TOWARD DIRECTING PHOTOSENSITIZERS TO KEY VIRAL TARGETS USING COMBINATORIALLY SELECTED OLIGONUCLEOTIDES

Kestutis G. Bendinskas

A Dissertation

Submitted to the Graduate College of Bowling Green
State University in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 1996

Committee:

W. Robert Midden, Advisor

Larry Small
Graduate Faculty Representative

Neocles B. Leontis

Michael Y. Ogawa

ABSTRACT

W. Robert Midden, Advisor

The objective of this project was to find double-stranded (ds) oligonucleotides that bind specifically to a single-stranded (ss) viral target under physiological conditions and to demonstrate inactivation of the target by the selected photosensitizer-oligonucleotide conjugates.

A random library of ds DNA sequences was screened to find oligonucleotides that bind to a 20 base long origin of lambda (λ) virus replication, Ori3. This was accomplished using Combinatorial Selection and Polymerase Chain Reaction (PCR) amplification procedures. Some of the selected sequences were cloned and sequenced. Two groups of consensus sequences were identified.

To assess binding specificity, double stranded probes were prepared using sequenced clones. The percentage of individual probes that bound to the target column was measured. For further studies of binding specificity, a ss M13 virus with a target sequence insert was constructed using standard cloning techniques.

An antisense photosensitizer dihydrodioxin-DNA conjugate was prepared and its selective inactivation properties were tested. While the photosensitizer alone was not sequence specific, its conjugate selectively

damaged up to 67% of target DNA, forming 51% of conjugate-target crosslinks and cleaving 16% of target directly.

These results will be useful in the development of a new generation of potent and efficacious photochemical agents.

This work is dedicated to my dear grandfather Kostas Znaidauskas.

His way of life has taught me to love people and work hard.

ACKNOWLEDGMENTS

Most of all I wish to thank Dr. W. Robert Midden, my friend and advisor, for his support, guidance, patience, and endless optimism that made the completion of this project possible. I express my gratitude to the committee, Dr. Neocles B. Leontis, Dr. Michael Y. Ogawa, and Dr. Larry Small, for their encouragement and beneficial suggestions. My special appreciation is to Dr. Andy Ellington, Dr. George Bullerjahn, Dr. Doris J. Beck, Dr. Kathy Specht, and Dr. Eileen Underwood, for they have taught and helped me at different stages of the project. I am very grateful to our collaborators Andy Harsh, Dr. Karlyn A. Schnapp, and Dr. Marshal R. Wilson for the synthesis work they performed. I also want to pay a tribute to the Graduate College of Bowling Green State University for the constant support. I thank my friends for their interest in this work, and the opponents of the project, for they gave me challenges to overcome. Sincere appreciation is expressed to my dear wife Elena and daughter Lida who patiently supported and inspired me throughout the years of research.

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

Crucial for replication of all living systems is the correct flow of genetic information from double-stranded DNA to messenger RNA and into proteins (Figure 1) [1]. Any interference or disorder in this process may lead to dramatic changes in the viability of the system. Taking advantage of this fact, scientists try to inhibit undesirable processes that take place when the foreign genetic programs of infectious agents or mutated DNA produce a disease.

Many conventional drugs are small molecules aimed at proteins. However, an increase in protein production may dramatically diminish the effects of such treatment. Therefore, it is reasonable to aim drugs at informational molecules such as DNA or messenger RNA (mRNA) which are the templates used to make proteins [2].

Oligonucleotides and ribozymes are used to target mRNA and viral RNA. Oligonucleotides employed in this approach are called "antisense probes" because they attack single-stranded nucleic acids that are used to make normal, "sense" proteins. However, production of mRNA may also be upregulated, making the treatments less efficient [3].

Double stranded DNA is the most potential and critical target. Inactivation of certain sections of DNA can inactivate the organism. In this process, called "antigene" strategy, triplex-forming oligonucleotides (TFOs) interfere with the transcription of genes, inhibiting the production of messenger RNA [4] (Figure 1).

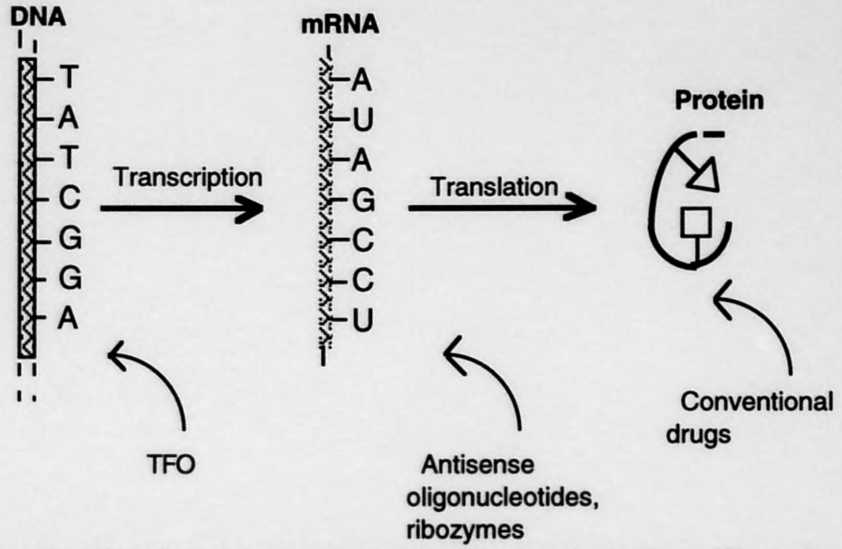


Figure 1. Novel and conventional treatment strategies. The inheritance information is transcribed from ds DNA into mRNA, and then translated into proteins. TFOs inhibit transcription. Ribozymes and antisense oligonucleotides inhibit translation. Conventional drugs attack proteins.

This project is based on the fact that DNA base sequences are highly characteristic for different organisms (virus *versus* human, for example) and are distinctive in normal and cancerous cells [1]. The challenge for scientists is to find selective agents which bind and selectively alter undesirable nucleic acids.

Unfortunately, small molecules and even restriction endonucleases are not capable of selective damage or inhibition of a single target in the genome of a reasonably large organism because they do not usually recognize more than eight bases in a row. The only species known to be capable of outstanding discriminating ability towards oligonucleotides are oligonucleotides themselves. To attack the viral genome or an oncogene without affecting the host human DNA, an attacking agent should recognize approximately 17 bases. Any agent capable of recognizing more than 16 bases should be specific [5].

Our project uses a new technique, called combinatorial selection (Chapter II), in the search for site-specific DNA sequences, that could inactivate

critical regions in viral or cancerous genomes. Combinatorial selection allows screening random libraries of compounds to select a few effectively binding species from the billions of available combinations. The attractiveness of this technique is the fact that binding might be discovered even without any previous knowledge of the binding mechanism. Many successes of this technique are reviewed later (Chapter II). We have used combinatorial selection to search for new modes of binding between double-stranded oligonucleotides and a single-stranded DNA target under physiological conditions. The problems with conventional Hoogsteen base pairing in the antigene strategy are limited recognition motifs and instability of binding complexes under physiological conditions. Many modifications of natural oligonucleotides have been explored to solve these problems, and many impressive achievements have been realized (Chapter I). We have exploited combinatorial selection to overcome these problems. To our best knowledge, this is a novelty that we are introducing independently of others in the field.

We set a general goal of selective inhibition of a critical viral target using a DNA-photosensitizer conjugate made of oligonucleotides that were combinatorially selected to bind to the viral target. Two specific goals that lead towards the realization of the general goal are:

— to find double-stranded (ds) oligonucleotides that bind specifically to a single-stranded (ss) viral target under physiological conditions and

— to demonstrate inactivation of the target by the selected photosensitizer-oligonucleotide conjugates.

The following chapters will summarize the cumulative knowledge of antisense and antigene therapies, outline the major achievements in the research employing oligonucleotide conjugates, demonstrate the utility of combinatorial selection, and present results of this project.

CHAPTER I. LITERATURE REVIEW: SEQUENCE SPECIFIC RECOGNITION AND MODIFICATION OF DNA BY OLIGONUCLEOTIDES AND OLIGONUCLEOTIDE CONJUGATES

The size of the genome in small organisms varies from a few thousand base pairs in viruses (9200 ribonucleotides in human immunodeficiency virus [6]) to a few million in bacteria (4.7×10^6 base pairs of DNA in *Escherichia coli* [7]). It reaches a few billions base pairs in higher eukaryotic systems (3×10^9 in human DNA [8]).

For an oligonucleotide to recognize a single site among so many possible sites it must have a certain minimum length. This length, N, is theoretically related to the size of the targeted genome via an equation: 4^N=genome size. Therefore, an oligonucleotide which is about 11 bases long should recognize a single target in *Escherichia coli* ($4^{11}=4.2 \times 10^6$), while a 17-mer should be specific for a single site in the human genome ($4^{17}=1.7 \times 10^{10}$) [5].

There are two recognition modes known so far for oligonucleotides binding to DNA.

The recognition of adenine by thymine or uracil via two hydrogen bonds, and recognition of cytosine by guanine via three hydrogen bonds as proposed by J. D. Watson and F. H. C. Crick in 1953 is the basis for the design of antisense probes that bind to single-stranded (ss) nucleic acids [9].

Discovered in 1957 [10], the triple-helical DNA binding motif is involved in all conventional antigene therapies. It was first described in detail by K. Hoogsteen two years after the discovery of triplexes. This so called Hoogsteen base pairing involves two hydrogen bonds between the third base of the triplet and two bases that pair using Watson-Crick hydrogen bonding (Figure 2).

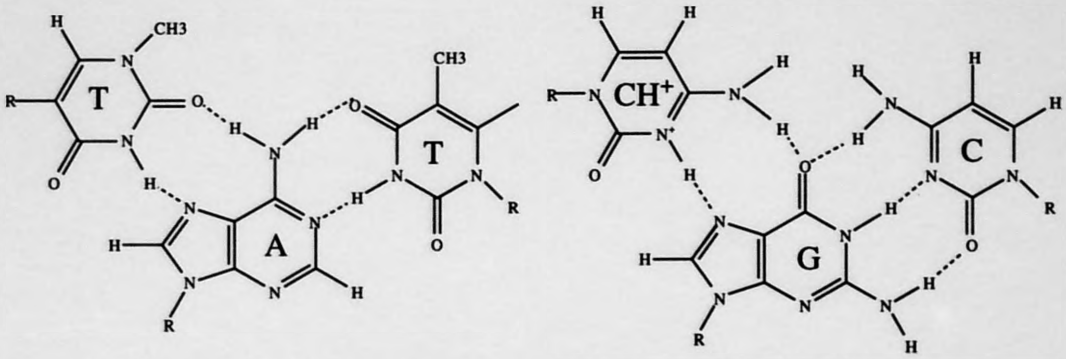


Figure 2. T×T·A and C+×G·C triplets. The far left base of each triplet comes from the third strand and is Hoogsteen bonded to the purine base of the Watson-Crick duplex.

There are only eight distinct base triplets that can be formed with the natural bases: four when the third strand is parallel with respect to the homopurine sequence of target DNA (T×T·A, C+×G·C, G×G·C, Inosine(I)×A·T) and four when it is antiparallel (first three triplets are the same, the fourth is A×A·T) [11]. The latter four are also referred to as reverse Hoogsteen base triplets. Triple helices preferentially form at homopurine·homopyrimidine sequences of target DNA [12], and this is a limitation in the recognition patterns of antigene strategy.

Antisense and antigene therapies involve control or modulation of information flow from DNA to mRNA to proteins, and are in this respect related to gene therapy. There are several human gene marking and human gene therapy trials going on [13]. The most noticeable achievement is a preliminary gene therapy trial which involved transfer of the adenosine deaminase (ADA) gene into the lymphocytes of the patients. ADA deficiency leads to high levels of 2'-deoxyadenosine in the circulation which is toxic to T and B cells and results in lethal immunodeficiency. Two patients have been treated by infusions of their own cells that have been transduced with a retroviral vector expressing human ADA. Significantly, the ADA level in circulating cells increased from

<2% to 20% of normal and the circulating lymphocyte count rose from below normal (570 μl^{-1}) with only traditional therapy into the normal range (2100 μl^{-1}) immediately after this treatment. Many challenges, such as potential helper virus production, possible oncogene activation, and limited longevity of the introduced cells, are still facing this therapy [13].

The antisense strategy has just scored a considerable achievement. A team of Colorado State University researchers, led by Barry J. Beaty, "took first steps toward genetically modifying mosquitoes to disable them from transmitting infectious viruses" [14]. They injected *Aedes aegypti* mosquitoes with Sindbis virus that produced antisense RNA directed against DEN-2 RNA. This RNA codes a dengue fever virus assembly protein. Mosquitoes are the carriers of the fever. Infected with the DEN-2 fever, mosquitoes that had been simultaneously treated with Sindbis virus could not infect other mosquitoes in 96% of cases and did not contain DEN-2 antigen in the saliva, the virus transmitting liquid from the infected mosquitoes to humans and from the infected humans to mosquitoes [15]. The researchers also disrupted replication of human encephalitis virus in *Aedes triseriatus* mosquitoes while directing antisense RNA towards LaCrosse RNA [16]. This is the first successful effort to express an exogenous gene that confers resistance to an important human pathogen [17].

The antigene strategy allowed researchers to direct a site-specific triplex-forming oligodeoxyribonucleotide to the promoter of the human *c-myc* gene. The binding of non-modified oligonucleotide to this site repressed transcription of the gene in vitro [18] and in vivo (living HeLa cells) while it did not inhibit expression of other genes, and no inhibition was observed by scrambled and complementary oligonucleotides [19].

Though very promising, both therapies, antisense and antigene, face several obstacles that must be overcome before widespread therapeutic use in humans is possible. These include cell wall permeability, nuclease sensitivity, uncertainties regarding optimum choice of a target, lack of information concerning pharmacology and pharmacokinetics [20], and limited recognition motifs (for antigene therapy only).

The efforts to overcome these challenges are tremendous. We will look over the diverse possibilities of oligonucleotide modifications, and the different reactive, transporting, and anchoring groups linked to oligonucleotides to see how they help to solve the above mentioned problems. It is reasonable to classify chemical changes in the structure of natural DNA while considering the effects of these changes on properties of oligonucleotides and the efficiency of oligonucleotides in antigene and antisense strategies. It also very important to know what changes can be introduced into oligonucleotides that bind to the target to increase binding affinity and specificity, to give them special cleaving or resistance properties once they have been selected.

Effects of Modifications of Oligonucleotides in the Phosphodiester Backbone

The negative charge of the phosphodiester linkages in nucleic acids is a barrier to the delivery of oligonucleotides to the cell through the lipophilic membrane. It also adds an unfavorable repulsive effect between phosphodiester backbones in triple-stranded DNA formation. One needs to modify the backbone to overcome these problems.

There are six groups of modified oligonucleotides: methylphosphonates, phosphotriesters, phosphorothioates, phosphoramidates, other phosphate modifications, and non-phosphate internucleoside linkages [21].

Methylphosphonates are nonionic derivatives with a PO_3-CH_3 linkage between sugars which can be easily synthesized using phosphoramidites as a substrate. At least five derivatives of methylphosphonates have been synthesized [21]. Since these compounds do not as readily assume a right-helical structure, they are less stable in duplexes than phosphodiesters with a melting temperature (T_m) about 8°C lower than that of unmodified backbones [22]. At 75μM concentration trimer and heptamer methylphosphonates were targeted to bacterial 16S rRNA, and inhibited *E. coli* (ML 308-225) protein production and colony formation by 98% [23]. Eleven-mer methylphosphonates, linked to a psoralen, inhibited, upon irradiation, the translation of rabbit globin mRNA (via cross-linking to it) in cell-free assays in reticulocyte and in wheat germ by 50-100% at 5μM oligomer concentrations. The authors stated that methylphosphonates show high resistance to nucleases. The stability of the duplex formed by an oligomer and RNA increased with oligomer chain length. Since methylphosphonates are not recognized by natural enzymes, the life-time of the oligomer was considerably increased ($T_{1/2}$=48 hours) in the serum containing cell culture medium and sufficient for use in experiments involving cells growing in culture [24].

Phosphotriesters have the general linkage formula of PO_3-O-R between sugars. At least 8 derivatives have been synthesized [21]. Phosphotriesters are very resistant to nucleases. They hybridize better than methylphosphonates because they are capable of readily adopting the right-handed helical form. Phosphotriesters are chemically more stable than oligodeoxyribonucleotides and oligoribonucleotides. Their duplexes with RNA are thermally more stable than the corresponding oligodeoxyribonucleotide-RNA duplexes due to the lack of electrostatic repulsion [25]. 2-O-allyl oligoribonucleotides were used as antisense probes targeted to an internal domain in U2 snRNP, and they

specifically inhibited the second step of pre-mRNA splicing [26]. This splicing process is the second mode of the antisense oligonucleotide action. Here, instead of a direct mRNA inactivation (I Mode), they inhibit pre-mRNA splicing (II Mode), so mRNA is not produced, and translation does not take place.

Phosphorothioates are anionic species: PO_3-S^-. Phosphorothioates have several very attractive features.

They are recognized by T7 RNA polymerase and *E. coli* polymerase as substrates. All the substitutions at the phosphorus in a backbone lead to S_p and R_p diastereomers. So there are 2^n stereoisomers of oligonucleotides in a sequence with n modifications in a backbone. On the basis of stereospecific digestion by nucleases, the authors stated that phosphorothioate-substituted RNA transcribed by T7 RNA polymerase had the same configuration as that transcribed by *E. coli* polymerase, i.e. R_p. Since they used chirally pure S_p diastereomer as a substrate, they concluded that reaction proceeded with inversion of configuration at phosphorus [27].

Oligodeoxynucleotides with a high proportion of unresolved phosphorothioate groups are almost totally resistant to digestion by snake venom phosphodiesterase (SVPD) and show increased 2-45 times resistance to nucleases P1 and S1 in comparison to that of natural diesters [21].

Phosphorothioate analogs are the only negatively charged derivatives. They, therefore, are the only ones recognized by RNase H, which identifies RNA-DNA heteroduplexes, and selectively cuts only the ribonucleotide strand. This enhances the antisense oligonucleotide action by 10-20 times [28].

Phosphorothioates are the only analogs recognized by deoxyribonucleotide transporting proteins in mammalian cells [29]. However, they are transported into HL 60 cells 7 times slower than the nonmodified oligonucleotides [30].

Because of many attractive properties, phosphorothioates were used in studies of autolytic processing of RNA, interactions with proteins, oligonucleotide directed mutagenesis, and as antiviral agents [28].

The most striking result is their anti-HIV-1 activity [31]. At 1 μM concentration, phosphorothioates exhibited potent antiviral activity and inhibited *de novo* viral DNA (not protein) synthesis but not by the expected antisense mechanism. The noncomplementary analog, S-dC_{14}, but neither the antisense nor sense-analogs exhibited the highest potential. The action of S-dC_n was length dependent (I(n=24)>I(n=14)>I(n=5)), and synergistically increased the anti-HIV activity of 2', 3'-dideoxyadenosine. The mode of binding is not known yet. It seems to involve inhibition of replication [31].

Phosphoramidates have the greatest opportunity for structural variation: PO_3-NR_2. At least 10 modifications have been made [21]. Phosphoramidates form less stable complexes with mRNA than phosphodiesters, and even less stable than methylphosphonates. T_m's for these complexes are 10°C lower for secondary amines, and 20°C lower for the tertiary ones in comparison to those of unmodified oligonucleotide complexes. T_m's are independent of salt concentration because there is no electrostatic repulsion between the complementary strands which would be masked by the higher salt concentrations in normal duplexes [22].

The phosphoramidate linkages have been shown to be resistant to SVPD and RNase H. For this backbone, it is easy to introduce intercalating or reactive agents into oligonucleotides. Another attractive feature of phosphoramidates is their ready conversion by hydrolysis with 10% isoamyl nitrate at 45°C to phosphodiesters for subsequent characterization by usual means [22].

In non-phosphate internucleoside linkages the entire sugar phosphate backbone is replaced. Three modifications appeared to have very interesting properties [32,33]. Neutral, non-diastereomeric pharmacetal (-O-CH_2-O-CH_2- between sugars), and thiopharmacetal (-O-CH_2-S-CH_2-) linkages were introduced into oligonucleotides [32]. They are capable of forming triple-helical DNA at physiological salt conditions. Pharmacetals form these structures better than R,S-methylphosphonates, thiopharmacetals, phosphodiesters, and R,S-methoxy ethyl amidates (in the order of decreasing triplex stability).

A polyamide nucleic acid (PNA) was designed by computer as a model with an achiral, chargeless polyamide backbone (-CO-CH_2-N(CO-Base)-CH_2-CH_2-NH-) [33], which resembles very closely the phosphodiester structure. When it was synthesized, its hybrid with a single DNA appeared to be more stable than normal double-stranded B-DNA. The PNA-DNA affinity is so high that a duplex is reformed only after denaturation in 80% formamide. T_{10} PNA bound to double stranded $(AT)_{32}$ DNA but did not form the expected classical triplex. The 300 nm radiation photocleavage, diazoacridine photofootprint, and cleavage by Staphylococcus and S1 nucleases were consistent only with a proposed T-strand displacement binding mode, which is initiated through inherent DNA breathing, and proceeds in a zipperlike fashion [33]. PNA possess antisense and antigene properties. Binding to either a T_{10} PNA or mixed sequence 15-mer PNA to the transcribed strand of a G-free transcription cassette caused 90 to 100 percent site-specific termination of Pol II transcription elongation [34].

Effects of Modifications of Bases in Oligonucleotides

Modification of bases as well as modifications of the backbone are designed to enhance resistance to nucleases, increase binding affinity and specificity at physiological conditions, and extend the range of recognition sequences.

One of the most often used modifications is the methylation of the 5-carbon in cytosine (^{me}C). Oligonucleotides, containing this modified base, exhibit substantially more efficient binding to duplex DNA than oligonucleotides containing cytosine [35]. Triplex formation is a highly discriminating process. After searching through almost 14 mega base pairs of yeast DNA, the Fe-EDTA linked ^{me}C containing oligonucleotide bound to and cleaved the 20 base pair (bp) long target site while leaving the other chromosomes largely intact [36].

The deoxyribonucleoside called P1 (1-(2-deoxy-β-D-ribofuranosyl)-3-methyl-5-amino-1*H*-pyrazolo[4,3-d]pyrimidin-7-one) bound to GC base pairs within the triple-helical motif as selectively and strongly as CH^+ but over an extended pH range [37]. N6-methyl-8-oxo-2'-deoxyadenosine (M) was found to be even more efficient as a TFO. M allows triplex formation at 37°C, pH 7.6, and physiological salt solution because it does not need to be protonated to form two hydrogen bonds with G [38].

An extremely popular modification is α-oligonucleotide in which bases are bound from the α side of sugar rings, while the β side is the natural side of binding. It is resistant to exonucleases and RNase H, as most modified backbone oligonucleotides are. It hybridizes with β-RNA and DNA stronger than β strands by themselves [39]. Therefore α-oligonucleotides have been used in many conjugates carrying photocrosslinking groups [40-42].

It is useful to know that natural phosphodiesters are resistant to exonucleases if they have "capping" groups at the ends of oligonucleotides A couple of modified linkages or bases at the 3'- and 5'-ends protect against rapid degradation. Bulky groups also preempt exonuclease attack [43].

Effects of Intercalating, Transporting, and Reactive Groups in Conjugates

Detailed synthesis, purification, and analysis of some conjugates is reviewed in [44]. It usually involves the introduction into the ligand structure of a functional group such as a hydroxyl or halogeno alkyl *via* a linker. Then these compounds are linked to oligonucleotides using an S-alkylation reaction although other synthetic schemes are also possible [45,46]. We will examine only the diversity and efficiency of the attached groups.

Intercalating groups help to increase the binding of the conjugates to their targets. Probably the most used stabilizing agent is acridine (Acr) [42, 44, 46-48]. It is usually linked to an oligonucleotide through a polymethylene bridge, and stabilizes the complex. The stabilization effect depends on the length of the polymethylene fragment, with a pentamethylene stretch giving maximal effect. 3', 5'-di-Acr derivative is not more stable than either 5'- or 3'-mono-Acr derivatives. A positive charge and aromaticity of acridine are the factors responsible for the duplex and triplex stabilization [46]. Oligomers bearing acridine at the 3'-end and Fe-EDTA at the 5'-end were targeted to a 27 bp long DNA fragment. The cleavages induced by Acr-linked and Acr-lacking oligonucleotides were compared, and revealed a 3- to 7-fold higher efficiency of the Acr-containing reagent due its intercalation at 0-10°C [48].

