>
<td></td>
</tr>
<tr>
<td>W-2123</td>
<td>GAG AGA GGT AGG AAT GGC ACG TGT GGG GGG ATA CCC CAG TCG TGC AGC ACT AGG TAC GTT GAG TCG AAC ACA GCC TTG CCT GGA AT</td>
<td></td>
</tr>
<tr>
<td>W-2134</td>
<td>GAG AGA GGT AGG AAT GGC ACG TGT AGT AGT GGT GTC AGT GGT CTC AGT GCG GTG GTT CTG GGG TGA ACA CAG CCT TGC CTG GAA T</td>
<td></td>
</tr>
<tr>
<td>928</td>
<td>GGG GGT AAA GTA TTC CGG GGC GGG A</td>
<td></td>
</tr>
</tbody>
</table>

Table 1 Sequences of tested oligonucleotides including full length version and two kinds of truncated versions. W= full length, T1= truncated version 1, T2 = truncated version 2, which includes CACGTG followed by random region.
The full length and the two truncated versions of the three selected oligonucleotides were end labeled by $^{32}$P and tested for binding to c-Myc:Max by EMSA. BSA was used as the non-specific binding control for c-Myc:Max. As shown in figure 17 and 18, all of the sequences could bind to the c-Myc:Max target protein; the W-2118, T2-2118, T2-2123 and T2-2134 could bind to BSA. The end labeled W-2123 and W-2134 were then tested for the minimum concentration required for the binding to c-Myc:Max. Results (Figure 19) showed that both of the two sequences had c-Myc:Max binding ability at the range from 1μM to 10nM, but no BSA binding activity.

To test if the synthesized oligonucleotides can compete with the E box probe to bind to c-Myc:Max, competition EMSA was performed by using end-labeled E box probe and unlabeled oligonucleotides as the competitors (Figure 20a). The ratios of each competitor to the labeled E box probe were 50:1 and 100:1. Quantitative analysis showed that none of these oligonucleotides could effectively compete with the E box for binding to the target protein (Figure 20a and 20b). The results of the competition EMSA suggested that the selected oligonucleotides could bind to c-Myc:Max complex but not to the E box binding domain on the protein.

The filter binding assay was used to test the binding affinity of the two full-length oligonucleotides W-2123 and W-2134 at concentrations ranging from 1nM to 1μM. The concentration of c-Myc:Max used in this analysis was 150nM throughout. The amount of $\gamma^{32}$P labeled oligonucleotides bound to the target protein was measured at various concentrations of each oligonucleotide. From the binding curve of each oligonucleotide, it was estimated that the $K_d$ of the full-length oligoW-2134 was around 180 nM, The whole
sequence of oligo W-2123 had $K_d$ of about 236nM. These filter binding assays were done by Marjan Mokhtarian (Figure 21).

For developing an aptamer, it is necessary to determine the minimum size of the aptamer and to test the binding affinity of aptamer for its target. Two truncated versions (T1-2118 and T1-2123) had 29 and 47 nucleotides and could bind to c-Myc:Max. The lengths are in the range of most aptamers; further deletion to get the shortest version and check of the binding ability may be necessary.

The sequences of the current putative aptamers do not include the whole E box region, and the data from the competition EMSA demonstrated that the putative aptamers could not compete with E box to binding to the c-Myc:Max. These results suggest that the binding site of the putative aptamers on the protein target is not in the E box binding domain.
CONCLUSIONS AND FUTURE DIRECTIONS

Our original plan was to select one or more DNA aptamers containing a primary aptamer sequence and an E box sequence. The data we obtained showed that there was no E box sequence in any of the selected putative aptamers. But from the results of EMSA and filter assay, these selected putative aptamers could bind to the c-Myc:Max target protein. Also, the EMSA data showed that none of the selected oligonucleotides could compete with E box to bind c-Myc:Max target. These results suggest that the selected oligonucleotides bind to some other sites on the c-Myc:Max heterodimer instead of the E box binding site. *In vivo* c-Myc cooperates with other cofactors, such as TRRAP, Tip60 through its N-terminal domain to regulate gene expression. So these putative aptamers with high binding affinity to c-Myc could interfere with the interaction of c-Myc with these cofactors.

Our current data shows that the $K_d$ of the aptamer W-2123 is about 240 nM and W-2134 is 180 nM (The $K_d$ of E box binding to c-Myc:Max is around 100nM). The EMSA data showed that W-2123 and W-2134 could bind both c-Myc and Max, so the procedure for improving the binding affinity of the aptamers will include optimizing the binding specificity by separating the c-Myc and Max binding activity of the aptamers.

