IN-VITRO SELECTION OF CHARGE CONDUCTING DNA SEQUENCES

by

Yi-Jeng Huang
B.Sc., Simon Fraser University, 2002

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

In the Department of
Molecular Biology and Biochemistry

SIMON FRASER UNIVERSITY

Spring 2006

APPROVAL

Name: **Yi-Jeng Huang**

Degree: **Master of Science**

Title of Thesis: **In-Vitro Selection of Charge Conducting DNA sequences**

Examining Committee:

Chair: **Dr. L. Craig**

Assistant Professor, Department of Molecular Biology and Biochemistry

---

Dr. D. Sen

Senior Supervisor
Professor, Department of Molecular Biology and Biochemistry

Dr. P. Unrau

Committee Member
Assistant Professor, Department of Molecular Biology and Biochemistry

---

Dr. M. Leroux

Committee Member
Associate Professor, Department of Molecular Biology and Biochemistry

---

Dr. H. Yu
Internal Examiner
Associate Professor, Department of Chemistry

Date Defended/Approved: April 6, 2006

ABSTRACT

Double helical DNA transports electrical charge, via hole transport, over distances of over 200Å (58 base pairs). Because the exact mechanism of hole transport is under debate, an alternative method to the current "rational" design approach used by investigators may shed new light on the problem. *In vitro* selection was used to isolate DNA sequences most efficient at hole conduction out of a random pool. Several selections were carried out, each providing valuable information for subsequent selection design. Tests carried out on the Round 7 of the SYB series confirmed that selection of superior conductors was indeed occurring. 93 clones sequenced from Round 7 were analyzed against 89 clones from the initial round and found to have statistically elevated levels of specific purine motifs segregated to only one specific strand of the duplex. These sequences, termed "transport permissive candidate motifs", suggested that Phonon-assisted Polaron-like hole hopping occurs in duplex DNA.

Keywords: Hole Transport, Charge Transport, Multi Step Hole Hopping, Phonon assisted Polaron like Hopping, Purine Segregation, In vitro selection.

DEDICATION

This thesis is dedicated to my wonderful family and friends. Thank you for walking with me during life's toughest hour. I will always remember my origins, and grateful.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude and respect for Dr. Dipankar Sen for his unwavering support and guidance during the course of my graduate studies, and for finally teaching me the correct spelling of "piperidine". Dr. Sen is worldly, inquisitive and well-spoken, a role model for budding scientists everywhere. I am grateful for the opportunity to work his laboratory.

To my committee members: Dr. Peter Unrau for his useful and often humorous discussions regarding a wide range of topics from quantum flux to statistics. Dr. Michel R. Leroux for his insightful critique and his biologically oriented perspective. To my examiner Dr. Hogan Yu, for his helpful discussions, encouragement, and support in troubleshooting the thesis.

Members of the Sen Lab past and present for their friendship and insightful discussions: Richard Fahlman, Anat Feldman, Dennis Wang, Dan Chinnapen, Bryan Andrews, Yong Liu, Ed Leung, Carlo Sankar, Becky Thorne, Hyun-Wu Lee, Janet Huang, and Gupreet Sekhon.

To my friends, the pillars of my life: Dave Toews, Caroline Wang, Violet Cheung, Nadia Kasenda, Bobby Park, Desiree Essen, Winnie Fung, Jacqueline Woo, the Round Table Gang (Dean Wang, Scott Emery, Brad Oldham, Hong Yan, Chris Huang, Tyr Fothergill) and the ACE SFU Team 04-05. We may be scattered geographically, but we are together in our hearts.

Last but not least, my family, to whom I owe everything to: my father Chia-Shian Huang, my mother Chun-Ya Huang, my brother Yi-Hsin Huang, and my grandparents. Thank you all.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

8-oxoG	7,8-dihydro-8-oxoguanine, the primary guanine oxidation product.
Å	Angstrom (10^{-10} meter)
A	Adenine
Bp	Base Pairs abbreviation
C	Cytosine
CA	"Carlo Arms", a test series made from sequences used successfully by Carlo Sankar in the past for his experiments.
CL	Confidence Level (usually 95%)
ddH_2O	Distilled and Deionized water
ETS	ETS test construct to compare the use of GGG vs 8-oxoG for detection stem.
G	Guanine
HOMO	Highest Occupied Molecular Orbital
MSH	Multi-Step (Hole) Hopping Mechanism
N15	Random Region 15 nucleotides in length
nt	Nucleotides
PAGE	Poly-Acrylamide Gel Electrophoresis
PPH	Phonon-Assisted Polaron-like Hopping Mechanism
R0	Round 0 selection abbreviation, initial pool.
R7	Round 7 selection abbreviation.
RT	Room temperature, usually from 19°C~21°C
SE	Standard Error (obtained by dividing standard deviation by square root of n, where n is sample size)
SX	Selection X Series, first selection done, worked out construct design problems.
SY	Selection Y Series, second selection done, had positive control contamination.
SYB	Selection Y-B Series, third selection done, isolated transport permissive candidate motifs.
SYK	"SY Known". Name of SY Series positive control construct. Had defined sequence instead of random region.
SYN	"SY N15". Name of SY Series selection construct. Consisted of a pool with a variety of sequence in the random region.
T	Thymine
TD	Thymine Dimers

1 BACKGROUND OF CHARGE CONDUCTION IN DOUBLE-STRANDED DNA

DNA is generally thought of as the biological molecule used for the storage and propagation of genetic information in all free-living organisms. The accepted biological paradigm states that the flow of genetic information occurs from DNA→ RNA→ PROTEINS. Even a cursory glance at the double helical structure proposed by Watson and Crick (Watson & Crick, 1953) shows how the strict base paring of Adenines to Thymines (A-T), Guanines to Cytosines (G-C), can facilitate replication of complementary strands to a high degree of accuracy. This fidelity and stability make DNA an ideal molecule for storing hereditary information, or genes, in living organisms.

As a polymer, DNA is fascinating for reasons other than simply storing genetic information. In an *in vitro* setting, without the usual constraints of associated proteins present in live organisms, single stranded DNA can fold into complex 3-dimensional structures and can be used to construct multi-unit arrays (Fahlman, 2001). Through the use of *in vitro* selection (also known as Systematic Evolution of Ligands with Exponential enrichment – SELEX), DNA molecules with specific ligand binding properties, termed "aptamers" (Famulok et al. 1993) or even catalytic abilities (DNAzymes and RNAzymes) have been found (Silverman, 2005).

In the last 15 years, considerable effort has been invested in understanding the mechanisms of electrical charge transport in DNA (Section 1.3). There are two kinds of charge transport; positive charge carried by an electron deficiency (radical cation), known

as "hole" transport and negative charge carried by electrons, (radical anion) called "excess electron" transport. There has been considerably more experimental work done on the nature of "hole" transport compared with excess electron transport. Hole transport is thought to have an important physiological role in aging and cancer, and is used extensively in the field of nanotechnology (Section 1.4).

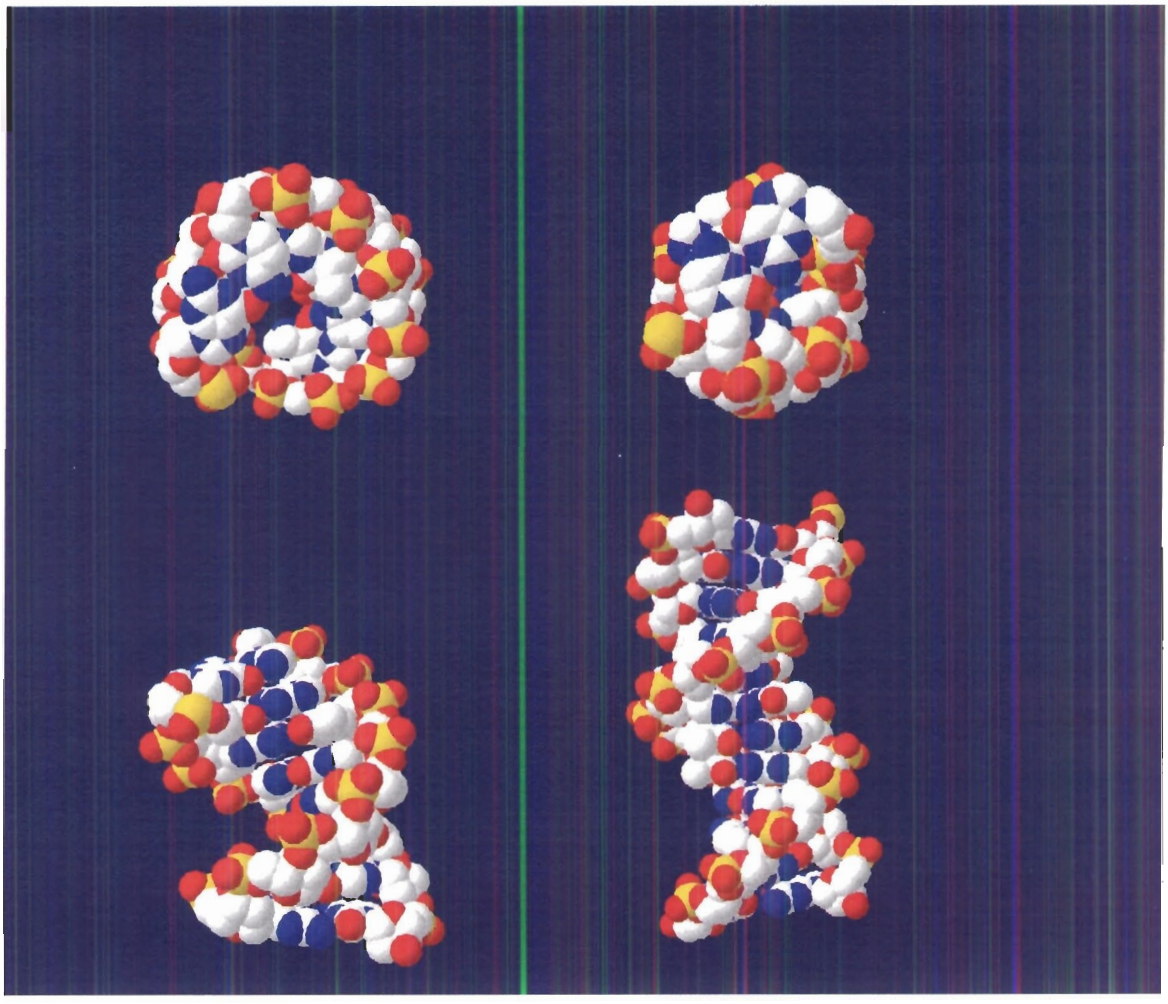
1.1 Structure and Property of Double-Stranded DNA

The right-handed B-Type (Fig 1.1) double-helix DNA structure is the most common form, found in most organisms. It forms under neutral pH and physiological salt concentrations. The B-Type helix has a major groove and a minor groove, which in a biological setting provides easily differentiated specific binding sites for proteins and other ligands. It has an average base per turn value of ~10.4.

Another form of right-handed double-helix that DNA can form when hybridized with RNA, is the A-Type helix (Fig 1.1). RNA is constrained by the 2' hydroxyl group, which discourages the formation of B-Type. The A-Type helix is more compact, the major groove is less accessible to solvent and the individual bases are more deeply buried within the helix. Both A and B-Type helices can be comprised of any combination of bases as long as Watson-Crick complementarity is observed.

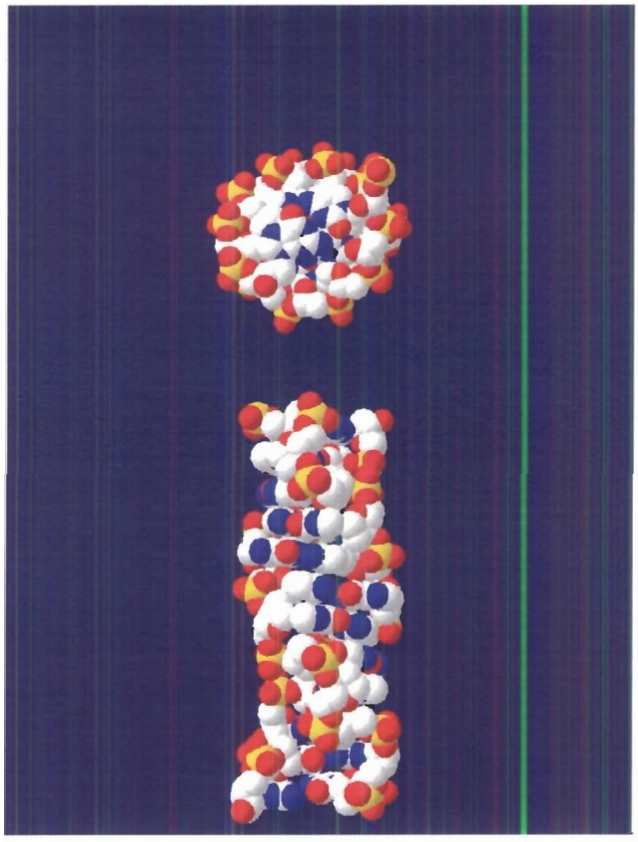
The Z-type helix (Fig 1.2) is formed *in vitro*, under high salt conditions (Herbert and Rich, 1999).It is highly distinct from A and B-type helices because it is left handed. Only specific sequences of bases permit Z-type helices, namely alternating purine-pyrimidine repeats.

Figure 1.1: DNA Helices. On the right is B-type helix, DNA (5-D(CpGpCpGpApApTpTpCpGpCpG)-3) PDB: 2BNA. On the left is A-type helix, Dna (5'-D(GpCpGpApApTpTpCpG)-3') PDB: 208D



Source: DeepView/Swiss-PdbViewer 3.7 and PDB Database

Figure 1.2: Z-Type DNA, (5'-D(M5cpGpG)-(Gpm5cpG)-3' 5'-D(M5cpGpCpCpm5cpG)-3') PDB: 145D



Source: DeepView/Swiss-PdbViewer 3.7 and PDB Database

1.2 History of Charge Transport

Electrical conduction is defined as the movement of electrically charged particles through a transmission medium. The method of transmission is highly dependent upon the medium, or the material used to conduct electricity. Molecular orbitals formed from the valence electronic orbitals of metal atoms allow electrons to move freely around the metal crystal, thus, making metals good conductors (Zumdahl, 1995). In contrast to metals, which only have partially filled valence electrons, carbon based molecules have filled valence orbitals. Under some conditions, electrons may be thermally excited from

the valence band to the next higher conduction band, making it conductive (Rakitin et al., 2000).

The notion of conductivity through DNA (a carbon based polymer) was raised more than 40 years ago (Eley & Spivey, 1962) not long after the discovery of DNA helical structure by Watson and Crick. It was noted that the base pair stacking within the double helix showed a resemblance to one-dimensional aromatic crystals. Aromatic crystals have been reported to have high charge carrier mobility (Schoonveld, et al. 1998). There are however, some differences between aromatic crystals and DNA. Unlike aromatic crystals, where the aromatic disks are all exactly the same, DNA is comprised of nonperiodic stacks of A:T and G:C base pairs. Due to the anti-parallel helical nature of duplex DNA, 3'->5' base stacking differs from 5'->3' stacking, again different from the ordered patterns of aromatic crystals.

Initial experiments in the '90s studying hole transport through double stranded DNA showed contradictory results. On one hand, DNA aromatic groups were reported as an efficient conductor by Barton's group (Murphy et al. 1993). On the other hand DNA was determined to act as an insulator (Lewis, 1997, Debije, 1999). Later reports observed a range of conductivity, but mostly acknowledged that hole transport did occur to some extent (Giese et al. 2002, Schuster et al. 2000).

Hole transport is studied most conveniently by covalently attaching an electron acceptor to a well defined, double stranded DNA (or to other complex DNA constructs). A radical cation generated after chemical or optical excitation then travels through the double stranded DNA, until it is quenched by an electron donor. The resulting radical cation is reported to travel from 5'->3' distances of 200Å (58 base pairs) or more

(Schuster, 2004). In the case of Barton's 1993 paper, a pair of metallointercalator electron donor/acceptors (Rhodium [donor] and Ruthenium [acceptor] complexes, Fig. 1.3) were used to measure conduction efficiency. If the electron donor is DNA itself, then the oxidative damage is measured by inducing strand breakage at oxidation sites, and measuring the relative yield at the various positions along the DNA strand (Sanii et al., 2000, Giese et al., 2002, Nakatani et al., 2000, Sankar et al., 2004). Although the extent and exact mechanism of the hole transport is still under debate, investigators agree that the electrical conductivity of DNA is highly dependent on its conformational state; DNA is efficiently conductive only when the π-stacking of the helix is intact.

1.3 Significance of Charge Transport

Oxidative damage has been shown to play an important role in several types of cancer and in aging (Piette, 1991). Radicals generated by environmental factors such as pollution or radiation, and physiological factors, such as respiration, cause severe damage to DNA by inducing base and sugar modifications, or DNA-protein crosslink (Friedberg et al, 1995). 8-oxoG, for example, is a common product of guanine oxidation, and can mispair with adenine, inducing a potentially cancer-causing, G to T mutations in mammalian cells (Besaratinia et al. 2004). Oxidative damage to mitochondrial DNA, lipids and proteins by cellular respiration (which takes place in mammals inside the mitochondria) was thought to cause progressive cellular dysfunction, and was closely associated with age related human pathologies (James et al., 2005). Since charge transport causes some oxidative damage to bases, an interesting question to pose, is whether or not charge transport has a role to play in cancer and aging.

Even as early as 1998 (Núñez et al., 1998), Barton's group has postulated the existence of biologically relevant G-rich "sinks" to draw oxidative damage away from critical DNA coding regions, and possibly "insulating areas" of the genome to shield these regions against hole transport. They furthered this hypothesis by demonstrating that charge transport occured in the DNA of Hela cells (Núñez, et al. 2001). Using an intercalating rhodium photooxidant capable of inducing hole transport, Núñez et al showed that even at protein-bound sites (which prevented intercalation by rhodium) oxidative damage was observed on the 5'guanine in a double guanine, a hallmark of hole transport (See Section 1.4.3 for more details regarding guanine oxidation). However, there is no current evidence to conclusively prove the existence of the postulated "sinks" and "insulation" elements in the genomes of living organisms. A recent paper (Kanvah, 2005) also suggested that while special sequences of DNA containing the 8-oxoG charge sink may protect distal GG from being oxidized, this does not work in a mixed DNA sequence, limiting its application in life to act as a natural damage "sink".

Charge transport also has applications in the field of nanotechnology and molecular electronics. Sen's group created DNA based molecular switches and sensors (Fahlman et al. 2002). Braun et al deposited silver on DNA as a template, creating silver wires; this has applications in nanoscale circuit construction (Braun et al. 1998). Ben-Jacob's group (Hermon, et al, 1998) suggested that phosphate groups can act as tunnel junctions, and DNA could be used as a transistor. For a more comprehensive review, refer to DNA Electronics in the in Encyclopedia of Nanoscience and Nanotechnology (Ventra & Zwolak, 2004).

1.4 Monitoring Hole Transport.

Hole transport studies generally used a covalently tethered electron acceptor, a length of double stranded DNA, and an electron donor, usually DNA itself. Section 1.4.1 described some of the large varieties of photoactivated electron acceptors used in hole transport studies. Section 1.4.2 described the mechanism of Anthraquinone, the electron acceptor chosen for our experiments. Section 1.4.3 described various oxidation products of the final electron acceptor guanine.

1.4.1 Photoactivated Electron Acceptors

Key to the study of photoinduced one-electron oxidative damage and hole transport mechanisms are light-activated oxidizing agents (termed photosensitizers or photooxidants). These photosensitizers are usually tethered to the end of a DNA duplex terminus, or intercalated between base pairs. Upon irradiation at a wavelength corresponding to the difference between the ground state and excited state of that particular sensitizer, light is absorbed and the photosensitizer undergoes electronic excitation. Many photosensitizers have a singlet ground state so that, upon irradiation, they are excited into the next most energetic singlet level. This excited state energy can be dissipated in many ways including fluorescence (re-emitting of a lower energy light) and conversion to heat (increased vibrational motion). If the singlet state survives long enough, it may be converted to the lower energy "excited triplet state" through what is called an intersystem crossing. Simple conversion from the triplet state back to ground singlet state through radiative processes (light or heat) is highly improbable since it is forbidden under quantum spin conservation rules. This triplet however, may convert back to ground state by transferring the energy to a *second* molecule provided the energy

levels match. In the case of DNA, light energy absorbed by the sensitizer can be passed on to double stranded DNA as a radical cation. The photosensitizer, having returned to its ground state, is available for re-excitation.

Barton's group specialized in the use of rhodium and ruthenium (Fig 1.3) intercalators for their studies (Núñez et al., 1998). However it has been shown that under certain conditions, their appended DNA duplexes aggregated and the resulted complex kinetics of hole flow obscured the experimental results (Fahlman et al., 2002). Trioxatriangulenium (TOTA+) ion is a GC pair-preferring intercalator (Fig 1.4) (Giese, 2002), used by Giese to inject radical cations into double-stranded DNA. Irradiated TOTA+ is however, relatively inefficient as a sensitizer because it reacts from its singlet state. Excited singlet states have a tendency of dropping back down to ground state by releasing heat or light instead of passing the energy onto DNA (Krebs et al., 1998).

Anthraquinone (AQ) (Fig 1.5) covalently linked to 5'-end of DNA via a C6 linker provides a convenient way of studying long-range charge transport because the charge injection site is clearly defined. Spectroscopic evidence suggests that such AQ is "end-capped", which means the π-orbitals of AQ is in within π-contact with the base pairs at the end of a duplex DNA (Gasper & Schuster, 1997). AQ is a convenient photosensitizer for many reasons. It is not overly sensitive to ambient light, which makes handling easier, it absorbs at 334nm, away from the 200-300nm UV region that is absorbed by and damaging to DNA, and is relatively easy to covalently couple to double-stranded DNA. For all the above reasons, it was chosen by us for the exploration of sequence effect of long-distance charge transport.

Figure 1.3: Rhodium (III) complex (top) and Ruthenium (II) complex (bottom).

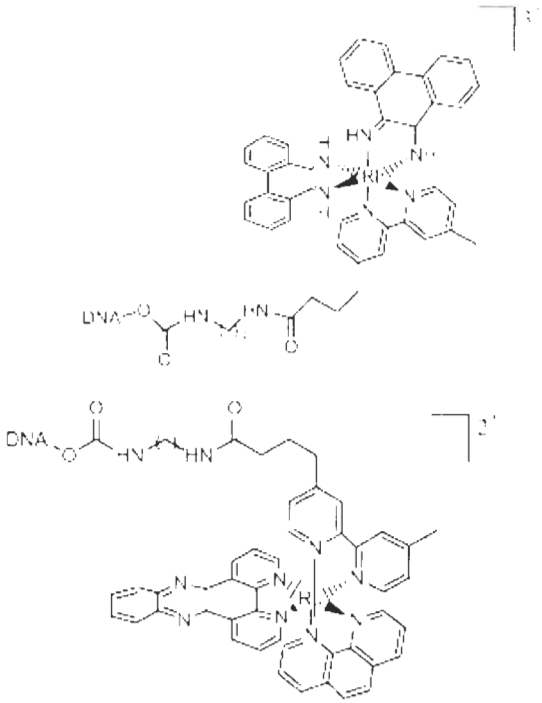


Figure 1.4: Trioxatriangulenium (TOTA+) ion.

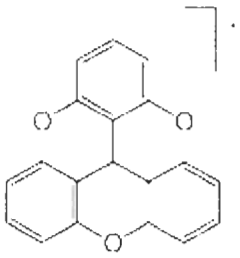
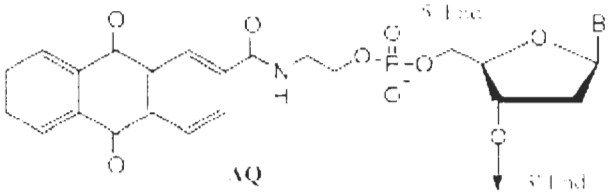


Figure 1.5: Anthraquinone derivative.