Phenazinium (Phn) derivatives have different, more stabilizing properties. The coupling of N-(2-hydroxyethyl)phenazinium to the 5'-terminal

phosphate group is more advantageous than to the 3' end: $\Delta G_{binding}(37^oC)$=-3.59±0.04 kcal/mol compared to -2.06±0.04 kcal/mol for 3'-Phn derivatives. The elongation of the linker from a dimethylene to a heptamethylene leads to destabilization of the complex. The complementary structure formed by the 3',5'-di-Phn derivative of the heptanucleotide was found to be the most stable among all the duplexes investigated. Relative to the unmodified complex, the free energy was -4.96±0.04 kcal/mol. The association constant (K_{ass}) of this modified complex at 37°C is $9.5 \cdot 10^6$ M^{-1}, whereas the analogous value for the unmodified complex is only $3 \cdot 10^3$ M^{-1} [49].

A benzo[e]pyridoindole derivative (BePI) strongly stabilizes triplex by binding to the major grove of duplex DNA. K_{ass} for BePI binding in triplex formation at 37°C is $1.3 \cdot 10^6$ M^{-1} (ΔG=-8.8 kcal/mol). The binding of BePI provides an additional free energy of -5.9 kcal/mol. Ethidium bromide, for example, has K_{ass} $0.8 \cdot 10^6$ M^{-1} to the triplex, and is displaced by BePI [50].

Some stabilizing effects by oxalopyridocarbozole, 9-amino-ellipticine, aflatoxin B, daunomycin, and psoralen also have been reported [46].

Transporting groups help oligonucleotides to enter cells. Poly(L)lysine (PLL) conjugates have 5-10 times higher uptake into L929 cells in comparison to that of oligonucleotides without these 14,000 molecular weight polymers. PLL stabilizes conjugates against exonucleases. As measured by fluorescence, the uptake was maximum at 4-6 hours. Poly(D)lysine was degraded by proteases while PLL was not. It is astonishing that concentrations of oligonucleotides necessary to inhibit vesicular stomatitis virus (VSV) by 2 logs in the case of the PLL conjugates dropped from 50-100μM range to the 0.1-1.0μM range. The mechanism of the uptake involves an endocytic pathway. Unfortunately, the uptake of PLL conjugates was very specific to this cell line [43].

Cholesterol also increases the oligonucleotide's penetration into the cell. It was mentioned above that 1 μM concentration of S-dC_{14} showed potential anti-HIV activity [31]. The cholesteryl-conjugated oligomers were able to inhibit HIV-1 100% at 0.25 μM concentrations in *de novo* assays and 80-100% (depending on monitoring) in postinfectional assays, when oligomer is added 24 hours after viral infection [51].

Streptavidin substituted with 12 mannose residues increases 20 times the intracellular concentration of biotinilated dodecakis (α-deoxythymidilate) in macrophages by comparison to that of free oligonucleotide [52].

A range of polycations form interpolyelectrolyte complexes (IPECs) with DNA [53]. Although not being true conjugates, IPECs enhance DNA uptake into cells and increase its transfection activity in both prokaryotic and eukaryotic cells. For example, polyvinylpyridinium salts increased transformation of *Bacillus subtilis* cells more than 10 times [54]. Lipospermines facilitated transfection of melanotrope cells with pCAT-4XB plasmid to the levels that are comparable to the calcium precipitation or lipofection methods while no significant cytotoxic effects were observed [55].

The enhanced uptake by the cells is also observed for oligonucleotide derivatives bearing octyl, dodecyl, and octadodecyl residues, resulting in a 3-, 4-, and 10-fold increase, respectively to the unmodified oligonucleotides [29].

It might be expected that any of these transporting groups will be anchored in a cell's wall. An alternative approach to the delivery of oligonucleotides to the cells is the application of liposomes, viral envelopes, and "erythrocyte ghosts." The authors [29] developed an efficient technique providing a 25-30% yield of entrapment of oligonucleotide derivatives in reconstituted envelopes of Sendai virus (RSVE). This technique resulted in an 80-fold increase of the binding of conjugates to the Krebs 2 Ascites carcinoma

cells and yielded a 26-fold increase of the poly(A)+RNA modification inside the cells. "Erythrocyte ghosts" also delivered alkylating agents to the cells very efficiently. Binding to the cells increased 100 times and the modification of poly(A)+RNA increased 66 times. It should be noticed that the ratio of the concentration of a conjugate in a cell and outside the cell was 0.6, 42.5, and 52.5 for the oligonucleotide free, in RSVE, and in "erythrocyte ghosts" respectively [29].

DNA damaging reagents are often conjugated with antisense and antigene probes. Alkylating agents coupled to oligonucleotides give them a principal advantage in the chemical modification of nucleic acids. They irreversibly crosslink a conjugate to the target. These agents were systematically studied in Russia [29,46]. Alkylation proceeds *via* the formation of a highly reactive ethyleneimmonium cation as an intermediate, the rate limiting stage of the overall process (S_N1 mechanism). In a complementary complex with a target, 95-99% of the cations formed react with the bases adjacent to the alkylating group. If these reagents are not part of the complex, the cations are distributed among all nucleophiles of the reaction mixture, and the extent of modification of the specific target is several orders of magnitude lower. Alkylating reagents mostly affect the nucleophilic centers of RNA and DNA, primarily guanine residues [46]. Up to 70% of oligonucleotides are released from the cells. The rest of the material seems to be bound irreversibly, and in 2-4 hours degrades inside the cell. When mice were repeatedly intraperitoneally injected with oligonucleotide derivatives with alkylating groups targeted against ns3 and ns5 genes of tick-borne encephalitis virus (TBEV), one third of them demonstrated resistance to TBEV, and developed specific immunity toward repeated infection. Control groups of mice which received

physiological salt solution and noncomplementary antisense oligonucleotide didn't survive the infection [29].

Another alkylating probe, DNA methyl thioether, has been shown to be able to cleave ϕX174 DNA with a 9% yield [56].

Oxidative metal-chelating groups attached to the oligonucleotides are able to cleave selectively both single and double stranded DNA. Fe-EDTA induce two regions of cutting sites, each seven nucleotides wide. This reflects an extended conformation of the linker in two different orientations which brings the OH·-generating chelate either close to the single stranded region outside the complementary sequence or within the major groove of the mini-double helix [48].

The Fe-EDTA-$T_4C^+T_6C^+T_4C^+$T-3' oligonucleotide was directed to a single purine-rich cluster at λ's origin of replication. 0.8μM of the conjugate was mixed with 4μM λ DNA at pH 7.0 and 24°C. The cleavage was initiated by 4mM dithiothreitol (DTT). λ double strand DNA was cleaved with 25% efficiency in a buffered solution as well as in 1% low melting agarose. The authors estimated that all the secondary sites (having one or two mismatches) were cleaved at 0°C at least 25 times less efficiently than the primary site [47].

Another reasonably good result with Fe-EDTA conjugates is the 6% scission yield for double stranded DNA in a 340 kbp chromosome III of *S. cerevisiae* at an engineered homopurine target site [36].

Phenanthroline-copper chelate was covalently attached to the 5' end of an 11 bp long oligonucleotide and targeted to the Simian virus SV-40 DNA which contains a single target site for this conjugate. In the presence of copper ions and reducing agents, a single specific double strand cleavage site was observed at 20°C. The efficiency of the cleavage was >70% at pH 7.4. Secondary cleavage sites were observed only at low temperature. The site of

cleavage in this reaction was only 4 nucleosides long, much shorter than in the case of Fe-EDTA derivatives [57].

In the presence of Fe-porphyrin derivatives and oxidizing or reducing agents, the target DNA is efficiently cleaved. Methylpyroporphyrin linked to an oligonucleotide in a mild reducing environment at 10°C produces 1-2% cleavage, and in an oxidizing environment up to 15%. A single base is the target for more than 50% of the damage [58]. In the presence of dithiothreitol (DTT) 37% cleavage at 10°C is achieved. Upon increasing temperature, the reaction yield decreases in agreement with the dissociation of the complex. It was noticed that the porphyrin ring was converted in 2-3 hours to a single unidentified product at 7°C in DTT solution [59]. The yield of hemin-pN_{14} is 20% in the presence of N-methylimidazole, and 40-50% in the presence of 1-hydroxybenzotriazole. The derivatives react efficiently and in a sequence-specific manner with a 303 bp target DNA. They induce the crosslinking reaction, which is sensitive to piperidine treatment. The total yield of cleavage for hemin-pN_{14} is 52%. For deuterohemin-pN_{16}, it is 60%. Five percent of the latter reaction yield is a direct cleavage of DNA by the reagents [60].

The biggest family of agents attached to oligonucleotides is the group of photoreactive agents. However, light is heavily absorbed by tissues in a wavelength-dependable manner. 10% of light, for example, penetrates only 1.7 mm of human skin at 700 nm light, 0.34 mm at 450 nm light, and 0.14 mm at 350 nm light [61]. Thus, treatment with these reagents is limited to tissue surfaces.

Photoadducts of psoralens (Pso) and DNA have been known since 1970. Psoralens are able to form interstrand crosslinks or monoadducts in double stranded DNA [62]. Pso-pN_{13} was covalently linked to target DNA and irreversibly inhibited *bla* gene transcription in *E. coli in vitro* [63]. Psoralen-methylphosphonate conjugate crosslinked with rabbit globin mRNA 50-100% at

oligomer concentrations of 5 μM [24]. Pso-pN_{16} inhibition of Dra I digestion of HIV-I DNA was 95% at 7 μM oligonucleotide concentration [64,65].

Investigation of porphyrin-DNA interactions has been stimulated by use of hematoporphyrin derivatives in photodynamic therapy (PDT) of malignant tumors. Different modes of binding, the mechanism of DNA cleavage, and the photochemical, spectroscopic, kinetic, and pharmacokinetic properties of porphyrins have been thoroughly studied [66-69]. Porphyrins linked to oligonucleotides produce various types of photodamage on a complementary target DNA. This includes oxidation of G and crosslinking of the porphyrin conjugate to the target sequence. Both reactions give 57% yield of total modification (20% oxidized G, and 37% crosslinking). Guanines located close to the porphyrin macrocycle are the most altered. No specificity is observed above the complex dissociation temperature [70].

λ phage was inactivated with 658 nm light using DNA-binding 5,10,15,20- tetrakis-(1-methyl-4-pyridyl)-21H, 23H-porphyne tetra p-tosylate [71]. This cationic commercially available compound leads to complete (>7 logs) phage inactivation at concentration of 10 μg/mL with a total light dose of only 20mJ/cm^2. Notably, this process is oxygen independent. It would be interesting to study this compound conjugate's activity and specificity towards λ phage DNA.

A benzo[e]pyridoindole derivative (BePI) exhibited the triplex-stabilizing property mentioned above [50]. Upon UV irradiation this compound at 0.1 μM concentration modified 90% of the polypyrimidine strands in a duplex of 10 nM concentration. It would be also interesting to observe the properties of its conjugate.

A few other photoactive compounds were linked to oligonucleotides to test their ability to induce irreversible modifications in a target DNA. 3-

azidoflavine [41] and p-azidophenacyl [42] conjugates were targeted to the single and double stranded 27-mers, and induced cleavage after piperidine treatment. A small fraction of ellipticine derivative photocrosslinks with double stranded DNA [72]. This case was reported as the first example of a sequence-specific artificial photoendonuclease although its site of action was 4 nucleosides long. Proflavin linked oligonucleotide was able to induce 12% cleavage after irradiation followed by piperidine treatment [40]. Uranyl UO_2^{2+} photooxidation induced cleavage at the ends of oligodeoxynucleotides. However, its conjugates failed to do so [73].

The question regarding the presence of triple-stranded structures within cells has been often raised. To test this, scientists prepared a monoclonal antibody by injecting a mouse with a triplex that contained methylated cytosine instead of cytosine in the polypyrimidine strands and, therefore, was stable at physiological conditions [74]. This antibody was later used in an immunofluorescence microscopy experiment to demonstrate the presence of triplexes in fixed mouse myeloma [74], *Chironomus tentans,* and *Drosophila melanogaster* cells [54]. The scientists concluded that triplexes are an inherent feature of the structure of eukaryotic DNA.

As an alternative to the homopurine·homopyrimidine triplexes, as we know them, are RecA-mediated three stranded structures. This protein induces the formation of a synaptic complex in which three strands of DNA are bound to RecA during genetic recombination. When protein is released, the triplex that is not limited by the narrow variety of base recognition patterns remains stable until the supercoiled DNA substrate is linearised by a restriction endonuclease [55,75]. Unfortunately these results are not supported by some scientists who offer a classical duplex explanation for this naturally occurring phenomena of homologous recombination of two chromosomes [76].

Many scientists have investigated the ability of modified oligonucleotides to provide sequence specific binding. We have chosen an alternative approach, using combinatorial selection in attempt to discover new motifs for sequence specific binding using natural DNA oligonucleotides.

CHAPTER II. LITERATURE REVIEW: COMBINATORIAL SELECTION

Combinatorial chemistry has been such an "innovative force that it sweeps through the discipline like a fast-moving brushfire through dry chaparral" [77]. This technique generates vast numbers of chemical compounds that can be screened for potential biological activity (Figure 3). The following examples should prove that no matter how unexpected and different, targets for the selection are, combinatorial selection is capable of finding individual molecules from the initial library of compounds that have high affinity and discrimination ability towards the target, provided that there are such molecules in the library and that screening conditions favor selective binding. RNA's, DNA's, peptides, carbohydrates, and, recently, small-molecule libraries have been used in combinatorial research.

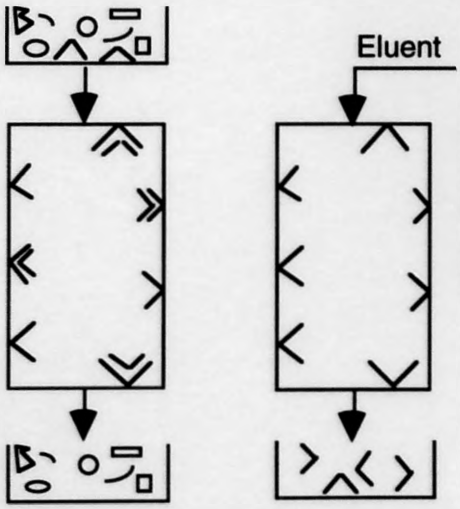


Figure 3. Combinatorial selection. A large number of chemical variants are tested for binding with the target. The most promissing compounds are isolated and identified for futher development.

First, methods for the synthesis of large numbers of random sequence RNA were developed. As early as 1977, scientists developed an RNA evolution system that selected for splicing activity [78]. But it wasn't until 1990 that efficient techniques for combinatorial selection were developed. A pool of DNA that has several random bases in-between defined sequences was prepared.

Defined sequences are needed so that this DNA can be amplified by transcription to RNA by T7 RNA polymerase. Then the pool of RNA's was subjected to selection for binding. The selected RNA's were reverse transcribed into DNA which was amplified by Polymerase Chain Reaction (PCR). This DNA was transcribed into RNA by T7 RNA polymerase for the next cycle of selection [79]. Since the process of RNA screening and library enrichment mimics the natural selection of surviving species, the technique was named "in vitro genetics" [80].

Two RNA's that tightly bind to T4 DNA polymerase (gp43) were selected from the calculated pool of 65,536 species by the research group at University of Colorado at Boulder [79]. After four rounds of selection of RNA-protein complexes that bound to nitrocellulose filters, 20 clones were sequenced. Nine of them appeared to be the wild-type sequence found in bacteriophage that is a host for gp43. Another eight clones were identical but varied from the wild type at four positions. These two RNA's had the same binding constant. Three other clones were all different and bound to the target about 20 times weaker than both consensus sequences. Still they were about 5 times better binders than a randomer [79].

A significantly larger pool (~10^{13} different sequences) of RNA's was screened to select subpopulations of molecules that bind specifically to seven different organic dyes. The binding to the columns carrying dyes increased from less than 0.1% binding at first screening to 85% after the fifth screening. That resulted in pools of 10^{2}-10^{5} different RNA's capable of binding that specific dye. Two of the pools were binding well not only to their columns but also to other columns. Sequenced pools were quite diverse but seemed to evolve similar ligand binding sites [81].

Three dyes out of seven that were used in the research described above were selected as the targets for the selection of single stranded (ss) DNA randomers. ss DNA was amplified by an asymmetric PCR. After five cycles of selection (including two negative selections eliminating sequences binding to the support) the fraction of bound DNA reached 29%. No significant sequence similarities were found among the 17 clones of DNA that bound to one of the dyes. Five out of eight clones that bound to the second dye had an 18-nucleotide consensus sequence that formed a stem-loop structure. The consensus sequence of the molecules was different from the RNA pools that bound to the same dyes. When transcribed into RNA, the consensus sequence did not bind to the column. Experiments proved that the 2' hydroxyls in the RNA interfered with binding to the ligand [82]. The authors also stated that by analogy with catalytic antibodies [83,84], selected oligonucleotides could generate new water soluble catalysts if they bound with sufficient preference to the transition state of a reaction as opposed to product or substrate binding.

An article describing a combinatorial selection of RNAs that bind to a 16 bp homopurine-homopyrimidine DNA sequence through triplex formation is closely related to our project [85]. The 25 bp ds DNA fragment was immobilized onto a thiol-Sepharose support via a disulfide bond. The randomized library of 10^{10} to 10^{12} RNAs was screened five times (plus three negative screenings). Out of seventeen clones sequenced, fourteen contained homopyrimidine sequences that were highly complementary to the target but were significantly truncated. Selection was performed in 2M NaCl, pH 5.5 buffer; elution used 50 mM Tris-HCl, pH 8.0 buffer. None of the clones had an "ideal" triplex forming sequence. Some had interior loops, some had nonstandard triplets or hairpin loops. However most of them bound to the target as specifically as the "ideal" third strand. This fact was determined when the conjugates of the selected

clones with K84C mutant staphylococcal nuclease cleaved the radioactively labeled target at the designated site [85].

Instead of conventional hybridization, triplex affinity capture was used to purify $(dT\text{-}dC)_n \cdot (dG\text{-}dA)_n$ dinucleotide repeats from a human genomic library using magnetic streptavidin-coated beads with a triplex-forming biotinylated oligonucleotide attached to them [86]. The same triplex affinity capture technique was employed in many similar projects. It was used to isolate a single copy clone from a yeast genomic library [87]. Cosmid insert DNA was recovered with a yield of up to 95% and a purity of at least 95% using the same procedure [88]. When screening was performed using an oligonucleotide synthesized over urethane linkage polymer support (avoiding the bulky streptavidin·biotin complex formation), triplex could discriminate between two sequences that differed in a single point mutation [89]. Affinity enrichment of an AluI-digested human chromosomal DNA was performed using biotinilated $(dG \cdot dA)_{17}$ in the presence of Mg^{2+}. Twelve out of fourteen isolated clones contained at least one polypurine·polypyrimidine tract. Most had an ideal 30-84 bp long $(dG \cdot dA)_n$ repeat, while some contained variants such as $(dC\text{-}dT)_{10} \cdot (dC) \cdot (dC\text{-}dT)_9$ [90].

Single stranded DNA was combinatorially selected to bind to human thrombin. After five cycles of selection, DNA aptamers (Latin "aptus"-"to fit") exhibited K_d of about 200 nM, whereas the original pool showed insignificant affinity for thrombin. Sequencing of selected clones that initially contained a randomer of 60 nucleotides displayed a highly conserved 14-17 base region. Some of the 32 selected aptamers were able to inhibit thrombin-catalyzed fibrin-clot formation in vitro using human plasma [91]. Later the structure of the consensus binding sequence was determined by NMR. It forms a unimolecular DNA quadruplex consisting of two G-quartets connected by one TGT loop and

two TT loops [92]. This example is particularly notable because human thrombin has no known nucleic acid binding function.

The optimal DNA substrate for T4 RNA ligase was selected using a 66-base long oligonucleotide with 10 bases in the middle completely randomized. Five rounds of screening resulted in a pool of DNA that was ligated 10 times more efficiently than the initial one. The majority of sequences approximated a well-defined consensus sequence [93].

ATP binding RNA's were selected from an estimated pool of 10^{10}-10^{11} random sequence molecules. After eight rounds of selection, RNA's that bound only to ATP but didn't bind to GTP, ITP, CTP, or UTP were cloned and after sequencing revealed an 11-nucleotide consensus sequence. The consensus RNA folded into a two stem-two loop structure that bound to ATP with K_d~14μM. After optimization K_d dropped to 0.7 μM, and the RNA bound ~0.7 equivalents of ATP [94].

The fact that RNA molecules exhibit an extremely high degree of ligand discrimination was demonstrated when they recognized bronchodilator theophylline 10,000 times better than caffeine, which differs from theophylline only by a methyl group at nitrogen atom N-7. The K_d was 0.1 μM in the presence of magnesium ions. NMR spectra supported a three stem, three loop structure suggested by computer modeling [95].

RNA's with dual specificity (binding both, arginine and guanosine) and RNA's binding only arginine or only guanosine were selected from a pool of RNA with 25 randomized nucleotides. The selected RNA's were identical in some regions and different in others. Researchers compared these RNA's to peptides that "seem to function similarly without being similar in primary structure" [96].

RNA's were selected to bind and inhibit reverse transcriptase of HIV-1. The consensus optimized RNA inhibited only HIV reverse transcriptase while it was ineffective for Moloney murine leukemia virus and avian myeloblastosis virus transcriptases. The initial pool containing 32 randomized nucleotides had to be enriched by combinatorial selection 9 times to get a preliminary consensus sequence [97].

RNA molecules have been selected to catalyze (like enzymes) a particular reaction. Such RNA's are named ribozymes [80, 98]. For example, a ribozyme that functions as an excellent template and as a catalyst in self-copying reaction has been derived from an estimated pool of $2x10^{13}$ different molecules [99]. The same group also isolated a ribozyme that ligated a 3',5'-phosphodiester bond with a rate 7,000,000 times faster than the uncatalyzed reaction rate [100].

Random sequences of DNA were inserted into herpes simplex virus type 1 (HSV-1) genome to make thymidine kinase mutants. Out of the library of $2x10^{6}$ mutants, 1540 were selected on special media as being able to phosphorylate dT. Two of the above clones had catalytic activity towards 3'-azido-3'-deoxythymidine (AZT); one was more active than wild type thymidine kinase [101]. Such proteins may have clinical potential because of the evolution of human gene therapy [13].

All of the methods above used solid support parallel synthesis of combinatorial libraries. These compounds were cleaved from the support and used for screening. Selected molecules had to be cloned and sequenced to be identified. Another method of synthesis, called spatially addressable synthesis has been developed [102]. Every synthesized compound is located at a defined place on a 1.6 cm^2 100 µm by 100 µm checkerboard allowing direct immediate determination of active molecules. An array of 1024 peptides was

synthesized using photolabile protecting groups and photolithography, and its interaction with a mouse monoclonal antibody directed against β-endorphin was assayed. An active peptide as well as 15 closely related peptides were identified by epifluorescence microscopy. The group also demonstrated that oligonucleotides can be synthesized using this method [102].

Another type of oligomer studied combinatorially are the peptoids, peptide-like molecules that have side chains different from natural peptides, which are not recognized by peptide-cleaving enzymes. Some selected molecules from the diverse N-(substituted) glycine peptoid library bound tightly to 7-transmembrane G-protein-coupled receptors [103].

Carbohydrates [104], benzodiazepines, mercaptoacyl prolines, and polymers also have been used to synthesize random libraries [105]. Small molecules have the advantage for drug development of being more stable and more deliverable.

As a combinatorially selected orally active drug from Eli Lilly for the treatment of central nervous system problems enters clinical trials, combinatorial chemistry becomes "...the core technology at drug discovery companies" [106].

CHAPTER III. APPROACH AND METHODS

We have chosen the bacteriophage λ and bacteria *Escherichia coli* (*E. coli*) as a model for our studies because they are relatively harmless and easy to handle. A strain of bacteria which contains λ phage within its own genome is necessary for the demonstration of the potential inactivation of retroviruses. Such a strain of *E. coli* exists, and is called λ+*E. coli*. In particular, we obtained a strain of *E. coli* from Dr. David Friedman (Department of Microbiology and Immunology, University of Michigan, Medical School, Ann Arbor, Michigan) that has a temperature activated virus. This strain (K5698=recA λ+ cI857) grows at 32°C, but when the temperature is raised to 39-42°C, virus (λ cI857) is activated and lyses the lawn bacteria (K37) forming plaques that are easy to count. This system is very attractive because it can be used to quantitate the effect of virus inhibition by the conjugates: if all virus is killed specifically, after temperature activation, bacteria will grow as if nothing happened, and no plaques will be formed.

