CONTROLLING GENE EXPRESSION WITH SYNTHETIC MOLECULES

APPROVED BY SUPERVISORY COMMITTEE

_____

Thomas Kodadek, Ph.D. (Supervisor)

_____

David Corey, Ph.D. (Chair)

_____

David Mangelsdorf, Ph.D.

_____

Hongtao Yu, Ph.D.

DEDICATED TO MY PARENTS

CONTROLLING GENE EXPRESSION WITH SYNTHETIC MOLECULES

by

PRASANNA G. ALLURI

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

July, 2006

ACKNOWLEDGEMENTS

I sincerely thank my research advisor Dr. Tom Kodadek for being an outstanding mentor and providing an environment in the lab that was both stimulating and fun. His guidance, support and encouragement have made my time in the lab a memorable one. I would also like to express my deep gratitude to my committee members Drs. David Corey (Chair), David Mangelsdorf and Hongtao Yu for their insight, valuable advice and constructive criticism. I have been very fortunate to have worked in the company of several talented and generous members of the Kodadek laboratory, past and present, from whom I learned a great deal. I thank each and every one of them for their help and for being such a fun gang to work with. I particularly treasure my association with Hernando, Kiran and Reddy with whom I spent countless hours in joy and despair during our work on peptoid libraries. They have not also been my teachers but continue to be great friends. I appreciate my collaborators Bo and Peng, who built an excellent assay system for evaluating the biological activity profiles of my peptoid ligands. I am particularly indebted to Bo who mentored me during my rotation and continued to do so for the rest of my time in the lab. I have benefited immensely from discussions with Kevin Luebke and other members of CBI (now DTR) who brought diverse perspectives to various scientific problems. I gratefully acknowledge technical assistance and instrumentation services from various UT Southwestern associates in the north as well as south campuses. I am

also thankful to Ray Doss of Peter Dervan lab at Caltech for assistance with hairpin polyamide synthesis.

Finally, I cannot fully express my admiration and gratitude to all my family members; they have been a great source of inspiration and pride and have always stood by me. They are responsible for every success I have experienced so far and I owe a great deal to them.

CONTROLLING GENE EXPRESSION WITH SYNTHETIC MOLECULES

Publication No. _____

Prasanna G. Alluri, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2006

Supervising Professor: Thomas J. Kodadek, Ph.D.

Aberrant gene expression patterns have been implicated in several pathological states. Synthetic molecules capable of functionally mimicking native transcription factors and regulating gene expression in a specific and predictable manner may represent a new paradigm in drug development. Native transcription factors are minimally composed of two domains, a DNA-binding domain (DBD) and an activation domain (AD). Several synthetic DBDs capable of recognizing DNA in a sequence specific manner have been reported in the literature.

Furthermore, studies have demonstrated that coupling of these synthetic DBDs to peptides that are capable of acting as activation domains results in chimeric molecules that are capable of activating target gene expression. Since peptides and other biomolecules generally have poor cell-membrane permeability and are prone to rapid enzymatic inactivation inside cells, it is highly desirable to develop artificial molecules that are capable of mimicking native ADs. Towards this goal, a comprehensive methodology for the synthesis, screening and characterization of large peptoid libraries has been developed. Peptoids are a new class of peptidomimetic compounds that are resistant to proteolytic cleavage and are relatively simple and cheap to synthesize. One of the combinatorial libraries was screened against CBP (CREB-binding protein), an important transcriptional co-activator, and three novel, low micromolar affinity ligands were isolated. A cell-based reporter gene assay was employed to assess the cell permeability and transcription activation potential of the synthetic ligands in live mammalian cells. The assay consists of transfecting into HeLa cells a luciferase reporter gene harboring Gal4 binding sites and a construct in which the ligand binding domain of the Glucocorticoid receptor has been fused to Gal4 DBD. The cells are treated with the CBP-binding peptoids that have been chemically coupled to a dexamethasone derivative. Among the three peptoids tested, one of the molecules as a steroid conjugate, has been found to activate the transcription of a reporter gene nearly 1000-fold suggesting that it may be acting as an activation domain

surrogate. The mechanistic aspects of the observed transcriptional activity of the peptoid-steroid conjugate remain to be elucidated.

TABLE OF CONTENTS

PUBLICATIONS

(1) **Alluri, P.G**., Liu, B., Yu, P., Xiao, X., and Kodadek, T., "Isolation and characterization of coactivator-binding peptoids from a combinatorial library." *Submitted.*

(2) Liu, B., Yu, P., **Alluri, P.G**., and Kodadek, T., (2005) "Transcriptional activation by an imidazole-pyrrole hairpin polyamide-dexamethasone conjugate in living mammalian cells." Mol. BioSyst. **1**: 307-315.

(3) Liu, B., **Alluri, P. G**., Yu, P., and Kodadek, T., (2005) "A potent transactivation domain mimic with activity in living cells." J. Amer. Chem. Soc. **127**: 8254-8255.

(4) Kodadek, T., Reddy, M. M., Olivos, H. J., Bachhawat-Sikder, K., and **Alluri, P. G**., (2004) "Synthetic molecules as antibody replacements." Acc. Chem. Res. **37**: 711-718.

(5) **Alluri, P. G**., Reddy, M. M., Bachhawat-Sikder, K., Olivos, H. J., and Kodadek, T., (2003) "Isolation of protein ligands from large peptoid libraries." J. Amer. Chem. Soc. **125**: 13995-14004.

(6) Olivos, H. J., **Alluri, P. G.**, Reddy, M. M., Salony, D., and Kodadek, T., (2002) "Microwave-assisted solid-phase synthesis of peptoids." Org. Lett. **4** :4057-4059

LIST OF FIGURES

LIST OF TABLES

LIST OF SCHEMES

ABBREVIATIONS

AD:	Activation Domain
AEEA:	Acetyl-ethyleneglycol-ethylamine
AH:	Amphipathic Helix
aPP:	Avian Pancreatic Polypeptide
AR:	Androgen Receptor
ATF:	Artificial Transcription Factor
ATP:	Adenosine triphosphate
t-Boc:	tert-Butoxy carbonyl
BSA:	Bovine Serum Albumin
cAMP:	Cyclic -Adenosine mono-phosphate
CBP:	CREB-Binding Protein
CCD:	Charge-Coupled Device
COX:	Cycloxygenase
CREB:	cAMP-Response Element-Binding Protein
DBD:	DNA-Binding Domain
DCM:	Dichloromethane
Dex:	Dexamethasone
DHFR:	Dihydrofolate reductase
DHT:	Dihydrotestosterone

DIC:	N, N'-Diisopropylcarbodiimide
DIPEA:	Diisopropylethylamine
DMAP:	N, N-Dimethylaminopyridine
DMF:	N, N-Dimethylformamide
DMSO:	Dimethyl sulfoxide
DNA:	Deoxyribonucleic acid
DP:	Differential Power
ESX:	Epithelial-restricted with Serine Box
ETS-1:	E-Twenty-Six-1
EXD:	Extradenticle
FKBP:	FK-506-Binding Protein
Fmoc:	Fluorenylmethoxy-carbonyl
Gal:	Galactose
Gal80BP-A:	Gal80-Binding Peptide-A
GCN4:	Gene Control of amino acid synthesis Non-de-repressible mutant 4
GR:	Glucocorticoid Receptor
GRE:	Glucocorticoid Response Element
GST:	Glutathione S-Transferase
HAT:	Histone Acetyl Transferase
HBD:	Hormone-Binding Domain

HBTU: O-(Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate

HEK: Human Embryonic Kidney

HeLa: Henrietta Lacks

Her: Human Epithelial growth factor Receptor

HIV-I: Human Immunodeficiency Virus-I

HOBt: 1-Hydroxybenzotriazole

HOX: Homeobox

HP: Hairpin Polyamide

HPLC: High Pressure Liquid Chromatography

HSP: Heat Shock Protein

IL: Interleukin

IPTG: Isopropyl-β-D-1-thiogalactopyranoside

ITC: Isothermal Titration Calorimetry

KBPo: KIX domain-Binding Peptoid

K_D: Equilibrium Dissociation Constant

KID: Kinase Inducible Domain

LBD: Ligand Binding Domain

LEF-1: Leukocyte Enhancer Factor-1

MALDI-TOF: Matrix-Assisted Laser Desorption Ionization-Time of Flight

MBHA: Methylbenzhydrylamine

MBP: Maltose Binding Protein

MCIP: Modulatory Calcineurin-Interacting Protein

Mif: Mifepristone

MS: Mass Spectrometer

MVP: Major Vault Protein

c-Myc: Cellular-Myelocytomatosis

NF-κB: Nuclear Factor-κB

NHS: N-Hydroxy succinimide

NMM: N-Methyl morpholine

NMO: N-Methylmorpholine-N-oxide

NMR: Nuclear Magnetic Resonance

OD: Optical Density

PAGE: Polyacrylamide Gel Electrophoresis

PBS: Phosphate Buffered Saline

PCR: Polymerase Chain Reaction

PDD: Protein-DNA Dimerizer

PEG: Polyethylene glycol

PKA: Protein Kinase A

PNA: Peptide Nucleic Acid

PolII:	Polymerase II
PR:	Progesterone Receptor
RNA:	Ribonucleic acid
RPM:	Revolutions Per Minute
RT:	Room Temperature
SDS:	Sodium dodecyl sulfate
SH3:	Src Homology 3
SREBP:	Sterol Regulatory Element-Binding Protein
STAT:	Signal Transducer and Activator of Transcription
Sur:	Sulfonylurea Receptor
SV40:	Simian Virus 40
SWI / SNF:	Switch / Sucrose Non-Fermenting
TBP:	TATA-Binding Protein
TBST:	Tris-Buffered Saline with Tween
TF:	Transcription Factor
TFA:	Trifluoroacetic acid
TFO:	Triple-helix-Forming Oligonucleotide
THF:	Tetrahydrofuran
TPAP:	Tetrapropylammonium perruthenate
tRNA:	Transfer RNA

UAS: Upstream Activator Site

VEGF: Vascular Endothelial Growth Factor

VP16: Herpes simplex Viral Protein 16

CHAPTER 1: INTRODUCTION

1.1. General introduction

Aberrant gene expression patterns have been implicated in several pathological states[1, 2]. Since regulation of gene expression primarily occurs at the level of transcription, intense efforts have been focused on understanding the basic mechanisms underlying transcriptional regulation. Moreover, there has been considerable interest in developing strategies for regulating the expression of genes by controlling their transcription. Among various such strategies, the use of synthetic, cell-permeable molecules for modulating gene expression remains highly attractive and promising for various reasons[3]. Such molecules have the potential to serve not only as a novel class of therapeutic agents, but also as powerful tools in functional genomics and basic biology.

1.2. Mechanism of transcriptional activation

Eukaryotic transcriptional activation is a highly complex and regulated event that involves the establishment of a multitude of protein-DNA and protein-protein interactions[4]. Activation of transcription is mediated by a class of proteins referred to as transcription activators[5]. Upon receiving the appropriate stimulus, transcription activators bind to their target genes in the vicinity of the promoter, usually upstream of the TATA box. This is followed by the recruitment of RNA polymerase II holoenzyme and other components of the basal

transcription machinery such as TFIID, a multi-protein complex which include the TATA box binding protein (TBP), and to form the pre-initiation complex[6]. The pre-initiation complex is composed of the RNA PolII, various general transcription factors such as TFIIA, TFIIB, TFIID etc., and components of a complex of about twenty proteins referred to as the mediator. Once the entire pre-initiation complex has assembled, the promoter region of the DNA is locally "melted" in an ATP-dependent manner to facilitate the polymerase to associate with and read the template strand. This is followed by promoter escape and elongation steps where the pre-initiation complex dissociates from the promoter and traverses along the gene transcribing it[7]. It has been suggested that the mediator remains at the promoter during this process and assists in the re-assembly of a pre-initiation complex for another round of transcription[8].

1.3. Role of activators in transcription activation

Transcriptional activators promote transcription activation through multiple mechanisms. For instance, they make direct contacts with various components of the mediator complex and, thus, play a role in the recruitment of the pre-initiation complex to the promoter[6]. They also help in making the DNA more accessible to various transcription factors by recruiting chromatin modifying / remodeling machinery such as histone acetyl transferases (HATs)[9] and SWI/SNF protein complexes[10, 11]. HATs acetylate lysine residues on histone

tails and loosen the chromatin[9], while SWI/SNF use energy derived from ATP hydrolysis to make the DNA more accessible[10, 11]. There is also evidence to suggest that transcription activators play an important role in facilitating promoter escape, elongation and re-initiation[7, 8]. Finally, recent experiments have unearthed yet another role of activators in transcription – recruitment of components of the proteasomal complex to activated promoters[12], a necessary step in the activation of the Gal genes in yeast and probably many others.

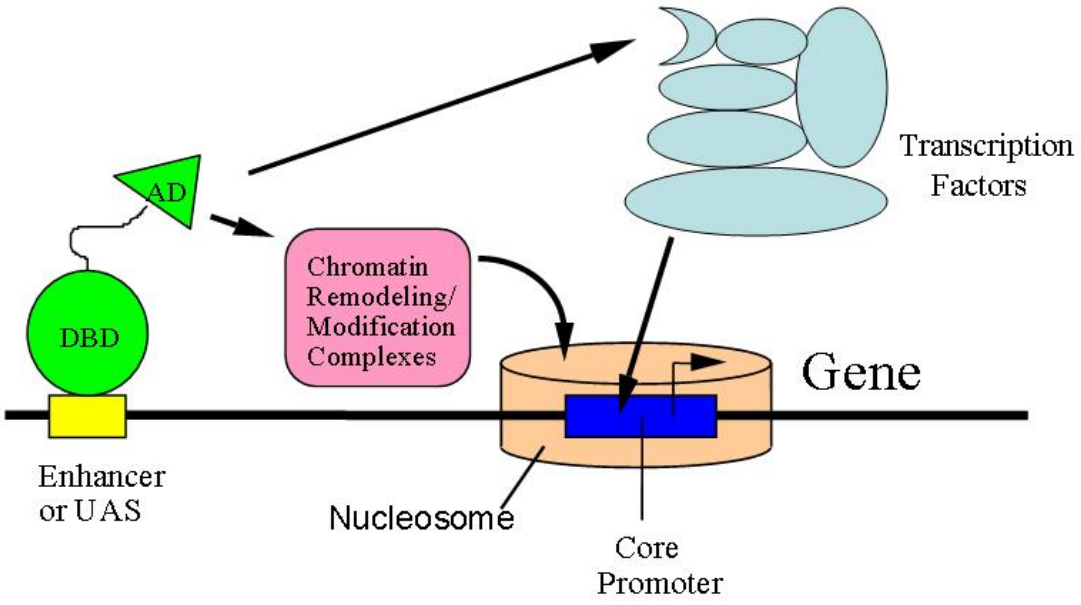


Figure 1.1: A schematic depiction of the role of transcription activators in stimulating target gene expression. (UAS = Upstream Activator Site).

1.4. Domain organization of activators

Transcription activators are modular in nature and are minimally comprised of two distinct and separable domains: the DNA-binding domain (DBD) and the activation domain (AD). The DBD is responsible for conferring gene specificity to the activator by recognizing small stretches (6-8 base pairs) of DNA near the promoter of one or a small number of genes. The AD has a more general function and is responsible for establishing various protein-protein interactions necessary for the induction of the gene. Domain swapping experiments have illustrated that the DBD and AD of transcription activators are exchangeable. For instance, fusion of VP16, a potent viral AD, to Gal4 DBD results in a new hybrid protein that retains the ability to activate Gal4 responsive genes[13].

The modular nature of transcription activators points towards a potential strategy for construction of artificial transcription activators; synthetic molecules capable of mimicking the native DBD and AD, when linked together appropriately, may result in an artificial transcription activator that may retain some or all the functions of a native activator. Thus, it is not surprising that much efforts have been focused over the last several years in developing synthetic molecules with DBD- and AD-like activities[14].

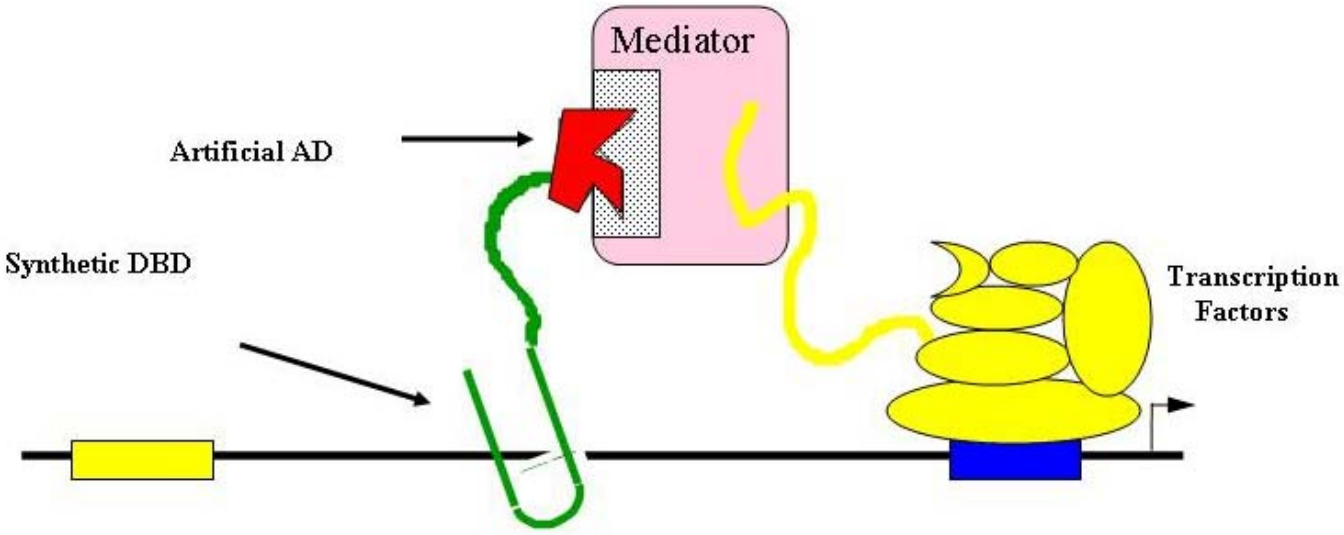


Figure 1.2: A general strategy towards development of artificial transcription activators. (AD = Activation Domain; DBD = DNA-Binding Domain).

1.5. Synthetic DBDs

Synthetic molecules that can recognize target DNA sequences with high affinity and specificity may serve as powerful tools for regulation of gene expression. In addition to their potential to target AD-mimetic molecules to specific genes to up-regulate their expression, they could also down-regulate gene expression by binding to specific sites near the promoter and interfering with the binding of transcription factors. In fact, giant strides have been over the last decade or so in the design and synthesis of such molecules and their utility in regulating gene expression has been well documented in the literature. Hairpin polyamides[15] and peptide nucleic acids (PNA)[16] are two of the most popular and extensively characterized groups of synthetic DBD mimetic molecules.

1.5.1. Hairpin Polyamides

Hairpin polyamides are oligomers of N-methylpyrrole and imidazole derivatives developed largely through the pioneering efforts of Dervan and his colleagues[15]. These molecules bind to DNA in the minor groove in an anti-parallel fashion and are capable of distinguishing all the four DNA base pairs (A/T, T/A, G/C and C/G) based on well developed pairing rules. Thus, it is possible to design polyamides to recognize pre-determined DNA sequence with high affinity and specificity in vitro. Typical polyamides are eight rings long and recognize stretches of DNA that are 6 base pairs long. A γ-aminobutyric acid

residue introduced between rings four and five induces a “hairpin” turn in the molecule and aligns the rings in an anti-parallel orientation. Another extremely attractive feature of polyamides is that they are sometimes cell permeable, at least in certain cell lines[17].

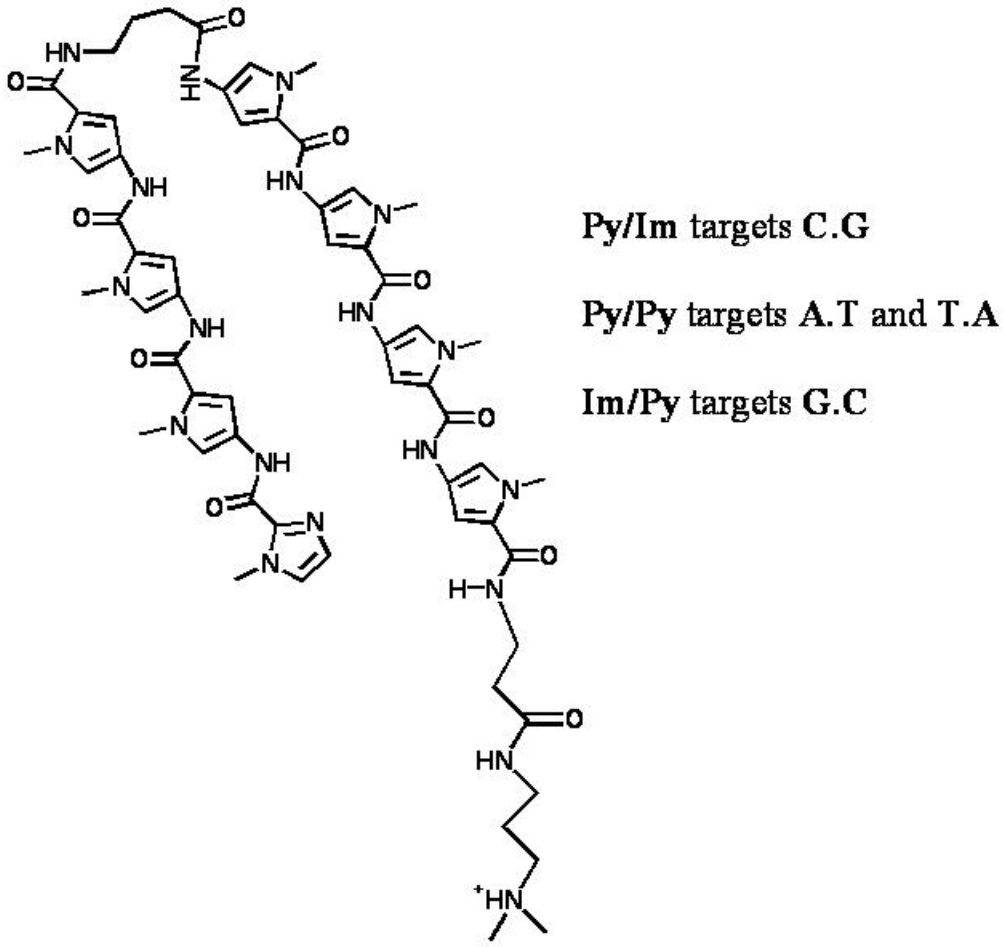


Figure 1.3: The chemical structure of a pyrrole (Py) and imidazole (Im) containing polyamide is shown along with the pairing rules for the recognition of duplex DNA.

Several studies have demonstrated the utility of polyamides in modulating gene expression. For instance, Dervan and co-workers designed an eight ring polyamide to target the recognition site of transcription factor TFIIIA, a minor groove binding transcription factor essential for the transcription of 5S RNA genes[17]. When fibroblasts derived from *Xenopus* kidney were treated with this molecule, marked inhibition of 5S RNA gene expression was observed without significant inhibition in tRNA transcription. In another study, a series of polyamides were designed to recognize DNA sequences immediately adjacent to the binding sites of transcription factors TBP, Ets-1 and LEF-1 on the HIV-I promoter. Dervan and colleagues went on to show that the polyamides inhibited both the binding of these transcription factors to their cognate sites and RNA polymerase II-mediated transcription of HIV-I transcription in cell-free systems. Impressively, the polyamides caused >99% inhibition of HIV-I replication in human lymphocytes[18].

1.5.2. Peptide Nucleic Acids

Peptide Nucleic Acids (PNAs), introduced by Nielsen and co-workers, represent the second class of molecules that hold promise to serve as synthetic DBD surrogates[16]. PNAs are essentially modified oligonucelotides, where the phosphodiester backbone is replaced with amide linkages, resulting in abrogation of unfavorable electrostatic repulsive forces between the negatively charged

phosphate groups found in oligonucleotides. This structural modification also makes PNAs resistant to nucleases / proteases and improves their affinity towards complementary single stranded DNA and RNA relative to a standard nucleic acid[19].

Among the various classes of PNAs described, bis- PNAs have been shown to be the most facile in invading relaxed, double stranded DNA resulting in the formation of a tetra-stranded complex in which one strand of PNA invades the duplex DNA, while the second strand establishes Hoogstein base pair interactions[20]. Freier and her colleagues used this approach to inhibit gene transcription by designing a bis-PNA to interfere with transcription factor binding[21]. Thus, they targeted a homopurine / homopyrimidine site on the IL-2Rα promoter which overlapped with NF-κB binding site. They found that the binding of PNA to this site correlated well with inhibition of NF-κB binding. However, inhibition of transcription by PNA necessitated pre-incubation of the PNA with DNA under low salt conditions. Thus, one of the limitations of bis-PNAs is that efficient strand invasion of relaxed, duplex DNA occurs only at polypurine-polypyrimidine sites under non-physiological (low ionic strength) conditions. However, recent efforts by Corey and co-workers suggest that the efficiency of strand invasion can be substantially improved even at physiologically relevant ionic strength conditions by appending positively charged amino acids to PNAs[19]. Furthermore, polypyrimidine PNAs containing

mixed base extensions, referred to as tail clamp PNAs have been employed to extend the range of DNA sequences accessible to PNAs[19, 22]. Future experiments are necessary to determine if the strand invasion rate of bis-PNAs can be sufficiently enhanced to enable targeting relaxed, duplex DNA inside cells. Nevertheless, there have been a few reports of antigene PNAs with activity in tissue culture and animal studies. For instance, Tyler and coworkers have reported an antigene PNA-mediated downregulation of neurotensin receptor expression in rats up on microinjection of the molecule directly into the brain[23]. However, the lack of mismatch control molecules in this study makes it difficult to determine the specificity of observed effect. In an effort to develop a general strategy to enhance the cell permeability of PNAs, Morris and coworkers covalently attached a dihydrotestosterone (DHT) molecule to the amino terminus of an antigene-PNA designed to inhibit the expression of c-myc, a gene implicated for its role in proliferation and malignant transformation of human cells[24]. They demonstrated that the nuclear localization of PNA-DHT was dependent on the expression of androgen receptor. The DHT-mediated nuclear uptake of the PNA has been shown to be associated with significant suppression of c-myc gene expression. The generality of this approach as well as the impact of the PNA-DHT-androgen receptor complex on the DNA-binding specificity of PNAs needs to be determined in carefully designed experiments. Finally, Corey and coworkers have recently reported a clever strategy for PNA-mediated silencing of gene

transcription by targeting sequences in the open complex in target genes. The formation of open complexes is observed in all actively transcribing genes and is induced by assembly of pre-initiation complex on the promoter of the target gene. More importantly, the open complex contains a stretch of single stranded DNA which can be readily targeted with high efficiency and specificity with PNAs. Thus, most of the difficulties associated in targeting relaxed, double stranded DNA are circumvented by this approach. The authors employed this attractive strategy to not only inhibit the transcription of human progesterone receptor gene[25], but also several other genes such major vault protein (MVP), androgen receptor (AR) and cyclooxygenease-2 (COX-2)[26], illustrating the generality of this approach. The promising results from these studies indicate that PNAs may serve as highly efficient synthetic DBD replacements. The development of efficient and general strategies for cellular delivery of PNAs could further enhance the broad utility of this remarkable class of molecules.

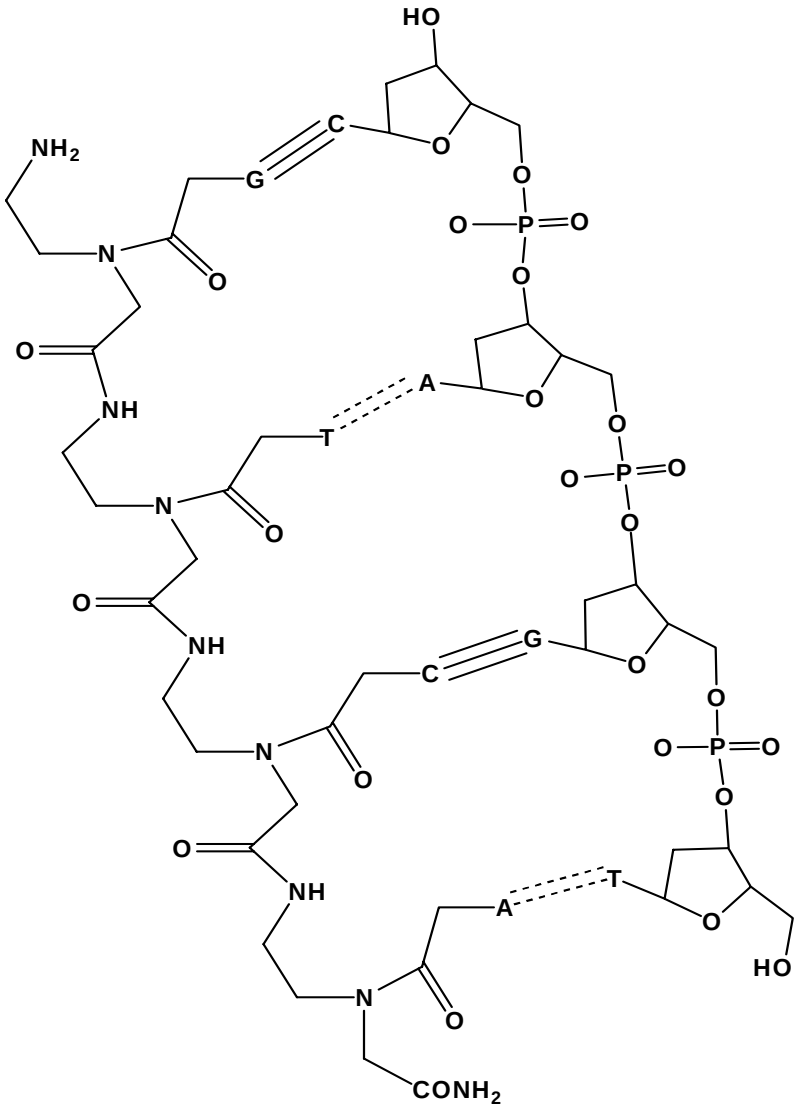


Figure 1.4: Watson and Crick base pairing between a PNA and DNA strand.

(A = Adenine, T = Thymine, G = Guanine and C = Cytosine)

1.6. Synthetic activation domains

As described earlier, the role of native activation domains in activating transcription is complex and involves multiple mechanisms. The precise target of most native activation domains such as Gal4 and VP16 is still obscure. Moreover, the structural characteristics of most activation domains remain unclear. This has led to classification of various activation domains based on the preponderance of certain types of amino acid residues. Thus, transcription activators such as Gal4, GCN4 and VP16 are referred to as acidic activators. Similarly, glutamine rich[27] and proline rich[28] activators have also been described. The lack of precise structural and molecular target information of activation domains presents a difficulty in the development of synthetic activation domain mimetic molecules. A brief review of various strategies devised to overcome this difficulty towards development of synthetic activation domain mimetic molecules follows.

1.6.1. Genetic assays

Genetic assays have been used as powerful tools for sifting through large peptide and other bio-molecule libraries and isolating molecules with desired biological functions. Such an approach is also useful for obtaining activation domain mimetic molecules, without having to worry, at least initially, about the mechanism by which they cause transcription activation. Several groups have

demonstrated the utility of this approach in obtaining activation domain surrogates.

In one of the earliest reports of such an approach, Ma and Ptashne screened a peptide library encoded by *E .coli* genomic fragments fused to Gal4 DBD for activation of transcription in yeast[29]. The induction of β-galactosidase, a reporter gene that confers blue color to yeast colonies, was used as a read-out for activation of transcription. From among 154 colonies that tested as positive out of approximately 15,000 transformants, the authors characterized fifteen of the shortest DNA fragments and determined the corresponding amino acid sequences. Varying in length from 12 to 81 amino acids, the peptides were rich in acidic residues and the activation potential of these molecules correlated with their net negative charge. It should be noted that native transcription activators in yeast such as Gal4 and GCN4 also have a large number of negatively charged amino acid residues in their sequences.

As an extension of this work, the same group also screened random libraries encoding 8 and 6 amino acid residues fused to Gal4 DBD in a format similar to the previous study[30]. Several hits isolated from this screen were shown to be capable of activating transcription comparable to and, in same cases, higher levels than Gal4, a potent yeast transcription activator. In marked contrast to their previous study, however, the activation domains isolated in this screen had few or no acidic residues. In fact, P201, an eight residue peptide and the most

potent among the activation domains isolated in this study, did not bear any acidic (or basic) residues at all. Subsequent studies have established Gal11, a yeast coactivator protein, as an essential target of P201.

The ease of isolation of activation domain mimetic peptides from naive libraries prompted some groups to explore if non-protein based bio-molecules such as aptamers are capable of functioning as activation domains. In two independent studies, co-workers of Ptashne and Liu employed a similar approach to address this question[31, 32]. A random sequence of 10-80 residues was displayed in the loop region of a RNA molecule that has been shown to adopt a stable hairpin like secondary structure. They also engineered into the RNA molecule a sequence that specifically binds to bacteriophage coat protein MS2. They, then, used a LexA-MS2 fusion protein to target the RNA molecules to the proximity of a reporter gene promoter, the induction of which was used as a readout for the selection. Thus, this method which uses a non-covalent strategy for the delivery of a non-peptidic molecule to a native DBD could be employed, in general, for testing the activity of non-native activation domains. Both groups reported several RNA molecules that activated transcription in this assay suggesting that it should be possible to isolate non-peptidic molecules with activation domain-like properties. However, the target(s) mediating the activity of these activation domain-mimetic molecules has not been characterized.

Although genetic assays offer an attractive and convenient means for isolating activation domain mimetic molecules, they introduce several limitations. For instance, one is restricted to bio-molecule-based activation domains, which, in general, tend to have poor cell permeability and bio-availability, and are prone to enzymatic degradation inside cells. Furthermore, identifying the targets of the activation domains isolated in these assays is relatively difficult. Therefore, other methodologies for isolating activation domain mimetic molecules are desirable to overcome these limitations.

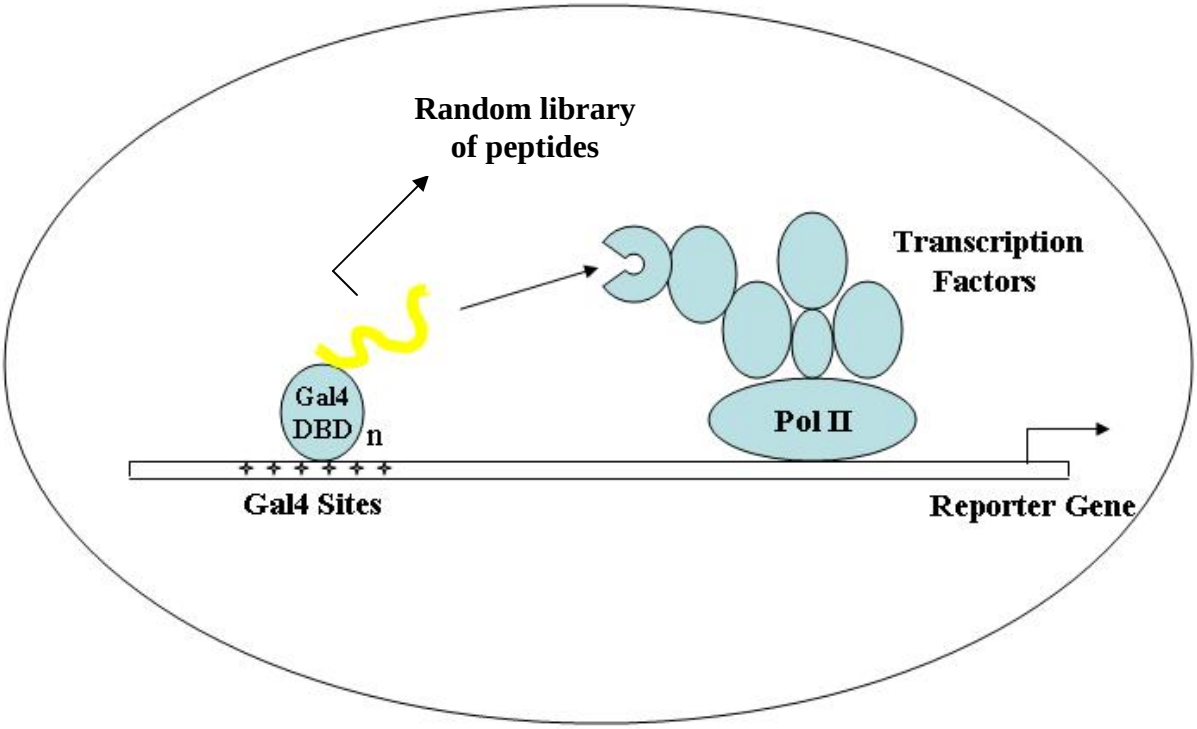


Figure 1.5: A schematic view of the genetic assay employed by Ptashne and co-workers to isolate novel peptidic activation domains.

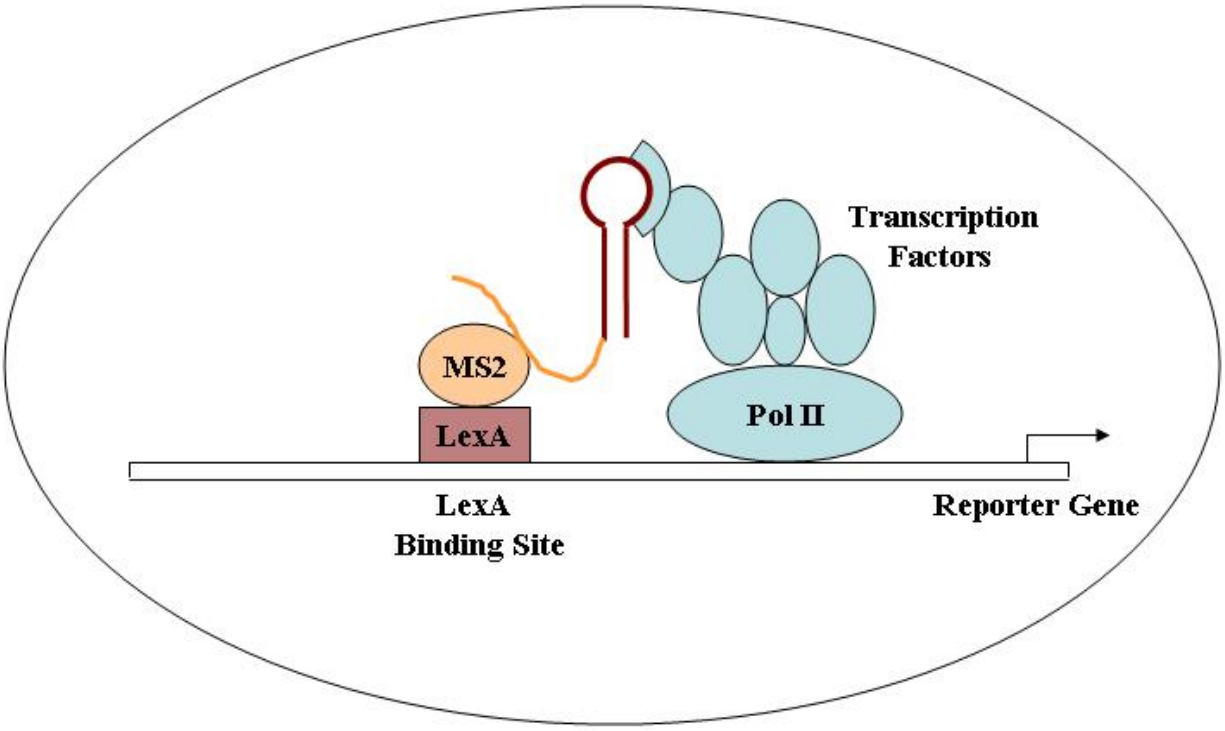


Figure 1.6: A schematic view of a genetic assay employed by Ptashne, Liu and co-workers to isolate RNA-based activation domains.

1.6.2. Binding assays

Recognizing the inherent limitations of genetic assays, Han and Kodadek explored the utility of binding assay-based screening experiments in the isolation of activation domain mimetic molecules[33]. Such an approach where compound collections of any class can be screened against appropriate transcription factors can be quite powerful as it could pave the way for the isolation of enzymatically stable, low molecular weight, synthetic molecule-based activation domains.

In a proof of principle experiment, the authors panned a 20mer phage display library of approximately 10^9 peptides against Gal80 over six rounds of selection under stringent conditions to demand specific binders to the target protein[33]. Gal80 is a yeast repressor protein that binds to Gal4 and muzzles its transcription activity. The authors hypothesized that molecules isolated to bind to Gal80 may compete with native activation domains for binding to the repressor and may themselves serve as artificial activation domains. They isolated two unique peptides, both of which were rich in acidic residues, a feature common in native activation domains such as Gal4. One of the peptides subjected to detailed study, Gal80BP-A, bound to Gal80 with an affinity of 300nM, as determined by a fluorescence polarization assay. Furthermore, the peptide not only competed with Gal4 for binding to Gal80, but also appeared to be capable of binding to co-activator protein Gal11, albeit with lower affinity (~ 1μM). More importantly, the peptide activated transcription robustly in vivo when fused to Gal4 DBD

validating that molecules selected solely on their ability to bind to appropriate transcription factors can indeed function as activation domains.

The generality of this protein-targeted approach to the isolation of artificial ADs has since been validated by two other groups. Montminy and his colleagues screened a phage display peptide library of 8mers against the KIX domain of CREB-binding protein (CBP), a mammalian coactivator, and isolated several peptides that shared a consensus motif[34]. A second library incorporating this consensus motif was constructed and another round of selection was performed to isolate several peptides. The most potent binder among them bound to KIX domain with a modest affinity of ~16μM and supported a 40-fold increase in the transcription activity of a reporter gene in mammalian cells upon fusion to Gal4 DBD.

More recently, Mapp and coworkers synthesized two split and pool combinatorial peptide libraries of 8mers with four randomized positions; one was designed to mimic native activation domains, while the second was biased against them[35]. The authors screened the libraries against Gal11 and identified three classes of peptides that bound to the target non-competitively and, presumably, at three distinct sites on the protein. Two of the peptides selected for further study were fused to LexA DBD and shown to activate transcription to levels comparable to VP2, a fragment of the potent viral activation domain VP16. In stark contrast to the Gal80-binding peptide, however, the AHYYPSE peptide,

which was the most potent among the peptides isolated in this study, failed to activate transcription in mammalian cells[36]. This observation suggests that the Gal80-binding peptide and the Mapp peptide recognize different targets (although Gal80-binding peptide has also been shown to bind to Gal11 in vitro) or distinct sites on the same target.

Target	*Sequence* of activation domain*
Gal80	YDQDMQNNTFDDLFWKEGHR
KIX domain of CBP	WAVYELLF
Gal11	AHYYYPSE

* Only the sequence of the most potent among the various activation domains isolated in each study is shown.

Table 1.1: A table showing the sequences of various peptide activation domains isolated from screening the protein targets shown in the table.

1.6.3. Designer activation domains

Another approach towards generation of artificial activation domains involves design of such molecules based on rational principles and empirical observations. In one of the earliest examples of such an approach, Ptashne and co-workers rationally designed a 15-residue peptide comprising of acidic and hydrophobic residues, which in principle should adopt an amphipathic alpha helical orientation[37]. The peptide, when fused to Gal4 DBD, activated transcription in yeast, while a scrambled version of the peptide which is not anticipated to adopt a helical structure did not.

In another study, Verdine and co-workers synthesized an unnatural transcription activator by simply replacing the natural amino acids of a 29-mer peptide derived from N-terminus of VP16 with the corresponding D amino acids[38]. The artificial peptide was then appended via a flexible linker to the immunosuppressive drug FK506, which has been derivatized to abrogate its immunosuppressive properties while retaining high affinity to its target receptor, FKBP. To facilitate targeting of the artificial activation domain to the promoter of a reporter gene construct, the authors co-transfected Jurkat cells with a reporter plasmid bearing five Gal4 DBD sites and a plasmid encoding a fusion of Gal4 DBD with three tandem repeats of FKBP12. To enhance the permeability, the authors incorporated the drug conjugates into liposomes. Treatment of cells with the drug conjugate resulted in > 500-fold induction of the reporter gene. This

study highlights one of the characteristics critical for artificial activation domains, protease resistance. For instance, the L-peptide-drug conjugate failed to activate transcription in their assay, presumably due to proteolytic inactivation inside cells. Another desirable feature for activation domain mimetic molecules is cell membrane permeability. Therefore, activation domain surrogates with better cell membrane permeability properties must be developed to extend their practical utility.

In an attempt to generate small molecule based activation domain mimetic molecules, Mapp and co-workers recently designed a series of isoxazolidines substituted with functional groups commonly found in native activation domains such as phenyl, hydroxyl, isobutyl and carboxylic acid[39]. These molecules were then chemically conjugated to methotrexate, a high affinity ligand of dihydrofolate reductase (DHFR). In a strategy similar to the Verdine study, the authors then expressed a LexA DBD-DHFR fusion protein to facilitate the targeting of the small molecule activation domains to a reporter gene bearing LexA binding sites. The activation potential of the small molecule drug conjugates was measured in an in vitro assay by mixing each molecule with the LexA DBD-DHFR fusion protein and the reporter gene in the presence of HeLa nuclear extracts. For comparison purposes, a conjugate of methotrexate and ATF14, a fragment of VP16 capable of activating transcription, was synthesized and subjected to the in vitro transcription assay. Among the five isoxazolidines

tested, two structurally related compounds showed remarkable activity with levels approaching ATF14. The authors speculated that the proteolytic stability and the cyclic nature of the isoxazolidines, which likely resemble the helical structure of most target bound activators, may have contributed to the potency of these molecules.

1.6.4. Hybrid approaches

Recently, Schepartz and co-workers used their previously reported "protein grafting" strategy, which incorporates both rational design principles and combinatorial chemistry, to design low micromolar affinity mini-protein ligands for KIX domain of CBP[40, 41]. In this approach avian pancreatic polypeptide (aPP), a small, well-folded protein is used as a scaffold to display a binding epitope. Thus, the sequences of aPP and the helix B of KID domain of CREB, which is involved in interacting with the KIX domain, were aligned and a new hybrid protein is designed by selecting residues from aPP required for formation of α-helix and residues from the KID domain responsible for recognizing KIX. Some of the positions in the hybrid protein were randomized and a library of 5 x 10^6 mini-proteins was generated in a phage display format to facilitate selection of high affinity ligands for KIX. Recognizing that phosphorylation of KID by protein kinase A (PKA) plays an important role in contributing to the binding affinity of the KID-KIX interaction, the recognition site of PKA in KID was retained in the mini-protein and the library was subjected to in vitro phosphorylation by PKA prior to selection. Two ligands selected in this study, PPKID4 and PPKID6 were found to be competent of activating the transcription of a reporter gene (containing Gal4 binding sites) in HEK 293 cells when tethered to Gal4 DBD. Furthermore, their transcriptional response was found to be

dependent on forskolin treatment, which presumably induces phosphorylation of the two ligands by PKA.