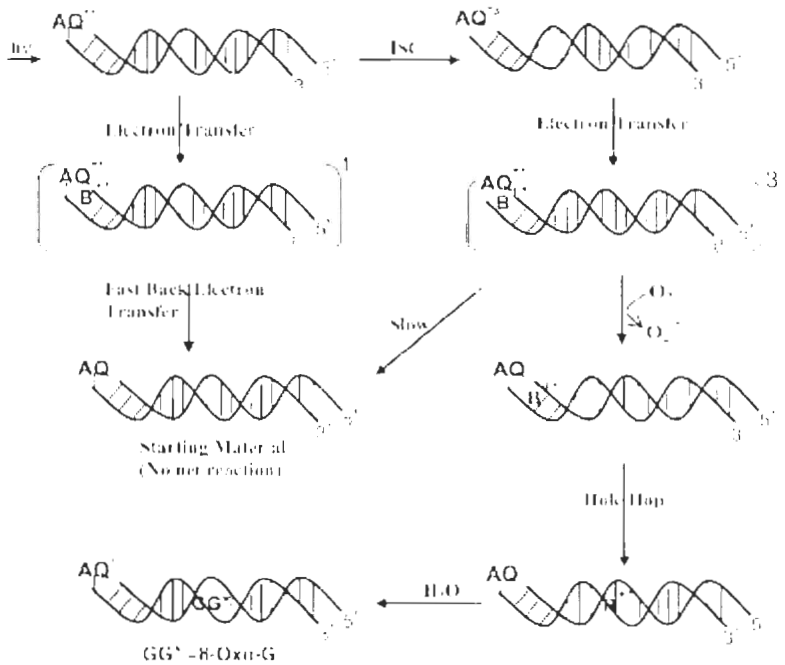


Source: Fig. 1.3, 1.4, 1.5 were reproduced with permission from G. B. Schuster and the Royal Chemistry Society, Chem. Commun., 2005, 2778–2784.

1.4.2 Anthraquinone Hole Transport Investigation System.

As Fig 1.6 shows, after excitation with 334nm UV light, the covalently tethered AQ either forms a singlet exited state, which may drop to the ground singlet state or intersystem cross to form a triplet state. The excited singlet state has a lifetime of mere nanoseconds. The triplet state however, is long-lived (microsecond timescale) (Görner, 2002) since back electron transfer is forbidden by quantum spin conservation rules Fig 1.6. This extra time allows the excited AQ radical anion to react with molecular oxygen. AQ then reverts back to ground state, the solution gains an extra superoxide molecule, and an electron hole is injected into the duplex DNA.

Figure 1.6: AQ charge injection mechanism. The singlet excited state (top left) either reverts back to ground state (second row left, third row left) or intersystem crosses to form the triplet state (top right). The excited AQ radical anion reacts with molecular oxygen (second row right). AQ then reverts back to ground state and an electron hole is injected into the duplex DNA (third row right). Hole migrates along double helix (bottom left). A small proportion of guanine radical cations may react with water, causing irreversible oxidative damage to the guanine base (bottom right).



Source: Reproduced with permission from G. B. Schuster and the Royal Chemistry Society, Chem. Commun., 2005, 2778–2784.

DNA sequence effects on the initial charge injection efficiency of an excited photosensitizer have been studied extensively. A run of adenines at the charge injection site (called an adenine bridge) allowed end-capped photosensitizers such as AQ to inject the charge into the double helix efficiently more efficiently (Majima et al., 2004, Sanii & Schuster, 2000). Sanii et al, showed the effects of varying the sequence of the double stranded DNA just after a 5' covalently-tethered AQ. Out of four different sequences placed adjacent to the AQ (AQ~TTTA/AAAT, AQ~AAAT/TTTA, AQ~GAAT/CTTA, AQ~CAAT/GTTA), the sequence determined to allow the most efficient charge injection (and subsequent hole transport to distal GG regions) is "AQ~AAAT/TTTA". Kawai et al provided evidence that, if a photosensitizer was able to oxidize adenine, a rapid adenine-hopping mechanism (similar to the guanine-hopping mechanism explained later) would enforce a charge separation (Kawai et al., 2003). Once the charge proceeded past the adenine bridge (and on to a guanine), the resulting guanine radical cation formed cannot readily re-oxidize the adenine bridge, preventing hole transport back to the original photosensitizer.

1.4.3 Electron Donor, Guanine and its Oxidation Products

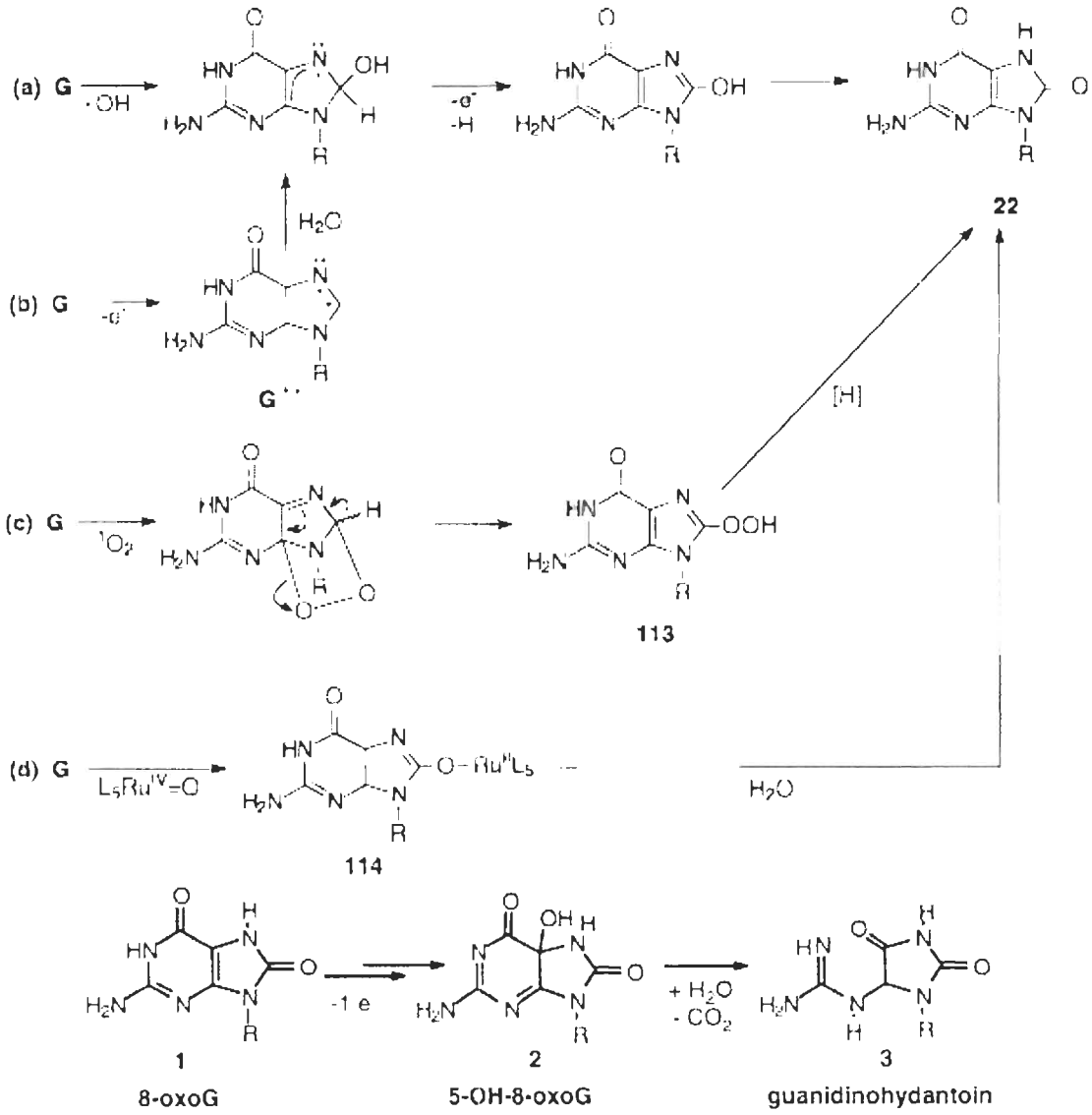
After a photosensitizer such as AQ successfully injects a positive charge into the DNA helix, the hole travels up and down the helix in the form of a radical cation (the specific mechanism of hole transport is discussed in Section 1.5). The reduction potential of free nucleotide bases are: dG = 1.29V, dA = 1.42V, dC = 1.6V and dT = 1.7V (V vs NHE) (Seidel et al, 1996) (oxidation order is expected to remain the same in duplex DNA

as well). As the base with the lowest oxidation potential, guanine radical cations readily forms during hole transport. As Fig. 1.6 bottom, right illustrates, a small proportion of guanine radical cations may react with water, causing irreversible oxidative damage to the guanine base.

When the guanine radical cation reacts with water, 7,8-dihydro-8-oxoguanine (8-oxoG) may be formed (Fig, 1.7 Top (b)) (Hickerson & Burrows, 1999). If 8-oxoG is further oxidized, the results yield 5-OH-8-oxo-G and guanidinohydantoin (Fig. 1.7 Bottom) (Hickerson & Burrows, 1999), both of which are highly piperidine labile. These damaged bases can be identified following a piperidine mediated cleavage of the DNA and damage at specific sites quantified (Sankar et al., 2004, Sanii et al 2000) using a denaturing PAGE. The location and extent of guanine oxidation form the basis for quantifying hole transport.

However, not all guanines react to the same extent. Experimental evidence on double-stranded DNA has shown that a 5' G would react more if its 3' neighbour was a purine rather a pyrimidine. The observed order of reactivity of G sequences (from 5'->3') is G<u>GG</u>G > <u>GG</u>G > <u>G</u>G > <u>G</u>A > <u>G</u>T = <u>G</u>C, where the underlined bases are the principal sites of cleavage (Spassky & Angelov, 1997, Muller , 1997, Nakamura, 1996). Saito's group explained that 5' Gs of GG sequences are more reactive than single Gs, as being due to the π-stacking effect of purines in which the HOMO resides mostly in the 5'G. Not coincidentally, 5' guanines of double or triple Gs have lower oxidation potential than single Gs (dG<u>dG</u>= 0.82V, dGdG<u>dG</u>= 0.64V, G oxidation product, 8-oxoG = 0.58V) (Burrows & Muller, 1998). This is the reason triple guanines are useful as a hole transport detector.

Figure 1.7: **Products of Guanine Oxidation. Top, path (b): After guanine radical cation (G+•) reacts with H_2O, another one-electron oxidation event leads to the formation of 8-oxoG (22). Bottom: Another one-electron oxidation event yields 5-OH-8-oxoG and guanidinohydantoin, two piperidine labile products.**



Source: Reproduced with permission from C. J. Burrows, Chem. Rev. 1998, 98, 1109-1151.

1.5 Charge Transport Theories

The two most significant models describing hole transport theories have emerged over the last decade. First is the Unistep Superexchange mechanism (Section 1.5.1), and the other is the Multi-step Hole Hopping mechanism (Section 1.5.2). A refinement of this latter model is the Phonon Assisted Polaron-like mechanism (Section 1.5.3). Section 1.5.4 describes what is known about excess electron transport.

1.5.1 The Unistep Superexchange Mechanism

The Unitstep Superexchange mechanism was advanced by Barton et al. when they discovered rapid fluorescence quenching using organometallic intercalators groups (Murphy et al., 1993, Turro & Barton, 1998). The photoexcited ruthenium (II) electron donor showed luminescence when intercalated into DNA. Rhodium (III) intercalator, capable of hole injection into DNA, was used as the electron acceptor. *When both compounds were intercalated in a defined sequence of duplex DNA, the photactivated rhodium complex was able to quench the luminescence of ruthenium, mediated by stacked DNA bases.* Since most of the luminescence was quenched in nanoseconds, they proposed that long distance hole transport could occur in time frames as fast as picoseconds, which would be consistent with quantum tunnelling in "electronically coupled, stacked bases".

They therefore proposed, that DNA acted as a 'molecular wire', where the π-orbitals from the aromatic bases form a delocalized continuous molecular orbital. Hole transport from donor to acceptor would occur by rapid single-step quantum tunnelling (up to 40Å distances). Electrons can be described as a wave function, where the wave represents the probability of finding the electron at a certain location. When the mobile

electron is not energetic enough to access the unoccupied delocalized molecular orbits of the bridge (in this case, the multi-base bridge), tunnelling describes the probability that some electrons will be propagated to the other side. This process is expected to decay rapidly as a function of distance because double-stranded DNA is dynamic over longer distances (Borer et al, 1994).

Although initially provocative, this mechanism does not explain most recent observations regarding charge transfer over long stretches of DNA over 200Å. Using a similar donor quenching system, Barton et al. found that different base sequences had a dramatic effect on charge transport (Barton et al, 2000). Both the time scale differences (10ps-in fast sequences and up to 512ps in slower sequences), as well as the sequence dependent hole transport efficiency suggested the invalidity of the 'molecular wire' as a viable explanation for long range charge transport in DNA. This mechanism however, is still useful in explaining and measuring short distance quantum tunnelling.

1.5.2 The Multi-Step Hole Hopping Mechanism and the Phonon Assisted Polaron-like Hopping Mechanism

Two primary mechanisms have been proposed to account for long-range radical cation transport.

The Multi-Step Hole Hopping Mechanism (MSH) advanced by Giese (Bixon et al., 1999) proposed that electron holes hop between sequential guanines by tunnelling through the AT bridge sequences in between. The process begins with an oxidation event, where a guanine radical cation is formed. Since guanine has the lowest oxidation potential, it would only oxidize guanines, not other bases. The ability of the guanine radical cation to oxidize neighbouring guanines allows holes to hop randomly from

guanine to guanine in a thermally activated process. Giese's group irradiated a modified base containing a 4' acylated group (t-butyllithium ketone) to create a sugar radical cation, which then travelled through a defined double stranded DNA to a GGG detector. Guanine oxidation was measured using piperidine mediated DNA strand cleavage. Probing constructs with varying distances from charge injection site to the detector (ranging from 5bp to 16bp) they found that the rate of hole transport was dependent upon the length of guanine gaps, but relatively independent of the identity of the bridging sequences.

The Phonon-assisted Polaron-like Hopping Mechanism (PPH) advanced by Schuster's group (Schuster et al., 2000) was a refinement of the multi-step hopping idea. But while multi-step hopping described the *guanine radical cation* as the sole carrier of charge, polaron-like hopping describes the presence of local energy-minimizing DNA distortions to distribute the charge. Unlike a true polaron, which traps a radical cation indefinitely, the structural distortions in this case cannot trap radicals indefinitely, hence the name "polaron-like". Schuster used phonons in the context that thermal motions of base pairs helped create the energy-minimizing polaron; hence "phonon-assisted". Using a tethered AQ system much like what was described in Section 1.4.2, Schuster's group probed a series of well defined DNA constructs with various repeating sequences. A row of double guanine (GG) detectors along the entire length of the construct was used to evaluate the hole transport progress through the duplex DNA. They found that, purine segregation onto either strand of a duplex promoted hole transport much more effectively than mixed purine/pyrimidine strands (Liu. et al., 2003).

The critical difference between the two long-distance charge transport theories is the process by which a radical cation centered upon a guanine migrates to the next guanine-centered location. The multi-step hopping assumes that the radical cation is localized only on individual guanines; never on the bridging AT sequences. From the context of this mechanism, distance between guanines is the most important parameter for hole conductivity. The identity of the intervening AT bridges or even which strand the guanine resided on, is of minimal importance. In contrast, the Polaron-like hopping mechanism proposes that the radical cation exists as a detectable low-energy distortion in the bridging sequences. Thermal motions of the base pairs in or near the polaron distortion would cause the polaron to migrate randomly to the next guanine containing region. The extent of the distortion and delocalization of charge depends on the trade-off between the energy saved from the conformation change and the energy used to create the conformational change. Not all distortions are created equal in effectiveness; the polaron-like mechanism predicts hole conductivity efficiency as being highly sequence dependent.

There have been several recent publications detailing evidence in support the of polaron-like hopping model. High level quantum simulations identified an important role for cationic counterions and tightly bound solvating water molecules (Barnett & Schuster, 2001, Barnett & Schuster, 2003). This was consistent with the polaron-like mechanism. More illuminating are the N6-cyclopropyldeoxyadenosine probe experiments (Dohno & Saito, 2003), which concluded that radical cations resided on adenines long enough to cause ring opening of a cyclopropyl group attached to the nucleobase. This was evidence against the multi-step hopping mechanism because quantum tunnelling (as predicted by

multi-step hopping) would not have paused over the intervening AT bridges long enough to cause oxidative damage. Finally, according to the multi-step hopping model, it was possible for a guanine radical cation to oxidize the 5' guanine of a double or triple guanine sequence. But the newly formed 5'*guanine radical cation* should be *unable* to oxidize single guanines further down the duplex because it now has a lower oxidation potential. Many experiments by various groups as well as observations from our lab had seen charge transport proceed past doublets (Sankar et al., 2004, Liu & Schuster, 2003). In such cases the multi-step hopping mechanism states: a 'stepping stone' was formed that acted to reduce the ionization potential of the bridge and allow superexchange to occur between the two GG runs (Nakatani, 2000). This idea sounds remarkably like the energy minimisation strategies proposed by the polaron-like mechanism.

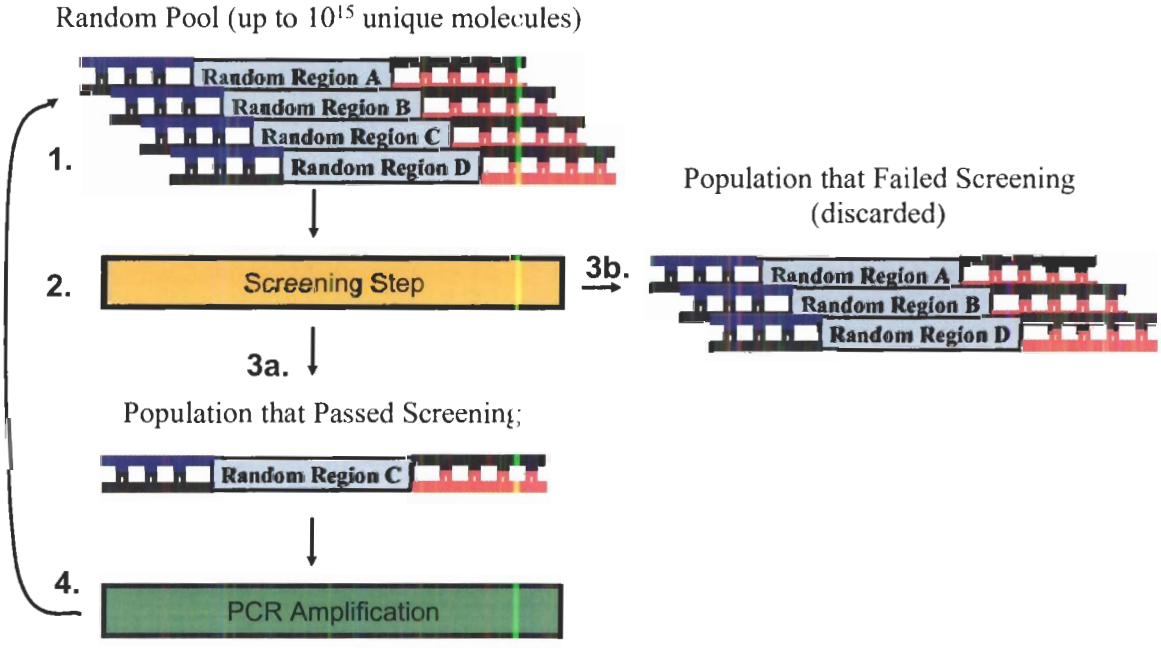
1.5.3 Excess Electron Transport

Much less is known about excess electron transport through DNA. Labs that have studied this phenomenon use light activated flavins or gamma irradiation of DNA for electron injection (Anderson & Wright 1999, Debije et al.,1999, Schwogler et al., 2000). Excess electron transport reduce C and T preferentially because of their accessible reduction potentials, dG = 1.29V, dA = 1.42V, dC = 1.6V and dT = 1.7V (V vs NHE) (Steenken & Jovanovic, 1997). Because a pyrimidine bases is present in every base pair, excess electron transport should proceed efficiently (Bixon et al. 1999). As an interesting note, thymine dimers have been known to be photoreversed by excess electron transport (Behrens et al. 2003).

1.6 *In vitro* Selection (SELEX) Theory

In the 1990s, the laboratories of Joyce (Joyce, 1989), Szostak (Ellington & Szostak, 1990), and Gold (Tuerk & Gold 1990) all independently developed a technique allowing parallel screenings of 10^{15} plus nucleic acids for specific functions (by contrast, the mouse antibody diversity is between 10^{9} to 10^{11}) (Famulok & Szostak, 1992, Famulok & Szostak,1993, Klug & Famulok, 1994). This method, known as *In vitro* Selection or SELEX (**S**ystematic **E**volution of **L**igands by **EX**ponential enrichment), has become an important tool in molecular biology. A simple schematic of an *in vitro* experiment is shown in Fig. 1.8.

Figure 1.8: ***In vitro* Selection Schematic. (1) A random pool is generated. (2) The random pool is subjected to a screening step to find molecules capable of a specific functionality. (3a) Molecules that are able to perform the function are PCR amplified and proceeds to the next round of selection (4). (3b) Molecules that are unable to perform the function is discarded.**



The method is conceptually straightforward. (1) Using automated DNA synthesis, a DNA oligonucletide is generated consisting of a random sequence region of defined length (called random region) flanked by two defined primer binding sites (called the conserved regions). A pool of these constructs will have a wide diversity of sequences in the random region, up to 10^{15} unique DNA or RNA molecules. This is called a selection pool. Past 10^{15} molecules, the volume and concentration required cannot be conveniently manipulated. (2) The selection pool is subjected to a screening test to screen for a desired functionality. The order of complexity of the selection pool justifies the assumption that a few molecules in the pool may fold in such a way as to possess the desired function. The current pool's overall ability to pass the screening test may be evaluated and used as a measurement of progress. (3a) The molecules able to pass the screen are selectively amplified in step (4). Those that do not pass the screening test are removed from the screening pool (3b). After many successive cycles, molecules able to pass the screening test are enriched enough to be cloned and sequenced. Typically the middle rounds of an *in vitro* selection experiment, one often sees an exponential increase in the overall pool performance, followed by a plateau in later rounds. The plateau indicates that the selection is ready for cloning and sequencing. These individual clones can then be analyzed and characterized.

2 *IN VITRO* SELECTION OF CHARGE TRANSPORT, MATERIALS AND METHODS

The materials and methods section encompasses four major topics. (Section 2.1) The *in vitro* selection methodology, the design of the selection library; (Section 2.2) procedures for oligonucleotide processing and purification; (Section 2.3) AQ coupling and HPLC purification; and (Section 2.4) the selection experimental protocols. More detailed descriptions can be found in the Appendix (Section 5.1).

2.1 *In vitro* Selection Methodology

The size of the random stretch of DNA within the selection construct is important in determining the scale of the initial library. The ETS and SX series (see Fig. 5.4 and Fig. 5.9 respectively) used by us had a random region of 10 base pairs (N10). The SY and SYB series had a random region of 15 base pairs (N15). Using N15 as an example, there are 4^{15} or 1.07×10^{9} different sequences of DNA possible. In order to encompass the whole range of combinations, and to ensure the greatest possible diversity, 1.8×10^{-15} mol (or 1.084×10^{9} molecules) of the starting random library is needed (since there are 6.023×10^{23} molecules per mole). To give a better perspective, even 1μL of 1μM (or 6.023×10^{11} molecules) initial library, for example, would cover all the possible combinations over 500 times over ([1×10^{-6}L x 1×10^{-6}Molar] x $6.022 \times 10^{23}/1.07 \times 10^{9}$ = 563). Note that a 0.2μmol (2×10^{-7}mol) DNA synthesis scale was used to create the initial library, whose diversity is over seven orders of magnitude higher than the minimum

required to retain all the possible combinations given N15. The overall length of the selection construct is constrained the by the 200Å (58bp) observed limits of charge conduction and the sizes of the conserved flanking regions used for PCR amplification (17 nucleotides).

2.2 DNA Oligonucleotides

All DNA oligonucleotides used for the selection was purchased from the University of Calgary Core DNA Services (http://dnaservices.myweb.med.ucalgary.ca/). The 0.2 μM chemical synthesis scale was used exclusively owing to superior quality control of DNA obtained at this scale. Random region synthesis was achieved by mixing 0.1M solutions of each of the four phosphoramidites into one bottle before injection into the automated DNA synthesizer (private communication with Dr. Pons from University of Calgary Core DNA services). We produced and coupled AQ to DNA. Thymine dimers were also made in-house.