We have selected the origin of replication as a critical target in the λ virus genome. RNA Polymerase and two regulatory proteins bind to the sequence AATCTATCACCGCAAGGGAT (Ori3) of promoter P_{rm} and operon Or3, thus regulating production of cI mRNA [107]. Inhibition of the binding of RNA Polymerase to Ori3 should stop production of mRNA, leading to the change in the viability of the virus.

The primary step in our project's realization is the combinatorial selection of specific sequence oligonucleotides which bind to the Ori3 of λ virus. Identification of ds sequences that bind at physiological conditions to ss DNA containing all four natural bases is a worthwhile goal to pursue using combinatorial technique because there is no known binding motif that could be

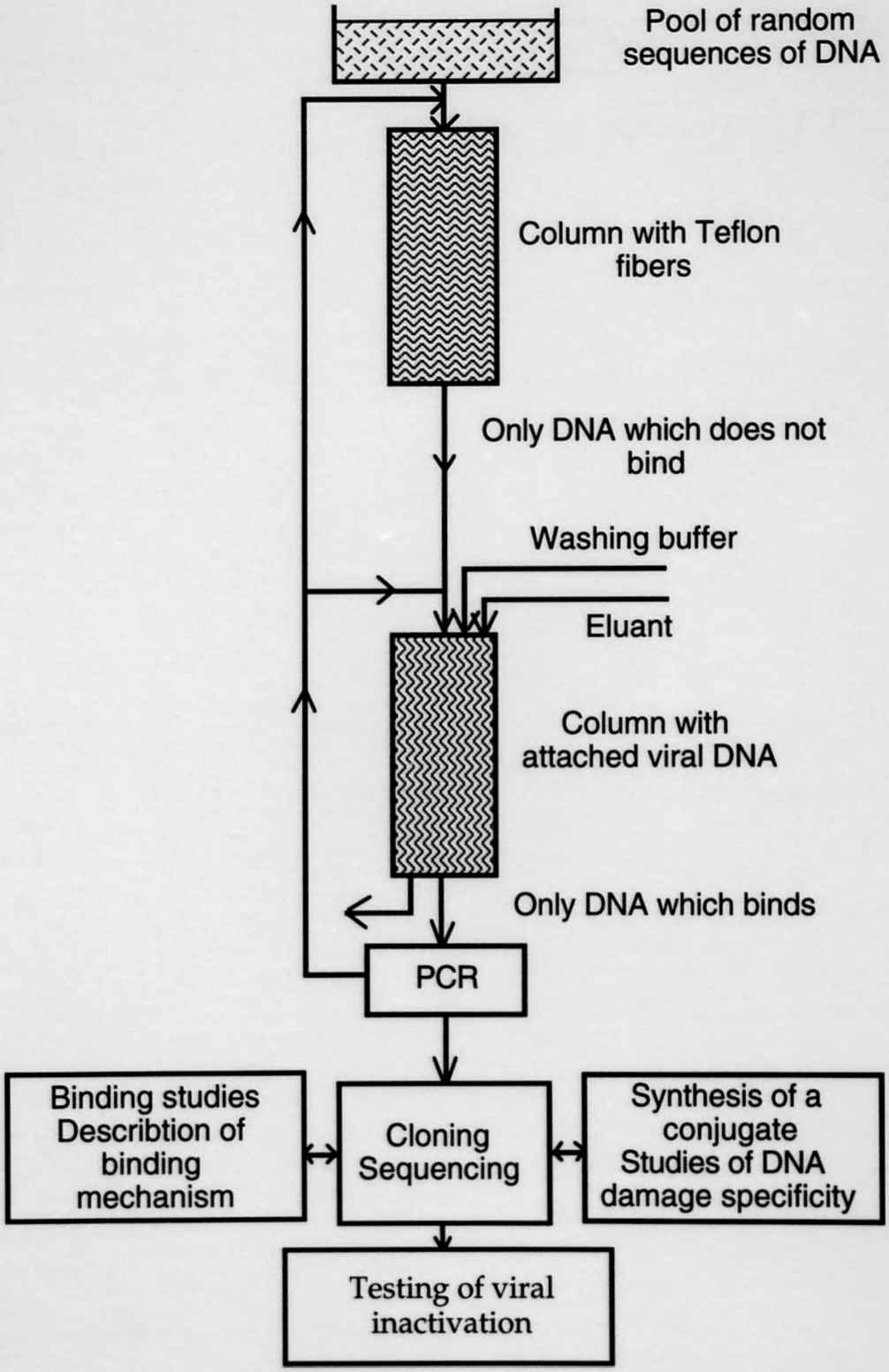


Figure 4. Schematic representation of the selection strategy.

used to design such a binding probe. It has been demonstrated by Dr. Peter G. Schultz's group that conventional triplex forming oligonucleotides are selected under high salt, low pH screening conditions using combinatorial selection [85]. Avoiding modifications of natural bases or of the backbone of oligonucleotides (Chapter I) and acknowledging the power of combinatorial selection (Chapter II), we took the next step and applied this technique to find ds-ss DNA binding under physiological conditions, where conventional triplex is not stable [12].

The following steps were performed during the combinatorial selection process (Figure 4):

A library of double-stranded random-sequence DNA was made by 10-fold Polymerase Chain Reaction (PCR) amplification of the single stranded 99 base-pair randomer GGGAGAATTCCGACCAGAAGCTT-N_{50}-CATATGTGCGTCTACATGGATCCTCA in the presence of two PCR primers (AGTAATACGACTCACTATAGGGAGAATTCCGACCAGAAG and TGAGGATCCATGTAGACGCACATA) (Figure 5). These primers enable PCR and add T7 promoter to the ds library, that can be used to produce a corresponding RNA library, if needed.

```
5'-AGTAATACGACTCACTATAGGGAGAATTCCGACCAGAAG-3'
                    5'-GGGAGAATTCCGACCAGAAGCTT-N50-CATATGTGCGTCTACATGGATCCTCA-3'
                                                  . . . . . . . .
                                                3'-ATACACGCAGATGTACCTAGGAGT-5'
```

Figure 5. Alignment of the primers and the template for PCR.

The viral target (AATCTATCACCGCAAGGGAT, the origin of replication Ori 3) was synthesized on and left attached to a Teflon fiber column. (Synthesis steps were performed in Dr. Andy Ellington's laboratory, Chemistry Department, Indiana University, Bloomington, Indiana.)

The library of random double stranded sequences in the screening buffer (CB, NCB or TCB) was applied to a column with Teflon fibers only. Only DNA

which did not bind to the column was collected. This step is called "negative screening."

DNA was then screened through the column with Ori3 attached to the Teflon fibers. Unbound DNA was washed out by the screening buffer. Only DNA which bound to the viral target was collected. It was eluted from the column using an elution buffer. This is "positive screening."

DNA was amplified by PCR.

Screening and amplification steps were repeated until at least 50% of DNA applied in the positive screening was retained on the column.

Oligonucleotides of the enriched library were inserted into an M13 vector, cloned, and sequenced. Consensus sequences were identified.

Selected oligonucleotides were prepared in large quantities (10 μg) and their binding to the target analyzed using gel shift assays and by measuring the percent of material bound to the column.

Since the target, λ virus, is double stranded but our procedure selected double stranded probes binding to a single stranded target, we prepared M13mp19 virus containing ss AATCTATCACCGCAAGGGAT target in its multiple cloning site. Inactivation of this target by ds probes should inhibit production of M13 plaques in its host *E. coli* JM101 cells.

Simultaneously, we have prepared a photosensitizer-oligonucleotide conjugate from a phosphoramidate precursor, made by Andy Harsh from Dr. Marshal Wilson's group (Chemistry Department, University of Cincinnati, Cincinnati, Ohio). Para-substituted DiHydroDioxin (PDHD)-TTTTCTTTT conjugate was tested for ability to induce sequence specific DNA damage in a Watson-Crick and triplex compiementary target.

CHAPTER IV. RESULTS AND DISCUSSION*

Synthesis of the Random Sequence Double Stranded Library

A double stranded random sequence library was prepared by 10-fold PCR amplification of 800 nanograms of a pool of 99-mers that had a 50 base pair random sequence flanked by defined sequences at the 5' and 3' ends (GGGAGAATTCCGACCAGAAGCTT-N_{50}-CATATGTGCGTCTACATGGATCCTCA). Theoretically, the 50 base random single stranded oligomer has $4^{50}=1.3x10^{30}$ possible sequences. This would require $2.1x10^{6}$ moles or $6.8x10^{7}$ kg of DNA (MW=99X325=32,175g/mole). Eight hundred nanograms corresponds to$1.5x10^{13}$ sequences. Only ~5% of synthetic template can be amplified by PCR because of the presence of chemical lesions in the majority of synthesized molecules [82]. The pool complexity is also reduced by an additional factor of three or more because PCR amplified molecules tend to mutate in the GC-rich regions [81]. This reduces the complexity of our initial double stranded pool to $2.5x10^{11}$ different sequences or 1/ $5x10^{18}$ portion of all possible sequences.

PCR amplification of the single stranded 99-mer yielded 8 micrograms of double stranded 120-mer (Figure 5). This represents $6.2x10^{13}$ molecules ($1x10^{-10}$ moles, MW=77880g/mole). Since the complexity of the pool was estimated as $2.5x10^{11}$ different sequences, about 250 copies of each double stranded molecule were present in the unscreened library.

Synthesis and Detection of the Target DNA

* For the details of a particular procedure, please, refer to Appendixes A and B.

The 20 base-long single stranded viral target (AATCTATCACCGCAAGGGAT, the origin of replication Ori 3) was synthesized on an Oligo Affinity Support 0.5μM column. It was deprotected by treatment with 30% ammonium hydroxide at 55°C overnight. The column was washed with 10 ml of 8M urea, 30 ml of water and dried under vacuum to prepare it for the screening.

To determine the amount of 20-mer attached to the column, 7.8 mg (18%) of the Teflon fibers were removed. The removed fibers were treated with 100mM sodium periodate, 100mM sodium phosphate (pH 6) solution, followed by N-propylamine:acetonitrile:water 1:2:8 (PAW) solution at 50°C to cleave DNA from the fibers. The resulting 3ml of PAW solution contained 108 μg of DNA (A_{260}=1.2AU). The remaining 502 μg (20% overall yield of the synthesis) of 20-mer represent 82.4 nmoles of DNA (MW=6091g/mole). That is an 800 fold excess over the $1x10^{-10}$ moles of double stranded library that was applied to the column for combinatorial screening.

First Round of Combinatorial Selection

For the first cycle of selection, 4 μg of the ds DNA library was diluted in 0.5 ml 1X Column Buffer (1X CB: 20mM Tris-Cl, pH 7.6, 1mM $MgCl_2$, 0.25M NaCl). The target column was washed with 15 ml (125 volumes of the column, V_{column}=120μl) of the same buffer and equilibrated with the random library at room temperature for 30 minutes. The non-binding sequences were eluted with 2 ml (17 volumes of the column) of CB buffer. 2 ml of the Elution Buffer (EB: 5mM EDTA, sodium salt, pH 7.6) were used to remove bound DNA which was later precipitated in ethanol with 40 μg of glycogen as a coprecipitant.

One tenth of DNA was used to determine the optimal number of thermal cycles (ONTC) for PCR (see Figure 6). The ONTC is a valuable number that is exponentially related to the initial amount of template material present in the mixture (the higher the ONTC, the less of the original template was present before PCR). It took 20 cycles (a theoretical amplification of $2^{20}=10^6X$) to increase the amount of DNA present in the tube to the maximum concentration achievable in PCR (1μg DNA per 100 μl reaction volume) for the first selection cycle. The remaining DNA was amplified by 20 PCR cycles, chloroform extracted and ethanol precipitated. After being dissolved in CB, it was ready for the next screening.

Lanes	1	2	3	4	5	6	7	8	9
Cycles	5	10	15	15+	20	25	30	30-	30+
OV	0.0	5.0	22.0	34.7	28.5	35.0	34.5	0.0	39.7

Figure 6. Determination of the optimal number of thermal cycles (ONTC). Lanes:1, 2, 3, 5, 6, 7- ten μl of PCR mixture after 5, 10, 15, 20, 25, and 30 cycles of PCR of the eluted DNA respectively, 4 and 9-ten μl of PCR mixture after 15 and 30 cycles of 10 ng of the known template (positive control for PCR), 8-ten μl of PCR mixture after 30 cycles without the template (negative control for PCR). Optical volume (OV) of each band was determined by scanning densitometry. After 25 PCR cycles amount of DNA that is directly proportional to OV did not increase, therefore ONTC is 25 (data of the second cycle of selection).

After the second cycle of selection, the ONTC increased to 25. The third selection also required 25 PCR cycles to amplify DNA after screening. Since ONTC reached 30 for the fourth screening, "negative" screening of the library to exclude DNA binding to the column and the DNA support material, Teflon fibers, was performed. The amplified fourth screening DNA was applied to the column with Teflon fibers only (no DNA attached). DNA that was not retained in this

column was used for the fifth screening. Negative screening was performed prior to all subsequent positive screenings. The amount of DNA applied to negative and positive screenings and DNA that is washed out or eluted from the column was monitored spectroscopically (see Figure 7).

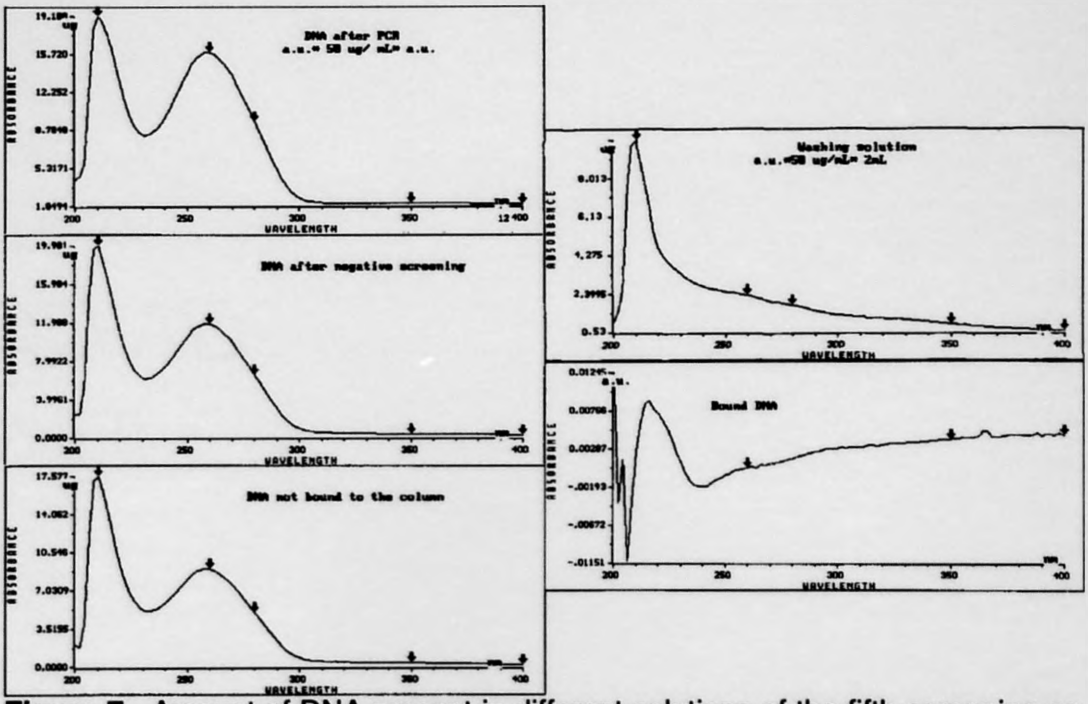


Figure 7. Amount of DNA present in different solutions of the fifth screening as determined spectroscopically ($1AU_{260}$=50µg ds DNA/ml): 15.9 µg after PCR before negative screening, 11.9 µg after negative before positive screening, 9 µg after positive screening (not bound to the column), 2.3 µg in the washing solution, and none detectable (A_{260}=0.0004) in the elution solution (bound DNA).

The coprecipitant glycogen was replaced by linearized polyacrylamide because the latter does not absorb at 260 nm. Prior to spectroscopic measurements the amount of DNA was monitored by agarose gel electrophoresis and photography of ethidium bromide stained samples (see Figure 6). To confirm that all bound DNA was eluted by elution buffer EB, the column was finally washed with 2 ml of 8M urea, and this wash was also checked spectroscopically and by PCR (after ethanol precipitation).

Six additional cycles of selection were completed in CB buffer, resulting in ten combinatorial selection cycles. The optimal number of thermal cycles never decreased to less than 25, and DNA was never detected spectroscopically in the elution buffer. These two facts indicate that the library was not being enriched. Since enriched tightly binding libraries have been isolated after an average of four to five cycles (see Chapter II), we concluded that CB is not a buffer that sufficiently favors binding of double stranded DNA to the single stranded DNA target and/or EB did not adequately elute specifically bound DNA.

Second Round of Combinatorial Selection

For the second round of selection, 4 μg of the initial ds DNA library was diluted in 0.5 ml of New Column Buffer (NCB: 20mM Tris-Cl, pH 7.3, 1mM $MgCl_2$, 3mM spermine tetrahydrochloride). This buffer differs from CB by 0.3 units of pH, and instead of 0.25M sodium chloride, it has spermine tetrahydrochloride, that is known, in contrast to monovalent ions, to favor triplex formation [108]. The time of equilibration of the library with the target was increased from 1/2 to 1 hour. The Elution Buffer (5mM EDTA, pH 7.6) was replaced with Alkaline Elution Buffer (AEB: 1mM EDTA, 10mM NaOH, pH 13.6) in order to elute all bound DNA, disrupting strong DNA-DNA interactions in the presence of spermine.

The first, second and third cycles of selection gave an ONTC of 20 (no increase, as was observed with the original selection conditions.) Although the spermine decreases absorbance of DNA and slows and blurs the DNA bands in agarose gel electrophoresis, this did not affect relative measurements because the same concentration of spermine was used in all measured solutions.

The urea elution after the third selection cycle contained some DNA that could be amplified by PCR. Moreover, it was amplified at a faster rate than AEB-eluted DNA. There is a possibility that AEB eluted some non-specifically binding DNA that was not washed out by 2 ml of NCB wash. Therefore, the number of 2 ml NCB washes was increased. 19 µg of DNA was prepared for two different fourth screenings (4A and 4B). 3X 2 ml NCB washes (6 ml = 50 volumes of the column) were used for screening 4A, and 5X 2 ml NCB washes (10 ml = 83 volumes of the column) were used for screening 4B. DNA was detected in the third wash of 4A screening. There was almost no DNA in wash five of 4B screening. Consequently, 10 ml of NCB was used to remove all non-specifically binding DNA before AEB elution. This removed most non-specifically binding DNA's. 30 thermal cycles were required to amplify DNA after the fourth screening. Fifth and sixth selection cycles had an ONTC of 30. The AEB elution always gave higher DNA yield after amplification than the fifth 2 ml wash. After the seventh selection cycle, selected sequences became significantly shorter and moved closer to the primer band on gels than the positive 120 base pair control. The ONTC for this screening was also 30.

The eighth screening brought ONTC to 10. The eluted DNA was spectroscopically detected in AEB, and could be seen on an ethidium bromide stained agarose gel before PCR amplification.

11.4 µg of DNA was applied to the negative column of the eighth screening (Figure 8). 9.9 µg was not retained and was applied to the positive column, and incubated with the target for 1 hour. 2.9 µg of DNA was detected in the solution after the positive eighth screening, indicating that about 7 µg of DNA was bound to the target. No DNA was detected spectroscopically in any of five washes, but there were at least 1.2 µg in the elution. This number is

artificially lowered because of the strong negative peak at 216 nm (A_{216}=-0.17, Figure 8).

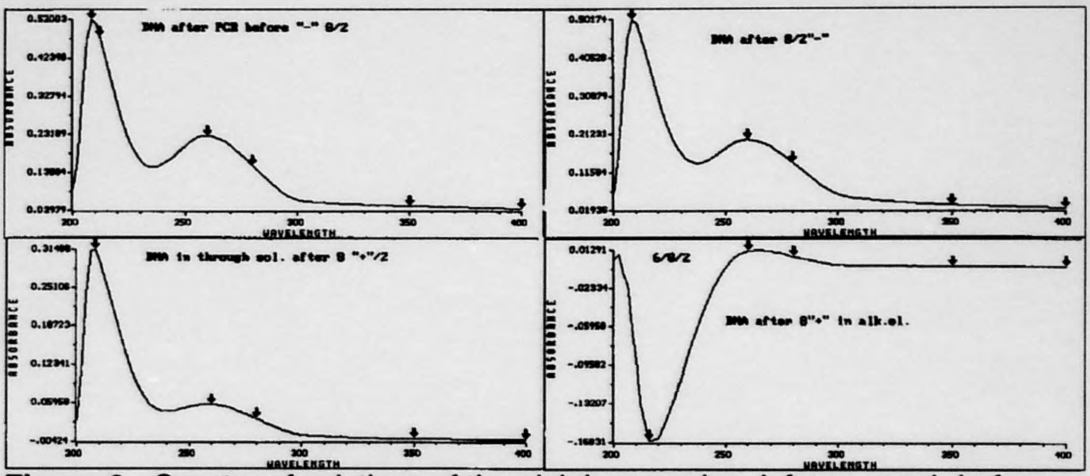


Figure 8. Spectra of solutions of the eighth screening: left top panel- before screening A_{260}=0.23, right top panel- after negative screening A_{260}=0.20, left bottom panel- after positive screening A_{260}=0.06, right bottom panel- alkaline elution solution A_{260}=0.012.

All five washes and both, AEB and urea elutions were ethanol precipitated and redissolved in 0.5 ml of water. 8 µl of DNA eluted by AEB before PCR amplification together with 6, 8, 10, and 12 µl of 18.4µg/ml DNA were loaded on the agarose gel (Figure 9).

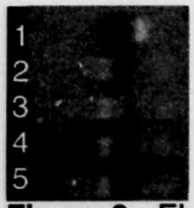


Figure 9. Eluted DNA on the ethidium bromide stained 1% agarose gel. Lanes 1- 8µl of AEB eluted DNA before PCR, lanes 2, 3, 4 and 5- six, eight, ten and twelve µl of 18.4 µg/ml of 120-mer respectively. There is more DNA in lane one, and it is of lower molecular weight than in any other lanes.

Since there is more DNA detected in the lane with AEB eluted DNA than any of the lanes with 18.4µg/ml DNA, it was estimated that the 0.5 ml of selected DNA solution contained all 7 µg of DNA that bound to the column (71% of the amount applied to the column).

Ten cycles of PCR were performed with all five washes, AEB, and urea elutions. Figure 10 shows that there is little DNA in all washes, and that 10 cycles are enough to amplify DNA from the AEB elution to the highest concentration achievable in the positive PCR control after 10 cycles. Figure 10 also demonstrates that selected DNA is shorter than original double stranded 120-mer.

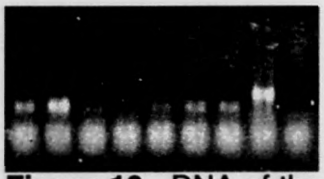


Figure 10. DNA of the eighth screening after 10 cycles of PCR. Lanes: 1-urea eluted, 2- AEB eluted, 3- fifth wash, 4- fourth wash, 5- third wash, 6- second wash, 7-first wash, 8- positive PCR control, 9- negative PCR control. Bottom band- PCR primers, upper band-amplified DNA.

AEB eluted DNA from the eighth screening of the second round (called 6/8/2) of combinatorial selection was amplified to 34μg /ml concentration. Since more than 50% of DNA applied in the positive screening was retained on the column (set criteria), this DNA was cloned and sequenced.

Cloning

The M13 cloning system was used to select individual sequences from the pool of DNA after screening. One μg of M13mp19 cloning vector and 0.5 μg of DNA from 6/8/2 selection was digested with *Bam*H I and *Eco*R I restriction endonucleases. The primers on the DNA from the selection were designed in such a way that they had one of these restriction sites on each end of the double stranded sequence after PCR. The digested segments of selected DNA were ligated into M13 virus using T4 DNA ligase. An important property of M13mp19 virus is its ability to produce a 30 amino acid long polypeptide that is capable of α-complementation with a mutated protein from the host JM101 *E.*

Coli cells, thus producing a functional β-galactosidase. This protein uses X-gal as a substrate to make a dark blue chromophore. Insertion of the DNA fragment into the region of M13 that is responsible for the production of the short polypeptide prevents α-complementation so the blue metobolite is not made. An average of 17% of plaques formed on the cell lawn were white when the ratio of the amount of insert to the amount of plasmid used in the ligation reaction was 14:1 (449 white, 2705 blue plaques). 33 clones were isolated and purified. DNA from 13 randomly selected clones was carefully purified using multiple extraction and precipitation steps for further use in sequencing reactions.