In an approach that integrates aspects of all the strategies discussed above-rational design, combinatorial chemistry and cell-based screening, Uesugi and co-workers developed a small molecule surrogate for transcription activation domain of ESX (epithelial-restricted with serine box), which interacts directly with Sur2 (a component of the human mediator complex) and causes the over-expression of the oncogene Her2 in malignant breast cancer cells[42]. Recognizing that a tryptophan residue in the activation domain of ESX is critical for its interaction with Sur2, the authors designed a 2422-compound small molecule library rich in indole-like moieties. The library was screened in a cell-based assay designed to measure the ability of the compounds to inhibit the transcription activity of ESX and a compound, referred to as adamanolol, was isolated. In addition to suppressing the ability of various Her2-positive breast cancer cell lines, adamanolol competed in vitro with a peptide fragment derived from ESX for binding to Sur2 with an IC_{50} of 8 μM. Therefore, the authors suspected that adamanolol exerts its effects by disrupting protein-protein interactions between ESX and Her2. Subsequently, the authors synthesized several derivatives of adamanolol to carry out structure-activity relationships and identified a second generation molecule they termed wrenchnolol, which was water soluble and more potent in disrupting ESX-Sur2 interaction in vitro[43]. Furthermore, wrenchnolol

diminished the expression of Her2 protein in breast cancer cells SK-BR3 and a biotinylated version of the molecule retained sur2 protein from nuclear extracts.

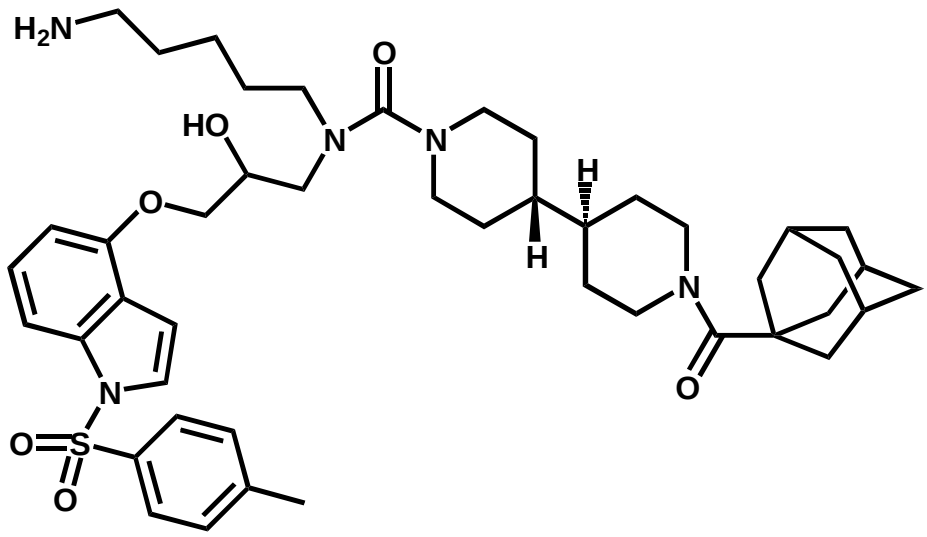


Figure 1.7: Wrenchnelol.

The chemical structure of an activation domain-mimetic small molecule developed by Uesugi and co-workers.

1.7. Artificial transcription activators: state of the art

Initial proof-of-principle experiments to demonstrate the feasibility of using small molecules to control gene regulation were carried out by Schreiber, Crabtree and co-workers. In a clever strategy the authors used FK1012, a dimeric version of FK506, to create a non-covalent interaction between the AD of NF3V1 (NF3V1 = transactivation domain of VP16 + three FKBP12 domain repeats + N-terminal nuclear localization sequence from SV40 large T antigen) and DBD of GF3 (GF3 = Gal4 DBD + three FKBP12 domain repeats)[8]. Thus, the authors expressed fusions of the two domains, AD and DBD, with FKBP12 in Jurkat cells before adding the chemical "glue" (FK1012) to "stick" the two domains together. The physical association of the AD with DBD reconstituted a functional transcription activator, as shown by the induction of a reporter gene. While this approach is limited to engineered cells, it nevertheless, demonstrated the principle that small molecules can, indeed, serve as powerful tools for controlling gene expression. The impressive advances in design of synthetic AD and DBD mimetic molecules, however, opened up the possibility of developing a general strategy for controlling gene expression even in non-engineered cells. This strategy, in essence, consists of chemically conjugating, via a suitable linker, chemical moieties that have been designed or selected to functionally replace the native activation and DNA-binding domains. A handful of groups have made promising breakthroughs in this rather ambitious goal.

In a pioneering effort in this area, Dervan, Ptashne and co-workers synthesized a hairpin polyamide-peptide conjugate designed to activate the transcription of a reporter gene bearing multiple polyamide binding sites[44]. Thus, an eight ring polyamide was designed to bind to a palindromic site of 5'-TGTTAT-3' separated by seven base pairs. The equilibrium dissociation constant of the polyamide DNA complex, as determined by foot printing analysis was ~ 1nM. A 20 residue peptide, AH, previously designed by Ptashne and co-workers and shown to have activation potential, was employed as a replacement for the native activation domain. Finally, dimerization of the artificial transcription activator was induced by including residues 251-281 from yeast transcription activator GCN4, a previously characterized dimerization sequence. The polyamide was synthesized by solid phase synthesis, released from the beads and coupled to the peptide via native ligation. The equilibrium dissociation constant of the synthetic polyamide-peptide • DNA complex was found to be 11nM. The ability of the polyamide-peptide conjugate to activate transcription was studied in an in vitro system on a DNA template with three palindromic polyamide binding sites employing yeast nuclear extracts. The artificial activator, at a concentration of 200nM, induced a 13-fold increase in transcription over basal levels. In control experiments, when the activating peptide was deleted or when a DNA template with a single base pair mismatch in the polyamide binding site was employed, no significant activation of transcription was observed. The authors also

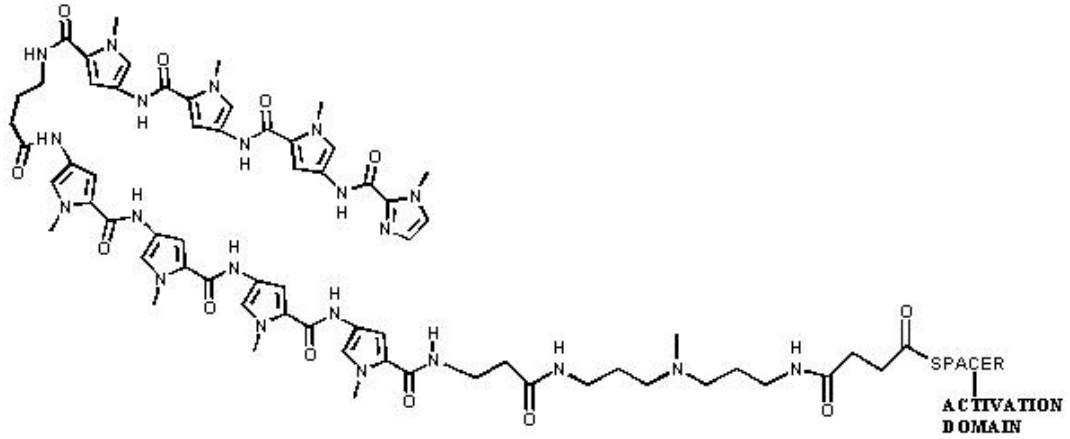
demonstrated that the dimerization domain is dispensable in the construction of a synthetic transcription activator. Thus, when the dimerization domain was replaced with a flexible linker, the molecule still retained its ability to activate transcription, with the length of the linker playing in a role in the level of activation.

In an attempt to reduce the size of the synthetic transcription activator, the same group employed a repeating eight amino acid motif from potent viral activation domain VP16 as a minimal activation domain (VP1)[45]. This synthetic construct, which had a much lower molecular weight compared to the polyamide-AH conjugate, was nearly as potent as the polyamide-AH conjugate in activating transcription in an in vitro transcription assay.

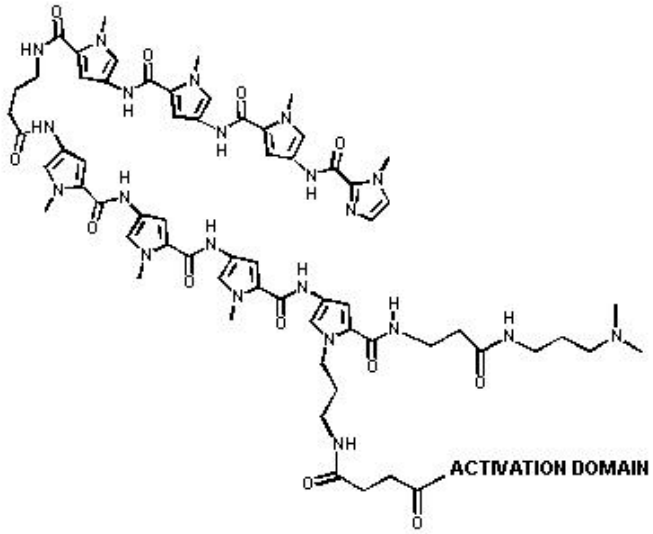
The authors, in a subsequent study, also probed the impact of the linker region on activation potential by introducing a variable number (6, 9, 12 or 15) of proline residues between the polyamide and the peptide activation domain[46]. The polyprolines form rigid helices allowing accurate prediction of the distance between the polyamide and the peptide in each synthetic construct based on the number of prolines between them. The authors found that optimal transcription activity was observed when the separation between the polyamide and the peptide was $36A^{o}$ (12 proline residues), although the differences between various synthetic constructs was quite modest (~ 2 fold).

In a recent report of a strategy towards construction of artificial transcription factors that function as protein-DNA dimerizers (PDD), Ansari, Dervan and co-workers appended a short conserved peptide fragment (YPWM) from the HOX-family of transcription factors, via a flexible linker, to a polyamide designed to bind to a DNA site which was in close proximity to the Extradenticle (Exd) recognition site[47]. Exd is a DNA binding protein that also interacts with the short peptide, albeit weakly. This Exd-peptide interaction, while weak, was significantly stabilized by the Exd-DNA interaction, resulting in efficient formation of the Exd-peptide-polyamide ternary complex. The co-operative nature of these interactions was illustrated from control experiments in which a polyamide lacking the peptide or conjugated to an inactive, mutated peptide (YPAA) failed to recruit Exd. In an attempt to further optimize the linker length, the authors synthesized a series of polyamide peptide conjugates separated by linkers of variable length, ranging from ~ 2.5 – 33A^{o}. They noticed that at low temperatures (4^{o}C), conjugates with linker lengths ranging from ~ 4 - 33A^{o} were all effective in recruiting Exd[48]. However, at physiological temperatures (37^{o}C), the conjugate bearing longest 33A^{o} linker was significantly impaired in its ability to recruit Exd, as a result of unfavorable entropic costs. The authors suggested that synthetic transcription factors such as these, which are functional at lower, but not higher temperature, may serve as temperature sensitive switches for controlling gene expression.

In an attempt to overcome the limitations associated with artificial transcription factors bearing peptide (or other bio-molecule) based activation domains, Uesugi, Dervan and co-workers developed a completely unnatural transcription activator by coupling a DNA binding-hairpin polyamide, via a flexible linker, to wrenchnolol[49]. Wrenchnolol, as described earlier, is a small molecule ligand for Sur-2, a component protein of the human mediator complex. The polyamide was designed to bind to the DNA sequence 5'-TGACCAT-3', six repeats of which were present on a reporter gene construct. An in vitro transcription assay was then used to show that the synthetic transcription activator was able to turn a reporter gene on in a dose dependent manner. Two base pair mismatch in the polyamide binding site of the reporter gene, as well as immuno-depletion of Sur-2 protein effectively abrogated the transcriptional activation response. Finally, in an attempt to determine the mechanism by which the synthetic transcription factor activates transcription, the authors immobilized a biotin labeled reporter gene on avidin-agarose resin and incubated it with nuclear extracts in the presence or absence of the synthetic transcription factor. Western blot analysis indicated that Sur-2 protein and RNA PolII were recruited to the promoter only in the presence of the synthetic transcription factor. The ability of the polyamide-wrenchnolol conjugate to activate transcription inside cells was unclear, as it had poor cell membrane permeability.



ACTIVATION DOMAIN = AH = PEFPGIELQELQELQALLQQ



ACTIVATION DOMAIN = VP2 = DFDLDMLGDFDLDMLG or
VP1= DFDLDMLG

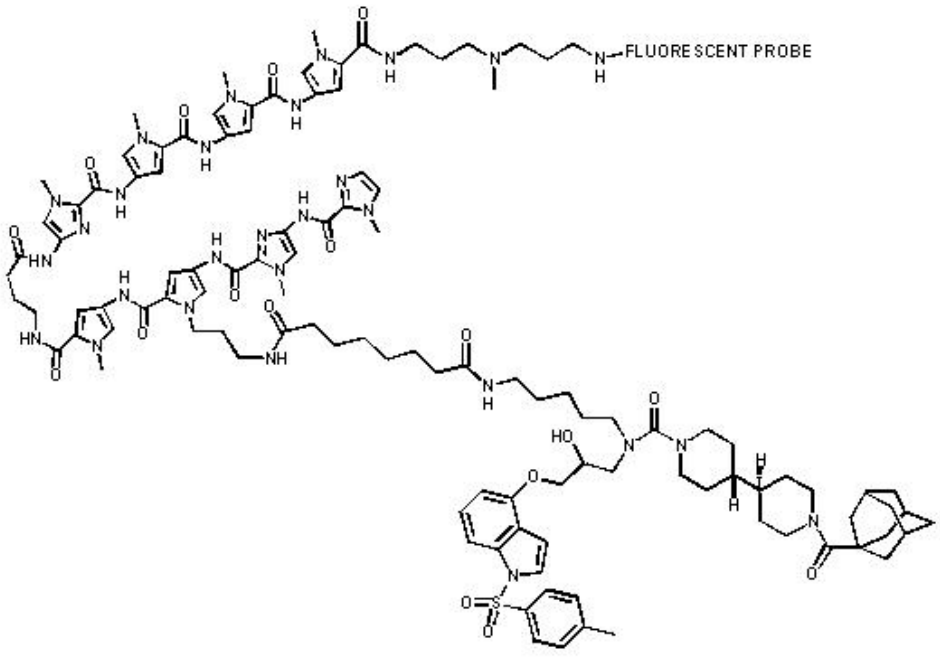


Figure 1.8: Structural and design features of various polyamide-based artificial transcriptional activators.

Synthetic transcription factors based on PNA-peptide conjugates have also been reported. In one such study, Kodadek, Corey and co-workers synthesized a bis-PNA to target a DNA sequence, 5'-AAGGAGGAGA-3' and tethered it via a flexible linker to Gal80BP, a 20mer peptide derived from panning a phage display library against yeast repressor Gal80[50, 51]. This peptide has been shown previously to be a potent activator of transcription in yeast (and subsequently in mammalian cells), when fused to Gal4 DBD. The PNA-peptide conjugate was shown to be competent in binding to the target DNA site as well as in recruiting Gal11, a yeast co-activator protein to the promoter. In a subsequent study, the authors employed an in vitro transcription assay using human nuclear extracts to assess the ability of the PNA-peptide conjugate to activate transcription. They noticed that the basal transcription level, which was high under the conditions of the assay, was effectively repressed with PNA alone, providing a cleaner background. Addition of the PNA-peptide conjugate resulted in a robust activation of transcription, whereas control peptides failed to elicit such a response when fused to the same PNA.

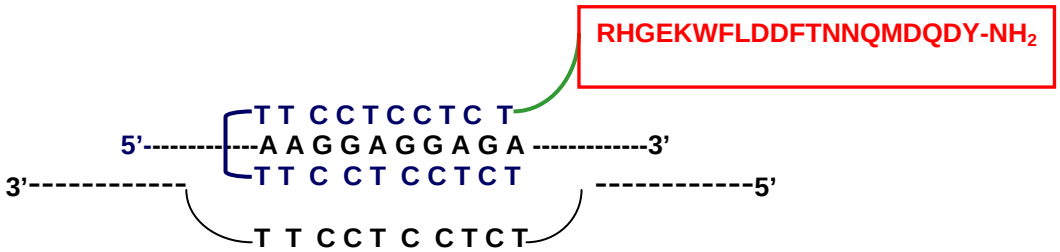
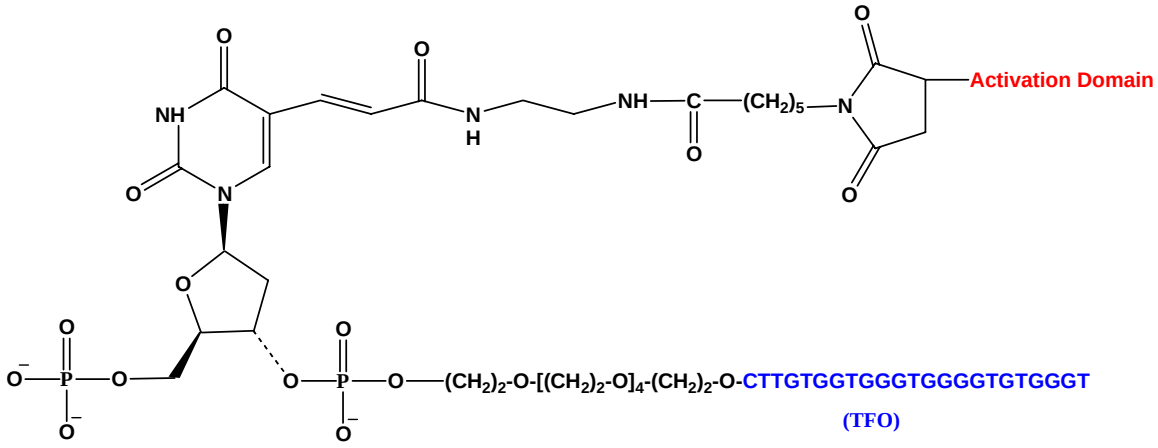


Figure 1.9: Design features of a PNA-based transcriptional activator.

In perhaps the only report to date of an artificial transcription factor that was capable of robustly activating transcription in tissue culture, Young and co-workers conjugated a 22 residue triple-helix-forming oligonucleotide (TFO) to a 29mer or 14mer peptide fragment derived from potent viral activation domain, VP16[52]. The TFO and the peptide activation domain were separated by a long, flexible poly (ethylene glycol) linker. The DNA-binding properties of the TFO employed in this study have previously been characterized. The ability of the TFO-peptide conjugate to activate transcription in vitro was tested using a transcription template containing five TFO binding sites. Gel mobility shift assays indicated the TFO retained its ability to bind to the DNA template even after conjugation to the peptide. Furthermore, the synthetic transcription factor was found to be quite potent in activating transcription of the template DNA with maximal levels achieved similar to those of Gal4-VP16 from a template containing five Gal4 binding site. Finally and quite remarkably, the TFO-peptide conjugates was shown to activate transcription of a reporter gene containing five TFO-binding sites in cultured cells when the reporter and the synthetic transcription factor were transfected into cells. While the results from the study were quite promising, future studies are necessary to determine the generality of this approach. Moreover, the poor membrane permeability and stability characteristics of TFO and peptide based molecules may hamper the wide-spread use of such molecules for controlling gene expression.



ACTIVATION DOMAIN = CGSDALDDFDLDMLGSDALDDFDLDMLGS (ATF29)
or CGSDALDDFDLDML (ATF14)

Figure 1.10: Design features of a TFO-based activation domain.

1.8. Brief perspective

In spite of the impressive advances in the design / development of synthetic DBD and AD mimetic molecules, the goal of obtaining a stable, cell permeable artificial transcription activator with activity inside non-engineered cells has not been achieved. Given that it has been possible to design polyamides to many DNA sequences and the fact that they are cell permeable, at least in certain cell lines, the lack of non-peptidic, cell permeable ligands with activation domain-like properties appears to be the bottle neck in achieving this important milestone. While some small molecule-based activation domain surrogates have been reported[39, 43] and shown to activate transcription in vitro, no example of a non-peptidic, cell permeable molecule with activation potential inside cells exists. Thus, development of such a molecule remains an important challenge in meeting the formidable goal of controlling gene expression with synthetic molecules.

1.9. Proposal summary

Among the various classes of peptidomimetic molecules reported in the literature, peptoids represent an attractive alternative to peptides[53]. Peptoids, introduced by Zuckermann approximately fifteen years ago, are oligomers of N-substituted glycines and differ from peptides in that the side chain group is attached to the amide nitrogen, rather than the alpha carbon atom. This structural modification confers several unique characteristics to peptoids. For instance,

unlike peptides, peptoids are completely resistant to proteolytic enzymes[54]. Furthermore, a recent study by Yu, Liu and Kodadek has suggested that a significant proportion of compounds in random peptoid libraries was cell permeable, when coupled to a steroid conjugate in a novel gene switch based high-throughput, cell permeability assay[55].

These favorable characteristics (protease resistance and cell membrane permeability) suggested that peptoids may serve as excellent candidates for developing artificial activation domains. Therefore, a comprehensive study was initiated to carry out the synthesis, characterization and screening of large combinatorial peptoid libraries against protein targets, especially transcription factors. The protein ligands isolated from these screens were validated and their ability to function as artificial activation domains was assessed. These studies indicate that peptoid libraries serve as an excellent source for isolating protein ligands in general, and possibly as artificial activation domains.

1.10. Bibliography

1. Unger T., Nau, M.M., Segal S., and Minna JD., *p53: a transdominant regulator of transcription whose function is ablated by mutations occurring in human cancer.* EMBO J., 1992. **11**: p. 1383-1390.
2. Pandolfi, P.P., *Transcription therapy for cancer.* Oncogene, 2001. **20**: p. 3116-3127.
3. Denison, C., and Kodadek, T., *Small molecule-based strategies for controlling gene expression.* Chem. & Biol., 1998. **5**: p. R129-R145.
4. Orphanides, G., Lagrange, T., and Reinberg, D., *The general transcription factors of RNA polymerase II.* Genes & Dev., 1996. **10**: p. 2657-2683.
5. Ptashne, M., and Gann, A., *Activators and targets.* Nature, 1990. **346**: p. 329-331.
6. Ptashne, M., and Gann, A., *Transcriptional activation by recruitment.* Nature, 1997. **386**: p. 569-577.
7. Reines, D., Conaway, J.W., and Conaway, R.C., *The RNA polymerase II general elongation factors.* Trends Biochem. Sci., 1996. **351-355**.
8. Ho, S.N., Biggar, S.R., Spencer, D.M., Schreiber, S.L., and Crabtree, G.R., *Dimeric ligands define a role for transcriptional activation domains in reinitiation.* Nature, 1996. **382**: p. 822-826.

9. Brownell, J.E., and Allis, C.D., *Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation.* Curr. Opin. Gen. Dev., 1996. **6**: p. 176-184.

10. Quinn, J., Fyrberg, A.M., Ganster, R.W., Schmidt, M.C., and Peterson, C.L., *DNA-binding properties of the SWI/SNF complex.* Nature, 1996. **379**: p. 844-847.

11. Burns, L.G., and Peterson, C.L., *The yeast SWI-SNF complex facilitates binding of a transcriptional activator to nucleosomal sites in vivo.* Mol. Cell. Biol., 1997. **17**: p. 4811-4819.

12. Gonzalez, F., Delahodde, A., Kodadek, T., and Johnston, S.A., *Recruitment of a 19S proteasome subcomplex to an activated promoter.* Science, 2002. **296**: p. 548-550.

13. Chasman, D.I., Leatherwood, J., Carey, M., Ptashne, M., and Kornberg, R.D., *Activation of yeast polymerase II transcription by herpes virus VP16 and GAL4 derivatives in vitro.* Mol. Cell. Biol., 1989. **9**: p. 4746-4749.

14. Ansari, A.Z. and Mapp, A.K., *Modular design of artificial transcription factors.* Curr. Opin. Chem. Biol., 2003. **6**: p. 765-772.

15. Dervan, P.B. and Edelson, B.S., *Recognition of the DNA minor groove by pyrrole-imidazole polyamides.* Curr. Opin. Struct. Biol., 2003. **13**: p. 284-299.

16. Nielsen, P.E., Egholm, M., Berg, R.H., and Buchardt, O., *Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide.* Science, 1991. **254**: p. 1497-1500.

17. Gottesfeld, J.M., Neely, L., Trauger, J.W., Baird, E.E., and Dervan, P.B., *Regulation of gene expression by small molecules.* Nature, 1997. **387**: p. 202-205.

18. Dickinson, L.A., Gulizia, R.J., Trauger, J.W., Baird, E.E., Mosier, D.E., Gottesfeld, J.M., and Dervan, P.B., *Inhibition of RNA polymerase II transcription in human cells by synthetic DNA-binding ligands.* Proc. Natl. Acad. Sci. USA, 1998. **95**: p. 12890-12895.

19. Kaihatsu, K., Janowsky, B.A., and Corey, D.R., *Recognition of chromosomal DNA by PNAs.* Chem. & Biol., 2004. **11**: p. 749-758.

20. Griffith, M.C., Risen, L.M., Greig, M.J., Lesnik, E.A., Sprankle, K.G., Griffey, R.H., Kiely, J.S., and Freir, S.M., *Single and bis peptide nucleic acids as triplexing agents: Binding and stoichiometry.* J. Amer. Chem. Soc., 1995. **117**: p. 831-832.

21. Vickers, T.A., Griffith, M.C., Ramasamy, K., Risen, L.M., and Freier, S.M., *Inhibition of NF-kappa B specific transcriptional activation by PNA strand invasion.* Nuc. Acids Res., 1995. **23**: p. 3003-3008.

22. Bentin, T.L., Larsen, H. J., and Nielsen, P. E., *Combined Triplex/Duplex Invasion of Double-Stranded DNA by "Tail-Clamp" Peptide Nucleic Acid.* Biochemistry, 2003. **42**: p. 13987-13995.

23. Tyler, B.M., Jansen, K., McCormick, D.J., Douglas, C.L., Boules, M., Stewart, J.A., Zhao, L., Lacy, B., Cusack, B., Fauq, A., and Richelson, E., *Peptide nucleic acids targeted to the neurotensin receptor and administered i.p. cross the blood-brain barrier and specifically reduce gene expression.* Proc. Natl. Acad. Sci. USA, 1999. **96**: p. 7053-7058.

24. Boffa, L.C., Scarfi', S., Mariani, M.R., Damonte, G., Allfrey, V.G., Benatti, U., and Morris, P.L., *Dihydrotestosterone as a selective cellular/nuclear localization vector for anti-gene peptide nucleic acid in prostatic carcinoma cells.* Cancer Res., 2000. **60**: p. 2258-2262.

25. Janowsky, B.A., Kaihatsu, K., Huffman, K.E., Schwatz, J.C., Ram, R., Hardy, D., Mendelson, C.R., and Corey, D.R., *Inhibiting transcription of chromosomal DNA with antigene peptide nucleic acids.* Nat. Chem. Biol., 2005. **1**: p. 210-215.

26. Janowsky, B.A., Huffman, K.E., Schwatz, J.C., Ram, R., Hardy, D., Shames, D.S., Minna, J.D., and Corey, D.R., *Inhibiting gene expression at transcription start sites in chromosomal DNA with antigene RNAs.* Nat. Chem. Biol., 2005. **1**: p. 216-222.

27. Courey, A.J., Holtzman, D.A., Lackson, S.P., and Tjian, R., *Synergistic activation by the glutamine-rich domains of human transcription factor Sp1.* Cell, 1989. **59**: p. 827-836.

28. Mermod, N., Oneill, E.A., Kelly, T.J., and Tjian, R., *The proline-rich transcriptional activator of Ctf/Nf-I is distinct from the replication and DNA-binding domain.* Cell, 1989. **58**: p. 741-753.

29. Ma, J., and Ptashne, M., *A new class of yeast transcriptional activators.* Cell, 1987. **51**: p. 113-119.

30. Lu, X.Y., Ansari, A.G., and Ptashne, M., *An artificial transcriptional activating region with unusual properties.* Proc Natl Acad Sci U S A, 2000. **97**: p. 1988-1992.

31. Buskirk, A.R., Kehayova, P.D., Landrigan, A., and Liu, D.R., *In vivo evolution of an RNA-based transcriptional activator.* Chem. & Biol., 2003. **10**: p. 533-540.

32. Saha, S., Ansari, A.Z., Jarrell, K.A., and Ptashne, M., *RNA sequences that work as transcriptional activating regions.* Nuc. Acids Res., 2003. **31**: p. 1565-1570.

33. Han, Y., and Kodadek, T., *Peptides selected to bind the Gal80 repressor are potent transcriptional activation domains in yeast.* J. Biol. Chem., 2000. **275**: p. 14979-14984.

34. Frangioni, J.V., LaRiccia, L.M., Cantley, L.C., and Montminy, M.R., *Minimal activators that bind the KIX domain of p300/CBP identified by phage display screening.* Nat. Biotechnol., 2000. **18**: p. 1080-1085.

35. Wu, Z., Belanger, G., Brennan, B.B., Lum, J.K., Minter, A.R., Rowe, S.P., Plachetka, A., Majumdar, C.Y., and Mapp, A.K., *Targeting the transcriptional machinery with unique artificial activators.* J. Amer. Chem. Soc., 2003. **125**: p. 12390-12391.

36. Majumdar, C.Y., Lum, J.K., Prasov, L., and Mapp, A.K., *Functional specificity of artificial transcriptional activators.* Chem. & Biol., 2005. **12**: p. 313-321.

37. Giniger, E., and Ptashne, M., *Transcription in yeast activated by a putative amphipathic α helix linked to a DNA binding unit.* Nature, 1987. **330**: p. 670-673.

38. Nyanguile, O., Uesugi, M., Austin, D.J., and Verdine, G.L., *A nonnatural transcriptional coactivator.* Proc. Natl. Acad. Sci. USA, 1997. **94**: p. 13402-13406.

39. Minter, A.R., Brennan, B.B., and Mapp, A.K., *A small molecule activation domain.* J. Amer. Chem. Soc., 2004. **126**: p. 10504-10505.

40. Volkman, H.M., Rutledge, S.E., and Schepartz, A., *Binding mode and transcriptional activation potential of high affinity ligands for the CBP KIX domain.* J. Amer. Chem. Soc., 2005. **127**: p. 4649-4658.

41. Rutledge, S.E., Volkman, H. M., and Schepartz, A., *Molecular Recognition of Protein Surfaces: High Affinity Ligands for the CBP KIX Domain.* J. Amer. Chem. Soc., 2003. **125**: p. 14336-14347.

42. Asada, S., Choi, Y., and Uesugi, M., *A gene-expression inhibitor that targets an alpha-helix-mediated protein interaction.* J. Amer. Chem. Soc., 2003. **125**: p. 4992-4993.

43. Shimogawa, H., Kwon, Y., Mao, Q., Kawazoe, Y., Choi, Y., Asada, S., Kigoshi, H., and Uesugi, M., *A wrench-shaped synthetic molecule that modulates a transcription factor-coactivator interaction.* J. Amer. Chem. Soc., 2004. **126**: p. 3461-3471.

44. Mapp, A.K., Ansari, A.Z., Ptashne, M., and Dervan, P.B., *Activation of gene expression by small molecule transcription factors.* Proc. Natl. Acad. Sci. USA, 2000. **97**: p. 3930-3935.

45. Ansari, A.Z., Mapp, A.K., Nguyen, D.H., Dervan, P.B., and Ptashne, M., *Towards a minimal motif for artificial transcriptional activators.* Chem. & Biol., 2001. **8**: p. 583-592.

46. Arora, P.S., Ansari, A.Z., Best, T.P., Ptashne, M., and Dervan, P.B., *Design of artificial transcriptional activators with rigid poly-L-proline linkers.* J. Amer. Chem. Soc., 2002. **124**: p. 13067-13071.

47. Arndt, H.-D., Hauschild, K.E., Sullivan, D.P., Lake, K., Dervan, P.B., and Ansari, A.Z., *Toward artificial developmental regulators.* J. Amer. Chem. Soc., 2003. **125**: p. 13322-13323.

48. Hauschild, K.E., Metzler, R.E., Arndt, H-D., Moretti, R., Raffaelle, M., Dervan, P.B., and Ansari, A.Z., *Temperature-sensitive protein-DNA dimerizers.* Proc Natl Acad Sci USA, 2005. **102**(14): p. 5008-5013.

49. Kwon, Y., Arndt, H.-D., Mao, Q., Kawazoe, Y., Dervan, P.B., Uesugi, M., *Small molecule transcription factor mimic.* J. Amer. Chem. Soc., 2004. **126**: p. 15940-15941.

50. Liu, B., Han, Y., Ferdous, A., Corey, D.R., and Kodadek, T., *Transcription activation by a PNA-peptide chimera in a mammalian cell extract.* Chem. & Biol., 2003. **10**: p. 909-916.

51. Liu, B., Han, Y., Corey, D.R., and Kodadek, T., *Towards synthetic transcription activators: Recruitment of transcription factors to DNA by a PNA-peptide chimera.* J. Amer. Chem. Soc., 2002. **124**: p. 1838-1839.

52. Stanojevic, D. and Young, R.A., *A highly potent artificial transcription factor.* Biochemistry, 2002. **41**: p. 7209-7216.

53. Simon, R.J., Kania, R.S., Zuckermann, R.N., Huebner, V.D., Jewell, D.A., Banville, S., Ng, S., Wang, L., Rosenberg, S., Marlowe, C.K., Spellmeyer, D.C., Tan, R., Frankel, A.D., Santi, D.V., Cohen, F.E., and Bartlett, P.E.,

Peptoids: A modular approach to drug discovery. Proc. Natl. Acad. Sci. USA, 1992. **89**: p. 9367-9371.

54. Miller, S.M., Simon, R.J., Ng, S., Zuckermann, R.N., Kerr, J.M., and Moos, W.H., *Comparison of the proteolytic susceptibilities of homologous L-amino acid, D-amino acid, and N-substituted glycine peptide and peptoid oligomers.* Drug Dev. Res., 1995. **35**(1): p. 20-32.

55. Yu, P., Liu, B., and Kodadek, T., *A high-throughput assay for assessing the cell permeability of combinatorial libraries.* Nat. Biotechnol., 2005. **23**: p. 746-751.

CHAPTER 2: SYNTHESIS AND SCREENING OF LARGE PEPTOID LIBRARIES FOR ISOLATION OF PROTEIN LIGANDS

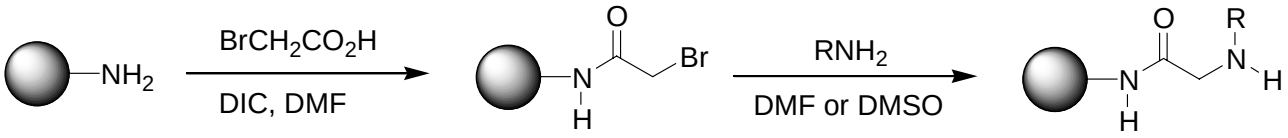
2.1 Introduction

Since introduced by Lam and co-workers in 1991, split and pool / one-bead-one-compound combinatorial libraries[1] have served as extremely powerful tools for rapid isolation of ligands against proteins and other target molecules. Peptoids, which can be conveniently synthesized by "sub-monomer" approach[2] (Scheme 2.1), are high amenable to split and pool synthesis. The sub-monomer method for peptoid synthesis consists of two steps:

(a) an acylation step, which is accomplished by the addition of bromoacetic acid and N,N'-Diisopropylcarbodiimide (DIC) and

(b) a nucleophilic displacement step, in which the bromide is displaced with a primary amine.

Thus, the primary amines serve as diversity generating elements in the synthesis of peptoid libraries. The commercial availability of hundreds of primary amines allows facile synthesis of large and structurally diverse peptoid libraries in a cost effective manner, often without the need with protection groups (depending on the nucleophilicity of other functional groups, if any, in the amine molecule).



DIC = N,N'-Diisopropylcarbodiimide
R = any alkyl group

Scheme 2.1: Sub-monomer approach for the synthesis of peptoids[2].

These favorable characteristics of peptoids, along with resistance to proteases[3] and potential for cell-permeability[4], make them excellent tools for manipulating various biological functions, including transcription. In spite of such favorable properties, especially for library construction, there was only one report of synthesis and screening of peptoid libraries[5] at the time this study was initiated. Zuckermann and co-workers reported two α1-adrenergic receptor ligands with K_i's in low nanomolar range isolated from screening a biased two- and three- residue peptoid library of modest size (~ 5000 compounds). Subsequently, a very small (12-compound) library of peptide-peptoid hybrids was screened and an excellent ligand for the SH3 domain was isolated[6]. More recently, peptoid ligands with analgesic[7], anti-microbial[8], and multi-drug resistance reversal[9] properties have been reported. While these studies were encouraging, screening of much larger libraries would be necessary in order to isolate good quality ligands from completely naïve libraries. Therefore, a pilot study was launched in which large (~ 100,000 - 500,000 compound-) peptoid libraries were synthesized, characterized and screened against a target protein.

To this end, a simple and general scheme for isolation of protein ligands was envisioned (Scheme 2.2.). The scheme consists of the following steps:

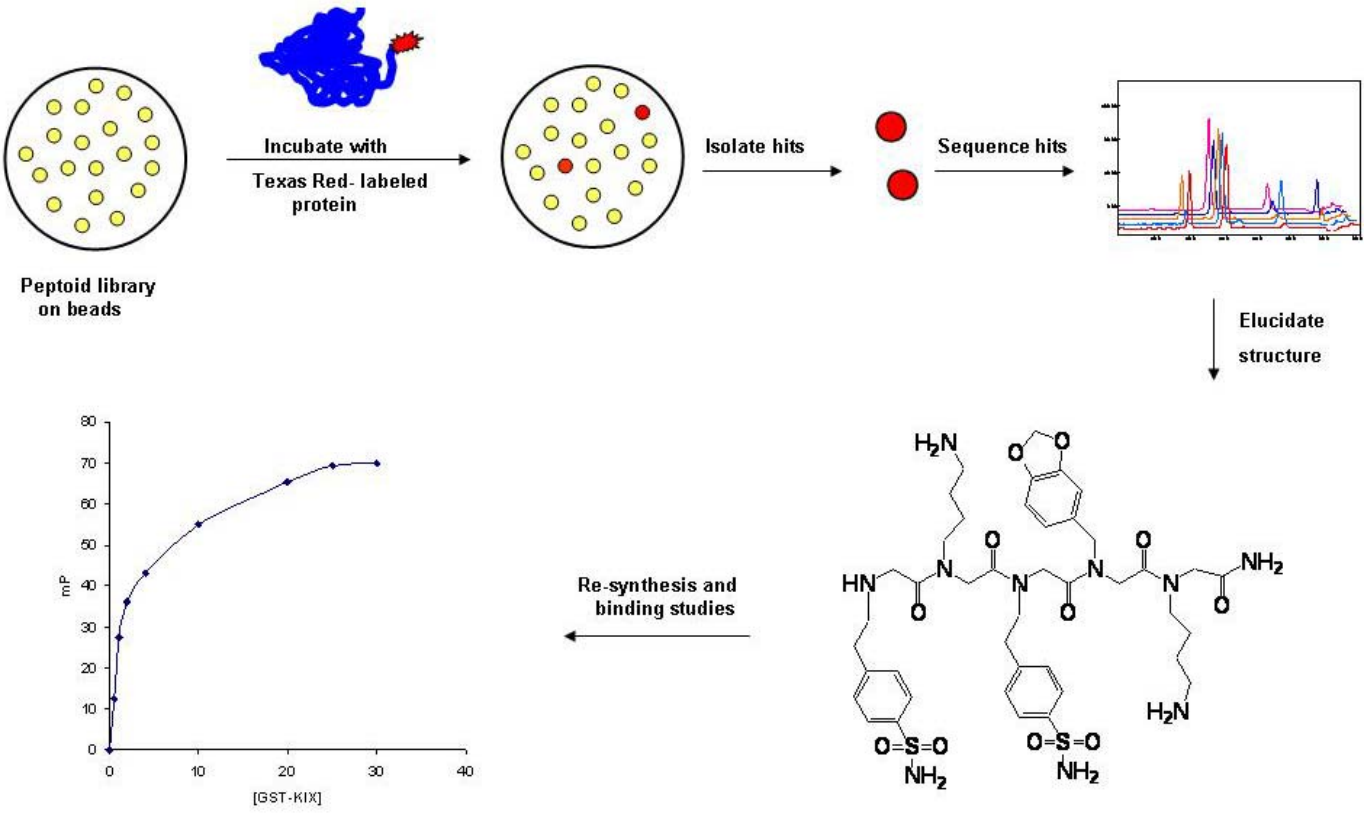
(a) synthesis of large peptoid libraries on solid support (resin) employing split and pool synthesis,

(b) incubation of the bead-bound library with fluorescently tagged protein target,

(c) isolation of a small fraction of library beads that are brightly fluorescent well above the background by manual or automated means,

(d) accurate determination of chemical structure of bead bound peptoids selected in the screening using sensitive analytical techniques (such as Edman sequencing) and

(e) validation and quantitation of ligand . protein interactions.



Scheme 2.2: A general scheme for isolation of protein ligands from peptoid libraries.

As reported here, the synthesis and screening of large peptoid libraries is facile and efficient. The data from this study suggest that peptoid libraries can serve as general and excellent sources for specific and inexpensive protein ligands. Therefore, screening of peptoid libraries against appropriate transcription factors may lead to peptoid-based artificial activation domains.

2.2 Materials and methods

2.2.1 Reagents and instrumentation

All the reagents and solvents were purchased from commercial suppliers and used without further purification. TentaGel macrobeads (140-170 micron diameter; substitution: 0.51mmol/g) were obtained from Rapp Polymere. Analytical HPLC was performed on a Biocad Sprint system with a C18 reverse-phase HPLC column (Vydac, 5μM, 4.6 mm i.d. x 250 mm) using the following solvents for elution; solvent A: H_2O / 0.1% TFA; solvent B: CH_3CN / 0.1% TFA. MALDI-TOF MS was performed on a Voyager-DE PRO biospectrometry workstation (Applied Biosystems) using α-hydroxy cinnamic acid as the matrix. A New Brunswick Scientific Innova 4400 incubator shaker was used to perform the peptoid syntheses at 37^{o}C. Microwave-assisted peptoid syntheses[10] were performed on a 1000W Whirlpool microwave oven (model MT1130SG) set to deliver 10% power. Edman sequencing of peptoids[11] was performed on an ABI

476A Protein Sequencer (Applied Biosystems). The fluorescence spectra of the beads were recorded with a hyperspectral imaging microscope constructed in the laboratory of Prof. Harold Garner (UT-Southwestern). The on-bead fluorescence assays were visualized with a Nikon Eclipse TE300 fluorescence microscope equipped with a Chroma 61002 triple band filter set and a CCD camera. MetaMorph software was used to acquire and process the photomicrographs. Isothermal titration calorimetry (ITC) experiments were performed on a MicroCal VP-ITC instrument.

2.2.2 Synthesis of peptoid libraries at 37°C

The synthesis of the 8-mer library was performed in standard 25ml glass peptide synthesis reaction vessels (Chemglass) in an incubator shaker at 37°C. 1.5g of TentaGel macrobeads (140-ı□̃ □m; substitution:0.51mmol/g) were distributed equally into 5 peptide synthesis reaction vessels, 5ml of DMF was added and the beads were allowed to swell at room temperature for 60 minutes. The DMF was drained and to each vessel was added 1.5 ml of 2M bromoacetic acid and 1.5 ml of 3.2M diisopropylcarbodiimide (DIC). The reaction vessels were placed on an incubator shaker set at 37°C and 225 r.p.m. for 40 minutes. The vessels were drained and the beads were thoroughly washed with DMF (8 x 3 ml). The beads in each of the vessels were treated with one of the five primary amines (see Table 1) at 2M concentration and allowed to react in the shaker at 37°C for

60 minutes. All the amines were dissolved in DMF, except 4-(2-aminoethyl)benzene sulfonamide which was dissolved in DMSO. The vessels were drained and washed thoroughly with DMF (8 x 3 ml). The beads in each of the reaction vessels were pooled into a large 250ml peptide synthesis vessel, drained, suspended in 50ml of dichloromethane/DMF (2:1) and randomized by bubbling argon for 15 minutes. The beads were distributed equally into five 25 ml peptide synthesis vessels and the procedure was repeated. The protocol was slightly modified for the final 4 residues of the library, where the displacement of the bromide by the primary amine was carried out for 90 minutes, instead of 60 minutes. At the end of the library synthesis, the beads were washed thoroughly with DMF (8 x 3 ml) and dichloromethane (3 x 3 ml), drained and treated with 6 mL of 95% TFA, 2.5% water and 2.5% anisole for 2 hours. The cleavage cocktail was drained and the beads were washed thoroughly with dichloromethane (8 x 3 ml). The beads were neutralized by treating with 10% diisopropylethylamine in DMF for 5 minutes, washed with dichloromethane (5 x 3 ml), and dried until further use.

2.2.3 Microwave-assisted peptoid library synthesis[10]

The synthesis of the 5-mer and the 6-mer libraries were performed employing a microwave-assisted protocol (see next section for details) on 1g and 2g of beads, respectively. In this protocol, both the acylation and bromide

displacement by the primary amine were performed twice for 15 seconds in a 1000W microwave oven set to deliver 10% power. The beads were shaken manually for 30 seconds between microwave pulses to ensure proper mixing. All the other steps were identical to the 37^{o}C procedure.

2.2.4 Microwave-assisted synthesis of peptoids for validation

Synthesis of peptoids for characterization, (by HPLC and MALDI-TOF analysis), was carried out on 50 mg of Fmoc-Rink amide MHBA resin (substitution: 0.73 mmol/g; Nova Biochem). The resin was swollen in DMF (2.5 mL) for 30 minutes, drained and 20% piperidine/DMF (2 x 2.5 mL) was added and shaken for 10 minutes. The resin was then drained and washed with DMF (8 x 2.5 mL). 2M Bromoacetic acid in anhydrous DMF (1 mL) and 2 M DIC in anhydrous DMF (1 mL) were added and the beads were gently shaken for about 30 seconds. The vessel was placed in a beaker inside the center of the microwave oven for 15 seconds and the power was set at 10%. The vessel were taken out and manually agitated for 30 seconds before placing again inside the microwave for another 15 seconds with the power set at 10%. The beads were then washed with DMF (8 x 4 mL) and a solution of 1M amine in anhydrous DMF or DMSO (2 mL) was added and the vessel was shaken for about 30 seconds. Then it was placed in a beaker inside the microwave oven for 15 seconds with the power set at 10%. The vessel was manually shaken for about 30 seconds and placed inside the

microwave for another 15 seconds with the power set at 10%. The beads were drained and washed with DMF (8 x 4 mL) and subjected to the addition of the next residue. At the end of the synthesis the beads were washed with methylene chloride (8 x 2.5 mL), dried under nitrogen and cleaved from the resin (with concomitant removal of side chain protection groups) by treating with 5 mL of 95% TFA, 2.5% water and 2.5% anisole for 2 hours. The suspension was filtered and the filtrate concentrated by blowing nitrogen over the solution. The concentrated filtrate was dissolved in 2 mL of 1:1 acetonitrile/water and lyophilized. The resultant solid was subjected to HPLC and MALDI-TOF analysis.

2.2.5 Protection of additional functional groups in primary amines

The secondary amino group in tryptamine[12] and one of the primary amino groups in 1,4-diaminobutane[13] were protected with a t-Boc groups. O-t-Butyl-2-amino ethanol[12], a hydroxyl-protected version of ethanolamine, was purchased from a commercial source (CSPS Pharmaceuticals).