2.2.1 Initial Round Libraries, Primers

Newly synthesized oligonucleotides in lyophilized form were first dissolved in 50 μL of denaturing gel loading buffer, heated for 90°C, then loaded on an 8% denaturing polyacrylamide gel (loading buffer consists of ~100% formamide, 10mM Tris-Cl (pH 8.0), 1mM EDTA, Xylenecyanol FF and Bromophenol Blue). Using UV shadowing, full length DNA gel bands were cut out and eluted into 10mL of 10mM Tris-Cl (pH 8.0), and 0.1mM EDTA. This size purification was especially important for synthesized oligonucleotides longer than 20 mer, since a significant amount of non-full length (error

prone) products were produced during chemical synthesis. After elution, the eluent was filtered through a 0.2 μm microfilter (Gelman Sciences), and concentrated to 1mL by 2-butanol extractions. The 1mL of purified DNA was then ethanol precipitated by making the solution to 300mM sodium acetate and the addition of 2.5x by volumes of 100% cold ethanol (standard procedures). Special processing of C6-NH_2 5'-modified oligonucleotides is explained in the Section 2.2.3 Anthraquinone Coupling and HPLC Purification section.

2.2.2 7,8-dihydro-8-oxo-deoxyguanine Processing

The synthesis of oligonucleotides containing 8-oxoG was carried out on an automated DNA synthesizer by University of Calgary DNA Core labs, using a standard synthesis protocol. The deprotection step for 8-oxoG was carried out within our lab. The DNA synthesis column was opened and the beads removed into a 1.3mL screwcap tube. 1mL of 30% NH_4OH (Ammonium Hydroxide) and 0.25 M 100% 2-mercaptoethanol was added into the screwcap tube, the tube sealed and heated for 17 hours at 55°C. The sealed tube was then cooled to freezing temperature to reduce internal pressure before it was opened and the supernatant removed (and discarded). Beads were washed with 300μL of ddH_2O (distilled and deionized) and this supernatant was transferred to a new tube. The supernatant was lyophilized to get DNA.

2.2.3 Thymine Dimer Formation

For thymine dimer formation, unlabeled T2 and 5' P^{32}-labeled T1 oligonucleotides were annealed, at 400μM concentrations, each to a complementary splint oligonucleotide (also 400μM) in buffered 40 mM $MgCl_2$. The DNA was irradiated in the presence of a 5

mM concentration of the triplet sensitizer acetophenone (shown specifically to favor the formation of cyclobutane thymine dimers (Wang, 1976) and 5% acetone for 3 h in a quartz cuvette (with a Photon Research Associates LN-1000 nitrogen laser with spectral output at 337 nm and a pulse rate of 7 Hz). The TD construct was size-purified on 12% denaturing polyacrylamide gels and tested for the presence of the cyclobutane [rather than the (6—4)] thymine dimer by direct photoreversal experiments at 254 nm only (the cyclobutane photoreverses at this wavelength). When irradiated thus, almost quantitative yield of the regenerated T1 and T2 strands was obtained.

2.3 Anthraquinone Coupling and HPLC Purification

Generation of anthraquinone-modified DNA sequences was accomplished by reacting the N-hydroxysuccinimide ester of anthraquinone-2-carboxylic acid with the 5'-C6-amino functionality (C6-NH_2) on the DNA (Fahlman & Sen, 2002). Prior to coupling, the DNA was treated to remove nitrogenous contaminants from the DNA synthesis. The dried DNA samples were first suspended in 100 μL of ddH_2O, and were extracted three times with 100 μL of chloroform. The DNA remaining in the aqueous phase was then precipitated by the addition of 30 μL of 1 M NaCl and 340 μL of 100% EtOH. Following mixing, the sample was chilled on dry ice for ~10 min, and then centrifuged in a microfuge for 20 min to pellet the DNA. The pellet was washed once with 150 μL of 70% aqueous ethanol (v/v). Following air-drying the pellet was dissolved in 100 μL of ddH_2O, the DNA concentration of the solution was determined in a standard fashion using UV absorbance measurements.

The AQ-NHS ester (4.8 mg) was dissolved in 238 μL of dimethylformamide. For each coupling reaction, 7 μL of this stock suspension was added to 75 μL of a 100 mM sodium borate solution (pH 8.5). To the resulting mixture was added 8-15 μL (5-10 nmol) of the purified amino-labeled DNA. The tubes containing the coupling mixtures were covered in aluminum foil and shaken overnight at room temperature. The DNA was then ethanol-precipitated by addition of 27 μL of 1 M NaCl and 280 μL of 100% ethanol (the solution was chilled in dry ice and the precipitated DNA collected and washed as described above). The large pellet obtained (containing a significant amount of the uncoupled anthraquinone) was now suspended in 50 μL of 100 mM aqueous triethylamine acetate (pH 6.9) to which was added 100 μL of chloroform. The uncoupled anthraquinone partititioned into the chloroform phase, and the aqueous phase was now extracted two more times with 100 μL of chloroform, prior to partial drying under vacuum to remove any residual chloroform. The DNA obtained was then purified by reverse phase chromatography on an HPLC using a C18 Bondpack column (Waters).

The HPLC protocol was as follows: the solvent flow was continuous at 1 mL/minute, and the column was heated to 65 C. The initial conditions were: 100% Solvent A (20:1 of 100 mM triethylamine acetate, pH 6.9:acetonitrile) changing to 30% Solvent B (100% acetonitrile), over 30 min and with a linear gradient. After this period, the solvent was rapidly changed to 100% Solvent B, for 15 min, before reconditioning the column to the starting conditions.

The concentrations of the various products of the coupling reactions could be determined spectroscopically. Absorbance values for the conjugate were made at 260 nm, using extinction coefficients for the individual bases obtained for single stranded DNA:

(260 nm, M-1 cm-1) adenine (A) = 15 000, guanine (G) = 12 300, cytosine (C) = 7400, thymine (T) = 6700, and anthraquinone (AQ) = 29 000 (Maniatis et al., 1989).

2.4 *In Vitro* Selection Series: ETS, SX, TD, SY, SYB

For reference and for future convenience, the protocols for the selection are listed in the following sections. The pre-experiment preparations includes Kinasing, Loading Buffer Preparations, and G, CT ladder creation. Detailed material and methods are included in the Appendix Section 5.1.2.

2.4.1 Kinase Reaction and Piperidine pre-treatment

Unmodified DNA sequences purified by PAGE before use. Sequences to be P^{32}-5' end-labeled were pretreated with 10% piperidine (90°C for 30 min followed by lyophilization) prior to 5'-P^{32}-kinasing and PAGE purification. The pretreatment of cleaved DNA molecules damaged during synthesis leads to lower background cleavages from photoirradiation experiments. The kinase reaction was done according to standard conditions using commercial T4 Kinase products and incubated for 30min at 37°C. Exact reaction conditions are listed in the Appendix Section 5.1.

2.4.2 PCR Amplification, Primer Extension

Master mix protocol: 10uL 10x Buffer (Custom: 500mM KCL, 100mM TrisCl(pH 8.3), 15mM MgCl2, 1% Triton X-100, 10mM β-Mercaptoethanol. Or QIAGen buffer), 4uL (25mM NTPs), 10uL (10uM Primer pETSGGO), 10uL (10uM Primer pETSCC), 1uL Taq Polymerase (QIAGen), 1uL Kinased pETSCC, 64uL ddH_2O. Total

Volume is 100uL. The PCR programming uses 95 °C -3 min (denaturation step), 50 °C – 2min (primer annealing step), 72 °C - 2min (extension).

2.4.3 Photoirradiation

Samples were irradiated using a UVP Black-Ray lamp (model UVL-56) with a 366nm wavelength maximum at 18 W (0.16AMPs). The distance between the bulb and the solution surface was 4.0cm. All irradiations were carried out at room temperature (~19-21°C). Temperature was maintained with a room temperature water bath. The lamp was warmed up 15 min ahead of use for consistency of light output.

2.4.4 Piperidine Treatment and 12% Denaturing Polyacrylamide Gel Electrophoresis

Following photoirradiation, samples were ethanol precipitated, lyophilized and then treated with 10% piperidine. The conditions were incubation at 90°C for 30 min, followed by lyophilization, but several experiments optimizing piperidine treatment used varying temperature and duration. The piperidine-treated DNA samples were then dissolved into denaturing loading buffer (as listed before), and loaded on to 12% 19:1 Bis:Acryl sequencing gel.

2.5 Phosphorimaging, Data Quantification and Analysis

Densitometry measurements were made using phosphorimaging plates exposed to the radioactive gels and read with the *Typhoon 9400 Variable Mode Imager*. The *Molecular Dynamics Image Quant 5.2* software was used for densitometry analysis. On a PAGE, each radioactive signal peak represented DNA containing the P^{32} label of a

certain mobility. DNA fragments were separated on a denaturing PAGE according to size. Image Quant was used analyze and integrate of the peaks of each gel lanes. The integration value of each peak was divided by the total integration values to get a percentage value called "percentage cleavage". This normalization method helps mitigate slight lane-to-lane radioactive count loading differences. The raw data is exported for further analysis as an Excel file. The method will be described in more detail in the Results and Discussion section.

2.6 Cloning and Sequencing of SY and SYB Series

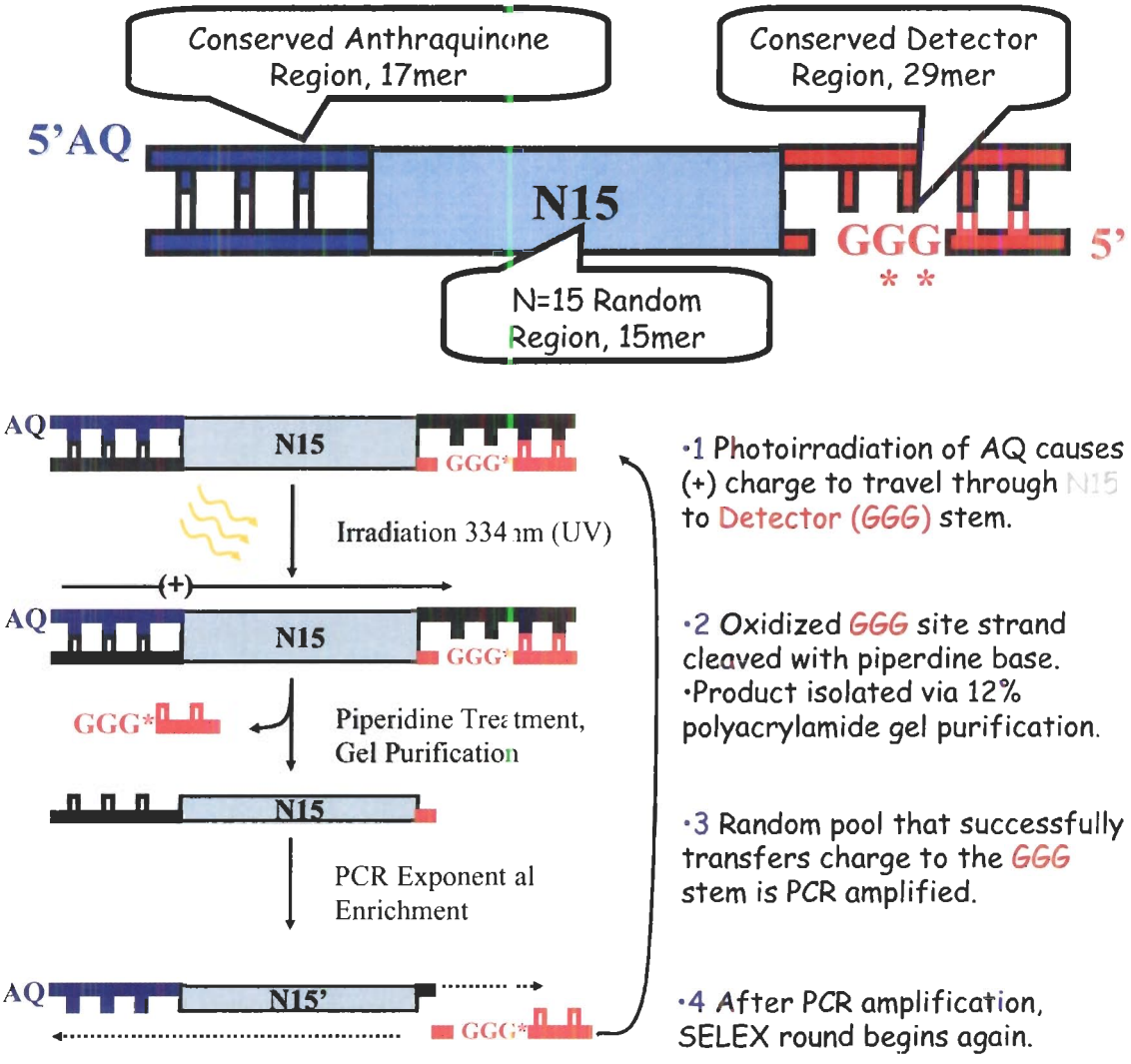
Cloning was accomplished using the Original TA Cloning Kit (Invitrogen), following standard methods as listed in the Invitrogen instruction manual. The insert to be sequenced was ligated to pCR2.1 plasmid and circularized. The plasmid was then inserted into One Shot ^{tm} E. Coli cells (INVαF') though heat shock. After the plating was carried out under permissive conditions, colonies with unique sequences inserted within the plasmids were white (as opposed to blue-green) and could be isolated and grown. The number of colonies collected varied from selection to selection. After growing each colony to a sufficient density of cells, the amplified plasmids were isolated using the Qiagen QIAprep Spin miniprep kit. The purified plasmids are sent to Macrogen Inc, a South Korean based commercial sequencing company, to be sequenced.

2.7 Experimental Error Discussion

The main experimental errors are from fluctuations in pipetting and weighing, and from data aquisition by the Typhoon 9400. Pipetting was done using the Gilson Pipetman P pipetor series. Specifically the P10, P20, P200, and P1000. These pipetors were calibrated according to Gilson specifications. The Sartorius CP2245S balance was used to measure dry compound weight to the fourth decimal place and had a linearity error of +/- 0.0002g. The BP2100S balance was used for cruder measurements such as Tris-Cl for various buffers. This has a linearity error of +/- 0.02g. The Typhoon 9400 variable mode imager specifications state that a uniformity error of +/- 5% occurs over the scan area.

3 *IN VITRO* SELECTION OF CHARGE TRANSPORT, RESULTS AND DISCUSSION

Figure 3.1: **Top: *In vitro* Selection Construct Schematic. Blue represents the conserved AQ stem, red represents the conserved detector stem, the light blue represents the random region (N15). Bottom: the selection schematic.**



As mentioned in the Chapter 1, the Multi-step Hole hopping mechanism and the Phonon-assisted Polaron-like Hopping mechanism are two major theories to explain long distance hole transport. Prior experiments detailing evidence to support the two mechanisms used rational design approaches to select sequences for experiments.

The purpose of our In vitro Selection of Charge Conducting Sequence project was to obtain sequences efficient at hole transport without using rational design methods, to see if the evidence gathered supported any of the major charge transport theories.

We aimed to study what constitutes a "good DNA sequence" for hole transport using an *in vitro* selection methodology precisely because it does not favour any theories in its design. This would mitigate some intentional or non-intentional bias inherent in rationally designing a construct to test it for conductivity. The biological and R&D importance of charge transport has been previously discussed in Section 1.3.

The design of the double stranded DNA selection construct had a few key features (Fig. 3.1 top). A stretch of random sequences was in the middle flanked by two conserved sequences. The length of the random region was 15 base pairs (N15) in the SY and SYB series. The conserved **AQ stem** had an Anthraquinone (AQ) covalently attatched to a 5'-C6-amino functionality on the 5' end. The conserved **Detector stem** contained a "detector element" to monitor charge transport that successfully passed to it through the random region. The sequences within the conserved (fixed-sequence) flanking regions

needed to be relatively efficient conductors, so that hole transport is able to proceed from the initial AQ hole injection site all the way to the detector.

Fig. 3.1 bottom depicted the *in vitro* selection scheme. (1) When irradiated at 334nm, AQ injected a positive charge (hole) into double stranded DNA. The hole migrated down the duplex via a long distance hole transport mechanism, through the N15 random region, until it reaches the detector, where it may oxidize guanines. (2) A subset of oxidized DNA lesions cause strand breakage when treated with hot piperidine. (3) The DNA molecules with the cleaved detector represent constructs successful at hole transport; they are isolated with polyacrylamide gel electrophoresis and PCR amplified to be used for the next selection round (4). *Note that the successful candidates are non-radioactive because the 5' P^{32} endlabel has been cleaved away.* Size markers were used to isolate where these successful candidates ran on the PAGE.

The design, testing, and implementation of the charge transport selection was carried out in three distinct phases. In the first phase, experiments were carried out to select the most appropriate detector for the selection (Section 3.1). Both 8-oxoG and thymine dimers were tested as an alternative to the conventional GGG detector. In the second phase, the conductivity of the flanking arms were tested in the CA series (Section 3.2). In the third and last phase the selection constructs were made, and selection began. The SX series, the first full scale selection, as such, was prone to many experimental challenges to be overcome. The experimental details are included in the Appendix (Section 5.4.3). When I began selection from the SY series (Section 3.3), the construct design questions had essentially been solved, however an unforeseen contamination of the selection pool with a positive control DNA duplex made it necessary to repeat the

slection. Full details of the SY series is in the Appendix (Section 5.4.4). The SYB series (Section 3.4) was a repetition of the SY series with careful safeguards in place to prevent contamination of the selection pool. After experiments quantifying conduction efficiency differences between initial round and round 7, (Section 3.4.2) it became apparent that the SYB series was successful in acquiring good charge conductors.

3.1 Selecting the Detector Stem: 8-oxoG, Thymine Dimer or Triple Guanine Motif

As mentioned above, the initial random library holds more than $1.07x10^9$ unique 15 base pair sequences. The purpose of the detector stem was ultimately to tag candidates with good hole transport sequences. Good candidates were defined as sequences able to allow hole transport down the AQ stem, through the random region, until it is detected at the detector stem. As such, the detector stem was designed to cause DNA strand cleavage that can be visualized by PAGE.

Double and Triple guanine detectors (5'GG and 5'GGG) are commonly (Sanii et al., 2000, Fahlman, et al 2002, Sankar et al. 2004) used for hole transport purposes. As alluded to in Section 1.5.3, the most 5' guanines in double and triple guanine motifs are highly prone to oxidation via hole transport. Upon oxididation, they become prone to piperidine-mediated strand cleavage, and are thus useful for hole transport detection purposes. However, owing to the fact that two oxidations are needed for guanine to become properly labile to piperidine-mediated cleavage (Burrows and Muller, 1998), and the fact that only a fraction of guanine radical cations actually react with water, only a small population of damaged guanines can be visualized.

We sought to make use of two new hole transport detectors to alleviate this problem: 8-oxoG (Appendix Section 5.4.1) and Thymine Dimers (Appenidx Section 5.4.2). As a major product of guanine oxidation, 8-oxoG is one step closer to being fully piperidine labile, which makes it a good potential detector stem candidate. However, using 8-oxoG as a detector, we found signal to noise ratio to be low (~2:1) compared

with conventional guanine doublets (~4:1) because even non-irradiated samples had very high piperidine cleavage rates.

Figure 3.2: 8-OxoG products, the piperidine labile 5-OH-8-oxo and guanidinohydantoin

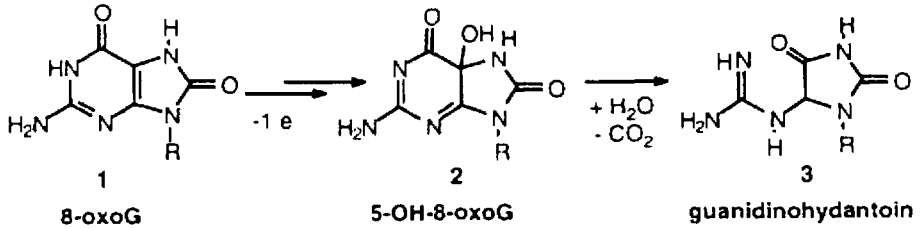
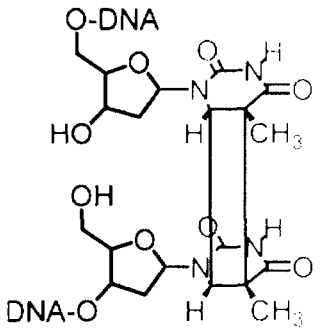


Figure 3.3: Thymine Dimer as detector stem candidate.



Two DNA strands linked via a thymine dimer (with no sugar-phosphate backbone linkage) was another detector stem candidate. While excess electron transport has been known to photoreverse thymine dimers (Behrens et al. 2003), there has been more controversy regarding hole transport's ability to photo-reverse thymine dimers (Dandliker et al. 1998, Ramaiah D et al. 1998). The thymine dimer detector stem was designed to result in strand cleavage if hole transport photo-reversed thymine dimers. Unfortunately, we did not detect cleavage of thymine dimers attributed to hole transport.

In the end, the triple guanine detector (5'GGG) was used for the selection.

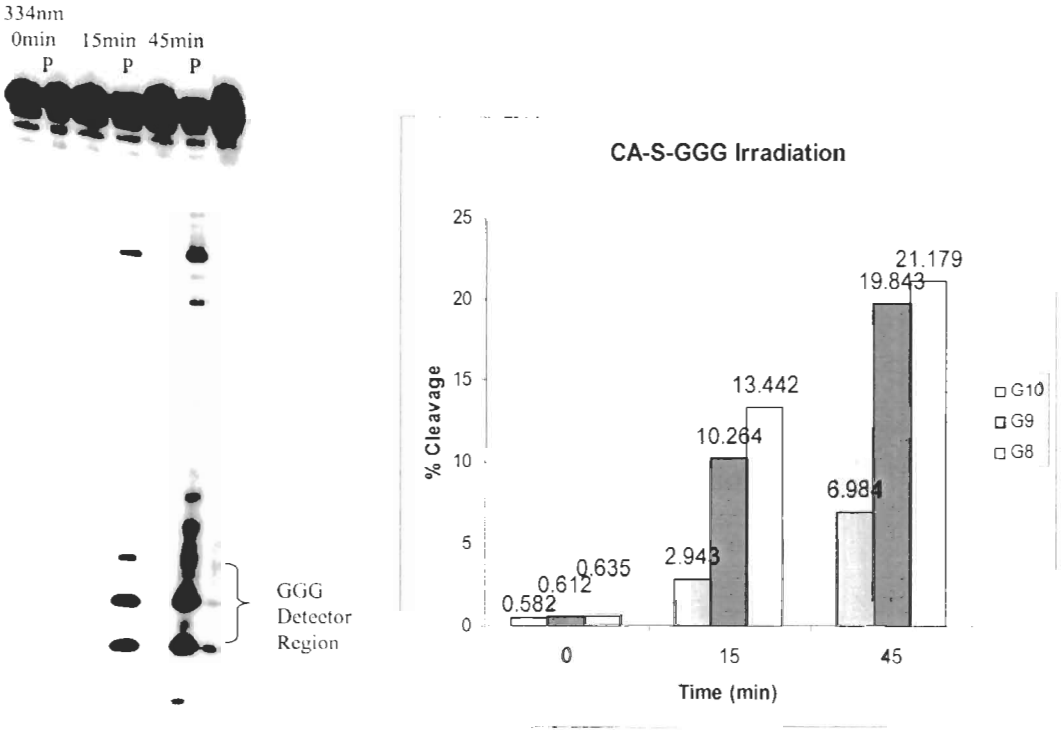
3.2 Constructing Conductive Flanking Arms, CA Validation Construct

Having chosen triple guanines as charge transport monitoring tools, the next step was to test the hole conduction efficiency of the sequences used for the constant flanking regions. The CA sequence (Fig. 3.4 top) was chosen because it was shown to be effective for detecting charge transport in the past (Sankar et al., 2004). Shown in the sequence in Fig. 3.4 top, with the triple guanine detector (GGG) and a P^{32} 5' end-label, this short sequence was used to validate the constant flanking regions for the use in actual selection constructs. An irradiation time-course experiment was set up where CA sequence samples were exposed to differing doses of 334nm UV irradiation. Following piperidine treatment, the lyophilized DNA samples were loaded on to a denaturing PAGE and the results are shown on Fig. 3.4, left. Dark controls and non-piperidine treated controls were included as reference. Denaturing PAGE separates single stranded DNA fragments by size. Each radioactive fragment (except for the full-length strand at the top of the gel) represented piperidine-mediated cleavage resulting from hole transport oxidative damage. Densitometry procedures were described in Section 2.5. Densitometry data graphed on Fig. 3.4 right shows that after 45min irradiation, G9 and G8 in the GGG detector (two 5'most guanines) showed significant damage (20% and 21% respectively). These results validated the constant sequences chosen for use in the SX selection series.