The next step of our future work would be to optimize the binding affinity and specificity. To realize this purpose, we could do SELEX on a doped library, which means we would synthesize a degenerate aptamer based on the current putative aptamer sequences (W-2123 and W-2134). During doping, 10% of the contributing bases in each synthetic reaction would be randomly changed to a base other than the original, and 90% would be kept the same as in the aptamer sequence. Thus some random mutations would be introduced to the
synthesized pool. The two pools of degenerate aptamers (of W-2123 and W-2134) would be combined to create a diversified oligonucleotide library. This library would be used in SELEX experiments to develop anti-c-Myc or anti-Max aptamers with higher binding affinity than the current ones.

To separate c-Myc and Max binding activity, we would do two parallel SELEX for c-Myc:Max (SELEX A) and Max (SELEX B) using the doped library. For the SELEX A, c-Myc:Max heterodimer would be used as the target for positive selections, the Max and the cellulose nitrate filter support would be used for the negative selections. For SELEX B, Max would be used for positive selection and the cellulose nitrate filter support would be used for the negative selection. After getting the “winner” target binders with high binding affinity and specificity, the minimal binding sequences could be determined by Mfold and then tested by EMSA and filter assay to estimate the binding affinity.

Our next step to further developing an anti-c-Myc:Max aptamer would be create a bivalent oligonucleotide by adding an E box (or a Max-binding aptamer if available) to the selected aptamer. In this bivalent oligonucleotide, the primary aptamer portion would bind to a site on c-Myc outside the E box binding domain and the added E box sequence would bind the c-Myc:Max DNA binding domain (or the added Max-binding aptamer would bind to the Max protein). In the bivalent molecule, the binding of one of the binding sites to the target protein could make the second binding an intramolecular reaction instead of a bimolecular reaction, therefore the aptamers would have a higher binding affinity to the c-Myc:Max than an aptamer that only has one binding site for its target.

For making the c-Myc-E box bivalent aptamer, a series of polyAs with different length
could be added to the optimized aptamer followed by one complement of the E box (5'CACGTG3') which would be annealed with the complementary E box sequence to become a double-stranded E box linked to the 3' end of the aptamer. The binding ability of the extended aptamers to c-Myc:Max would be tested by EMSA. Competition studies with various concentrations of extended aptamers and a certain concentration of radio labeled E box could be used to test the ability of the extended aptamers to compete with the E box to bind c-Myc:Max. The extended aptamer with the shortest polyA linker and the strongest competitive ability to bind to c-Myc:Max would be used for future studies. Once the length of polyA linker is optimized, the bivalent aptamer would be synthesized as a single-stranded molecule with an E box-hairpin structure at the end. If a Max aptamer were selected, the same strategy described above could be used to develop a bivalent c-Myc-Max aptamer by add the Max aptamer to the 3' end linked by a proper length of polyA (Figure 22).

Possible approaches to increasing the bivalent aptamers' stability are circularization and chemical modifications. Usually circular DNAs are more stable than linear ones in serum and plasma because exonucleases cannot digest circular DNA. Several different modifications could be considered to modify the bivalent aptamers, such as 2'O-methoxy-ethyl and fluoro-arabino. The main considerations here should be the influence of the modifications on DNA stability and on toxicity.

After a stable bivalent aptamer is available, we would need to develop strategies to transfer the single-stranded DNA molecules into proper cell lines. Two delivery strategies could be developed. One is to use liposomes as the carriers in an endocytic-dependent way. Liposome-mediated delivery is frequently used for transferring oligonucleotides into cells.
The delivery procedure can be tested by monitoring the uptake of fluorescence signals from fluorescein labeled oligonucleotides. Another strategy is to use peptide-mediated delivery in an endocytic-independent way. Several basic peptides have been utilized to transfer oligonucleotides into cells. An advantage of this method is that it is possible to deliver DNA molecules into both the nucleus and the cytoplasm depending on whether a nuclear localization signal is included on the peptide. Also, the delivery procedure will be tested via fluorescence signals. An example of this strategy is MPG peptide. It consists of a hydrophobic domain, which is from the fusion sequence of HIV gp41, and a hydrophilic domain, which is from the nuclear localization signal sequence of the SV40 T-antigen. The MPG peptide can tightly bind to single-stranded or double-stranded DNA and ferry the oligonucleotides into cells in an endosome-independent way (Simeoni 2003). The effect of the bivalent aptamer on the activity of c-Myc:Max would be tested in cultured cells. The continued action of c-Myc will affect some cancer-related genes' expression, cell proliferation and cell death procedure. So the study of the bivalent aptamer's function could be focused on these aspects.