2.2.5.1 Protection of tryptamine[12]

16.06 grams of tryptamine was dissolved in 100 mL of methylene chloride and 40 mL of pyridine was added. The solution was stirred on an ice bath for 10 minutes and 15 mL of trifluoroacetic anhydride was added drop by drop over 10

minutes. The reaction was then allowed to warm back to room temperature and stirred for another 2 hours. The reaction mixture was concentrated in vacuo and the resultant dark brown oil was dissolved in ethyl acetate and washed with 1 M potassium bisulfate, saturated sodium bicarbonate and saturated brine. The organic layer was dried over sodium sulfate and concentrated in vacuo. The crude residue was subjected to flash chromatography (methylene chloride / hexanes) to afford 22.6 g (88% yield) of N-trifluoroacetyl tryptamine as a light cream colored solid.

22.6 g of N-trifluoroacetyl tryptamine was dissolved in 60 mL of THF and 22.5 g of di-tert-butyl dicarbonate and 550 mg of N,N-dimethylaminopyridine were added and stirred at 37^{o}C for 2 h. The reaction mixture was then concentrated in vacuo, residue dissolved in ethyl acetate and washed with 1M potassium bisulfate, saturated sodium bicarbonate and saturated brine solutions. The organic layer was dried over sodium sulfate, concentrated in vacuo and purified by flash chromatography (methylene chloride / hexanes) to yield 25.4 g (80% yield) of N^{in}-t-Boc N-trifluoroacetyl tryptamine as a light cream colored solid.

25.4 g of N^{in}-t-Boc N-trifluoroacetyl tryptamine was dissolved in 225 mL of methanol and aqueous potassium carbonate (15 g in 90 mL water) and stirred at room temperature for 24 hours. The solvent was then removed and the aqueous

layer was extracted with ethyl acetate. The organic layer was washed with saturated brine, dried over sodium sulfate and concentrated in vacuo to afford ~ 18g of N^{in}-t-Boc tryptamine as a light yellow colored oil.

2.2.5.2 Protection of 1, 4-diaminobutane[13]

100 mmoles of $(Boc)_2O$ in 100 mL of chloroform was added drop wise to a stirring solution of 500 mmoles of 1, 4-diaminobutane in ~ 2.5 L of chloroform at 0^oC. The reaction mixture was then allowed to warm back to room temperature, stirred for an additional two hours and filtered. The filtrate was diluted with ethyl acetate, washed with brine solution and the organic layer was concentrated in vacuo, after drying over anhydrous sodium sulfate, to yield the desired product.

2.2.6 Sequencing peptoids by Edman degradation[11]

The sequencing of peptoids was performed on an ABI 476A protein sequencer, using the FSTNML program and a standard gradient (Gradient 1). The FSTNML program was slightly modified by adding a 60 second "wait" step at the end of the cycle to enable the gradient to run slightly longer than normal.

2.2.7 Protein purification

Glutathione-S-Transferase (GST) was expressed in *E. coli* BL21-RIL from the commercially available plasmid pGEX-2T (Amersham Biosciences). The

cells were grown until an OD_{600} of 0.8 was reached, at which time 1mM IPTG was added to the medium to induce protein expression. After further growth at 37°C for 3hrs, the cells were harvested, sonicated and centrifuged at 22,000 rpm. The cleared lysate was then incubated with glutathione-agarose beads equilibrated with PBS at 4°C for 1hr. The beads were washed with 10-12 volumes of PBS, packed into a column and further rinsed with PBS. GST bound to the beads was eluted with 10mM reduced glutathione/PBS and fractions were collected and analyzed on a 12% denaturing polyacrylamide gel. The fractions containing highly purified GST were pooled and dialyzed against PBS + 10% glycerol. The protein concentration was estimated using Cooomassie Plus Protein Assay Reagent Kit using BSA as a standard.

2.2.8 Protein labeling with Texas Red

The protein solution (preferably 2mg/ml) was adjusted to pH 8.3 with 0.2M $NaHCO_3$ buffer. To this 5µl of 50mg/ml Texas Red-NHS ester solution in DMF was added with mild vortexing to mix the sample. This solution was incubated with tumbling at room temperature for one hour, after which the reaction was quenched with 1.5M hydroxylamine. Dye-conjugated protein was separated from excess dye using a desalting column. Measurement of the absorbance of the sample at 280 nm and 595 nm indicated that, on average, these

conditions resulted in each protein molecule acquiring one molecule of Texas Red.

2.2.9 Preparation of *E. coli* lysate for screening experiments

The *E. coli* (BL21-RIL strain) cells were grown overnight at 37^{o}C in Luria broth. The cells were harvested by low speed centrifugation, washed and resuspended in sonication buffer (50mM NaH_2PO_4 pH 8.0, 300mM NaCl, 0.1% Tween20 + protease inhibitor). The cells were then sonicated and centrifuged at 22,000 rpm to remove cell debris and provide the cleared cell lysate. The concentration of the lysate was estimated using the Bradford assay with BSA as a standard.

2.2.10 Library screening and identification of hits

Approximately 100, 000 beads from the 8mer library were swollen in PBS for 1 hour, after which they were blocked with BSA at room temperature for one hour to block any non-specific binding sites. The beads were then incubated with 2 μM solution of Texas Red-conjugated GST in PBS in the presence of large excess of BSA as competitor protein. The beads were washed with TBST (6 X 1 mL) and visualized under a fluorescence microscope fitted with a Texas Red filter. The brightest beads were isolated manually with a pipette tip.

In another experiment, approximately 50,000 beads from the 5mer library were screened against Tex Red-labeled GST under more stringent conditions. Thus, the beads were blocked with 5% milk/ TBST and then incubated with 1 μM Texas Red-labeled GST in the presence of 1000-fold excess of *E. coli* lysate. The beads were then washed with TBST (6 X 1 mL) and visualized under the microscope.

In each case, after picking the putative "hits", each bead was heated in a 1% SDS solution for 20 minutes, followed by three washes with 1X PBS. They were then sequenced by Edman degradation.

2.2.11 Isothermal titration calorimetry (ITC)

ITC experiments were conducted on a MicroCal VP-ITC instrument. For the titration, 30 μM GST in PBS + 10% DMSO was taken in the sample cell. To this, 15μl aliquots of the peptoid solution in the same buffer were added from a computer-controlled 250μl rotating syringe. The syringe was set at 400 rpm with intervals of 3 minutes between injections to attain baseline stabilization. The heat absorbed or released accompanying the titration was recorded as differential power (DP) by the instrument software. Experiments were carried out with C values between 1 and 400. The total heat recorded was then fitted via a non-linear

least-squares minimization method. Titration of the ligand solution with the buffer alone gave the heats of dilution.

2.2.12 Protein capture assays using TentaGel-displayed peptoids

5mg of TentaGel beads (displaying the respective hit sequences, Nlys-Nbsa-Nlys-Nser-Nbsa-Npip-Nbsa-Npip-$CONH_2$ and Nbsa-Nlys-Nbsa-Npip-Nlys-$CONH_2$ or a random sequence Npip-Nser-Nbsa-Nall-Nlys-Npip-$CONH_2$) were equilibrated in phosphate-buffered saline (PBS) for 60 minutes. The buffer was removed and the beads were blocked with 2% BSA for 60 minutes to saturate any non-specific binding sites. The beads were then washed with PBS (x3 times), and incubated with 500nM (unless indicated otherwise) of a Texas Red-labeled GST or MBP in 2% BSA (in 1X TBST buffer), in a 300□L volume for 60 minutes. The beads were washed with TBST six times to remove any unbound protein and photographed under a fluorescence microscope.

Experiments that employed native (unlabeled) proteins were performed as follows. 10mg beads displaying the peptoid were exposed to 1μM protein in the presence of 1000-fold excess *E. coli* lysate, 0.2% Tween20 and 0.2M NaCl in a total volume of 2ml at room temperature for 2hrs. The beads were washed thrice with TBST (20mM Tris buffered saline + 0.1%Tween20). 10μl of 2X SDS-PAGE loading dye was then added directly to these beads and boiled for 10

minutes. The entire supernatant was loaded on to a 12% denaturing polyacrylamide gel and analyzed by Western blot using an anti-GST antibody.

2.2.13 Dilution experiment

15 mg of TentaGel beads displaying Nbsa-Nlys-Nbsa-Npip-Nlys-$CONH_2$ were equilibrated in PBS for 60 min. The buffer was removed and the beads were incubated with *E. coli* lysate for 60 minutes to block any non-specific binding sites. The beads were washed with PBS three times and split into three Eppendorf tubes. The beads were incubated with 1μM, 500nM or 100nM respectively, of Texas Red-labeled GST in the presence of a 100-fold excess of *E. coli* lysate in a 300 μL volume for 60 minutes. The beads were washed with TBST six times to remove any unbound protein and visualized under a fluorescence microscope. The experiment was also done at 10 nM protein, but little or no fluorescence above background was observed (not shown).

2.3 Results

2.3.1 TentaGel macrobeads serve as excellent solid support for peptoid library synthesis

One of the critical issues in the generation of a split and pool combinatorial library is the selection of a solid support / resin that is optimal for both synthesis and screening experiments. Thus, the solid support should have:

(1) excellent swelling properties in both organic solvents and water so that it supports efficient synthesis and subsequent display of the bound peptoids in aqueous buffers during screening experiments,

(2) sufficiently high loading capacity to enable unambiguous determination of the structure of peptoids from individual beads by direct Edman or mass spectrometry- based sequencing, eliminating the need for complicated encoding strategies,

(3) low fluorescence background to facilitate fluorescence based on-bead screening assays, and

(4) physical robustness to retain bead integrity over several rounds of split and pool synthesis and screening.

After considerable experimentation with various commercially available resins, TentaGel macrobeads (140-170 μm in diameter from Rapp Polymere) were selected as the resin of choice, since they satisfied most of the criteria mentioned above. Thus, the hydrophobic core of the resin made up of polystyrene

makes the beads physically robust and affords excellent swelling properties in organic solvents. The derivatization of the resin with poly (ethylene) glycol chains vastly improves the swelling properties in aqueous solutions and provides a "non-sticky" surface that is ideal of reducing non-specific binding during screening experiments. Each gram of the resin contains approximately 500,000 high capacity beads enabling facile synthesis of large libraries starting only from a few grams of the resin. The capacity of each bead of ~ 1 nmole is well above the detection limit of Edman and mass spectrometry based sequencing techniques. Finally, the background fluorescence of the resin, while less than optimal, is lower than most commercially available resins tested.

2.3.2 Expanding the repertoire of peptoid building blocks

As mentioned earlier, the diversity generating elements in peptoids are primary amines. Therefore, assembling a large repertoire of primary amines suitable for peptoid synthesis is essential for the generation of large and structurally diverse peptoid libraries. Prior studies have established several primary amines that work well in peptoid synthesis[5, 15-18] (Figure 2.1). To expand this collection, several new amines (2, 7, 9, 10 and 11 in figure 2.1) were tested for their suitability for incorporation into peptoids.

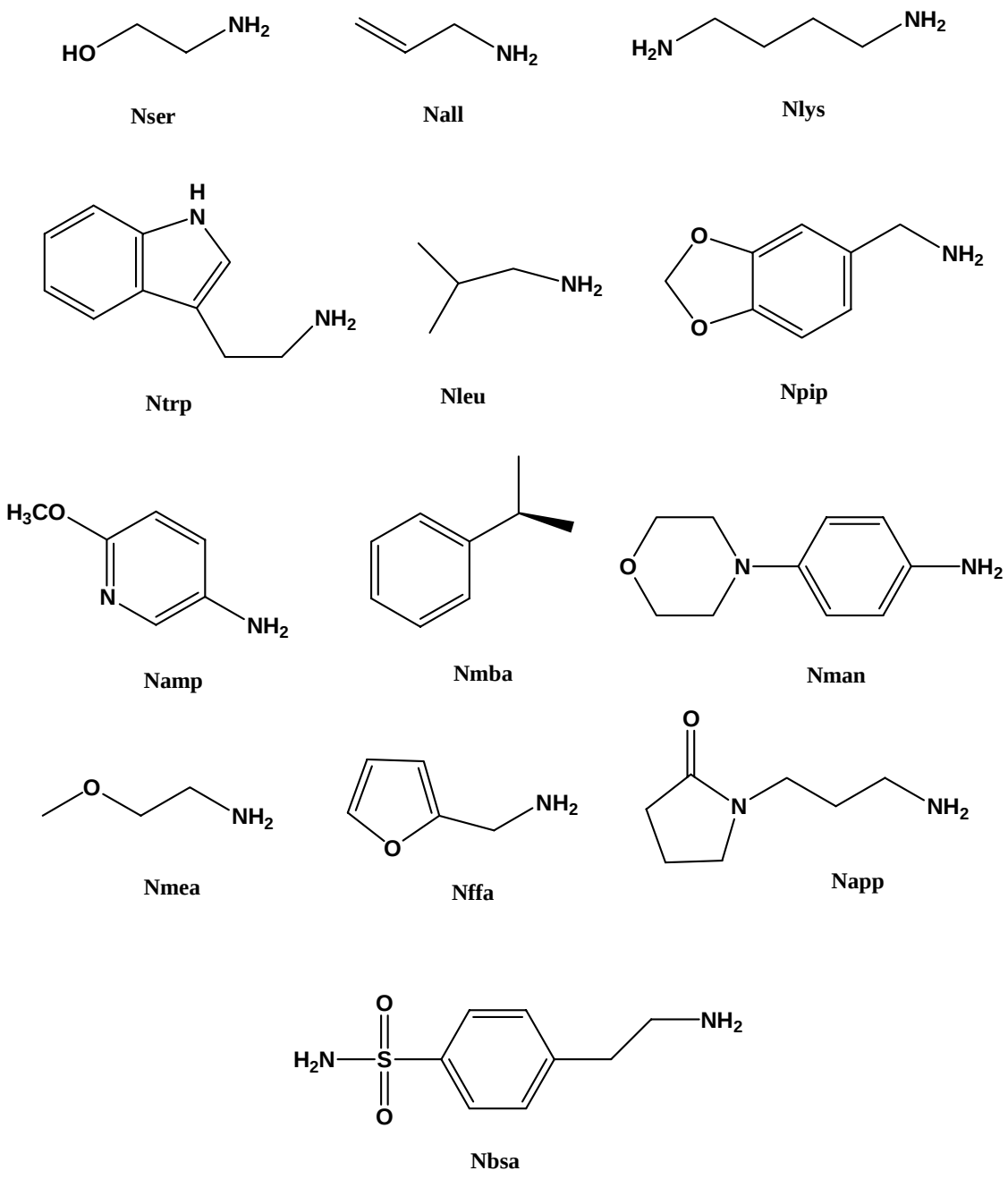
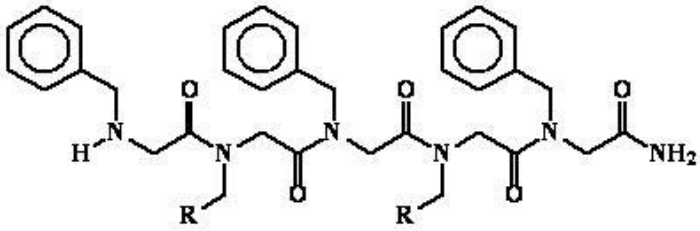


Figure 2.1: Building blocks.

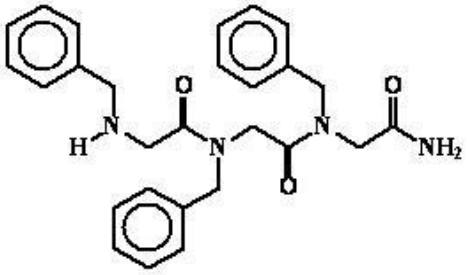
List of amines used in the synthesis of various peptoids and peptoid libraries. The corresponding three letter nomenclature of the peptoid units is also shown.

The testing of new monomers was accomplished by synthesizing a test pentameric sequence in which the test amine was used in steps two and four, while the residues 1, 3 and 5 were derived from well behaved benzylamine[16] (Figure 2.2). If the reactivity of the test amine is diminished by steric hindrance or electronic effects of other proximal functional groups, the displacement of the bromide in steps two and four will be compromised and a trimeric peptoid derived from benzylamines would result as the major product. One the other hand, a well behaved amine would yield the full-length pentameric peptoid in high yields with little, if any, of the trimer. All the new amines tested in this assay gave the full-length peptoid in greater than 85% yield validating them as "good" sub-monomers for peptoid synthesis and library construction.

(a)



(b)



(c)

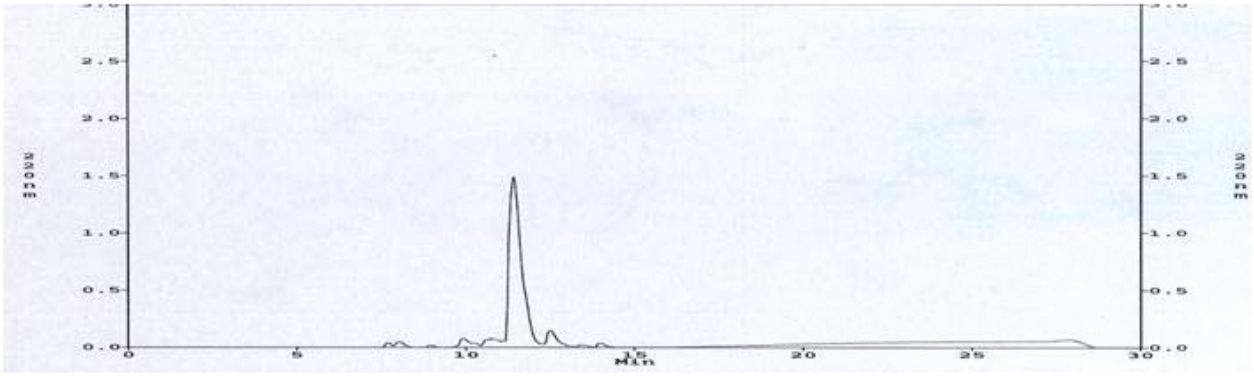


Figure 2.2: Monomer validation.

A schematic depiction of a positive result (a) and negative result (b) during the testing of a new monomer.

(c) HPLC trace of a crude test pentamer synthesized to validate a new monomer (Nman). The major peak in the HPLC has been found to correspond to the mass of the full-length pentamer, suggesting that Nman is a good monomer for peptoid synthesis.

2.3.3 Microwave enhanced peptoid synthesis

Although the sub-monomer approach[2] introduced by Zuckermann and others works well and allows the synthesis of high quality peptoids, the coupling time per residue at room temperature under standard literature conditions is about 2.5 to 3 hours. Increasing the temperature to 35 ^{o} C reduces the coupling time to approximately 80 minutes[15], but still means that the synthesis of a 10mer peptoid takes about a day. In a high throughput approach such as that required for the synthesis of large peptoid libraries and eventual re-synthesis of several hits for validation, it is desirable to accelerate the reaction rates to facilitate rapid isolation of bonafide ligands against protein targets.

Microwave irradiation has been reported to accelerate the rates of several chemical reactions[19], including the solid phase synthesis of peptides[20]. Therefore, a study was initiated to explore the possibility of using microwave irradiation to accelerate the solid phase synthesis of peptoids[10]. Thus, a series of 9-residue peptoids comprising of homo-oligomers and a hetero-oligomer were constructed using several amines in domestic microwave oven (Table 2.1.). All the peptoids were synthesized on Rink MBHA amide resin, which facilitates the TFA-mediated cleavage of the peptoids from the resin at the end of the synthesis and subsequent analysis of their quality using HPLC and MALDI-TOF mass spectrometry. For comparison purposes, all the peptoids were synthesized under standard literature conditions (of room temperature and 37 ^{o} C) as well as the

following microwave enhanced synthesis conditions. In each case, stock solutions of bromoacetic acid and DIC prepared in DMF at 2 M concentration were employed, while the concentration of amines in DMF or DMSO were maintained at 1 M.

The microwave synthesis was carried out in a 1000-W commercial microwave oven with power adjusted to deliver at 10%. Systematic studies have revealed that reaction times as short as 30-40 seconds were sufficient to support efficient acylation and amine displacement reactions resulting in the desired full-length peptoids in high yields and purities. Since there was not provision to stir the solutions inside the microwave oven, the reaction vessels were irradiated for 15 seconds, gently stirred by manual agitation, and irradiated again for another 15 seconds. The temperature of the solution at the end of 15-s irradiation step, as measured by a thermometer, did not exceed beyond 35^{o}C. However, sophisticated in situ measurements of reactions temperatures during irradiation may be required to answer if the enhanced reaction rates are as a result of microwave effects or simply efficient “flash heating” of reaction solutions. Nevertheless, the results clearly indicate that microwave irradiation supports rapid and efficient synthesis of peptoids. The yields and purities of peptoids synthesized by this protocol were comparable or superior to those obtained from standard literature conditions (Table 2.1.). The major products obtained from all the methods had identical HPLC retention times and masses. The main impurities in each case consisted of

shorter residue sequences (8-mers, 7-mers etc.), and in some cases, intermediates where a hydroxyl group displaced the bromide resulting in termination of chain elongation. Presumably, it is a consequence of presence of traces of water in the reagents and relatively inefficient acylation of the α-hydroxyacids.

Thus, a fast and efficient protocol for the synthesis of peptoids using a readily available domestic microwave oven has been developed. The time required for steps such as bead swelling and washing obviously remains unaltered. Nevertheless, the dramatic acceleration of the coupling steps translates into significant saving in the overall time required for the synthesis of peptoids. For instance, a 10-residue peptoid can be synthesized on pre-swollen beads in less than ~ 3.5 hours using the microwave enhanced procedure described here, compared to 22-35 hours using existing literature protocols (depending on the temperature). Given the fact that overall purities and yields of these peptoids were comparable or superior to those synthesized from standard literature protocols, the microwave enhanced protocol for the synthesis of peptoids should be of great utility in the high-throughput synthesis of peptoids and peptoid libraries.

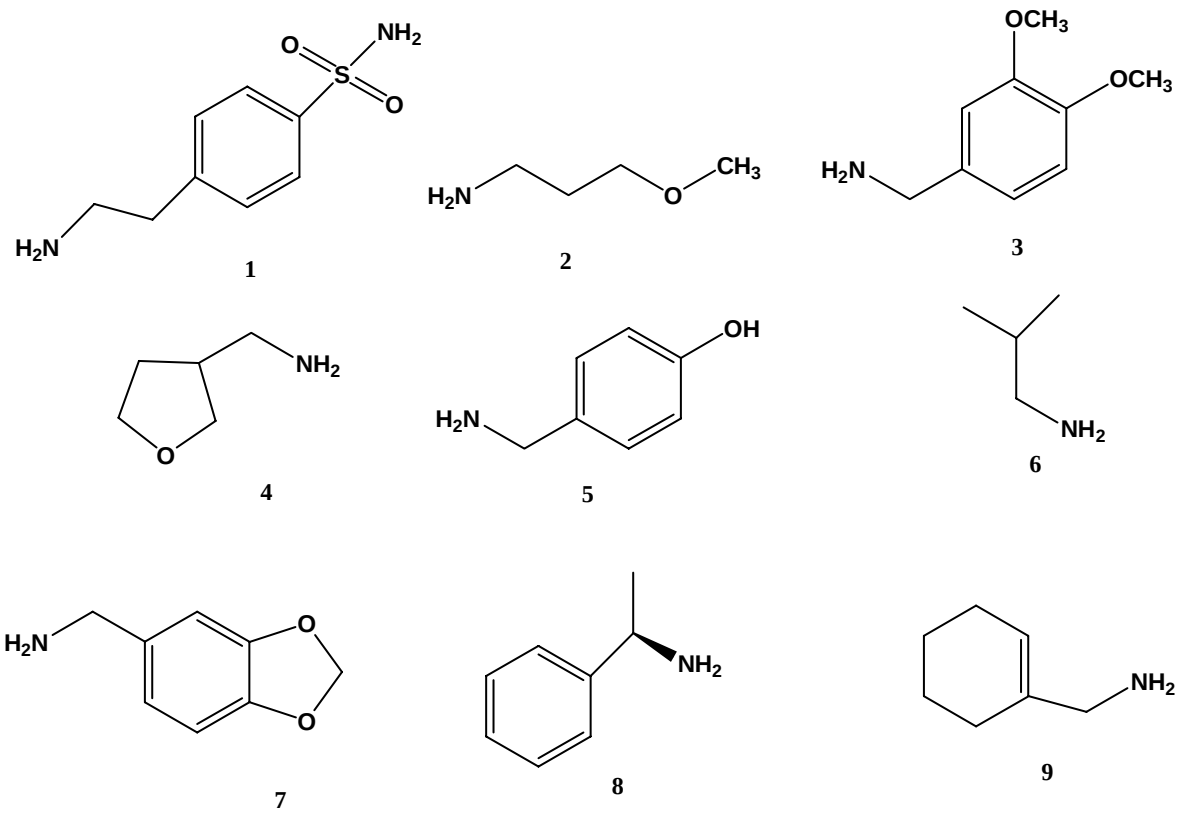


Figure 2.3: Amines used in microwave-assisted peptoid synthesis.

A list of amines used in the synthesis of various peptoids shown in Table 2.1.

		Microwave		37 °C		Room temperature	
Peptoid	Sequence	% Purity	% Yield	% Purity	% Yield	% Purity	% Yield
	(a)						
1	$(1)_9$	90	70	70	70	25	70
2	$(2)_9$	90	88	80	90	70	80
3	$(3)_9$	60	75	63	80	45	60
4	$(4)_9$	75	65	65	65	50	81
5	5-8-6-3-2--1-9-7-4	80	85	80	88	70	60

Table 2.1: Crude product characteristics.

The percentage purity and yield data of various peptoids synthesized using the microwave protocol in comparison with standard literature methods. (Figure courtesy of Dr. Hernando Olivos).

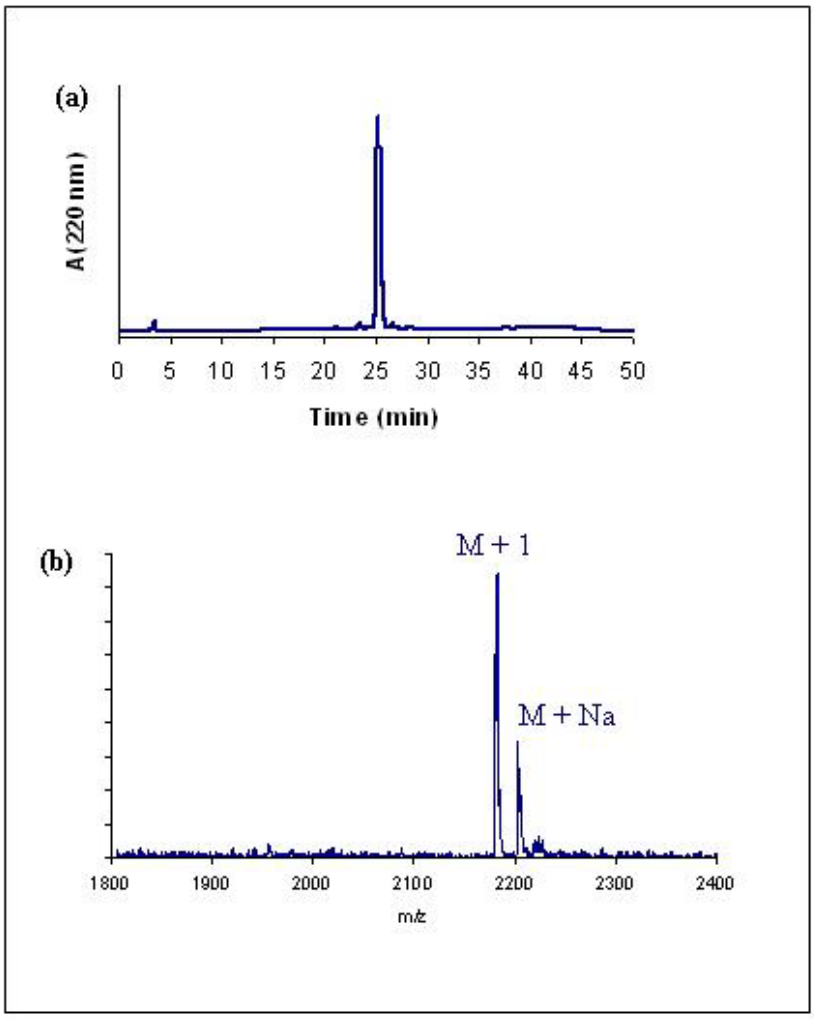


Figure 2.4: Peptoids synthesized using the microwave-assisted protocol are of excellent quality.

(a) A representative HPLC of one of the peptoids (10) synthesized using microwave chemistry. (b) MALDI-TOF mass spectrum of the main peak. (Figure courtesy of Dr. Hernando Olivos)

2.3.4 Edman sequencing of peptoids

One of the challenges associated with split and pool synthetic combinatorial libraries is the difficulty of unambiguously determining the chemical structure of compounds displayed on individual beads. The amount of compound present on each bead is too small to apply traditional techniques such as NMR for structural elucidation. Peptide libraries do not suffer from this drawback since extremely sensitive techniques such as MS-based C-terminal sequencing and N-terminal Edman sequencing are available, facilitating facile characterization of peptide libraries and hits derived from such libraries. It is perhaps one of the primary reasons (along with ease of synthesis) that led to the rampant use of split and pool peptide libraries among the scientific community.

Peptoids, fortunately, because of their structural proximity to peptides are also amenable for Edman sequencing. In fact, Liskamp has reported a protocol for the manual sequencing of peptoids using several beads as input[21]. Thus, the peptoid containing beads were treated with phenyl isothiocyanate, followed by TFA to cyclize the amino terminal residue into thiazolinone and release it from the beads. The filtrate was then heated in 1:1 water at 70 ^{o} C to convert the thiazolinone into thiohydantoin. This process was repeated until all the residues in the peptoid were converted into the corresponding thiohydantoins and manually injected into HLPC to obtain their retention times.

Although the Liskamp protocol provides an excellent starting point for sequencing peptoids by Edman degradation, it is of limited utility while dealing with large split and pool peptoid libraries, where it is critical to be able to sequence individual beads. The reason for Liskamp using multiple beads as input for Edman sequencing perhaps stems from the inefficiency and practical difficulty involved in manually handling individual beads and extremely small reaction volumes over multiple cycles of chemistry. While it is possible to employ larger (400-500 μm) TentaGel Macrobeads to overcome these sensitivity issues and even facilitate spectroscopic analysis of compounds derived from a single bead, such beads introduce a practical limitation on the size of the library that can be constructed. Therefore, an automated Edman degradation protocol for sequencing peptoids capable of overcoming all the shortcomings mentioned above was developed by adapting a commercial peptide sequencer (ABI 476A)[11].

The standard chemistry conditions for sequencing peptides in the commercial sequencer have also been found to be extremely effective with peptoids. However, the default HPLC protocols for separating the peptide derived-hydantoins have been found unsuitable for peptoids. Therefore, the program files have been suitably modified to facilitate efficient separation and identification of hydantoins derived from peptoid sequencing (Figure 2.5). A more detailed description of the changes has been provided in the experimental section. The optimized protocol was then used to determine the retention times of

all the monomers under investigation and a reference chart was generated. These retention times can then be used as standards for comparison and identification of unknown peptoid sequences (Table 2.2).

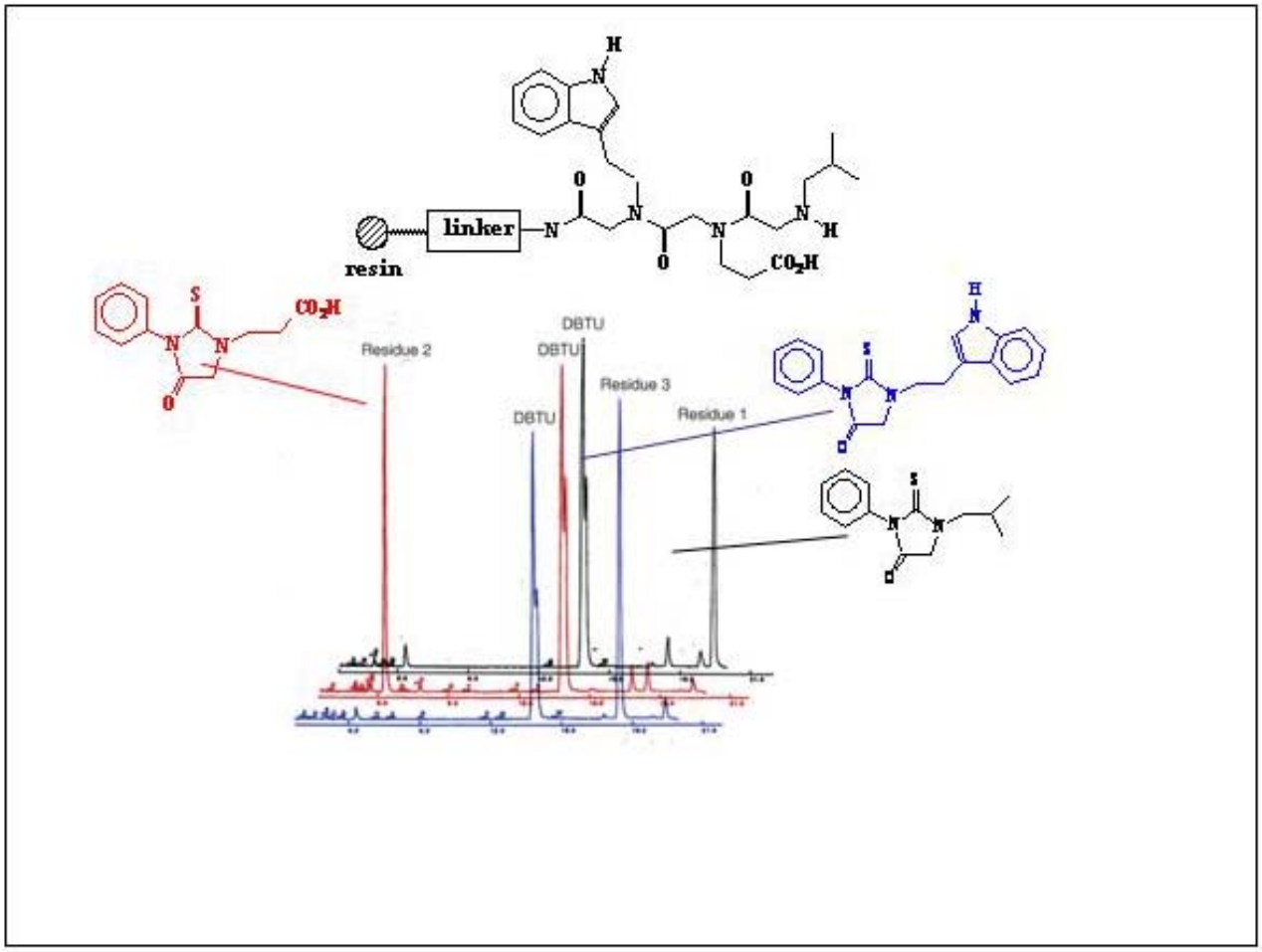


Figure 2.5: Automated Edman sequencing of peptoids.

Edman traces derived from sequencing of a three residue peptoid showing the well separation of the hydantoin-derived peaks, enabling ready and unambiguous identification of the sequence.

Amine	Retention time (min.)
Nser	6.4
Nall	13.9
Nlys	18.4
Npip	20.9
Nbsa	12.5
Nleu	19.5
Ntrp	21.3
Nman	16.1
Namp	15.1
Nmba	21.8
Napp	10.0
Nffa	17.0
Nmea	11.3

Table 2.2: Retention time standards for Edman sequencing.

The retention times of various peptoid-derived hydantoins obtained using the automated Edman degradation protocol.

2.3.5 Optimization of conditions for on-bead screening

Most of the on-bead screening assays reported in the literature employed fluorescein-tagged target proteins to allow visualization of bound protein to the ligand displayed on the bead. However, virtually all the commercial beads tested in this study had high "green" background fluorescence. This introduced a difficulty in screening experiments since high protein concentrations had to be employed in order to obtain signals above background levels. However, high protein concentrations also increased non-specific binding of the protein to the bead surface, making the screening process difficult and inefficient. In an attempt to overcome this difficulty, hyperspectral microscopy[22] was employed to study the fluorescence emission spectrum of the bead surface (Figure 2.6). Consistent with earlier observations that the beads had high background fluorescence in the green region, fluorescence emission was maximal in this region. However, the intensity of the bead fluorescence dropped off quite dramatically in the red portion of the spectrum.

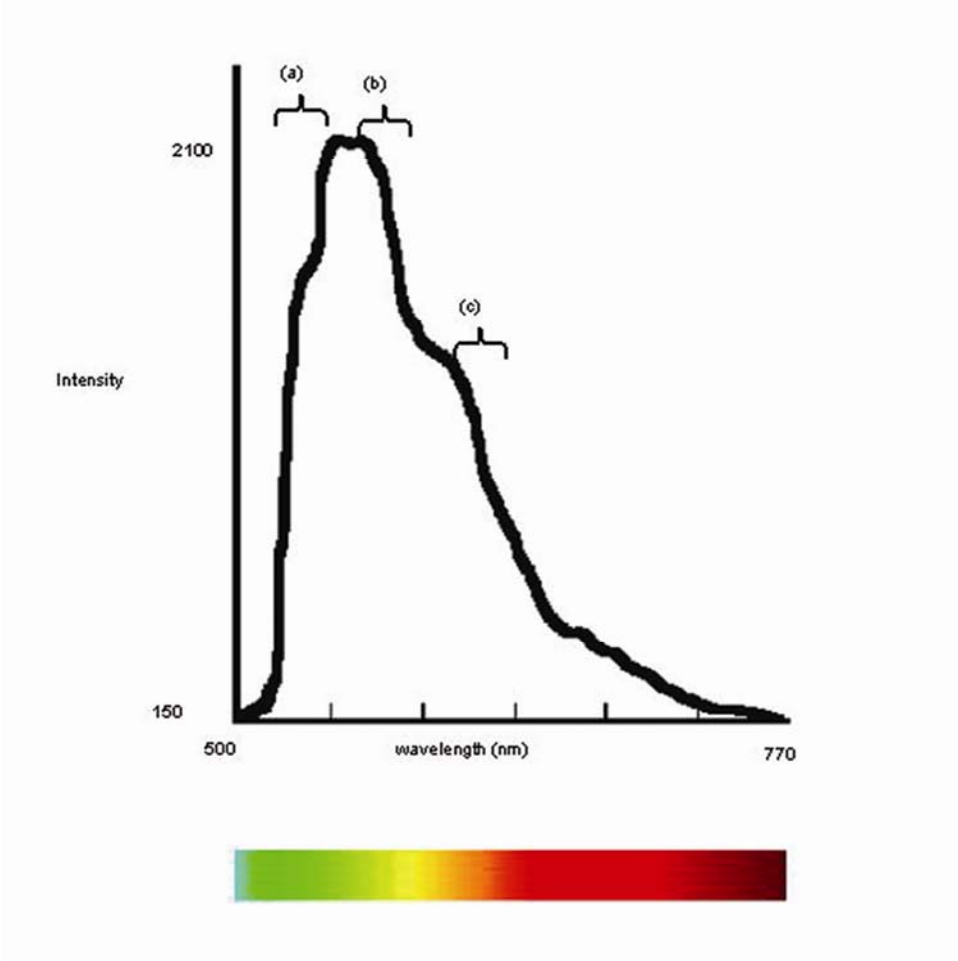


Figure 2.6: Bead fluorescence emission spectrum.

Fluorescence emission spectrum of a TentaGel Macrobead obtained using a hyperspectral imaging microscope. Also shown is the emission of some fluorescent dyes: (a) fluorescein, (b) tatramethylrhodamine and (c) Texas Red.

The emission spectrum of the bead surface suggested that organic dyes with emission in the red region of the spectrum will offer the best contrast in on-bead screenings. To test this hypothesis, Gal80BP[23], a previously characterized bonafide ligand for yeast transcription factor Gal80, was synthesized on TentaGel beads on an automated peptide synthesizer and incubated with a Texas Red-labeled Gal80 at 1 μM concentration for one hour in the presence of excess of BSA as competitor protein. As a control, a random 20 residue peptide was also synthesized and incubated with the protein under identical conditions. After thoroughly washing the beads with PBS buffer, both sets of beads were visualized under a fluorescence microscope. As shown in Figure 2.7, the beads displaying the Gal80BP were brightly fluorescent, while random peptide beads exhibited little fluorescence above background level. Therefore, Texas Red was used as the dye of choice in all subsequent on-bead screenings.

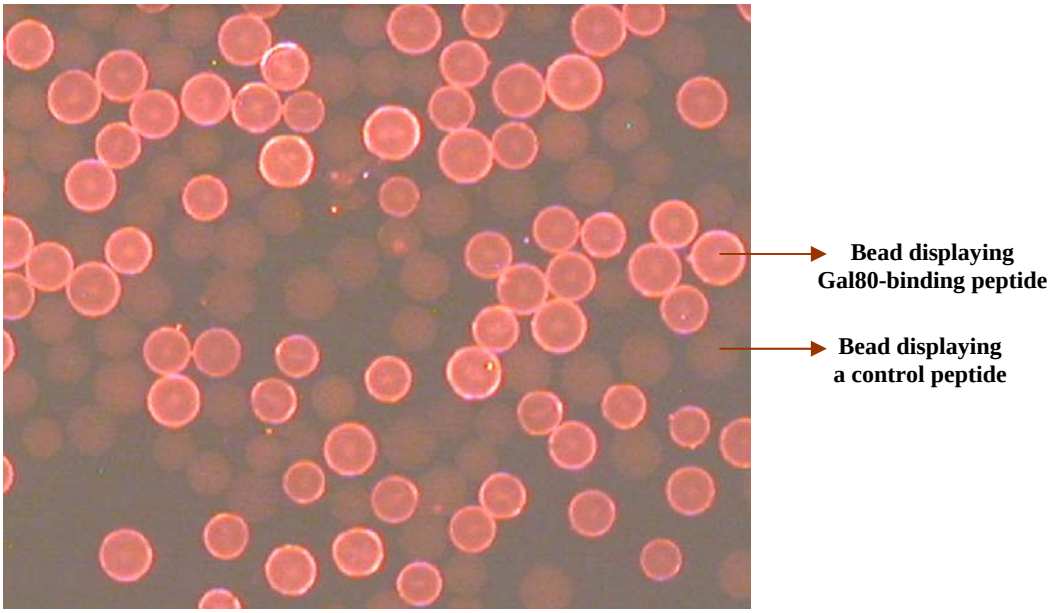


Figure 2.7: On-bead screening assay.

A photomicrograph of a model on-bead screening assay in which TentaGel beads displaying Gal80-binding peptide or a control peptide were incubated with Texas Red-labeled Gal80. (Figure courtesy of Dr. Kiran Sikder)

While manual microscopic visualization and collection of bright beads is certainly possible, an automated approach for bead screening presents several advantages. In addition to convenience, such an approach allows rapid pre-sorting of library beads, prior to incubation with protein, to remove any auto-fluorescent beads. The number of such auto-fluorescent beads, while small, is still significant resulting in false positives in screening experiments, if pre-sorting is not performed. Therefore, a commercial bead sorter (COPAS from Union Biometrica) was adapted for pre-sorting and screening of beads.

To carry out pre-sorting or screening, the library beads were dispersed in sheath buffer and passed through a syringe needle (to break clumps) into the sample cup. The beads then flow through a narrow capillary in single file and the fluorescence emanating from the beads following excitation with a laser is read by a fluorescence detector. The threshold of the gates can be suitably adjusted to only collect beads above a certain fluorescence level. Figure 2.8 depicts a model experiment in which approximately 100,000 library beads were incubated with Texas Red labeled- Gal 80 protein (100 nM) in the presence of other competitor proteins and subjected to automated sorting with COPAS. The plot generated by the machine depicts time of flight (TOF) on the x-axis and fluorescence intensity on the y-axis. As shown in the top panel of the figure, the gating (rectangular box) was suitably adjusted to exclude beads which had times of flight that were abnormally high (likely represent bead aggregates) or low (bead fragments or

other dust particles). The bottom panel depicts gating based on fluorescence intensity, which is adjusted to collect only beads with fluorescence intensities well above the background. Thus, a convenient and high throughput assay for screening large peptoid libraries was established.

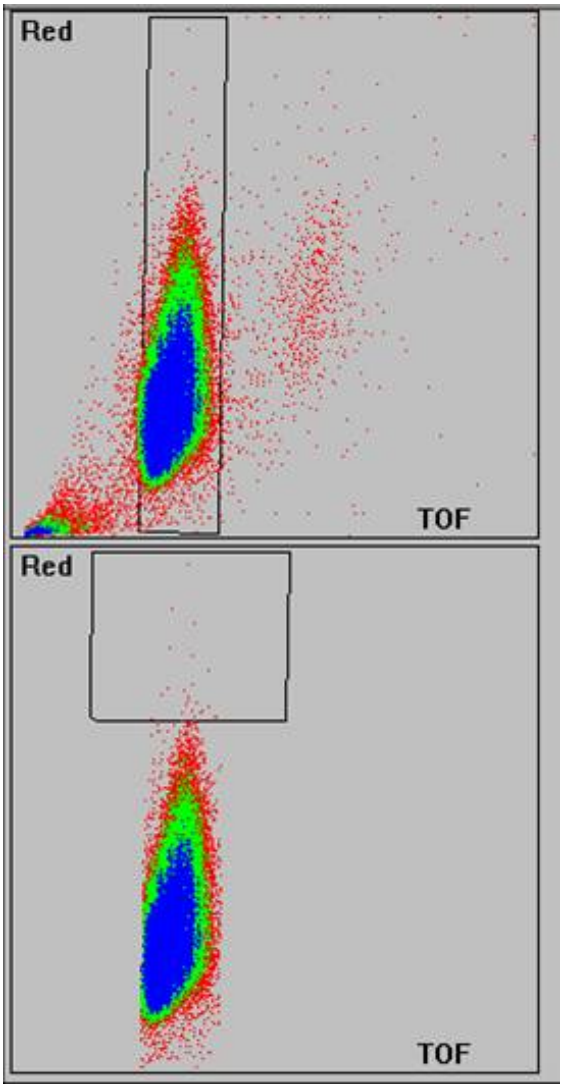


Figure 2.8: Automated screening of peptoid libraries.

A model screening experiment in which a peptoid library incubated with 100 nM Texas Red-labeled Gal80 was sorted using an automated bead sorter. The top panel shows gating based on time of flight (TOF), a measure of bead size and the bottom panel shows gating based on fluorescence intensity.

2.3.6 Synthesis and characterization of large peptoid libraries

Having thoroughly optimized the protocols for the synthesis, screening and sequencing of peptoids, a variety of libraries varying in size and diversity were constructed (Table 2.3). The first library utilized five amines and consisted of eight residue peptoids, providing a theoretical diversity of 390,625 compounds. The second library was only five residues long, but employed ten different monomers, providing a theoretical diversity of 100,000 compounds. Finally, an extremely large library of hexamers was made using nine different amines, providing a theoretical diversity of 531, 441 compounds. All the libraries were synthesized by standard split and pool synthesis[1] employing a slightly modified literature procedure[15] (8mer library) or microwave assisted procedure[10] (5 and 6mer libraries). All the amines were dissolved in DMF, except 4-(2-aminoethyl)benzene sulfonamide which was dissolved in DMSO. After the coupling of the first residue, the resin from all the vessels was pooled into a 250 mL glass peptide synthesis reaction vessel, mixed by bubbling argon through the suspension for 15 minutes and split before each acylation step. At the end of the library synthesis, the side chain protection groups were removed by treating with 95% TFA, 2.5% water and 2.5% anisole for 2h. The resin was then neutralized in 10% DIEA in DMF, washed with DCM and dried until further use.