Figure 3.4: **Top shows the sequence of the CA test construct. After irradiation time-course (0min, 15min, and 45min at 334nm, room temperature) and the subsequent piperidine-mediated cleavage, samples were loaded and run on a 12% denaturing PAGE shown on the left. The GGG region is shown at the bottom of the gel. Right is the densitometry data of the G10, G9, G8 expressed as % cleavage.**

CA Series

```
AQ~TTTAG CTCAC GAGAC GCTCC CATAG TGA 3′
3′ AAATC GAGTG CTCTG CGAGG GTATC ACT 5′′
```



3.3 SY Series, Lessons Learnt

The SY Series was designed in part from the lessons learnt from the SX series (Section 5.4.3). The distance from the charge injection site to the detector was 41 bases in the SX series. Because of the significant cleavage at the detector even at that distance, the SY series was made longer; 50 bases in length from charge injection site to detector. This enabled the constant flanking regions (primer annealing sites) to be extended to 17bp, the preferred PCR primer size. The random region was expanded to 15bp (from 10bp in the SX series) to gather more information. Zuker's secondary structure predictor "mfold" (Zuker, 1999) was used to help avoid constant region single-strand sequences that were prone to forming self-secondary structures.

Figure 3.5: Top: SY Positive control duplex. Bottom: selection construct duplex

SYK (positive control, 59BP)

```
   1          10          20          30          40          50          59
AQ~TTTAG CTGAC TGACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA 3'
3' AAATC GACTG ACTGA GCTCA CACGT CTGAG CAGCT CACGC TCTGC GACTG GGTAT CACT 5*'
  59          50          40          30          20          10 98       1
```

SY and SYB Series (Selection Construct, 59BP)

```
   1          10          20          30          40          50          59
AQ~TTTAG CTGAC TGACT CGNNN NNNNN NNNNN NNCGA GTGCG AGACG CTGAC CCATA GTGA 3'
3' AAATC GACTG ACTGA GCXXX XXXXX XXXXX XXGCT CACGC TCTGC GACTG GGTAT CACT 5*'
  59          50          40          30          20          10 98       1
```

The SY selection construct is shown above in Fig 3.5 top. The bolded areas are the two primers to be used to amplify selection sequences after each round. One primer contained anthraquinone attached by a C6 linker to its 5' end, the other contained the underlined G10, G9, G8 triple guanine detector. The boxed region represented the location of the random bases. The AAA bridge used for effective charge injection by AQ

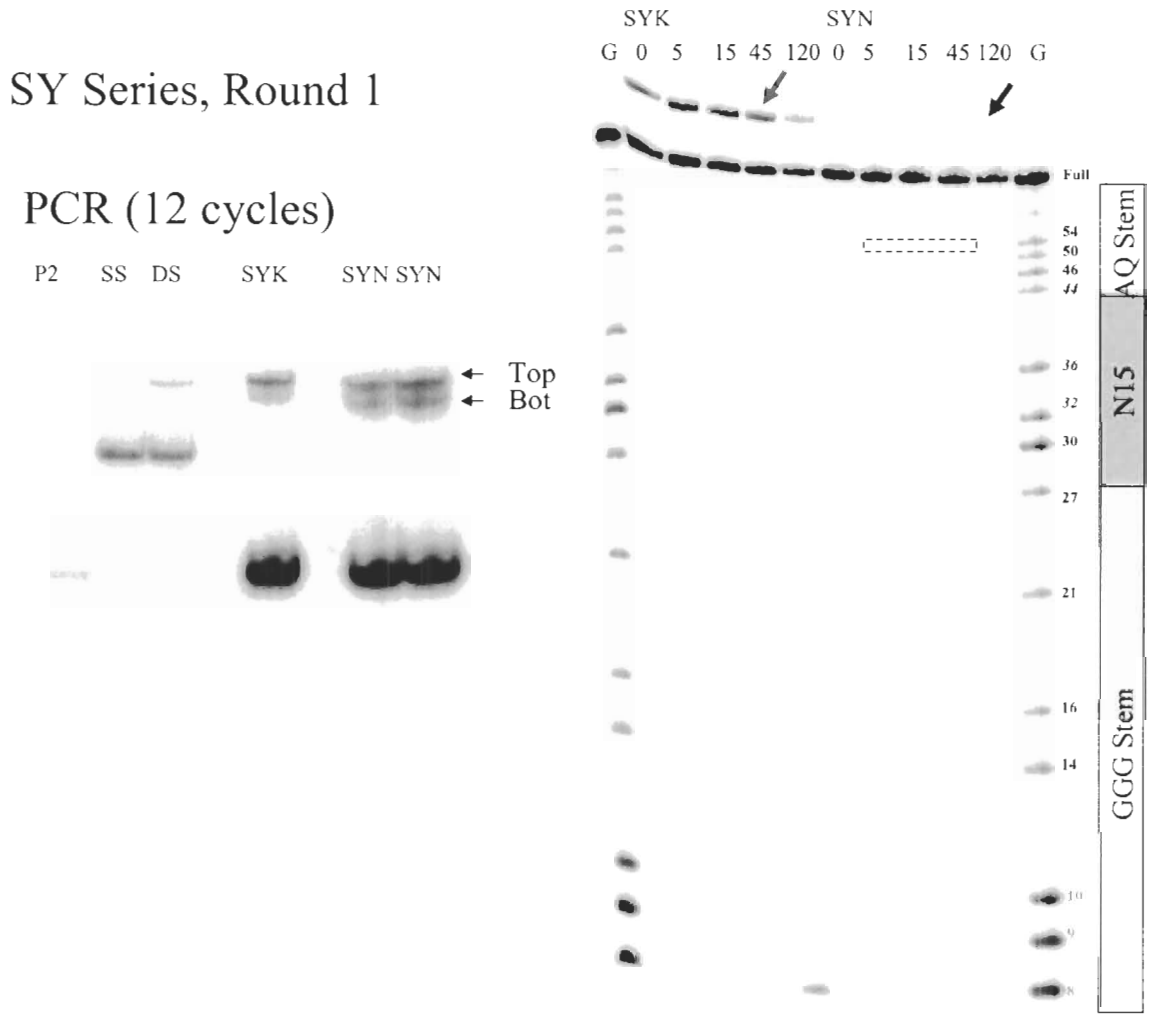
is highlighted in grey (Sanii & Schuster, 2000) The SYK construct was a fixed sequence positive control used for the SY series.

The number of possible unique sequence in this selection of a N15 random region was relatively small (4^{15} =1.07x10^9 combinations in all). Even a suboptimal PCR amplification would yield 60-100 copy numbers of the each unique sequence. The details of the round 1 selection is outlined below.

3.3.1 SY Selection Round 1

Features of the SY round 1 selection are shown in Fig. 3.6. Selection constructs from the initial pool were PCR amplified 12 cycles and run on a native 12% PAGE (Fig. 3.6, left) for purification purposes. Trial PCR runs (not shown) suggested 10-12 cycles to be the most appropriate. Size referencing controls were run alongside the PCR products. P2 represented the size of the 5' P^{32} radiolabeled primer. SS control represented the size of a single stranded product. DS represented the size of the double stranded product. SYK lane used the positive control as template. SYN lane used the initial selection pool as template. There were two PCR products found on the native gel (called top and bottom products). Top product was consistent with the double stranded control and is considered to be the product to be used for selection purposes. The faster product (called Bottom product) was found by DEPC probing (Appendix Section 5.4.4.3) experiments (Maniatis et al., 1989) to be mostly single stranded. Most likely it resulted from secondary structure formation of PCR products. Since the formation of this product only interfered with the overall yield of the double-stranded hole conduction substrates, but not with the selection itself, the presence of faster product was eliminated during gel purification.

Figure 3.6: **SY *in vitro* selection Round 1 Results: Left is a 12% native PAGE used to purify PCR products. Lanes include P2 primer, single stranded size control, double stranded size control, SYK, and the SYN selection sequence. Top band was the double stranded product to be cut and eluted. After the selection construct was purified by native PAGE, irradiated, piperidine treated to induce strand cleavage, the lyophilized results were loaded and ran on a 12% denaturing PAGE and shown on the right. SYK is positive control construct, and SYN is selection sequence. Numbers 0 to 120 represents irradiation time in minutes. Flanking the constructs are chemically produced G ladders (GGG detector is labeled G8, G9, G10). Dotted box represents the size of the successfully cleaved candidate sequence to be cut, eluted, and amplified for subsequent selection rounds.**



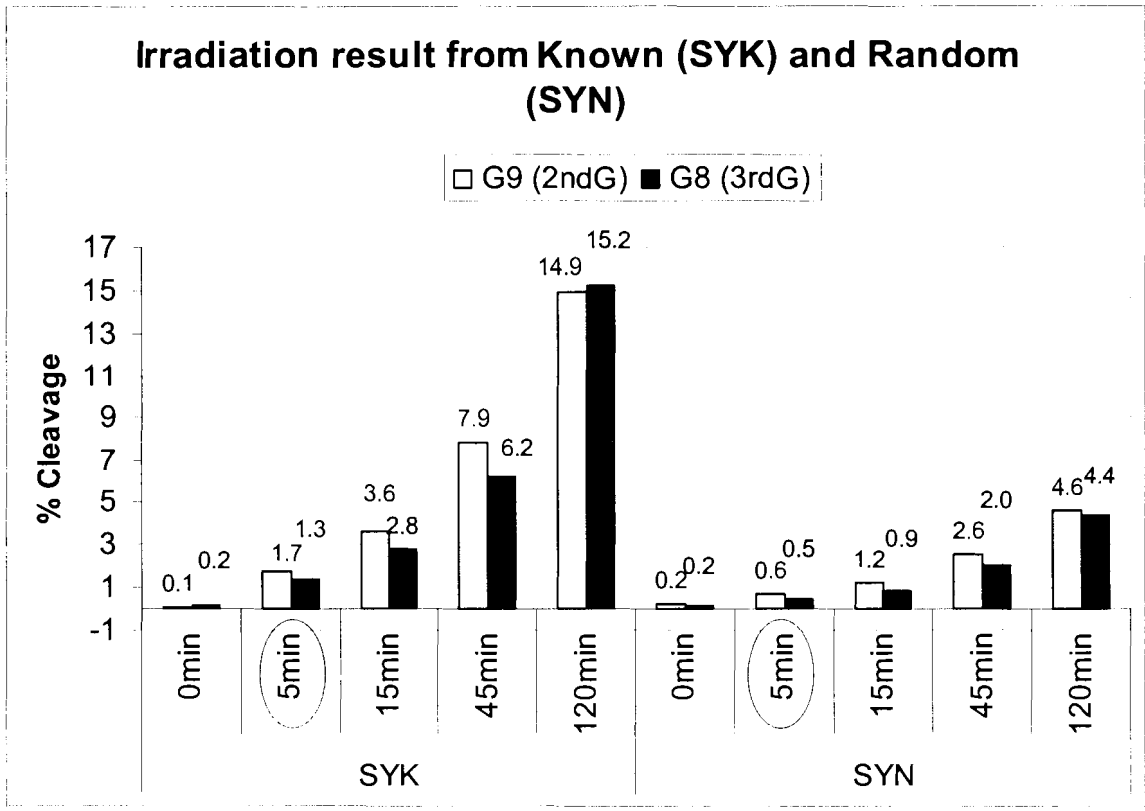
After the PCR amplified constructs were purified, an irradiation time-course was carried out at 334nm, and the irradiated DNA was piperidine treated to induce strand cleavage at oxidized DNA sites. The results are shown on 12% denaturing polyacrylamide gel in Fig. 3.6, right. Densitometry data at the G8, G9 detector are tabulated below, in Fig. 3.7. Two G-ladders (chemical cleavage at G sites using DMS) were included in Fig. 3.6 right to identify oxidatively damaged guanines. Guanines 10, 9, 8 corresponds to the triple G detector.

As expected, densitometry data showed dose-dependent (irradiation time), hole transport-mediated cleavage patterns. The SYK positive control was damaged considerably more compared to the SYN selection pool at equivalent irradiation times, which was expected because SYK was designed to be relatively conductive. The SYN selection pool contained a large sequence variety in the N15 region, not all of which would be expected to be good conductors. The dotted box region showed the area of the gel where the non-radioactive band of the successfully cleaved selection candidate were cut out and eluted for amplification. The location of these successful candidates (50-52 nucleotide long) was inferred using the G-ladder as a size marker. 5 min irradiation time was chosen for subsequent selection rounds because even at 5 min, a significant signal was seen in GGG detector.

One curious anomaly observed in these denaturing gels was the presence of an unknown band larger than full length in the SYK construct shown by an arrow (this was termed "Slow SYK band" for future reference). SYN in round 1 did not exhibit this band. Further denaturation experiments found this "slow" band to be double stranded constructs not properly denatured in the denaturing PAGE (Appendix Section 5.4.4.4).

Subsequently, this problem was remediated by a thorough boiling of samples prior to loading on the denaturing gel.

Figure 3.7: Densitometry data of detector G9, G8 for Round 1 expressed as % cleavage. Note SYK was damaged more than SYN results of equivalent irradiation dose.



3.3.2 SY Selection Round 7 Sequencing and Cloning.

By round 7, the *in vitro* selection did not appear to be selecting better conductors compared with the positive control and initial round controls (Appendix Section 5.4.4.5), so a decision was made to clone and sequence the DNA population of rounds 5 and 7. Thirty three clones were sequenced from round 7 using procedures outlined in the Materials and Methods (Appendix Section 5.1), and the results are shown in Fig. 3.8.

Figure 3.8: **Round 7 Sequencing Results, sample of population. At the top the SYK construct sequence. Shown at the bottom is 33 clones sequenced from round 7 pool.**

SYK (Known sequence, positive control 59BI')

```
   1          10          20          30          40          50          59
AQ~TTTAG CTgAC tgact cgAGT GTGCA GACTC GTcga gtgcG AGACG CtgaC CCATA GTGA 3'
3' AAATC GAcTG actga gctca cacgt ctgag cagct cacgC TCTGC GactG GGTAT CACT 5*'
  59          50          40          30          20          10 98       1
```

Round 7 Sequencing Results

```
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGCAG CGTCA CTGCG GGCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGACG ATCAT ACCCA ATCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGGTG ATGTT AGCGA CTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGGCG ACGAA ACAAT ATCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGCAA TGTGA AGCGC CGCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGGCC ATCAG CGCCT CCCGA GTGCG AGACG CTAAG CCGAA TTCC
TTTAG CTGAC TCACT CGTCC CGATG TCTGA ATCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGTTA GAGGG GGAGG AGCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGGTA TGTCT TAAGG GGCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGGAC CGGAT TTAGG GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGTAA ACTTA CTCCG CACGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGACC CGGAT CAAAC AGCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGGAA TAGAC AACTG CACGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGCAA ACTAA ATTTG CCCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGACG ACGCC TGGTA CCCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGCAG CTTGT GATAG ACCGA GTGCG AGACG CTGAC CCATA GTGA
```

An unfortunate consequences of using a positive control side-by-side with the selection pool was the contamination, by the positive control, of the selection pool. The round 7 cloned sequences showed the presence of the positive control sequence, amounting to half of the population (Fig. 3.8 underlined sequences). Even by round 5 (clones listed in Appendix Section 5.2) a large proportion of the pool had been contaminated by positive control. The parasitic-like invasion of the positive control sequences were probably due to cross-contamination from various steps of the selection. Specifically, running positive control and selection lanes side by side on native and denaturing PAGE may allow cross-diffusion of the positive control into the selection lanes. Since the positive control was also a relatively good conductor, the selection itself failed to filter it out. *In vitro* selection works on the premise of magnifying small competitive advantages, we felt that the cross-contamination by SYK might have compromised the selection pool. Although there were still sequences in the selection pool that could be used for analysis, it was decided to start a new selection without the positive control. The experimental protocols worked out with the SY series and the experience with that particular construct were essential for the next selection, the SYB series.

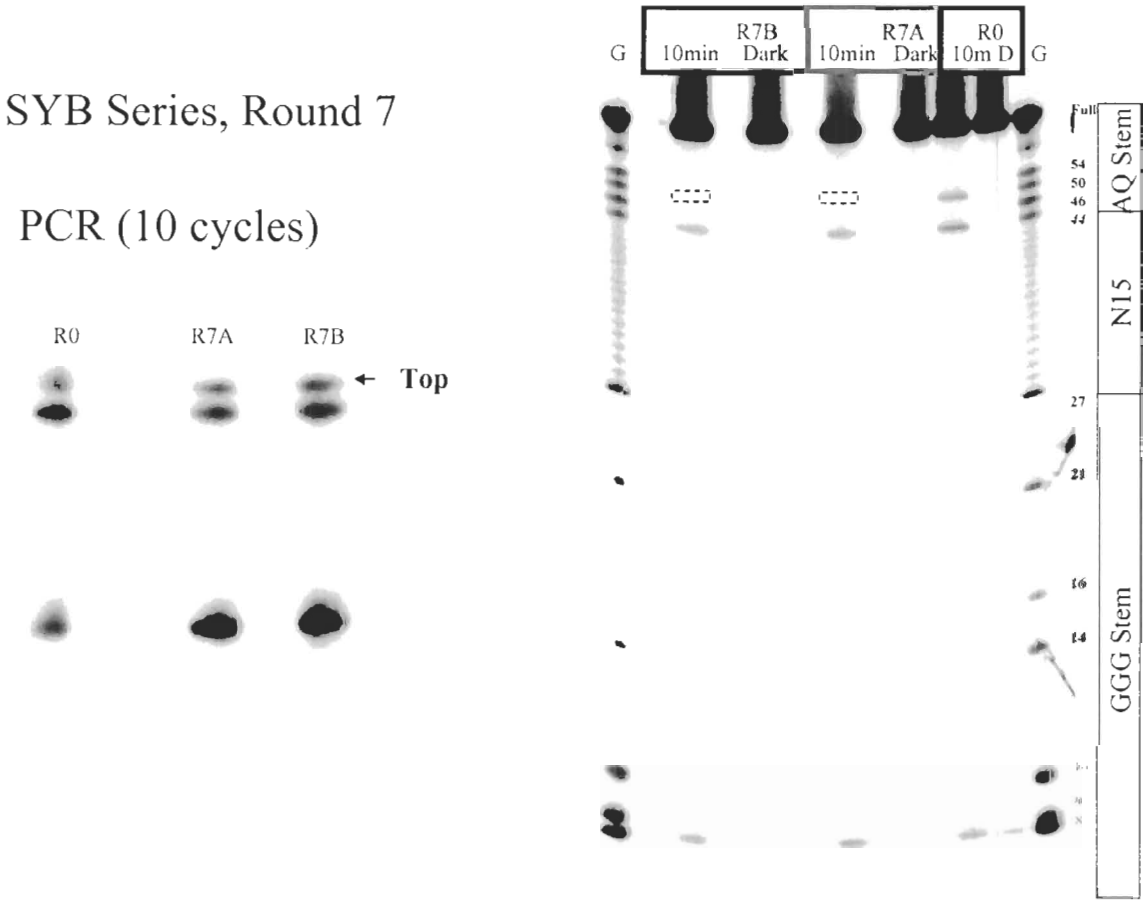
3.4 SYB Series

The SYB selection was in essence a repeat selection of the SY library. Since a goal of this selection was speed, the weeklong SY procedures were condensed carefully into two day experiments. Each week consisted of one day of pre-selection preparation (kinasing experiments for example) followed by two rounds of selection. The exact experimental procedures are listed in the Appendix, Section 5.1. New oligonucleotide was ordered, processed, AQ were coupled to the primer and the conjugate HPLC purified; then 7 rounds of selection was carried out. Initial round and round 7 pools were sequenced and cloned. All within a span of two months.

Besides expediency, several experimental procedure differences distinguished the SYB selection from the SY selection. As described above, the positive control construct SYK was not used for the experiment. Two identical, but independent, parallel selections were carried out, labeled the 'A' and 'B' pools to guard against the contamination of one. The irradiation conditions were 30mins for the first two rounds, followed by 10min irradiation for the subsequent five rounds. Extensive care was taken to ensure that there was no trace of SYK in this selection. Fig 3.9 shows results of round 7 in more detail.

Fig. 3.9 left showed the native PAGE used to purify the PCR results. Owing to past experience with the SY series, we knew that the selection construct we wanted to purify was the top band. The initial selection pool was used for generation of the G-ladder and for conduction comparisons with the current round. After irradiation, piperidine treatment and denaturing PAGE (Fig. 3.9 right), the successful candidate sequences were cut out of the gel and eluted for subsequent rounds. The densitometry data from the G8 and G9 detector for rounds 1 to 7 are tabulated in Fig. 3.10.

Figure 3.9: **SYB *in vitro* selection Round 7 Results: Left is a 12% native PAGE used to purify the PCR products. Lanes include R0 selection, Round 7 A and B pool. Top band was the double stranded product to be cut and eluted. After the selection construct was purified by native PAGE, irradiated, piperidine treated to induce strand cleavage, the lyophilized results were loaded and ran on a 12% denaturing PAGE and shown on the right. Lanes listed from right to left is: R0 dark, R0 10min, R7A pool dark, R7A 10min, R7B pool dark, R7B 10min, and chemically produced G-ladders (GGG detector is labeled G8, G9, G10). Dotted box represents the size of the successfully cleaved candidate sequence to be cut, eluted and amplified for subsequent selection rounds.**



3.4.1 SYB Round 1-7 Summary

Figure 3.10: SYB Series Round 1 -7 Densitometry data expressed as % cleavage at G9-G8 site of GGG detector. Bar graph is listed in the same order as legend *(i.e. first bar from the left represents Dark G8 % cleavage for round 1, second bar from the left represents Dark G9 % cleavage for round 1 and so on)*. Black bars represents dark reaction. White bars represents the SYB-A pool, Grey bars represents the SYB-B pool. Round 1 and 2 had longer irradiation times, so its values are normalized to give proper comparisons to later rounds.