To check the effect of the bivalent molecule on inhibiting the c-Myc activity in increasing gene expression, we would use PC3 cells that express c-Myc as the host cells and then co-transfect the cells with pBTdel208 expression construct and phRL-TK internal control plasmid. The pBTdel208 plasmid includes a reporter gene of firefly luciferase and its transcription is driven by the truncated 208 bp human telomerase (hTERT) promoter that is positively regulated by c-Myc. The phRL-TK internal control plasmid includes a reporter gene of sea pansy luciferase that is regulated by the constitutive HSV-thymidine kinase promoter in a c-Myc-independent manner. We would treat the host cells with the bivalent
c-Myc:Max aptamer and then transfect the cells with the reporter constructs. The influence of the aptamer on c-Myc's activity would be detected by measuring the firefly luciferase activity relative to sea pansy luciferase activity in the cell lysates.

To test the bivalent aptamer's effect on cell proliferation, we would use D145 neuroblastoma cells with high level of c-Myc expression as target cells and use human foreskin fibroblast cells without c-Myc overexpression as a control. We would treat the cells with optimized concentration of the bivalent aptamer and the cell proliferation rate would be determined by counting the cells until the cell reach confluence.

As a potential anti-cancer drug, the developed bivalent aptamer can be useful in the cooperation with other anti-cancer strategies including antisense, RNAi etc. The combination of different strategies should be more powerful than single treatment.

An obvious advantage of using an aptamer compared to other anti-cancer drugs is that it can be developed to an allosteric aptamer, which means two different aptamers recognizing different targets could be linked together and function simultaneously. For example, the c-Myc:Max aptamer could be linked to another aptamer that can specifically bind to an anti-cancer drug, such as Gleevec. With this configuration, the binding of the c-Myc:Max aptamer to its target will positively regulate the activity of the anti-Gleevec aptamer, which means the anti-Gleevec aptamer will be allosterically regulated by the c-Myc:Max aptamer, therefore the anti-Gleevec aptamer will work in a c-Myc:Max-dependent way to concentrate the anti-cancer drug in cells. The synergistic activity of these two anti-cancer treatments is expected to be more effective than only one treatment.
To express the human c-Myc and Max protein and isolate c-Myc:Max heterodimers that bind to E box-involved probes, we tried several different expression vectors and host cells. First, we obtained the c-Myc cDNA-containing plasmid pHM6 (pcMYCpHM6), the Max cDNA-containing plasmid pHM6 (pcMAXpHM6), and the Max cDNA-containing plasmid pSG5puro (pcMaxpSG5puro) from Dr. Clement Lee.

After checking the sequence of the cDNAs, we subcloned these two cDNA sequences into different expression vectors for expression and purification. The c-Myc cDNA in the pcMYCpHM6 was subcloned into plasmid pET-21b(+) between BamHI and XhoI to get pcMYCpET21b(+) for expression. Primer 171 containing BamHI cutting site and primer 172 containing XhoI cutting site were individually used as the 5’ and 3’ primers for PCR to amplify the c-Myc cDNA in pcMYCpHM6 (Table 2). It was found that the His-tag in pET-21b(+) was after the stop codon and thus, this construct could not be used to purify the c-Myc protein. Thus, the c-Myc cDNA in plasmid pET21b(+) was then subcloned into pET41b(+) to get pcMYCpET41b(+) with His-tag and GST tag coding sequences upstream of the start codon of c-Myc cDNA (Figure 23). The Max cDNA was subcloned from pcMaxpSG5puro into a pTrcHisA expression vector between the BamHI and EcoRI sites to get pcMAXTrcHisA with His-tag coding sequences at 5 prime (done by Tony Peterson) (Figure 24).

After preparing the expression vectors for c-Myc and Max, the two plasmid constructs containing c-Myc or Max cDNA were individually transformed into BL21(DE3) and XL1-blue host cells for expression and purification (done by Samir Mehanovic). The
insoluble c-Myc protein was contained in inclusion bodies and purified under denaturing conditions with 6M guanidine-HCl. The Max protein was soluble and directly purified from the supernatant of the lysed cells (refer to “Materials and Methods” for details). Ni-NTA was used to purify both the proteins. The purified human c-Myc and Max protein were checked by Coomassie blue stain and Western blot with the primary mouse anti-his antibody and the secondary goat anti-mouse IgG-HRP (Figure 25, 26).

In figure 25, the left panel shows the Coomassie blue stain of c-Myc protein. The expected molecular weight was around 100 Kd including the GST-tag and other sequences involved in the pET41b(+) vector. The Western blot result shown on the right panel indicated that around 70% of the purified protein was c-Myc. In figure 26, the purified Max from XL-1 blue cells showed around 26 Kd including 6his-tag and some other sequences expressed from pTrcHisA vector.