Library 1

Length: 8 residues
Monomers used: 5 (Nbsa, Nleu, Nlys, Nmba, Ntrp)
Theoretical diversity: 390, 625

Library 2

Length: 5 residues
Monomers used: 10 (Namp, Napp, Nbsa, Nffa, Nleu, Nlys, Nman, Nmba, Nmea, Npip)
Theoretical diversity: 100, 000

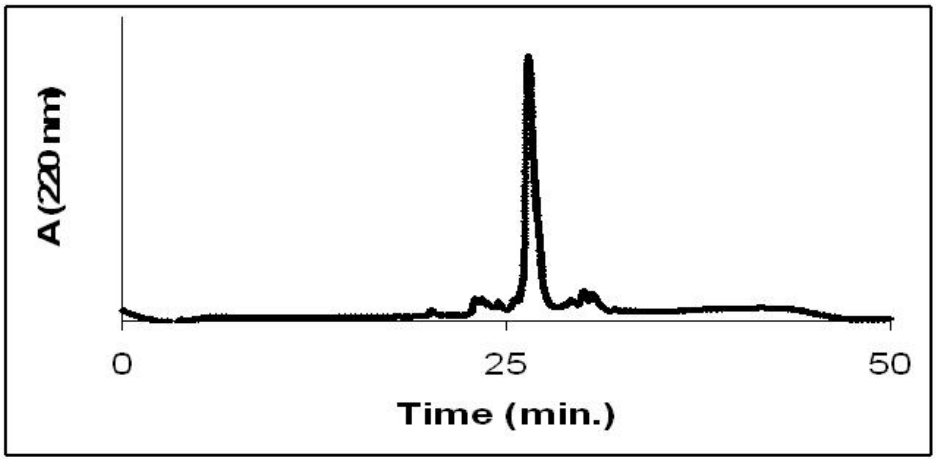
Library 3

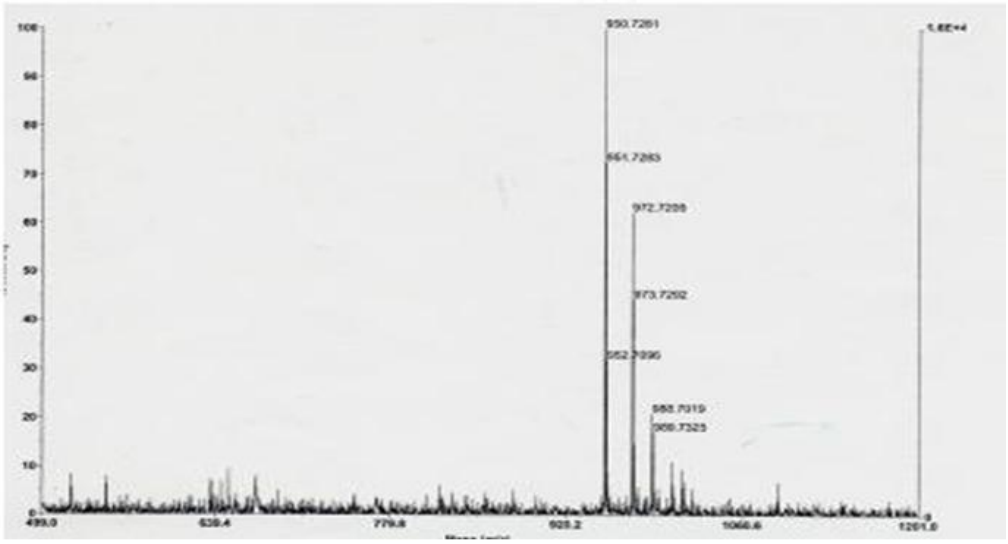
Length: 6 residues
Monomers used: 9 (Napp, Nbsa, Nffa, Nleu, Nlys, Nmba, Nmea, Npip, Ntrp)
Theoretical diversity: 531, 441

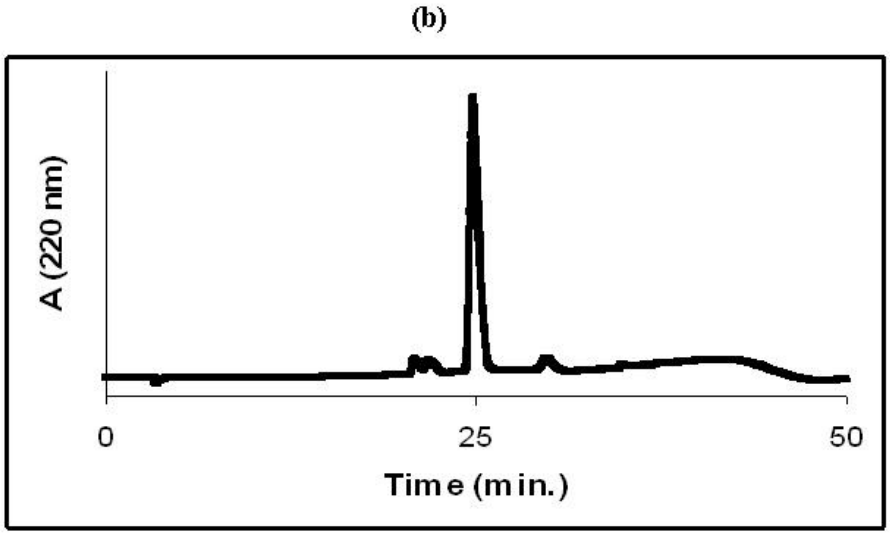
Table 2.3: The design characteristics of three peptoid libraries of varying diversity.

Detailed analyses were carried out to test the quality of each of these libraries. The data for the largest of the libraries are presented here. To address the likely purity of the library members, two mixed sequence hexamers, Ntrp-Nmea-Npip-Nlys-Nffa-Nmba-$CONH_2$ and Nbsa-Nleu-Napp-Nmea-Npip-$CONH_2$ were synthesized on Rink amide MHBA resin. The final product was released from the beads using 95% TFA, 2.5% water, 2.5% anisole, and the material was characterized HPLC and MALDI-TOF mass spectrometry. The data from the analysis suggest that the desired compounds were obtained in >85% purity (Figure 2.9). Between them, these hexamers contain all the monomers that were subsequently used in the library construction. The results suggest that in the absence of unexpected context effects, all of the coupling steps proceed in high yield.

(a)


A (220 nm)
0
25
50
Time (min.)


950.7261
1.8E+4
661.7263
972.7208
973.7252
952.7095
988.7319
989.7325
100
90
80
70
60
50
40
30
20
10
0
499.0
639.4
779.8
920.2
1060.6
1201.0



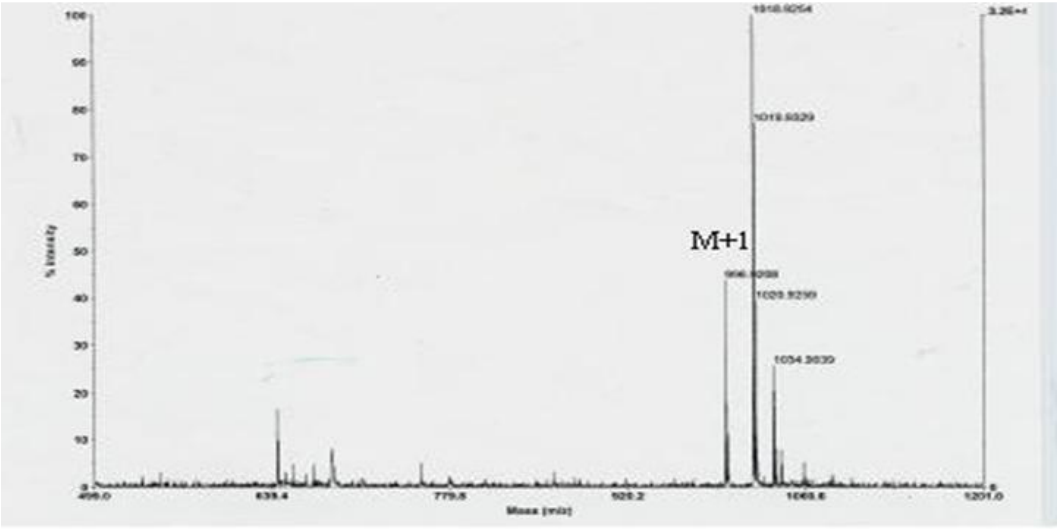
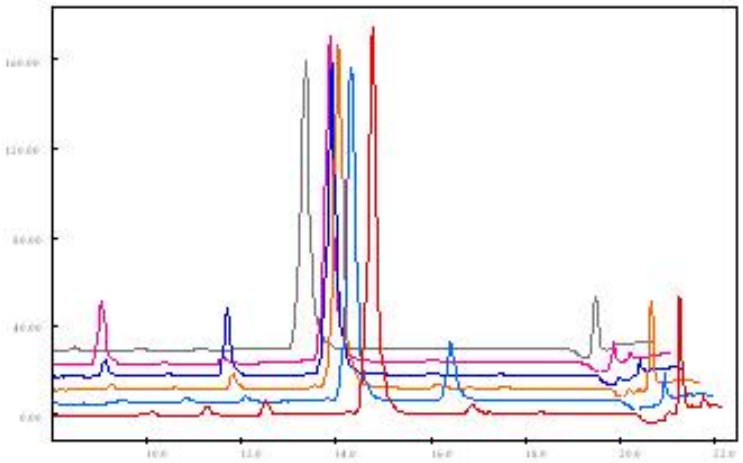


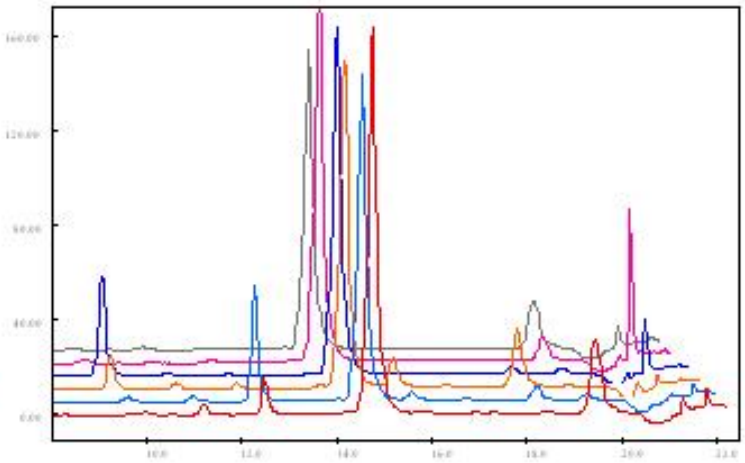
Figure 2.9: Characterization of peptoid libraries by HPLC and MALDI-TOF mass spectrometry.

HPLC traces and mass spectra of two hexameric peptoids, (a) Ntrp-Nmea-Npip-Nlys-Nffa-Nmba and (b) Nbsa-Nleu-Napp-Nffa-Napp-Npip, which, between them contain all the amines employed in the construction of the library.

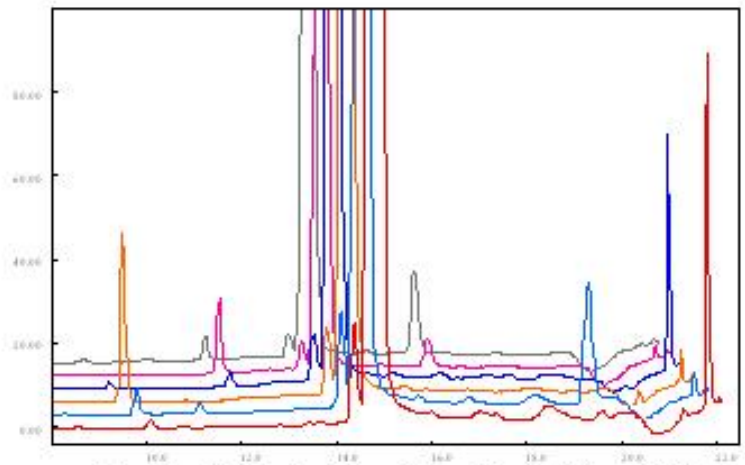
To get a more direct estimate of the quality of the library, ten beads were randomly picked from the library and subjected to Edman degradation analysis[11] (Figure 2.10). All the sequences obtained were different, as would be expected for a large, diverse library. The Edman traces also suggested that a full-length sequence was obtained in each case. Finally, analysis of the statistical distribution of each of the monomers in the ten random sequences suggested a relatively even distribution of all the monomers in the library, with no obvious bias for or against any monomers (Figure 2.11). Based on these data, it was concluded that the quality of the libraries synthesized was excellent and suitable for carrying out screening experiments.



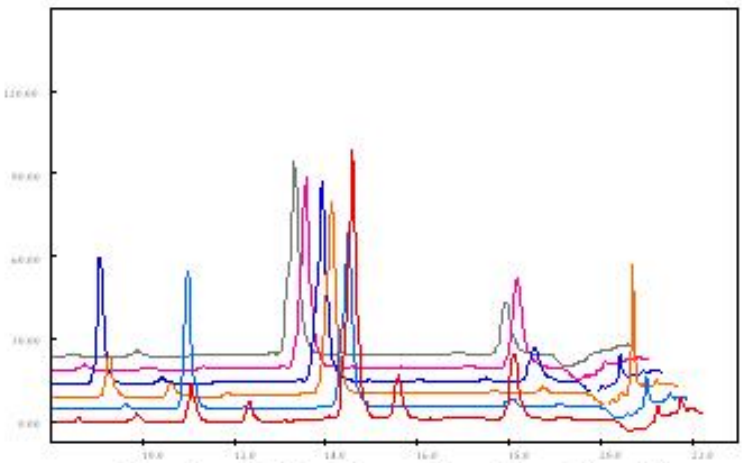
Random1: Ntrp-Nffa-Ntrp-Nbsa-Napp-Npip



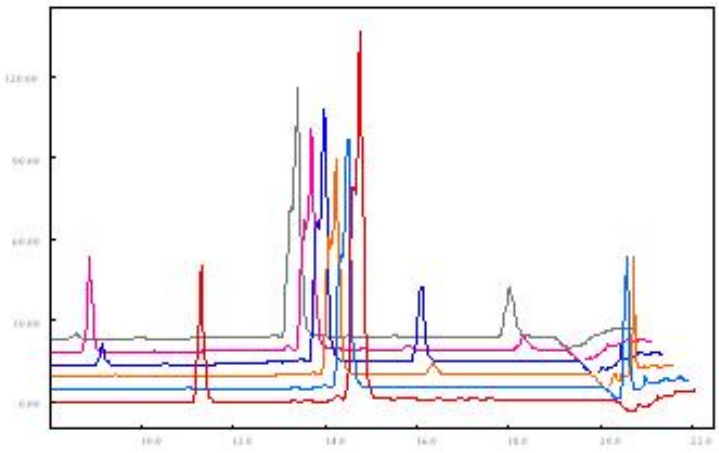
Random2: Nleu-Nbsa-Nlys-Napp-Ntrp-Nleu



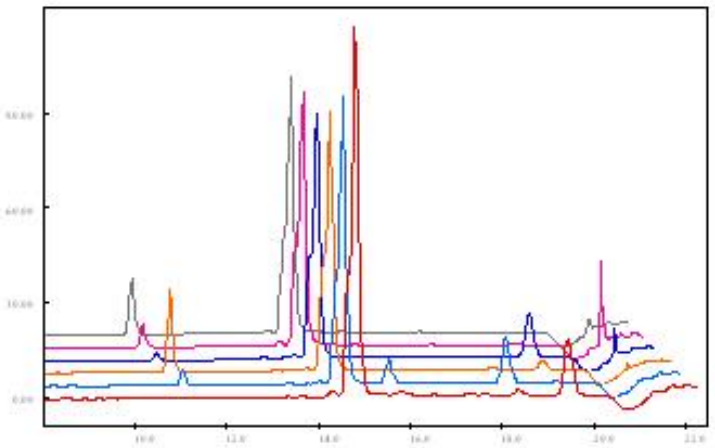
Random3: Nmba-Nleu-Napp-Nmba-Nbsa-Nffa



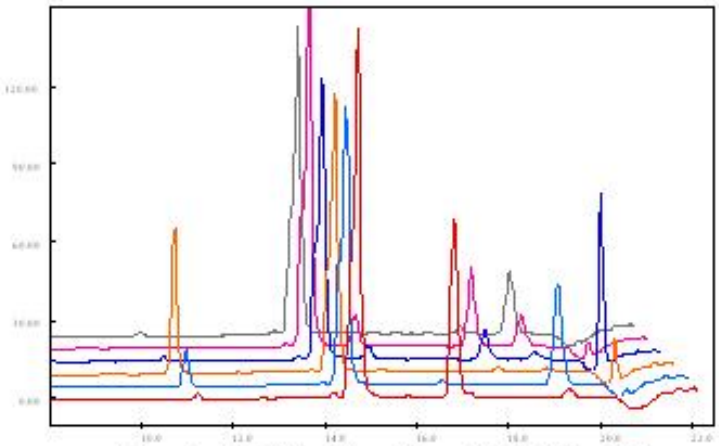
Random4: Nlys-Nmea-Ntrp-Napp-Nleu-Nleu



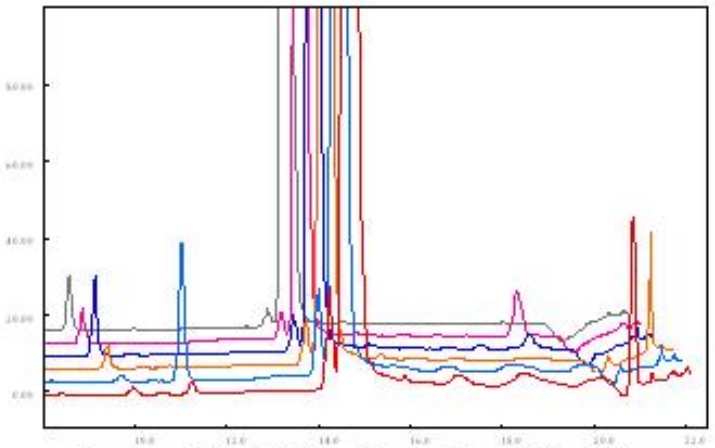
Random5: Nmea-Npip-Nmba-Nffa-Napp-Nleu



Random6: Nleu-Nlys-Nmea-Nleu-Ntrp-Nmea



Random7: Nffa-Nleu-Nmea-Npip-Nlys-Nleu



Random8: Npip-Nmea-Nmba-Napp-Nleu-Napp

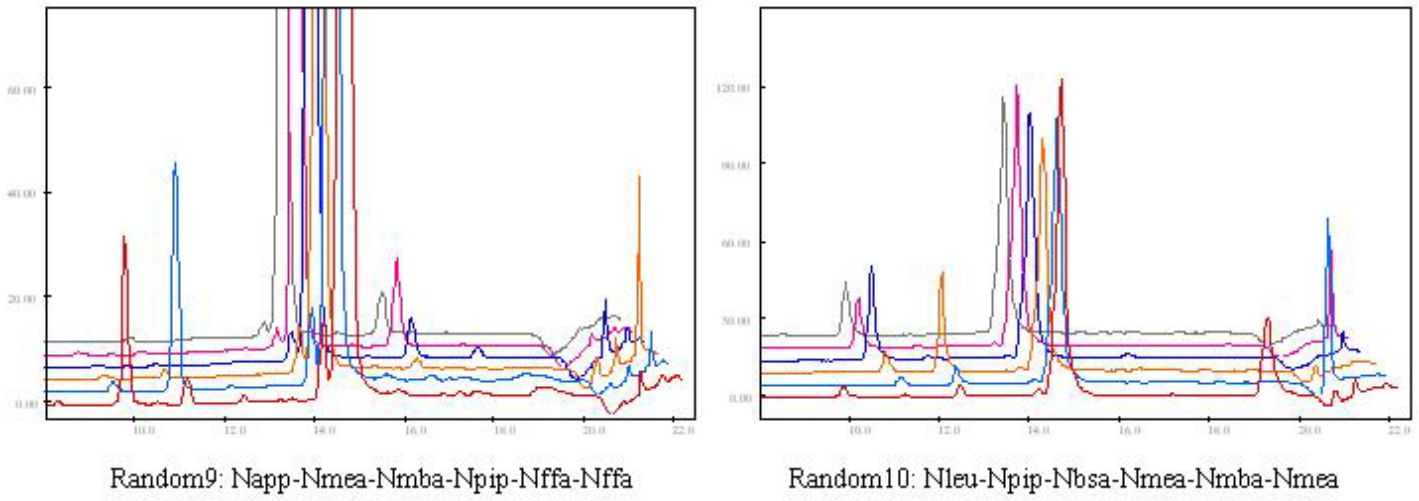


Figure 2.10: Characterization of peptoid libraries by Edman sequencing.

Edman traces from ten beads randomly picked from the hexameric 531,441-compound peptoid library. The corresponding text sequence is shown below each trace. A full-length sequence was obtained in each case.

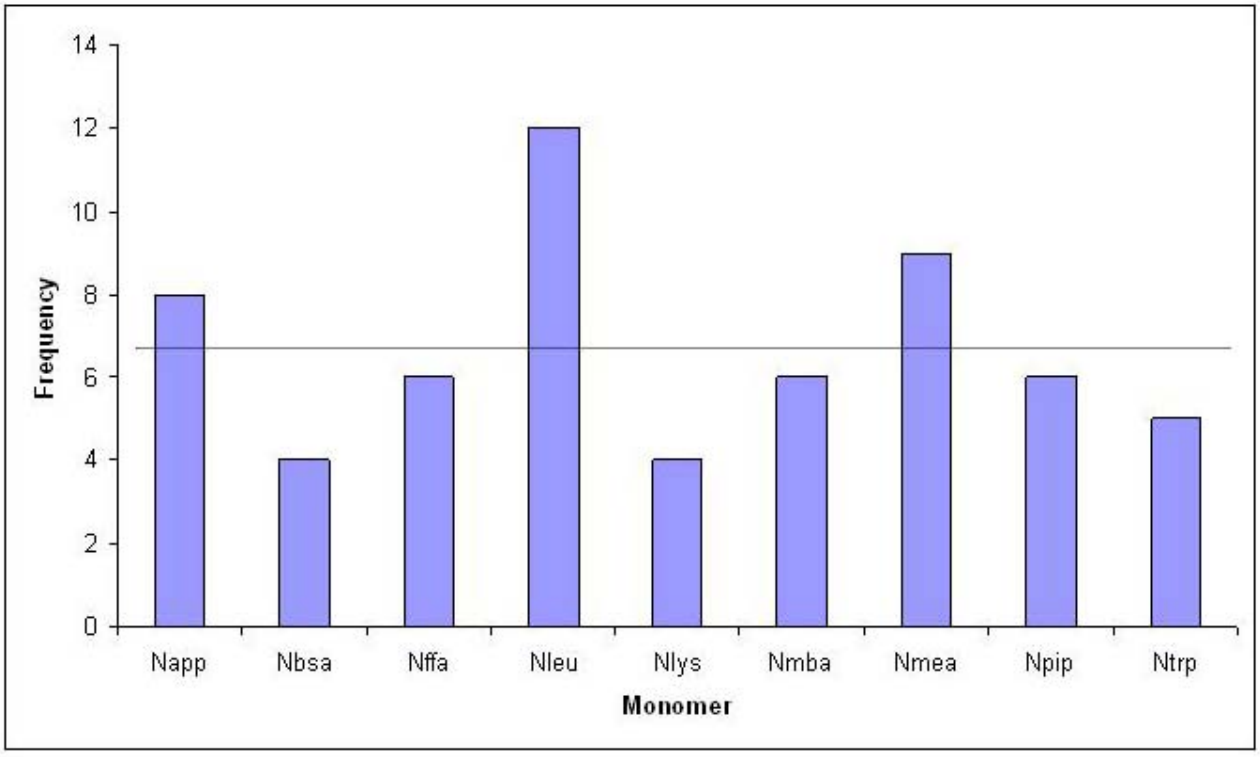


Figure 2.11: Frequency distribution of various monomers in a peptoid library.

An estimate of the frequency distribution of each monomer in the hexameric library was computed from ten random sequences. The horizontal line represents the expected theoretical mean distribution for each monomer.

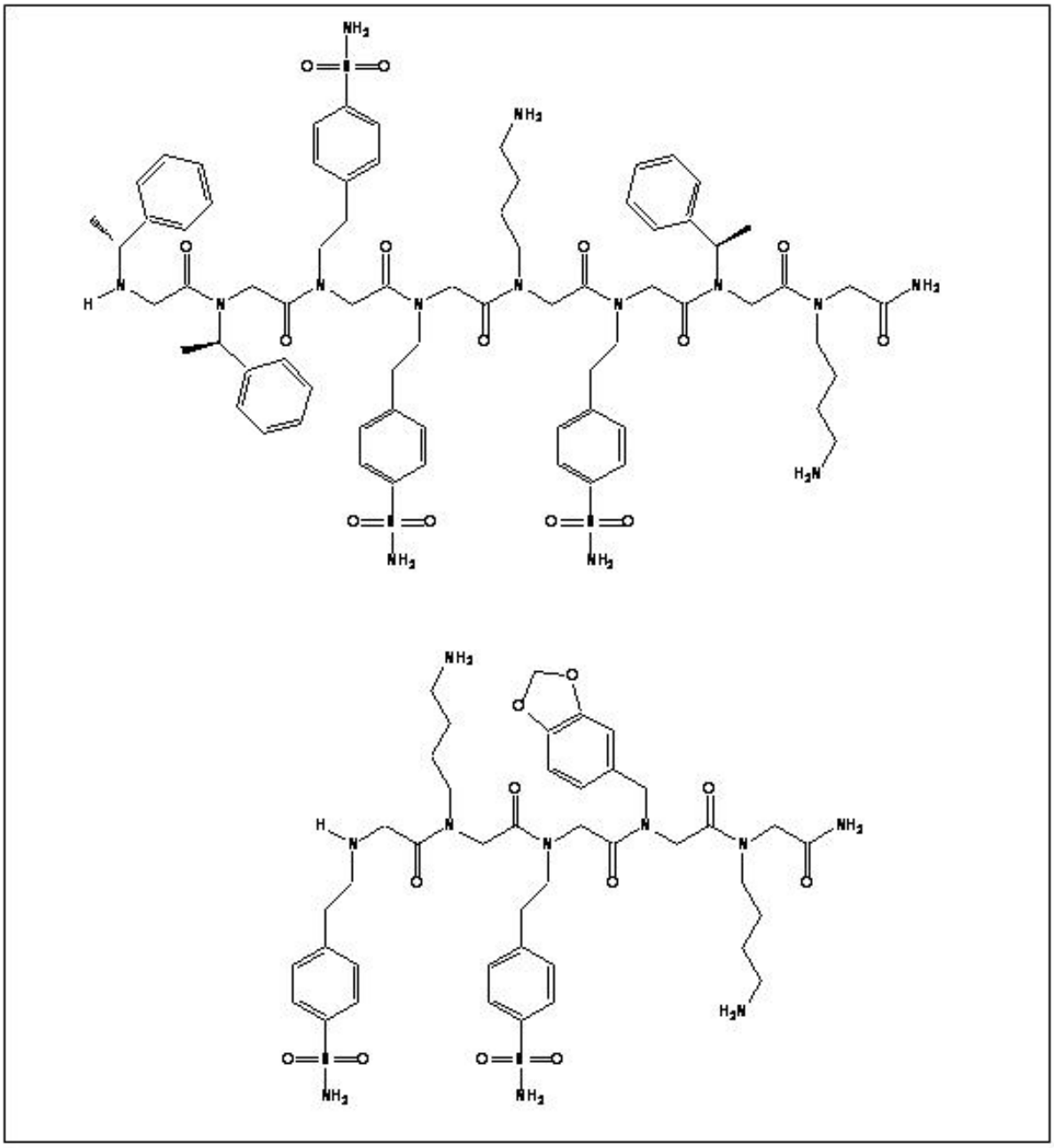
2.3.7 Screening of peptoid libraries against GST

To determine if these libraries would be facile sources of protein ligands, parts of the 5mer and 8mer libraries were screened against Texas Red-labeled glutathione S-transferase (GST). GST was chosen as a target because of its ready availability and with the hope that this test screen will enable developing the right conditions for future screenings against relevant transcription factors.

In the first screening experiment, GST was labeled with Texas Red, by the treating the protein with Texas Red-NHS ester in a pH 8.3 sodium bicarbonate buffer under conditions that resulted, on an average, in each protein molecule receiving one molecule of the dye. The excess dye was separated from the protein using a desalting column. A 2 μM solution of the labeled protein was then incubated with approximately 100, 000 beads of the 8mer library in the presence of a large excess of unlabeled BSA, which serves as a competitor protein. In a second experiment, to demand more specific and tighter binders, approximately 50,000 beads from the 5mer library were incubated with 1 μM solution of the labeled protein in the presence of a 1000-fold excess of *E. coli* lysate for one hour. The beads were then washed thoroughly to remove any unbound protein and visualized under a fluorescence microscope fitted with a Texas Red filter. Approximately 1% of beads in the 8mer library and 0.5% of beads in the 5mer library exhibited fluorescence above background levels. One of the brightest beads from each library was isolated, washed with PBS buffer, heated in 1% SDS

solution for 20 minutes at 95^{o}C, rinsed with distilled water and subjected to on-bead Edman degradation analysis. The SDS treatment was necessary since beads that have been exposed to lysates did not sequence well. Presumably, this was caused by components of the lysate coating the beads and interfering with the sequencing process. Consistent with this view, nascent beads that have not been exposed to lysates sequenced quite well. The retention times of hydantoins from each cycle of Edman sequencing of the hits were compared to the reference standards and the chemical structures of the hits were deduced unambiguously. A comparison of the two structures revealed they were closely related. In fact, four out of the five positions in the 5mer were identical to the 8mer peptoid (Figure 2.12). It was quite remarkable that the screening of two libraries of different lengths and diversities yielded hits that were almost identical.

(a)



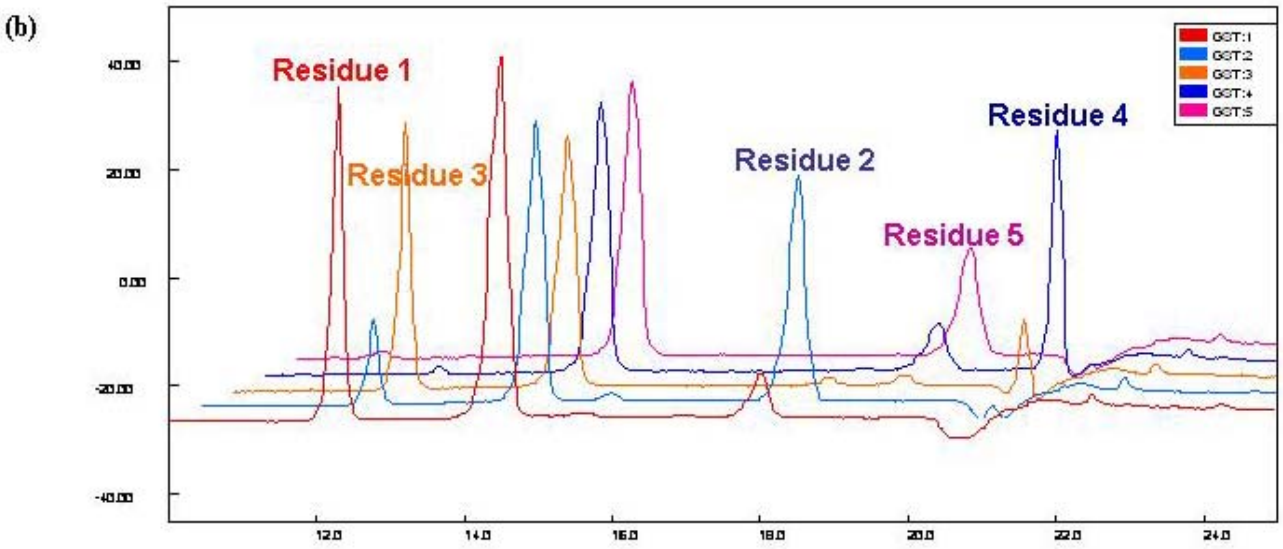


Figure 2.12: Identification of GST-binding peptoids.

(a) The chemical structures of two novel GST ligands isolated from eight- and five- residue peptoid libraries. (b) The Edman trace used to elucidate the chemical structure of the short (5mer) peptoid.

In order to validate the two novel GST ligands, the compounds were re-synthesized on TentaGel beads and incubated with 500 nM solution of Texas Red-labeled GST or MBP, an unrelated protein, in the presence of 2% BSA as competitor. As a control, a random 6mer peptoid was also treated with labeled GST. The beads were then thoroughly washed and visualized under fluorescence microscope. As can be seen in figure 2.13, the hit beads captured the GST protein efficiently, while no detectable binding of the control protein was observed. The control peptoid also failed to capture any detectable levels of GST. In order to see if the observed binding is dependent on the concentration of the protein added, the 5mer peptoid was treated with various concentrations of the protein in the presence of 100-fold excess of *E. coli* lysate. As seen in figure 2.14, dose dependent binding of the protein was observed from concentrations of 1 μM to 100 nM. Little or no fluorescence above background was noticed at a protein concentration of 10 nM.

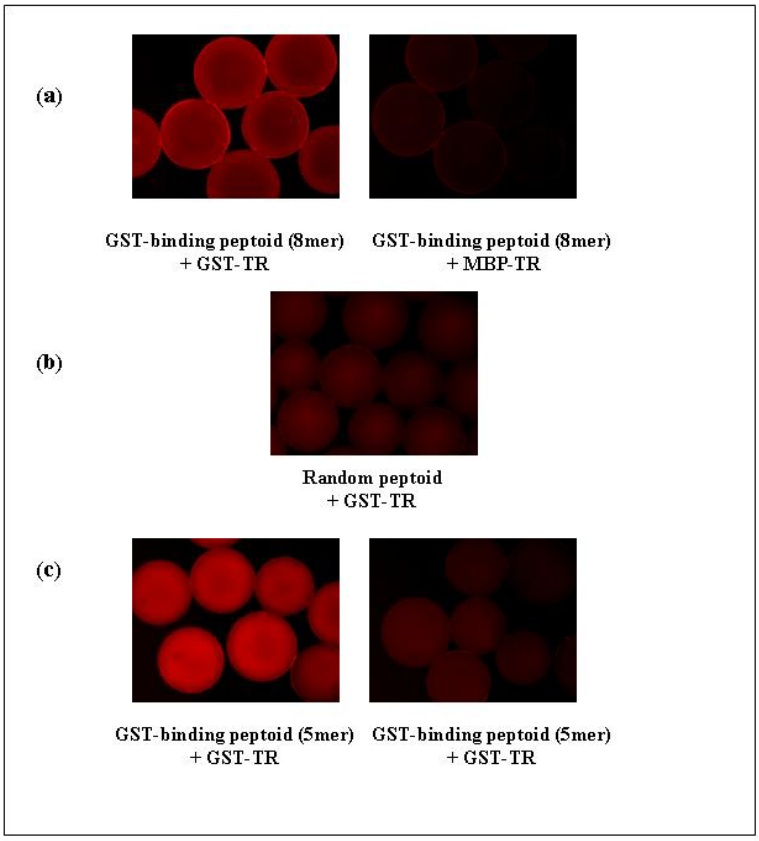


Figure 2.13: Characterization of on-bead binding properties of GST ligands.

(a) Photomicrographs obtained after incubation of TentaGel beads displaying the 8mer GST ligand with 500 nM Texas Red-labeled GST (left panel) or Texas Red-labeled MBP (right panel).

(b) Similar experiment in which a bead displaying a random peptoid (Npip-Nser-Nbsa-Nall-Nlys-Npip) was incubated with 500 nM Texas Red-labeled GST.

(c) Same experiment as in (a), but with TentaGel beads displaying the 5mer GST ligand.

2% BSA solution was included in all the experiments to prevent non-specific interactions.

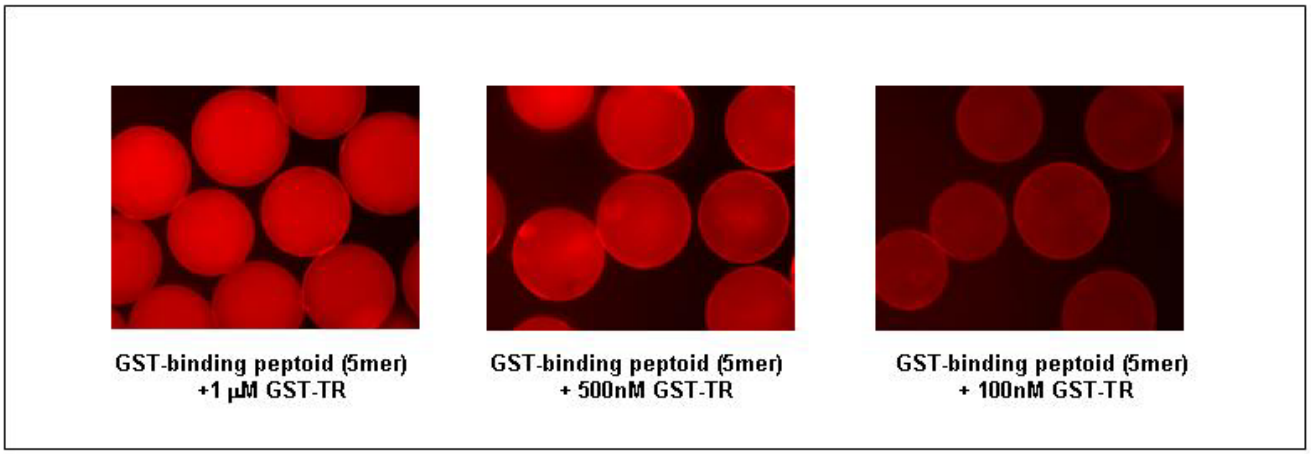


Figure 2.14: Dilution experiment.

Photomicrographs showing the capture of Texas Red-labeled GST by the 5mer GST ligand at the protein concentrations indicated. All solutions included a 100-fold excess of *E. coli* lysate.

To eliminate the possibility that the hits recognize the Tex Red labeled-GST and not the native protein, a pull-down experiment was performed using unmodified GST protein. Thus, TentaGel beads displaying the hit peptoids were incubated with unlabeled GST in the presence of 100-fold excess of *E. coli* lysate. The beads were thoroughly washed, bound protein eluted and subjected to SDS-PAGE / Western blot analysis. As shown in Figure 2.15, both the peptoids retained the native GST protein effectively, while the control peptoid or beads alone failed to show any detectable pull-down of the protein. Based on these results, it was concluded that the bead bound peptoid ligands were capable of selectively recognizing GST in milieu of vast excess of other proteins.

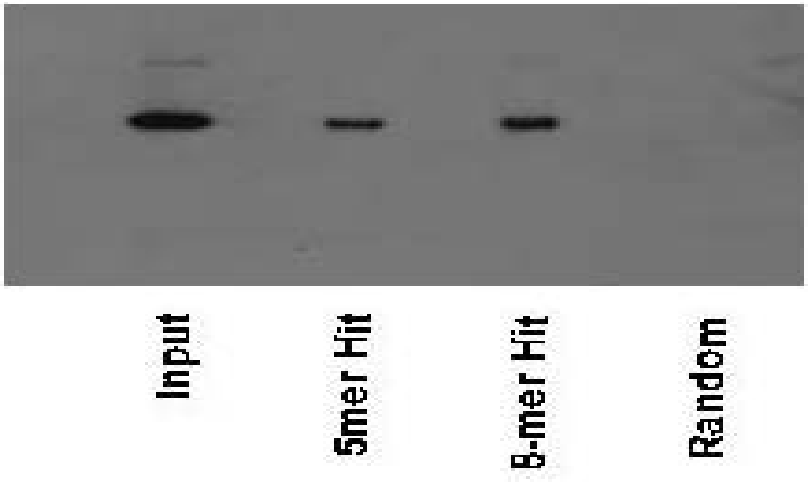


Figure 2.15: Capture of native GST by TentaGel displayed peptoid ligands. Western blot obtained using anti-GST antibodies to measure the capture of GST, by the bead bounds ligands, from a solution containing the protein at 1 μM concentration, in presence of a 100-fold excess of *E. coli* extract. A random peptoid (Nmba-Nbsa-Nleu-Nlys-Npip-Nmba-Nleu-Nleu) was also included in the experiment as a control. (Figure courtesy of Dr. Kiran Sikder)

Finally, to investigate the solution binding properties, the 5mer peptoid was re-synthesized on Rink resin, cleaved from beads and purified to homogeneity by HPLC. Solution binding studies of the peptoid with the protein were conducted using Isothermal Titration Calorimetry (ITC)[14]. The data from the study (Figure 2.16) revealed the equilibrium dissociation constant of the peptoid **.** protein complex to be approximately 60 μM.

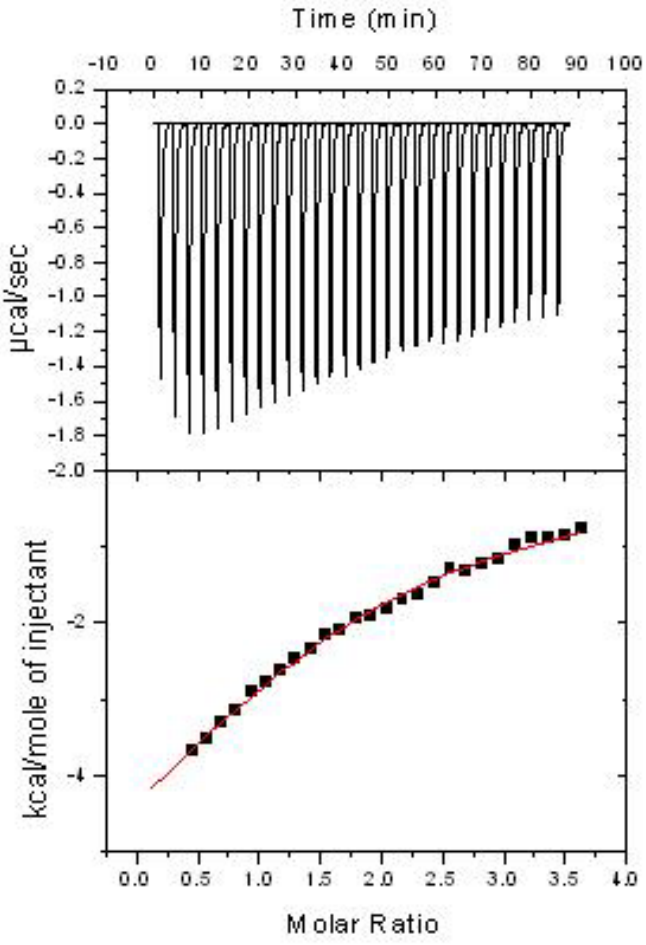


Figure 2.16: Characterization of peptoid-protein interaction by ITC.

ITC trace for binding of the five residue peptoid to GST showing the raw data (top panel) and the best fit of the data points (bottom panel). The K_D value of the peptoid-protein complex was found to be 62 µM. (Figure courtesy of Dr. Kiran Sikder)

2.4 Discussion

Peptoid libraries are a promising and convenient source of protease resistant protein ligands. The sub-monomer approach[2] to peptoid synthesis introduced by Zuckermann and colleagues provides a facile route for the synthesis of split and pool libraries[1] of peptoids using inexpensive and commercially available primary amines as the diversity generating elements. In an attempt to expand the collection of amines available for peptoid synthesis, several new monomers were tested and found to be suitable for incorporation into peptoid sequences (Figure 2.1). Similarly, in order to increase the efficiency and speed of peptoid synthesis, a microwave assisted protocol for peptoid synthesis was developed. This protocol allows for tremendous acceleration of individual coupling steps, making it possible to synthesize relatively large peptoid libraries in less than a day. Furthermore, the quality of peptoids synthesized in this route was found to be comparable or superior to reported literature procedures (Table 2.1).

In spite of the several favorable properties of peptoids, there have been only a few reports of construction of peptoid libraries[5-9]. In order to overcome the various bottlenecks associated with peptoid library synthesis, a systematic study was carried out to optimize the conditions necessary for the synthesis and screening of large peptoid libraries[11]. After considerable exploration, TentaGel Macrobeads were chosen as the solid support of choice for various reasons discussed earlier.

Similarly, Texas Red was selected as the flour of choice for labeling proteins since the bead surface had minimum background fluorescence in the red region of the fluorescence spectrum (Figure 2.6). Furthermore, an automated Edman sequencing protocol that allows unambiguous elucidation of chemical structure of peptoids bound to individual beads was developed (Figure 2.5). This cleared a major roadblock for construction and characterization of relatively large (>100, 000) peptoid libraries. Finally, optimized conditions were developed for manual as well as automated screening of bead bound peptoid libraries using Texas Red-labeled target proteins (Figures 2.7 and 2.8).

With the infra-structure in place, three peptoid libraries ranging in diversity from ~100, 000 compounds to > 500, 000 compounds were constructed (Table 2.3). A variety of quality control tests indicated that the libraries were of excellent quality (Figures 2.9 – 2.11). Two of the libraries were screened against Texas Red-labeled GST protein and the chemical structure of two of the brightest hits (one from each library) was determined by Edman sequencing (Figure 2.12). A strong consensus was found between the two peptoids, with four out of five positions in the 5mer peptoid being identical to the 8mer peptoid. A variety of on-bead assays were carried out using Texas Red-labeled and unmodified GST to assess the specificity of binding of the ligands to their target protein (Figures 2.13 - 2.15). The data from these studies indicated that both the ligands were capable of capturing the target protein in the presence of large excess of unrelated proteins

even at protein concentrations as low as 100 nM. Finally, solution binding studies were carried out using ITC to ensure that the binding characteristics of the peptoids on the bead surface translate into binding in solution. The data from the ITC trace revealed the equilibrium dissociation constant of the peptoid-protein complex to be ~ 60 μM (Figure 2.16). Overall, the data from this study suggest that peptoid libraries can serve as a general source of protein ligands. This study, thus, suggests that isolation of a peptoid-based activation domain surrogate could be conveniently achieved by screening a peptoid library against appropriate transcription factors.

A general protocol for isolation of non-peptidic, potentially cell permeable synthetic ligands could be envisioned to have several applications beyond modulation of transcription. Thus, relatively short (3-5mer) peptoid libraries could screened against various enzymes, receptors and other drug targets to facilitate discovery of ligands that serve as a starting point for drug discovery endeavors. Such molecules could also serve as valuable tools in basic biology by manipulating various signal transduction pathways and protein-protein interactions. The data from this study[11] and others[24] also indicate that peptoids ligands could also be used to capture / detect individual proteins from a complex milieu and, hence, could be of value for construction of protein detecting microarrays. Finally, in a recent study[25], Reddy and Kodadek used a peptoid library printed on glass side to develop “fingerprints” of individual proteins in

complex protein mixtures, enabling their convenient identification. Thus, it may also be possible to use peptoids arrays to “fingerprint” complex physiological or pathological states in an attempt to develop new generation diagnostic devises, which can detect disease onset before symptoms appear. It remains to be seen if the tremendous potential of these molecules will be fully realized.