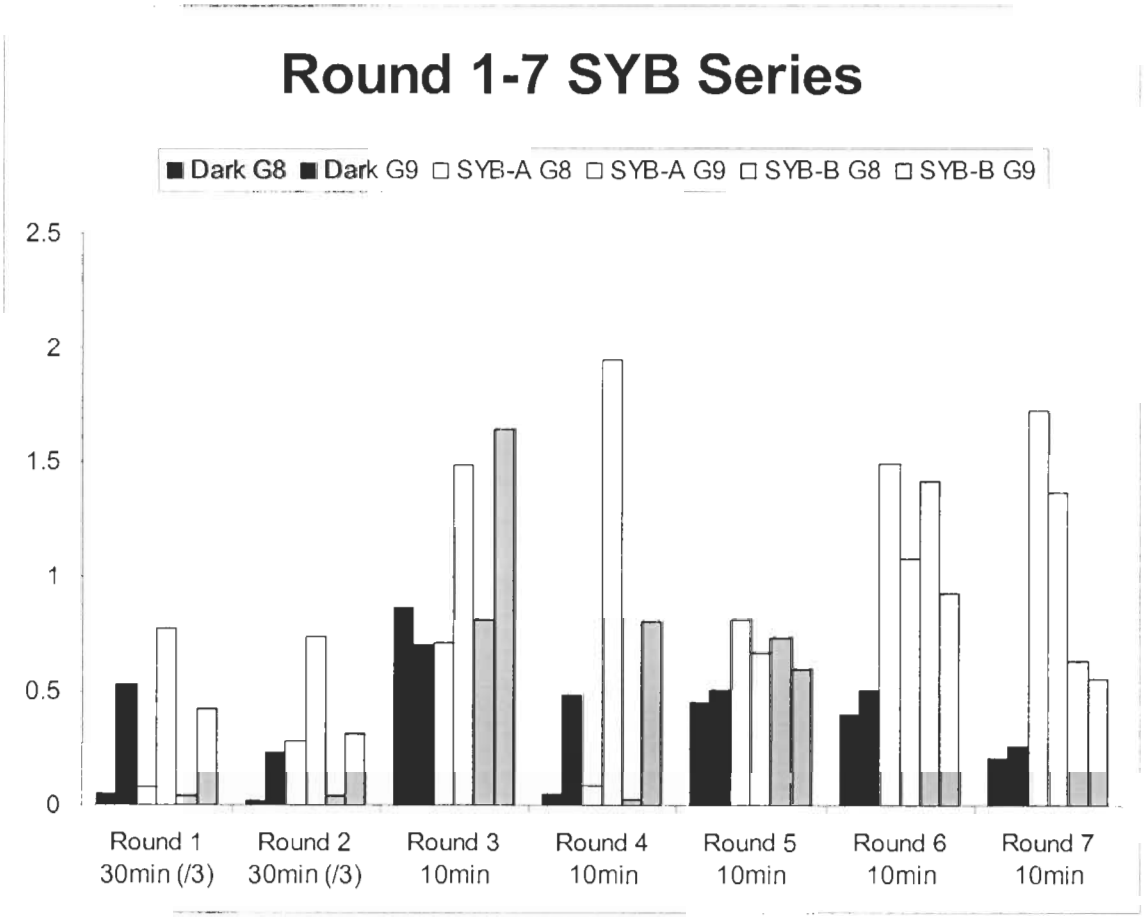


Fig. 3.10 showed the densitometry results of damage at the G8, G9 site during the course of the selection (rounds 1 through 7). The black bars represented dark controls, the white bars represented the A-pool, the grey bars represented the B-pool. By round 7, improvements in conduction in both A and B pools compared with initial round can be seen. A-pool exhibited a three-fold a higher signal than initial round, while in B-pool, the improvement seemed more modest.

3.4.2 Quantifying Selection Success, SYB Irradiation Time-course comparing R0 and R7

Before analysis of the round 7 selection pool, it was important to ascertain whether a real selection for conductive sequences had indeed been achieved. Fig 3.11 bottom left showed a denaturing PAGE result of a time-course that was carried out to compare the performance of the initial round selection pool (SYB-R0) with the round 7 selection pool (SYB-R7). Independent triplicates of each irradiation time-point (0min, 10min, 30min, and 90min) were carried out in the strictest of light conditions to minimize any background. Since ambient light exposure was one of the main reasons contributing to round-to-round background cleavage fluctuations, it was reduced by keeping samples in dark whenever possible, and by working without sunlight under reduced laboratory lighting. These steps resulted in clean bands and very low background. Densitometry results (Fig. 3.11 bottom, right) on the denaturing gel (Fig. 3.11 bottom, left) showed that on average, % cleavage at the detector site in round 7 was 3-fold higher than in round 0. Standard deviation error bars were used to quantitate data variation.

To determine whether the differences seen between round 0 (R0) and round 7 (R7) were statistically significant, the two-tailed Student's T-test was used on the G8

triplicate densitometry data comparing 90min irradiation (Fig. 3.12). The stricter two-tailed t-test was used because we were interested in genuine differences of either R0 having higher signal or R7 having higher signal. The probability that the difference between R0 and R7 was due to chance was less than 5% ($p=0.049$); which showed beyond a reasonable doubt that the increase in conduction seen in R7 was real.

Having established that a significant improvement in charge conductivity of the R7 pool existed over the initial pool, the next step was to sequence the two pools and identify molecules responsible for this improvement.

Figure 3.11: SYB Series comparison between Round 0 and Round 7. Shown at the top is the SYB construct (same as SY). After construct was purified by native PAGE, irradiated in a time-course, piperidine treated to induce strand cleavage, the lyophilized results were loaded and run on a 12% denaturing PAGE shown bottom left (one example from a triplicate of such gels, each running a repeat of the experiment). Bottom Right are histograms summarizing the SYB Series comparison between Round 0 and Round 7. Densitomery data of G9 –G8 are expressed as % of total signal (% Cleavage).

```
  1          10          20          30          40          50          59
AQ~TTTAG CTGAC TGACT CGNNN NNNNN NNNNN NNCGA GTGCG AGACG CTGAC CCATA GTGA 3'
3' AAATC GACTG ACTGA GCXXX XXXXX XXXXX XXGCT CACGC TCTGC GACTG GGTAT CACT 5*'
  59          50          40    N15   30          20          10 9 8      1
```

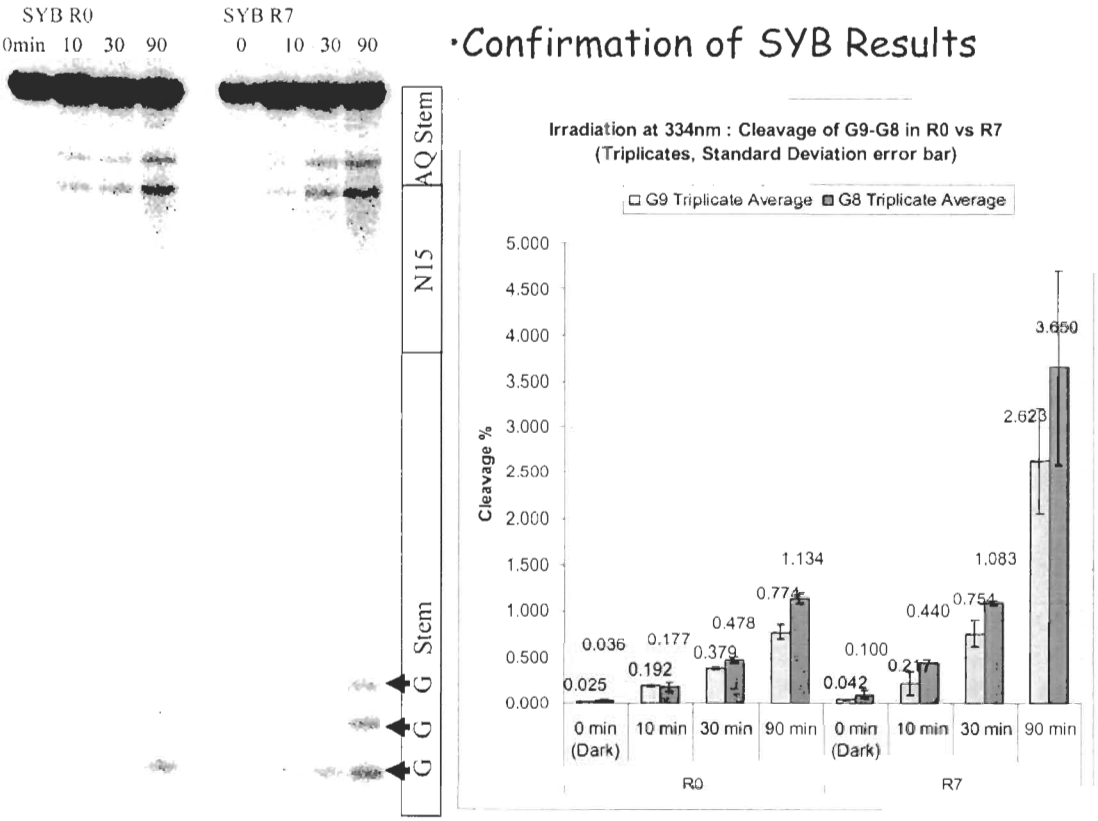


Figure 3.12: SYB Series comparison between Round 0 and Round 7. Densitomery data of G9 –G8 and two tailed Student's T-test.

	G9 Triplicates					G8 Triplicates					
	G9a	G9b	G9c	**G9 Triplicate**	*STDEV*	G8a	G8b	G8c	**G8 Triplicate**	*STDEV*	StdError
0 mir	0.022	0.026	0.026	**0.025**	*0.0023*	0.048	0.030	0.03	**0.036**	*0.0104*	0.006
10 m	0.209	0.183	0.183	**0.192**	*0.0150*	0.233	0.149	0.149	**0.177**	*0.0485*	0.028
30 m	0.361	0.388	0.388	**0.379**	*0.0156*	0.437	0.499	0.499	**0.478**	*0.0358*	0.020667
90 m	0.866	0.728	0.728	**0.774**	*0.0797*	1.063	1.170	1.17	**1.134**	*0.0618*	0.035667
0 mir	0.049	0.038	0.038	**0.042**	*0.0064*	0.051	0.124	0.124	**0.100**	*0.0421*	0.024333
10 m	0.361	0.145	0.145	**0.217**	*0.1247*	0.437	0.442	0.442	**0.440**	*0.0029*	0.001667
30 m	0.914	0.674	0.674	**0.754**	*0.1386*	1.110	1.069	1.069	**1.083**	*0.0237*	0.013667
90 m	1.958	2.955	2.955	**2.623**	*0.5756*	2.427	4.262	4.262	**3.650**	*1.0594*	0.611667
						2-tail t-test = 0.049					

3.4.3 SYB Selection Round 7 Sequencing and Cloning.

Owing to the specific primer chosen for sequencing, the N-strand (as depicted in Fig. 3.13 bolded) was sequenced. Note that the original single stranded selection library synthesized at the University of Calgary was the complementary strand (called the X-strand).

43 clones were picked from SYB-7A, 50 from SYB-7B, and 89 from SYB-0 (initial pool) and sequenced. The high number of clones was required because there may not be one predominant species of DNA good at hole conduction, but a motif, or a pattern. As expected, the N15 region of the initial pool appeared to be relatively random in sequence (see Appendix Section 5.3). Round 7 clones also exhibited no predominant sequence at a first glance.

Figure 3.13: Sequencing nomenclature. The 5'->3' N-strand (bolded) was sequenced. The 3'->5' was the complementary X-strand (not bolded).

```
   1         10         20         30         40         50         59
5' TTTAG CTGAC TGACT CGNNN NNNNN NNNNN NNCGA GTGCG AGACG CTGAC CCATA GTGA 3'
3' AAATC GACTG ACTGA GCXXX XXXXX XXXXX XXGCT CACGC TCTGC GACTG GGTAT CACT 5''
  59         50         40   N15   30         20         10 9 8       1
```

3.5 SYB Series, Qualitative and Statistical Analysis of Sequences from SYB Clones

The purpose of the statistical analysis was to search for data to evaluate the two major charge transport theories, Multi-step Hole Hopping mechanism (MSH) and Phonon Assisted Polaron-like Hopping mechanism (PPH). Base composition and sequence alignment exercises were used to search for obvious sequence changes or patterns.

Because MSH relies on quantum tunneling of electrons from guanine to guanine, distance between neighboring guanines is of paramount importance. MSH predicts that as long as guanines are spaced one or two bases apart, consistently, the identity of the bridging bases are not excessively important (Illustrated hypothetically in Fig. 3.14, top). The placement of guanines on either strand of the duplex DNA was thought to be sufficient to promote g-hopping. We postulated that guanine-step counting might be a useful statistical indicator for supporting MSH. By contrast, PPH predicts that both sequence identity and purine segregation are of great importance for hole transport (Liu et al., 2002) (Illustrated hypothetically in Fig. 3.14 bottom). In this context "purine segregation" is an idea that continuous stretches of purines segregated onto either strand of a duplex DNA contribute more favorably to hole transport than do mixed purine/pyrimidine strand (Liu et al, 2002). Purine stretch counts and neighbour studies (counting occurrence of 2-3nt combinations) may provide statistical evidence for PPH.

Because there were no obvious N15 sequence that accounted for the increase in the measured signal in R7 compared with the initial pool R0, the next logical step for us was statistical analysis of the N15 sequence regions to try and elucidate subtle differences.

Figure 3.14: A hypothetical depiction of MSH vs PPH. Top: MSH predicts regularly spaced guanines punctuated by bridging sequences. Bottom: PPH predicts specific sequences and purine stretches to be good for hole conduction.

MSH

5′ N**G**NN**G**NCNN**G**NC**GG**N**G**N**G** 3′

3′ XCXXCX**G**XXCX**G**CCXCXC 5*′

PPH

5′ T**GAAG**TCTT**GA**C**GGAGAG** 3′

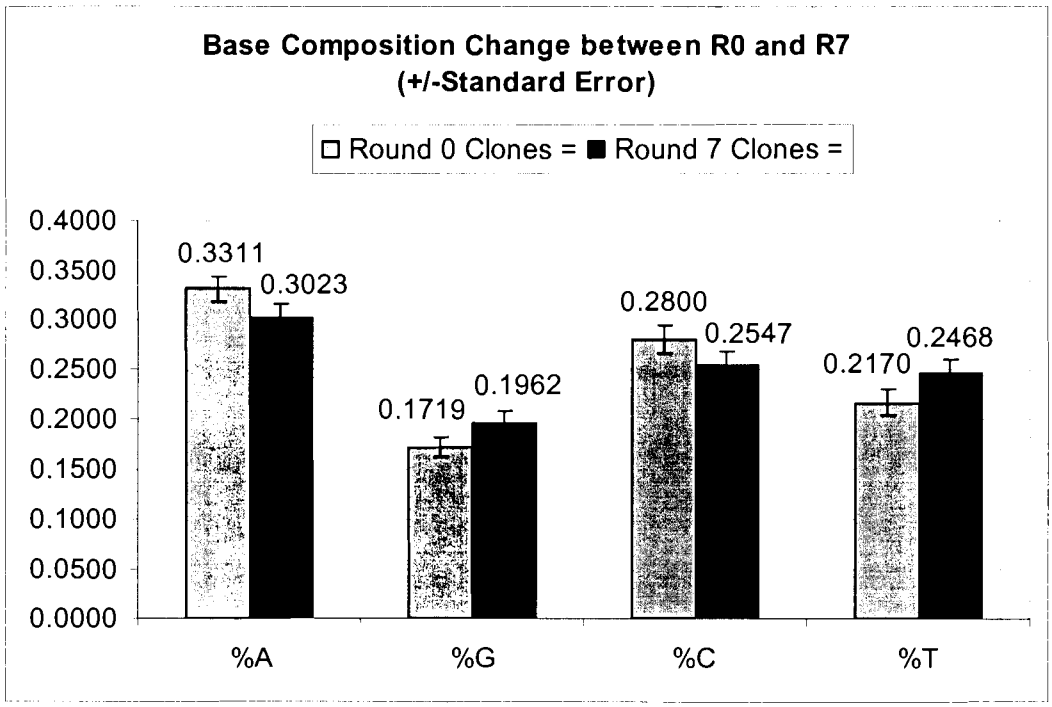
3′ ACTTC**AGAA**CTGCCTCTC 5*′

3.5.1 Base Composition Analysis

Base composition analysis was done to see if there were any significant differences between the N15 regions of R0 and R7. The observed probabilities of each of the four bases occurring over N15 were calculated for each of the 89 R0 and 93 R7 clones. The population of these probabilities were averaged and standard errors calculated (shown in Fig. 3.15). Analysis of round 0 clones showed that during the original DNA synthesis, the four bases were not incorporated in perfectly equimolar fashion; in fact the percentages of base occurrence on the N-strand were as follows: A%=33.1$\pm$1.3, G%=17.1$\pm$1.1, C%=28.0$\pm$1.4, T%=21.7$\pm$1.3.. Keeping in mind that while the N-strand was sequenced, the original synthesized strand was the complementary X-strand, base composition showed that original synthesis had significant T and G biases. The phosphoramidites used for automated DNASynthesis were probably not precisely calibrated to achieve 25% incorporation during actual DNA synthesis. Analysis of round

7 clones however, showed A%= 30.2±1.4, G%=19.6±1.1, C%=25.5±1.3, T%=24.68±1.3. Base composition values were confirmed with GeneDoc 2.6.002 Multiple Sequence Alignment Editor & Shading Utility. As the selection progressed from R0->R7, A and C composition decreased significantly (>95%CL) and G, T increased significantly (>95%CL). This meant that in the complementary X-strand, A and C abundance increased while those of G and T decreased.

Figure 3.15: Base Composition % on the N strand.



3.5.2 Sequence Alignment

Having made an analysis of the initial and selected base composition, the next logical step was to carry out sequence alignments to find obvious patterns. The cloned sequences were converted to a FASTA format, and simple base identity alignments were

carried out using ClustalX1.83 and viewed using GeneDoc 2.6.002. The complete results of his analysis are presented in the Appendix Section 5.3.4. Neither R0 or R7 alignments show any obvious pattern nor are enriched with any predominant sequence. A number sequences were found to be duplicated in R7 however, and these are listed in Fig. 3.16. There were no sequences found to be duplicated in R0.

Figure 3.16: Duplicated N15 sequences identified in R7. Left column lists the names of the clones, right column list the N15 sequence.

7A-2, 7A-18	TAGGGAGATAGTAAA
7A-12, 7A-20	GCGAGGTCAAAACGT
7A-8, 7A-37	GCACAACCATCACGG
7A-34, 7A-19	TGACGAGATTAGGAC
7A-15, 7A-4	CACAGGTTGGTATGA
7A-30, 7A-9	GATGTACGTACTTCA
7B-34, 7B-26	TTCTTTTTAGCCGTA
7A-36, 7A-7	CTGATCATCTGCGGG
7B-33, 7B-14	TATGATACGGGGGCA
7A-14, 7A-11	CGGTTGTCCGGGTTA
7B-55, 7B-41	CGTGTTGGAGATGTC

3.5.3 Guanine-Gap Statistics, Evidence for MSH

The Multi-Step Hole Hopping model suggests that long distance hole transport is highly dependent on the spacing between neighbouring guanines in a duplex (called guanine "gaps," more details found in Section 3.5 introduction). Our analysis first counted the frequency of guanine-gaps and sorted them according to the guanine-gap length. In other words, "0-gap" means guanines are adjacent to each other, "3-gap" means there are three A/T base pairs between successive guanines. Since inter-strand hopping was no impediment to MSH, the presence of cytosine bases was considered to be equivalent to the presence of guanine. *The frequency of a particular guanine-gap was*

then divided by the total number of guanine-steps observed (in other words, every guanine-gap found). This was termed an "observed probability".

Fig. 3.17 shows the tabulated results of the analysis. The model represents the probability of guanine-gap occurring due to random chance during synthesis (normalized with base composition). This model was used as a standard to calculate the ratio of observed probability/model probability. This ratio (histograms shown in Fig. 3.18) amplifies the differences between observed situation and model calculated. The fluctuations seen in guanine-gaps >6 nucleotides cannot be considered reliable because they can be skewed by a single event. For example, one 10 nucleotide guanine-step in R7 caused a 450% increase relative to the with model.

Since the R0 pool had not undergone selection, we expected the R0 results to roughly mirror the model. This was found to be generally true, with a few fluctuations. In R7 however, the 2-nucleotide guanine-gap observed an increased trend of 37%. While this trend is interesting, the fluctuations observed in R0 results when compared to the model sheds doubt on the statistical significance of the 2-nucleotide guanine-gap trend.

Figure 3.17: On the x-axis is the number of bases between G or C called guanine-gap. On the y-axis is the frequency of a specific gap size divided by total guanine-gap occurrences. White Bar represents the model probability; the chances of random occurrence (given the base composition). Black bar represents observed probability of round 0, grey bar represents observed probability of round 7.

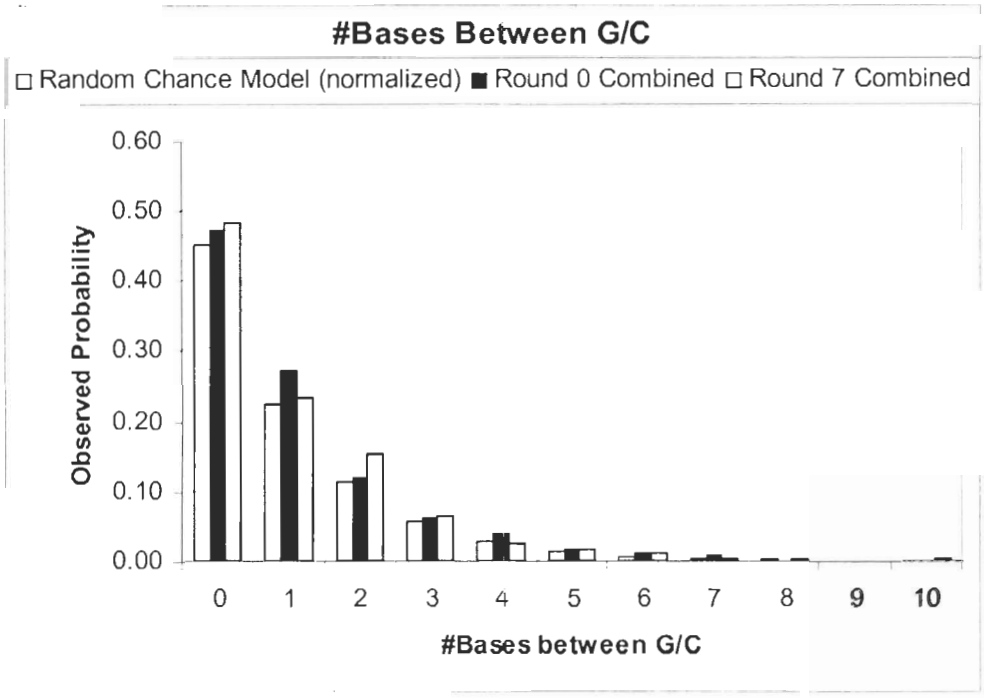
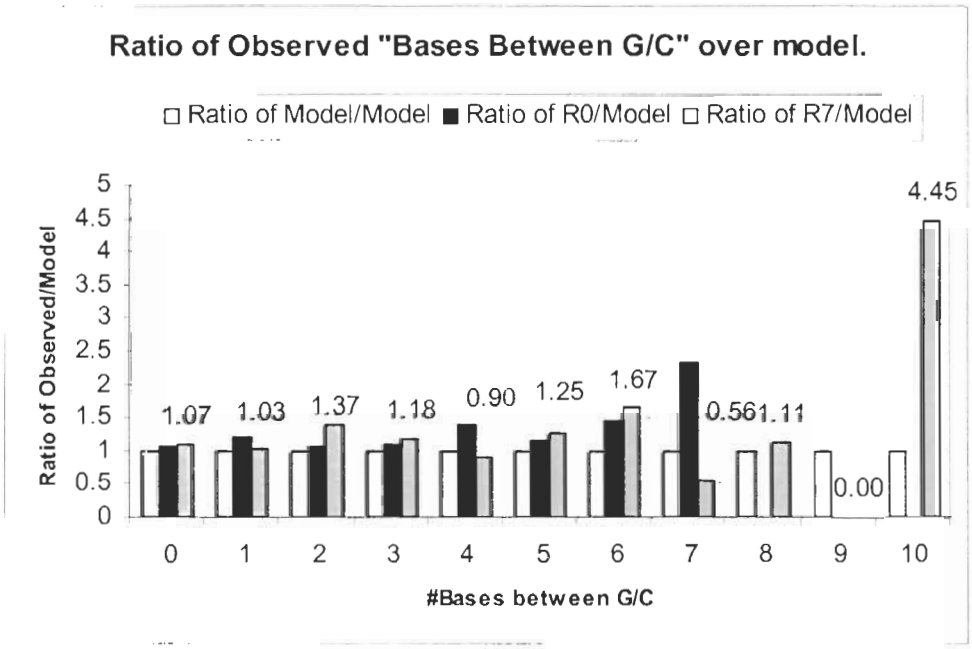


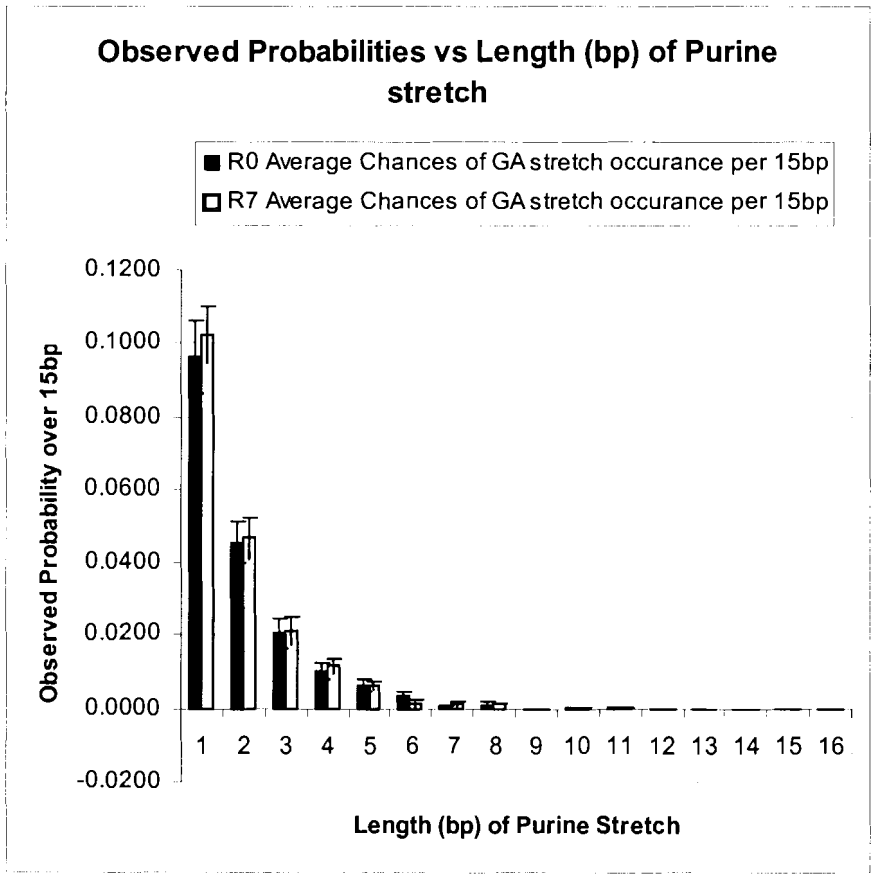
Figure 3.18: Ratio of observed probability/model. White bar (value of 1) means similar to model. Black bar represented observed R0/model, grey bar represents observed R7/model (ratio values included).