To get active c-Myc:Max heterodimer, we co-refolded the two purified proteins against BC 500 (20 mM Tris-HCl pH7.9, 20% glycerol, 500 mM KCl, 0.05% NP40, 10 mM 2-mercaptoethano) including 4M urea, BC500 including 2M urea, BC500 including 1M urea, BC500 including 0.5M urea, BC500, BC100 (for twice) with 2 hours for each step. The activity of the refolded c-Myc:Max heterodimer was checked by EMSA with 6% native PAGE. Unfortunately, the refolded product didn't show c-Myc:Max activity but only a blur area (figure 27). One possibility was that the GST c-Myc fusion protein folded incorrectly and/or that the large GST tag (which included 209 amino acids) may block the dimerization domain in the C-terminus of the c-Myc protein so it could not correctly dimerize with the Max protein.

We switched to another strategy to express the two proteins by using the pETDuet-1
expression vector (Novagen) as shown in figure 28. pETDuet-1 includes two multiple cloning sites and can co-express two different target proteins in *E. coli*. Firstly we subcloned the Max cDNA from pcMAXpHM6 to pETDuet-1 between SacI and HindIII to get pcMaxpETDuet-1. Primer 185 containing the SacI cutting site and primer 186 containing HindIII cutting site were individually used to amplify the Max cDNA cut from pcMAXpHM6 (Table 2), then the c-Myc cDNA was subcloned from pcMYCpET41b(+) to pcMaxpETDuet-1 between BglII and XhoI to get pcMax-Myc-pETDuet-1. The Bgl II-containing primer 187 and XhoI-containing primer 188 were used to amplify the c-Myc cDNA, which had been cut from pcMYCpET41b(+) (Table 2). The expression vector including c-Myc and Max cDNA was then transformed into BL21-CodonPlus (DE3)-RP host cells. Because the Max protein was a 6his-tag fusion protein and the c-Myc was a GST-tag fusion protein in pETDuet-1 vector, we used TALON resin and Glutathione resin in series to purify the co-expressed heterodimer. The Coomassie blue staining results (Figure 29) showed two bands at around the 98 Kd and 26 Kd position, which should be the c-Myc and Max proteins respectively. The primary rabbit anti-c-myc (N-262) polyclonal IgG and the secondary goat anti-rabbit IgG-HRP were used in the Western blot for the c-Myc, the primary mouse anti-his antibody and the goat anti-mouse IgG-HRP were used in the Western blot for the Max. The results indicated these two proteins were correctly expressed in host cells (figure 29). EMSA was utilized to check the activity of c-Myc:Max to bind the E box, but it only showed the Max:Max homodimer instead of c-Myc:Max heterodimer activity (figure 30). Again the GST tag maybe had some effects on the proper folding of c-Myc and Max in this case.

Finally, we obtained the expression vectors pRSET-His-c-Myc and
pET15b(+)-His-Max from Dr. Ernest Martines to individually express the c-Myc and Max proteins. We achieved refolding of these two proteins with activity to bind E box probe (refer to “Results and Discussion” for details).

<table>
<thead>
<tr>
<th>Oligo</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>171</td>
<td>5'-CCC AAG CTG <strong>GAT CCC</strong> ATG TAC CC-3' (BamH I)</td>
</tr>
<tr>
<td>172</td>
<td>5'-CC+-G CGA ATT <strong>CTC GAG</strong> GGA GGT ACC-3' (Xho I)</td>
</tr>
<tr>
<td>185</td>
<td>5'-CC <strong>CCG AGC TCC</strong> TTG ATG AGC GAT AAC GAT G-3' (Sac I)</td>
</tr>
<tr>
<td>186</td>
<td>5'-C CCC <strong>AAG CTT</strong> TGA TGC ACG TGT AGC G-3' (Hind III)</td>
</tr>
<tr>
<td>187</td>
<td>5'-GG GAA <strong>GAT CTA</strong> GGA GAT ATA CAT ATG-3' (Bgl II)</td>
</tr>
<tr>
<td>188</td>
<td>5'-CCG <strong>CTC GAG</strong> ACC AAG CTT CTA CGC ACA AG-3' (Xho I)</td>
</tr>
</tbody>
</table>