2.5 Bibliography

1. Lam, K.S., Salmon, S.E., Hersh, E.M., Hruby, V.J., Kazmierski, W.M., and Knapp, R.J., *A new type of synthetic peptide library for identifying ligand-binding activity.* Nature, 1991. **354**: p. 82-84.
2. Zuckermann, R.N., Kerr, J.M., Kent, S.B.H., and Moos, W.H., *Efficient method for the preparation of peptoids (oligo N-substituted glycines) by submonomer solid-phase synthesis.* J. Amer. Chem. Soc., 1992. **114**: p. 10646-10647.
3. Miller, S.M., Simon, R.J., Ng, S., Zuckermann, R.N., Kerr, J.M., and Moos, W.H., *Comparison of the proteolytic susceptibilities of homologous L-amino acid, D-amino acid, and N-substituted glycine peptide and peptoid oligomers.* Drug Dev. Res., 1995. **35**: p. 20-32.
4. Yu, P., Liu, B., and Kodadek, T., *A high-throughput assay for assessing the cell permeability of combinatorial libraries.* Nat. Biotechnol., 2005. **23**: p. 746-751.
5. Zuckermann, R.N., Martin, E.J., Spellmeyer, D.C., Stauber, G.B., Shoemaker, K.R., Kerr, J.M., Figliozzi, G.M., Goff, D.A., Siani, M.A., et al., *Discovery of nanomolar ligands for 7-transmembrane G-protein-coupled receptors from a diverse N-(substituted)glycine peptoid library.* J. Med. Chem., 1994. **37**: p. 2678-2685.

6. Nguyen, J.T., Porter, M., Amoui, M., Miller, W.T., Zuckermann, R.N., and Lim, W.A., *Improving SH3 domain ligand selectivity using a non-natural scaffold.* Chem. & Biol., 2000. **7**: p. 463-473.

7. García-Martínez, C., Humet, M., Planells-Cases, R., Gomis, A., Caprini, M., Viana, F., De la Peña, E., Sanchez-Baeza, F., Carbonell, T., Felipe, C.D., Pérez-Payá, E., Belmonte, C., Messeguer, A., and Ferrer-Montiel, A., *Attenuation of thermal nociception and hyperalgesia by VR1 blockers.* Proc Natl Acad Sci U S A, 2002. **99**: p. 2374-2379.

8. Humet, M.C., T., Masip, I., Sanchez-Baeza, F., Mora, P., Canton, E., Gobernado, M., Abad, C., Perez-Paya, E., and Messeguer, A., *A positional scanning combinatorial library of peptoids as a source of biological active molecules: Identification of Antimicrobials.* J. Comb. Chem., 2003. **5**: p. 597-605.

9. Masip, I., Cortes, N., Abad, M-J., Guardiola, M., Perez-Paya, E., Ferragut, J., Ferrer-Montiel, A., and Messeguer, A., *Design and synthesis of an optimized positional scanning library of peptoids: identification of novel multidrug resistance reversal agents.* Bioorg. Med. Chem., 2005. **13**(6): p. 1923-1929.

10. Olivos, H.J., Alluri, P.G., Reddy, M.M., Salony, D., and Kodadek, T., *Microwave-assisted solid phase synthesis of peptoids.* Org. Lett., 2002. **4**: p. 4057-4059.

11. Alluri, P.G., Reddy, M.M., Bachhawat-Siker, K., Olivos, H.J., and Kodadek, T., *Isolation of protein ligands from large peptoid libraries.* J. Amer. Chem. Soc., 2003. **125**: p. 13995-14004.

12. Uno, T., Beausoleil, E., Goldsmith, R.A., Levine, B.H., and Zuckermann, R.N., *New submonomers for poly N-substituted glycines (peptoids).* Tet. Lett., 1999. **40**: p. 1475-1478.

13. Pons, J.-F., Fauchere, J-L., Lamaty, F., Molla, A., and Lazaro, R., *A constrained diketopiperazine as a new scaffold for the synthesis of peptidomimetics.* Eur. J. Org. Chem., 1998: p. 853-859.

14. Leavitt, S., and Freire, E., *Direct measurement of protein binding energetics by isothermal titration calorimetry.* Curr. Opin. Struct. Biol., 2001. **11**(5): p. 560-566.

15. Kirshenbaum, K., Barron, A.E., Goldsmith, R.A., Armand, P., Bradley, E.K., Truong, K.T.V., Dill, K.A., Cohen, F.E., and Zuckermann, R.N., *Sequence-specific polypeptoids: a diverse family of heteropolymers with stable secondary structure.* Proc. Natl. Acad. Sci. USA, 1998. **95**: p. 4303-4308.

16. Figliozzi, G.M., Goldsmith, R., Ng, S.C., Banville, S.C., and Zuckermann, R.N., *Synthesis of N-substituted glycine peptoid libraries.* Methods Enzymol., 1996. **267**: p. 437-447.

17. Wender, P.A., Mitchell, D.J., Pattabiraman, K., Pelkey, E.T., Steinman, L., and Rothbard, J.B. *The design, synthesis, and evaluation of molecules that enable or enhance cellular uptake: peptoid molecular transporters.* Proc. Natl. Acad. Sci. USA, 2000. **97**: p. 13003-13008.

18. Simon, R.J., Kania, R.S., Zuckermann, R.N., Huebner, V.D., Jewell, D.A., Banville, S., Ng, S., Wang, L., Rosenberg, S., Marlowe, C.K., Spellmeyer, D.C., Tan, R., Frankel, A.D., Santi, D.V., Cohen, F.E., and Bartlett, P.A., *Peptoids: A modular approach to drug discovery.* Proc. Natl. Acad. Sci. USA, 1992. **89**: p. 9367-9371.

19. Lew, A., Krutzik, P.O., Hart, M. E., and Chamberlin, A. R., *Increasing rates of reaction: microwave-assisted organic synthesis for combinatorial chemistry.* J. Comb. Chem., 2002. **4**: p. 95-105.

20. Yu, H.M., Chen, S.T., and Wang, K.T., *Enhanced coupling efficiency in solid-phase peptide synthesis by microwave irradiation.* J. Org. Chem., 1992. **57**: p. 4781-4784.

21. Boeijen, A., and Liskamp, R.M.J., *Sequencing of peptoid peptidomimetics by Edman degradation.* Tet. Lett., 1998. **39**: p. 3589-3592.

22. Schultz, R.A., Nielsen, T., Zavaleta, J.R., Ruch, R., Wyatt, R., and Garner, H.R., *Hyperspectral imaging: A novel approach for microscopic analysis.* cytometry, 2001. **43**: p. 239-247.

23. Han, Y., and Kodadek, T., *Peptides selected to bind the Gal80 repressor are potent transcriptional activation domains in yeast.* J. Biol. Chem., 2000. **275**: p. 14979-14984.

24. Reddy, M.M., Bachhawat-Sikder, K., and Kodadek, T., *Transformation of low-affinity lead compounds into high-affinity protein capture agents.* Chem. & Biol., 2004. **11**: p. 1127-1137.

25. Reddy, M.M., and Kodadek, T. *Protein "fingerprinting" in complex protein mixtures with peptoid microarrays.* Proc. Natl. Acad. Sci. USA., 2005. **102**: p12672-12677.

CHAPTER 3: PEPTOID-BASED ACTIVATION DOMAIN--MIMETIC MOLECULES

3.1. Introduction

Synthetic, cell permeable molecules capable of modulating transcription have the potential to serve as powerful tools in biology and medicine. The development of such artificial transcription factors which are functional inside cells has been hampered, in part, by the lack of non-peptidic molecules with activation domain-like properties inside cells. The availability of a general method for isolation of non-peptidic protein ligands[1] opened up the possibility of generating activation domain mimetic peptoid molecules by employing binding assay-based screening of peptoid libraries against appropriate transcription factors. The effectiveness of this strategy has previously been validated with peptides. Thus, peptides that have been selected from phage display or split and pool libraries, based solely on their ability to bind to certain transcription factors (Gal80[2], Gal11[3] and KIX domain of CBP[4]), have been shown to work as activation domains in yeast and / or mammalian cells when tethered to Gal4 DBD. Peptoids, in addition to protease resistance[5], have other favorable properties. For instance, a study by Yu, Liu and Kodadek has suggested that a significant proportion of members in random peptoid libraries were cell permeable[6]. While additional studies are required to see if such a trend is general, these preliminary

results were nevertheless encouraging. Therefore, efforts were focused on developing a peptoid-based activation domain mimetic.

CBP (CREB-binding protein) is a mammalian transcriptional co-activator[7], which acts as a bridge between activators and the transcription machinery and causes induction of activator responsive genes. CBP has also been reported to have intrinsic histone acetyl transferase activity[7], which is responsible for acetylation of specific lysine residues within histone tails and loosening of the chromatin to make DNA more accessible to various transcription factors. It was initially discovered as the co-activator for the cAMP regulated transcription activator, CREB (cAMP Response Element-Binding protein)[8].

The interaction between CREB and CBP has been well characterized at a molecular level[9]. The domains within CREB and CBP responsible for the interaction are referred as KID and KIX, respectively. The association between the two domains is stimulated by phopshorylation of a serine residue (Ser 133) within the KID domain of CREB by protein kinase A (PKA) in response to external stimuli[8]. The solution structure of the KIX domain in complex with phosphorylated KID (pKID) was solved by Wright and his colleagues using NMR spectroscopy[9]. The data from this study indicated that the pKID exists as an unstructured peptide in the unbound state. However, binding to KIX leads to a major conformational change and folding of pKID into two helices (αA and αB)

oriented perpendicular to each other. The KIX domain, on the other hand, is structured and consists of three helices (α1, α2 and α3).

The well defined nature of the KIX-KID interaction makes the KIX domain of CBP an excellent target for development of synthetic activation domain replacements. The fact that it is a mammalian transcription factor is also desirable since the long term goal of this study is to develop artificial transcription factors which are functional in mammalian cells. Furthermore, as described earlier, the KIX domain has been validated as a bonafide target for generating synthetic activation domains by a study by Montminy and others[4]. Finally, studies by Greenberg and co-workers have revealed that a tight relationship exists between CREB-dependent transcriptional response and binding affinity of the KIX-KID interaction[10]. This result suggests that the activity of the artificial transcription activator could perhaps be "tuned" by adjusting the affinity of the synthetic ligand for the KIX domain.

The identification of non-peptidic activation domain surrogates based on a binding assay also necessitates a suitable assay for assessing the transcriptional activation potential of co-activator binding ligands. For peptides, this is easily accomplished by making a genetic fusion of the newly isolated activation domain with a native DBD and measuring the ability of this construct to activate a reporter gene bearing binding sites for the DBD. This is obviously not possible

with non-peptidic molecules. Therefore, Yu, Liu and Kodadek developed a cell based assay for assessing the cell permeability[6] and activation potential[11] of synthetic molecules, taking advantage of the highly specific interaction between the ligand binding domain of glucocorticoid receptor (GR) and its synthetic ligand dexamethasone.

The glucocorticoid receptor (GR) is a members of the nuclear hormone receptor super-family[12]. The GR is a ligand-inducible transcription factor that modulates the expression of glucocorticoid response genes and serves several important physiological functions[13]. Glucocorticoid hormones, which are the endogenous ligands for GR, have effects on metabolism, the immune system and the central nervous system[14]. Glucocorticoids also induce apoptosis in immature T cells. The ability of the GR to mediate several important biological functions has made it an attractive therapeutic target and several synthetic ligands have been developed for this receptor (Figure 3.2). One such compound, Dexamethasone (Dex), is used widely in the clinic as an immunosuppressive and anti-inflammatory agent, as well as in the treatment of multiple myeloma[15]. Mifepristone (RU-486), a glucocorticoid receptor antagonist is used clinically to induce abortion[16].

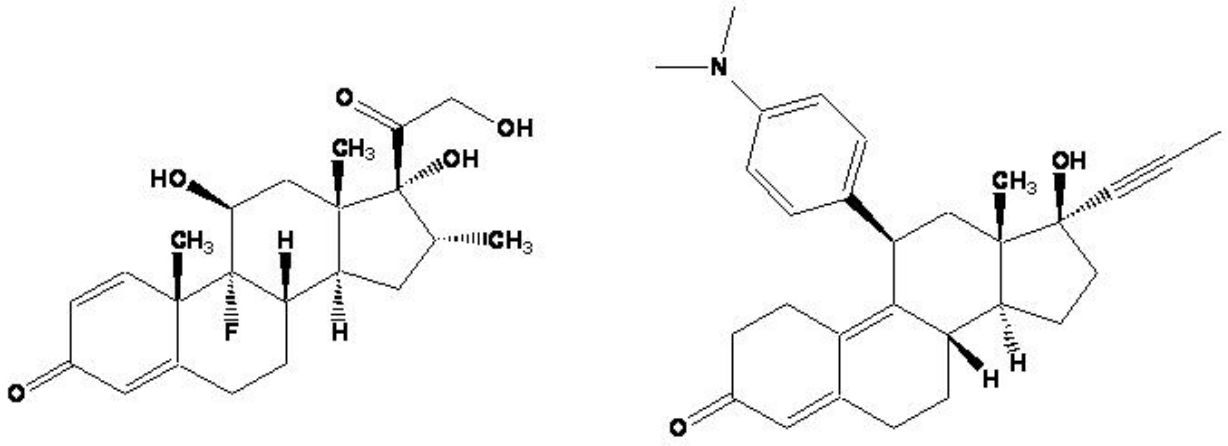


Figure 3.1: Dexamethasone and Mifepristone.

Chemical structures of dexamethasone (left), a potent agonist and mifepristone (right), a potent antagonist of the glucocorticoid receptor.

Similar to other nuclear hormone receptors, GR exhibits a modular structure, comprised of several domains with independent function [17](Figure 3.2). The domains in GR of primary significance are the DNA binding domain (DBD), the hormone binding domain (HBD), and the transactivation domains. The DBD is located in the central region of the receptor and is rich in cysteines and basic residues. It is comprised of two zinc fingers, each of which results from the coordination of four cysteine residues to one zinc atom. The DBD of the receptor is responsible for recognizing and binding to GREs with high affinity. It is believed to be important for the dimerization of the receptor as well. The LBD is located in the carboxy terminal region of the receptor. As the name suggests, this region mediates binding of the ligand to the receptor. The crystal structures of the LBD of several nuclear hormone receptors have been solved, both in the unliganded and ligand-bound states[12]. These structures revealed several major conformational changes between the two states. Deletion of the LBD of GR has been shown to result in a constitutively active form of the receptor, whose activity is comparable to the wild-type receptor and is independent of the ligand. The GR has two transactivation domains, TF1 and TF2, both of which are capable of functioning independently[18]. They are also capable of activating transcription, when fused to unrelated DBDs like the GAL4-DBD. The transactivation domains are believed to recruit transcription machinery to proximal promoter regions and facilitate transcription of target genes.

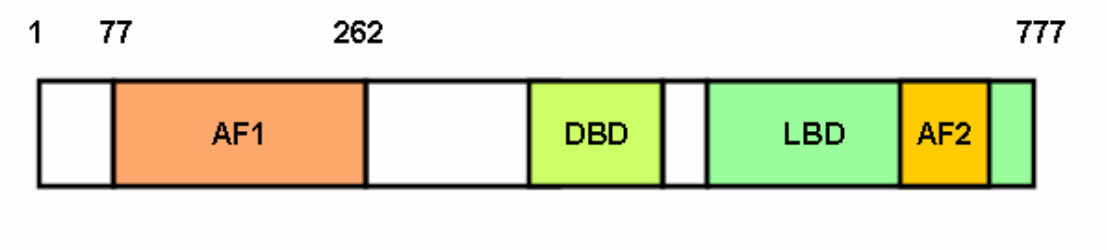


Figure 3.2: Domain architecture of the glucocorticoid receptor.

A schematic illustrating the modular nature of the glucocorticoid receptor consisting of distinct and separable domains (AF1 and AF2: Trans-activation Domains; DBD: DNA Binding-Domain; LBD: Ligand Binding Domain).

(Adapted from reference18)

In the absence of hormone, GR remains in the cytoplasm in an inactive form, sequestered to heat shock proteins (hsp), most notably hsp90[19]. Binding of the hormone to GR results in a conformational change, which results in the dissociation of the heat shock proteins from GR, receptor dimerization and translocation to nucleus, where it binds to specific DNA sequences with the consensus sequence AGAACAnnnTGTTCT (where n is any nucleotide) named glucocorticoid response elements (GRE) and causes activation or inhibition of hormone sensitive genes[13].

Yu, Liu and Kodadek took advantage of this inducible nature of GR response and devised a cell permeability assay[6], which measures the relative cell permeabilities of dexamethasone-conjugated synthetic molecules (Figure 3.3). This assay employs mammalian cells transfected with three plasmids. One directs expression of the Gal4 DNA-binding domain (Gal4 DBD) fused to the glucocorticoid receptor ligand binding domain (GRLBD) and the VP16 activation domain (VP16; residues 413-490). The other two carry distinguishable luciferase reporter genes, one of which is Gal4-responsive and one which is not (transfection control). In the absence of a steroid agonist, the Gal4 DBD-GR LBD-VP16 fusion protein is trapped in the cytoplasm by interaction with Hsp90[19] and thus cannot activate transcription of the Gal4-responsive luciferase gene. When the steroid is added, it passes through the cell membrane, binds to the LBD and frees it from the heat shock protein, resulting in nuclear translocation of the activator

and high-level luciferase transcription. Thus, this system can be used to compare quantitatively the permeability of various steroid conjugates to a steroid-only control molecule.

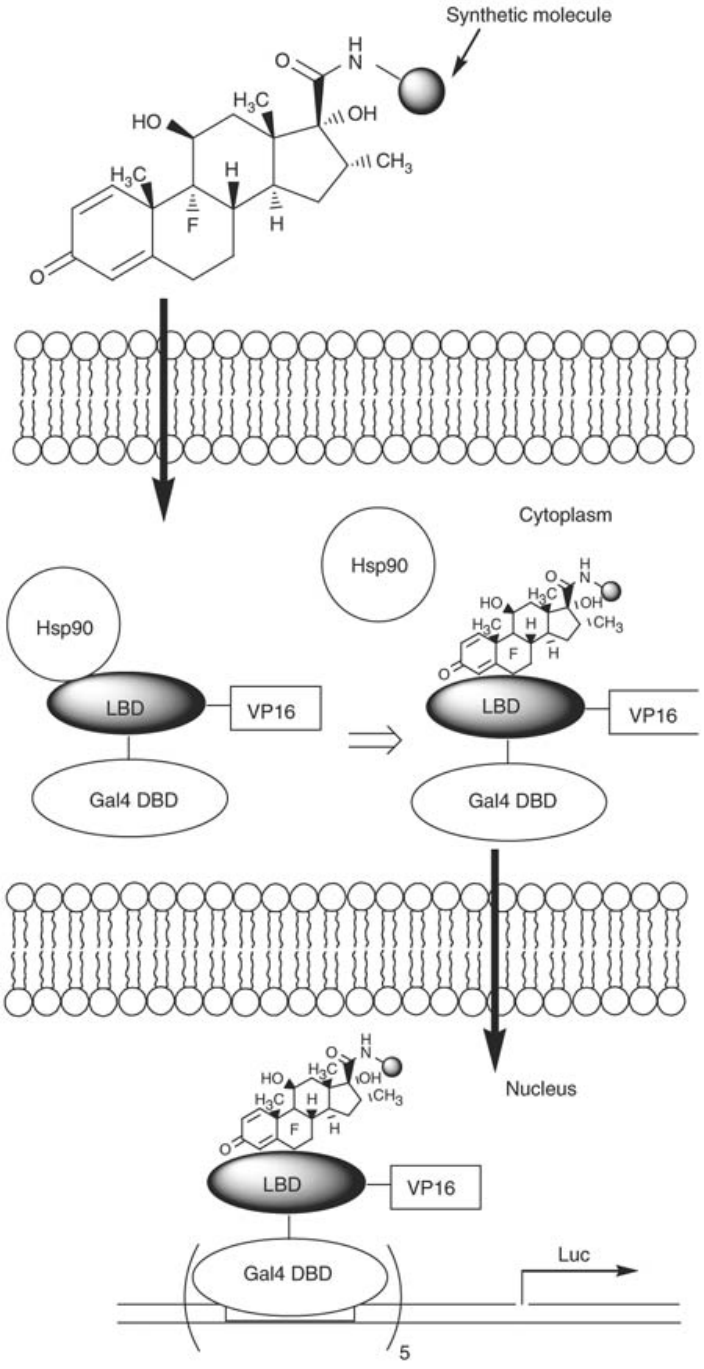


Figure 3.3: Assay for evaluating the relative cell permeability of steroid conjugates. (Yu, Liu and Kodadek 2005)

See text for details.

The assay to measure transcription activation potential[11] is also a derivative of the methodology used to monitor cell permeability (Figure 3.4). In this case however, a plasmid encoding Gal4DBD-GRLBD instead of Gal4DBD-GRLBD-VP16 is transfected into the cells. Since this construct lacks the potent VP16 activation domain, it does not activate the Gal4-responsive luciferase reporter gene. However, the steroid-GRLBD association will allow this protein to move into the nucleus and bind the reporter gene promoter. This would result in the delivery of the putative synthetic activation domain to the target promoter. If the test molecule can function as artificial activation domain, then one should observe significant induction of luciferase expression.

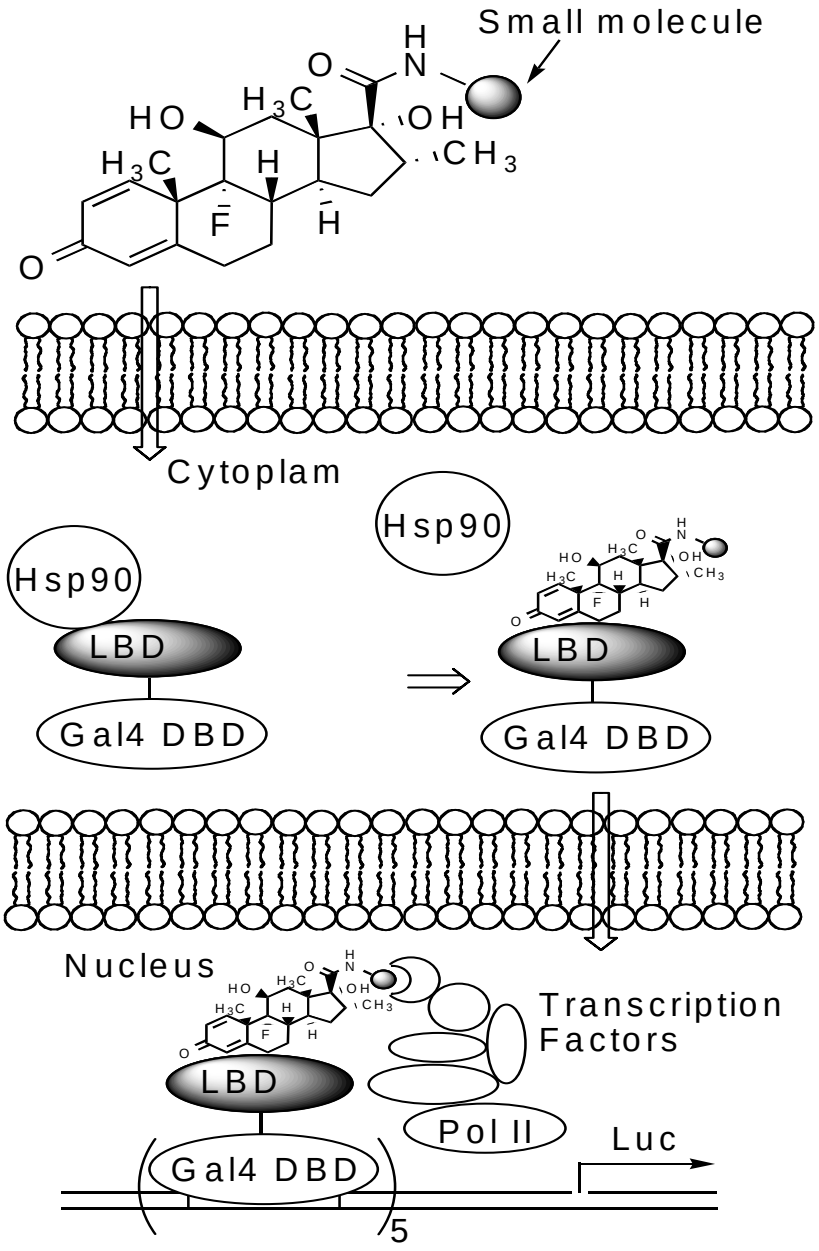


Figure 3.4: Assay for evaluating the transactivation potential of synthetic molecules in the context of steroid conjugation.

See text for details.

In this study, a cell permeable, non-peptidic molecule with apparent transcriptional activity in living cells was discovered[11]. This molecule was obtained by screening a combinatorial library of peptoids for ligands to the KIX domain of CBP. As discussed below, three hexameric KIX domain-binding peptoids were derived from screening a library of approximately 100,000 compounds. One of these molecules proved to have the noted activity. When delivered to the promoter of a reporter gene non-covalently as a steroid conjugate, this molecule supported a nearly 1000-fold increase in transcription over the basal level in mammalian cells.

3.2. Materials and methods

3.2.1. Chemicals

All reagents and solvents were obtained from commercial sources. TentaGel macrobeads (140-170 micron diameter, 0.51 mmole/g capacity) were from Rapp Polymere. Rink Amide MBHA resin (0.69 mmole/g capacity), Fmoc-D-Serine and HBTU and Fmoc-Glycine were from NOVAbiochem. All the reagents employed in peptoid synthesis (amines, bromoacetic acid and 1,3-diisopropylcarbodiimide), with the exception of 1,4-diaminobutane and O-tert-butyl ethanolamine, were from Aldrich. 1,4-diaminobutane was purchased from

Acros, while O-tert-butyl ethanolamine was obtained from CSPS Pharmaceuticals.

3.2.2. Peptoid library synthesis

2 g of TentaGel beads was swollen in DMF for 30 minutes. A D-serine and glycine amino acid spacer was appended to the beads using standard Fluorenylmethoxy-carbonyl (Fmoc) chemistry on a Rainin 12-channel peptide synthesizer (Protein Technologies). Briefly, the resin was divided into seven equal portions, the Fmoc-group removed with 20% piperidine in DMF for 10 minutes (2x 2.5 mL) and the beads thoroughly washed with DMF (8 x 5 mL). The beads were then incubated with 2.5 mL of 0.4 M amino acid and 2.5 mL of 0.4 M HBTU in NMM/DMF for 60 minutes. The beads were thoroughly washed with DMF (8 x 5 mL) and the same protocol repeated for coupling the second amino acid. The Fmoc-group on the second amino acid was removed as described previously and the beads were prepared for library synthesis by thoroughly washing with DMF (8 x 5 mL). The acylation step was carried out in each reaction vessel for 1 hour at 37 °C using 1.5 mL each of 2.8 M bromoacetic acid and 3.2 M diisopropylcarbodiimide. After washing the beads thoroughly with DMF (8 × 3 mL), a 2 M solution of amine (see Figure 3) was added and the reaction vessels were shaken at 37 °C for 90 minutes. All the solutions of the amines were prepared in DMF, except 4-(2-aminoethyl) benzenesulfonamide,

which was prepared in dimethylsulfoxide (DMSO). The beads from all the seven vessels were pooled, randomized by bubbling argon and split again. The protocol was repeated until the desired length was achieved. At the end of the synthesis, the protective groups were removed by treating the beads with a 5mL solution of 95% trifluoroacetic acid (TFA), 2.5% water and 2.5% anisole. The beads were then washed thoroughly with dichloromethane, followed by DMF, neutralized with 20% DIPEA in DMF, rinsed with dichloromethane again, and stored at 4 °C until further use. Four beads were randomly picked from the library and sequenced by Edman degradation using a previously reported protocol[1].

3.2.3. Syntheses of fluorescein-labeled peptoids

Peptoids KBPo1 and KBPo2 (see figure 4 for sequences) were synthesized on Rink AM amide resins using a microwave assisted protocol[20]. In each synthesis, 36 mg of resins (25 μmole scale) was used. The acylation step was carried out as previously described in 1.0 mL each of 2M bromoacetic acid and 3.2M diisoprpylcarbodiimide in DMF for 30 seconds. Amine displacement was done in 2 mL of 2M amine solution in DMF or DMSO. When the peptoid synthesis was completed, the beads were shaken in 1.25 mL each of 0.2M 5(6)-carboxyfluorescein in DMF and 0.2M HBTU in NMM/DMF for 60 minutes at room temperature. The beads were washed thoroughly with DMF, dichloromethane, and cleaved in 5 mL of cleaving cocktail (95 % TFA, 2.5 %

water, and 2.5 % anisole) for 2.5 hr. TFA was removed by blowing argon and the products were purified to homogeneity by reverse phase HPLC using a C18analytical column.

3.2.4. Synthesis of OxDex-AEEA-KBPo1 and OxDex-AEEA-KBPo2

KBPo1 and KBPo2 were synthesized on Rink AM amide resins at 25 μmole scale. *N*-(9-Fluorenylmethoxy-carbonyl)-acetyl-ethyleneglycol-ethyl-amine (Fmoc-AEEA-OH) and OxDex-COOH were added to the peptoids using standard Fmoc chemistry. For each coupling step, equal amount (100 μmole) of Fmoc-AEEA-OH (or OxDex-COOH), HBTU, HoBt, 32 μL of DIPEA, and 14 μl of 2,6-lutidine in 2 mL of anhydrous DMF were applied to the beads. Each coupling step lasted 45 min at room temperature. The products were cleaved off the beads and purified as described above.

3.2.5. Demethylation of mifepristone

A previously reported literature procedure was adapted for this reaction[21]. To a solution of 1g of mifepristone in 5mL of anhydrous methylene chloride was added a filtered solution of 2.6 g of NMO in 5 mL of anhydrous methylene chloride. The resultant solution was stirred under argon in an ice bath and a solution of 80 mg of TPAP in 1 mL was added drop wise over 10 minutes. The solution was allowed to warm back to room temperature and stirred

overnight. The reaction was quenched with 10 mL of 10% sodium bisulfite and stirred for another 10 minutes. The product was then partitioned between ethyl acetate and water. The combined organic layers were washed with pH 7 phosphate buffer and brine, filtered through celite and concentrated in vacuo. The crude formamide thus obtained was dissolved in 12.5 mL of methanol and 10 mL of 10% aq. HCl was added slowly and the reaction mixture was stirred at room temperature for 48 h. The pH of the reaction mixture was adjusted to 7 by addition of 10% sodium carbonate solution and resultant precipitate was collected by filtration. The product thus obtained was dissolved in methylene chloride, dried over sodium sulfate and concentrated in vacuo.

3.2.6. Synthesis of Mif-AEEA-KBPo2

KBPo2 was synthesized on Rink amide MHBA resin using microwave assisted protocol as described earlier[20]. Fmoc-AEEA spacer was then appended to the amino terminus under standard Fmoc chemistry conditions (i.e., 0.2 M solution of the spacer and 0.2 M solution of HBTU in NMM/DMF and mixed for 1 h). The Fmoc group was removed with 20% piperidine in DMF (10 min. X 2), beads thoroughly washed and treated with 1 mL of 2M bromoacetic acid and 1 mL of 3.2 M DIC and subjected to microwave irradiation (2x 15 s with manual agitation for 30 s in between). The beads were then thoroughly washed with DMF and treated with 0.5 mL of 0.5 M demethyl mifepristone. The reaction vessel was

placed in an incubator shaker for 12 h at 225 rpm. The beads were then thoroughly washed with DMF, followed by DCM and the beads were dried. The dried beads were then cleaved with 95% TFA/2.5% triisopropylsilane and 2.5% water for 2 h. The cleavage mixture was concentrated and the crude product was purified by HPLC. Fractions containing the desired product were lyophilized and stored at -20^{o}C until further use.

3.2.7. Plasmids

E. coli glutathione-S-transferase (GST) expression vector pGEX-2T was purchased from Pharmacia Biotech. *E. coli* expression vector pGEX-2T-CBP(378-817), which encodes the GST-KIX fusion protein was from Kodadek lab stock. The mammalian cell expression plasmid encoding the Gal4(1-147)-hGR(499-777)-VP16(413-490) fusion, pEDBD-hGRLBD-VP16 was from Kodadek lab stock. Luciferase reporter pG5B was a kind gift from Dr. Marc R. Montminy (Salk Institute). *Renilla* luciferase expression vector pRL-SV40 was from Promega.

The mammalian cell expression plasmid encoding the Gal4(1-147)-hGR(499-777) fusion was constructed by replacing the reading frame of plasmid pEDBD-hGRLBD-VP16 with Gal4(1-147)-hGR(499-777) fusion oligonucleotide. pEDBD-hGRLBD-VP16 was used as a template in a polymerase chain reaction (PCR) with primers PYP007 (5'-CGT CAG ATC CGC TAG CAT GAA GC-3')

and PYP008 (5'-CCT GGC GCG GCC GCT CAC TTT TGA TGA AAC AGA AG-3'). The PCR product was purified using Qiaquick PCR purification kit (Qiagen) and was digested with *NheI* and *NotI*. The oligonucleotide was inserted in the *NheI/NotI* cleaved pEDBD-hGRLBD-VP16 vector. The final plasmid was named pEDBD-hGRLBD.

The Quickchange protocol (Strategene) was followed to make a site-directed mutation in hGRLBD that translated to an E755A amino acid substitution. In the PCR reaction, plasmid pEDBD-hGRLBD was used as the template. The two mutagenic primers were PYP009 (5'-CCC GAG ATG TTA GCT GCA ATC ATC ACC AAT CAG-3') and PYP010 (5'-CTG ATT GGT GAT GAT TGC AGC TAA CAT CTC GGG-3'). The final plasmid was named pEDBD-hGRLBD (E755A).

3.2.8. Proteins

GST and GST-KIX proteins were prepared as described[1]. Texas Red labeled GST-KIX protein was prepared following the same procedure published before[1]. Bovine serum albumin (BSA) was from PIERCE.

3.2.9. Selection of peptoids against KIX domain of mouse CREB-binding protein (CBP)

Approximately 100,000 beads (190 mg) from the peptoid library were swollen in TBST (50 mM Tris, pH 7.4, 150 mM NaCl, 0.1% Tween 20) for 1 hr. The buffer was removed and the beads were soaked in a prepared *E. coli* lysate at room temperature for 2 hr. The *E. coli* lysate was removed and beads were washed with PBS buffer. The beads were then incubated in 5 mL TBST buffer containing 500 nM Texas Red labeled GST-KIX protein, 25 μM GST protein, 0.35 mM (total protein concentration) *E. coli* cell lysate, and 0.15 mM BSA at room temperature for 1 hr. The beads were thoroughly washed with TBST and visualized under a fluorescence microscope. Bright fluorescent beads were picked manually with a micropipette. Each bead was heated in 1% SDS at 95 °C for 5 min and washed thoroughly with PBS buffer. The sequences of peptoids were solved by single-bead Edman degradation following a previously published protocol.

3.2.10. Fluorescence polarization assays

In 200 μL of PBS buffer, the indicated amount of proteins (GST-KIX, GST, or BSA) and approximately 10 nM fluorescein-labeled peptoids were mixed at room temperature for 20 min. The fluorescence polarization values of samples were measured on a Beacon 2000 fluorescence spectrometer.

3.2.11. Glucocorticoid receptor binding competition assay

The binding affinities of OxDex derivatives for glucocorticoid receptor were determined using GR Competitor Assay Kit from Invitrogen. The complete procedure was described previously[6].

3.2.12. Cell line

HeLa cells were obtained from the American Type Culture Collection, (CCL-2).

3.2.13. Tissue culture, transfection, and luciferase assays

Cell maintenance, transfection and luciferase assays were conducted following procedures described before[6]. OxDex-AEEA-KBPo1, OxDex-AEEA-KBPo2, OxDex, and Dex samples at various concentrations were prepared in DMSO. After transfection, samples were added directly in the tissue culture media. The DMSO concentration in the media was kept under 2 %.

3.3. Results

3.3.1. Isolation of KIX-binding peptoids

Initial efforts to isolate a KIX ligand were focused on screening the large ~ 500, 000 peptoid library described in the previous chapter[1]. In spite of the excellent quality of the library, screening of this library was complicated by the presence auto-fluorescent beads in the library, resulting in the isolation of several false positives. Although, pre-sorting of the beads with the bead sorter diminished this problem considerably, it was not completely eliminated. Therefore, some of these monomers suspected to contribute to auto-fluorescent properties were discarded and new library of hexameric peptoids with a theoretical diversity of 117,649 compounds was constructed by split and pool synthesis[22] on 2 g of TentaGel Macrobeads (Rappe Polymere). A constant D-Serine-Glycine dipeptide sequence was introduced as a spacer prior to library synthesis. Seven amines were employed in the library synthesis, which employed the "sub-monomer" route for peptoid synthesis[23] (Figure 3.5). The primary hydroxyl group in ethanolamine was protected as t-butyl ether, while the secondary amine in tryptamine and one of the primary amines in 1,4-diaminobutane were protected with t-Boc groups[24, 25]. Upon completion of the split and pool synthesis, these acid-labile protecting groups were removed without disturbing the linkage of the peptoids to the beads, which was not acid-sensitive.

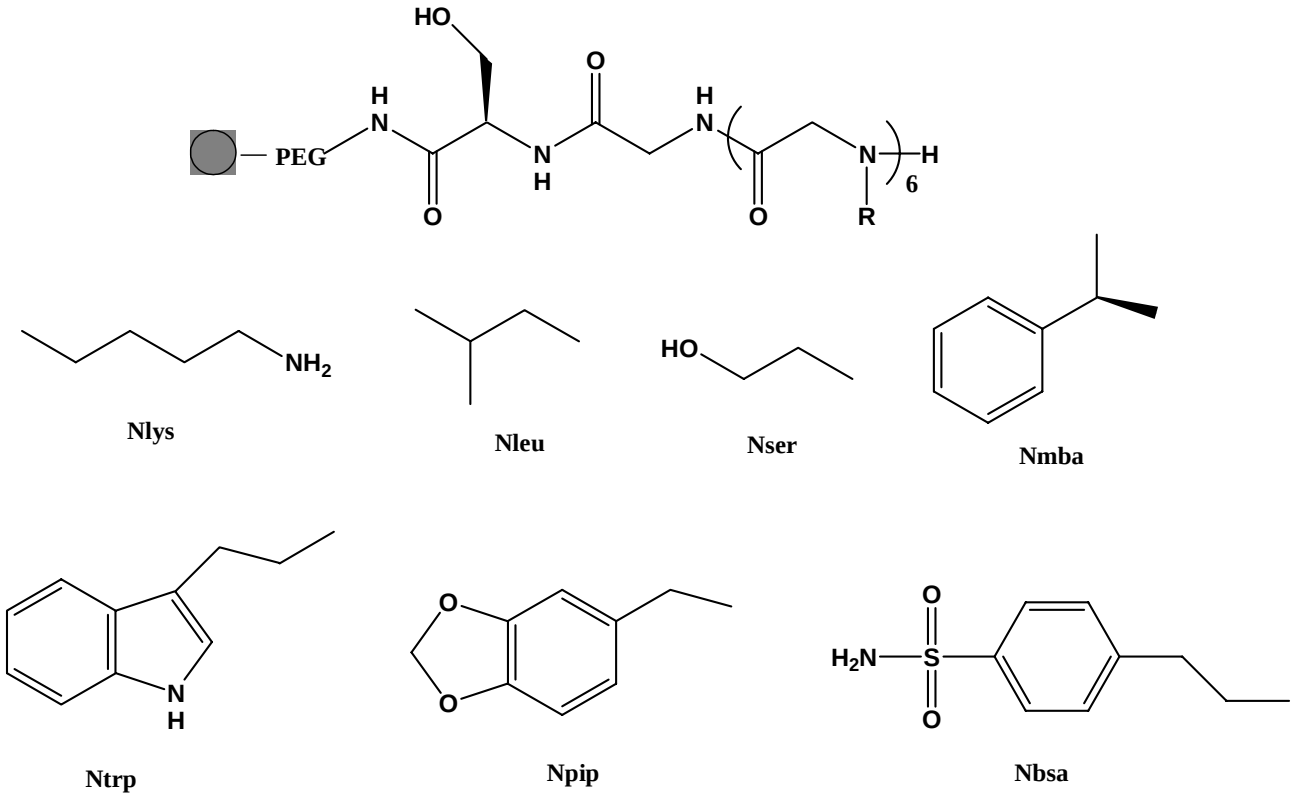
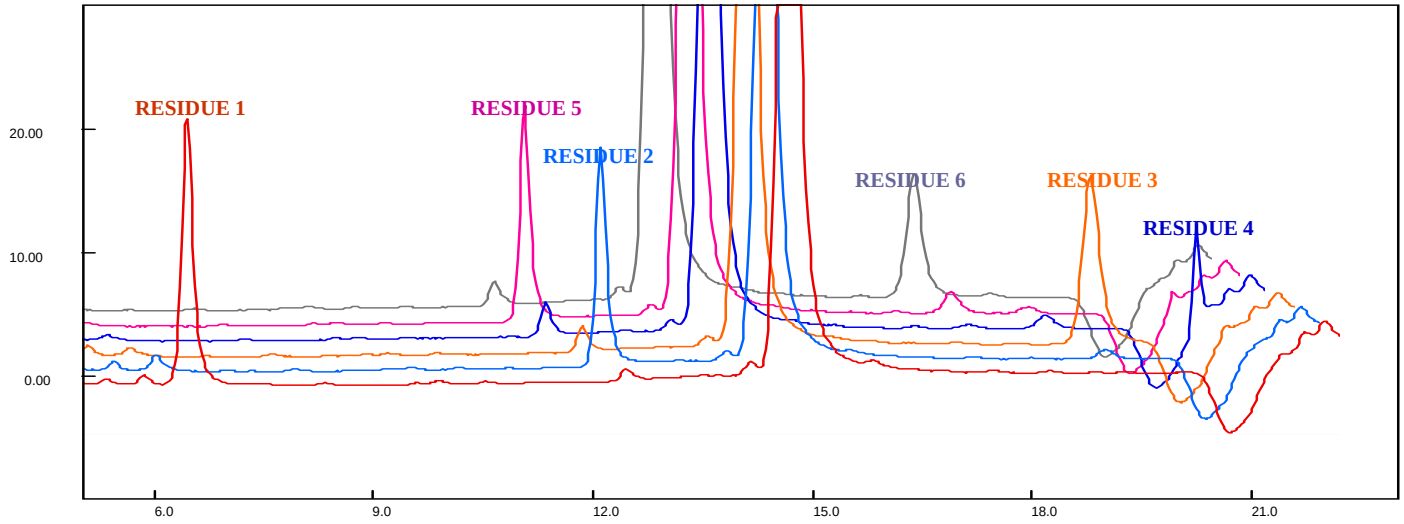


Figure 3.5: Library design.

The general structure of the library employed in this study. The chemical structures of the side chains are shown at the bottom of the figure.

To assess the quality of the library, four beads were picked randomly and subjected to Edman degradation using a previously reported protocol[1]. A complete sequence was obtained in each case and all seven amines employed in the synthesis were represented in at least one of the sequences (Figure 3.6). Based on these data, it was concluded that the library was of suitable quality to undertake screening experiments.



Random 1: Nser-Nbsa-Nleu-Ntrp-Nbsa-Nlys
Random 2: Npip-Nleu-Nleu-Nmba-Nmba-Nleu
Random 3: Nleu-Nmba-Nlys-Nlys-Nbsa-Nbsa
Random 4: Nser-Nmba-Nleu-Nlys-Nbsa-Nser

Figure 3.6: Library characterization by Edman sequencing.

Text sequences of four beads randomly picked from the library (bottom), along with the representative Edman trace of one of the beads (top). The four sequences, taken together, have all the monomers used in the construction of the library.

A recombinant, purified Glutathione-S-transferase (GST) fusion protein containing a fragment of murine CBP (residues 378-817) was labeled with Texas Red-succinimidyl ester under conditions that resulted in each protein molecule acquiring, on average, one dye molecule. Excess fluor was removed using a desalting column. Approximately 100, 000 library beads were then pre-sorted using the automated bead sorter to remove any auto-fluorescent beads, pre-swollen in TBST buffer, and blocked with *E. coli* lysate to saturate any non-specific binding sites on the bead surface. The labeled protein was then incubated with the bead-displayed peptoid library. The incubation solution contained, in addition to the Texas Red-labeled GST-KIX domain fusion protein (500 nM), a 700-fold excess of proteins present in an *E. coli* lysate and a 300-fold excess of bovine serum albumin (BSA) as competitor proteins. Previous studies have indicated that the presence of a heterogeneous mixture of competitor proteins greatly reduces the frequency with which "sticky", relatively non-specific ligands are obtained[1]. Finally, a 50-fold excess of unlabeled GST protein was also included in the solution to minimize the possibility of isolating hits against the GST part of the fusion protein. After incubation for one hour at room temperature, the beads were washed thoroughly with TBST to remove any protein not bound tightly. The entire collection of beads was then added to a Petri dish and examined under a fluorescence microscope. While a number of beads exhibited fluorescence above background levels, three beads that appeared exceptionally

bright compared to the rest of the library were collected using a micropipette. Figure 3.7 shows a micrograph that illustrates the high contrast between one of these beads and a number of others scored as “negatives” in the same field. The three brightest beads were treated with 1% SDS to remove bound proteins and the identity of each of the three peptoid “hits” was determined by single-bead Edman degradation (Figure 3.8).

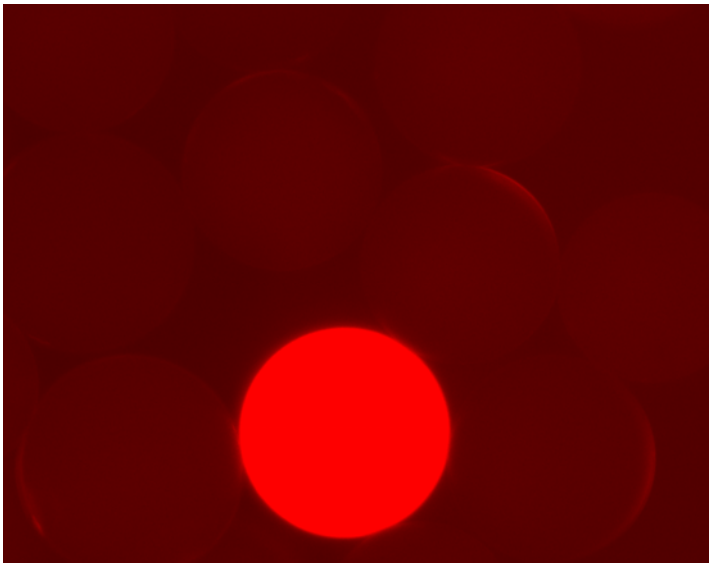
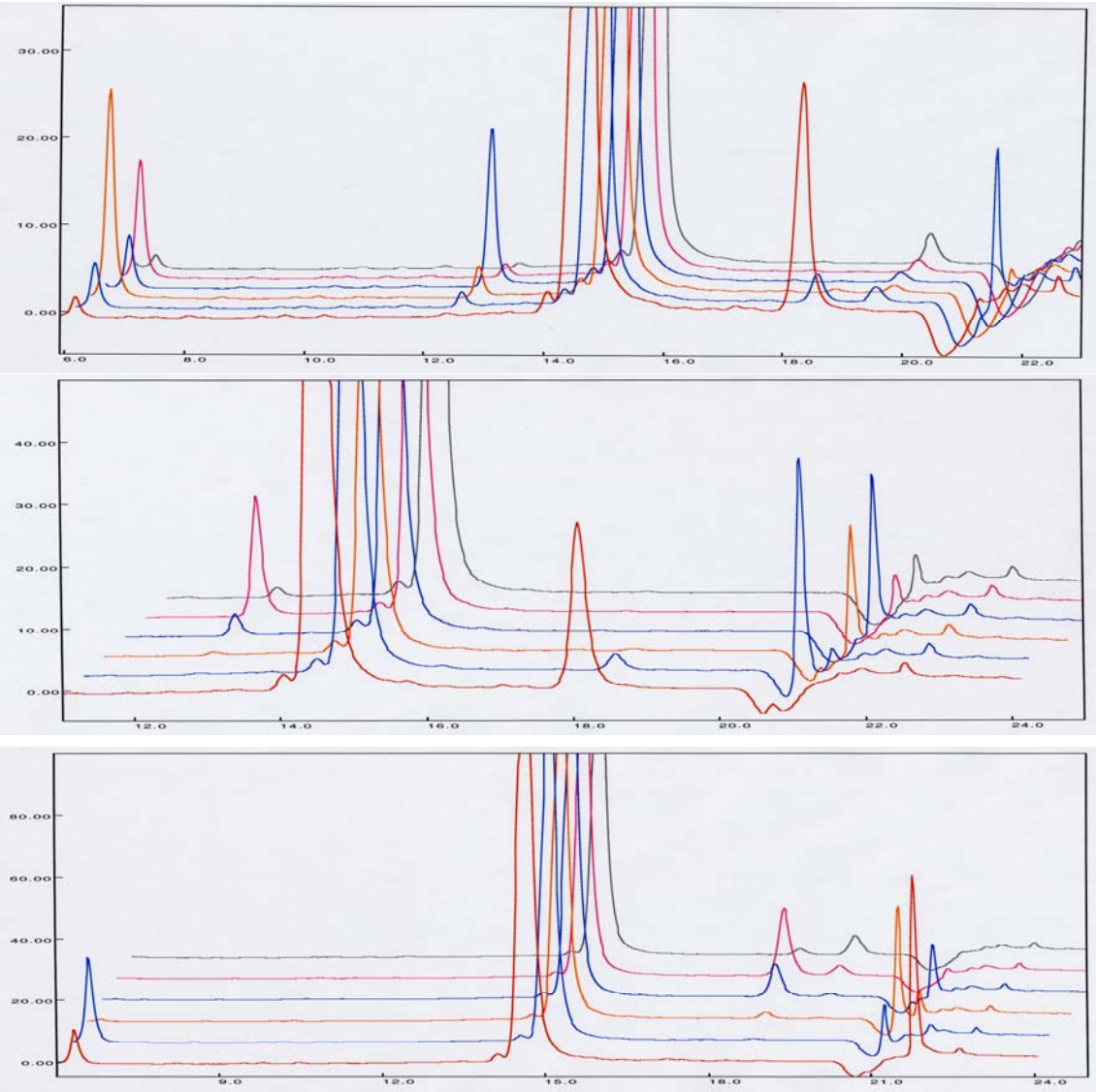


Figure 3.7: Isolation of GST-KIX fusion protein-binding peptoids.