3.5.4 Purine Segregation Statistics, Evidence for PPH

The phonon-assisted polaron-like hopping mechanism posits that both sequence identity and purine segregation are important for hole transport. In the next analysis we carried out, the frequency and length of purine stretches were counted and sorted by length. *The frequency of the observed lengths was divided by the length of the N15 random region to give the observed probability.* As tabulated in Fig. 3.19, R7 purine stretches (grey bar) does not differ significantly from R0 (black bar), within standard error. It appeared that during the course of the selection, no significant purine segregation onto either of the duplex DNA strands was occurring.

Figure 3.19: Purine Stretch Occurrence Frequency. Standard Error (SE) bars shown.



3.5.5 2-bp Neighbour Test statistics, Evidence for PPH

At the core of the Phonon-assisted Polaron-like hopping mechanism lies the premise that specific sequences are important in forming the polarons to distribute the charges. The 2-bp Neighbour test simply counts how often a unique 2-bp species (4^2 or 16 combinations possible) appears with in a given N15 sequence. *This count was divided by the length of the random region (usually 15bp) to give a "probability of occurrence per 15bp".* The calculation was carried out for each of the 16 two-nucleotide combinations for each of the 89 R0 and 93 R7 clones. The average probability, the standard deviation, and the standard error (SE), were then calculated. Fig. 3.20 tabulates the R0 and R7 comparisons, along with standard error bars. Also included is a model representing the probability of a 2-bp base combination occurring due to random chance during DNA synthesis (i.e. normalized with initial base composition).

Figure 3.20: **2-BP Neighbour Statistics (16 combinations). Obs R0 is observed probability of 2bp sequence occurring in Round 0. Obs R7 is for round 7. Error bars are +/- Standard Error (SE). "Down" means significant decrease in sequence (>95%CL) in R7 compared R0. "Up" means increase.**

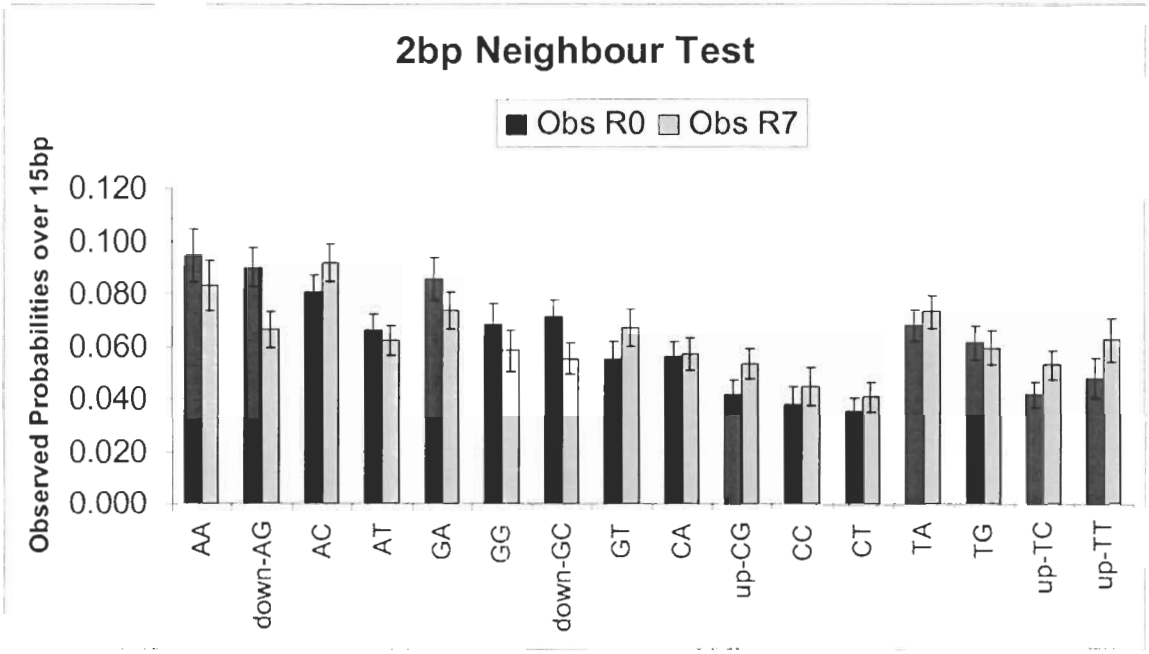


Figure 3.21: **2-bp Neighbour test in Context. Top is SY selection construct. Bolded Middle are sequences that increased significantly compared with starting pool (and their grayed complimentary strand). Note that the 3-bp sequences are not in any particular order order. Bottom are sequences that decreased significantly. Locations are arbitrarily assigned for illustration only, specific locations of sequences are listed in the Appendix.**

```
  1          10          20          30          40          50          59
AQ~TTTAG CTGAC tgact cgNNN NNNNN NNNNN NNcga gtgcG AGACG CtgaC CCATA GTGA 3'
3' AAATC GACTG actga gcXXX XXXXX XXXXX XXgct cacgC TCTGC GactG GGTAT CACT 5'
  59          50          40    N15   30          20          10 9 8       1
```

In R7, CG, TC, TT increased compared with R0 (95%CL).

```
AQ~TTTA---|-CG-TC-TT--|---CCC---3'<-Sequenced "AQ strand"
    AAAT  | GC AG AA |     GGG
```

In R7, AG, GC decreased compared with R0 (95%CL).

```
AQ~TTTA---|-AG-GC-----|---CCC---3'
    AAAT  | TC CG     |     GGG
```

Given the large sampling size in both R0 and R7, statistically significant differences could be detected at 95% Confidence Levels (CL). In R7, **AG** and **GC** species were found with significantly lower frequency relative to R0. **CG, TC, TT** were found more significantly more abundant.

What does these data mean in the context of the original construct? Fig 3.21 illustrates the changes noted between R0 and R7 in the context of the SYB selection construct. *Note that the sequences shown were placed in no particular order; placement was made only to illustrate the location of various features of the construct relative to the sequences of interest.* The DNA strand that was sequenced was called the N-strand. In the context of the original construct, the N-strand contained the 5' C6 linked AQ. For clarity in the Neighbour test analysis, the 5'->3' single strand containing AQ, will be called the "AQ strand". The 3'->5' single strand containing the GGG will be called the "GGG strand". In Fig. 3.21, middle, the 5'TC 3', 5' TT 3' and 5' CG dinucleoside motifs were shown to be significantly more abundant in R7 compared with R0. This means the complementary bases are purines (3'AG 5' and 3'AA 5', respectively). Fig. 3.21, bottom shows that 5'AG 3' and 5'GC 3' was decreased in abundance in R7. This means the complementary bases are pyrimidines (3'TC 5'). The reasons for an increase of 5'CG 3' and a decrease of 5'GC 3' was not immediately obvious. However, it is important to consider that the two are structurally different, owing to the 5'->3' directionality of DNA.

In conclusion, it appears that R7 selection favored the addition of some purines (AG and AA) onto the GGG strand while disfavoring pyrimidines TC. A more complex 3-bp Neighbour test was carried out to acquire more information. On a final note, some

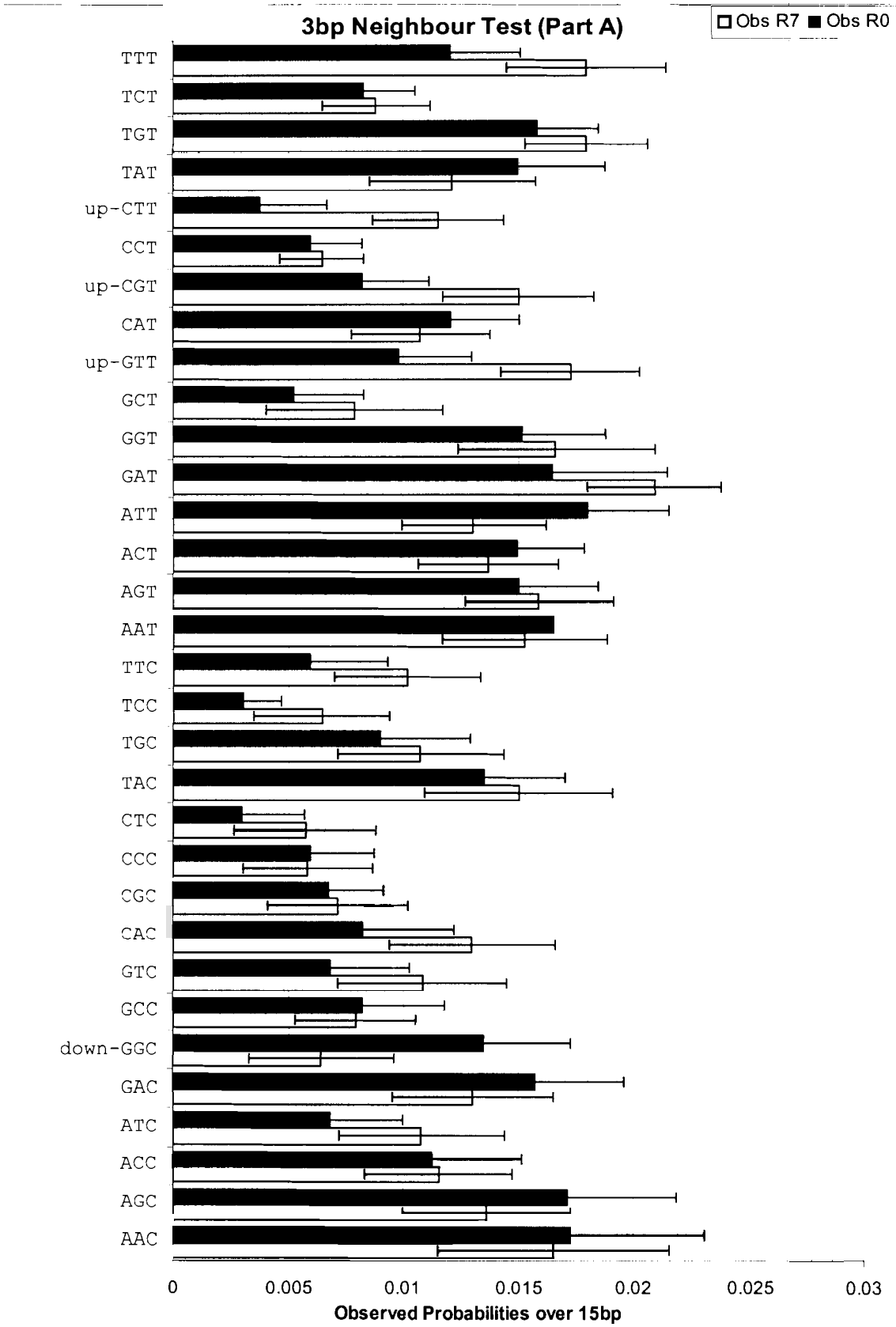
purines such as 3'GG 5' and 3'GA 5'did not seem to have participated in the selection to a significant degree.

3.5.6 3-bp Neighbour Test statistics, Evidence for PPH

The 3-np Neighbour test was more complex because of the higher number of possible unique 3-bp species (4^3 or 64 possible combinations, called motifs). The increased number of combinations also meant a lesser sampling size, which meant that even a confidence level of 95% was on the borderline of significance and could only be called a "trend". A 99% confidence level was used to determine significance. The "observed probabilities over 15bp" were calculated in the same manner as in the 2-bp Neighbour test. The results are tabulated in Fig. 2.34 (two pages) with SE bars.

Under the stringent 99% CL criteria, 5'GTT 3', 5'TCA 3', 5'CGT 3', and 5'CTT 3' motifs had significantly higher abundance in R7 compared with R0. 5'AGA 3', 5'GGC 3' and 5'GCA 3' decreased significantly in R7 relative to R0. "Runners up" motifs that met only 95% CL (indicates trends) were as follows: 5'ACG 3', 5'CAC 3', 5'CGA 3', 5'TCC 3', 5'TTC 3' increased in abundance. 5'GAG 3', 5'ACG 3' decreased in abundance.

Figure 3.22: (Next Two Pages) 3-BP Neighbour Statistics (64 combinations). Obs R0 is observed probability of 3bp sequence occurring in Round 0. Obs R7 is for round 7. Error bars are +/- Standard Error. Differences between R0 and R7 greater than 99%CL is labeled "up" for increased occurrence in R7 or "down" for decreased occurrence in R7.


3bp Neighbour Test (Part A)
Obs R7
Obs R0
TTT
TCT
TGT
TAT
up-CTT
CCT
up-CGT
CAT
up-GTT
GCT
GGT
GAT
ATT
ACT
AGT
AAT
TTC
TCC
TGC
TAC
CTC
CCC
CGC
CAC
GTC
GCC
down-GGC
GAC
ATC
ACC
AGC
AAC
0
0.005
0.01
0.015
0.02
0.025
0.03
Observed Probabilities over 15bp

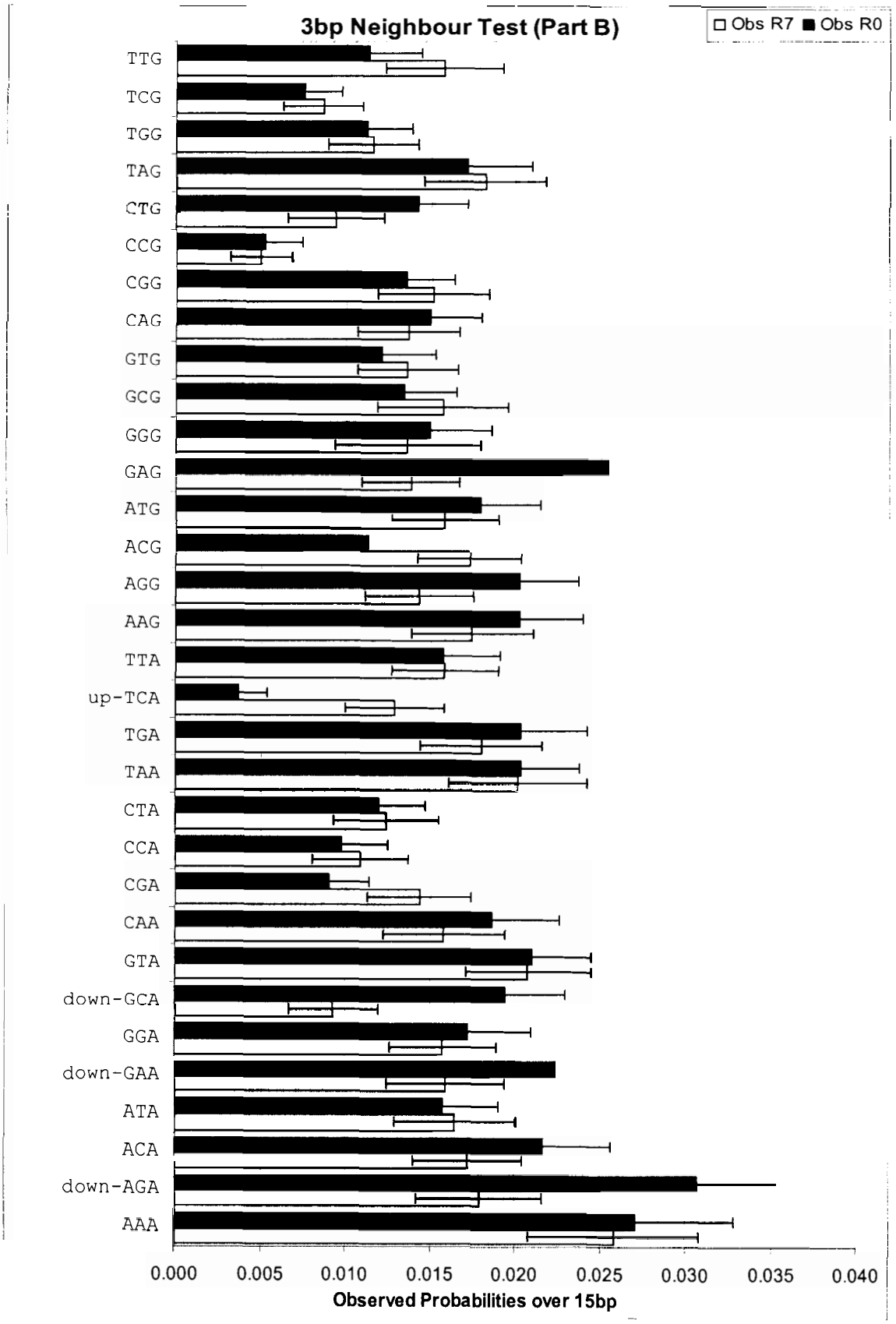

3bp Neighbour Test (Part B)
Obs R7
Obs R0
TTG
TCG
TGG
TAG
CTG
CCG
CGG
CAG
GTG
GCG
GGG
GAG
ATG
ACG
AGG
AAG
TTA
up-TCA
TGA
TAA
CTA
CCA
CGA
CAA
GTA
down-GCA
GGA
down-GAA
ATA
ACA
down-AGA
AAA
0.000
0.005
0.010
0.015
0.020
0.025
0.030
0.035
0.040
Observed Probabilities over 15bp

Figure 3.23: 3-bp Neighbour test in context of original selection construct. At the top is the selection construct. In the middle (boxed and bolded) are sequences that were found in significantly greater abundance when compared with starting pool (boxed and grayed sequences represent complimentary strand). Bottom are sequences that had decreased significantly. Locations of 3bp motifs are arbitrarily assigned for illustration only, specific sequences are listed in the Appendix 5.3.1.

```
  1          10          20          30          40          50          59
AQ~TTTAG CTGAC tgact cgNNN NNNNN NNNNN NNcga gtgcG AGACG CtgaC CCATA GTGA 3'
3' AAATC GACTG actga gcXXX XXXXX XXXXX XXgct cacgC TCTGC GactG GGTAT CACT 5*'
 59          50          40   N15    30          20          10 9 8       1
```

In R7, GTT, TCA, CGT and CTT increased in abundance compared with R0 (99%CL).

```
AQ~TTTA---|-GTT TCA CGT CTT-|---CCC---3'<-Sequenced "AQ strand"
    AAAT    | CAA AGT GCA GAA |   GGG
```

In R7, AGA, GAA, GGC and GCA decreased in abundance compared with R0 (99%CL).

```
AQ~TTTA---|-AGA GAA GGC GCA-|---CCC---3'
    AAAT    | TCT CTT CCG CGT |   GGG
```

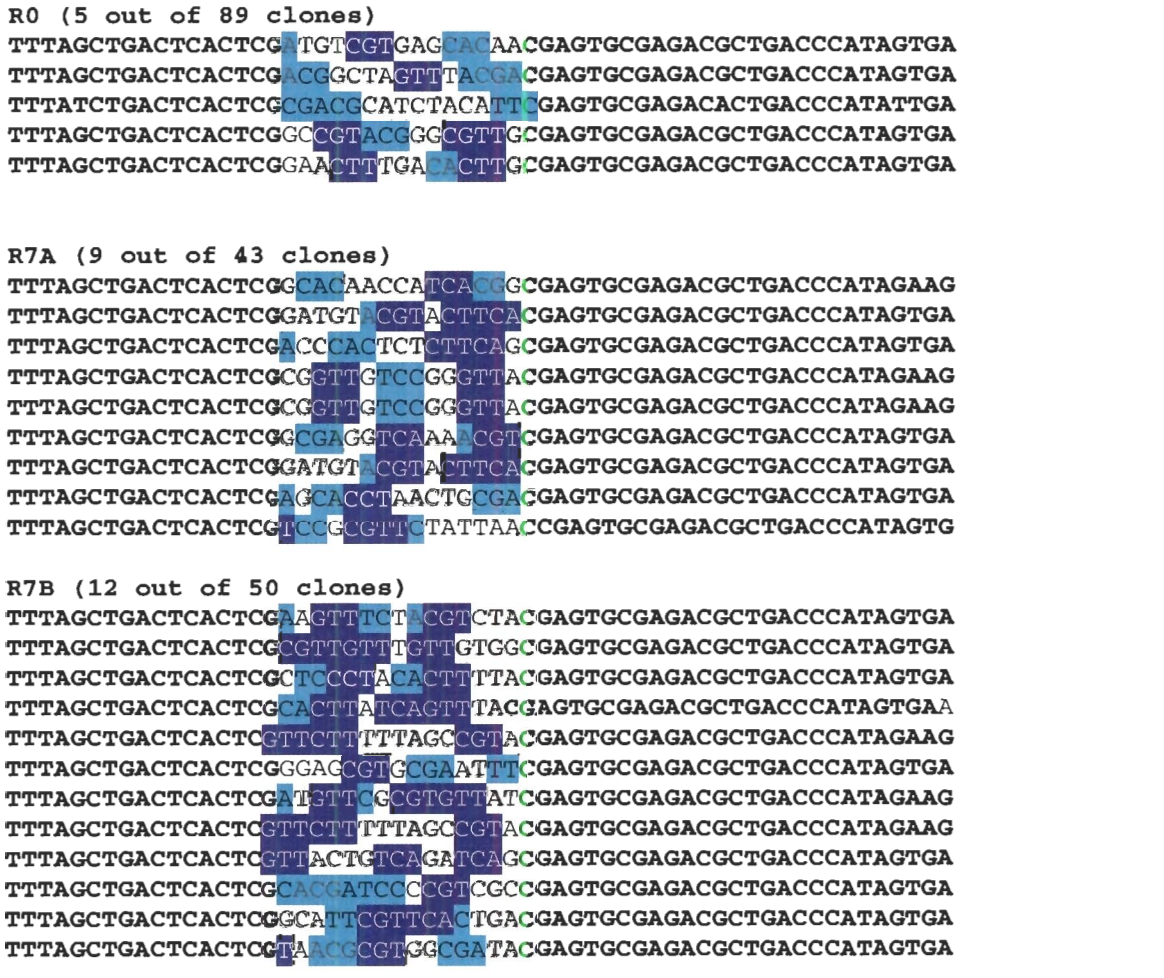
Fig. 3.23 illustrates the changes noted in R7, compared with R0, in the context of the SYB selection construct. Again, the sequences shown are placed in no particular order; placement is only meant to illustrate the location of various features of the construct relative to the motifs of interest. The sequenced strand was the 5'->3' AQ strand. The GGG strand was the complementary strand present during the selection. All four motifs that increased in abundance in R7, (Fig. 2.35 Middle) (5'GTT 3', 5'TCA 3', 5'CGT 3', and 5'CTT 3') had complementary strands with at least two purines (3'CAA 5', 3'AGT 5', 3'GCA 5', 3'GAA 5'). In contrast, all four motifs that decreased in abundance in R7 (Fig. 2.35 Bottom) (5'AGA 3', 5'GAA 3', 5'GGC 3', 5'GCA 3') had at least two pyrimidines (3'TCT 5', 3'CTT 5', 3'CCG 5', 3'CGT 5'). The segregation of purines on to the GGG strand in round 7, was also supported by trends observed from the "runners up" motifs.