Table 2. Oligonucleotide sequences used in amplification and subcloning of c-Myc and Max cDNAs. Several primers containing proper enzyme digestion sites were designed to amplify the c-Myc and Max cDNA and subclone them into proper expression vectors. The enzyme cutting sites are emphasized by bolding and under lines.
Full length c-Myc cDNA was expressed in *E.coli* BL21(DE3)CodonPlusRP as a 6his-tagged fusion protein and purified under denaturing conditions with 6M guanidine-HCl. The lysed cells, the supernatant obtained after centrifugation of lysed cells and the eluted fractions were analyzed by SDS-PAGE. The left panel shows the Coomassie blue staining results, the molecular weight of c-Myc protein was around 64 Kd in 15% PAGE, which was the same as in the published literature (Farina et al. 2004). The right panel shows the Western blot. Dilution of 1:3000 for the primary mouse anti-his antibody and 1:500 for the goat anti-mouse IgG-HRP were used for the Western blot.
Figure 3 Purification and analysis of human Max protein. Human Max protein was expressed in *E. coli* BL21 (DE3) pLysS as a 6his-tagged fusion protein and purified under non-denaturing conditions. The lysed cells and the elution fraction were analyzed by SDS-PAGE. The left panel shows the Coomassie blue staining results, the molecular weight of Max protein was determined to be around 18 Kd in 15% PAGE. The right panel shows the Western blotting results. Dilution of 1:3000 for the primary mouse anti-his antibody and 1:500 for the goat anti-mouse IgG-HRP were used for the Western blot. The arrows show the position of the purified Max protein (Done by Natalya Peretyatko).
Figure 4 EMSA of c-Myc:Max and Max:Max complex. The purified human c-Myc and Max protein were co-refolded by step-wise dialysis starting with a molar ratio of 3:1. The protein mixture was dialyzed in 8 successive steps of 2 hours each against BC500, 4M urea, BC500, 2M urea, BC500, 1M urea, BC500, 0.5M urea, BC500, BC100 (for twice) then SELEX buffer. The Max protein alone without c-Myc was also dialyzed through the same procedure. The activity of the human c-Myc:Max and Max:Max complexes were analyzed by 6% non-denaturing PAGE using $\gamma^{32}$P end-labeled E box containing dsDNA (refer to Table 1 for sequence). The upper arrow points to the c-Myc:Max heterodimer and the lower one points to the Max:Max homodimer. The relative mobility of both were the same as published by others (Farina et al. 2004).
Figure 5 The effect of temperature on c-Myc:Max activity. 2μM c-Myc:Max was incubated in SELEX buffer at 37°C or 24°C for 3 hours and then was diluted into EMSA binding buffer. The activity of c-Myc:Max heterodimer was tested by EMSA using γ-32P end-labeled E box containing dsDNA. The protein complex showed more c-Myc:Max activity after incubation at 24°C (upper arrow) than after incubation at 37°C. More Max:Max activity was observed at 37°C (lower arrow).
Figure 6 The influence of KCl concentration on binding of c-Myc:Max to the E box. c-Myc:Max was incubated with $\gamma^{32}\text{P}$ end-labeled E box probe at a molar ratio of probe to protein of 17:1 in SELEX buffer including 100mM KCl or 300mM KCl at ~24°C for 3 hours. The activity of c-Myc:Max was tested by EMSA using a 6% native acrylamide gel. The protein heterodimer didn’t show ideal binding when incubated in 300mM KCl (the first two lanes), but showed good binding activity when incubated in 100mM KCl (the fourth and fifth lanes). The upper arrow points to the c-Myc:Max heterodimer and the lower one points to the Max:Max homodimer. These results are consistent with previously reported data (Farina et al. 2004).
Figure 7 Diagram of the oligonucleotides in the random pool used in SELEX. In the middle of each oligonucleotide in the pool were 42 random nucleotides with equal probability of A, T, C, or G at each position. The 5’ and 3’ segments containing fixed sequences could be annealed with the correct primer for PCR. In the 5’ segment, there was also one complement of the E-box sequence, which was CACGTG.
In vitro selection of anti-c-Myc DNA aptamer
(Filter binding assay)

Positive selection

Random pool (γ-32P labeled) + c-Myc/Max protein

Elution (Hot urea)

or

Purify ss DNA

Nitrocellulose membrane

(unbound ss DNA) (FT)

Separate by streptavidin beads

13 (18) rounds

Negative selection

Nitrocellulose membrane

(18) rounds

Nitrocellulose membrane

(18) rounds

Cloning and Sequencing (TOPO TA plasmid)
Figure 8 Diagram of the SELEX procedure. A filter capture protocol was used for SELEX. One round of SELEX included a positive selection against c-Myc:Max and a negative selection against the nitrocellulose membrane alone or Max:Max. The ssDNA pool in each round of selection was end labeled with $^{32}$P. The binding efficiency of the selected pool was measured by cpm (counts per minute) value. After each cycle of selection, the selected pool was amplified by low fidelity PCR. The 5 prime end of 3 prime primer was biotinylated. After the 13th rounds of selection (the binding efficiency was ~27%), the ss DNA pool was cloned into the TOPO TA plasmid for sequencing.