Fluorescent micrograph showing a field including a bead scored as a "hit" and several scored as "negatives" after incubation of the bead library with Texas Red-labeled GST-KIX.



Hit#1 **Nlys-Npip-Ntrp-Ntrp-Nbsa-Ntrp**
Hit#2 **Nlys-Ntrp-Nser-Nbsa-Nser-Nleu**
Hit#3 **Nmba-Nser-Npip-Ntrp-Nlys-Nleu**

Figure 3.8: Identification of GST-KIX-binding peptoids by Edman degradation.

Edman sequencing traces of three "hits" isolated in the screen. The corresponding text sequences are shown at the bottom.

To validate the results of the bead screening effort in a solution binding experiment, two of the KIX domain-binding peptoids (named as KBPo1 and KBPo2) were re-synthesized on Rink Amide MBHA resin and capped at their N-terminus with 5(6)-carboxyfluorescein. The fluorescein-labeled peptoids (named as F-KBPo1 and F-KBPo2) were then cleaved from the resin, purified by HPLC and their binding affinity for the GST-KIX domain fusion protein was determined in a titration experiment monitored by fluorescence polarization spectroscopy[26]. The data (Figure 3.9) show that the fluoresceinated KBPo1 and KBPo2 derivatives bind to GST-KIX domain protein with equilibrium dissociation constants (K_Ds), of approximately 5 μM. These titrations were then repeated using GST or BSA in place of GST-KIX to assess the specificity of binding. The fluoresceinated KBPo2 molecule had little affinity for these proteins, as expected for a high specificity KIX ligand. However, this was not the case for fluoresceinated KBPo1, which bound the non-specific proteins as well as, if not better than, the KIX fusion protein. This result was surprising, given the stringent screening conditions employed, which included a buffer containing excess unlabeled GST and BSA as competitor proteins. Nonetheless, it was concluded that that KBPo1 is a relatively "sticky" molecule that binds promiscuously to many proteins. This may be due to the presence of aromatic, hydrophobic residues at five of the six variable positions in this molecule, whereas KBPo2 is a more polar molecule (Fig. 3.10).

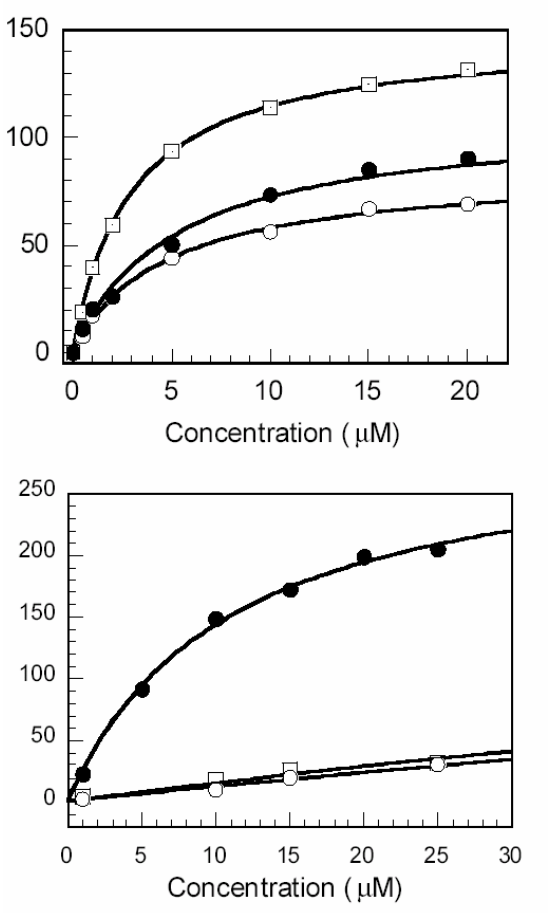


Figure 3.9: Binding affinities and specificities of the KIX-binding peptoids. Fluorescein labeled- KBPo1 (top panel) or KBPo2 (bottom panel) was titrated with the indicated concentration of GST-KIX (●), GST (○) or BSA (□).

3.3.2. Steroid conjugates of the KIX-binding peptoids are cell permeable

In order to function as part of a pharmacologic agent for specific gene activation, the KIX-binding peptoids obviously must be cell permeable. Preliminary microscopy experiments using the fluorescein-conjugated molecules indicated that this is the case (data not shown). However, detailed permeability studies were carried out in the context of peptoid-steroid conjugates employing the cell based permeability assay[6] described earlier (Figure 3.3). The GR agonist dexamethasone was oxidized with periodic acid[27] to yield OxDex (Fig. 3.9), which was then conjugated to KBPo1 and KBPo2 via an ethylene glycol linker (AEEA, see Figure 3.10). A control molecule which includes the OxDex and the linker, but lacks a peptoid and terminates in a primary amide (OxDex-AEEA-$CONH_2$, Figure 3.10) was also prepared.

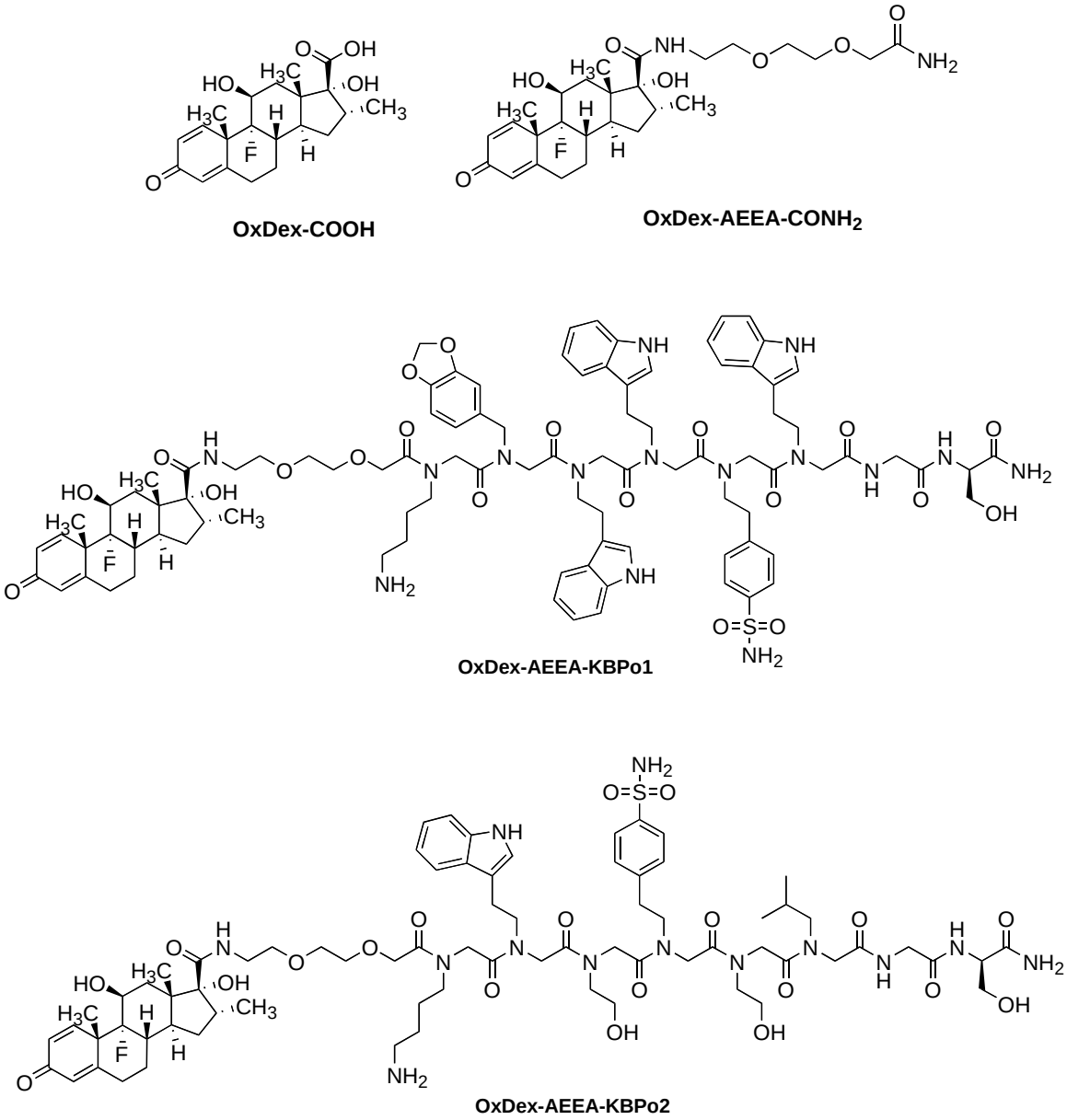


Figure 3.10: Chemical structures of molecules used in cellular assays.

To measure the cell permeability of these conjugates, cells transfected with the three plasmids (Gal4 DBD-GR LBD- VP16, reporter gene and transfection control) were incubated with different concentrations of the OxDex-AEEA-$CONH_2$ control or the OxDex-AEEA-peptoid conjugates and the level of luciferase gene expression was measured. The results are shown in Figure 3.11.

In each case, a dose-dependent induction of Gal4-responsive luciferase expression was observed, indicating that all of the molecules entered the cells. At a concentration of 100 μM, OxDex-AEEA-KBPo1 and OxDex-AEEA-KBPo2 induced 500- and 1,000-fold increases in luciferase expression, respectively. The titration curve revealed EC_{50} values of 37 μM and 8 μM, respectively, for these peptoid conjugates. As a comparison, OxDex-AEEA-$CONH_2$ gave a maximum induction of 950-fold with an EC_{50} value of 48 μM. Thus, it was concluded that both Ox-Dex-AEEA-peptoid conjugates are cell permeable.

To interpret these results more quantitatively, it is also important to measure the affinity of each compound for the GRLBD since this will also affect the level of luciferase induction in cells. This was determined using an in vitro fluorescence polarization assay, in which OxDex derivatives were employed to compete with a commercially available fluorescent GR ligand, Fluormone™ (Invitrogen), for binding to the purified human GR. The IC_{50} values of OxDex-AEEA-$CONH_2$ and OxDex-AEEA-KBPo2 were 1.3 μM and 0.26 μM, respectively (Figure 3.11). Assuming that this five-fold higher affinity of the

peptoid conjugate in vitro reflects the relative affinities for the GRLBD inside cells and given that OxDex-AEEA-KBPo2 was about six-fold more potent than OxDex-AEEA-CONH2 in the cellular assay, it was concluded that the steroid control molecule and the steroid-KBPo2 conjugate are approximately equally cell permeable. This is a remarkable result in that it argues that the peptoid does not compromise the cell permeability of the steroid.

The in vitro binding data obtained in the experiment using the other peptoid conjugate, OxDex-AEEA-KBPo1 could not be fit neatly to a sigmoidal curve. The estimated IC_{50} value of ≈10 μM should thus be taken with a grain of salt. The non-ideal behavior of OxDex-AEEA-KBPo1 in this experiment may reflect aggregation of the molecule or other undesirable events.

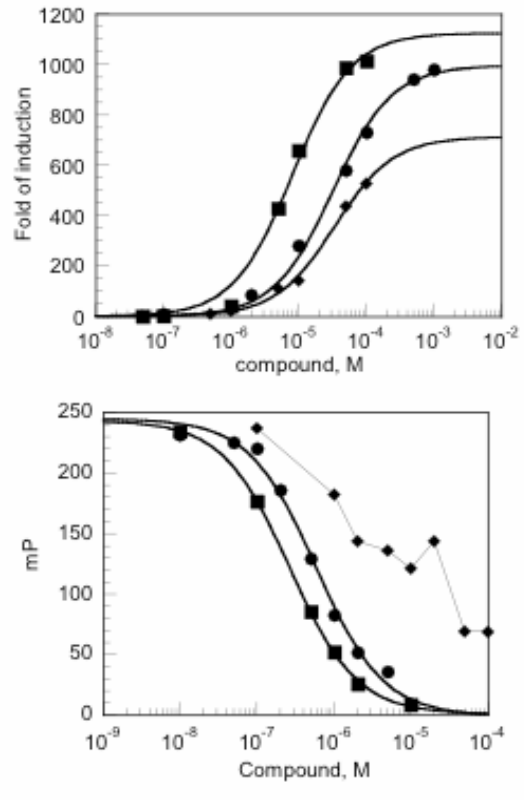


Figure 3.11: Evaluation of cell permeability of peptoids KBPo1 and KBPo2.

This experiment monitors the activation of a Gal4-responsive luciferase reporter gene when OxDex-AEEA conjugates (Figure 3.9) are incubated with HeLa cells that express a Gal4 DBD-GR LBD- VP16 AD fusion protein.

Top Panel: Dose dependence of the OxDex-AEEA-KBPo1 (▪)-, OxDex-AEEA-KBPo2 (♦)-, and OxDex-AEEA-$CONH_2$ (□)- mediated induction of luciferase expression.

Bottom panel: Affinities of the OxDex-AEEA conjugates {OxDex-AEEA-KBPo1 (▪), OxDex-AEEA-KBPo2 (♦), and OxDex-AEEA-$CONH_2$ (□)} for glucocorticoid receptor in vitro as monitored by a competitive fluorescence anisotropy assay. (Figure courtesy of Drs. Bo Liu and Peng Yu).

3.3.3. A KIX domain-binding peptoid activates the transcription of a reporter gene when fused to a steroid conjugate

To determine if the KIX-binding peptoids could function as activation domains in mammalian cells, the transcription activation assay described earlier was employed (Figure 3.4). The results of this experiment are shown in Figure 3.12. As mentioned above, titration of the cells expressing Gal4 DBD-GRLBD with the control steroid OxDex-AEEA-$CONH_2$ did not induce detectable luciferase transcription at any concentration tested. The same result was obtained with OxDex-AEEA-KBPo1. No induction of transcription was detected even at a concentration of 100 μM, even though this is well above the apparent binding constant for this conjugate with the GRLBD in cells inferred from the EC 50 of 37 μM observed in the permeability assay. Taken together, these data suggest that while the OxDex-AEEA-KBPo1 conjugate can enter cells and bind the GRLBD, it does not function as an activation domain. In stark contrast, titration of Gal4-GRLBD-expressing cells with the other peptoid conjugate, OxDex-AEEA-KBPo2 resulted in a dose-dependent increase in the expression of the reporter gene (Figure 3.12). At an OxDex-AEEA-KBPo2 concentration of 100 μM, expression of the Gal4-responsive luciferase reporter gene was induced about 900-fold above the basal level. Half-maximal stimulation was observed at about 10 μM OxDex-AEEA-KBPo2. These data suggest that KBPo2 may be capable of functioning as an activation domain surrogate in living mammalian cells.

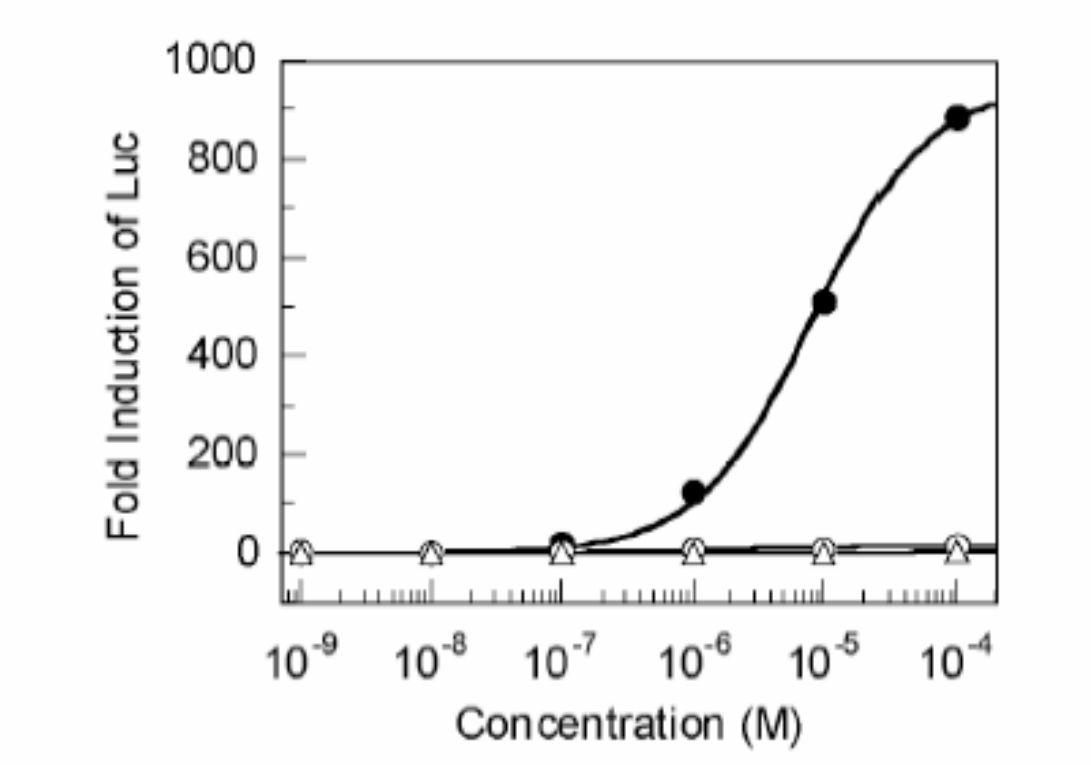


Figure 3.12: The OxDex-AEEA-KBPo2 conjugate supports activated transcription in HeLa cells.

Dose dependence of the level of Gal4-responsive luciferase gene expression when cells expressing the Gal4 DBD-GR LBD fusion protein was incubated with OxDex-AEEA-KBPo1 (○), OxDex-AEEA-KBPo2 (●), or OXDex-AEEA-CONH2 (Δ).

(Figure courtesy of Drs. Bo Liu and Peng Yu).

3.3.4. Limitations of the transcription activity assay

While the assay described for measuring transactivation potential of synthetic molecules offered a convenient route for validating the activity of putative activation domain surrogates, the interpretation of the data from this study was complicated by the observation that dexamethasone itself, in unmodified form, caused activation of transcription (Figure 3.13). This is attributable to the presence of an activation domain (AF2) in close proximity to the GR LBD (Figure 3.2). Conversion of Dexamethasone to OxDex completely abrogated this activity. The molecular basis for the difference in activity between these two molecules is unclear. In any case, conjugation of test molecules to OxDex provided a convenient route to measure activation potential in an environment free from any background reporter gene activity. While unlikely, one cannot rule out, however, the possibility that conjugation of the test molecule to OxDex may lead to activation of transcription via a mechanism similar to that seen with unmodified Dex. In other words, conjugation of peptoids could, in principle, alter the binding conformation of OxDex resulting in it behaving like unmodified dexamethasone. However, conjugation of various molecules, in general, did not lead to restoration of transcription activity of OxDex. Thus, OxDex-AEEA-CONH2 and OxDex-AEEA-KBPo1 were completely inactive in the transcription activation assay, while retaining cell permeability. The transcription activity of OxDex-AEEA-KBPo2, therefore, likely results from the

ability of KBPo2 to function as an activation domain surrogate. However, further experimental evidence would be necessary to verify this scenario.

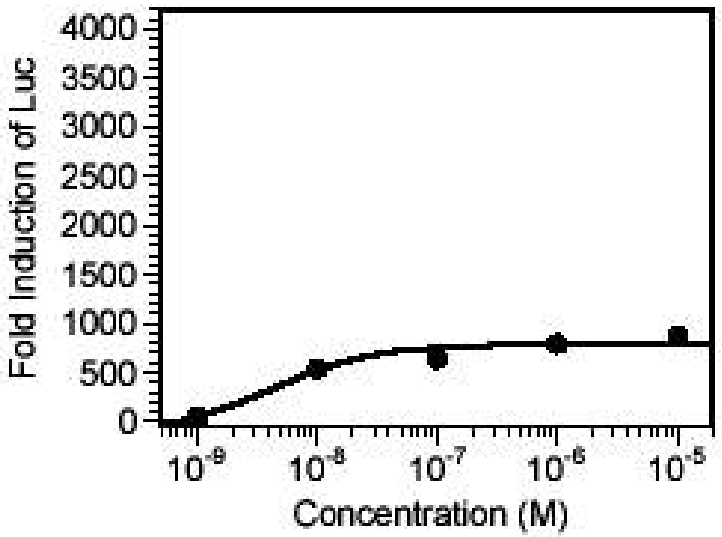


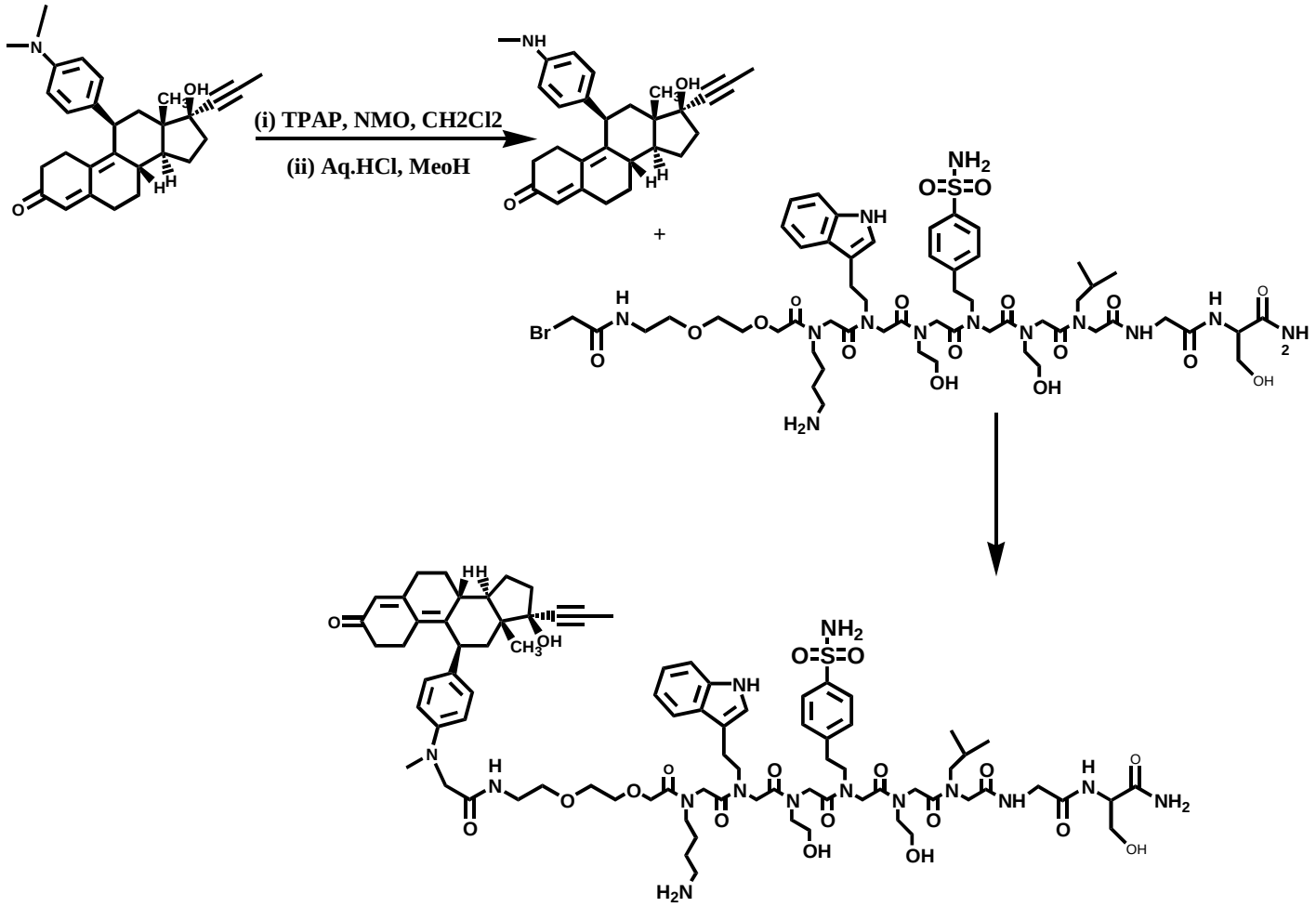
Figure 3.13: Dexamethasone-induced transcription activation.

Dose dependence of the level of Gal4-responsive luciferase gene expression when cells expressing the Gal4 DBD-GR LBD fusion protein were incubated with dexamethasone. (Figure courtesy of Drs. Bo Liu and Peng Yu).

In an attempt to rule out the unlikely scenario described above, a slightly different scheme for measuring transcription activity of KBPo2 was envisioned. In this scheme, OxDex was replaced with Mifepristone, a potent glucocorticoid receptor antagonist[16, 21]. Similar to dexamethasone, mifepristone also binds to GR with high affinity, resulting in the translocation of the receptor-ligand complex to the nucleus and binding of GR to DNA. However, mifepristone, being an antagonist, prevents GR from activating transcription. Thus, it offers an attractive means to target KBPo2 to a promoter without, in itself, inducing transcription.

In order to facilitate coupling to KBPo2, mifepristone was subjected to a demethylation reaction[21], converting the tertiary amino group to a secondary amine. This structural modification has previously been reported in the literature and has been shown not to significantly compromise with the affinity of the ligand to GR. In fact, von Geldern and co-workers used this approach to conjugate mifepristone to bile acids to develop liver-selective glucocorticoid receptor antagonists[21]. Coupling of KBPo2 to the mifepristone derivative via the AEEA linker was accomplished on solid support as shown in scheme 3.1. As a control, a mifepristone-AEEA conjugate without the peptoid was also synthesized. The products were released from the beads by treatment with a cleavage cocktail consisting of 95% TFA / 2.5% water / 2.5% triisopropylsilane, lyophilized and purified to homogeneity by HPLC. The purified product was

characterized by MALDI-TOF mass spectrometry to verify the identity of the desired product (Figure 3.14 and 3.15).



Scheme 3.1: Mif-AEEA-KBPo2 synthesis scheme.

A depiction of the scheme used for the synthesis of the Mif-AEEA-KBPo2.

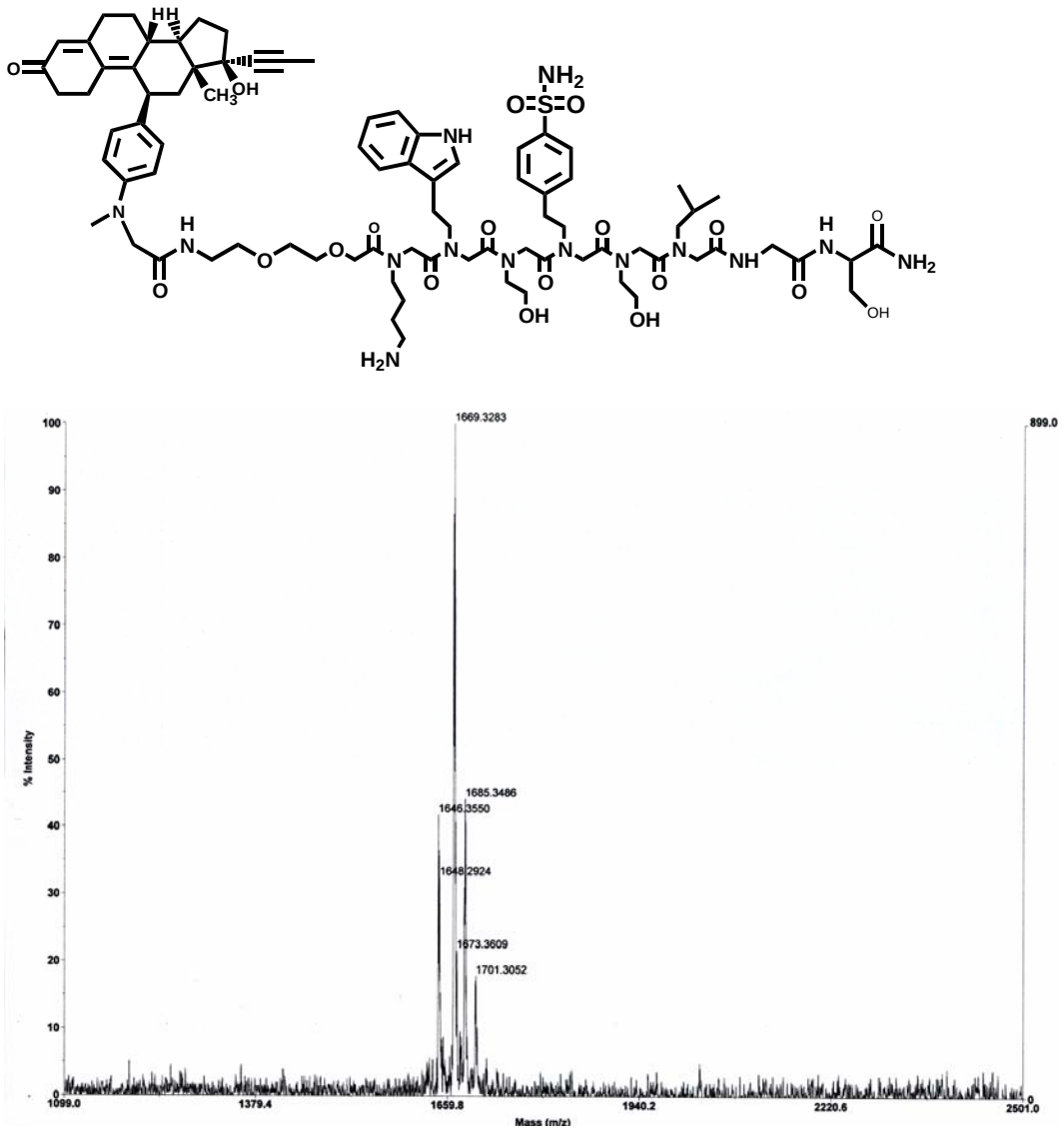


Figure 3.14: Mif-AEEA-KBPo2.

Chemical structure (top panel) and MALDI-TOF mass spectrum (bottom panel) of Mif-AEEA-KBPo2.

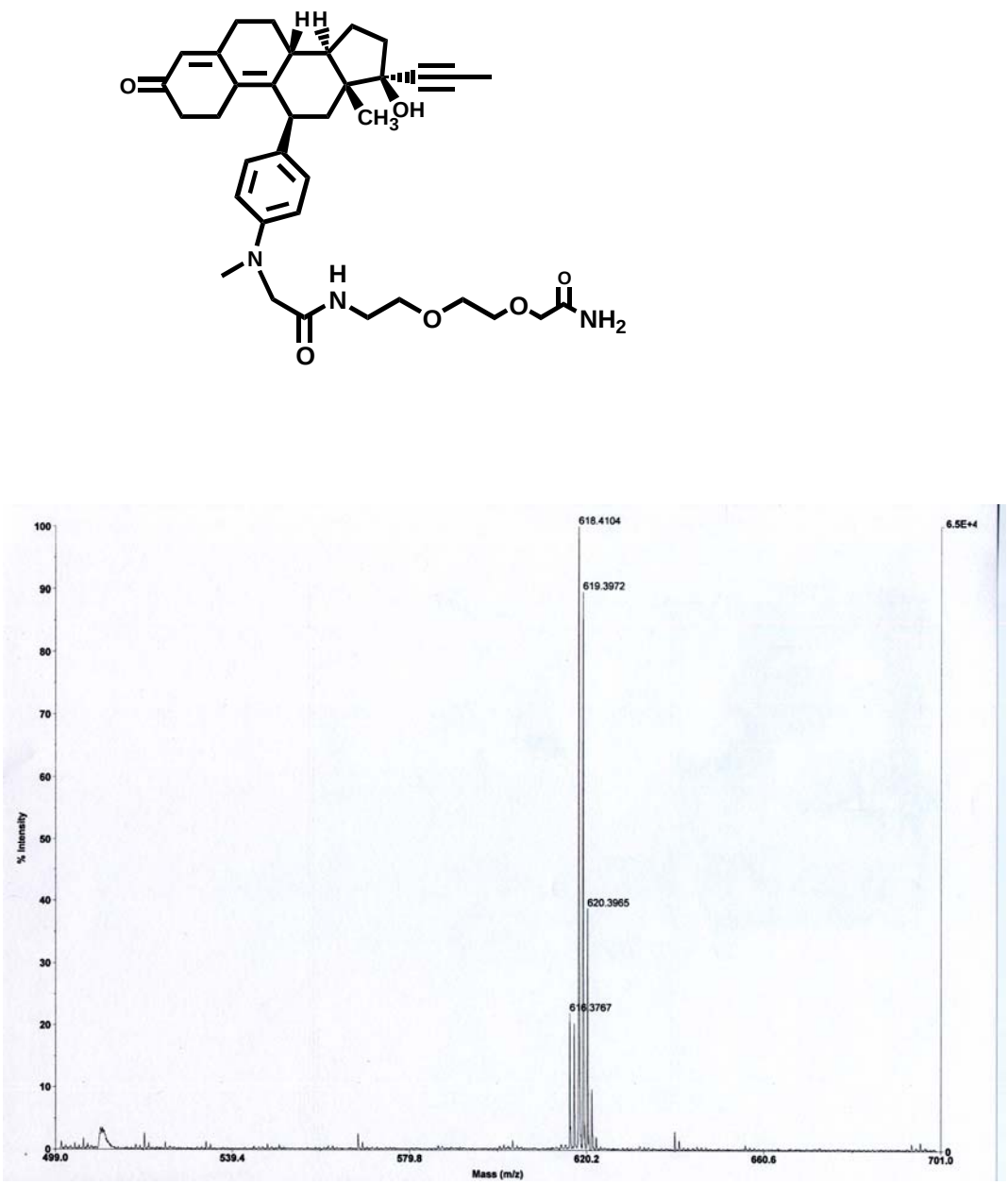


Figure 3.15: Mif-AEEA.

Chemical structure (top panel) and MALDI-TOF (bottom panel) mass spectrum of Mif-AEEA.

The previously described assays for cell permeability and transcriptional activity were then performed on the drug conjugates. The data from these assays are presented in Figure 3.16. As expected, mifepristone was found to be cell permeable as shown by its ability to induce luciferase transcription in the permeability assay. It was found to be largely inactive in the transcription assay, although slight induction of the reporter gene was observed at very high drug concentrations. Unfortunately, and quite surprisingly, the mifepristone-AEEA conjugate and the mifepristone-AEEA-KBPo2 conjugates were found to be cell impermeable. Thus, in the permeability assay, no induction of the luciferase gene transcription was noticeable under any concentrations of the conjugates tested. This precluded measuring of transactivation potential of mifepristone-AEEA-KBPo2 conjugate. Therefore, the transcription activation potential of KBPo2 could not be unequivocally established. Further experiments are necessary to gain insight into the mechanistic aspects of OxDex-AEEA-KBPo2-mediated transcription activation.

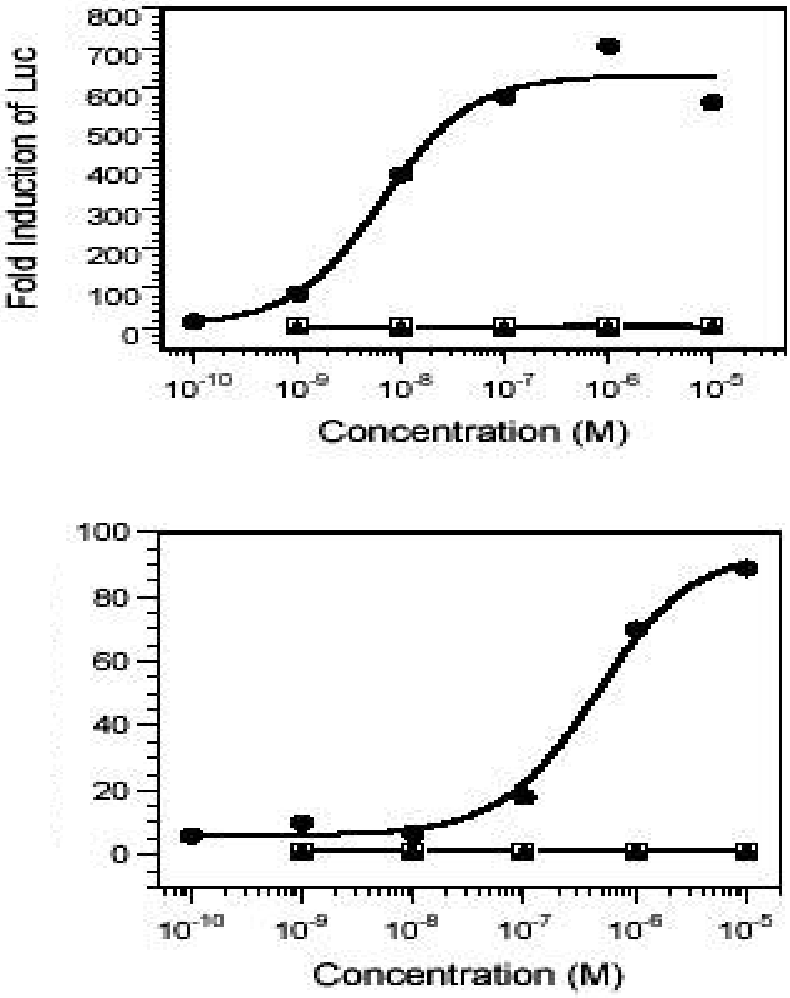


Figure 3.16: Mif-AEEA and Mif-AEEA-KBPo2 are not cell permeable.

Top panel: Dose dependence of the level of Gal4-responsive luciferase gene expression when cells expressing the Gal4 DBD-GR LBD- VP16 fusion protein were incubated with Mifepristone (●), Mif-AEEA (▲) or Mif-AEEA-KBPo2 (□). Bottom panel: Dose dependence of the level of Gal4-responsive luciferase gene expression when cells expressing the Gal4 DBD-GR LBD fusion protein were incubated with Mifepristone (●), Mif-AEEA (▲) or Mif-AEEA-KBPo2 (□). (Figure courtesy of Drs. Bo Liu and Peng Yu).

3.4 Discussion

There is great interest in the development of cell permeable pharmacologic agents capable of activating gene expression in a specific fashion. Essentially all efforts towards this goal aim to replicate two seminal activities of native gene-specific activators, which are to bind near the target gene sequence specifically and to recruit the RNA polymerase II transcription machinery to the core promoter[28, 29]. The latter activity will require the development of cell permeable molecules that bind to accessible surfaces of appropriate proteins such as coactivators, thus allowing them to function as artificial activation domains. Prior to this work, there had been two reports of non-peptidic molecules capable of acting as activation domains when fused covalently or non-covalently to a sequence-specific DNA-binding moiety. Uesugi and co-workers developed a molecule they called wrenchnolol by structure-aided design that binds the coactivator Sur2[30]. When fused to a hairpin polyamide, wrenchnolol supports significant transcriptional activation in vitro in experiments employing nuclear extracts[31]. Wrenchnolol is cell permeable, however, the polyamide-wrenchnolol chimera is not. Therefore, the activity of this molecule in vivo remains to be addressed. The second example, reported by Mapp and co-workers, was derived from an in vitro functional screen of a small collection of isoxazolidine-containing molecules modeled loosely after an activating peptide derived from the VP 16 activation domain[32]. In this study, the synthetic

activation domain candidates were linked to methotrexate and then incubated with a nuclear extract containing a LexA DBD-dihydrofolate reductase (DHFR) fusion protein and a reporter gene with LexA binding sites in the promoter. High affinity binding of methotrexate to DHFR delivered the isoxazolidine-containing molecule to the promoter, a strategy similar to that employed in this study to join a synthetic activation domain substitute with a native DNA-binding domain. Using this assay, two molecules that stimulated transcription about five-fold in vitro were discovered. However, there are no reports of the activities of these molecules in vivo.

In this study, a different approach was taken to the isolation of a non-peptidic activation domain substitute by screening a library of ≈ 100,000 hexameric peptoids for ligands to the KIX domain of the mammalian coactivator CBP. The KIX domain is the target of the native activators, including CREB[9], and has been validated as a target for the isolation of peptides with activation domain function by Montminy and co-workers[4]. Three peptoids were isolated that appeared to bind a GST-KIX domain fusion protein with much higher affinity than the other compounds in the library. Two of these molecules (KBPo1 and KBPo2) were resynthesized and studied further. Both were found to bind GST-KIX domain in vitro with low micromolar dissociation constants (Figure 3.8). However, whereas KBPo2 did not bind appreciably to GST or BSA, suggesting it was a specific KIX domain ligand, KBPo1 exhibited promiscuous binding

behavior. Both peptoids scored as cell permeable in a functional assay that measures the ability of dexamethasone conjugates of the molecules (see Figure 3.9 for structures) to enter the cytoplasm and activate a Gal4 DBD-GRLBD-VP16 AD fusion protein (Figure 3.10).

The critical experiment was to determine if the KIX domain-binding peptoids are capable of functioning as an activation domain in cells. To do this, an assay (Fig. 3.4) in which the peptoid was coupled non-covalently to the Gal4 DBD-GRLBD by virtue of the OxDex-GRLBD interaction was used. This resulted in delivery of the peptoids to the promoters of the reporter genes. As shown in Figure 3.11, the KBPo2 molecule exhibited considerable activity, supporting a 900-fold increase in transcription at the highest concentration tested (100 μM). There was no induction of transcription when the cells were treated with the control steroid (OxDex-AEEA-$CONH_2$) or the KBPo1 steroid conjugate. These data suggest that the KBPo2 peptoid may be functioning as an activation domain in this system. Finally, in order to rule out the possibility that conformational changes in the binding of OxDex to GR are responsible for the observed transcriptional activity, a Mifepristone-AEEA-KBPo2 and Mifepristone-AEEA conjugate were prepared and subjected to cell permeability and transcriptional activation studies (Figure 3.12 and 3.13). Unfortunately, both the conjugates were completely impermeability, which precluded gleaning of any meaningful information from these assays. Thus, the precise mechanism of

OxDex-AEEA-KBPo2-mediated transcription activation remains to be established. It is clear that the observed activation potential is unique to KBPo2 since conjugation of spacer (AEEA alone) or KBPo1 to OxDex resulted in no observable transcription activation even at high concentrations.

Several experimental possibilities exist to determine whether or not KBPo2 has bonafide activation potential. To rule out the possibility that the observed transcription activation is mediated by the AF2 domain (see Figure 3.2) present within the LBD of the glucocorticoid receptor, one could remove the AF2 domain and re-assess the ability of the OxDex-AEEA-KBPo2 conjugate to activate transcription. Any observed activity should, in principle, be attributable to the inherent transactivation potential of the peptoid. The ability of OxDex-AEEA-KBPo2 to activate transcription in the presence of a huge excess of KBPo2 alone can also be assessed. Since most transcription co-activators are believed to be present in limiting concentrations, excess of activation domain alone by competing for the co-activator(s) should diminish the transcription activation activity of DBD-conjugated activation domain. On the other hand, if KBPo2 is inducing a conformational change in OxDex resulting in it adopting an agonist-like conformation, excess of KBPo2 alone should have no impact on the transcriptional activity of the peptoid-steroid conjugate. The success of this experiment, however, would be dependent on achieving high nuclear concentration of the peptoid without causing any toxicity to the cells.

Furthermore, the ability of KBPo2 to inhibit the ability of CBP-dependent gene expression can also be assessed. Thus, if KBPo2 binds to CBP at a site that is critical for its role in transcription, treatment of cells with KBPo2 should result in the repression of CBP-dependent gene expression. It should be noted, however, that such a result would only suggest a possible transactivation potential of the molecule and, in itself, is not a proof of transcriptional activity. Another possibility is to measure the transcriptional activity of the peptoid-steroid conjugate in cells that do not express CBP, the putative target of the peptoid. There has in fact been a report of RNAi-mediated silencing of CBP expression in SW480 cells[35]. However, the results of such an experiment should be interepreted with caution since transcription activators have been known to interact with multiple transcription factors. For instance, Gal80BP-A, which was selected based on its ability to bind to Gal80 in an in vitro binding assay has been shown to interact with Gal11, a yeast transcription co-activator, although no selection pressure for binding to Gal11 was applied in the selection of this peptide[2]. Finally, it should also be possible to use other receptor-ligand pairs in order to unequivocally establish the transctivation potential of KBPo2. For instance, the transcription activity of putative small molecule and peptide activation domains have been established by coupling them to FK506[34] or methotrexate[32]. Similar to dexamethasone-GR system, the high affinity interaction between FK506-methotrexate and FKBP-dihydrofolate reductase

(DHFR) has been used to deliver the synthetic activation domain to a promoter. Since FKBP and DHFR and completely devoid of any transcription activity, such a system should provide an unambiguous answer regarding the transcription activity of KBPo2.

In conclusion, a cell permeable, coactivator-binding peptoid was identified and shown to activate transcription in living mammalian cells as a steroid conjugate. Future studies are necessary to establish the mechanism by which this molecule causes transcription activation.