Sequencing

The Universal Cycle Primer 5'-GTTTTCCCAGTCACGACGTTGTA-3' was radioactively labeled with γ-^{35}S ATP employing T4 polynucleotide kinase and *Taq* Polymerase Cycle sequencing was performed. After gel electrophoresis and autoradiography only sequences that were positioned above the insert region in M13 could be resolved. Shorter times of electrophoresis gave very poor resolution. Therefore, a sequencing primer 5'-AGGGGGATGTGCTGCAAGGCGA-3' was designed that aligned to M13 42 bases further (-82 position, 0 at *Eco*R I restriction site) from the insert than the Universal Primer does (-40 position). This primer produced good results. However strong compressions were encountered in the insert region because selected inserts were GC rich. 7-deaza-dGTP Termination Mixes were used to sequence these regions.

Sequencing of 13 "white" clones showed that 12 of them had a single insert, while one (#33) had a double insert in a virus that was single-digested. Results of the sequencing are summarized in Figure 11.

Two consensus sequence groups of clones were identified. One consensus sequence is CTACCCCGGCCCA (clone 26), another CTGGTGGCCA (clone 24).

Clone 7 is identical to 26. Clone 33A is just missing one G to fit the consensus. Clones 11 and 12 are much shorter but have six bases that fit the first group of similar sequences.

Clones 13, 20, and 29 differ only by one base from the consensus clone 24. Clone 25 misses one base and differs by one other base from 24's sequence. Clone 18 has 7 out of its 9 bases identical to the consensus. Clone 33B has just one extra T and misses CA at the end. Clone 32 is only 60% homologous to the second consensus group but it is still more related to the second than to the first group. It is actually the only clone that conserved the expected CTT sequence at the beginning of the randomer while 86% of sequences conserved CA at the end.

An average length of the randomer was 10 bases. It is, as suggested by gel electrophoresis results, much shorter than the initial 50 bp random sequence. Only clone 27, which does not resemble any of the consensus sequences, is 22 bases long. Such an extreme shortening of all other sequences can be explained by the stronger binding of these sequences to the target and their combinatorial selection, by the formation of very tight GC-rich loops that were skipped by Taq polymerase during PCR and by the fact that polymerase favors short sequences. This phenomena of the shortening of combinatorial libraries also has been observed by other scientists [85]. The

size selection of DNA on denaturing gels is suggested to avoid this problem for all further studies.

```
Target:
5'-AATCTATCACCGCAAGGGAT-3'

M13mp19 multiple cloning site:
5'-GCCAGTGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAG-3'
3'-CGGTCACTTAAGCTCGAGCCATGGGCCCCTAGGAGATGTC-5'

ds 120 bp randomer:
5'-AGTAATACGACTCACTATAGGGAGAATTCCGACCAGAAGCTT-N50-CATATGTGCGTCTACATGGATCCTCA-3'
3'-TCATTATGCTGAGTGATATCCCTCTTAAGGCTGGTCTTCGAA-N50-GTATACACGCAGATGTACCTAGGAGT-5'

M13mp19 after cloning:
5'-GCCAGTGAATTCCGACCAGAAGCTT-N50-CATATGTGCGTCTACATGGATCCTCTAGAG-3'
3'-CGGTCACTTAAGGCTGGTCTTCGAA-N50-GTATACACGCAGATGTACCTAGGAGATCTC-5'

Individual clones:
26   5'-CTACCCCGGCCCA-3'
7    5'-CTACCCCGGCCCA-3'
33A  5'-CTACCCCG.CCCA-3'
11   5'-...TGCGGG.CCA-3'
12   5'-.....CGGG.CCA-3'

24   5'-CTGGTGGCCA-3'
13   5'-CTGCTGGCCA-3'
20   5'-CT.GTGGCCA-3'
29   5'-CGGGTGGCCA-3'
25   5'-CTG.TGGACA-3'
18   5'-CTGCTGT.CA-3'
33B  5'-CTGTGTGG..-3'
32   5'-CTTCTCGTTGGA-3'

27   5'-GAGAAGATGGTAAATAGTATCA-3'
```

Figure 11. Sequences of the target, the M13mp19 multiple cloning region, the 120 bp randomer, the M13 vector after cloning, and the individual sequenced clones. The latter are also double stranded but are shown as single stranded sequences for an easier comparison. Small boxes highlight *Bam*H I and *Eco*R I restriction sites. The larger box surrounds an expected randomer that is not defined by the PCR primers.

None of the selected clones is complementary to more than four bases of the target. This fact is significant because any complementary sequence would only prove that some of the library was single stranded when applied to the column, and that Watson-Crick duplex-forming sequences were selected.

No secondary structures except normal duplex seem possible with these double stranded probes. For example, the duplex of clone 24 has ΔG=-120 kcal/mol, while the lowest energy predicted by the DNA-mFold program (Dr. Michael Zuker, Washington University Medical School, using Dr. John SantaLucia Jr. free energies) for the loop-stem structure of the ss sequence of this clone is -12 kcal/mol (Figure 12).

These calculations, however, do not mean that the alternative folding motifs for ss DNA do not exist. For example, this same program cannot predict how the ss DNA sequence that bound to human thrombin GGTTGGNNNGGTTGG [91] folds. Only NOESY spectra revealed a hard to imagine unimolecular DNA quadruplex consisting of two G-quartets connected by two TT loops and one NNN loop [92].

Figure 12. Loop-stem structure of single stranded clone 24.

Third Round of Combinatorial Selection

Since many different clones were selected in NCB, containing 3mM spermine tetrahydrochloride, additional screening was performed with more stringent screening conditions. The concentration of spermine tetrahydrochloride was lowered to 1mM, the concentration that is found in the human body. This was done to determine whether long sequences like clone 27 will survive these stringent conditions. *Taq* Start Antibody, that has just been made available by Clontech to enhance specificity of PCR, was used to inhibit *Taq* polymerase activity during set-up of PCR reactions and to prevent it from skipping GC-rich loops that are stable at room temperature.

Third Column Buffer (TCB: 20mM Tris-Cl, pH 7.3, 1mM $MgCl_2$, 1mM $SpmCl_4$) was used for three additional combinatorial selection cycles of the library after the second round of selection. The ONTC was 20 for cycle 9, and increased to 30 for cycle 10. 0.77 µg of DNA was detected in the elution solution after cycle 11. This DNA was cloned, 36 "white" plaques selected, and 11 sequenced. Unfortunately, all eleven appeared to be mutated by a single base or sequence-shifted original M13mp19 that did not contain any inserts. Consequently, attention was shifted to binding studies of selected clones after the second round of combinatorial selection rather than continuing with additional cycles of library enrichment.

Binding Studies

Purified DNA of clones 12, 18, 20, 24, 25, 26, 27, and 32 were amplified by PCR. After an optimization of the annealing temperature (40°C), double stranded individual probes were obtained with defined sequences between

PCR primers. These probes are a small selected part of the initial random library. Each probe was increased in 10 separate PCR amplifications using the products of the previous PCR for the templates to generate sufficient quantities for further testing.

Single strand target DNA was purified and labeled with γ-^{35}S ATP for use in gel shift assays. Bromphenol blue and xylene cyanol dyes that are used in loading buffers were found to interfere and distort low molecular weight (<30) DNA bands. Both dyes were subsequently excluded from loading mixtures. Since probes were selected to bind to the target in TCB, this buffer was used in binding studies. The target was premixed with probes at various concentrations (10^{-10}-10^{-6}M probe, 10^{-9}-10^{-7}M target) in 1X TCB and incubated at 5°C for an hour. They were analyzed on an 8% polyacrylamide electrophoresis gel at 4°C with circulating 1X TCB electrode buffer. DNA did not move from the wells, most likely because of the high concentration (3mM) of spermine tetrahydrochloride present in the system. This tetravalent cation neutralizes the negative charge on DNA preventing normal gel electrophoresis. Even when TCB in the gel and electrode buffer was substituted by Tris-Borate Magnesium buffer (TBM: 90mM Tris-Borate, 10mM $MgCl_2$, pH 7.3) which did not have spermine but instead had a higher concentration of magnesium ions to stabilize the binding, the minute amounts of spermine that were present in the loaded mixtures diffused DNA bands, making them of very little value.

Since gel shift assays were apparently not feasible, binding affinity was estimated by measuring retention of individual probes on the selection column. All probes were applied directly to the column and their binding to the immobilized target DNA measured. The common positive screening procedure was performed, absorbance of solutions measured (disregarded if less than 0.03AU) and percentages of bound DNA calculated. Absorbencies of the

applied, flow-through and eluted solutions from the third screening and the eighth screening of the second round of selection (3/2 & 8/2) are shown for comparison (Table 1).

Screening 3/2 still had a complex library that did not bind well, while 8/2 demonstrated strong binding. The first consensus sequence clone 26 bound to the column well, while the much shorter clone 12 of the same consensus group exhibited 2X weaker binding. Sequences of the second consensus group showed overall better binding than sequences of the first group. This is consistent with the observation that the second group consensus sequences were found in greater number after 6/8/2 screening. The clone 24 probe has a consensus sequence of the second group. Even though it bound better than clones 18 and 20, it was a weaker binder than clones 25 and 32. Clone 27 that is not related to any of the consensus groups also bound well. These results suggest that selected sequences may bind to the target by forming a low-affinity (weak binding) long-lived (when formed, it is not easily washed away) target-probe complex.

One can also compare the amounts of DNA eluted from the column to assess binding affinity. Spectroscopically they are almost the same for all clones. We ethanol precipitated DNA of all clones from first washes and elutions, redissolved in 40 µl of TE buffer, and ran electrophoresis on 1% agarose gel. In contrast to spectroscopic measurements, we detected most DNA in the elutions of clones 24, 20, and 25, less in the elutions of the rest of the clones, and very little DNA in the washes of all clones. Assuming complete precipitation, we concluded that we can not completely rely on spectroscopic measurements, and that we need a more sensitive assay to quantify binding.

Table 1. Summary of the binding studies

Clone #	Applied	Not bound	Bound	% bound	First wash	Other washes	Elution	% in wash 1	% eluted
12	1.13	0.80	0.33	30	0.09	0.00	0.08	27	24
18	1.27	0.98	0.29	23	0.09	0.00	0.08	31	28
20	1.77	1.39	0.38	22	0.14	0.00	0.11	37	29
24	1.74	1.37	0.37	21	0.12	0.00	0.13	32	35
25	2.11	1.77	0.34	16	0.20	0.00	0.14	59	41
26	0.76	0.51	0.25	33	0.06	0.00	0.12	24	48
27	0.25	0.07	0.18	72	0.00	0.00	0.10	0.0	56
32	0.25	0.12	0.13	52	0.00	0.00	0.12	0.0	92
3/2	0.15	0.09	0.06	40	0.00	0.00	0.00	0.00	0
8 /2	0.20	0.06	0.14	70	0.00	0.00	0.14	0.0	100

One way to quantify this binding would be by measuring site specific cleavage of the target by a photosensitizer-probe conjugate at different concentrations. The association constant can also be measured in numerous other ways: using DNAse I quantitative footprint titration experiments [109], titration experiments while observing shifts in absorbance [110], and using mapping reagents such as diethyl pyrocarbonate, OsO_4, $KMnO_4$ [111], methidiumpropyl-EDTA-Fe(II), or nuclease S_1 [112].

An M13 single stranded virus containing the target sequence was prepared for the measurements of the association constant. Double stranded synthesized target was inserted into a multiple cloning site of replicative form of M13mp19. Ten "white" clones were selected, purified, and sequenced to see which ones had the inserted target. All ten had the target in an appropriate place and inserted in the right direction. ss M13 with the target insert was purified for further studies.

Cleavage of the Target DNA with the Photosensitizer-Oligonucleotide Conjugate

Before testing the photosensitizer conjugates of the selected probes, the determined photosensitizer was tested for its ability to selectively cleave target DNA when conjugated with complementary ss oligonucleotide. First, this could be used for a measurement of binding constants between the target and probes, and secondly, this would prepare an easy to control antiviral agent.

M13 was chosen as a target for these studies Its genome has many critical target sites. The sequence AAAAGAAAA (between bases 5983 and 5992 of M13, traditional numbering) was selected for several reasons. It is a potential site not only for double strand but also for triple strand formation. It is a

part of the Lac i' gene which enables blue/white plaque screening. It is a unique site in the genome. It is symmetrical allowing measurement of the direction of binding of a conjugate. It is a short sequence, and therefore any binding to it should not be very strong, mimicking the binding of our probes to lambda's Ori3.

A 30-mer 5'-TCTCGCTGGTGAAAAGAAAAACCACCCTGG-3' containing the target was synthesized, purified by PAGE and labeled with ^{32}P at the 5' end. The conjugate should have a polypyrimidine sequence 5'-TTTTCTTTT-3' to be capable of both, double and triple strand formation.

We considered many chemical groups as potential DNA damaging agents for a conjugate (Chapter 1). We were interested in testing the new class of efficient, photochemically activated reagents, recently developed by Dr. Marshal Wilson at the University of Cincinnati and Dr. Karlyn Shnapp at the University of Northern Kentucky. This class of photosensitizers is based on the DiHydroDioxin (DHD) structure. They are activated by 356 nm light (Figure 13) and effectively nick and cleave DNA [113]. DHD's are masked ortho-quinones: they can be made from and they produce PhenAnthreneQuinone (PAQ) when irradiated with 356 nm light. This release of parent PAQ is believed to be responsible for DNA damage and cleavage.

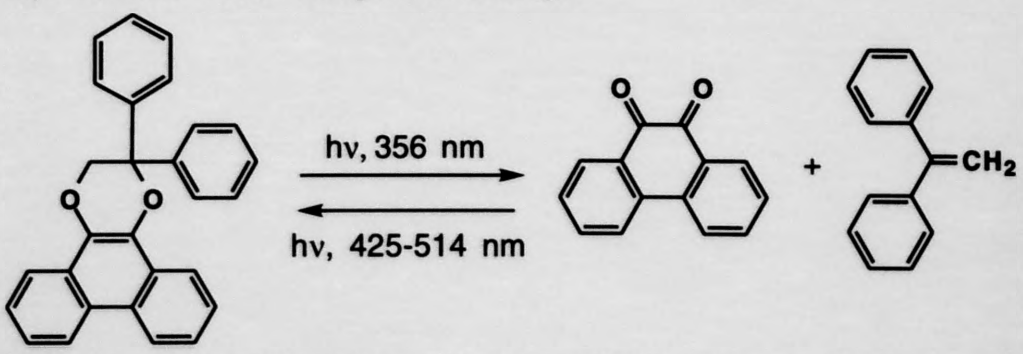


Figure 13. Reverse Schönberg Reaction. DHD is photolysed by 356 nm light forming 9,10-PAQ and an olefin.

Several DHD analogs carrying ionic groups to make them soluble in water have been prepared. DHD's with cationic groups on the olefinic moiety bind to DNA with typical binding constants $K_{ass}=2.5x10^4$ M^{-1}. They show hyperchromism and a 3-7 nm bathochromic shift in the presence of DNA which is consistent with intercalation and groove binding. Cationic DHD's efficiently relax supercoiled ΦX 174 FR I DNA upon irradiation with low doses of 365 nm light and inactivated up to 4 logs of *Vaccinia* virus. Results of transient spectroscopy, quenching studies of the DNA cleavage with dithiothreitol (DTT) and experiments of irradiation of DHD's in the presence of cytochrome-c and superoxide dismutase support a mechanism of DNA damage which involves release of excited PAQ that is capable of hydrogen abstraction or electron transfer and the involvement of radicals and superoxide [113]. No information was available about sequence specificity of DHD's. Because of the small size limiting any recognition patterns, little if any specificity was expected. Because of the relatively low DNA binding affinity (10 times lower than ethidium bromide, 1000 time lower than tripyridyl ruthenium) DHD's are perfect candidates for conjugates since they should allow sequence-specific recognition by the targeting oligonucleotide, increasing its binding slightly, but not so much as to overwhelm site-specificity.

We designed a conjugate that could be prepared by automated synthesis techniques. It should have a sequence recognizing $\overline{TTTTCTTTT}$ oligonucleotide, as well as a polylinker such as $(CH_2)_6$ (to give the photosensitizer some degree of freedom to interact with the target), and the photosensitizer DHD.

The cyanoethyl N,N-diisopropyl phosphoramidite of the para-substituted DHD derivative (PDHD) with a polylinker was prepared for the synthesis of a conjugate. The PDHD dication was made for comparison of the effects of the

free photosensitizer versus the effects of the conjugate (Figure 14). Both of these compounds were designed, prepared and characterized by Andy Harsh using organic chemistry techniques developed in Dr. Marshal Wilson's laboratory (Chemistry Department, University of Cincinnati) [114].

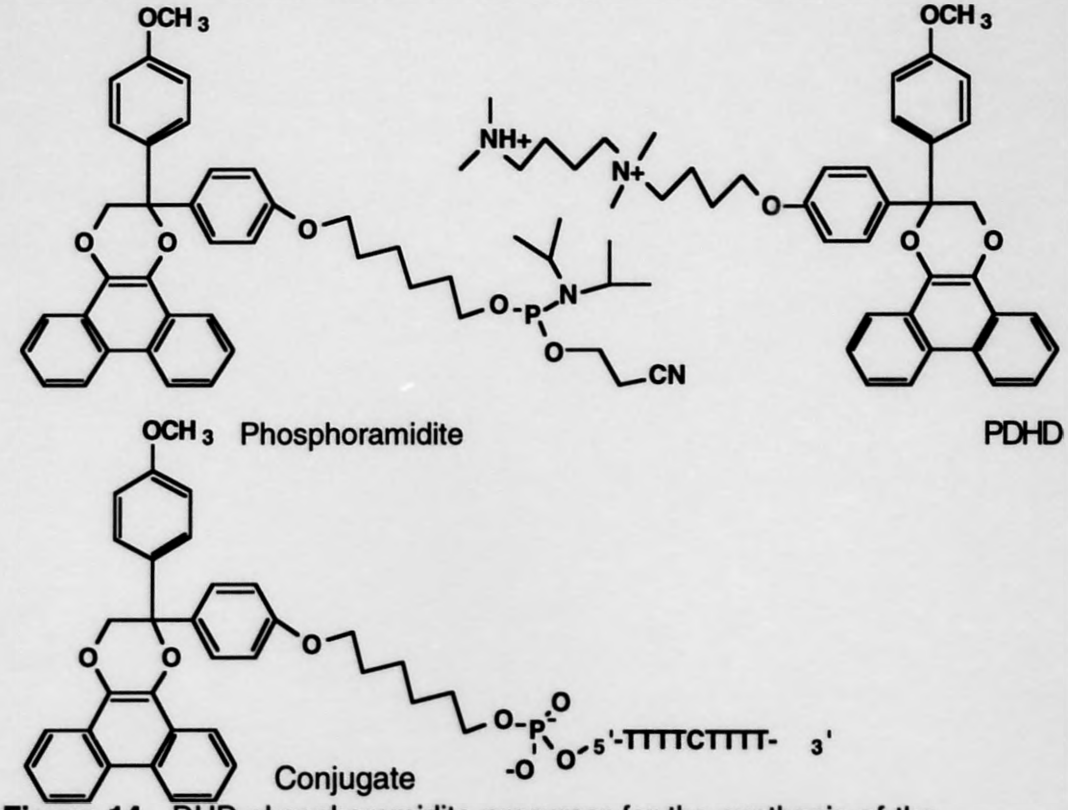


Figure 14. DHD-phosphoramidite precursor for the synthesis of the conjugate, water soluble photosensitizer PDHD, and the PDHD conjugate.

The conjugate was prepared in the DNA Core Facility at the University of Cincinnati by standard synthesis of TTTTGTTTT at 1μM scale and the coupling of this oligonucleotide to PDHD-phosphoramidite on the ABI 394 synthesizer using special precautions and the manual adjustment of standard synthesis procedures to avoid any damage of the conjugate by light, acid, or harsh treatment during deprotection (see Appendix A. Procedures: Synthesis of the DNA target, randomer, and the DNA-PDHD conjugate).

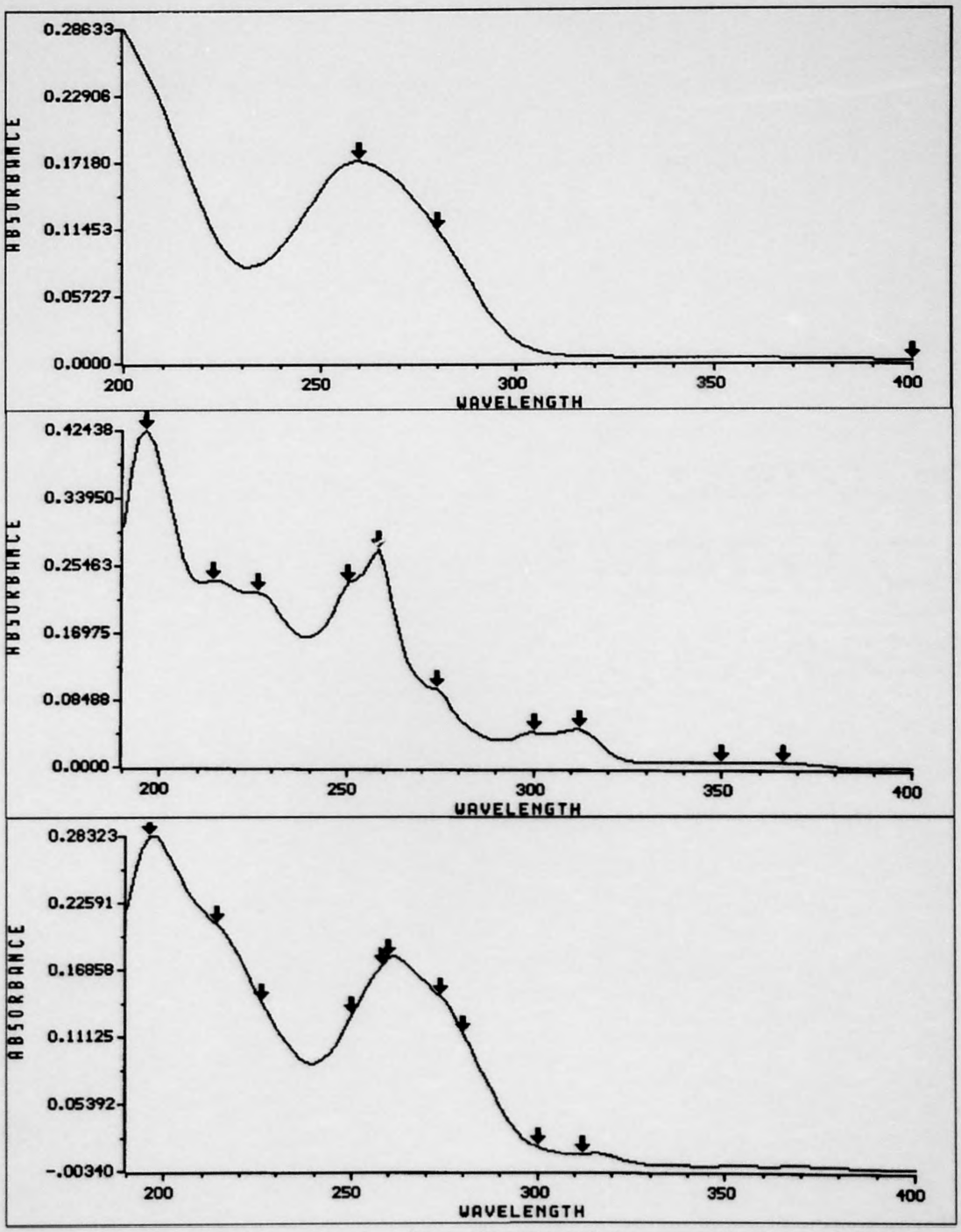


Figure15. Absorbance spectra. One μM 9T DNA (top panel), five μM PDHD (middle panel), and one μM conjugate (bottom panel) are presented.

The yield after purification through a reverse phase column was 191 µg. The target 5'-TCTCGCTGGTGAAAAGAAAAACCACCCTGG-3' (30A), the complementary sequence to the target 5'-CCAGGGTGGTTTTTCTTTTCACCAGCGAGA-3' (30T) and the oligonucleotide 5'-TTTTCTTTT-3' (9T) were also kindly provided by the DNA Core Facility at the University of Cincinnati.

The synthesis of the conjugate was confirmed using UV-VIS, fluorescence and PAGE. The 200-400 nm spectra of 9T DNA alone, PDHD, and the conjugate were compared. The spectra of the conjugate has the peaks of both, 9T DNA and PDHD (Figure 15). The fluorescence spectra of PDHD exhibited very strong emission at 420 nm (Figure 16).