The blue balls denote the c-Myc:Max complex for positive selection, and the red balls stand for Max:Max for negative selection. The small red diamond means biotin labeled at the 5’ end of 3’ primer for PCR.
Figure 9 PCR amplification of a selected single stranded DNA pool. The size of the PCR products of the selected pool ran as 86 bp on a 12% PAGE that was staining with ethedium bromide. The original random pool was used as the positive control and ddH2O was used for the negative control of the selected pool. The DNA marker used was phiX174 DNA/Hinf I marker. The 5 prime end of the 3 prime primer was biotinylated to allow the PCR products to be captured by Dynabead-280 Streptavidin.
Figure 10 Results of SELEX for selection of an anti-c-Myc aptamer. The basic procedure of SELEX is as described in figure 9. Sixteen rounds of selection were performed beginning with a random pool of single stranded DNA including 42N in sequence: “GAGAGAGGTAGGAATGGCACGT(42N)GAACACAGCCTTGCTGGAAT”. The binding efficiency of each cycle of the selection is shown in the Table under the X axis. The blue bars denote the positive selections, dark red bars stand for the negative selections against the nitrocellulose filter membrane and the yellow bar denotes the negative selection against the Max protein. The ratios on the tops of the bars represent the ratio of protein to single-stranded DNA pool used for that round of selection.
Figure 11 Results of SELEX for isolation of anti-c-Myc aptamers (Done by Marjan Mokhtarian). The basic procedure of SELEX was as described in figure 9. Thirteen rounds of selection were performed starting with a random pool of single stranded DNA including 42N in sequence: “GAGAGAGGTAAGGAATGGCACGT (42N) GAACACAGCCTTGCTGGGAAT”. The starting pool was from 12 cycles of low fidelity PCR and included around $10^{10}$ DNA molecules.
Figure 12 Diagram of the sequences of the synthesized oligonucleotides. The codes on the left denote the name of each oligonucleotide. The brown region represents the 42 random core sequences that were referenced in figure 10, the green segment denotes one complement of E box sequence, CACGTG, and the black lines stand for the 5’ and 3’ primers. The sequence of each oligonucleotide was described in Table 2. W=whole, T1=truncated version 1, T2=truncated version 2.
Figure 13 The binding affinity of five cloned oligonucleotides for c-Myc:Max. The five oligonucleotides with binding efficiencies above 15% were end-labeled by polynucleotide kinase using $\gamma^{32}$P, and tested by EMSA using a 6% native PAGE. Three of them (T2 2118, T2 2123 and T2 2134) could bind to c-Myc:Max. The $\gamma^{32}$P end labeled E box probe was a positive control for these labeled oligonucleotides. The right arrow shows the position of DNA-c-Myc:Max complex, and the left arrow is the position of DNA-Max:Max complex.
Figure 14 Alignment of the full lengths of the three oligonucleotides (W-2118, W-2123 and W-2134) that could bind to c-Myc:Max according to EMSA (figure 13). The blue and yellow highlighting show some matches among these three oligonucleotides, and the black segments identify gaps inserted for alignment. Several Gs were involved in the matches.
Figure 15 Structure predictions of the full lengths of the three selected oligonucleotides (W-2118, W-2123 and W-2134). The full lengths of the three oligonucleotides that could bind to c-Myc:Max based on the results shown in figure 13 were folded using the Mfold program to predict the two-dimensional structures. For each oligonucleotide, the structure with the lowest free energy is shown. The blue dots denote the matches among the oligonucleotides based on the alignment results shown in figure 14, and the red circles show the loop or a stem-loop structure in these oligonucleotides in which accumulated matches were found. The sequence in the red circle of each oligonucleotide is shown in red letters beside the proposed oligonucleotide structure.
W-2118