In conclusion, the 3-bp neighbour analysis of the N15 region in R7, suggested that the selection favoured the segregation of purine rich 3-bp motifs onto the GGG strand while disfavouring pyrimidine species. This is consistent with the PPH prediction of hole transport favoring purine strands. However, while the PPH theory predicts purine stretches on both strands of the duplex (Liu et al., 2002), we observed segregation of purines to only one strand of the duplex DNA, specifically on the GGG strand. This specific purine segregation we observed, may be partially explained by the A-hopping mechanism advanced by groups such as Majima et al (Majima, 2004). Initial hole injection by photooxidants such as AQ is said to occur at the adenine bridge through A hopping. According to this theory, a hole injection would occur through the adenine bridge of our construct, which is on the 3' end of the GGG strand. Therefore, it does not appear coincidental when the selection of purines occurred on this same strand.

The notion that all purines are favoured on GGG strand however, is overly simplistic. Many motifs such as TCT, CCT, CTA, CCC with purine rich complementary strands did not show any significant change at all in R7, when compared with R0. Therefore, the 3-bp motif found to be significantly increased in R7 (**5'GTT 3'**/CAA, **5'TCA 3'**/AGT, **5'CGT 3'**/GCA, and **5'CTT 3'**/GAA), probably did not occur by accident, and could be termed "transport permissive candidate" motifs. 8 transport permissive candidate motifs were found, 4 of which were statistically significant to 99% CL.

3.5.7 Overall Picture, SYB Analysis Combined

Figure 3.24: **Clones with transport permissive candidate motifs making up more than half (>8/15) of the N15 region. Dark blue illustrates 99%CL transport permissive candidate motifs, light blue illustrates 95%CL permissive motifs, bold represents the constant flanking regions. Note that R7A+B had 21/93 of these clones with high proportion of candidate motifs while R0 only had 5/89. At the bottom are motifs found in R7 and frequency of occurrence.**



Transport Permissive Motifs and frequency of occurrence in R7

5'CTTCA3' /GAAGT x3 (found three times in R7 clones)
5'GTTC 3' /CAAG x6
5'GTTG 3' /CAAC x12
5'ACGT 3' /TGCA x8

Fig. 3.24 succinctly illustrated the results of the SYB analysis. Clones with transport permissive candidate motifs taking up more than half (8/15) of the N15 region were selected. R0 clones did not contain many of such clones (5 out 89 clones). In contrast, R7 clones had a much higher (21 out of 93 7A+7B) number of such clones. Transport permissive candidate motifs in R7 also tended to cluster together, sometimes joining to form easily recognized sequences. These clones were the starting point of finding longer motifs important to charge transport. Promising candidates of these longer motifs were included in Fig. 3.24, bottom. The complete list of the raw data and with the highlighted transport permissive candidate sequences were included in the Appendix Section 5.3.1.

4 CONCLUSIONS

Hole transport has been known to occur in duplex DNA for the last fifteen years. Given its potential importance in cancer and aging studies, and applications in the field of nanotechnology, the mechanism of hole transport has been extensively investigated. While there has been many prior experiments in support of both prevalent theories of long distance hole transport, namely the Multi-Step Hole Hopping mechanism (MSH) and the Phonon-assisted Polaron-like Hopping mechanism (PPH), these experiments have been conducted using rationally designed DNA sequences.

In this study we present an *in vitro* selection project aimed at obtaining efficient charge conducting sequences without the use of rational design methods. While MSH predictes that conductivity was based only on the regular spacing of guanines (on either strand of the duplex), PPH predictes that the exact sequence context and purine segregation (on either strand) is important for hole conductivity. The *In vitro* selection we proposed, was used precisely because it did not favour either theory in its design.

When R7 selection sequences (Section 3.4.2) showed a 3-fold increase in hole transport efficiency compared with initial round, the selection pool were cloned and sequenced to identify these efficient conductors. Base composition studies (Section 3.5.1) showed that R7 population was significantly different from R0. Alignment (Section 3.5.2) of N15 region of R0 and R7 clones did not at first, reveal any obvious patterns. Statistical methods were employed to find subtle differences. G-gap statistics (Section 3.5.3) was used to study the distances between guanines, an important measure in the

Multi-step Hole Hopping theory. While it did not find any significant changes in R7, a 2-gap increased trend was observed. Purine segregation statistics (Section 3.5.4), based on the Phonon-assisted Polaron-like hopping mechanism, did not detect significant differences in R7, suggesting that having sequential purines on either strand was not a sufficient precondition in facilitating hole conduction. Ultimately, the 2-bp and 3-bp Neighbour test (Section 3.5.5, 3.5.6), found specific purine motifs segregated to the specific strand where initial charge injection occurred. 8 transport permissive candidate motifs were found, 4 of which were statistically significant to 99% CL.

Finally, when these specific purine motifs, termed "transport permissive candidate motifs," were identified on the individual R7 clones (Section 3.5.7). Clones with transport permissive candidate motifs taking up half their random sequences were identified and longer common motifs were found (Fig. 3.24). In R0 there were only 5/89 of such clones found, (possibly due to random chance). In R7 there were 21/93 of such clones found. The characterization of the longer motifs present within many of these clones may shed light on why R7 selection pool were better hole conductors than R0. The best explanation for the evidence presented was PPH, since it predicted the to some extent, the purine segregation as well as the importance of sequence context.

Future work in this project requires follow up experiments, to characterize the conductivity of the transport permissive candidate motifs. An example of such experiments is shown in Fig. 4.1 The test construct would consist of a middle test region flanked by two GG detectors (proximal and distal), with an adenine bridge section near AQ to facilitate charge injection. A comparison of proximal and distal detector damage would characterize the relative efficiency of hole transport between different sequences.

The test sequences to be included would consist of varying repeats of transport permissive candidate motifs. These experiments would characterize the relative conductivity of the permissive sequences and compare them to sequences known to be poor or excellent conductors (such as Liu et al, 2002 sequences). Our observations noted that the purines were only segregated to the strand containing the adenine bridge. Further tests comparing (Fig 4.1) purine sequences on the X-strand with the N-strand could confirm this observation.

Figure 4.1: Future Work: Characterizing the transport permissive candidate motifs. Top shows a hypothetical construct where XXXX/NNNN can be various repeats of test sequences. A list of possible test sequences are suggested at the bottom.

AQ~TTTTCCTT**NNNNNNNN**CCATATAT-3′

3′AAAA**GG**AA**XXXXXXXXGG**TATATA-5′

Where XXXX/NNNN are:
Transport Permissive Sequences (test sequences)
5′CTTCA3′ /GAAGT
5′GTTC 3′ /CAAG
5′GTTG 3′ /CAAC
5′ACGT 3′ /TGCA
Examples of known poor conductors (Liu et al 2003)
5′TATA 3′ /ATAT
Examples of known good conductors (Liu et al 2003)
5′CCTT 3′ /GGAA

5 APPENDIXES

5.1 Additional Materials and Methods Detail

5.1.1 Pre-Experiment Preparations (Kinase, Loading Buffers, and G, CT Ladder creation)

KINASE REACTION (40min)
Prepared fresh every third round.
-In Screwcap Tube
+16μL H2O
+5μL FWD RXN Buffer
+2μL 100uM p2 (primer name)
+1μL T4 Kinase
+1μL Gamma P^{32} ATP
TotVol=25uL

-Incubate 37C bath for 30min

Ethanol precipitation (45min)
-To screwcap tube (TotVol 25uL):
+3.00μL 2M NaCl (250mM effective conc)
+70μL 100% EtOH
-Vortex
-Cool in dry ice 5min
-Cold room centrifuge 20min @ 13000 (max) rpm.
-Discard hot EtOH in liq hot waste
-+70μL 70% EtOH
-Discard hot EtOH in liq hot waste
-Airdry few minutes
-Resuspend in 10μL Denaturing Loading Buffer

Purification of Kinase, 12% Polyacrylamide Gel (2hr)
In 50mL grad cylinder
-15mL 40% 19:1 Acryl:Bis (12% effective conc)
-25g Urea (8M effective conc)
-5mL 10x TBE (1xTBE effective conc)
Top off to 50mL with ddH2O
TotVol=50mL

-Tape short plates
-+ 500μL 10% APS / Mix

-+20μL TEMED / Mix quickly
-Pour into plates, wait 15min.

-Running Buffer (1x)
+100mL 10xTBE
+Fill to 1L ddH2O
Mix

-Prerun with denaturing buffer ~1hr.
-Flush wells
-Heat 10μL Denaturing dye with Kinase oligos 99C for 1min.
-Load 10μL into each lane.

Phospho-imaging plates Developing (20min)
-Dot 4 corners of gel with radioactive dye (leftover from loading kinase)
-Wrap gel with seran wrap.
-Expose phospho-imaging plate for 1min (or less)
-Develop on Typhoon.

Eluding DNA from Gel (2hr)
-Band cut and put into 1.3mL Eppendorf
-400μL 300mM NaCl (~250-300mM effective concentration)
-Shake 2hrs

Ethanol Precipitation after Kinase Purification (45min)
-Remove supernatant carefully and place into another 1.3mL Tube. (Roughly ~300μL solution left)
+1000μL 100% EtOH (~3x 300μL by volume)
-Vortex
-Cool in dry ice 5min
-Cold room centrifuge 20min @ 13000 (max) rpm.
-Discard EtOH in hot waste
-+200μL 70% EtOH
-Discard EtOH
-Airdry few minutes
-Store in 20μL of TE (ss oligos, so TE okay)

LOADING BUFFERS:
Denaturing Loading Buffer (10mL):
+10mL Formamide
+10mg Xylenecyanol FF (aka pinch)
+10mg Bromophenol Blue
+20μL 0.5M EDTA (1mM effective)
+100μL Tris (10mM effective)

Non-denaturing (Native) Loading Buffer (10mL):

+10mg Xylenecyanol FF
+10mg Bromophenol Blue
+3mL Glycerol
+20μL 0.5M EDTA (1mM effective conc)
+500μL 1M TrisCl (50mM effective)
Fill to 10mL ddH2O

G, CT LADDERS (do while PCR)
-In 0.6mL tube
+4μL hot p2GGG
+2μL 100uM p2GGG
+15μL ddH2O
TotVol=21uL
-Heat with PCR to 90C 1min. Cool on ice.

-G-Ladder: In 1.3mL tubes	**CT Ladder: In 1.3mL tubes**
+140μL ddH2O +50μL 200mM LiCacadilate (*CAUTION*) +10μL above heated oligos +1μL 100% DMS TotVol~201uL Incubate 10min @ RT +50μL Stop solution (400μL ddH2O + 500μL NaAcetate(3M) + 70μL b-mercaptoethanol) +750μL 100%EtOH	+10μL ddH2O +30μL Hydrazine (new stock) +10μL above heated oligos TotVol~50uL Incubate 10min @RT +20μL 3M NaAcetate +180μL ddH2O +750μL 100% EtOH

Ethanol precipitation after Ladder
(NaAcetate and EtOH already added)
-Vortex
-Cool in dry ice 5min
-Cold room centrifuge 30min @ 13000 (max) rpm.
-+100μL 70% EtOH x2
-Airdry few minutes

5.1.2 Selection Procedures (two day procedure)

DAY 1 **-PCR 3hrs**	**PCR Amplification of Selection** -In 0.6mL Tube (30min) +65uL H2O (OR 74µL H2O for SYK+/SYR0- Control) +10uL PCR Buffer (x10) +10µL dNTP (x10) +1uL 100uM p2GGG +1uL 100uM p1AQ --- +10uL Selection Template OR 1µL of 1uM SYK+/SYR0- Control +2uL p2GGG *HOT +1uL Taq Polymerase TotVol=100uL -12 Rounds of PCR (2hrs) ***WAIT***
***(GOTO Run Native Gel) **-Native Gel prep 2hrs**	**Ethanol precipitation (45min)** -To screwcap tube (TotVol 100uL): +12.5µL 2M NaCl (250mM effective conc) +300µL 100% EtOH -Vortex -Cool in dry ice 5min -Cold room centrifuge 20min @ 13000 (max) rpm. -Discard hot EtOH in liq hot waste -+70µL 70% EtOH -Discard hot EtOH in liq hot waste -Airdry few minutes -Resuspend in 10µL Native Loading Buffer **Purification of Selection, 12% Native Gel (6hrs)** -In 100mL grad cylinder. +30mL 40% 19:1 Acryl:Bis (12% effective conc) +10mL 10xTB (1x TB) Fill to 100mL ddH2O

	-Tape short plates using Thick Spacers -+ 500µL 10% APS / Mix -+20µL TEMED / Mix quickly -Pour into plates, wait 30min to fully polymerize. -Running Buffer +100mL 10xTB (1x) +Fill to 1L ddH2O Mix -Prerun with Non-denaturing (Native) Loading Buffer ~1hr -Flush wells -Load 10µL into each lane
-Run Native Gel 4hrs.	-Run Native Gel (4hrs) ***WAIT***
-Elution of PCR 3hrs	**Phospho-imaging plate Developing (20min)** **Eluding PCR products from 12% native gel (2hrs)** -Band cut, crushed and put into 1.3mL Eppendorf -Elute in 400µL 300mM NaCl -Elute (2hrs) ***WAIT OR STOP*** **Ethanol Precipitation after Native Gel Purification (45min)** -Remove supernatant carefully and place into another 1.3mL Tube. (Roughly ~300µL solution left) +1000µL 100% EtOH (~3x 300µL by volume) -Vortex -Cool in dry ice 5min -Cold room centrifuge 20min @ 13000 (max) rpm. -Discard EtOH in hot waste -+200µL 70% EtOH -Discard EtOH -Airdry few minutes -Store Purified PCR product in 20µL of Irradiation Buffer (50mM Tris Buffer, 50mM NaCl). (Storage of current round samples)
D1=10hrs	**"After" Phospho-imaging plate Developing (ON)**

DAY 2 **-Irradiation 1hr**	**Irradiation (30min)** +10µL Purified PCR Product in 0.6mL Tube. -Irradiation 333nm for 5min - 4cm from lamp. **Ethanol precipitation after Irradiation (45min)** -To 0.6mL tube with Irradiated Sample (TotVol 10uL): +1.5uL 2M NaCl (250mM effective conc) +30µL 100% EtOH -Vortex -Cool in dry ice 5min -Cold room centrifuge 20min @ 13000 (max) rpm. -+30µL 70% EtOH x2 -Airdry few minutes
-Piperidine Treatment 2.5hrs *****GOTO Run Sequencing Gel** **-Sequencing Gel Prep 2hrs**	**Piperidine treatment after Irradiation and Ladders (2.5hr)** -In 0.6µL tube, add 100µL 10% Piperidine (transfer ladder into 0.6mL tubes) -Heat 90C for 30min using PCR. -Quench in ice, Spin down -+200µL ddH2O (TotVol=300uL) -Speedvac overnight or (~2hrs) with heat. ***WAIT*** **Sequencing 12% Large Polyacrylamide Gel (5hrs)** In 500mL grad cylinder -60mL 40% 19:1 Acryl:Bis (12% effective conc) -100g Urea (8M effective conc) -10mL 10x TBE (0.5xTBE effective conc) Top off to 200mL with ddH2O TotVol=200mL -Tape long plates -In 100mL Grad cylinder, add 100mL above gel, -+ 500µL 10% APS / Mix -+50µL TEMED / Mix quickly

	-Pour into plates, wait (~30min). ***WAIT***
	-Running Buffer +50mL 10xTBE +Fill to 1L ddH2O Mix (enough for two gel boxes assuming no leakage)
	-Prerun with denaturing dye (~1hr) ***WAIT***
	-Flush wells -Add ~15µL denaturing dye to dried oligos. Heat 90C for 1min.
-Run Sequencing Gel 3hrs	-Load 30 cps worth of oligos into each lane (try for 5-10µL for sharp bands). -Run Sequencing Gel (3hrs)
Elution of Irradiation 5hrs	***WAIT***
	Phospho-imaging plate Developing (2hr)
	Eluding Irradiation products from 12% Denaturing gel (2hrs) -Band cut, crushed and put into 1.3mL Eppendorf -Elute in 400µL 300mM NaCl -Elute (2hrs) ***WAIT OR STOP***
	Ethanol Precipitation after Irradiation Denaturing Gel (45min) -Remove supernatant carefully and place into another 1.3mL Tube. (Roughly ~300µL solution left) +1000µL 100% EtOH (~3x 300µL by volume) -Vortex -Cool in dry ice 5min -Cold room centrifuge 20min @ 13000 (max) rpm. -Discard EtOH in hot waste -+200µL 70% EtOH -Discard EtOH -Airdry few minutes
D2=11hrs	-Store Irradiation Template product in 20µL of TE
	"After" Phospho-imaging plate Developing (ON)

5.2 SY Series Clones

Round 5

TTTAG CTGAC TCACT CGAGT GAGAG GACTC TTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGGCA AACAG TGCTA GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGATG TAAGT GGATA ATCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGA CGAAG CAGCC GGCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGACA AAGGG TACGT AACGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGCCA CAAGG CGCTT ACCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAAA CGTTT ACGGT TCCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGTAT CGACG CCGAG CGCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGCCA AGGTT GTCAT AACGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGTTC AAATA AGTAT TGCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAAA CGTTT ACGGT TCCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGGTA GCACG TCGAT AGCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGTAG TAAAC ACTAG TGCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGCAT TTGCG GCAAG CTACGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGA AAGAG GGACT ATCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGACG ATATT GAAGG GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGACG GAACG CCCCT AACGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGGGG GGGCA ACACC TCGCG GGGGG CAAAC ACCCT GCACC CTTA

Round 10

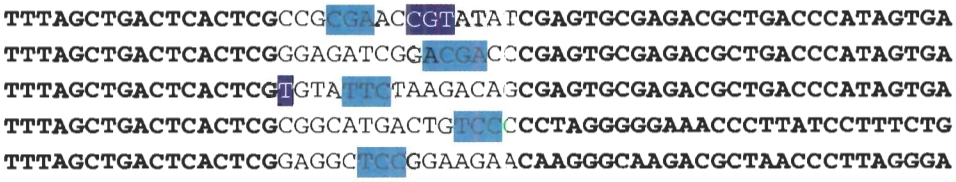
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGCAG CGTCA CTGCG GGCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGACG ATCAT ACCCA ATCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGGTG ATGTT AGCGA CTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGGCG ACGAA ACAAT ATCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGCAA TGTGA AGCGC CGCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGGCC ATCAG CGCCT CCCGA GTGCG AGACG CTAAG CCGAA TTCC
TTTAG CTGAC TCACT CGTCC CGATG TCTGA ATCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGTTA GAGGG GGAGG AGCGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CGAGT GTGCA GACTC GTCGA GTGCG AGACG CTGAC CCATA GTGA

TTTAG CTGAC TCACT CG**GTA TGTCT TAAGG GG**CGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CG**GAC CGGAT TTAGG GT**CGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CG**AGT GTGCA GACTC GT**CGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CG**TAA ACTTA CTCCG CA**CGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CG**ACC CGGAT CAAAC AG**CGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CG**AGT GTGCA GACTC GT**CGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CG**GAA TAGAC AACTG CA**CGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CG**AGT GTGCA GACTC GT**CGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CG**AGT GTGCA GACTC GT**CGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CG**CAA ACTAA ATTTG CC**CGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CG**ACG ACGCC TGGTA CC**CGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CG**AGT GTGCA GACTC GT**CGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CG**AGT GTGCA GACTC GT**CGA GTGCG AGACG CTGAC CCATA GTGA
TTTAG CTGAC TCACT CG**CAG CTTGT GATAG AC**CGA GTGCG AGACG CTGAC CCATA GTGA

5.3 SYB Series Clones

5.3.1 SYB Series Sequence, Raw Data.

Round 0, Light highlights represents 95% CL transport permissive candidate motifs, dark highlights represents 99% CL transport permissive candidate motifs.



Round 7A, Light highlights represents 95% CL transport permissive candidate motifs, dark highlights represents 99% CL transport permissive candidate motifs.

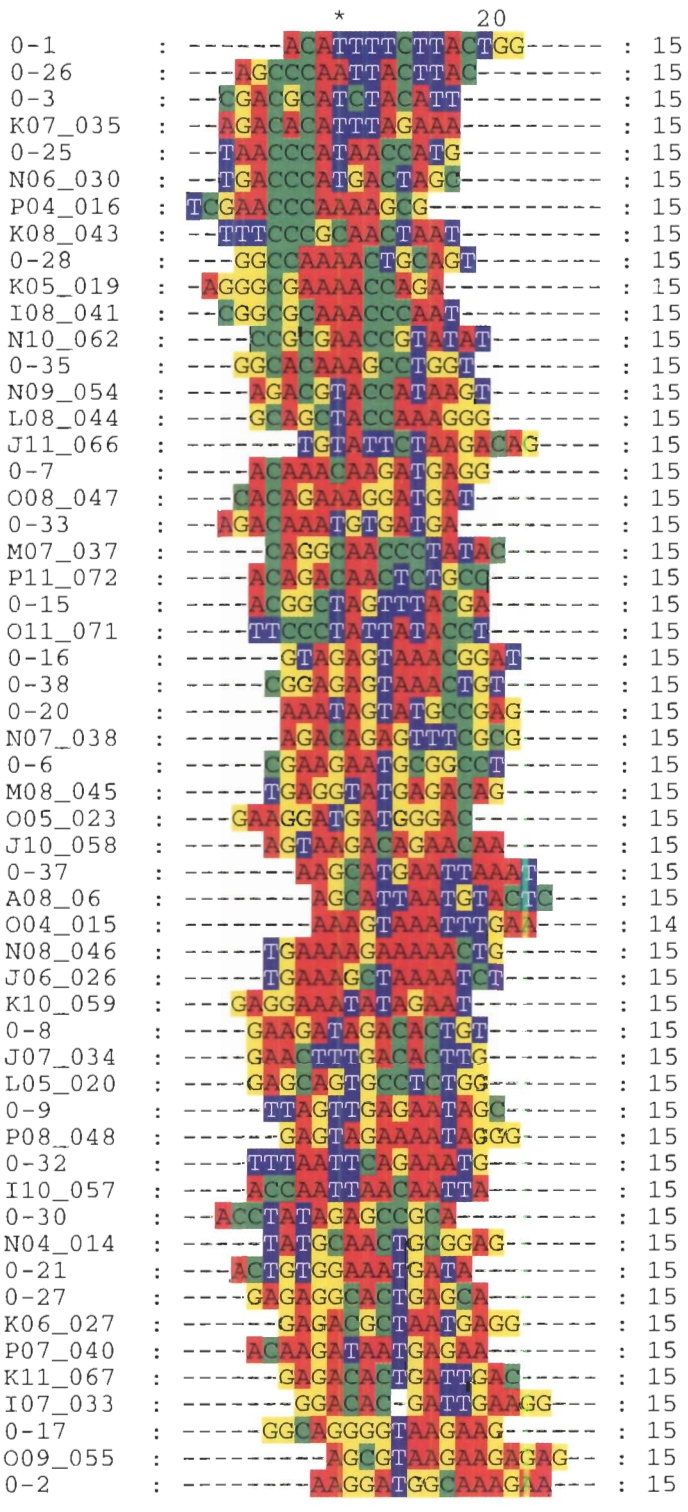
Round 7B Light blue highlights 95% CL transport permissive candidate motifs, dark blue highlights 99% CL transport permissive candidate motifs.