3.5. Bibliography

1. Alluri, P.G., Reddy, M.M., Bachhawat-Sikder, K., Olivos, H.J., and Kodadek, T., *Isolation of protein ligands from large peptoid libraries.* J. Amer. Chem. Soc., 2003. **125**: p. 13995-14004.
2. Han, Y., and Kodadek, T., *Peptides selected to bind the Gal80 repressor are potent transcriptional activation domains in yeast.* J. Biol. Chem., 2000. **275**: p. 14979-14984.
3. Wu, Z., Belanger, G., Brennan, B.B., Lum, J.K., Minter, A.R., Rowe, S.P., Plachetka, A., Majumdar, C.Y., and Mapp, A.K., *Targeting the transcriptional machinery with unique artificial activators.* J. Amer. Chem. Soc., 2003. **125**: p. 12390-12391.
4. Frangioni, J.V., LaRiccia, L.M., Cantley, L.C., and Montminy, M.R., *Minimal activators that bind the KIX domain of p300/CBP identified by phage display screening.* Nat. Biotechnol., 2000. **18**: p. 1080-1085.
5. Miller, S.M., Simon, R.J., Ng, S., Zuckermann, R.N., Kerr, J.M., and Moos, W.H., *Comparison of the proteolytic susceptibilities of homologous L-amino acid, D-amino acid, and N-substituted glycine peptide and peptoid oligomers.* Drug Dev. Res., 1995. **35**(1): p. 20-32.
6. Yu, P., Liu, B., and Kodadek, T., *A high-throughput assay for assessing the cell permeability of combinatorial libraries.* Nat. Biotechnol., 2005. **23**: p. 746-751.

7. Bannister, A., and Kouzarides, T., *The CBP co-activator is a histone acetyl transferase.* Nature, 1996. **384**: p. 641-643.

8. Chrivia, J.C., Kwok, R.P., Lamb, N., Hagiwara, M., Montminy, and M., Goodman, R.H., *Phosphorylated CREB binds specifically to nuclear protein CBP.* Nature, 1993. **365**: p. 855-859.

9. Radhakrishnan, I., Perez-Alvarado, G.C., Parker, D., Dyson, H.J., Montminy, M.R., and Wright, P.E., *Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions.* Cell, 1997. **91**: p. 741-752.

10. Shaywitz, A.J., Dove, S.L., Kornhauser, J.M., Hochschild, A., and Greenberg, M.E., *Magnitude of the CREB-dependent transcriptional response is determined by the strength of the interaction between the kinase-inducible domain of CREB and the KIX domain of the CREB-binding protein.* Mol. Cell. Biol., 2000. **20**(24): p. 9409-9422.

11. Liu, B., Alluri, P.G., Yu, P., and Kodadek, T, *A Potent transactivation domain mimic with activity in living cells.* J. Amer. Chem. Soc., 2005. **127**(23): p. 8254-8255.

12. Aranda, A., and Pascual, A., *Nuclear hormone receptors and gene expression.* Physiol. Rev., 2001. **81**: p. 1269-1304.

13. McEwan, I.J., Wright, A.P.H., and Gustafsson, J.-A., *Mechanism of gene expression by the glucocorticoid receptor: Role of protein-protein interactions.* Bioessays, 1997. **19**(2): p. 153-160.

14. Truss, M., and Beato, M., *Steroid hormone receptors: interaction with deoxyribonucleic acid and transcription factors.* Endocr. Rev., 1993. **14**: p. 459-479.

15. Chauhan, D., Pandey, P., Ogata, A., Teoh, G., Treon, S., Urashima, M., Kharbanda, S., and Anderson, K.C., *Dexamethasone induces apoptosis of multiple myeloma cells in a JNK/SAP kinase independent mechanism.* Oncogene, 1997. **15**: p. 837-843.

16. Schreiber C., and Mitchell, C., *Mifepristone in abortion care.* Semin. Reprod. Med., 2005. **23**(1): p. 82-91.

17. Giguere, V., Hollenberg, S.M., and Rosenfeld, M.G., *Functional domains of the human glucocorticoid receptor.* Cell, 1986. **46**(5): p. 645-652.

18. Hollenberg, S.M., and Evans, R.M., *Multiple and cooperative trans-activation domains of the human glucocorticoid receptor.* Cell, 1988. **55**: p. 899-906.

19. Feil, R., Brocard, J., Mascrez, B., LeMeur, M., Metzger, D., and Chambon, P., *Ligand-activated site-specific recombination in mice.* Proc. Natl. Acad. Sci, USA, 1996. **93**: p. 10887-10890.

20. Olivos, H.J., Alluri, P.G., Reddy, M.M., Salony, D., and Kodadek, T., *Microwave-assisted solid phase synthesis of peptoids.* Org. Lett., 2002. **4**: p. 4057-4059.

21. von Geldern, T.W., Tu, N., et al., *Liver-selective glucocorticoid antagonists: A novel treatment for Type 2 diabetes.* J. Med. Chem., 2004. **47**: p. 4213-4230.

22. Lam, K.S., Salmon, S.E., Hersh, E.M., Hruby, V.J., Kazmierski, W.M., and Knapp, R.J., *A new type of synthetic peptide library for identifying ligand-binding activity.* Nature, 1991. **354**: p. 82-84.

23. Zuckermann, R.N., Kerr, J.M., Kent, S.B.H., and Moos, W.H., *Efficient method for the preparation of peptoids (oligo N-substituted glycines) by submonomer solid-phase synthesis.* J. Amer. Chem. Soc., 1992. **114**: p. 10646-10647.

24. Uno, T., Beausoleil, E., Goldsmith, R.A., Levine, B.H., and Zuckermann, R.N., *New submonomers for poly N-substituted glycines (peptoids).* Tet. Lett., 1999. **40**: p. 1475-1478.

25. Pons, J.-F., Fauchere, J-L., Lamaty, F., Molla, A., and Lazaro, R., *A constrained diketopiperazine as a new scaffold for the synthesis of peptidomimetics.* Eur. J. Org. Chem., 1998. **1998**: p. 853-859.

26. Heyduk, T., Ma, Y., Tang, H., and Ebright, R.H., *Fluorescence anisotropy: Rapid, quantitative assay for protein-DNA and protein-protein interactions.* Methods Enzymol., 1996. **274**: p. 492-503.

27. Lecitra, E.J., and Liu, J.O., *A three-hybrid system for detecting small ligand-protein interactions.* Proc. Natl. Acad. Sci. USA, 1996. **93**: p. 12817-12821.

28. Denison, C., and Kodadek, T., *Small molecule-based strategies for controlling gene expression.* Chem. & Biol., 1998. **5**: p. R129-R145.

29. Ansari, A.Z., and Mapp, A.K., *Modular design of artificial transcription factors.* Curr. Opin. Chem. Biol., 2003. **6**: p. 765-772.

30. Shimogawa, H., Kwon, Y., Mao, Q., Kawazoe, Y., Choi, Y., Asada, S., Kigoshi, H., and Uesugi, M., *A wrench-shaped synthetic molecule that modulates a transcription factor-coactivator interaction.* J. Amer. Chem. Soc., 2004. **126**: p. 3461-3471.

31. Kwon, Y., Arndt, H.-D., Mao, Q., Kawazoe, Y., Dervan, P.B., Uesugi, M., *Small molecule transcription factor mimic.* J. Amer. Chem. Soc., 2004. **126**: p. 15940-15941.

32. Minter, A.R., Brennan, B.B., and Mapp, A.K., *A small molecule activation domain.* J. Amer. Chem. Soc., 2004. **126**: p. 10504-10505.

33. Cress, W.D. and Triezenberg, S.J., *Critical structural elements of the VP16 transcriptional activation domain.* Science, 1991. **251**: p. 87-90.

34. Nyanguile, O., Uesugi, M., Austin, D.J., and Verdine, G.L., *A nonnatural transcriptional coactivator.* Proc. Natl. Acad. Sci. USA, 1997. **94**: p. 13402-13406.

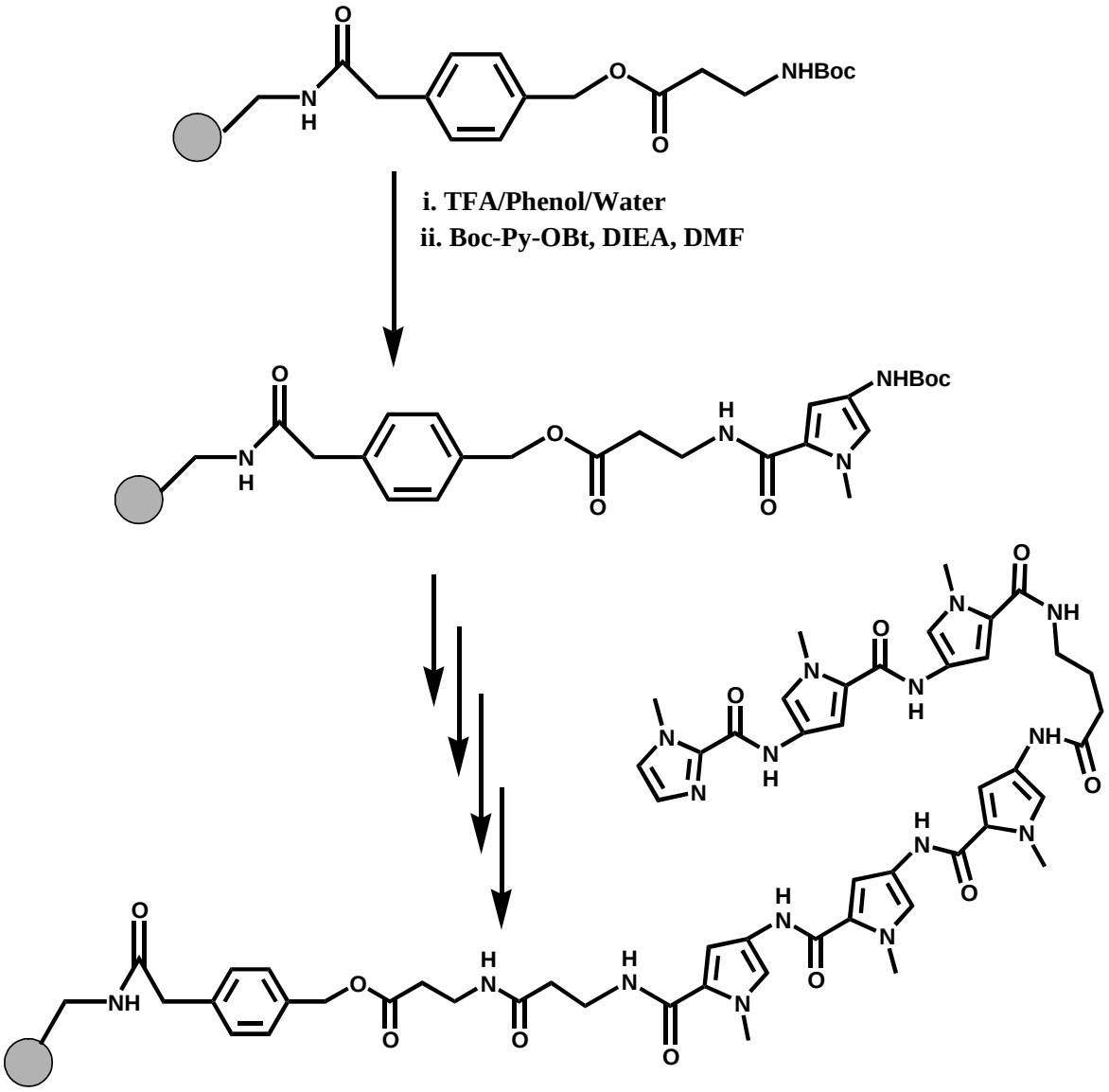
35. Emami, K.H., Nguyen, C., Ma, H., Kim, D.H., Jeong, K.W., Eguchi, M., Moon, R.T., Teo, J.-L., Oh, S.W., Kim, H.Y., Moon, S.H., Ha, J.R., and Kahn, M., *A small molecule inhibitor of β-catenin / CREB-binding protein transcription.* Proc. Natl. Acad. Sci. USA, 2004. **101**: 12682-12687.

CHAPTER 4: SYNTHESIS AND CHARACTERIZATION OF HAIRPIN POLYAMIDE DERIVATIVES

4.1 Introduction

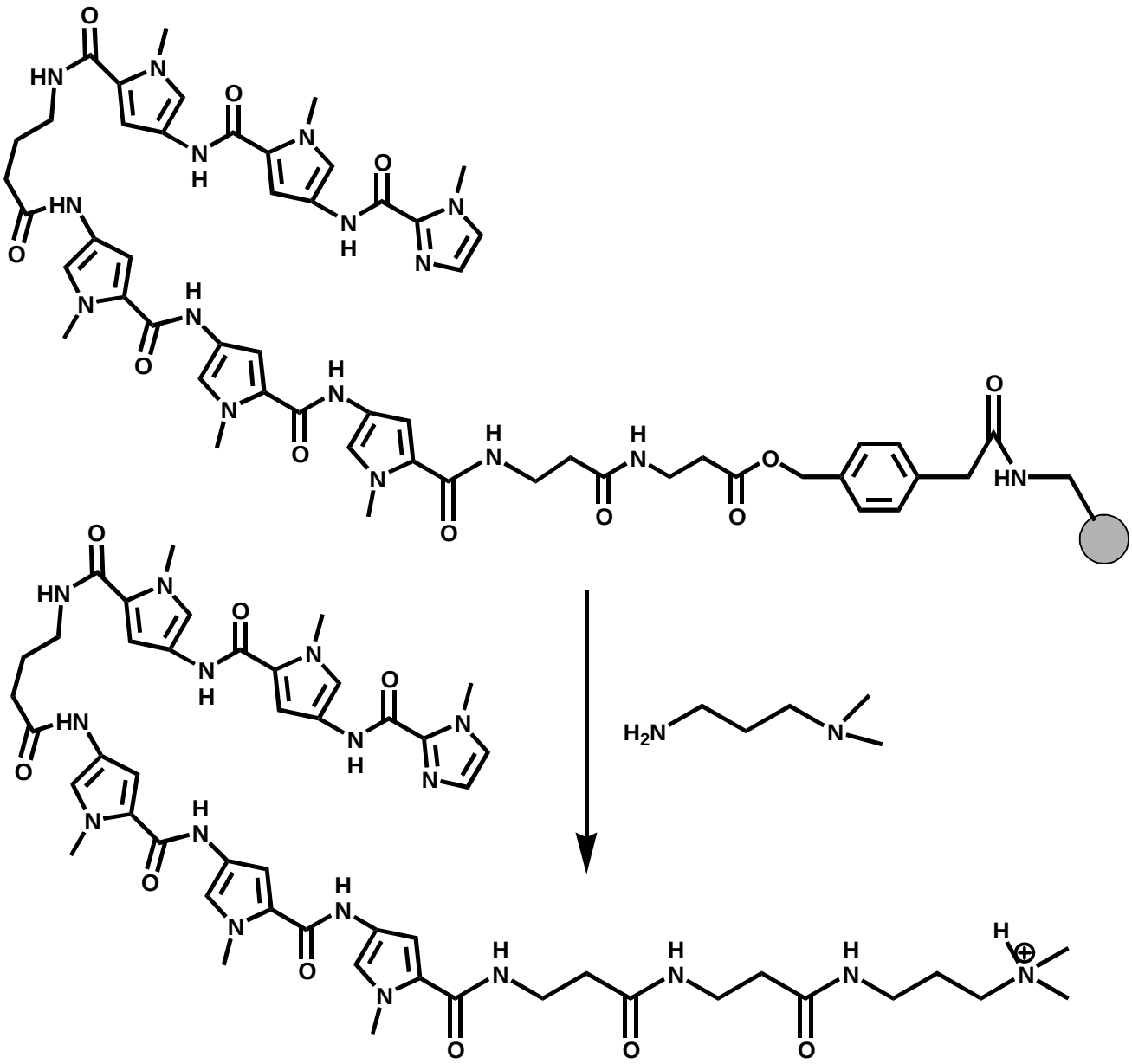
There is great interest in the development of synthetic molecules capable of activating the expression of specific genes in a predictable manner. Essentially all approaches to this problem have involved the construction of chimeric molecules comprised of a sequence-specific DNA binding unit linked to a moiety capable of recruiting the RNA polymerase II transcription complex [1]. This strategy has been made feasible by the extensive work done on synthetic, sequence-specific DNA-binding molecules. Hairpin polyamides (HPs) represent a highly attractive class of DNA-binding molecules since simple pairing rules allow design of HPs to bind to desired DNA sequence in vitro [2]. Furthermore, HPs have been found to be cell permeable in some cases [3]. While much needs to be learned regarding their activity in living cells, polyamides provide an exciting opportunity for development of artificial transcription activators. Furthermore, the identification of KIX-binding peptoids with apparent transcription activation potential suggested that HP-peptoid conjugates may have the ability to serve as synthetic transcription factors [4]. Therefore, efforts were focused on developing appropriate conditions for the synthesis of HPs and HP-peptoid conjugates.

The solid phase synthesis protocol developed by Dervan and co-workers provides a facile and convenient route for the synthesis of HPs [5]. This procedure was essentially an adaptation of t-Boc chemistry that has been extensively employed in the synthesis of peptides. Scheme 4.1 depicts salient features of solid phase HP synthesis. The synthesis is carried out on Pam resin, which has been pre-loaded with a β-alanine residue. The t-Boc group protecting the primary amine of β-alanine is removed by treatment with TFA and the first monomer is coupled by forming an amide bond between the amine group of β-alanine and the activated carboxylate group of the monomer (usually a pyrrole or imidazole derivative). The activation of the carboxylic acid group is carried out using HBTU or HOBt. The cycle of t-Boc deprotection and coupling of a new monomer are repeated until the desired length is achieved. At the end of the synthesis, the HP is released from the beads in an aminolysis reaction, in which the beads displaying the HP are heated with a neat solution of amine at 80^{o}C. One advantage of this type of cleavage is that by choosing an appropriate amine, one can install a variety of functional groups at the HP terminus that serve as handles for subsequent chemistry.



Scheme 4.1: Solid phase synthesis of polyamides.

A schematic depiction of various steps involved in the solid phase synthesis of HPs. [5]



Scheme 4.2: Release of HP from the resin.

A schematic depiction of the aminolysis reaction used to release the HP from the resin at the end of the synthesis. [5]

4.2 Materials and methods

All of the reagents and solvents were purchased from commercial suppliers and used without further purification. Analytical HPLC was performed on a Biocad Sprint system with a C18 reverse-phase HPLC column (Vydac, 5μM, 4.6 mm i.d. x 250 mm) using the following solvents for elution; solvent A: H_2O / 0.1% TFA; solvent B: CH_3CN / 0.1% TFA. MALDI-TOF MS was performed on a Voyager-DE PRO biospectrometry workstation (Applied Biosystems) using α-hydroxy cinnamic acid as the matrix.

4.2.1 Activation of Boc-Py-acid

Activation of Boc-Py-acid via formation of an ester with hydroxybenzotriazole (HOBt) was carried out following the literature procedure [5].

4.2.2 Synthesis of Im-Py3-γ-Py4-β-Dp_2

Synthesis of this molecule was carried out employing a slight variation in the reported literature procedure [5]. Thus, 100 mg of Boc-β-alanine-Pam-resin (0.76 mmoles / g) was swollen in DMF for 30 minutes. The solvent was drained and the Boc group was removed using a cleavage cocktail composed of 92.5% TFA, 5% phenol and 2.5% water. The deprotection step was carried three times for 2, 5 and 10 minutes, respectively in 3 mL of cleavage cocktail. The beads

were then washed thoroughly with DMF (6 X 3 mL) and neutralized in 20% DIEA/DMF for 5 minutes. The beads were washed again with DMF (3 X 3 mL) and treated with 3 mL of 0.2 M solution of Boc-Py-OBt ester in DMF containing 0.5 mL of DIEA and shaken in an incubator shaker for 60 minutes. At the end of the coupling step, the beads were thoroughly washed with DMF and deprotection of Boc group from the previous residue and coupling of the next Py residue was carried out using the same procedure described earlier. In the case of γ-aminobutyric acid and Imidazole-2-carboxylic acid couplings, 1 mmole of the acid and 0.95 mmoles of HBTU were dissolved in 2 mL of DMF and 1 mL of DIEA and added to the beads. The coupling was carried out as described earlier for 60 minutes in an incubator shaker. At the end of the synthesis, the beads were thoroughly washed with DMF and DCM (5 X 3 mL) and dried under vacuum. The dry resin was then transferred to an eppendorf tube, 500 μL of neat (N,N-dimethylamino)propylamine added and heated at 80^{o}C for 6 hours in an oven. The beads were manually agitated at least once every hour to mix the contents. The cleavage mixture was then diluted with 0.1% TFA/water, filtered and the filtrate was purified by HPLC using a C18 column. The HPLC fractions containing the pure product were lyophilized and stored at -20^{o}C until further use.

4.2.3 Synthesis of Im-Py3-γ-Py4-β-PPG

Im-Py3-γ-Py4-β was synthesized as described above, resin thoroughly washed with DMF and DCM (5 X 3 mL) and dried. 25 mg of the resin was then transferred to an eppendorf tube and 125 μL of neat propargylamine was added and incubated in an oven at 60^{o}C for 24 hours. The beads were manually agitated frequently to mix the contents. The cleavage mixture was then diluted with 0.1% TFA/water, filtered and the filtrate was purified by HPLC using a C18 column. The HPLC fractions containing the pure product were lyophilized and stored at -20^{o}C until further use.

4.2.4 Synthesis of KBPo1-AEEA$_2$-Cys

KBPo1 was synthesized on 50 mg Rink amide MHBA resin (0.69 mmoles/g) as described earlier using the microwave assisted protocol [4]. To the peptoid beads was then added 1.5 mL solution each of 0.2 M Fmoc-AEEA in DMF and 0.2 M HBTU in NMM/DMF (0.4M) and shaken at room temperature for 60 minutes. The beads were then washed thoroughly with DMF (6 X 3 mL) and treated with 20% piperidine/DMF for 10 minutes (2 X 3 mL) to deprotect the Fmoc group. The same procedure was repeated to couple the second spacer. After deprotecting the Fmoc group with 20% piperidine in DMF (as described earlier), the beads were thoroughly washed with DMF (6 X 3 mL) and treated with a 1.5 mL solution each of 0.2 M Fmoc-Cys-COOH in DMF and 0.2 M HBTU in

NMM/DMF (0.4M) and shaken for 60 minutes. The beads were washed thoroughly with DMF and the Fmoc group was removed with 20% piperidine / DMF as described earlier. The beads were thoroughly washed with DMF (5 X 3 mL) and DCM (5 X 3 mL) and dried. The product was released from the beads by treatment with a cleavage cocktail containing 95% TFA, 2.5% triisopropylsilane and 2.5% water for 2 h. The cleavage mixture was filtered, filtrate concentrated by blowing nitrogen and lyophilized. The crude lyophilized product was purified by reverse phase HPLC and the pure fractions were lyophilized and stored at -20^{o}C until further use.

4.2.5 Synthesis of KBPo1-PEG-N_3

KBPo1 was synthesized on 50 mg of Rink amide MHBA resin (0.69 mmoles/g) as described earlier under microwave conditions [4]. The peptoid beads were then treated with 1.5 mL solution each of 0.2 M solution of the PEG spacer (which terminates as a carboxylate on one end and as azide on the other end) and 0.2 M HBTU in NMM/DMF (0.4 M) for 60 minutes at room temperature. At the end of the coupling step, the beads were thoroughly washed with DMF (6 X 3mL) and DCM (6 X 3 mL), dried and cleaved using a cleavage cocktail consisting of 95% TFA, 2.5% triisopropylsilane and 2.5% water for 2 h. The cleavage mixture was separated from the beads by filtration and the filtrate was concentrated by blowing nitrogen and lyophilized. The crude lyophilized

product was purified by reverse phase HPLC and the pure product fractions were lyophilized and stored at -20^{o}C until further use.

4.2.6 Synthesis of thiolane-2,5-dione

Synthesis of thiolane-2,5-dione was carried out following a previously reported procedure [6].

4.2.7 Synthesis of Im-Py3-γ-Py4-β-Dp_2 thioester for native ligation

Synthesis of the HP thioester was carried out following a procedure, which was a slight adaptation of a previously reported protocol [7]. To a solution of 2 mg of polyamide (from 4.2.2) in 100 μL of NMM was added 2.5 μL of 1.0 M solution of thiolane-2,5-dione and 1 μL of DIEA. The reaction mixture was incubated for 30-45 minutes with occasional agitation to mix the contents. The reaction mixture was then diluted with 90 μL of 100 mM sodium acetate buffer (pH 5.2), cooled to 4^{o}C and 1 μL of benzyl bromide was added with mixing and allowed to sit for 45 minutes. The reaction mixture was then diluted with 0.1% TFA/water and purified by reverse phase HPLC. The HPLC fractions containing the desired product were lyophilized and stored at -20^{o}C until further use.

4.2.8 Native ligation conditions for attempted conjugation of Im-Py3-γ-Py4-β-Dp with KBPo1-AEEA$_2$-Cys

Native ligation mediated coupling of the peptoid terminating in cysteine (from 4.2.4) to the polyamide thioester (from 4.2.7) was attempted under previously reported literature conditions [7]. Thus, 1.8 mg of the peptoid derivative was dissolved in 20 μL of NMM and added to 1.5 mg of the polyamide derivative in 180 μL of 100 mM potassium phosphate buffer (pH 7.3) containing 5% thiophenol (10 μl). The reaction mixture was incubated at room temperature for 2-4 days with occasional agitation to mix the contents.

4.2.9 Click chemistry conditions for attempted conjugation of Im-Py3-γ-Py4-β-PPG to KBPo1-PEG-N$_3$

Click chemistry mediated coupling of the peptoid terminating in an azide (from 4.2.5) with the polyamide derivative (from 4.2.3) was attempted under previously reported literature conditions [8]. Thus, 1.6 mg of the peptoid derivative and 1.2 mg of the polyamide derivative were dissolved in 100 μL of 1:1 water and tert-butyl alcohol and 3 μL of freshly prepared 1M sodium ascorbate and 0.5 mg of copper (II) sulfate pentahydrate were added and tumbled at room temperature for 24-48 hours.

4.2.10 Dnase I footprinting

Dnase I footprinting titrations were carried out following the established protocol with slight modification [9]. A polymerase chain reaction (PCR) was carried out to get a 290 bp DNA fragment containing six Impy3-γ-Py4-binding sites (5'-TGTTAT-3'), using pGL3-6×HPB plasmid as template, H6TATA01(5'-GGC GCG GAA TTC TAG GCT GTC CCC AGT GCA-3') and H6TATA02(5'-GCG CGC GGA TCC AGC GGA TAG AAT GGC GCC-3') as primers. This fragment was inserted in the *EcoR1*/*BamH1* sites of plasmid pUC19 (Invitrogen). The constructed plasmid was named pUC19-H6. 7 μg plasmid pUC19-H6 DNA was digested by *EcoR1* and precipitated in 70 % ethanol. The linear DNA was dissolved in 10 μL of H_2O and to this solution was added 2 μL of sequenase reaction buffer (provided by USB Corporation), 1 mL of 0.1 M dithiothreitol, 0.5 μL of [^{32}P]dATP (10 mCi/mL, ~3,000 Ci/mmol, Amersham), 0.5 μL of [^{32}P]dTTP (10 mCi/mL, ~3,000 Ci/mmol, Amersham), and 2 μL of sequenase (USB Corporation). The reaction was carried out at room temperature for 25 min. 1 μL of 10 mM dNTP was added and the reaction continued for another 5 min. The reaction mixture was extracted with 50 μL of phenol/chloroform/isoamyl alcohol (25:24:1), 50 μL of chloroform, and purified with a G-50 gel filtration spin column (Princeton Separation). The labeled DNA was digested by *BamH1* to get a single-end labeled DNA fragment. The final single-end radio-labeled 272

bp DNA fragment was purified using 0.8% agarose gel electrophoresis and QIAGEN QIAquick gel extraction kit.

All binding reactions were executed in a total volume of 200 μL of solution containing 100 pM labeled DNA, 10 mM Tris (pH 8.0), 5 mM $MgCl_2$, 1 mM $CaCl_2$, 2 mM DTT, 100 mM KCl, and 0-300 nM ImPy7. The reactions were allowed at room temperature for 16 h. Footprinting reactions were initiated by adding 1 μL of a Dnase1 stock solution (Invitrogen, diluted to the concentration giving > 50% intact DNA). The reactions lasted 2 min at 37 °C. After each reaction, the solution was added immediately to a stop solution (648 μL of ethanol, 2 μL of 3 mg/mL tRNA, and 50 μL of 7.5 ammonium acetate) at –70 °C. The DNA pellets were washed with 75 % ethanol, dried in air, and resuspended in 7 μL of formamide loading buffer. The DNA loading samples were heated at 88 °C for 10 min and placed on ice before loading. Electrophoresis was carried on an 8% denaturing polyacrylamide gel (5% cross-linking, 7 M urea) at 2,000 V for 3.0 h. Urea was removed by soaking the gel in a solution containing 5% acetic acid and 15% methanol for 20 min. Gels were dried and exposed to a storage phosphor screen (Molecular Dynamics).

4.2.11 DNA sequencing

3.2 μg pUC19-H6 plasmid DNA was denatured in 27 μL of denaturing buffer containing 222 mM NaOH and 2.2 mM EDTA at 37 °C for 30 min. 3 μL of 3.0 M sodium acetate (pH 5.0) was added to neutralize the buffer, and the single strand DNA was precipitated in 70 % ethanol. The DNA pellet was resuspended in 7 μL of water and 2 μL of sequenase buffer (USB Corporation), and 1 μL of 5 μM sequencing primer H6TATA01seq (5'-AA TTC TAG GCT GTC CCC AGT-3') was added. The mixture was heated at 75 °C for 2 min, cooled slowly to room temperature, and placed on ice for 10 minutes before sequencing reactions. Sequencing reactions were carried out following the protocol provided by USB.

4.3 Results

4.3.1 Synthesis and characterization of Im-Py3-γ-Py4-β-Dp_2

The hairpin polyamide Py3-γ-Py4-β-Dp_2 (Figure 4.1) was chosen as the target for synthetic and DNA binding studies. It has been previously designed by Dervan and co-workers to recognize a 6 base pair sequence of 5'-TGTTAT-3'. It has also been used previously as the DNA-binding domain of artificial transcription activators in cell free systems [10]. Thus, it was thought to serve as

an excellent starting point towards development of an artificial transcription activator containing a peptoid as the artificial activation domain. Py3-γ-Py4-β was synthesized on solid phase following the published procedure with minor variations (as detailed in the experimental section) [5]. As described earlier, the polyamide synthesis involves standard Boc-chemistry, which was developed initially for the synthesis of peptides. The activation of the carboxylate groups of the monomers such as imidazole-2-carboxylic acid and γ-aminobutyric acid can be readily achieved in situ using HBTU. However, the efficiency of activation of Boc-Py-COOH is poor and in situ activation generally affords poor results. Therefore, an activated version of Boc-Py-COOH (Boc-Py-OBt ester) was prepared and used as a monomer in the polyamide synthesis to improve the overall efficiency of the synthesis. At the end of the synthesis, the polyamide was cleaved from the resin with 3,3'-diamino-*N*-methyldipropyl amine to leave a primary amine at the c-terminus of the sequence. The cleaved product was purified by HPLC. Figure 4.2 shows HPLC trace and MALDI-TOF mass spectrum of the polyamide.

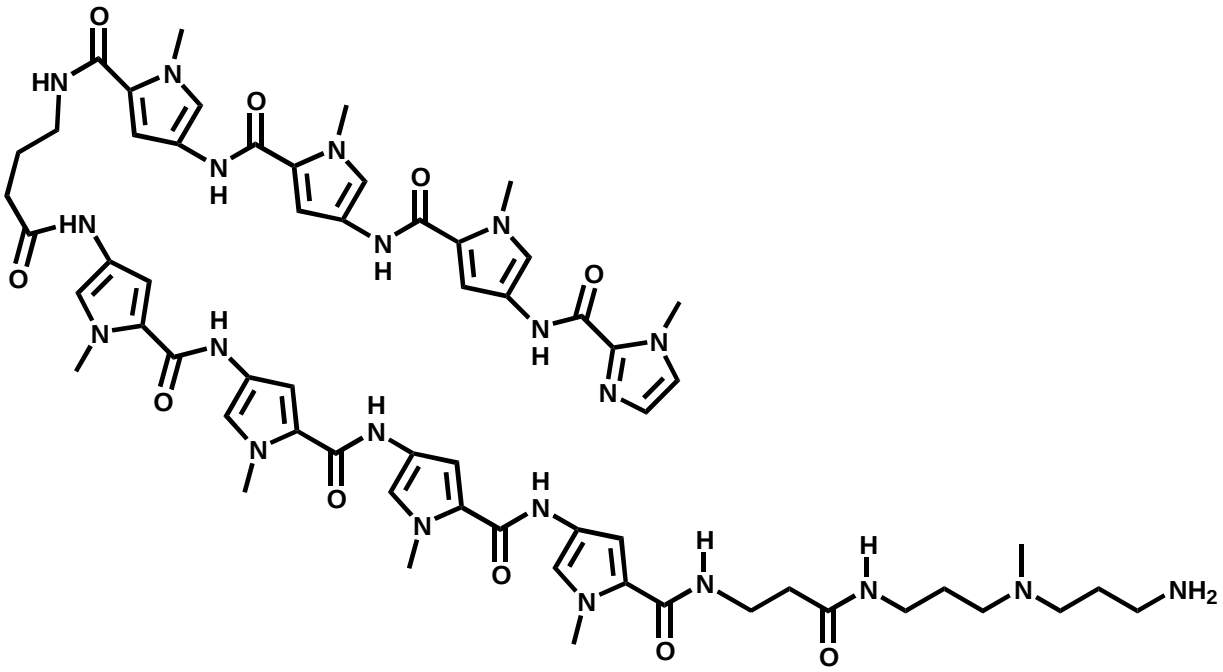


Figure 4.1: Im-Py3-γ-Py4-β-Dp$_2$.

Chemical structure of the HP designed to recognize a 6 base pair sequence of 5'-TGTTAT-3'.

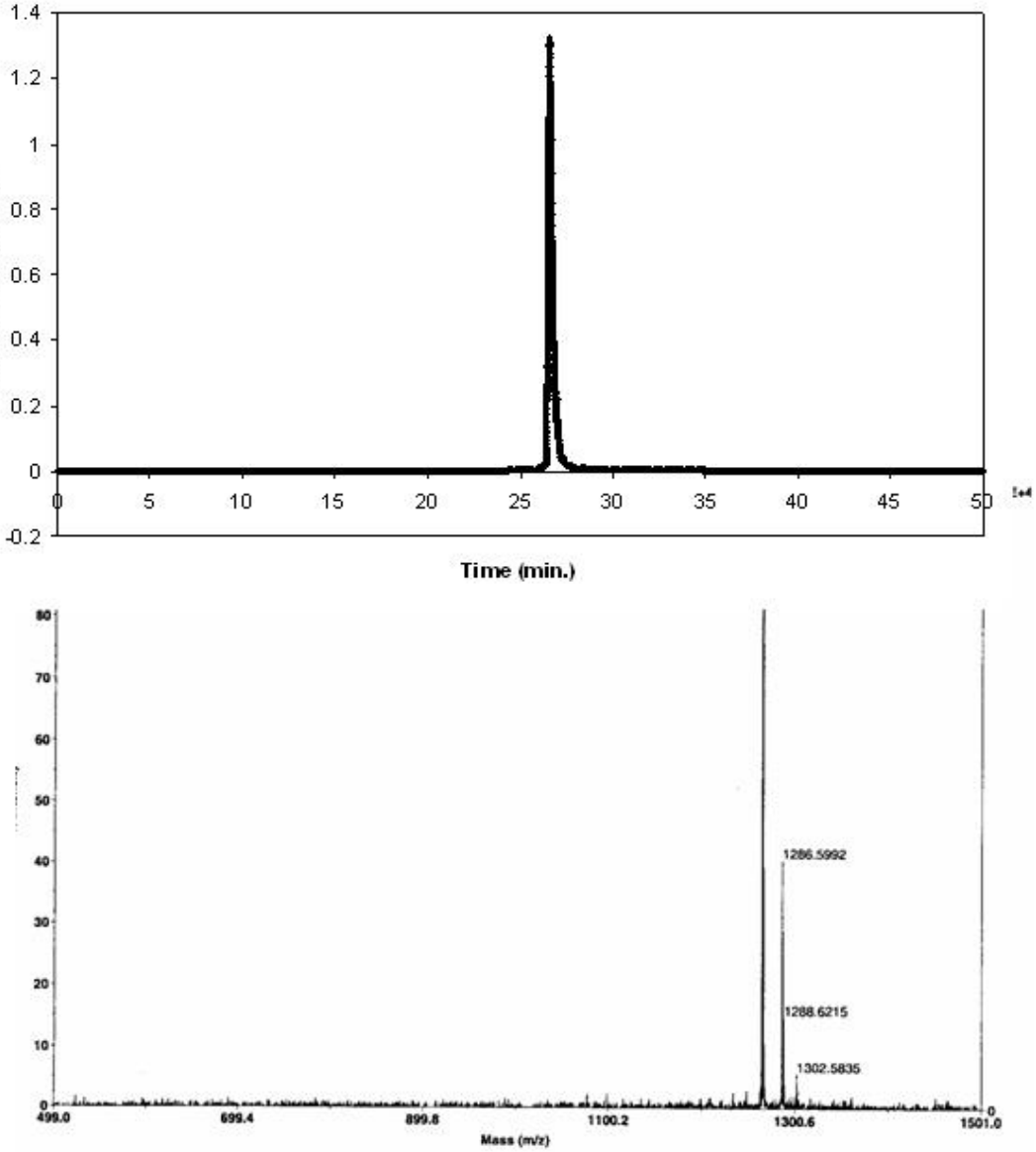


Figure 4.2: Characterization of Im-Py3-γ-Py4-β-Dp_2.

HPLC (top panel) and MALDI-TOF mass spectrum (bottom panel) of Im-Py3-γ-Py4-β-Dp_2.

4.3.2 DNA binding studies of Im-Py3-γ-Py4-β-Dp_2

With a high quality polyamide in hand, efforts were focused on studying the DNA binding properties of Im-Py3-γ-Py4-β-Dp_2 by DNase I footprinting. A 276 bp DNA fragment containing six Im-Py3-γ-Py4-β-Dp (5'-TGTTAT-3') and a minimal E1b promoter was singly 3' end-labeled with ^{32}P. About 100 pM of radio-labeled DNA was incubated with different concentrations of Im-Py3-γ-Py4-β-Dp for 16 hours followed by brief treatment with Dnase I. Figure 4.3 shows the experimental results. The binding of Im-Py3-γ-Py4-β-Dp_2 to the target sites was clearly seen. Complete binding was observed at 50 nM Im-Py3-γ-Py4-β-Dp_2, while no visible binding occurred at 5 nM concentration. This places the equilibrium dissociation constant (K_D) somewhere in this range and shows that Im-Py3-γ-Py4-β-Dp_2 binds to its target DNA site in vitro with high affinity.

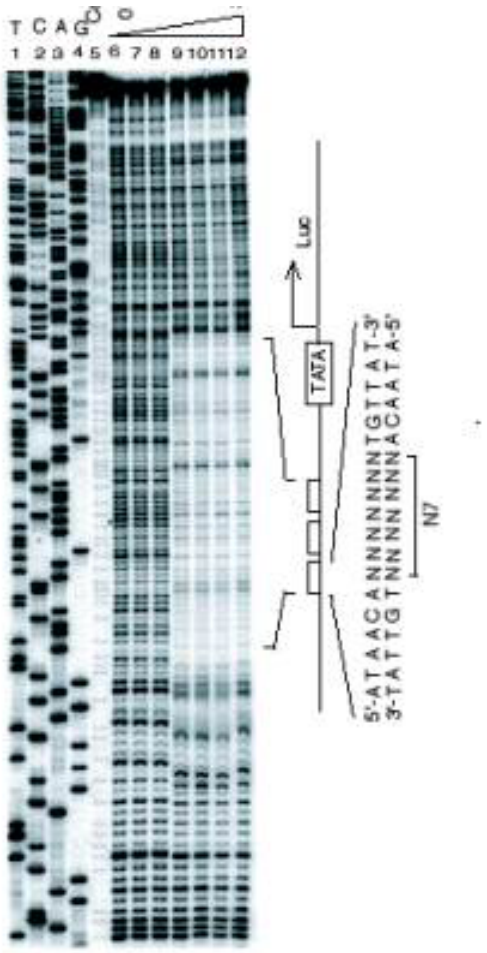


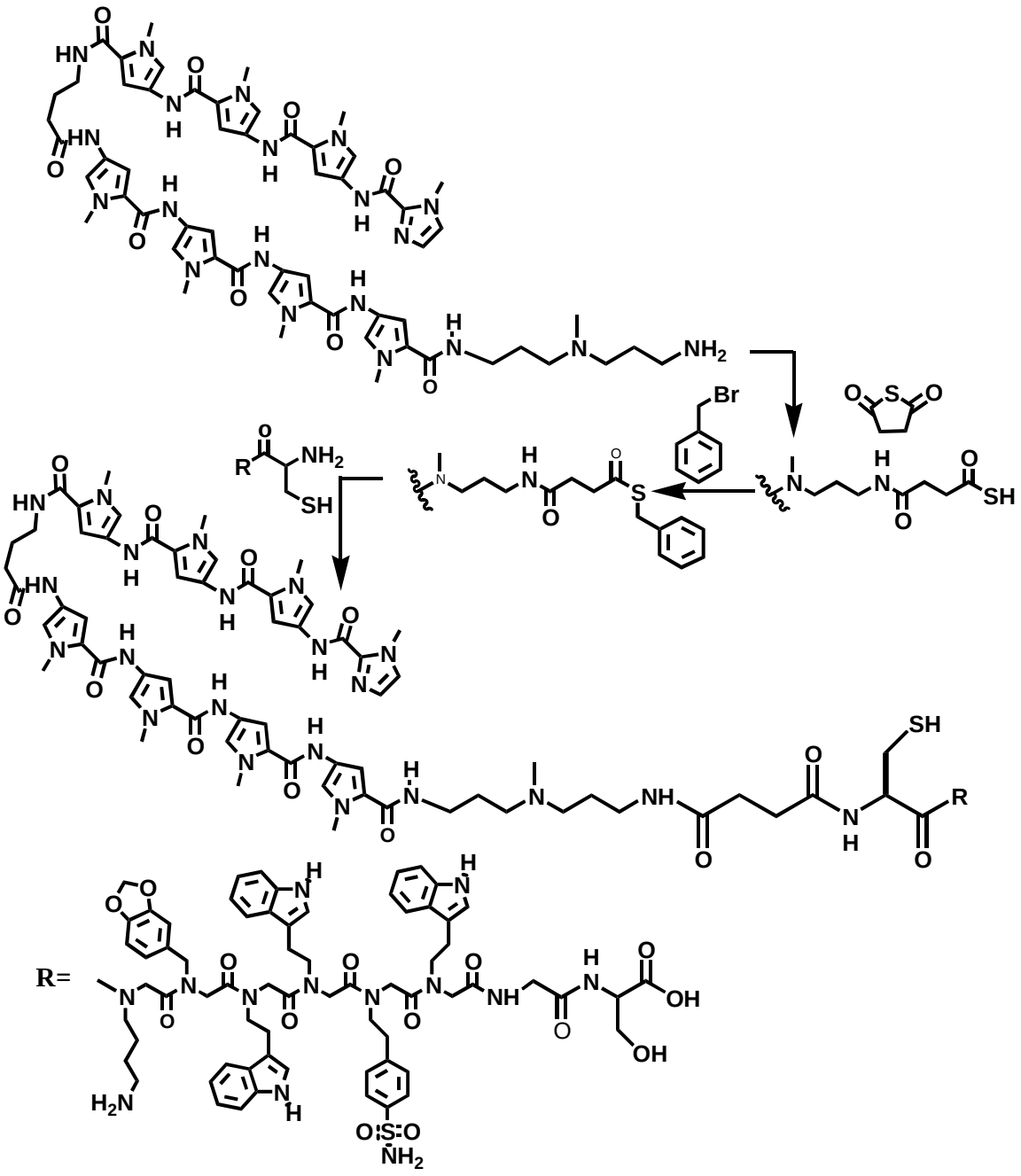
Figure 4.3: DNA binding studies of Im-Py3-γ-Py4-β-Dp$_2$.

Storage phosphor autoradiogram of a Dnase I footprinting titration of Im-Py3-γ-Py4-β-Dp$_2$ on the 3'-^{32}P-labeled 276 bp DNA fragment containing the promoter region of the reporter plasmid pGL3-6×HPB. Lane 1-4: sequencing lane T, C, A, G. Lane 5: undigested DNA. Lane 6-12: Dnase I digestion products in the presence of Im-Py3-γ-Py4-β-Dp at concentrations of 0 nM, 1 nM, 5 nM, 50 nM, 100 nM, 200 nM, and 300 nM. (Figure courtesy of Dr. Bo Liu)

4.3.3 Synthetic studies towards developing a general procedure for preparing peptoid-polyamide conjugates

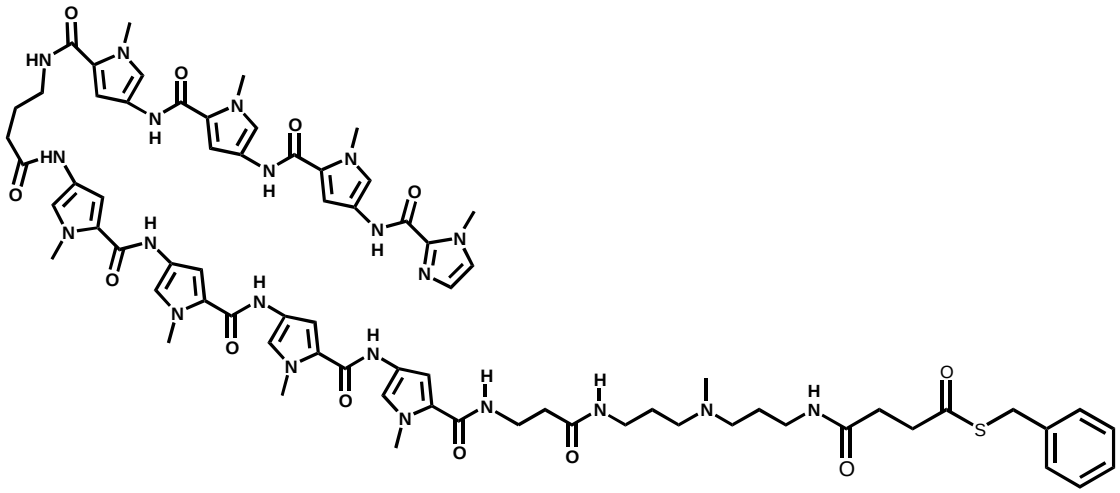
4.3.3.1 Native ligation

With a protocol for the synthesis and characterization of polyamides in place, the next step was to develop conditions for preparation of polyamide-peptoid conjugates to assess their transcriptional activation potential. Dervan and co-workers have described a native ligation-based procedure for coupling peptides to polyamides [7]. The same scheme, in principle, should be applicable for conjugating peptoids to polyamides. As shown in scheme 4.3, it consists of preparation of a polyamide-thioester derivative by reacting the free terminal amino group of the polyamide with thiolane-2,5-dione to result in the formation of a thioacid derivative of the polyamide. The thioacid is then treated with benzyl bromide to form of a thioester following displacement of the bromide by the thiolate. Incubation of the polyamide thioester with a peptoid terminating in cysteine under native ligation conditions should result in the desired polyamide-peptoid conjugate.



Scheme 4.3: Proposed scheme for native ligation-mediated synthesis of HP-peptoid conjugates.