T1-2118
dG = -4.08 T1218

T2-2118
dG = -4.65 M2118
Figure 16 Structure predictions and comparisons of the full length and truncated versions of the three putative aptamers. The full length and truncated versions of the three selected oligonucleotides that could bind to c-Myc:Max based on the results shown in figure 13 were folded using Mfold program. For each oligonucleotide, the structure with the lowest free energy is shown. The W-series show the full length version of each oligonucleotide, the T1-series denote the truncated version including the stem-loop structure of each oligonucleotide, and the T2-series stand for the truncated version including an E box complement and the random region. Diagram of these oligonucleotides are shown in figure 12.
Figure 17 The binding affinity of two truncated versions of the three selected oligonucleotides for c-Myc:Max. The two truncated versions of the three selected oligonucleotides were end labeled with $^{32}\text{P}$ and tested for binding affinity for c-Myc:Max by EMSA in a 6% native polyacrylamide gel. The $^{32}\text{P}$ end-labeled E box was a positive control and the end-labeled ss oligonucleotide “928” was a negative control for the labeled selected oligonucleotides, and the BSA was control for c-Myc:Max. The arrow shows the position of DNA-c-Myc:Max complex. The sequences of each oligonucleotide are described in Table 1.
Figure 18 The binding affinity of full length versions of the three selected putative aptamers for c-Myc:Max. The full length versions of the three selected oligonucleotides were end labeled with $^{32}$P and tested for binding affinity for c-Myc:Max by EMSA in a 6% native polyacrylamide gel. The $^{32}$P end-labeled E box was a positive control and the end-labeled ss oligonucleotide “928” was a negative control for the labeled selected oligonucleotides. BSA was the control for c-Myc:Max. The arrow shows the position of DNA-c-Myc:Max. The sequences of each oligonucleotide are described in Table 1.
Figure 19 Titration of the full length selected oligonucleotides W-2123 and W-2134. 32P end-labeled oligonucleotides W-2123 and W-2134 were diluted to 1.0μM, 100nM, 10nM, 1nM and 0.1nM. The diluted oligonucleotides were tested for binding affinity to c-Myc:Max by EMSA using a 6% native polyacrylamide gel. 32P end-labeled E box was a positive control for the labeled oligonucleotides and BSA was a non-specific binding control for c-Myc:Max. The upper arrow points at the position of DNA-c-Myc:Max complex, and the lower arrow points at the position of DNA-Max complex.
Figure 20a Ability of the full length and truncated versions of the three oligonucleotides from SELEX to compete for c-Myc:Max binding.

Figure 20b The selected oligonucleotides from SELEX do not bind to the E box binding domain of c-Myc:Max.
Figure 20a Ability of the full length and truncated versions of the three selected oligonucleotides from SELEX to compete for c-Myc:Max binding. $^{32}$P end-labeled E box was incubated with c-Myc:Max protein with or without the presence of other unlabeled oligonucleotides (competitors). The molar ratios of the labeled E box and each competitor were 1:50 and 1:100. The competition between each putative aptamer (competitor) and E box was tested by EMSA on a 6% native PAGE. Non-labeled API binding sequence was a control for non-specific binding, and non-labeled E box was a positive control for competition.

Figure 20b The selected oligonucleotides from SELEX do not bind to the E box binding domain of c-Myc:Max. Quantitative analysis of the data shown in figure 20a was performed by evaluating the amount of $^{32}$P end-labeled E box bound to the c-Myc:Max in the presence and absence of each non-labeled oligonucleotide. API was used as a control for non-specific binding. Non-labeled E box probe was a positive control for specific binding.
Figure 21 Binding curves of putative aptamers W-2123 and W-2134.

The oligonucleotides were end labeled by $^{32}$P at the 5 prime end and incubated at different concentrations from 1nM to 1μM with a fixed amount of c-Myc:Max (150nM). BSA was used to test binding background for each putative aptamer and the background binding were subtracted from the c-Myc:Max binding at the corresponding concentration of putative aptamer. All data were the average of triplicate values. The $K_d$ of each oligo was determined by Eadie-Hofstee analysis. (This experiment was done by Marjan Mokhtarian).
Figure 22 Developing a bivalent Myc aptamer. A, c-Myc:Max (red and green dimer) is a transcription factor that interacts with many proteins in addition to its DNA recognition element, the E box. B, A bivalent circular aptamer binds to the DNA binding site to block E box binding and to another site to block protein interactions. C, To create a bivalent aptamer a second, E box or Max aptamer sequence will be added to the aptamer. The trial extended aptamers will have the second sequence (blue) situated at various distances on a poly adenylate tail (tan) from the aptamer (black). D and E, The full E box will be created on the trial extended aptamers by annealing with the complementary sequence (blue) and will be tested for binding to c-Myc:Max. The cartoon in D shows an aptamer for which the extension is not long enough to create a bivalent aptamer. The cartoon in E shows an aptamer for which the polyA extension is sufficiently long to create a bivalent Myc-E box aptamer. F, A similar approach will be used to create a c-Myc:Max bivalent aptamer by adding a Max aptamer sequence at the end linked by a poly A linker of the proper length (By permission from a proposal of Dr. Marit Nilsen-Hamilton).
Figure 23 Subcloning of the c-Myc cDNA into pET41b(+) expression vector.
Figure 24 Subcloning of Max cDNA into pTrcHisA expression vector. The Max cDNA
Figure 23 Subcloning of the c-Myc cDNA into pET41b(+) expression vector. The c-Myc cDNA was subcloned from plasmid pHM6 into pET21b(+) and then from pET21b (+) into pET41b(+) between XhoI and BamHI. The primers used in subcloning were described in Table2. The expression vector pET41b(+) including the c-Myc cDNA was transformed into host cells BL21 (DE3) to express the c-Myc protein with a 6his tag at the N terminus.