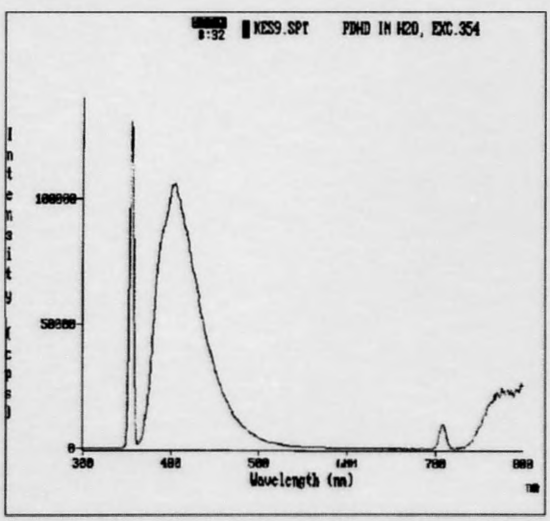


Figure16. Fluorescence spectra of 2.5µM PDHD, excitation at 354nm.

Covalent linkage of PDHD to the oligonucleotide was confirmed by electrophoresis of PDHD, PDHD+9T DNA and the conjugate on 8% polyacrylamide 8M urea gel. The DNA and PDHD were visualized by UV-shadowing (gel was placed on the fluorescent TLC plate; 254 nm light was shined from above; DNA is visible because it absorbs UV, and plate does not fluoresce at that place making a dark spot). As expected, the PDHD was moved

out of the gel by the electric current because it has a positive charge. Even though we knew that it binds to DNA, it was removed from the lane containing PDHD and 9T DNA. Only the lane containing conjugate exhibited the strong dark blue (about 400-440 nm) fluorescence that is also observed in Figure 16. This proves that DHD is covalently linked to 9T DNA in the conjugate. Notably, the conjugate is slower on the gel because of the higher molecular weight and because of the lower charge to molecular weight ratio (DHD is an approximate equivalent of three nucleotides with no charge) (Figure 17).

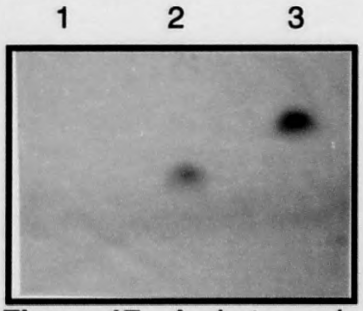


Figure 17. A photograph of the UV-shadowing experiment. Lane 1: one nmole of PDHD, lane 2: one nmole of PDHD + one nmole of 9T DNA, lane 3: one nmole of the conjugate.

To assess how PDHD in a conjugate affect DNA-DNA binding, UV-melting experiments were performed. 30A was premixed with 30T, 9T and a conjugate at 1:1 molar ratio in 10mM sodium phosphate, 1M NaCl buffer, heated to 98°C, and cooled slowly to 4°C to anneal. Multiple UV-melting curves (Absorbance (A) vs. Temperature (T), Figure 18) were collected, and melting temperatures of duplexes were determined at the maximum of the calculated dA/dT curve. The melting temperature for the 30A+30T duplex was 80.8±0.2 °C (±2σ). The duplex of 30A+9T melted at 20.1±0.9°C. The conjugate-30A duplex had a melting temperature of 21±1°C. This is the expected result indicating very little if any stabilization of the duplex by the conjugated PDHD.

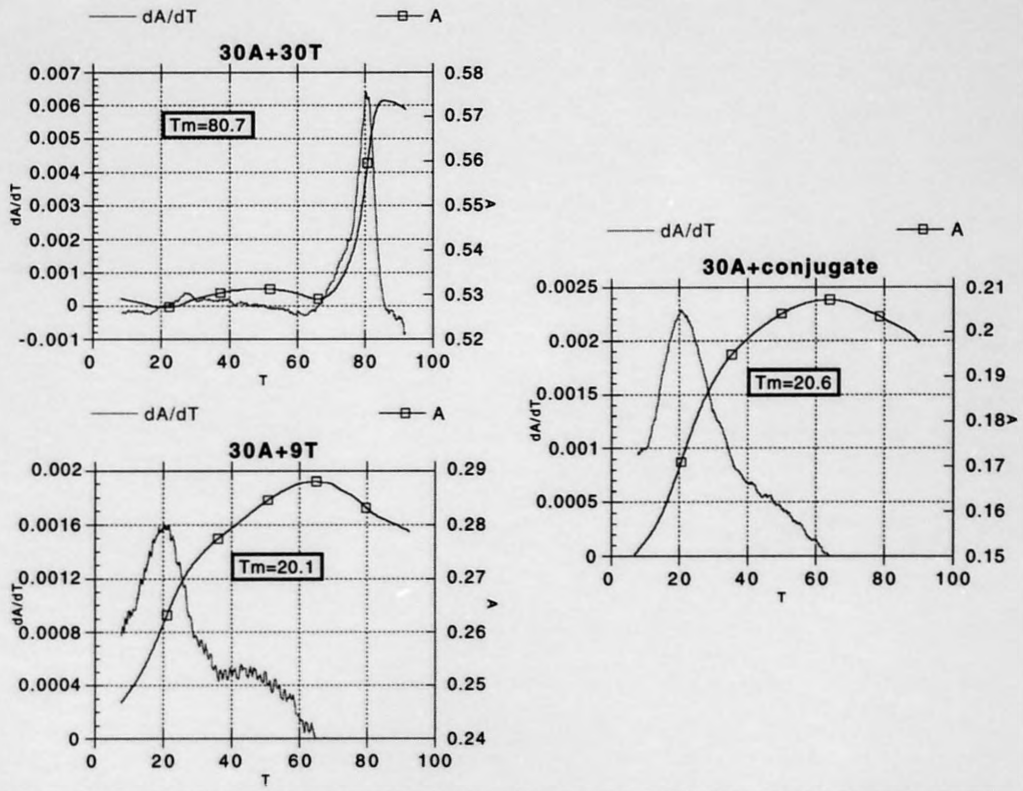


Figure 18. Melting curves. A vs. T and calculated dA/dT vs. T of 30A+30T, 30A+9T and 30A+conjugate in 10mM sodium phosphate, 1M NaCl, pH 7.4.

Gel shift experiments were performed with the same mixtures (30A+30T, 30A+9T, 30A+conjugate) and also with triplex-forming 30A+30T+9T in several different buffers:

TE: 10 mM Tris-Cl, 1mM EDTA, pH 8.0 (duplex favoring),

SDB (Standard Duplex Buffer): 10mM sodium phosphate, 1M NaCl, pH 7.4,

TMA: 50mM sodium acetate, 3mM magnesium acetate, pH 5.0 (triplex favoring),

SMN: 20mM MES, 3mM $MgCl_2$, 100mM NaCl, pH 6.0 (triplex favoring),

MMS: 20mM MES, 1mM $MgCl_2$, 1mM spermine tetrachloride, 10mM NaCl, pH 6.0 (strongly triplex favoring).

One pmole of ^{32}P labeled 30A was mixed with 1-1000 pmoles of the complementary strand, and 2-1000 pmoles of the third strand. With all buffers used, gel shift was observed only with 30A+30T. Neither the triplex mix nor 30A+9T or 30A+conjugate duplexes showed any shift of ^{32}P labeled 30A after electrophoresis in the cold on a 20% native polyacrylamide gel. 30A+9T (or conjugate) duplex has only 19 hydrogen bonds. The third strand binds to the duplex by 18 hydrogen bonds. Both structures apparently are not strong enough to cause a gel shift. However, all photochemical cleavage experiments indicate binding. This is an important observation which is analogous to the failure to observe gel shift of radioactively labeled ss Ori3 DNA (original target) mixed with ds probes from selected clones. The formation of stable binding there might also be confirmed by photochemical cleavage experiments.

Irradiation (for 1-12 hours) of ss 30-mer target 5'-^{32}P-TCTCGCTGGTGAAAAGAAAAACCACCCTGG-3' with and without a PDHD conjugate and dication PDHD was performed in SDB (10 mM sodium phosphate, 1M NaCl, pH7.4) at 350 nm at about 2mW/cm^2 in a Pyrex dish at 4°C. The concentration of the radioactively labeled target was constant at 10nM. The concentration of the PDHD conjugate and PDHD was varied from 0 to 10μM. The reactants were premixed, heated to 55°C and cooled to 4°C at a rate of 0.5°C/minute before irradiation. The reaction after irradiation was quenched by 10mM DTT and treated with 10% piperidine at 90°C to cleave the single stranded target at damaged bases. DNA was triple-lyophilized and redissolved in 80% formamide. Results were analyzed by electrophoresis on 20% polyacrylamide, 8M urea gels. To determine which bases were damaged Maxam-Gilbert chemical sequencing was performed. Two controls were also included: the target that was not irradiated and the target that was irradiated and

treated with piperidine to determine the background cleavage that is not due to PDHD and PDHD conjugate.

The following can be determined from Figure 19 (page 92):

Free PDHD damages specifically only G's in ss DNA.

The conjugate specifically cleaves A21 and C22. This can be explained only by ds formation. Notably, it also damages T10, a site that is consistent with triplex formation. It protects from non-specific cleavage, G16, which is in the middle of the sequence binding region. It non-specifically cleaves other G's. It also forms crosslinks of the conjugate with the target forming two new bands of higher MW (30A+a conjugate via ds formation and 30A+2 conjugates via triplex formation, Figure20).

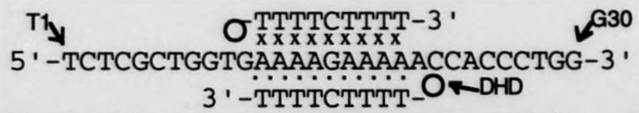


Figure 20. Alignment of the target and the conjugates. The bottom strand binds to the target in antiparallel orientation to the target via Watson-Crick base-pairing. The third strand binds in parallel orientation via Hoogsteen base-pairing. The bases in the target are numbered from the 5' end.

Damage to the target was quantitated by scanning densitometry. The total amount of crosslinks, undamaged 30-mer, specific cleavage in the region between bases A19 and C27, damage of G16 and damage in the region between bases G8 and G11 was estimated (Figure 21).

As time of irradiation was increased from 1/2 to 12 hours, percent of crosslinks increased from 4 to 26, percent of specific damage increased from 2 to 17. The irradiation time for all further studies was set for 12 hours.

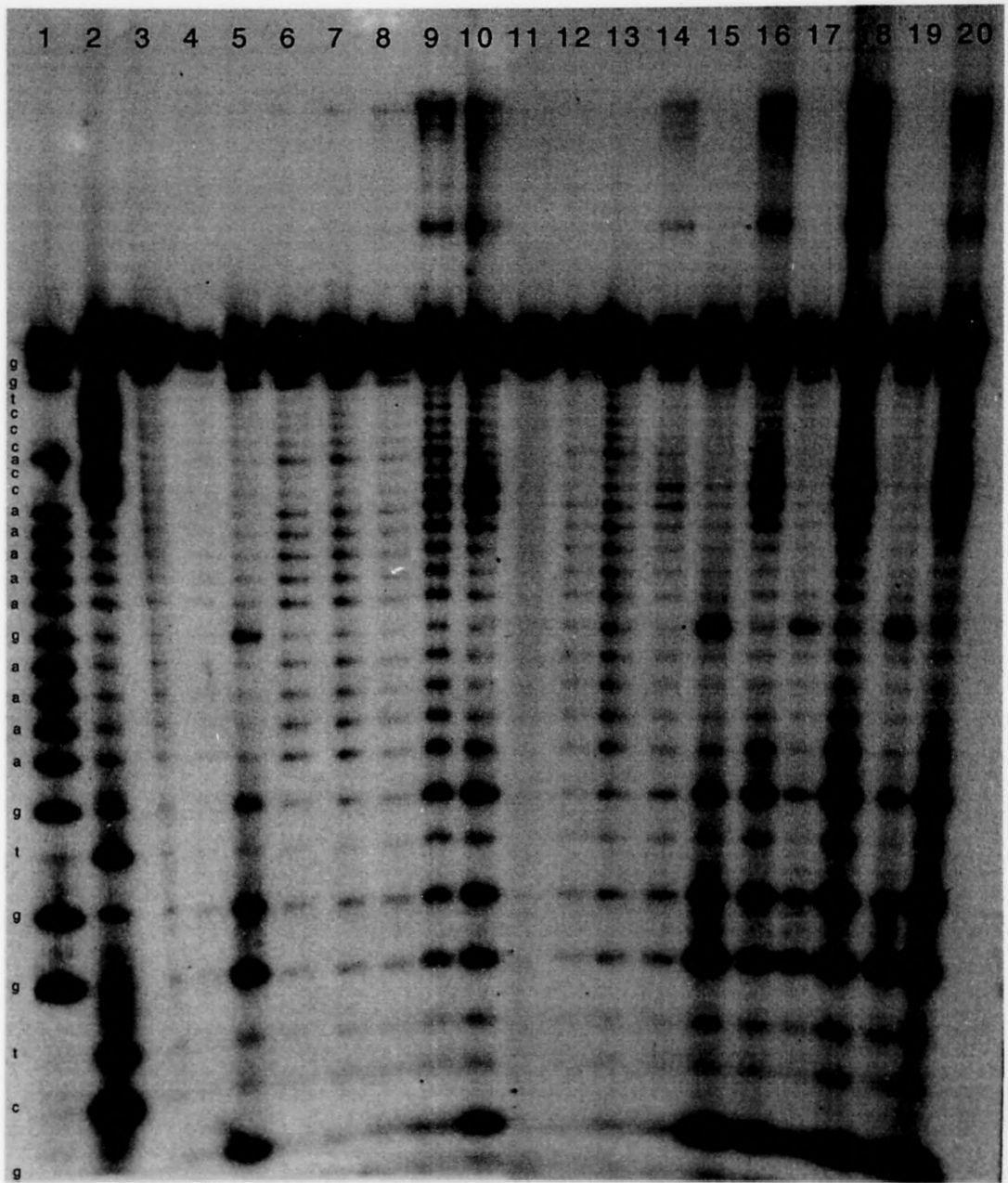
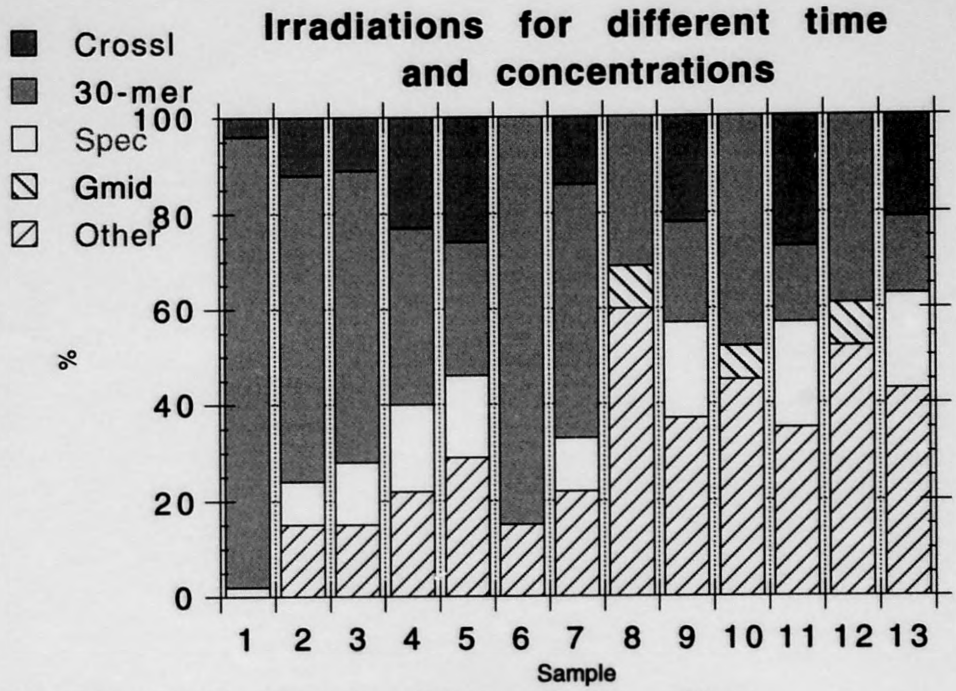


Figure 19. Irradiation of the target with/without PDHD/conjugate for different times and concentrations. Lanes: 1- GA Maxam-Gilbert reaction of the target, 2- TC Maxam-Gilbert reaction, 3- 10nM target (called D) irradiated for 12 hours and treated with piperidine (as the rest of the lanes unless otherwise indicated), 4- D not irradiated and treated, 5- target + 10μM PDHD (called P) irradiated for 0.5 hr, 6- target + 10μM conjugate (called A) irradiated for 0.5 hr, 7- A for 1 hr, 8- A for 3 hours, 9- A for 6 hours, 10- A for 12 hours, 11- as lane 4, 12- D, 13- P 2.5μM, 14-A 2.5μM, 15- P 10μM, 16- A 10μM, 17- P 25μM, 18-A 25μM, 19- P 50μM, 20- A 50μM.



Sample	Crossl	30-mer	Spec	Gmid	Other
1	4	94	2	0	0
2	12	64	9	0	15
3	11	61	13	0	15
4	23	37	18	0	22
5	26	28	17	0	29
6	0	85	0	0	15
7	14	53	11	0	22
8	0	31	0	9	60
9	22	21	20	0	37
10	0	48	0	7	45
11	27	16	22	0	35
12	0	39	0	9	52
13	21	16	20	0	43

Figure 21. Quantitation of irradiations for different time and concentrations. The amount of crosslinks (Crossl), undamaged 30-mer (30-mer), specific cleavage in the region between bases A19 and C27 (Spec), damage to G16 (Gmid), and damage in the region between bases G8 and G11 (Other) was measured scanning densitometry. Samples: 1- irradiation for 1/2 hr, 2- irradiation for 1 hr, 3- irradiation for 3 hours, 4- irradiation for 6 hours, 5- irradiation for 12 hours, 6- 2.5μM PDHD, 7- 2.5μM conjugate, 8- 10μM PDHD, 9- 10μM conjugate, 10- 25μM PDHD, 11- 25μM conjugate, 12- 50μM PDHD, 13- 50μM conjugate.

The concentration of PDHD and the conjugate were varied in the range 2.5μM-50μM. There was little specific damage and crosslinks formed at 2.5μM concentrations. They increased almost to a maximum at 10μM concentrations. Any further increase of concentrations only caused more non-specific damage (Figure 21). The concentrations of damaging agents were set at 10μM for all further experiments.

Figures 22 and 23 demonstrate that when ss target was irradiated with PDHD or PDHD conjugate but not treated with piperidine, little cleavage at G's was observed with free PDHD, and there were more crosslinks than cleavage with the conjugate. This indicates that PDHD is damaging G's, but not cleaving them, and that the conjugate primarily forms crosslinks, that are later partially cleaved by piperidine.

When ss target was irradiated with a large amount of carrier DNA (2,000 carrier:1 target), PDHD cleaved the target several fold less, indicating complete non-specificity, while the conjugate cleaved with almost the same efficiency as it cleaved in the absence of carrier DNA (Figures 23 and 24). Specific cleavage increased when the concentration of NaCl was increased (0, 0.1, 0.25, 0.5, 0.75, 1.0 M NaCl). This is consistent with the double strand stability at different salt concentrations. Varying pH (6.0, 7.4, 8.8) did not affect cleavage significantly (Figures 23 and 24).

The maximum site-specific modification achieved was 67% (51% of crosslinks and 16% of direct cleavage) in a sample that was not treated with piperidine.

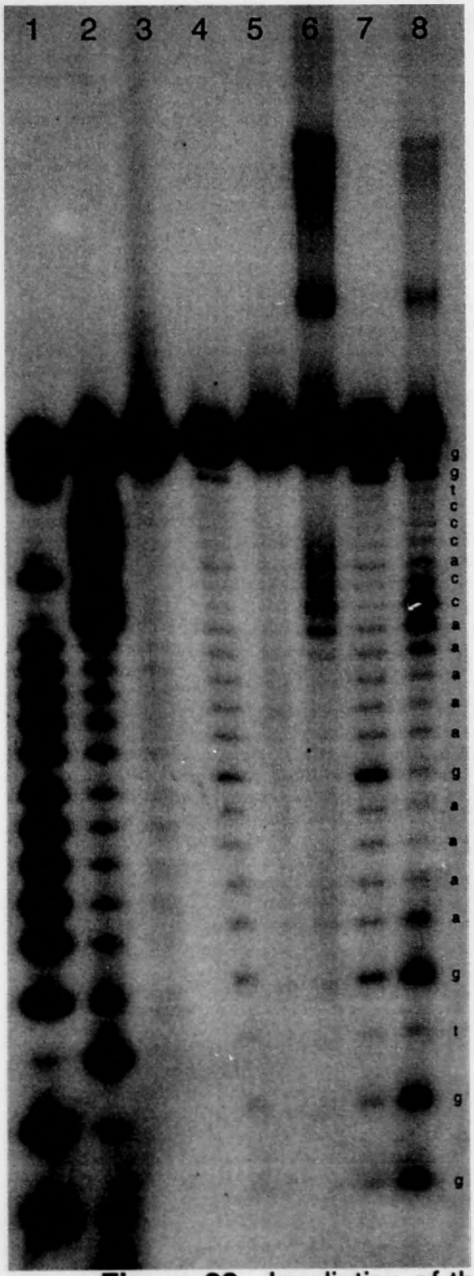
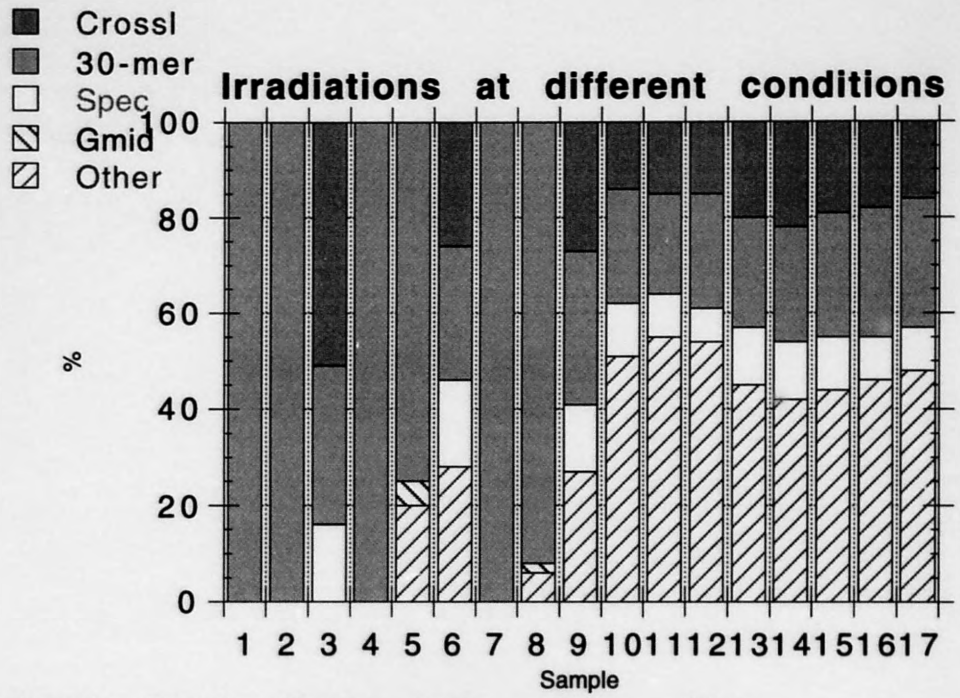


Figure 22. Irradiation of the target with/without PDHD/conjugate with/without piperidine treatment. Lanes: 1- GA Maxam-Gilbert reaction of the target, 2- TC Maxam-Gilbert reaction, 3- 10nM target (called D) not treated, 4- D irradiated for 12 hours (as the rest of the samples) and treated with piperidine, 5- target + 10μM PDHD (called P) irradiated, without piperidine treatment, 6- target + 10μM conjugate (called A) irradiated, without piperidine treatment, 7- P irradiated, with piperidine treatment, 8- A irradiated, with piperidine treatment.



Sample	Crossl	30-mer	Spec	Gmid	Other
1	0	100	0	0	0
2	0	100	0	0	0
3	51	33	16	0	0
4	0	100	0	0	0
5	0	75	0	5	20
6	26	28	18	0	28
7	0	100	0	0	0
8	0	92	0	2	6
9	27	32	14	0	27
10	14	24	11	0	51
11	15	21	9	0	55
12	15	24	7	0	54
13	20	23	12	0	45
14	22	24	12	0	42
15	19	26	11	0	44
16	18	27	9	0	46
17	16	27	9	0	48

Figure 23. Quantitation of irradiations at different conditions. Samples: 1, 2, 3- DNA alone, with PDHD and with the conjugate, no piperidine treatment, 4, 5, 6- DNA alone, with PDHD and with the conjugate, piperidine treated, 7, 8, 9- DNA alone, with PDHD and with the conjugate, with 2 μg of calf thymus carrier DNA present, 10-15-NaCl concentrations of 0M, 0.1M, 0.25M, 0.5M, and1M respectively, 15- pH=7.4, 16- pH=6, 17- pH=8.8.