To test the applicability of this route for preparing peptoid-polyamide conjugates, the derivatives of peptoid and polyamide shown in figure 4.4 and 4.5 were synthesized. The Dervan protocol for synthesis of HP thioesters was found to work quite well. Figure 4.4 shows the MALDI-TOF mass spectrum of the final product isolated from this synthetic route, which positively identified it as the desired HP thioester derivative. However, it was observed that presence of even trace amounts of (N,N-dimethylamino)propylamine, (which is used to release the polyamide from the beads), in the polyamide sample could significantly compromise the efficiency and yields of the thioester formation reaction. Ensuring complete removal of the free amine can be tricky as it is "invisible" in the HPLC traces. Using a shallow elution gradient during HPLC purification of the final product has been found to be effective in completely separating the amine from the product.



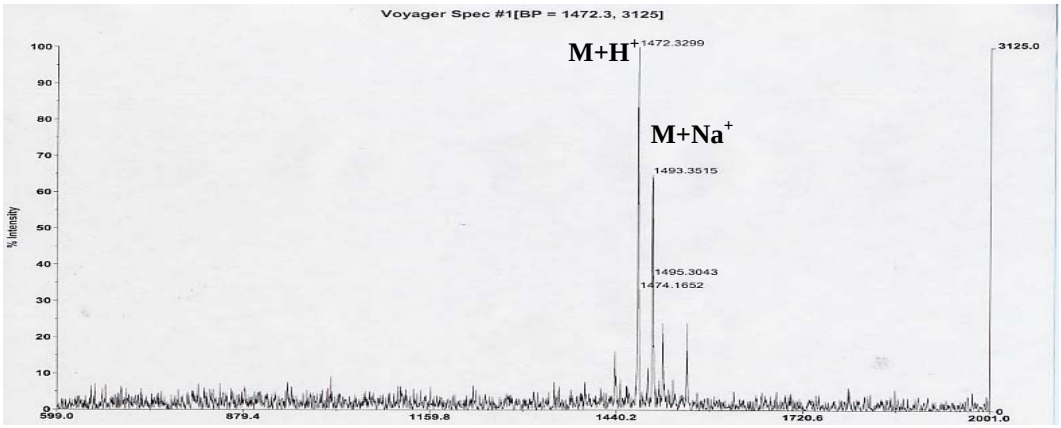


Figure 4.4: HP thioester.

The chemical structure and MALDI-TOF mass spectrum of thioester conjugate of HP.

The synthesis of KBPo1 conjugate with a terminal cysteine was readily synthesized using standard Fmoc chemistry. Thus, KBPo1 was first synthesized on Rink amide MHBA resin using a microwave-assisted protocol [4]. Two copies of the AEEA spacer and a terminal cysteine residue were then coupled to the peptoid on an automated peptide synthesizer under standard peptide coupling conditions. The final product was then released from the beads with concomitant removal of side chain protection groups by treating the beads with 95% TFA/ 2.5% triisopropylsilane/ 2.5% water for two hours. The crude product obtained was purified by HPLC and analyzed by MALDI-TOF mass spectrometer. As shown in Figure 4.5, the desired product was obtained in excellent purity.

In spite of the success in preparing the corresponding peptoid and polyamide derivatives for native ligation chemistry, several efforts at the synthesis of the peptoid-polyamide conjugate ended unsuccessfully. This was unexpected as the same HP thioester was used to successfully couple to a PNA with a terminal cysteine. Therefore, a different strategy for preparing the peptoid-polyamide conjugate was pursued.

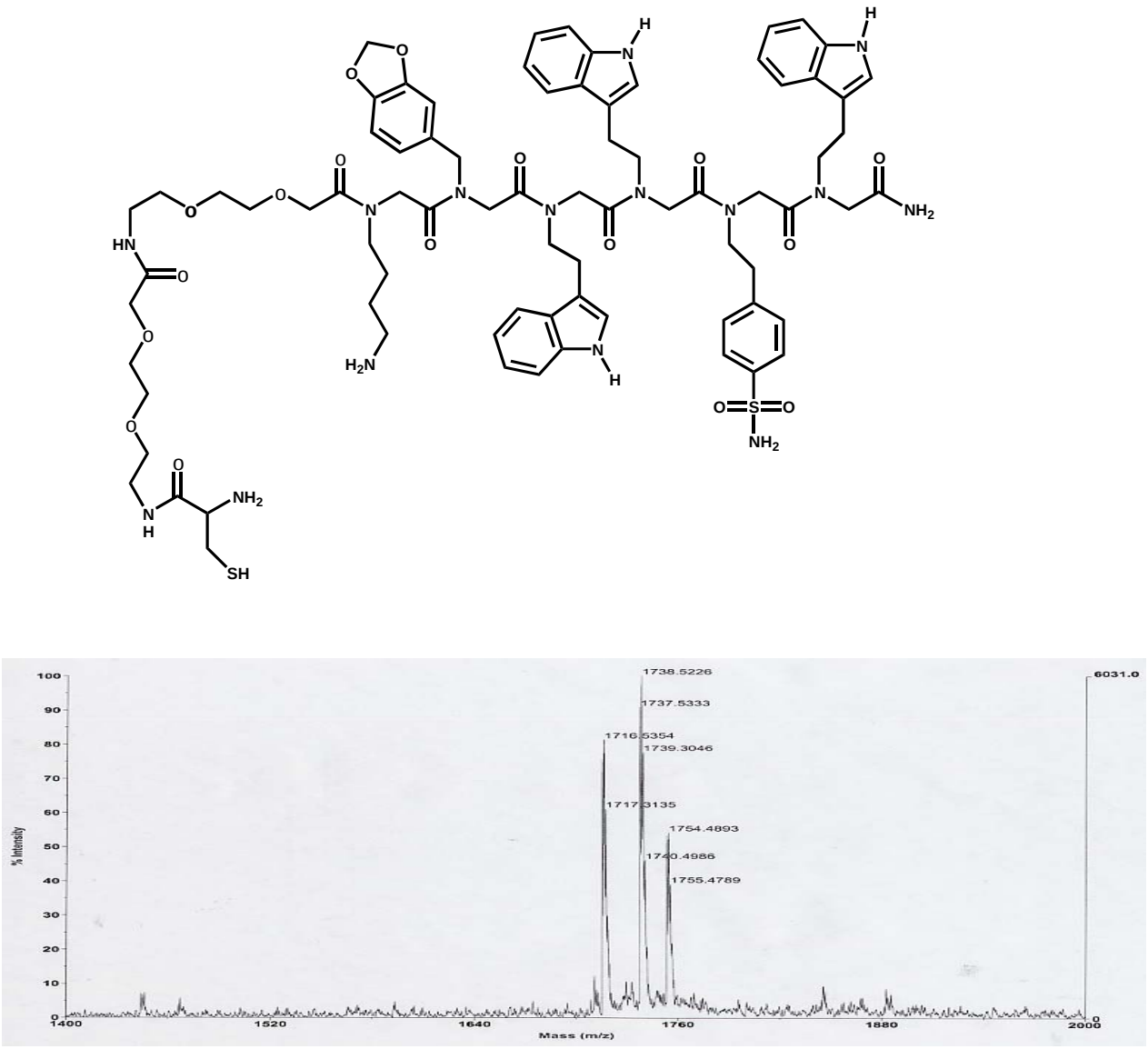
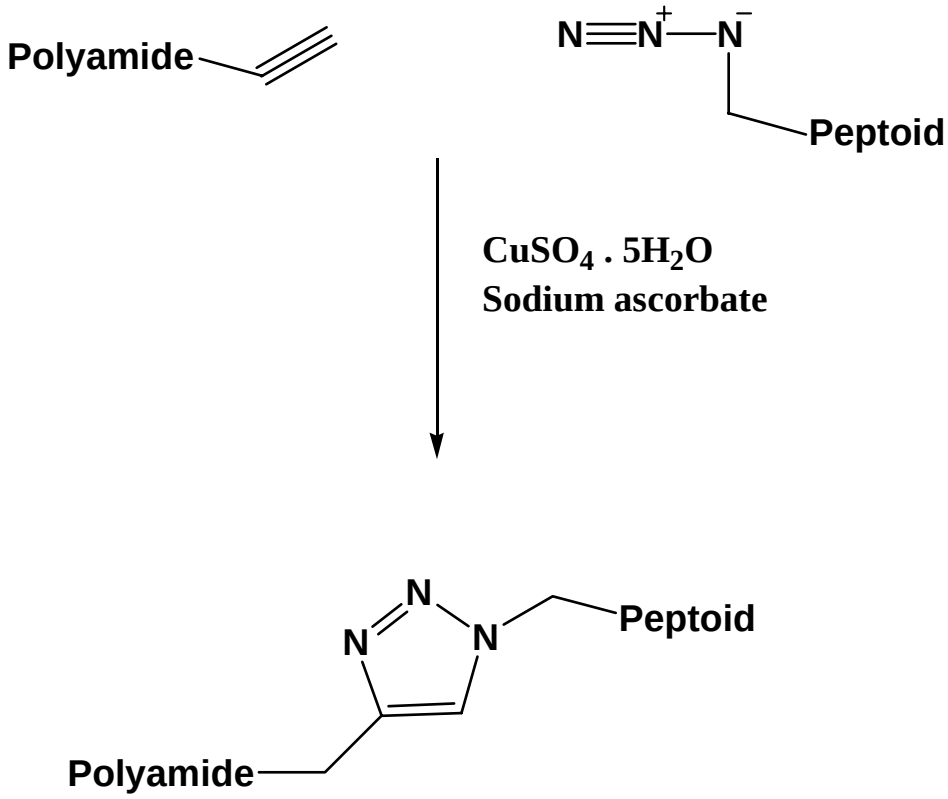


Figure 4.5: KBPo1-AEEA2-Cys.

The chemical structure (top panel) and MALDI-TOF mass spectrum (bottom panel) of the KBPo1 containing a terminal cysteine prepared for native ligation chemistry.

4.3.3.2 Click chemistry

Since introduced by Barry Sharpless, click chemistry has been widely used for forming cyclization via triazole formation of compounds containing an azide and an alkyne [11]. The reaction has been known to be extremely tolerant to the presence of other functional groups in the reacting molecules. Therefore, it was speculated that it could be a valuable tool for preparing peptoid-polyamide conjugates. However, this necessitated the preparation of suitable peptoid and polyamide derivatives.



Scheme 4.4: Click chemistry.

Proposed scheme for synthesis of HP-peptoid conjugates by click chemistry.

The synthesis of a peptoid with a terminal azide was readily accomplished by forming an amide bond between the terminal amino group of the peptoid and the carboxylate group of a commercially available poly(ethyleneglycol) (PEG) linker, which terminated in an azide. HPLC and MALDI-TOF mass spectrometry analysis indicated that the desired product was obtained in high purity (Figure 4.6).

To obtain a suitable conjugate of the polyamide for subjecting it to click chemistry, a new synthetic scheme was envisioned in which propargyl amine is used to release the polyamide from the beads at the end of the synthesis. This should conveniently install the propargyl group in the polyamide, making it ready for click chemistry. The aminolysis reaction to release the polyamide from the beads is typically carried out at 80^{o}C for 3 – 12 hours. However, the volatility of propargyl amine prevented the use of such temperatures for an extended time. Furthermore, the relatively low nucleophilicity of propargyl amine further decreased the efficiency of the aminolysis reaction. After considerable experimentation, it was found that carrying out the reaction at 60^{o}C for 24 hours affords efficient release of the polyamide from the beads with the concomitant installation of the propargyl group. The HPLC and MALDI-TOF mass spectrum of the polyamide conjugate indicated that the product obtained was indeed the desired one (Figure 4.7).

The polyamide and peptoid derivatives were then dissolved in 1:1 water / tert-butyl alcohol and mixed together in the presence of sodium ascorbate and copper (II) sulfate to facilitate their conjugation via the formation of a triazole. Unfortunately, repeated attempts have failed to isolate the desired peptoid-polyamide conjugate. Some recovery of the starting material was observed and the identity of the remaining side products could not be determined. It was speculated that the highly hydrophobic nature of KBPo1 may have caused aggregation or some other undesirable effect that impeded the chemistry.

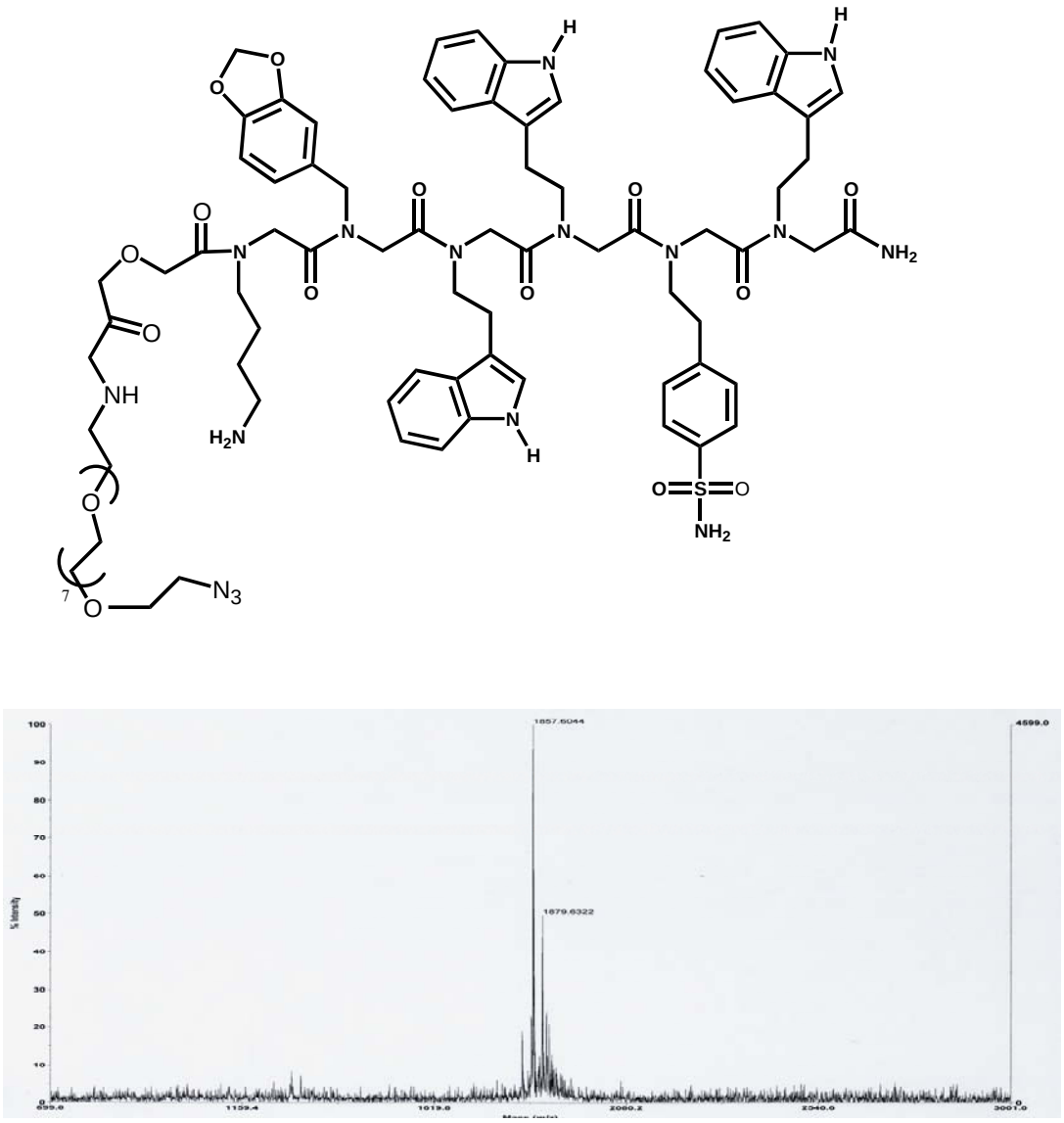


Figure 4.6: KBPo1-PEG-N_3.

Chemical structure (top panel) and MALDI-TOF mass spectrum (bottom panel) of KBPo1 terminating in an azide.

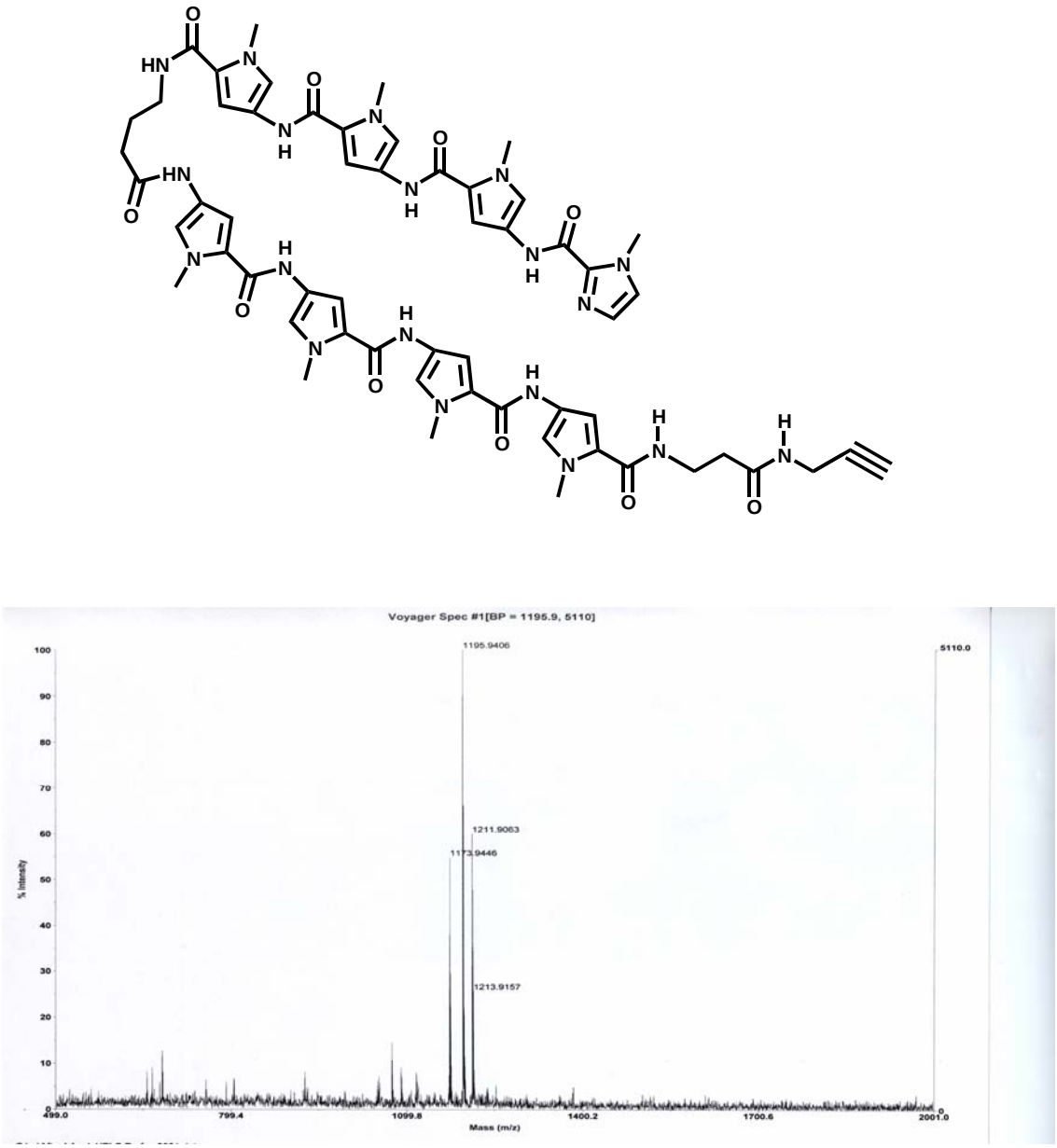


Figure 4.7: Im-Py3-γ-Py4-β-PPG.

Chemical structure (top panel) and MALDI-TOF mass spectrum (bottom panel) of HP conjugate with a terminal propargyl group.

4.4 Discussion

An attractive strategy for development of artificial transcription activators consists of joining together two synthetic molecules with activation domain-like and site specific DNA binding properties [12]. The previous chapter described a peptoid selected from a combinatorial library as a ligand for the KIX domain that activated transcription in living mammalian cells as a steroid conjugate [4]. Conjugation of this molecule to a site specific DNA-binding HP may result in an artificial transcription activator. Therefore, a previously well characterized HP, Im-Py3-γ-Py4-β-Dp_2 was synthesized and rigorously characterized by HPLC and MALDI-TOF mass spectrometry. DNA binding studies using DNase footprinting analysis revealed that the HP bound to its target DNA site, 5'-TGTTAT-3', with an apparent affinity of between 5 and 50 nM. The next step was to develop conditions for preparation of HP-peptoid conjugates. Initial efforts towards this goal were focused on taking advantage of native ligation chemistry to prepare the conjugates [7]. Thus, a thioester derivative of HP and a peptoid derivative terminating in cysteine were synthesized and characterized. However, the conjugation reaction of these derivatives was not successful. Therefore, a second strategy involving click chemistry was explored [8]. Thus, a peptoid derivative terminating in an azide was conveniently prepared by a coupling a PEG linker with a terminal azide group to the peptoid via amide bond formation. A propargyl derivative of the HP was prepared by cleaving the HP at the end of the synthesis

with neat propargyl amine. Both the conjugates were rigorously characterized by HPLC and MALDI-TOF mass spectrometry. However, efforts to obtain the HP-peptoid conjugate were, again, unsuccessful.

One interesting observation regarding native ligation was that the HP thioster was found to successfully and efficiently couple to a PNA molecule terminating in a cysteine. This result suggested that the problem with the coupling may be associated with the peptoid. In fact, KBPo1 is an extremely hydrophobic molecule. While it appeared soluble under the reaction conditions employed, the possibility that it may be aggregating cannot be ruled out, especially since the reactions were carried out in aqueous buffers. Moreover, KBPo1 was found to be transcriptionally inactive in the cell based assay described in the previous chapter, although it appeared to be cell permeable. Moreover, fluorescence polarization studies suggested that this molecule is highly promiscuous in interactions with proteins. Thus, it does not appear to be a good candidate for development of artificial transcription activators. Unfortunately, the synthetic studies reported here were conducted without the hindsight regarding the binding specificity or the transcriptional activity of this molecule. Given that native ligation has been found to be successful with other molecules, it is highly likely that the unusually high hydrophobicity of this molecule is responsible for the difficulties associated with the chemistry. If such were the case, preparation of HP-KBPo2 conjugates should

be facile given that KBPo2 is a relatively polar molecule. Future studies are necessary to address this issue.

4.5 Bibliography

1. Ansari, A.Z., and Mapp, A.K., *Modular design of artificial transcription factors.* Curr. Opin. Chem. Biol., 2003. **6**: p. 765-772.
2. Dervan, P.B., and Edelson, B.S., *Recognition of the DNA minor groove by pyrrole-imidazole polyamides.* Curr. Opin. Struct. Biol., 2003. **13**: p. 284-299.
3. Dickinson, L.A., Gulizia, R.J., Trauger, J.W., Baird, E.E., Mosier, D.E., Gottesfeld, J.M., and Dervan, P.B., *Inhibition of RNA polymerase II transcription in human cells by synthetic DNA-binding ligands.* Proc. Natl. Acad. Sci. USA, 1998. **95**: p. 12890-12895.
4. Liu, B., Alluri, P.G., Yu, P., and Kodadek, T., *A potent transactivation domain mimic with activity in living cells.* J. Amer. Chem. Soc., 2005. **127**: p. 8254-8255.
5. Baird, E.E., and Dervan, P.B., *Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids.* J. Amer. Chem. Soc., 1996. **118**: p. 6141-6146.
6. Kates, M.J., and Schauble, J.H., *Synthesis of small-medium ring thioanhydrides.* J. Heterocyclic Chem., 1995. **32**: p. 971-978.
7. Mapp, A.K., and Dervan, P.B., *Preparation of thioesters for the ligation of peptides with non-native substrates.* Tet. Lett., 2000. **41**: p. 9451-9454.

8. Rostovtsev, V.V., Green, L.G., Fokin, V.V., and Sharpless, K.B., *A stepwise Huisgen cycloaddition process: copper (I)-catalyzed regioselective "ligation" of azides and terminal alkynes.* Ang. Chem. Intl. Ed., 2002. **41**: p. 2596-2599.

9. Bremer, R.E., Baird, E.E., and Dervan, P.B., *Inhibition of major groove-binding proteins by pyrrole-imidazole polyamides with an Arg-Pro-Arg patch.* Chem. & Biol., 1998. **5**: p. 119-133.

10. Mapp, A.K., Ansari, A.Z., Ptashne, M., and Dervan, P.B., *Activation of gene expression by small molecule transcription factors.* Proc. Natl. Acad. Sci. USA, 2000. **97**: p. 3930-3935.

11. Kolb, H.C., and Sharpless, K.B., *The growing impact of click chemistry on drug discovery.* Drug Discov. Today, 2003. **8**: p. 1128-1137.

12. Denison, C., and Kodadek, T., *Small molecule-based strategies for controlling gene expression.* Chem. & Biol., 1998. **5**: p. R129-R145.

CHAPTER 5: PERSPECTIVE

5.1 Perspective and future directions

Synthetic, cell permeable molecules capable of controlling gene expression in a specific and predictable manner could have profound impact on biology and medicine. The modular nature of native transcription factors suggests that it may be possible to develop artificial transcription activators by linking together synthetic molecules with DBD- and AD-like properties. This idea has been validated in various formats in cell free systems[1-4]. In spite of impressive advances in the design and or isolation of AD- and DBD-mimetic synthetic molecules, a stable, cell permeable artificial transcription activator that can function in living cells has remained elusive. The difficulties in achieving this goal are two fold.

(1) Although several peptide-based activation domain-mimetic molecules have been designed/developed and validated inside cells, finding a cell permeable and enzymatically stable molecule with such activity has been difficult.

(2) While it has been possible to design synthetic molecules that recognize double stranded DNA in vitro in a sequence specific manner, much needs to be learned regarding the design of such molecules for manipulating gene function inside cells.

In this regard, it is quite encouraging that a peptoid isolated in this study as a ligand for the KIX domain of CBP has been shown to activate the transcription of a reporter gene in living cells in the context of a steroid conjugate[5]. However, it is obviously critical to re-validate the peptoid as a bonafide activation domain in an assay format that does not show any background transcription levels in the presence of unmodified dexamethasone alone. Several potential strategies for accomplishing this goal have been presented in chapter 3.

Once this goal is accomplished, the next step towards the construction of a synthetic transcription activator would be to design, synthesize and validate the activity of a suitable synthetic DNA-binding domain against a pre-determined DNA sequence. While this has been accomplished in vitro (as described in chapter 4), it is desirable to have a general assay that rapidly assesses the functional capability of the designed synthetic DNA-binding molecules in a complex intracellular environment. Once such a general assay is available, one could envision designing numerous synthetic DNA-binding molecules to target several sites within a gene and selecting only those molecules that show acceptable activity and specificity under physiologically relevant and challenging conditions. In fact, such an assay for testing the intracellular activity of synthetic DNA-binding molecules has been developed and reported recently[6]. It is essentially an adaptation of the cell permeability and transcription activation assay described in the previous chapter. Thus, the synthetic DNA-binding molecule

(such as a hairpin polyamide) is conjugated to modified dexamethasone and added to cells that have been transfected with a reporter gene containing HP-binding sites and a construct encoding the fusion protein GR LBD-VP16 AD. Efficient binding of the polyamide to its binding sites on the reporter gene results in the recruitment of the GR LBD-VP16 AD to the promoter by virtue of the high affinity interaction between dexamethasone and GR LBD. This, in turn, results in induction of the reporter gene activity. Thus, this assay could be used to quantitatively compare the performance of a series of polyamides under highly physiologically relevant conditions. The molecule with the best activity could then be selected for further development as an artificial transcription activator by coupling to a suitable artificial activation domain. Furthermore, the assay also provides valuable information regarding relative cell permeability of various synthetic DBDs under question, which is obviously a critical determinant.

Thus, the strategies described here could be potentially adapted to conveniently characterize the activities of both synthetic AD- and DBD- mimetic molecules under highly relevant physiological conditions. Chemical conjugation of such characterized AD- and DBD- mimetic synthetic molecules via an appropriate linker, in principle, should result in a functional artificial transcription activator that is stable in the intracellular environment. In this regard, several considerations that are critical for the success of this approach are worth mentioning. For instance, it is essential to determine the specificity of the artificial

activator-mediated effects on target gene expression. This can be addressed by employing a series of control molecules involving mismatch synthetic DBD for one or more base pairs of the target DNA sequence and a control AD that has not been selected to bind to a co-activator or relevant transcription factor. Furthermore, high concentrations of the synthetic DBD alone and AD alone should antagonize the activation caused by the artificial activator. Finally, DNA microarray analysis can be performed to indirectly measure the specificity of the artificial activator by studying the effects of the molecule on global gene expression patterns. In spite of such experiments, any observed effects on target gene expression can only be inferred to be directly caused by the synthetic transcription factor. This is especially true given that no direct methods are currently available for ascertaining the DNA-binding specificity of synthetic transcription factors in native, intracellular environment in the context of the whole genome. Furthermore, the impact of the chromatin structure on the DNA-binding properties of the synthetic DBD molecules is also poorly understood.

One possible strategy to get around this problem could be to adapt the so-called “ChIP to chip” assays for determining the DNA-binding specificity of the synthetic transcription factors on a genome wide scale[7]. Such an approach has successfully been employed as a powerful tool for analyzing native transcription factor-chromatin interactions. Thus, one could envision a similar approach for synthetic transcription factors where they are cross-linked to DNA in the native

environment, DNA sheared following cell lysis, and the synthetic transcription factor-DNA complexes selectively isolated from the cell lysate. The last step can be conveniently carried out by appending a biotin (or a similar probe) to the synthetic transcription factor during synthesis and employing the high affinity interaction between biotin and streptavidin for the purification of the DNA-synthetic transcription factor complex. Finally, the DNA can be eluted, labeled with a fluorescent probe in a ligation mediated polymerase chain reaction (LM-PCR) and analyzed by hybridization to a microarray. Such an approach will allow one to directly compare extent of specific DNA-synthetic transcription factor binding in the background of non-specific binding on a genome scale in the native environment.

Another important consideration is determination of the level of gene activation needed to elicit a given pharmacological effect. This will vary considerably depending on the gene target. For instance, if the target is an enzyme, one might notice a pharmacological effect even with relatively small change in levels of gene expression (provided that the activity of the protein is not regulated post-translationally). On the other hand, if the target gene encodes a protein (such as a structural protein) that does not have a high turn over (in terms of activity), higher levels of gene activation may be necessary to observe a significant pharmacological effect. In any case, it is likely that one will have to determine this empirically in a case by case manner. If the role of a gene in a

pathological state is already established, the level of gene activation necessary for a pharmacological effect can be inferred by comparing the expression levels of the gene in healthy and pathological states. This can be done using northern blotting analysis and the synthetic transcription factor can then be used to attempt to restore the expression levels of the target gene to normal levels.

Another issue that remains to be determined is the number of activator molecules that needs to be delivered in order to elicit the desired pharmacological effect. This would be difficult to estimate off hand, especially without any knowledge regarding the levels of activation achievable with one molecule of the activator. Nevertheless, it is possible that the level of activation achieved with a single molecule of a first generation synthetic transcription factor would be insufficient to elicit a meaningful pharmacological response. Therefore, it is desirable to develop ways to enhance the activation levels achieved with a first generation artificial transcription activator. Previous reports on transcription activators indicate that conjugating multiple copies of activation domains to a DBD dramatically improves the potency of transcription activators[8]. It is possible that such an approach might be useful in improving the potency of artificial transcription activators as well. However, the associated increase in molecular weight may compromise the cell permeability properties of the molecule limiting its practical utility. In an alternative strategy, one could envision designing several synthetic DBDs (such as hairpin polyamides) along the

promoter in close proximity so that a high local concentration of artificial transcription activator can be achieved without compromising cell permeability. This may, in turn, result in an increase in the efficiency of transcription machinery recruitment and induction levels of the target gene. One added advantage of a polyamide-based synthetic DBDs in this regard is that they bind to the minor grove of the DNA and hence unlikely to disrupt the binding of native transcription factors that may be necessary for activating gene transcription. This is because majority of native transcription factors bind to the major grove of the DNA. Similarly, one could also envision isolating ligands against more than one component of the transcription machinery and targeting them to the promoter of the target gene (via synthetic DBDs). Artificial activator-mediated recruitment of multiple components of the transcription machinery to the promoter should result in a more efficient and potent induction of the target gene. For instance, protein complexes that modify chromatin structure have been shown to play a key role in controlling gene expression. Therefore, synthetic molecules capable of recruiting such complexes to the promoter could synergistically enhance the potency of gene induction. In fact, small molecule inhibitors of histone deacetylases have already been reported in the literature. When conjugated to a synthetic DBD, such molecules could potentially play a significant role in increasing the expression levels of target genes to pharmacologically meaningful levels. Alternatively, one could also potentially set up unbiased cell- based assays to isolate synthetic

molecules that enhance the potency of first generation artificial transcription activators via potentially novel mechanisms.

Finally, target selection can play a huge role in determining the success of this technology. For instance, the local chromatin structure of target gene determines the accessibility of the promoter region to the transcription factors and hence could determine the effectiveness of the synthetic transcription factor in activating the gene. Therefore, it may be necessary to perform DNase I hypersensitivity assays to assess the accessibility of the target site on the promoter. Furthermore, it is perhaps a good idea to initially validate the utility of synthetic transcription activators against genes that have previously been activated by alternative methods such as those using zinc finger based-transcription activators. For instance, in a recent report Barbas III and co-workers employed such an approach to up-regulate the expression of endogenous γ-globulin gene in an erythroleukemia cell line (K562)[10]. Having found no DNase I hypersensitive sites they targeted sequences next to known cis-regulatory sites of the target gene promoter. They reasoned that sites on the promoter which are in close proximity to transcription factor binding sites are more likely to be situated in accessible regions of the chromatin. Given that the γ-globulin gene promoter has been well studied and even targeted for activation by zinc finger-based artificial activators, it serves as an excellent starting point for the design of synthetic transcription factors.

In addition to γ-globulin, the zinc finger-based activators have been used to up-regulate the expression of other genes as well. These genes should also serve as excellent targets for activation with synthetic transcription activators. For instance, Wolffe and coworkers designed such protein-based activators against Vascular Endothelial Growth Factor A (VEGF-A)[11]. VEGF-A is a potent angiogenic factor important in new blood vessel growth. In addition to other signals, hypoxia is a powerful inducer of VEGF-A expression. The authors first determined the accessible sites on VEGF-A promoter by performing DNase I hypersensitivity assays. A series of zinc finger-based transcription activators were then designed against both accessible and inaccessible DNA elements on the promoter. The authors went on determine the ability of the activators to activate VEGF-A transcription on a naked DNA template as well as on an endogenous chromosomal location. Not surprisingly, the efficiency of transcription was higher from accessible sites of the promoter compared to inaccessible sites. Furthermore, the pattern of activation from naked promoter and endogenous locus was quite distinct. This underscores the importance of taking chromatin structure into consideration while designing artificial transcription activators. In another example, Case and coworkers employed a similar approach to activate the expression of human erythropoietin gene[12]. Erythropoietin is a potent inducer of mammalian erythrocyte production and is used clinically in patients whose renal function has been compromised. The lessons learnt from these experiments,

especially in the selection of target sites on the promoter of the gene of interest, could be quite valuable in the design of synthetic transcription factors. Furthermore, targeting the same genes allows direct comparisons between these two techniques and increases credibility if similar results are observed in both cases. Thus, γ-globulin, VEGF-A and erythropoietin serve as excellent model systems for assessing the potential and generality of synthetic transcription factors in activating gene expression. Once the initial ground rules for the design of synthetic transcription factors are established using these well studied genes, the technology could than be applied to activate the transcription of more challenging genes that have not been previously targeted. In the following discussion, a list of genes that may serve as highly attractive targets for activation by synthetic transcription factors for potential therapeutic purposes is provided.

Certain pathological conditions such as heart failure are caused by abnormalities in complex signaling pathways making the precise understanding of molecular basis of such disorders challenging. Regardless of the identity of the culprits in such phenomena, certain master regulators that control the progression of the disease have been identified. For instance, overexpression of a class of proteins referred to as MCIPs (modulatory calcineurin-interacting proteins) has been shown to prevent cardiac hypertrophy[13]. Therefore, an artificial transcriptional activator-mediated overexpression of MCIPs may serve as an attractive strategy for the treatment of cardiac hypertrophy.

Mutations in some transcription factors have been shown to be associated with certain types of cancers. For example, majority of human cancers have been found to carry mutations in p53, particularly in the DNA-binding domain of the protein[14]. This results in the expression of a protein that is defective in DNA binding and transcriptional activation. Thus, synthetic transcription factors which can selectively control the expression of p53-responsive genes could be of utility in the treatment of various cancers.

Staf50 (Stimulated Tran-Activating Factor of 50 kDa) is an interferon-inducible factor that belong to the TRIM (tripartite motif) family of proteins, several members of which have been implicated for their role in anti-viral defense[15]. Interestingly, Staf50 has been shown to inhibit long terminal repeat (LTR)-mediated human immunodeficiency virus 1 (HIV-1) transcription. While quiescent T cells constitutively express high levels of Staf50, its levels have been shown to be highly repressed in activated T cells[16]. It has been speculated that repression of Staf50 allows HIV-1 to establish a productive infection in activated T cells. Therefore, sustained expression of staf50 may serve as an attractive strategy for inhibition of HIV-1 transcription and replication. Artificial transcription activators capable of up-regulating Staf50 may serve as attractive means for controlling HIV-1 replication.

TTD-A, a form of trichothiodystrophy characterized by defective nucleotide excision repair (NER), has been shown to be caused by deficiency in

the protein levels of TFIIH[17]. Thus, microinjection of purified TFIIH has been found to restore DNA repair function in HeLa cells following UV irradiation. Remarkably, the decreased levels of TFIIH have been found to have no bearing on transcription. While several factors could contribute towards deceased levels of TFIIH (impaired transcription, decreased mRNA or protein half life etc.), it may be possible to restore deficient protein levels to normal levels by delivering artificial transcription activators that selectively up-regulate the transcription of TFIIH polypeptides.

Leptin has in recent times received widespread attention for its potent effects on depleting adipocyte fat in animal studies[18]. Therefore, it was suggested as a possible therapeutic agent in the treatment of obesity. In spite of early promise, recent studies have indicated that high levels of leptin are ineffective in treating diet-induced obesity in animal models. That could be due, at least in part, to the observation that hyperleptinemia was associated with decreased or even undetectable levels of leptin receptor (lepr-b)[18]. Therefore, synthetic transcription factors that can upregulate the expression of the leptin receptor may serve as attractive therapeutic agents in the treatment of obesity and related disorders such as diabetes.

The above examples illustrate the potential of synthetic transcription factors as possible therapeutic agents for a wide variety of pathological conditions. While it should be possible to readily assess the effectiveness of these

molecules by measuring the extent of induction of target gene expression (by northern blot analysis), it is critical to determine if the observed increase in the expression of the target gene is significant enough to result in a therapeutic effect. Furthermore, there should be a reasonable therapeutic window such that doses of the molecule that are required to elicit the desirable pharmacological effect should not result in significant toxic or other undesirable effects. Therefore, it is desirable to have suitable animal models to assess the efficacy and safety of the molecules for the therapeutic benefit in question. As mentioned earlier, the importance of control molecules (for example, containing mismatch DBD) for assessing the specificity of the response cannot be overemphasized. The following discussion presents a few examples of animal models that could be valuable in assessing the effectiveness of artificial transcription factors designed to target some of the genes described earlier. For instance, synthetic molecules that can upregulate the expression of MCIP has been proposed as potential therapeutic agents for the treatment of cardiac hypertrophy. The effectiveness of such molecules could be tested in transgenic mice engineered to overexpress calcineurin[13]. Such mice have not only been shown to have hypertrophied hearts but are rescued by overexpressing MCIP transgene. Similarly, the effectiveness of an artificial activator designed to stimulate the expression of erythropoietin can be assessed in animal models by measuring any increase in red blood cell count in the peripheral blood. The effectiveness of synthetic activators of Staf50 gene expression could

similarly be readily assessed in tissue culture as well as well established Simian Immunodeficiency Virus (SIV) monkey models.

Finally, the use of synthetic transcription factors can perhaps be extended towards the treatment of certain genetic disorders that are otherwise untreatable. For example, some disorders such as α1-antitrypsin (AAT) deficiency are inherited in an autosomal dominant fashion[19]. This results from a mutation in one of the alleles of the AAT gene, which results in the expression of a defective protein that functions in a dominant negative fashion resulting in severe deficiency in the plasma levels of the normal protein[19]. The deficiency of AAT, which normally inhibits tissue proteases such as elastase, results in destruction of connective tissue of the lungs (due to overactive proteases) leading to emphysema. In normal individuals, AAT, after being synthesized in the liver, enters the ER and undergoes a series of post-translational modifications and achieves the correct tertiary structure in a chaperone-mediated process. It then enters the Golgi and is secreted as a 55 kDa glycoprotein. The abnormal protein, on the other hand, forms aggregates and is retained in the endoplasmic reticulum in the liver leading to liver damage and inflammation. Currently, there is no effective treatment of AAT deficiency and in severe cases can only be corrected by a liver and lung transplant. The difficulty in treating such a complex and multi-organ disorder stems from the fact any approach that increases the expression of the wild type gene will also increase the expression of the mutated gene negating

any beneficial effects. However, it would be interesting to look for single nucleotide polymorphisms (SNPs) upstream of the coding part of the gene that coevolve with the disease-causing mutation. If such polymorphisms exist, it may be possible to design synthetic transcription factors that preferentially upregulate the expression of the wild-type gene without significantly altering the expression of the mutated gene. Thus, the SNPs on the mutated gene could make it a "mismatch" for the synthetic DBD allowing the artificial transcription factor to preferentially upregulate only the expression of the 'good copy" of the gene. If such an approach turns out to be feasible and general, it may serve as a novel approach for the treatment of several other autosomal dominantly inherited diseases such as breast cancer (due to mutations in BRCA1 and 2 genes)[20].

In summary, the recent progress in the design/isolation of synthetic molecules with DBD- and AD-like functions suggests that construction of artificial transcription activators capable of controlling target gene expression may be a realistic possibility in the near future. While several hurdles remain in reaching this lofty goal, the tremendous potential of such molecules in the treatment of various human diseases makes such an endeavor exciting and worthwhile.

5.2 Bibliography

1. Mapp, A.K., Ansari, A.Z., Ptashne, M., and Dervan, P.B., *Activation of gene expression by small molecule transcription factors.* Proc. Natl. Acad. Sci. USA, 2000. **97**: p. 3930-3935.

2. Ansari, A.Z., Mapp, A.K., Nguyen, D.H., Dervan, P.B., and Ptashne, M., *Towards a minimal motif for artificial transcriptional activators.* Chem. & Biol., 2001. **8**: p. 583-592.

3. Kwon, Y., Arndt, H.-D., Mao, Q., Kawazoe, Y., Dervan, P.B., Uesugi, M., *Small molecule transcription factor mimic.* J. Amer. Chem. Soc., 2004. **126**: p. 15940-15941.

4. Liu, B., Han, Y., Ferdous, A., Corey, D.R., and Kodadek, T., *Transcription activation by a PNA-peptide chimera in a mammalian cell extract.* Chem. & Biol., 2003. **10**: p. 909-916.

5. Liu, B., Alluri, P.G., Yu, P., and Kodadek, T., *A Potent Transactivation Domain Mimic with Activity in Living Cells.* J. Amer. Chem. Soc., 2005. **127**: p. 8254-8255.

6. Liu, B., Yu, P., Alluri, P.G., and Kodadek, T., *Simple reporter gene-based assays for hairpin poly(amide) conjugate permeability and DNA-binding activity in living cells.* Mol. BioSys., 2005. **1**: p307-317.

7. Sikder, D., and Kodadek, T., *Genomic studies of transcription factor-DNA interactions.* Curr. Opin. Chem. Biol., 2005. **9**: p38-45.

8. Han, Y., and Kodadek, T., *Peptides selected to bind the Gal80 repressor are potent transcription activations domains in yeast.* J. Biol. Chem. 2000. **275**: p 14979-14984.

9. Taunton, J., Collins, J.L., and Schreiber, S.L., *Synthesis of natural and modified trapoxins, useful reagents for exploring histone deacetylase function.* J. Am. Chem. Soc. 1996. **118**: p10412-10422.

10. Graslund, T., Li, X., Magnenat, L., Popkov, M., and Barbas III, C.F., *Exploring strategies for the design of artificial transcription factors.* J. Biol. Chem. 2005. **280**: p3707-3714.

11. Liu, P-Q. et al., *Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions.* J. Biol. Chem. 2001. **276**: p11323-11334.

12. Zhang, L., Spratt, S.K., Liu, Q., Johnstone, B., Raschke, E.E., Jamieson, A.C., Rebar, E.J., Wolffe, A.P., and Case, C.C., *Synthetic zinc finger transcription factor action at an endogenous chromosomal site.* J. Biol. Chem. 2000. **275:** p33850-33860.

13. McKinsey, T.A., and Olsen, E.N., *Toward transcription therapies for the failing heart: chemical screens to modulate genes.* J. Clin. Invest. 2005. **115**: p538-546.

14. Greenblatt, M.S., Bennett, W.P., Hollstein, M., and Harris, C.C., *Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis.* Cancer Res., 1994. **54**: p. 4855-4878.

15. Tissot, C., and Mechti, N., *Molecular cloning of a new interferon-induced factor that represses human immunodeficiency virus type 1 long terminal repeat expression.* J. Biol. Chem. 1995. **270:** p14891-14898.

16. Gongora, C., Tissot, C., Cerdan, C., and Mechti, N., *The interferon-inducible Staf50 gene is downregulated during T cell costimulation with CD2 and CD28.* J. Interferon Cytokine Res. 2000. **20**: p955-961.

17. Vermeulen, W., Bergmann, E., Auriol, J., Rademakers, S., Frit, P., Appeldoorn, E., Jan, H.J., and Hoeijmakers, J.-M.E., *Sublimiting concentration of TFIIH transcription/DNA repair factor causes TTD-A trichothiodystrophy disorder.* Nat. Genet. 2000. **26**: p307-313.

18. Wang, M-Y., Orci, L., Ravazzola, M., and Unger, R.H., *Fat storage in adipocytes requires inactivation of leptin's paracrine activity: Implications for treating human obesity.* Proc. Natl. Acad. Sci. USA. 2005. **102**: 18011-18016.

19. Perlmutter, D.H., *Alpha-1 Antitypsin deficiency: diagnosis and treatment.* Clin. Liv. Dis. 2004. **8**: p839-859.

20. Palma, M., Ristori, E., Ricevuto, E., Giannini, G., and Gulino, A., *BRCA1 and BRCA2: the genetic testing and the current management options for mutation carriers.* Crit. Rec. Oncol. Hematol. 2006. **57**: p1-23.

VITAE

Prasanna G. Alluri was born in India on August 26, 1973, the son of Mrs. Subhadra Alluri and Dr. A.V.N. Raju. Prasanna graduated with a bachelor's degree in Pharmaceutical sciences from Mangalore University (India) in 1996 and received a Master's degree in Medicinal chemistry from Panjab University (India) in 1998. He entered the Biological chemistry graduate program of UT Southwestern in 1999 and joined the Kodadek laboratory in 2000.

Permanent Address: 6B-5-1, Southern Street,

Eluru - 534001, A.P., India