Figure 24 Subcloning of Max cDNA into pTrcHisA expression vector. The Max cDNA was subcloned from plasmid pSG5puro into pTrcHisA between BamHI and EcoRI. The expression vector including Max cDNA was transformed into XL-1 Blue cells to express Max protein with 6his tag at N terminal (done by Tony Peterson).
Figure 25 Comassie blue stain (left) and Western blot (right) analysis of purified c-Myc protein from BL21 (DE3). c-Myc cDNA subcloned into pET41b(+) was expressed in BL21(DE3). A Ni-NTA column was used for purification under denaturing conditions with 7 M urea. The Comassie blue stain was used to check the apparent molecular mass of the purified c-Myc protein using 15% PAGE. The estimated mass was around 100 Kd. This is a fusion protein of c-Myc with a GST tag and a 6His tag as well as some amino acids from the expression vector. Dilution of 1:3000 for the primary mouse anti-his antibody and 1:500 for the goat anti-mouse IgG-HRP were used for the Western blot.
Figure 26 Coomassie blue stain (left) and Western blot (right) analysis for the purified Max protein from XL-1 blue. Max cDNA subcloned into pTrcHisA was expressed in XL-1 blue. A Ni-NTA column was used for purification under non-denaturing conditions. The Commassie blue stain using a 15% PAGE was performed to check the apparent molecular mass of the Max protein. The arrows show the position of the purified Max protein; the apparent molecular mass was around 26 Kd. This protein was a fusion protein of max with a 6His tag and some amino acids from the expression vector. Dilution of 1:3000 for the primary mouse anti-his antibody and 1:50,000 for the secondary goat anti-mouse IgG-HRP were used for the Western blot.
Figure 27 EMSA of the refolded c-Myc:Max heterodimer. Purified c-Myc and Max proteins were co-refolded by sequential dialysis against 1) BC500 containing 4M urea, 2) BC500 containing 2M urea, 3) BC500 containing 1M urea, 4) BC500 containing 0.5M urea, 5) BC500, 6) BC100 (twice) with 2 hours for each step. $^{32}$P end-labeled E box containing probe was used for checking the activity of the refolded product by EMSA on a 6% native PAGE. BSA was a negative control for c-Myc:Max. Max was another control for c-Myc:Max.
Figure 28 Subcloning of the Max and c-Myc cDNA into the pETDuet-1 expression vector to co-express these two proteins. The Max cDNA was subcloned from plasmid pHM6 into pETDuet-1 between SacI and HindIII, and the c-Myc cDNA was subcloned from pET41b(+) into the same pETDuet-1 between Bgl II and XhoI to co-express the Max and c-Myc protein. The primers including the proper enzyme digestion sites for PCR and subcloning are described in Table 2.
Figure 29 Coomassie blue stain (left) and Western blot analysis for the purified c-Myc (middle) and the Max (right) expressed from pETDuet-1 expression vector in BL21(DE3)CodonPlus RP host cells. TALON and Glutathione resins were used in series to purify the co-expressed GST-tagged fusion protein with c-Myc and 6his-tagged fusion protein with Max. The Commassie blue stain was used to check the apparent molecular mass of the c-Myc and Max proteins after resolving them through a 15% PAGE. The c-Myc fusion protein appeared to be about 98 Kd, and the Max fusion protein appeared around 26 Kd including his-tag. Dilution of 1:100 for the primary rabbit anti-c-myc (N-262) polyclonal IgG and 1:50,000 for the secondary goat anti-rabbit IgG-HRP were used in the Western blot for c-Myc; dilution of 1:3000 for the primary mouse anti-his antibody and 1:50,000 for the goat anti-mouse IgG-HRP were used in the Western blot for the Max.
Figure 30 EMSA of the co-expressed c-Myc:Max heterodimer. $^{32}$P end-labeled E box containing probe (refer to Table 1 for the sequence) was used for checking the activity of the co-expressed c-Myc and Max by EMSA through a 6% PAGE. BSA was a negative control for c-Myc:Max. Max was another control for c-Myc:Max.
REFERENCES