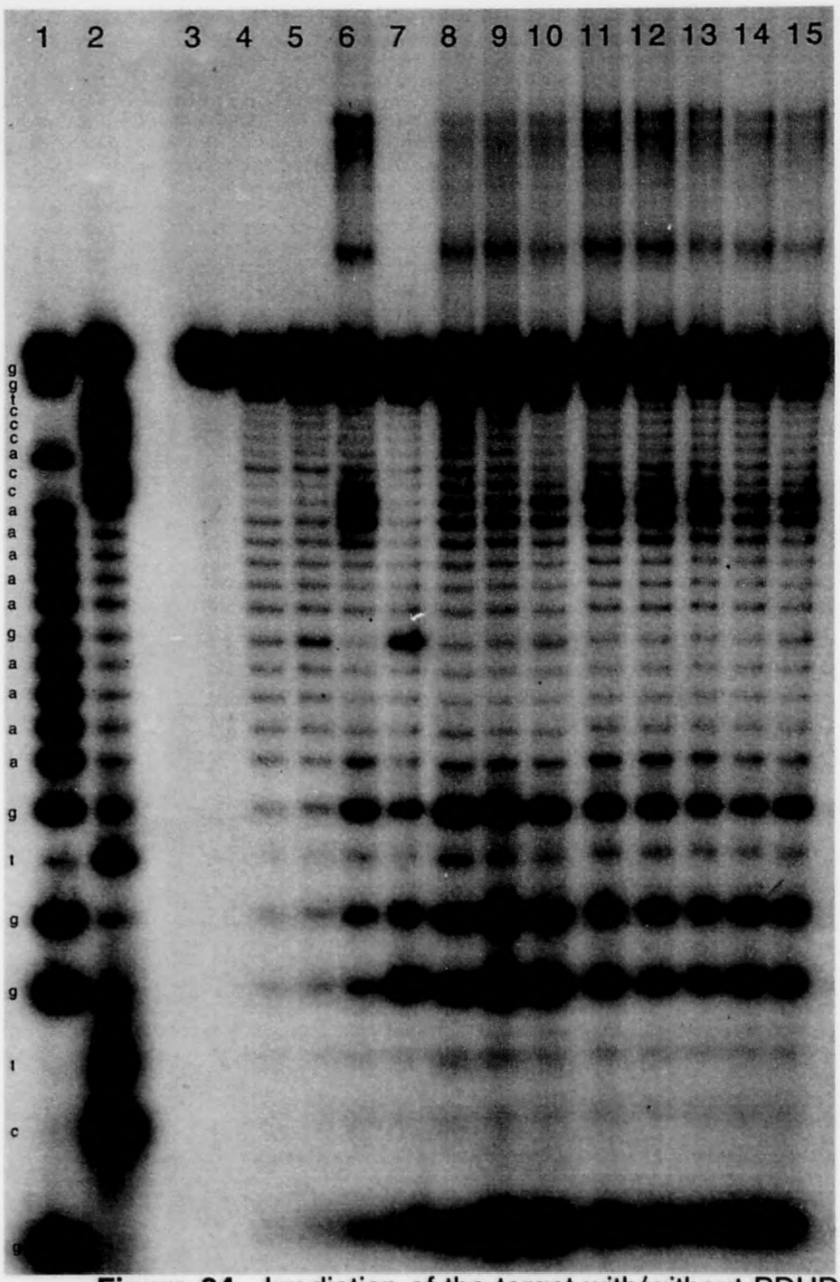


Figure 24. Irradiation of the target with/without PDHD/conjugate in presence of carrier DNA, at different concentrations of NaCl and different pH. Lanes: 1- GA Maxam-Gilbert reaction of the target, 2- TC Maxam-Gilbert reaction, 3- 10 nM target (called D) not treated, 4- D irradiated for 12 hours and treated with piperidine (as the rest of the lanes unless otherwise indicated), irradiation with 2μg of calf thymus carrier DNA, 5- target + 10μM PDHD (called P) with carrier DNA, 6- target + 10μM conjugate (called A) with carrier DNA, 7- P without carrier DNA, 13-A without carrier DNA, 8- A in 0M NaCl, 9- A in 0.1M NaCl, 10- A in 0.25M NaCl, 11- A in 0.5M NaCl, 12- A in 0.75M NaCl, 13- A in 1.0M NaCl, pH 7.4, 14- A in pH 6.0, 15-A in pH 8.8.

Four separate irradiations of the target with the conjugate at standard conditions (10nM target, 10μM conjugate, 12 hours of irradiation, piperidine treated) gave an average of 40% of site-specific modifications: 23% (σ=3%) of crosslinks and 17% (σ=3%) of direct cleavage.

The photochemical cleavage studies prove that the PDHD-oligonucleotide conjugate binds to the complementary target even though this was not detected by gel-shift assays. It efficiently and site-specifically cleaves the designated target, converting a non-specific photochemical reagent into a sequence specific one.

CONCLUSIONS AND FUTURE DEVELOPMENT

Screening of the random library of ds DNA sequences towards the 20 base long origin of lambda (λ) virus replication, Ori3, at optimized conditions yielded two groups of consensus sequences. The percent of binding of several selected sequences to the target column suggest that selected sequences may bind to the target by forming a low-affinity long-lived target-probe complex.

An antisense photosensitizer dihydrodioxin-DNA conjugate was prepared and its selective inactivation properties confirmed. Photochemical cleavage studies prove that the PDHD-oligonucleotide conjugate binds to the complementary target even though this was not detected by gel-shift assays. It efficiently and site-specifically cleaves the designated target, converting a non-specific photochemical reagent into a sequence specific one.

In future studies, the ss PDHD conjugate should be tested as a selective photochemical M13 virus inactivating agent by measuring inactivation of the lac i' gene. PDHD conjugates of the ds DNA sequences identified by combinatorial selection should be used to determine the ss target- ds probe binding constants and tested for their ability to selectively inactivate the ss M13 virus with inserted λ Ori3.

REFERENCES

1. Alberts, B., at al. *Molecular Biology of the Cell*; Garland: New York, 1994.

2. Cohen, J. S.; Hogan, M. E. *Scientific American* **1994**, 76-92.

3. Crooke, S. T., Lebleu, B. *Antisense Research and Applications*; CRC: Boka Raton, 1993.

4. Maher III, L. J.; Wold, B.; Dervan, P. B. *Antisense Research and Development* **1991**, *1*(3), 227-281.

5. Helene, C., Toulme, J.J. *Oligodeoxynucleotides: Antisense Inhibitors of Gene Expression*; McMillan Press: London, 1989; 137-172.

6. Wain-Hobson, S.; Sonigo, P.; Danes, O.; Cole, S.; Alizon, M. *Cell* **1985**, *40*, 9-17.

7. Kohara, Y.; Akiyama, K.; Isomo, K. *Cell* **1987**, *50*, 495-508.

8. Lewin, B. *Genes*; Wiley-Liss: New York, 1983.

9. Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737-738.

10. Feinsfield, G.; Davies, D.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023.

11. Helene, C. *Anti-Cancer Drug Design* **1991**, *6*, 569-584.

12. Thoung, N. G.; Helene, C. *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 666-690.

13. Miller, A. D. *Nature* **1992**, *357*, 455-460.

14. Brennan, M. *C&EN* **1996**, *May 13*, 4-5.

15. Olson, K. E.; Higgs, S; Gains, P. J.; Powers, A. M.; Davis, B. S.; Kamrud, K. I.; Carlston, J. A.; Blair, C. D.; Beaty, B.J. *Science* **1996**, *272*, 884.

16. Powers, A. M.; Kamrud, K. I.; Olson, K. E.; Higgs, S.; Carlston, J. A.; Beaty, B. J. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 4187.

17. James, A.A. *Science* **1996**, *272*, 829.

18. Cooney, M.; Czernuszewicz, G.; Postel, E. H.; Flint, S. J.; Hogan, M. E. *Science* **1988**, *241*, 456-460.

19. Postel, E. H.; Flint, S. J.; Kessler, D. J.; Hogan, M. E. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 8227-8231.

20. Stein, C.A. *Leukemia* **1992**, *6*(10), 967-974.

21. Goodchild, J. *Bioconjugate Chem.* **1990**, *1*(3), 165-187.

22. Froeler, B.; Matteucci, P. Ng. *Nucl. Acids Res.* **1988**, *16*(11), 4831-4839.

23. Jayroman, K.; McPorland, K.; Miller, P.; Ts'o, P. O. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*(3), 1537-1541.

24. Kean, J. M.; Murakami, A.; Blake, K. B.; Cushman, C. D.; Miller, P. S. *Biochemistry* **1988**, *27*, 9113-9121.

25. Sproat, B. S., Lammond, A. I. *Oligonucleotides and anologues*; IRL-Press: Oxford, 1991; 49-86.

26. Barabina, S. M. L.; Sproat, B. S.; Lammond, A.I . *Nucl. Acids Res.* **1992**, *20*(17), 4457-4464.

27. Griffiths, A. D.; Potter, B. V. L.; Eperon, I. C. *Nucl. Acids Res.* **1987**, *15*(10), 4145-4161.

28. Zon, G., Steck, W. J. *Oligonucleotides and anologues*; IRL-Press: Oxford, 1991; 87-108.

29. Vlasov, V. V., Yakubov, L. A. *Prospects for antisense nucleic acid therapy of cancer and AIDS*; Wiley-Liss: New York, 1991; 243-266.

30. Stein, C. A.; Mori, K.; Loke, S. L.; Subasinghe, C.; Shinozuka, K.; Cohen, J. S.; Neckers, L.M. *Gene* **1988**, *72*, 333-341.

31. Matsukura, M.; Shinozuka, K.; Zon, G.; Mitsuya, H.; Petz, M.; Cohen, J. S.; Broder, S. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7706-7710.

32. Matteucci, M.; Lin, K. Y.; Butcher, S.; Moulds, C. *J. Am. Chem. Soc.* **1991**, *113*, 7767-7768.

33. Nielsen, P. E.; Egholm, M.; Berg, R. H. *Science* **1991**, *254*, 1497-1500.

34. Hanvey, J. C.; Peffer, N. J.; Bisi, J. E.; Thomson, S. A.; Codilla, R.; Josey, J. A.; Ricca, D.J .; Hassman, F.; Bonham, M. A.; Au, K. G.; Carter, S. G.; Bruckenstain, D. A.; Boyd, A. L.; Noble, S. A.; Babiss, L. E.; *Science* **1992**, *258*, 1481-1485.

35. Mahel, L. J.; Wold, B.; Dervan, P. B. *Science* **1989**, *245*, 725-730.

36. Strobel, S. A.; Dervan, P. B. *Science* **1990**, *249*, 73-75.

37. Kohl, J. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 1470-1478.

38. Krawczyk, S. H.; Milligan, J. F.; Wadwani, S.; Moulds, C.; Froehler, B. C.; Matteucci, M. D. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 3761-3764.

39. Morvan, F.; Rayner, B.; Imbach, I. L.; Thenet, S.; Bertrand, J. R.; Paoletti, J.; Malvy, C.; Paoletti, C. *Nucl. Acids Res.* **1987**, *15*, 3421-3437.

40. Praseuth, D.; Doan, T. L.; Chassignol, M.; Decout, J.-L.; Habhoub, N.; Lhomme, J.; Thoung, N. T.; Helene, C. *Biochemistry* **1988**, *27*, 3031-3038.

41. Doan, T. L.; Perrouault, L.; Praseuth, D.; Decout, J.-L.; Habhoub, N.; Lhomme, J.; Thoung, N. T.; Helene, C. *Nucl. Acids Res.* **1987**, *19*, 7749-7760.

42. Praseuth, D.; Perrouault, L.; Doan, T. L.; Chassignol, M.; Thoung, N. T.; Helene, C. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 1349-1353.

43. Leonetti, J.-P.; Degols, G.; Lebleu, B. *Bioconjugate Chem.* **1991**, *1*, 149-153.

44. Thoung, N. T., Asseline, U. *Oligonucleotides and anologues* ; IRL-Press: Oxford, 1991; 283-308.

45. Rei, D.; Ulrich, H. D.; Shultz, P. G. *Science* **1991**, *253*, 1408-1411.

46. Knorre, D. G., Zarytova, V. F. *Prospects for antisense nucleic acid therapy of cancer and AIDS*; Wiley-Liss: New York, 1991; 195-218.

47. Strobel, S. A.; Moser, H. E.; Dervan, P. B. *J. Am. Chem.Soc.* **1988**, *110*, 7927-7929.

48. Boidot-Forget, M.; Chassignol, M.; Tacasugi, M.; Thoung, N. T.; Helene, C. *Gene* **1988**, *72*, 361-371.

49. Lokhof, S. G.; Podyminogin, M. A.; Sergeev, D. S.; Silnikov, V. N.; Kutyavin, I. V.; Shishkin, G. V.; Zarytova, V. P. *Bioconjugate Chem.* **1992**, *3*, 414-419.

50. Mergny, J. L; Duval-Valentin, G.; Nguyen, C. H.; Perrouault, L.; Faucon, B.; Rougee, M.; Montenay-Garestier, T.; Bisagni, E.; Helene, C. *Science* **1992**, *256*, 1681-1684.

51. Faraqui, F.; Sarin, P. S.; Sun, D.; Letsinger, R. L. *Bioconjugate Chem.* **1991**, *2*, 422-426.

52. Bonfils, E.; Mendes, C.; Roche, A. C.; Monsigny, M.; Midoux, P. *Bioconjugate Chem.* **1992**, *3*, 277-284.

53. Behr, J.-P.; Demeneix, B.; Loefer, J.-P.; Perez-Mutzel, J. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 6982-6986.

54. Burkholder, G. D.; Latimer, L. J. P.; Lee, J. S. *Chomosoma* **1991**, *101*, 11-18.

55. Rao, B. J.; Dutreix, M.; Radding, C. M. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2984-2988.

56. Iverson, B. L.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4615-4619.

57. Francois, J.-C.; Saison-Behmoaras, T.; Barbier, C.; Chassignol, M.; Thoung, N. T.; Helene, C. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 9702-9706.

58. Doan, T. L.; Perrouault, L.; Chassignol, M.; Thoung, N. T.; Helene, C. *Nucl.Acids.Res.* **1987**, *15*(21), 8643-8659.

59. Doan, T. L.; Perrouault, L.; Helene, C. *Biochemistry* **1986**, *25*, 6736-6739.

60. Frolova, E. I.; Ivanova, E. M.; Zarytova, V. F.; Abramova, T. V.; Vlasov, V. V. *FEBS* **1990**, *269*(1), 101-104.

61. Anderson, R. R.; Parrish, J. A. *Lasers in surgery and medicine* **1981**, *1*, 263-276.

62. Cimino, G. D.; Gamper, H. B.; Isaacs, S. T.; Hearst, J. E. *Ann. Rev. Biochem.* **1985**, *54*, 1151-1193.

63. Duval-Valentin, G.;Thoung, N. T.; Helene, C. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 504-508.

64. Giovannangeli, C.; Rougee, M.; Garestier, T.; Thoung, N. T.; Helene, C. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 8631-8635.

65. Giovannangeli, C.; Thoung, N. T.; Helene, C. *Nucl.Acids.Res.* **1992**, *20*(16), 4275-4281.

66. Byrnes, R. W.; Fiel, R. J.; Datta-Gupta, N. *Chem.-Biol. Interactions* **1988**, *67*, 225-241.

67. Gibbs, E. J.; Pasternak, R. F. *Seminars in Hematology* **1989**, *26*(2), 77-85.

68. Marzili, L. G. *New J. Chem.* **1990**, *14*, 409-420.

69. Pratviel, G.; Pitie, M.; Bernadou, J.; Meunier, B. *Nucleic Acids. Res.* **1991**, *19*(22), 6283-6288.

70. Doan, T. L.; Praseuth, D.; Perrouault, L.; Chassignol, M.; Thoung, N. T.; Helene, C. *Bioconjugate Chem.* **1990**, *1*, 108-113.

71. Kasturi, C.; Platz, M. S. *Photochemistry and Photobiology* **1992**, *56*(4), 427-429.

72. Perrouault, L.; Asseline, U.; Rivalle, C.; Thoung, N. T.; Bisagni, E.; Giovannangeli, C.; Doan, T. L.; Helene, C. *Nature* **1990**, *344*, 358-360.

73. Hill, A. R.; Orgel, L. E. *Bioconjugate Chem.* **1991**, *2*, 431-434.

74. Lee, J. S.; Burkholder, G. D.; Latimer, L. J. P.; Haug, B. L.; Braun, R. P. *Nucleic Acids.Res.* **1987**, *15*(3), 1047-1061.

75. Hsieh, P.; Camerini-Otero, C. S .; Camerini-Otero, R. D. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 6492-6496.

76. Stasiak, A. *Molecular Microbiology* **1992**, *6*(22), 3267-3276.

77. Baum, R. *C&EN* **1996**, *February 12*, 28.

78. Joyce, G. F. *RNA: Catalysis, Splicing, and Evolution*; Elsevier: Amsterdam, 1977; 83-87

79. Tuerk, C.; Gold, L. *Science* **1990**, *249*, 505-510.

80. Szostak, J. W. *TIBS* **1992**, 89-93.

81. Ellington, A.; Szoostak, J. W. *Nature* **1990**, *346*, 818-822.

82. Ellington, A.; Szoostak, J. W. *Nature* **1992**, *355*, 850-852.

83. Lerner, R. A.; Benkovoc, S. J.; Schultz, P. G. *Science* **1991**, *252*, 659-667.

84. Lesley, S. A.; Patten, P. A.; Schultz, P. G. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 1160-1165.

85. Pei, D.; Ulrich, H. D.; Schultz, P. G. *Science* **1991**, *253*, 1408-1411.

86. Ito, T.; Smith, C. L.; Cantor, C. R. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 495-498.

87. Ito, T.; Smith, C. L.; Cantor, C. R. *Nucleic Acids Research* **1992**, *20*(13), 3524.

88. Ji, H.; Smith, L. M.; Guilfoyle, R. A. *GATA* **1994**, *11*(2), 43-47.

89. Kumar, A.; Rustogi, S. *Nucleosides & Nucleotides* **1994**, *13*(9), 1855-1860.

90. Nishikava, N.; Oishi, M.; Kiyama, R. *The Journal of Biological Chemistry* **1995**, *270*(16), 9258-9264.

91. Bock, L. C.; Griffin, L. C.; Latham, J. A.; Vermaas, E. H.; Toole, J. J. *Nature* **1992**, *355*, 564-566.

92. Macaya, R. F.; Schultze, P.; Smith, F. W.; Roe, J. A.; Feogon, J. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 3745-3749.

93. Harada, K.; Orgel, L. E. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 1576-1579.

94. Sassanfar, M.; Szostak, J. W. *Nature* **1993**, *364*, 550-553.

95. Jenison, R. D.; Gill, S. C.; Pardi, A.; Polisli, B. *Science* **1994**, *263*, 1425-1429.

96. Connell, G. J.; Yarus, M. *Science* **1994**, *264*, 1137-1141.

97. Tuerk, C.; MacDougal, S.; Gold, L. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 6988-6992.

98. Burke, J. M.; Berzal-Herranz, A. *FASEB* **1993**, *112*(7), 106-112.

99. Green, R.; Szostak, J. W. *Science* **1992**, *258*, 1910-1915.

100. Bartel, D. P.; Szostak, J. W. *Science* **1993**, *261*, 1411-1418.

101. Munir, K. M.; French, D. C.; Loeb, L. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 4212-4216.

102. Fodor, S. P. S.; Read, J. L.; Pirrung, M.S.; Strayer, L.; Lu, A.T.; Solas, D. *Science* **1991**, *251*, 767-773.

103. Zuckermann, R. N., et al. *J. Med. Chem.* **1994**, *37*, 2678.

104. Kanie, O.; Barresi, F.; Ding, Y.; Labbe, J.; Otter, A.; Folsberg, L. S.; Ernst, B.; Hindsgaul, O. *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 2720.

105. Borman, S. *C&EN* **1996**, *February 12*, 29-54.

106. Thayer, A. M. *C&EN* **1996**, *February 12*, 57-64.

107. Ptashne, M. *A genetic switch-gene control and phage λ*; Cell Press: Cambridge, 1987.

108. Singleton, S. F.; Dervan, P. B. *Biochemistry* **1993**, *32*, 13171-13179.

109. Szewczyk, J. W.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, 6778-6779.

110. Asseline, U.; Bonfils, E.; Dupret, D.; Thoung, N. T. *Bioconjugate Chemistry* **1996**, *7*, 369-379.

111. Nielsen, P. E. *J. Mol. Recognition* **1990**, *3*, 1-25.

112. A. van Belkum; Blommers, M. J. J.; H. van den Elst; J. H. van Boom; Hilbers, C W. *Nucleic Acids Research* **1990**, *18*(16), 4703-4710.

113. Wilson, R. M.; Shnapp, K. A.; Harsh, A.; Keller, S.; Schlemm, D.; Hanson, J. F.; Jackson, N. L. unpublished work, **1996**.

114. Harsh, A.; advisor Wilson, R. M. dissertation, Chemistry Department, University of Cicinnaty, **1996**.

115. Gailard, C.; Staruss, F. *Nucleic Acids Research* **1990**, *18*(2), 378.

116. *M13 Cloning/Dideoxy Sequencing Instruction Manual*; Bethesta Research Laboratories Life Technologies: Gaithersburg, 1992.

117. *Step-By-Step Protocols for Taq Cycle Sequencing System*; United States Biochemical: Cleveland,1992.

APPENDIX A. PROCEDURES*

Synthesis of the DNA Target, Randomer, and the DNA-PDHD Conjugate

1. Set up for the standard automated phosphoramidite DNA synthesis on the DNA Synthesizer "Applied Biosystems 391" according to the manufacturer's instructions. Use "Glen Research" phosphoramidites, columns, solvents and reagents.

2. For the 20-mer target synthesis, prepare 40 10 x 100 mm tubes to monitor the progress of the synthesis: cleaved by trichloroacetic acid (TCA), 4',4''-dimethoxytrityl (DMTrO) protecting group produces an orange color. Comparing the absorbance at 625 nm of wash solutions after each deprotection step, estimate the efficiency of each coupling step and an overall yield of the synthesis.

3. Use the special "Oligo Affinity Support" 0.5μM synthesis column that allows the synthesis of an oligonucleotide which can be deprotected while remaining attached to the supporting Teflon fibers.

4. When the synthesis is complete, remove the fibers using tweezers and place them into a 1.5 ml microcentrifuge tube. Add 1 ml 30% NH_4OH and incubate at 55°C overnight to deprotect the oligonucleotide.

5. Place the fibers back into a column. Wash with 15 ml of water from both ends of the column. Dry the column under vacuum for 1 hour.

6. The 99-mer randomer is prepared on a standard prepacked disposable 1μM controlled pore glass (CPG) column. Type in "N" instead of G,

---

* In order to facilitate further progress of this large ongoing project, procedures are given in specific detail to aid other scientist in reproducing results and to help other students to progress rapidly with minimum errors since this was the first use in this Department of PCR, combinatorial selection and thermal cycling sequencing.

A, T, or C where random bases should appear. After the synthesis, precipitate DNA in ethanol (Step 17 of the Selection Procedure) and amplify by PCR (*Taq* Polymerase Chain Reaction Procedure).

7. Use special precautions when synthesizing the conjugate:

Keep the photosensitizer cyanoethyl phosphoramidite derivative in the dark at all times.

Test this derivative for stability in all the reagents that will be used in the synthesis, carefully dehydrate and dissolve in acetonitrile to make 0.1M solution.

Design the oligonucleotide for the conjugate to have as many T's as possible since T is not protected; for other bases use monomers that are easy to deprotect.

After the 5'-end deprotection step of the oligonucleotide to which the photosensitizer will be attached, wash TCA from the column with acetonitrile for 2 minutes (4 times longer than usual) to minimize exposure of the photosensitizer to acid.

Perform the coupling reaction in 0.5M tetrazole in acetonitrile for 3 minutes (6 times longer than usual) to increase coupling efficiency.

Use 20mM I_2 (not 0.1M I_2) in THF/pyridine/water for the oxidation step.

Exclude the detritylation step (since it involves the use of TCA) at the end of the program.

After cleaving from the column in aqueous ammonia, deprotect for 1 hour at room temperature.

Purify using reverse-phase cartridges.

Do not perform manual detritylation while purifying.

Order all other necessary oligonucleotides from commercial suppliers, such as Genosys Biotechnologies, Inc. or Oligos Etc., Inc.

Detection of DNA on Teflon Fibers

1. Remove approximately 10% of fibers from the column. Determine an exact percentage by weighing the fraction and the rest of the fibers.

2. Treat the fibers that have DNA which is to be removed with a 1 ml solution of 0.1M sodium periodate in 0.1M sodium phosphate (pH=6.0). Stir the mixture in the dark at room temperature for 6 hours. This oxidizes the diol of ribose that links the oligonucleotide to the fibers.

3. Remove the fibers and wash them with 10 ml of water.

4. Submerge the fibers in 1 ml of N-propylamine:acetonitrile:water 1:2:8 (PAW) solution and incubate for 3 hours at 50°C in a screw cap vial (to cleave the oxidized ribose). Recover the DNA containing supernatant and rinse the fibers twice with 1 ml PAW solution.

5. Join all three fractions. Dry DNA under vacuum. Dissolve in 1 ml of water. Determine the amount of DNA present spectroscopically.

Selection Procedure

1. Open the screening column, and place the Teflon fibers inside. Carefully install filters on both sides of the fibers. Attach the Luer-lock female connectors to the column. Place the supporting plates on both sides of the assembly. Insert the bolts into the holes on the plates. Tighten the whole set up with the nuts to prevent leaks.

2. Wash the fibers in the column 5 times with 3 ml of dd H_2O using a 3 ml Luer-lock syringe and reversing the direction of flow twice.

3. Wash the column with 15 ml of the screening buffer (CB, NCB or TCB). Remove all the buffer from the column by forcing air through it.

4. Leave one syringe attached to the column. Remove the plunger from the second syringe. Attach this syringe to the column.

5. Measure the absorbance of the DNA to be screened.

6. Transfer about 10 μg of DNA (in 2 ml of screening buffer) into the syringe barrel. Insert the syringe plunger slowly, and gently push the DNA solution through the column into the bottom syringe. Detach the top syringe, remove air from it and from the column, reattach the syringe.

7. Equilibrate the column with the DNA randomer for 1 hour, slowly pumping the solution from one syringe to the other every 5 minutes. This is a negative screening.

8. Collect all the solution that contains DNA that had not bound to the fibers in the column. Measure its absorbance.

9. Disassemble the column. Place the Teflon fibers in 8M urea solution.

10. Use the DNA 20-mer target attached to the Teflon fibers for the positive screening. Assemble the column and wash it with 15 ml of water and with 15 ml of the screening buffer (as in Step 2 and Step 3).

11. For the positive screening, use the solution of the DNA randomer from the negative screening. Equilibrate DNA with the target on the column for 1 hour, occasionally pumping the solution from one syringe to the other.

12. Collect all unbound DNA and measure its absorbance. Clear the remaining solution from the column by forcing air through it.

13. Wash the column five times with 2 ml of the screening buffer, applying each portion slowly over a 1 minute period. Collect all washes to measure their absorbance or save for PCR analysis later.

14. Elute bound DNA with 2 ml of 1mM NaEDTA, 10mM NaOH solution, and then with 2 ml of 8M urea. Apply solutions slowly over a 1 minute period. Collect them for further processing.

15. Disassemble the column. Place the Teflon fibers with a 20-mer on them in 8M urea solution. Store at -20°C.

16. Measure the absorbance of all 5 washes and both elutions.

17. Precipitate DNA in 13x100 mm round bottom centrifuge tubes by adding 3M ammonium acetate to a final concentration of 0.3M, 2.5 volumes of ice cold 95% ethanol (or 1.5 vol. of isopropanol), and 10 µl of 0.25% linear polyacrylamide for coprecipitant [115]. Mix and cool the solution at -80°C for 20 minutes or at -20°C for 2 hours. Centrifuge the tubes at 12,000xg for 15 minutes. Withdraw the supernatant with a transfer pipette, trying not to disturb the pellet. Add 1 ml of ice cold 95% ethanol, spin at 12,000xg for 5 minutes. Withdraw the supernatant. Dry under vacuum for 30 minutes.

18. Dissolve the DNA pellet in 10 µl of water.

Taq Polymerase Chain Reaction (PCR)

1. Prepare nine 0.5 ml microcentrifuge tubes. Prepare to PCR a positive (with a known template) and a negative (without template) controls, and solutions after the precipitation from five washes and two elutions. First, run 1/10 of each selection.

2. For one 100 µL PCR mixture use:

   1 µl of DNA template

   10 µl of 10X PCR Buffer

   0.5 µl of 2% gelatin solution (prevents template from adhering to the tube wall)

4 µl of 5mM DNTP's

1 µl of 100µM primer 39-mer

1 µl of 100µM primer 24-mer

81 µl of water

1 µl *Taq* Polymerase antibody (inhibits *Taq* Polymerase activity until "hot start" preventing the synthesis of truncated strands)

0.5 µl of 2units/µl Taq Polymerase

Mix and centrifuge in the microcentrifuge. Overlay with 2 drops of mineral oil.

3. Run the reaction in the thermal cycler to find the minimum number of cycles needed:

94°C for 5 minutes to hot-start the reaction (this inactivates *Taq* Polymerase antibody and any contaminating enzymes);

run six times for five cycles:

94°C for 45 seconds

40°C for 60 seconds

72°C for 135 seconds

remove 10 µl of the reaction mixture for the analysis

cool to 4°C to stop the reaction

4. Prepare two horizontal 1% agarose gels to analyze DNA after PCR. For that purpose, heat the solution of 1% agarose in 1X TBA buffer in a boiling water bath until all agarose grains completely disappear. For each gel, pour the hot solution into the mold, place the comb, and let it solidify at room temperature for 30 minutes. Place gels into the horizontal gel electrophoresis apparatus. Fill the buffer chambers with 1X TBA. Add enough buffer to submerge gels under a 1 mm layer of solution.

5. Add 2 µl of 6X XBS loading buffer to each 10 µl sample. Mix and carefully load solutions (five washes and two elutions after 5, 10, 15, 20, 25, and 30 cycles of PCR, positive and negative controls after 30 cycles of PCR) into 44 wells of the gels.

6. Run electrophoresis at 100 V until dyes are situated at about the center of the gel.

7. Soak gels in 0.5µg/ml ethidium bromide solution for 45 minutes at room temperature.

8. Place each gel on the UV box and turn on the long wavelength UV. Photograph the gel through the red filter using Polaroid film.

9. Determine the minimum necessary number of cycles to achieve maximum amplification of the eluted DNA. By differences between the brightness of the bands, estimate the amounts of DNA present in washes and elutions before PCR.

10. Prepare eleven 0.5 ml microcentrifuge tubes for PCR of positive and negative controls, and nine 1 µl samples of the elution solution. Use the recipe from Step 2. Run the minimum necessary number of cycles. Add 10 minutes of 72°C incubation at the end of the last cycle to ensure that all newly synthesized DNA is double stranded. Cool down to 4°C to stop the reaction.

11. Add 0.1 ml of chloroform: isoamyl alcohol 24:1 to each tube of amplified elution DNA to dissolve mineral oil. Centrifuge the tubes in a microcentrifuge.

12. Join all 10 aqueous fractions (top layer) in a 1.5 ml tube. Add 0.1 ml of chloroform: isoamyl alcohol 24:1. Centrifuge the tube. Remove aqueous DNA solution.

13. Precipitate DNA in ethanol as described in Step 17 of the Selection Procedure.

14. Dissolve DNA in 2 ml of screening buffer. It is ready for the next selection step.

Cloning

The purpose of this procedure is to select and purify individual sequences from the enriched pool of DNA that was screened. Part of the pool is digested with two restriction endonucleases producing fragments with two different "sticky ends". These fragments are ligated into an M13 virus with "forced" orientation. Virus is later transfected into JM101 *E. Coli* host bacteria. The M13mp19 cloning vector and its host JM101 have a feature allowing identification of viruses that have a DNA fragment inserted into the multiple cloning site of the vector: in the presence of IPTG and X-gal, these modified (with an insert) non-recombinant (not capable of α-complementation) phages produce white (colorless) plaques, while unmodified recombinant ones make blue plaques upon plating and growth on indicator plates. One needs to select the white plaques and purify the bacteriophage to obtain clones containing individual sequences from our pool of DNA.

Restriction Endonuclease Digestion

Prepare four 1.5 ml microcentrifuge tubes. One will be used to digest M13mp19 replicative form (RF) DNA with two restriction endonucleases. Two others should have M13mp19 RF DNA digested with one of the endonucleases. The last one is used to prepare a double-digested insert from the ds DNA from screenings.

1. Mix in a microcentrifuge tube:

2 µg of DNA

2 µl *Eco*R I 10x buffer

10 units of *Eco*R I (for tubes 1, 2 and 4)

10 units of *Bam*H I (for tubes 1, 3 and 4)

Water to 20 µl

2. Centrifuge the mixture for 5 sec in a microcentrifuge. Incubate at 37°C for one hour.

3. Add 3 µl of 0.25M Na_2EDTA to stop the reaction, mix, and add 27 µl of 10mM Tris-HCl (pH=7.5), 1mM Na_2EDTA (TE) buffer.

4. Precipitate DNA by adding 3M sodium acetate to a final concentration of 0.3M (5.6 µl) and add 2.5 volumes (140 µl) of 95% ice cold ethanol. Use 10 µl of 0.25% linear polyacrylamide as a coprecipitant.

5. Hold at -20°C for 2 hours or at -80°C for 20 minutes.

6. Centrifuge at 10,000xg for 10 minutes at 4°C.

7. Discard the supernatant. Carefully rinse the DNA pellet with 0.5 ml of ice cold 95% ethanol.

8. Centrifuge at 10, 000xg for 5 minutes at 4°C.

9. Discard the supernatant and dry the pellet for 30 minutes under vacuum.

10. Dissolve in an appropriate volume of 1X ligation buffer to make 10ng/µl (2fmoles/µl) of vector and 1.0ng/µl (20fmoles/µl) of insert assuming 100% efficiency of ethanol precipitation.

Ligation

Prepare several mixtures with the different ratios of amounts of vector to insert (1:2, 1:7, 1:15 in fmoles of DNA). For the controls use both single-

digested vectors and a double-digested vector using water instead of the insert DNA.

1. Mix the following in a 500 µl microcentrifuge tube:

   10 µl restriction endonuclease-digested M13mp19 vector (100 ng, 20 fmoles)

   0-15 µl restriction endonuclease-digested DNA insert fragments (0-15 ng, 0-300 fmoles)

   10 µl 5X ligation buffer

   1 µl 1unit/µl T4 DNA ligase

   water to 50 µl

2 Incubate at 15°C for 6 hours.

3. Stop the reaction by the addition of 200 µl of 0.025M Na_2EDTA.

Ligated DNA can be stored at 4°C until needed for transformation or kept at -20°C for long-term storage.

Preparation of Competent Host Cells

Use JM101 cells as a transient host and as an exponential culture to provide a suitable lawn for infection. Restreak the host cells every four weeks on fresh minimal agar M-9 plates. After growing them at 37°C, seal the plates and store in a cold room at 4°C.

1. Inoculate 3 ml of YT broth in a 13x100 mm culture tube with a loop from a single colony of JM101. Incubate at 37°C with shaking at 200rpm overnight.

2. Dilute an aliquot of the overnight culture 1:100 in YT broth in a growth flask with a glass arm for monitoring absorbance of the culture. Prepare 3 ml of

this mixture per transformation. Set the optical zero on the SPEC 20 at 550 nm. Grow the culture at 37°C with shaking at 250rpm.

3. Monitor the absorbance every 30 min.

4. Transfer 100 µl of the overnight culture to 10 ml of sterile YT media in the second growth flask 2 hours after you started growing competent cells. Incubate lawn host cells at 37°C with shaking at 250rpm.

5. When the absorbance of competent cells is about 0.5 AU, place the growth flask on ice for 20 minutes.

6. Transfer the culture to a chilled 50 ml conical centrifuge tube. Centrifuge at 700xg for 10 minutes at 4°C. Higher speed will pack the cells too tightly and they will disintegrate when resuspended.

7. Carefully decant the supernatant. Suspend the cell pellet in 1/2 of the original volume of sterile ice cold 50mM $CaCl_2$, 10mM Tris-Cl (pH 8.0) by gently swirling the tube. Do not vortex these cells since they are very fragile. Incubate on ice for 30 minutes.

8. Centrifuge the cells at 700xg for 10 minutes at 4°C. Decant the supernatant. Resuspend the cell pellet in 1/10 the original volume of ice cold 50mM $CaCl_2$, 10mM Tris-Cl (pH 8.0) by gently swirling the tube.

Transformation

It is important to include the following controls in the transformation procedure:

original M13mp19 to monitor transformation efficiency (maximum, 100% of plaques expected),

M13 digested with *Eco*R I to determine efficiency of digestion (20% plaques expected),

M13 digested with *Bam*H I to determine efficiency of digestion (20% plaques),

double digested M13 to determine efficiency of digestion (10%plaques)

ligated M13 digested with *Eco*R I to determine efficiency of ligation (90% plaques),

ligated M13 digested with *Bam*H I to determine efficiency of ligation (90% plaques),

ligated double digested M13 (20% plaques).

Use all three ligated mixtures with the different vector-insert ratios.

1. Label thirteen 500 µl microcentrifuge tubes and chill on ice (seven controls plus two for each ligated mixture with the different vector-insert ratios).

2. Gently mix competent cells. Aliquot 0.3 ml of competent cells into each tube.

3. Add 3 ng of control DNA to the appropriate tubes. Add 3 ng and 15 ng of ligated mixtures with the different vector-insert ratios to the appropriate tubes. Gently mix each tube.

4. Incubate on ice for 30 minutes. Heat shock the cells at 42°C for two minutes. Then plate cells immediately.

Top Agar Plating

Complete steps 1 and 2 of this procedure in advance, a day before plating.

1. Prepare YT agar plates. Use at least 25 ml of melted agar per plate. Incubate inverted plates at 37°C overnight or dry them at 50°C for 1 hour in the inverted open position. Label the bottoms of the plates.

2. Melt YT top agar in an autoclave, aliquot 3 ml into 10 x 100 mm sterile glass culture tubes, hold at 55°C in a dry bath.

3. Add 10 µl of 100mM IPTG and 50 µl of 2% X-gal to each tube containing top agar.

4. Add 200 µl of the exponential lawn culture (absorbance circa 0.5 at this time) to one of the tubes of the transformed cells that have just been heat-shocked. Gently mix.

5. Remove one tube containing top agar from the bath. Add the entire volume of mixed cells using transfer pipette. Quickly cover the tube with a piece of sterile parafilm, and invert the tube a couple of times, avoid generating bubbles.

6. Pour the entire contents of the tube onto the appropriate plate. Quickly rock the plate to evenly spread the top agar. Cover the plate and allow the top agar to solidify.

7. Return to Step 4 until all transformation mixes have been plated. Prepare also a control having no transformed cells.

8. After all of the top agar has solidified invert the plates and incubate at 37°C overnight.

Selection and Single Plaque Purification of Clones

1. Pick a separate well-isolated white plaque with a transfer pipette and transfer it to 100 µl of SM buffer. Mix well and allow phage to diffuse from the agar overnight at 4°C. Pick one blue plaque as a control.

2. Prepare 10-fold serial dilutions of bacteriophage stocks in SM for each clone. Dispense 0.1 ml of each dilution into 500 µl microcentrifuge tubes.

3. Add 0.2 ml of plating bacteria (JM101 *E. Coli* at stationary phase, A_{550}=2.0) to each tube. Mix by vortexing. Incubate at 37°C for 20 minutes to allow phage particles to adsorb. JM101 is grown to stationary phase to give consistent results for the infection with an intact phage, while JM101 at exponential phase is better when used for the transfection with viral DNA.

4. Add 10 µl of 100mM IPTG and 50 µl of 2% X-gal to each 10 x 100mm glass tube containing 3 ml of YT top agar at 55°C.

5. Transfer bacteria from one microcentrifuge tube into the top agar media, cover the glass tube with parafilm, invert the tube a couple of times, and immediately pour onto a labeled plate containing 25-35 ml of hardened bottom YT agar. Repeat this step with each of the microcentrifuge tubes and include a control without the phage.

6. Close the plates and let the top agar harden. Invert the plates and incubate at 37°C for overnight.

7. Pick a very well isolated single white plaque for each clone, and transfer it to 100 µl of SM buffer. Mix by vortexing. Incubate the phage overnight at 4°C. It can be stored this way for several days.

Prepare "blue-plaque" and "no-plaque" controls by performing Step 7 with a blue plaque and with the lawn cells.

Sequencing

This procedure determines an exact sequence of bases in an insert DNA. First, the single stranded DNA template for sequencing is prepared. Then the *Taq* Polymerase based cycle-sequencing procedure is performed using radioactively labeled primer. Polyacrylamide gel electrophoresis, exposure of

the gel to an X-ray film and film development lead to the final results of sequencing.

Preparation of Single-Stranded DNA Template

1. Inoculate 5 ml of YT media with a single JM101 colony. Grow overnight at 37°C with shaking at 200rpm.

2. Dilute the culture 1:50 in YT media.

3. Dispense 5 ml of diluted bacterial cells into sterile 13x100 mm culture tubes.

4. Add 10 µl (1/10 of the plaque culture) of SM solution containing the phage with an individual clone to the suspension of JM101 cells. Include "blue-plaque" and "no-plaque" controls at this point.

5. Incubate labeled culture tubes for 6 hours at 37°C with vigorous shaking at 250-300rpm. Do not incubate overnight to avoid contamination of template with bacterial nucleic acids.

6. Transfer 1.5 ml of the culture to two 1.5 ml microcentrifuge tubes.

7. Centrifuge for 5 minutes at 12,000xg. Transfer 1.4 ml of the supernatant into two fresh tubes. The supernatant in one of the tubes will be used in further purification, while the second tube should be saved as a virus stock of the clone. It is stable at 4°C indefinitely. Do not add chloroform and do not freeze since this inactivates M13.

8. Centrifuge the culture for the second time to improve the template quality (for 5 minutes at 12,000xg).

9. Transfer 1.2 ml of supernatant to a fresh tube. Add 300 µl of 20%PEG, 2.5M NaCl solution. Mix briefly, let stand at 4°C for 15 minutes.

10. Centrifuge for 10 minutes at 12,000xg. The visible pellet should be in all tubes except for the "no-plaque" control.

11. Completely remove all supernatant: PEG/NaCl will interfere with subsequent steps. Centrifuge for 5 minutes at 12,000xg and remove the residual supernatant if any is left.

12. Resuspend the phage in 0.2 ml of TE buffer and add 0.2 ml of phenol:chloroform 3:1 mixture. Prepare phenol in advance by 3-5 sequential extractions with TE buffer. Vortex the mixture. Centrifuge for 5 minutes at 12,000xg.

13. Remove the top aqueous layer, avoiding the interphase and lower phase, to a fresh tube.

14. Repeat steps 12 and 13 with this aqueous layer.

15 Add 0.2 ml of chloroform:isoamyl alcohol 24:1 mixture to the aqueous phase. Vortex vigorously for 20 seconds. Centrifuge for 5 minutes at 12,000xg to separate phases.

16. Remove 70-80% of aqueous phase to a new tube.

17. Add 3M sodium acetate to a final concentration of 0.3M and then add 2.5 volumes of ice-cold 95% ethanol.

18. Chill for 30 minutes in a -80°C freezer. Centrifuge for 15 minutes at 12,000xg.

19. Wash the DNA pellet twice with 95% ice-cold ethanol. Dry the sample under vacuum for 30 minutes.

20. Resuspend the DNA pellet in 50 µl of TE buffer.

21. Repeat steps 17-20 at least once more with this DNA to get a higher quality template DNA for sequencing reactions.

Labeling Primers

Universal M13 cycle primer, reverse M13 cycle primer and specially designed (for example: -82M13) primers can be used for cycle sequencing. Primers are labeled using T4 Polynucleotide Kinase. Preliminary phosphatase treatment is not needed for the synthetic oligonucleotides.

1. Prepare the following reaction mixture in a 0.5 ml microcentrifuge tube (use gloves, tweezers to handle samples and a radioactivity screen to protect yourself):

   2.5 µl γ-^{35}S ATP 1200Ci/mmol, 10mCi/ml (25 pmoles)

   1 µl 20 µM primer (20 pmoles)

   1 µl 10X Kinase buffer

   2 µl T4 Polynucleotide Kinase 10 units/µl (20 units)

   water to 10 µl

2. Vortex the mixture and centrifuge in an microcentrifuge.
3. Incubate at 37°C for 2 hours.
4. Heat at 95°C for 5 minutes to inactivate the kinase.
5. Centrifuge the mixture to collect any condensate.
6. Store at -80°C for up to a few months.

Check for Complete Labeling

1. Prepare 20% acrylamide:bis-acrylamide 20:1 containing 8M urea in 1X TBE buffer. Use 35 ml per gel.
2. Use a small vertical protein gel electrophoresis apparatus.
3. Wash both glass plates with water, then with 95% ethanol. Cover both plates twice with 50 µl of the water repellent Sigmacoat.

4. Place 1 mm spacers at both sides of the plates. Clamp the plates to secure the position of spacers. Use 3M tape to seal the bottom and the sides of an assembly. Fortify the corners with some extra tape and small clamps.

5. Add 0.2 ml of freshly prepared 10% ammonium persulfate and 20 µl of TEMED. Mix and pour immediately in the space between plates. Place the comb to make wells for the loading of samples.

6. Allow the gel to solidify, mark the wells, remove the comb and the seal from the bottom.

7. Prepare two 0.5 ml microcentrifuge tubes. Add 1.0 µl of loading buffer to both. Add 0.5 µl of labeled primer to one and 0.5 µl of 300 Ci/mmol 2.5mCi/ml γ-^{35}S ATP to the other.

8. Load both mixtures into the wells. Run electrophoresis in 1X TBE buffer at 3W for 2-3 hours.

9. Open the plates. Soak the gel in 10% glycerol for 20 minutes to remove urea. Place the gel on a piece of Whatman paper, cover it with Saran Wrap, and dry on a vacuum drier at 80°C for 40 minutes.

10. Expose the dried gel to Fuji or Kodak X-ray film in a cassette for 30 minutes. Develop the film. Determine whether labeling was complete. About 80% ATP should be consumed, and the DNA band should be about 4 times darker than the ATP band.

Cycle Sequencing

1. Label four 0.5 ml microcentrifuge tubes (G, A, T, C) for each clone and the control.

2. Fill the G, A, T and C tubes with 4 µl of the ddGTP, ddATP, ddTTP, and ddCTP Termination Mixes respectively. Cap the tubes to prevent evaporation.

Use the 7-deaza-dGTP Termination Mixes when sequence compressions (unresolved bands, common with G &C rich sequences) are encountered.

3. Pre-dilute the *Taq* Polymerase to 0.5unit/µl concentration in ice-cold Polymerase Dilution Buffer. Prepare enough Polymerase only for one day.

4. Mix the following master mix in a microcentrifuge vial for each clone:

5-12.5 µl template DNA

2 µl Cycle Sequencing 10X Reaction Buffer

1 µl labeled primer

2 µl diluted *Taq* DNA Polymerase

1 µl *Taq* Antibody

water to make 17.5 µl of the total volume

Vortex and centrifuge the mixture.

5. Remove 4 µl of the master mix and transfer it to each of the G, A, T, C tubes from step 2. Mix and centrifuge. Overlay each vial with 10-20 µl mineral oil, cap the vials and place them in the thermal cycler.

6. Start the cycling program:

95°C for 5 minutes to "hot-start" the reaction

30 cycles of:

95°C for 30 seconds

55°C for 30 seconds

72°C for 2 minutes

72°C for 5 minutes to finish all extension reactions

4°C incubation

7. Add 4 µl of Stop Solution to each of the termination reactions. Mix and centrifuge the tubes in a microcentrifuge. Keep the samples frozen at -20°C until ready to load on the sequencing gel.

Denaturing Gel Electrophoresis

1. Prepare 20% acrylamide:bis-acrylamide 20:1 containing 8M urea in 1X TBE buffer. Use 75 ml per gel. Filter and degas the solution.

2. Wash the base glass sequencing plate with water and 95% ethanol. Cover it twice or thrice with the following mixture:

10 µl silane

30 µl glacial acetic acid

0.3 ml water

3. Wash the "rabbit ear" plate with water and 95% ethanol. Cover it twice with 0.5 ml of the water repellent Sigmacoat. Use 0.25 ml if the same plate is used for all the gels.

4. Lay 0.2 mm thick Teflon spacers along the edges of the base plate. Place the "rabbit ear" plate, with the treated surface facing the treated surface of the base plate, on top of the spacers.

5. Clamp the edges of the plates using clips. Tape the sides and the bottom of the assembly with 3M (other brands leak) tape. Double enforce the bottom corners of the plates with the tape and small clamps.

6. Have the comb and three clips on hand. Lay the assembly on a flat surface, lift the comb end of it to about 30° angle to horizontal and place it on a support.

7. Add 0.4 ml of 10% ammonium persulfate and 40 µl of TEMED to the acrylamide solution. Mix and pour the solution in-between the plates, starting from one corner, slightly tilting the plates towards the same corner. Keep applying the gel solution continuously to avoid bubbles. When the gel solution reaches the top of the plates, lower the support and insert the comb. Clamp the comb with the clips.

8. Allow the gel to polymerize for 30 min. Remove the tape from the bottom of the gel. Mark the wells with a permanent marker. Carefully wash the wells after removing the comb.

9. Tighten the plates in the electrophoresis apparatus. Fill the upper and the lower buffer chambers with 1X TBE buffer. Pre-run the gel for 3 hours at 40W.

10. Wash out the wells using a transfer pipette to remove urea.

11. Heat the samples to 70°C for 2-5 minutes to denature the DNA and load 4 µl immediately on the gel in the "G, A, T, C" order for each clone. Load 7-deaza-dGTP mixtures in the same order in four adjacent lanes.

12. Run electrophoresis at constant power of 30W for about 6 hours (until Bromphenol Blue dye reaches the bottom of the gel)

13. Remove the gel from the electrophoresis apparatus and place on a flat surface. Remove all tape. Lift the "rabbit ear" plate using a spatula.

14. Place the base plate with the gel into a bath of 10% methanol, 10% acetic acid for 30 minutes to fix the DNA and to remove urea. Wash the gel with distilled water and dry at 70°C in an oven for 3-4 hours.

Autoradiography

1. Check the radioactivity with a hand-held counter to approximate time of exposure.

2. Expose the emulsion side of an X-ray film directly to the dried gel. Cover the film with a plastic plate. Wrap everything with heavy duty black plastic. Clamp the sides of the plates together using six clips.

3. After exposure develop using standard equipment in the dark room. Read the sequence of the clone using the ladder of bands on the film.

Assessment of Sequence Specific Cleavage

PAGE Purification of DNA Samples

1. Dissolve commercially synthesized oligonucleotides in the deionized, distilled and sterilized (dds) water to make 1mM solution.

2. Prepare a small 20% polyacrylamide, 8M urea gel in-between two water-repellent plates. Mix 2 µl of oligonucleotides to be purified with 3 µl of 80% formamide. Load them on the gel.

3. Run electrophoresis at 8W constant power of for 3 hours.

4. Cover a fluorescent thin layer chromatography (TLC) plate with Saran Wrap. Place the gel on the TLC plate. Visualize the DNA by shining short wavelength UV from above using a hand held lamp. Cut out the DNA bands with a sharp razor blade. Place the gel slice in a 1.5 ml microcentrifuge tube.

5. Add 0.2 ml of 0.5M ammonium acetate, 1mM Na_2EDTA. Shake the sample at 37°C for at least 4 hours. Remove 180 µl.

6. Repeat step 5 two more times. Join all three fractions.

7. Add 2.5 volumes of 95% ethanol, freeze at -80°C for 20 minutes. Centrifuge at 15,000xg for 15 minutes. Wash the pellet with 0.5 ml 95% ice cold ethanol. Dry the sample under vacuum for 30 minutes.

8. Dissolve DNA in 1 ml dds water. Measure its absorbance to determine the exact amount of DNA present.

9. Add 0.11 ml of 3M ammonium acetate. Repeat step 7. Dissolve DNA in dds water to make 20µM solution.

Labeling Target DNA

1. Take extreme care when working with [γ-^{32}P] ATP. Wear gloves and hide behind the screen, keep the hand-held counter on all the time. Handle all samples with tweezers and micro spatulas. Discard all the waste into designated labeled baskets.

2. Prepare the following mixture in 0.5 ml microcentrifuge tube:

   1 µl of 20µM target DNA

   1 µl 10X T4 Polynucleotide Kinase buffer

   6 µl dds water

   1 µl 6000Ci/mmoi, 150mCi/ml [γ-^{32}P] ATP

   1 µl 10units/µl T4 Polynucleotide Kinase

   Centrifuge the mixture, vortex it gently. This makes 2µM target solution.

3. Incubate at 37°C for 10 minutes. Heat at 95°C to inactivate the kinase.

4. Centrifuge the mixture to collect the condensate. Keep in the labeled lead vial at -20°C for up to 6 weeks.

5. Check the efficiency of the target labeling by running a small denaturing 20% polyacrylamide gel. Load 1 µl of 200nM labeled target solution premixed with 4 µl 80% formamide into the carefully washed wells. After 3 hours of electrophoresis at 8W of the constant power, place the gel on Whatman paper. Do not dry. Cover the gel with Saran Wrap and expose to an X-ray film for 5 minutes. Develop the film. Confirm the labeling. Overexpose another film overnight to ensure the labeled target purity. The target is ready for irradiation studies and chemical sequencing.

Maxam and Gilbert Chemical Sequencing of the Target

1. Label two 1.5 ml screw-cap microcentrifuge tubes as "GA" and "TC". The following procedure is for the GA reaction. Do the same for the TC reaction unless otherwise indicated. Keep in mind that ^{32}P-labeled DNA is very radioactive

2. Mix in a tube:

   4 µl 1µg/µl carrier calf thymus DNA

   5 µl 2µM end-labeled target DNA

   20 µl dd water (10 µl water for TC)

   30 µl 88% formic acid (30 µl of anhydrous hydrazine for TC)

   Incubate for 2 hours at room temperature.

3. Stop the reaction by adding:

   200 µl ice cold Hydrazine Stop Solution (0.3 M sodium acetate, 0.1 mM Na_2EDTA), (210 µl for TC)

   2 µl 20 µg/µl glycogen

   0.75 ml 95% ice-cold ethanol

4. Freeze at -80°C for 30 minutes. Centrifuge at 10,000xg for 20 minutes. Discard the supernatant, wash the pellet with 0.5 ml 95% ice-cold ethanol.

5. Dissolve in 0.2 ml 0.3M ice-cold sodium acetate. Add 0.5 ml 95% ice-cold ethanol. Repeat Step 4. Wash the pellet once again with 0.5 ml 95% ice-cold ethanol. Dry under vacuum for 30 minutes.

6. Add 100 µl 10% piperidine. Cap the tube tightly with the screw-cap and vortex it. Incubate at 90°C for 60 minutes (50 minutes for TC).

7. Dry under vacuum overnight. Add 40 µl of dd water, vortex, dry under vacuum. Add 20 µl of dd water, vortex, dry under vacuum. This step removes piperidine.

8. Dissolve in 200 µl of 80% formamide. Use 2 µl of this solution per lane in the denaturing 20% polyacrylamide gel electrophoresis to confirm the chemical sequencing efficiency.

Irradiating Target with the Photochemically Active Reagents

1. Dilute the labeled target DNA in dd water to 40nM concentration.
2. Mix in a 0.5 ml microcentrifuge tube:

   5 µl of 40nM ^{32}P-labeled target DNA

   8 µl of 2.5X Binding Buffer (usually 1X is: 10mM sodium phosphate, pH=7.4, 1M NaCl)

   2µl 100µM photosensitizer PDHD or 2 µl of 100 µM PDHD-DNA conjugate or 2 µl water for control

   2 µl of 1µg/µl carrier DNA (if decided to include)

   x µl dd water to a final volume of 20 µl

Dim the lights when adding photochemical reagents. Keep the samples from any direct irradiation by covering them with black fabric.

3. Vortex and centrifuge mixtures. Heat at 55°C for 1 minute. Cool down to 4°C at the rate of 1°C/minute in a thermal cycler. Incubate in dark at 4°C for 1 hour.

4. Build the 60 x 80 mm 8 mm high glass sample table in a round (d=120mm, 40 mm high) Pyrex dish. Fill 7 mm of the dish with 1X Binding Buffer. Place a 50 x 70 mm piece of parafilm on the sample table. Incubate in the cold room for an hour.

5. Place a 20 µl droplet of the mixture of the target DNA with photochemical reagents on the parafilm. Place six 100 µl 1X Binding Buffer droplets around to prevent evaporation of your sample. Put the glass cover on the dish.

6. Carefully place the dish in the irradiation reactor in the cold room at 4°C . Turn on the fan to keep samples cold.

7. Turn the 350 nm lights on and irradiate for an appropriate time, usually 12 hours. Check occasionally the size of the droplets. Add a few µl of water if necessary.

8. Stop the reaction by adding 2 µl of 0.1M dithiothreitol (DTT). Transfer the droplet to a 1.5 ml screw-cap tube. Vortex the mixture, incubate at room temperature for 20 minutes.

9. Add 120 µl of 10% piperidine, mix, incubate at 90°C for 1 hour to cleave strands at damage sites.

10. To remove piperidine, dry under vacuum for overnight. Add 40 µl of dd water, vortex, dry under vacuum. Add 20 µl of dd water, vortex, dry under vacuum.

11. Dissolve in 10 µl of 80% formamide. Use 3 µl of this solution per lane in the denaturing 20% polyacrylamide gel electrophoresis to assess the photochemical cleavage of the target. Use the chemically sequenced target and the non-irradiated target for controls.

12. Run electrophoresis for about 5 hours at 30W constant power. Do not dry the gel, cover it with Saran Wrap. Expose an X-ray film for a few hours or for overnight.

13. Determine the cleavage efficiency by quantitative scanning densitometry. Use Molecular Analyst software on a computer connected to a densitometer to control the scanning process and quantify the data.

UV-Melting

1. Prepare 1 ml of duplex or triplex DNA in an appropriate buffer. Recommended absorbance is 0.4-0.7 AU.

2. Turn the Carey 219 spectrophotometer, the circulating water bath and the computer on.

3. Allow instrumentation to warm up for 15 minutes.

4. Start the N_2 flow through the spectrophotometer. Set the pressure at 50 kPa.

5. Set the following on the Carey 219:

Source-UV, Power-UV	Wavelength-260 nm
Wavelength mode-timer only	Beam interchange-normal
Measurement-serial	Mode-auto gain
Gain-0.3 (set by control)	Range-0.5
Suppressions-0	Temperature display-on
Record time-20 seconds	Cycle time-0 minutes

6. Place the cuvettes filled with the buffer only. Adjust the optical zero with the Balance and Fine Balance dials.

7. Check that the cuvette with a thermocouple is filled with water.

8. Add 100 ml of glycerol to 2 l of water in the waterbath to prevent boiling.

9. Type "log" into the computer. Choose "yes" for "minimum logging" and "no" for "separate files". Name your file. Choose the starting and the ending temperatures. More consistent results are achieved by choosing the high

temperature (95°C) as the starting one, and cooling the sample slowly, at 0.5°C/minute rate, to 4°C.

10. Add DNA to the sample cuvette. Cover both cuvettes with a flexible strong caps made from the tops of rubber pipet bulbs to prevent evaporation. Wait until samples reach 94-95°C. Push the "Start" button on the spectrophotometer. Press "Enter" on the computer to start recording data. It takes 3-4 hours to finish the experiment.

11. Copy your file onto a 3 1/4' DOS formatted disc by typing "copy filename D:" on the screen. Plot the data using KaleidaGraph.

APPENDIX B. MATERIALS AND SUPPLIES

Instruments

A PerSeptive Biosystems Expedite Synthesizer was used to make regular oligonucleotides. The conjugate and the randomer were made on an ABI 394 Synthesizer.

Hewlet Packard 8452A Diode Array Spectrophotometer was employed to measure DNA and DNA conjugate's absorbance. It is a single beam, microprocessor controlled UV-VIS (190 to 820 nm) spectrophotometer. Spectra were printed by Hewlet Packard Desk Jet 500 printer.

SPEX 1680 0.22M Double Spectrometer equipped with a photomultiplier tube was used to record fluorescence spectra of the photosensitizer and the conjugate.

Cary219 Spectrophotometer (Varian Associates Model No. 95450) was used for UV-melting experiments. It is a direct ratio recording UV-VIS double beam spectrophotometer. It has a fitting for nitrogen purging. A Haake A80 circulating water bath provides the control of the temperature of the cell. An IBM computer controls the water bath and records the absorbance data from the Cary 219 and, simultaneously, the temperature measured by a thermocouple in the control cell located close to the experimental cell.

DNA irradiation was performed in a RAYONET Photochemical Reactor equipped with a cooling fan, GraLab Universal Timer 171 (0-60 min.) and heated-cathode 3500A emission lamps from Southern New England Ultraviolet Co., Middletown, Connecticut.

For gel electrophoresis, a Fisher Biotech/ Fisher Scientific DNA Sequencing System FB-SEQ3545 and Hoefer Scientific Instruments (Model

No. SE400) protein gel system was used. Dan-Kar Corp. (Model No. DK203) 0-2,000 Volts, 0-200 miliampers Electrophoresis Power Supply provided power for the DNA sequencing gels. Viral and plasmid DNA was analyzed on agarose gels poured in BIO-RAD Laboratories DNA SUB CELL.

PCR was performed in Thermolyne Temp-Tronic (Model No DB66925) 36 well 0.5 ml tube and Perkin Elmer Cetus P4973 54 well 0.5 ml tube DNA thermal cyclers. Both provided 1°C/min temperature gradient when heating and cooling. Temperature cycling is controlled by the programs.

We precipitated DNA after PCR using SS-34 rotor in Sorval Superspeed RC-2 Automated Refrigerated Centrifuge (0-20,000rpm). Beckman J2-21 refrigerated centrifuge (0-20,000rpm) with the JA-21 rotor was used in the precipitation of all radioactively labeled DNA. Bacterial cells were centrifuged in the International Equipment IECER-6000 refrigerated centrifuge using the 219 Rotor with different adapters.

Developed pictures or X-ray films were scanned using a Bio-Rad GS-670 imaging densitometer. This data was processed on an IBM computer using Molecular Analyst software.

General Apparatus

Cold room: Forma Scientific Model No. 75011 environment control system

Eberline Instrument Corporation Geiger Counter: Model No. E-120, 0-50 mR/hr, 0-60,000 CPM

Microcentrifuge: Brinkman Instruments Inc. Centrifuge 5414, 15,000rpm

Vortex: Vortex-Genie Model K-550-G with a touch start and speed control

Hybridization Incubator: Robbins Scientific Corporation with the temperature and speed controls

Hot Plate Stirrer: Corning PC-351 with the temperature and speed controls

pH meter: Sargent-Welch pH4090 with a Fisher SN204165 electrode, temperature adjustment, intercept and slope controls

Dry baths: Boekel 110011 with low and high temperature controls and Fisher Isotemp Model 145

Lyophilizer: Central Scientific Company 91505 Centro-Hyvac 2 vacuum pump connected to a Virtis Company 6211-024500 refrigerated condensation trap that leads to a Nalgene 250 mm Vacuum Desiccator

Hand-held UV lamp: Spectroline Longlife Model ENF-260C, 23Watt, with a short wavelength (254nm)-long wavelength (365nm) switch

Photo-UV-box: Fotodyne Inc. Model 3-3000, 204Watt, 365 nm

Camera: Fotodyne Inc. 5-5334 or 1-1740 with Polaroid665 or Polaroid667 film and the red filter

Gel dryer: Savant Speed Gel SG200 with temperature, time controls, and a vacuum switch

Gel-on-glass dryer: Blue M Stabil Therm dry type bacteriological incubator

Autoclave: Market Forge Sterilmate Model STM-E TYPE C, 121°C

Spectrocolorimeter for the cell cultures: Baush & Lomb Spectronic 20

Incubator: Fisher Isotemp 200 Series Model 255D with a temperature control

Shaking incubator: Lab-Line Instruments Environ-Shaker 3597 with the temperature and speed controls

Shaking water bath: New Brunswick Scientific G-76 with the temperature and speed controls

Balances: Ainsworth Fisher 200 (0-200±0.01g) and Metter HL52 (0-160±10^{-5}g)

Pipettors: Gilson Pipetman 5000µl, Gilson Pipetman 1000µl, Gilson Pipetman 200µl, Oxford Adjustable Sampler System 50-200µl, Oxford Adjustable Sampler System 2-10µl, Oxford Benchmate 0.5-10µl. All pipettors were calibrated at Paramount Precision Calibration and deliver ± 0.5% of the preset volume.

Vacuum filtering system: Fisher 300 ml glass top and support, Millipore clamps

Thermometer: FREAS, -10°C to 110°C, can be read to nearest ± 0.5°C

DNA alignment was performed using MacVector program on Power PC Power Macintosh 8100/80.

Data and word processing was done on Macintosh IIcx using Microsoft Word5.1, KaleidaGraph 2.1 and CSC ChemDraw programs.

Solutions

All solutions were autoclaved to sterilize unless otherwise indicated.

Column Buffer (CB): 20mM Tris-Cl, pH 7.6, 1mM $MgCl_2$, 0.25M NaCl

New Column Buffer (NCB): 20mM Tris-Cl, pH 7.3, 1mM $MgCl_2$, 3mM spermine tetrahydrochloride ($SpmCl_4$), filtered to sterilize

Third Column Buffer (TCB): 20mM Tris-Cl, pH 7.3, 1mM $MgCl_2$, 1mM $SpmCl_4$, filtered to sterilize

Elution Buffer (EB): 5mM EDTA, sodium salt, pH 7.6

Alkaline Elution Buffer (AEB): 1mM EDTA, 10mM NaOH

10X Taq Buffer for PCR: 500mM KCl, 100mM Tris-Cl, pH 8.5, 15mM $MgCl_2$

10X TAE Buffer: 0.4M Tris-acetate, 10mM EDTA, pH 7.8

10X TBE Buffer: 0.89M Tris-borate, 0.89M Boric acid, 20mM EDTA, pH 7.8

6X XBS Loading Buffer: 0.25% Bromphenol Blue, 0.25% Xylene Cyanol, 40% sucrose

DNA Staining Solution: Ethidium Bromide 1mg/L in 1X TAE, not sterilized

TE Buffer: 10mM Tris-Cl, 1mM Na_2EDTA, pH 7.5

TBM Buffer: 90mM Tris-Borate, 10mM $MgCl_2$, pH 7.3TE: 10 mM Tris-Cl, 1mM EDTA, pH 8.0

SDB (Standard Duplex Buffer): 10mM sodium phosphate, 1M NaCl, pH 7.4

TMA: 50mM sodium acetate, 3mM magnesium acetate, pH 5.0

SMN: 20mM MES, 3mM $MgCl_2$, 100mM NaCl, pH 6.0 (triplex favoring),

MMS: 20mM MES, 1mM $MgCl_2$, 1mM spermine tetrachloride, 10mM NaCl, pH 6.0

3M Sodium Acetate, pH 5.0

50mM $CaCl_2$, 10mM Tris-Cl, pH 8.0

100mM IPTG, not sterilized, kept at -20°C

2% X-Gal in N', N-dimethylformamide, kept in dark at -4°C

YT medium: 8g/l Bacto-Tryptone, 5g/l Bacto-Yeast Extract, 5g/l NaCl, pH 7.5

YT agar for plates: 12g/l Bacto-Agar in YT

YT top agar: 6g/l Bacto-Agar in YT

M13 Cloning Kit reagents [116]

Taquence Cycle-Sequencing Kit reagents [117]

Consumables and Chemicals

Acetic Acid Glacial: Malinckrodt, F.W. 60.0

Acrylamide: Sigma, 99%, F.W. 71.1

Bacto-Agar: Difco, for culture media

Agarose: Bio-Rad Laboratories, ultra pure DNA grade, gelling temperature 36°C

*Bam*HI Restriction Endonuclease: NEB, 20units/µl, 10X buffer included

Boric Acid: Aldrich, 99.99%, F.W. 61.8

N,N'-Methylene-bis-acrylamide: Sigma, F.W. 154.2

Cacodylic acid: USB, sodium salt, trihydrate, F.W. 214

Calcium Chloride: Fisher, dihydrate, F.W. 147

Chloroform: Fisher, 99.9%, F.W. 119.4

Deoxyribonucleic Acid: Sigma, from calf thymus, sodium salt, highly polymerized

Deoxynucleoside-5'-triphosphates: USB, dATP, dTTP, dCTP, dGTP, 25 µmoles each

N, N-Dimethylformamide: Malinckrodt, analytical reagent, F.W. 73.1

*Eco*R I Restriction Endonuclease: NEB, 20units/µl, 10X buffer included

EDTA (Ethylenediaminetetraacetic Acid): Fisher, 99.6%, F.W. 292.2

Ethyl Alcohol: Midwest Grain products Co., 95%, F.W. 46.0

Nylon-66 Alltech Membrane Filters, 47 mm diameter, 0.45 micron pores

Gelatin Powder: Aldrich, USP grade

Glucose: EM Science, monohydrate, F.W. 198.2

Glycerol: Baker, 95%, F.W. 92.1

Glycogen: USB, from rabbit liver

Hind III Restriction Endonuclease: NEB, 20units/µl, 10X buffer included

Hydrochloric Acid: Fisher, 36.5-38%, reagent A.C.S., F.W. 36.5

Isopropyl Alcohol: Malinckrodt, spectrophotometric grade, F.W. 60.1

Iso-amyl Alcohol: Fisher, 98%, F.W. 88.2

M13 Cloning Kit: Bethesda Research Laboratories Life Technologies, Inc. It included:

M13mp18 RF DNA, 5 µg

M13mp 19 RF DNA, 5 µg

T4 DNA ligase, 100 units

Ligase Dilution Buffer, 0.5 ml

5X Ligase Reaction Buffer, 1.0 ml

X-gal, 100 mg

IPTG, 1 g

E. coli strain JM101, 1 BACTI-DISK

Magnesium Chloride: Sigma, 1M solution, sterile filtered

MES: Sigma, F.W. 195.2, pK_a=6.1

Methanol: Fisher, HPLC grade, F.W. 32.0

0.5 and 1.6 ml microcentrifuge tubes and all pipette tips from Phoenix Research Products, autoclaved to sterilize

Nitric Acid: Fisher, 69-71%, reagent A.C.S, F.W. 63.0

[γ-^{32}P]ATP: Amersham, 37TBq/mmol, 1000Ci/mmol

AmericanCan Company Parafilm M laboratory film, sterilized for 15 minutes under 254nm UV

VWR Polystyrene Disposable Sterile Petri Dishes, 100x15mm

Phenol: Fisher, reagent grade, F.W. 94.1

Potassium Phosphate, Dibasic: Aldrich, F.W. 174.2

Potassium Phosphate, Monobasic: Malinckrodt, F.W. 136.1

[^{35}S]ATPγS: Amersham, 185TBq/mmol, 5000Ci/mmol

Sodium Acetate: Analytical Reagent, trihydrate, F.W. 136.1

Sodium Chloride: Baker, USP grade, F.W. 58.5

Sodium Hydroxide: Fisher, 50%, reagent grade, F.W. 40.0

Sodium Phosphate, Dibasic: Baker, USP grade, F.W. 142.0

Sodium Phosphate, Monobasic: Baker, USP grade, F.W. 138.0

Spermine Tetrahydrochloride: Sigma, 98%, F.W. 348.2

Sterile single use Becton Dickinson Luer-LOK 3cc Syringes

Sterile single use Samco 3 ml Transfer Pipettes

Sulfuric Acid: Malinckrodt, 96%, F.W. 98.0

Taq DNA Polymerase: USB, 5units/µl

Taq DNA Polymerase Antibody: Clontech, TaqStart Antibody

Taquence Cycle Sequencing Kit: USB 71075. It included:

- *Taq* Enzyme Dilution Buffer, 0.2 ml
- *Taq* Reaction Buffer, 0.3 ml
- T4 Polynucleotide Kinase (PNK), 500 units
- 10X PNK reaction Buffer
- Control DNA pUC19, 1µg
- Universal Cycling Primer, 400pmoles
- Reverse Cycling Primer, 400pmoles
- ddG, ddA, ddT, ddC Termination Mixes with dGTP, 0.42ml each
- ddG, ddA, ddT, ddC Termination Mixes with 7-deaza-dGTP, 0.42ml each
- Mineral Oil, light, 1.2 ml
- Stop Solution, 2.5 ml

Tris Base: USB, ultrapure, F.W. 121.1

Bacto Tryptone: Difco, for culture media

Urea: USB, enzyme grade, F.W. 60.1

Xylene Cyanole FF: Eastman Kodak, technical

Bacto Yeast Extract: Difco, for culture media