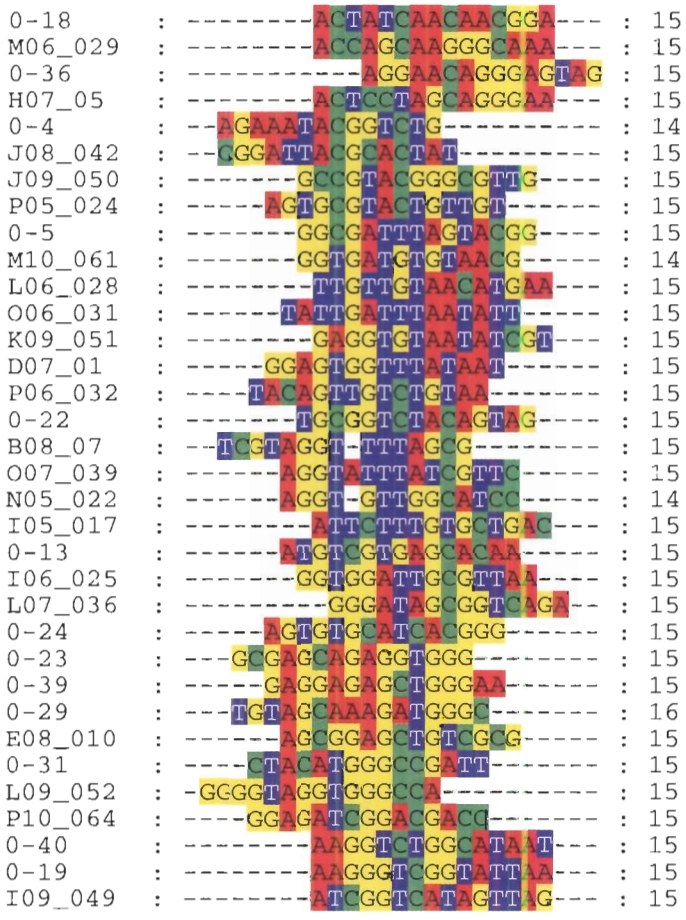
Duplicated Clones

>7A-2	>7A-18	3'	TAGGGAGATAGTAAA
>7A-12	>7A-20		GCGAGGTCAAAACGT
>7A-8	>7A-37		GCACAACCATCACGG
>7A-34	>7A-19		TGACGAGATTAGGAC
>7A-15	>7A-4		CACAGGTTGGTATGA
>7A-30	>7A-9		GATGTACGTACTTCA
>7B-34	>7B-26		TTCTTTTTAGCCGTA
>7A-36	>7A-7		CTGATCATCTGCGGG
>7B-33	>7B-14		TATGATACGGGGGCA
>7A-14	>7A-11		CGGTTGTCCGGGTTA
>7B-55	>7B-41		CGTGTTGGAGATGTC

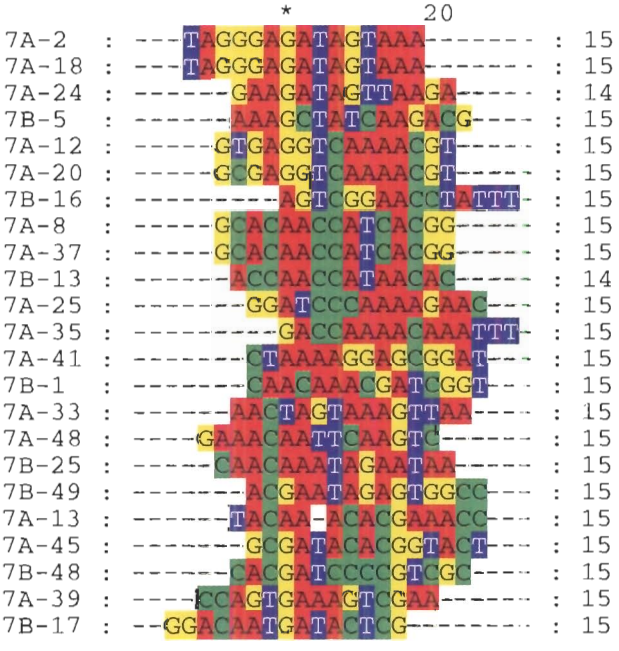
5.3.2 SYB Series Sequence Alignment

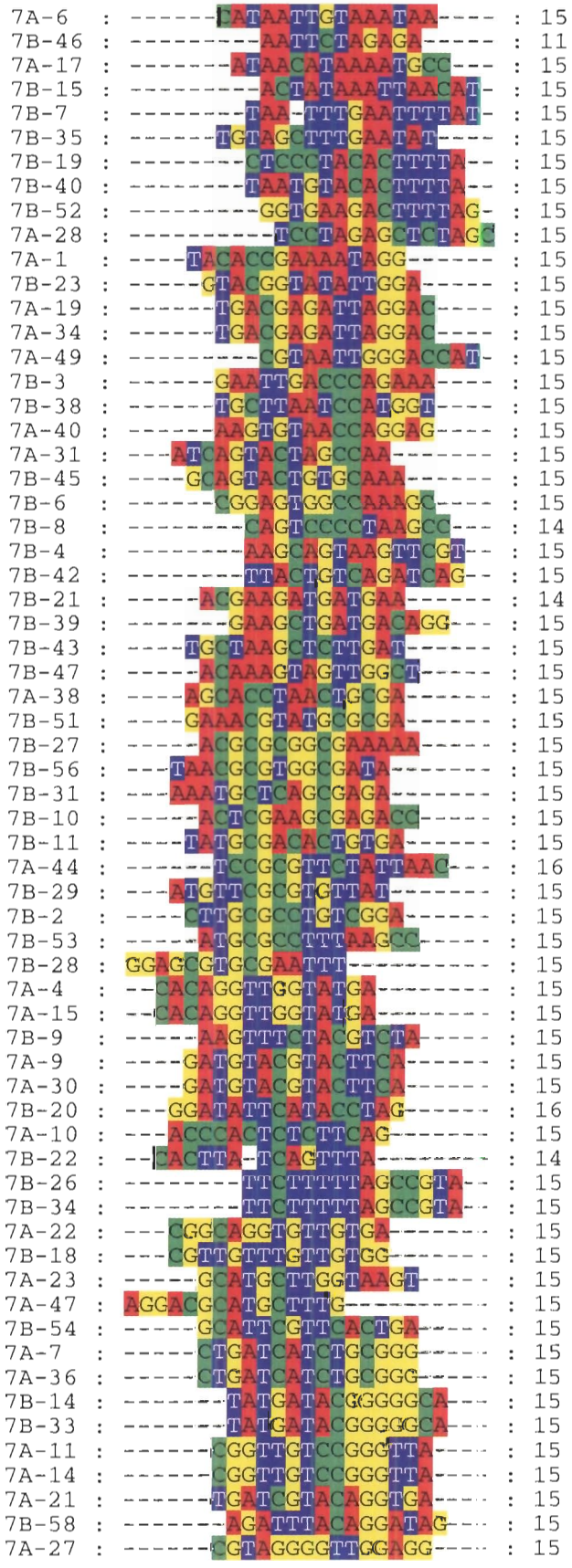
ClustalX alignment used for initial aligr ment of clones, viewed with GeneDoc Viewer.

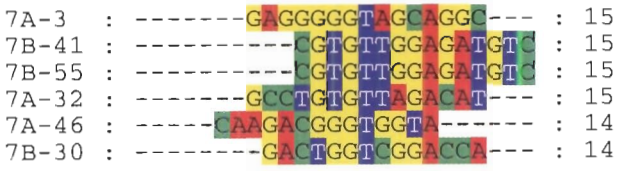




Round 7 Combined (SYB-7A, SYB-7B)




7A-6 : 15
7B-46 : 11
7A-17 : 15
7B-15 : 15
7B-7 : 15
7B-35 : 15
7B-19 : 15
7B-40 : 15
7B-52 : 15
7A-28 : 15
7A-1 : 15
7B-23 : 15
7A-19 : 15
7A-34 : 15
7A-49 : 15
7B-3 : 15
7B-38 : 15
7A-40 : 15
7A-31 : 15
7B-45 : 15
7B-6 : 15
7B-8 : 14
7B-4 : 15
7B-42 : 15
7B-21 : 14
7B-39 : 15
7B-43 : 15
7B-47 : 15
7A-38 : 15
7B-51 : 15
7B-27 : 15
7B-56 : 15
7B-31 : 15
7B-10 : 15
7B-11 : 15
7A-44 : 16
7B-29 : 15
7B-2 : 15
7B-53 : 15
7B-28 : 15
7A-4 : 15
7A-15 : 15
7B-9 : 15
7A-9 : 15
7A-30 : 15
7B-20 : 16
7A-10 : 15
7B-22 : 14
7B-26 : 15
7B-34 : 15
7A-22 : 15
7B-18 : 15
7A-23 : 15
7A-47 : 15
7B-54 : 15
7A-7 : 15
7A-36 : 15
7B-14 : 15
7B-33 : 15
7A-11 : 15
7A-14 : 15
7A-21 : 15
7B-58 : 15
7A-27 : 15

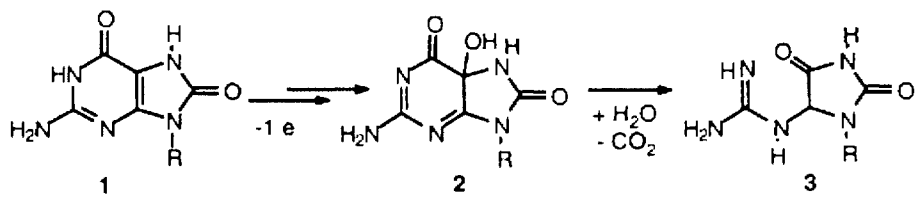

7A-3 : -------GAGGGGGTAGCAGGC--- : 15
7B-41 : ----------CGTGTTGGAGATGTC : 15
7B-55 : ----------CGTGTTGGAGATGTC : 15
7A-32 : --------GCCTGTGTTAGACAT--- : 15
7A-46 : -----CAAGACGGGTGGTA------ : 14
7B-30 : --------GACTGGTCGGACCA--- : 14

5.4 Additional Experiments

5.4.1 Detector Evaluation, ETS Series: GG-8-oxoG vs GGG Detector

The ability to monitor charge transport is directly related to how successful the selection would be. The goal of the ETS series is to test the suitability of 8-oxoG (GG-8-oxoG) as a detector compared with the conventional GGG detector. Previous experiments by this lab used double or triple guanine sequences for monitoring oxidative damage for reasons mentioned before in section 1.1.6. However, 8-oxoG has been shown to be a potentially good detector of charge transport, but has not been used in our lab before. In normal guanine oxidation, only a small proportion of the oxidatively damaged product is piperidine labile (susceptible to piperidine cleavage). This means only a small proportion of the damage can be visualized. The major guanine oxidation product 7,8-dihydro-8-oxoguanine (8-oxo-G) for example, is only moderately piperidine labile. Further oxidation of 8-oxo-G would yield 5-OH-8-oxo-G and guanidinohydantoin (Fig 2.2), both of which are highly piperidine labile. We theorize that if we replaced the 5' most guanine in a triple guanine sequence with an 8-oxoG, we can create a detection region that would be more sensitive to oxidative damage.

Figure 5.1: 8-OxoG products, the piperidine labile 5-OH-8-oxo and guanidinohydantoin



There are some challenges that needed to be overcome if 8-oxoG was to be used successfully as detector. Literature suggested that neither calf thymus DNA polymerase Beta nor E. Coli DNA Polymerase I (Klenow Fragment) were able to extend past 8-oxoG base (Duarte, 1999). Some literature also suggests that the base incorporated opposite to the 8-oxo-G can be the natural Cytosine or a misincorporation of Adenosine (Kornyushyna, 2002). My first experiments tested whether the permissive Taq polymerase used in PCR amplification is able to overcome these limitations and extend past the 8-oxoG site.

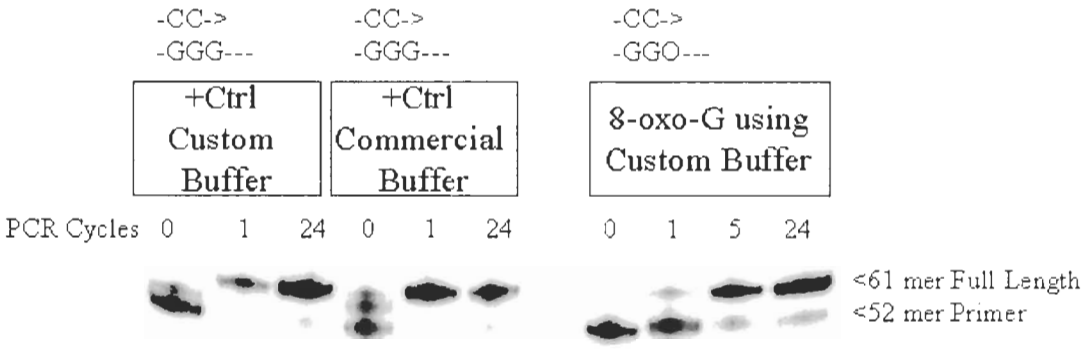
5.4.1.1 PCR Extension past 8-oxoG experiments

Taq PCR Extension past 8-oxoG experiments were carried out using a double stranded DNA construct made of two oligos, pETSCC(52mer) and pETSGGO(26mer). There are two 3' overhangs as well as a 17 base-pairing region. Primer ETSGGO contains an 8-oxo-G base (denoted by O in the sequence). Primer pETSCC was 5'end-labelled with P^{32} to monitor the incorporation extension of bases past 8-oxo-G.

Figure 5.2: pETSCC and pETSGGO together formed the double stranded DNA construct used to test if Taq Polymerase is able to extend past 8-OxoG site (denoted by symbol O). 5' end-label is shown with "*".

```
pETSCC (First primer with P32 5' end-label)
 5*AAATC ACTCT GCACA [TGACT GATCA GTTCA TCAAG] CTTGA AGTGT GTAAG CC ->          3'
 3'                                          <- GAACT TCACA CATTC GGOAC TCATC G 5'
pETSGGO (Second primer with 8-oxo-dG)
```

Figure 5.3: 12% polyacrylamide native gel result of the Taq PCR extension past 8-oxo-G experiment. Samples of PCR rounds 0 (initial, No PCR), 1 and 24 were taken under two buffer conditions. To the left and middle is the positive control using regular guanines. Also present in the middle is a 10bp ladder for reference. To the left is 8-oxoG series.



Shown in Fig 2.4, the experiment consists of two controls using different PCR buffers and one 8-oxo-G. The primers used in this experiment overlaps so it can only primer extend. However, since the actual selection would use PCR amplification, PCR amplification protocols was used in this experiment. The un-extended radio-labeled pETSCC was 52mer while the fully extended product was 61mer. Commercial Buffer provided by the QIAGen PCR Kit. The Custom Buffer uses standard Maniatis protocol for the most part modified only with the addition of 10μM β-Mercaptoethanol in the 10x Buffer to was used to reduce background oxidative damage to 8-oxoG (communications with Peter Unrau). Results showed that constructs with GGG as detector primer extended as expected. Taq polymerase was also able to extend past 8-oxoG using custom buffer, but longer extension times may be required.

5.4.1.2 GGG and GG-8-oxoG Irradiation Experiments

Having successfully tested 8-oxoG's ability to PCR amplify, the next step was a test irradiation of a construct similar to what the selection construct would look like. The two main purposes of this experiment are to test whether the 8-oxoG make a good detector and to test the constant flanking regions' ability to transport charge efficiently from the AQ stem to the detector stem.

The three constructs used (Fig. 5.4) were identical except in the detector. One strand contained a covalently attached 5'AQ, the other contained the detector and the P^{32} 5'end-label. The main charge conduction theories are in agreement that closely spaced guanines facilitate charge transport. In the ETS series, guanines were dispersed throughout the construct no more than two bases apart. The bracketed N20 Region is a representation of a typical random region.

Figure 5.4: Three constructs used to test the viability of 8-oxoG. GGG Construct is the positive control using conventional GGG detector. GGO 8-oxoG is the 8-oxoG construct.

GGG Construct

```
ETSAQCCC
AQ-AAATC ACTCT GCACA [TGACT GATCA GTTCA TCAAG] CTTGA AGTGT GTAAG CCCTG AGTAG C 3'
3' TTTAG TGAGA CGTGT [ACTGA CTAGT CAAGT AGTTC] GAACT TCACA CATTC GGGAC TCATC G 5'
ETSGGG                Simulated N20 Region                    Detector
```

GGO (8-oxo-G) Construct

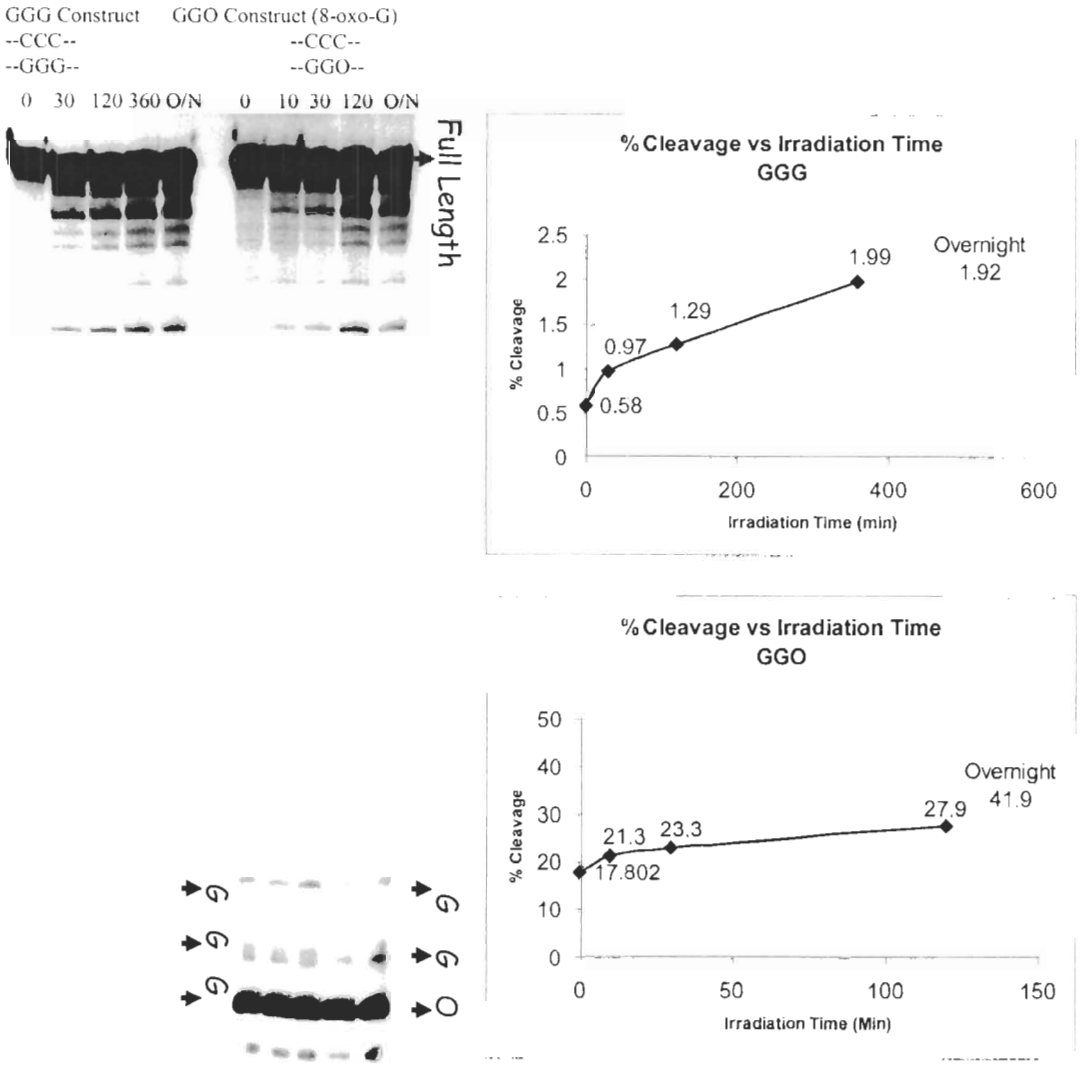
```
ETSAQCCC
AQ-AAATC ACTCT GCACA [TGACT GATCA GTTCA TCAAG] CTTGA AGTGT GTAAG CCCTG AGTAG C 3'
3' TTTAG TGAGA CGTGT [ACTGA CTAGT CAAGT AGTTC] GAACT TCACA CATTC GGOAC TCATC G 5'
ETSGGO                Simulated N20 Region                    Detector
```

The oligonucleotides were pre-treated with piperidine 30min at 90°C to minimize background cleavage of 8-oxo-G due to non-specific damage (that occurred perhaps during synthesis or DNA purification steps). Irradiation time-courses were carried out for both constructs, GGG and GGO according to protocol listed in Material and Methods.

GGG irradiation time was 0min (dark), 30min, 120min, 360min and overnight. GGO irradiation time was 0min (dark), 10min, 30min 120min and overnight. Post-irradiated samples were treated with piperidine 30min at 90°C and ran on 12% denaturing polyacrylamide gel. Post-irradiated piperidine treatment optimization (not shown) showed little effect in increasing signal to noise ratio so the standard 30min at 90°C was used.

Fig 5.5 left shows the raw densitometry results. The top band is the full length (60bp) band, while the bottom bands shows the location of the detector (GGG or GGO). Total distance traveled by the radical cation to reach the detector is 50bp. Damage at the 5' most guanine or 8-oxoG detector is expressed quantitatively in Fig 5.5 (Right Up, Right Down) as a percentage of total signal of each lane, called % cleavage (See 2.3.2 Typhoon and Data Analysis).

Figure 5.5: **8-oxoG vs GGG experiments. After being irradiated (time-course in min, 0 (dark), 30, 120, 360, Overnight), piperidine treated to induce strand cleavage, the lyophilized results were loaded and ran on a 12% denaturing PAGE and shown on the left. Densitometry data of G10, G9, G8 detectors is tabulated on the right. Note the high background cleavage % of 8-oxoG even without irradiation.**



Like the preliminary experiments, the GGG constructs still show very low cleavage even after overnight irradiation (1.29% at 120min). The background cleavage for GGG was low as well (0.58% at 0min). The low signal was expected because the distance traveled by the radical cation (50bp) was nearing the observed limits of charge conduction (200Å or 66bp). GGO results, although shows considerable cleavage (27.9% at 120min), it also showed a considerable background cleavage (17% at 0min). This means that a great proportion of 8-oxoG-containing oligonucleotides were damaged independent of irradiation. Ultimately, 8-oxoG containing oligonucleotides expressed a less than two-fold increase in signal compared to background, and was therefore deemed unsuitable to serve as a detector for the selection. GGG on the other hand expressed an above a two-fold signal to noise ratio.

8-oxoG has been shown by Schuster's group as (Liu & Schuster, 2003) a good potential detector stem, owing to its low oxidation potential. However in our experience 8-oxoG was difficult to use for the purposes of selection because of high levels of background cleavage. Therefore, we focused on finding suitable flanking sequences to increase the GG signal relative to background, since this seemed a more rewarding route.

5.4.2 Detector Evaluation, Thymine Dimer Series

It has been known for a few years that excess electron transport reverses thymine dimers. This phenomenon was recently investigated by Behrens et al (Behrens et al. 2003) using a flavin cap functioning as an electron donor, and thymine dimer hairpin as electron acceptor. When thymine dimers photoreverse, the hairpin is cleaved and the fragments can be detected with an HPLC.

Although there has been some controversy regarding hole transport's ability to photoreverse thymine dimers (Dandliker et al. 1998, Ramaiah D et al. 1998), we decided to explore the possibility of using it as a detector for hole transport. In the TD-Series constructs, the cyclobutane thymine dimers were not linked by the phosphodiester backbone, only through the covalent linkage between the dimerized thymines. If charge injection by AQ into the double stranded construct causes the dimers to photoreverse, then a breakage associated with this repair would be seen on denaturing polyacrylamide gel. Fig 5.6 shows the construct in more detail.

Figure 5.6: Thymine Dimer series (TD-Series). Note the T=T represents Thymine dimers linked only by the dimer region, with no phosophodiester backbone. "*" represents P^{32} 5' end-labeled strand.

```
AQ~tttac GTTCC TTcGT CA-aCG CACAT TCCTT TAtCT 3'
3' aaatg CAAGG AAgCA GT=TGC GTGTA AGGAA ATaTA 5*'
  Proximal GG    Thymine Dimer     Distal GG
```

Figure 5.7: **The diagram on the right depicts cyclobutane thymine dimers with no phosphate backbone linkage. Thymine Dimer series (TD-Series) irradiation results on 12% denaturing polyacrylamide gel is shown on right side. Lanes from left to right is: Direct photoreversal at 260nm, 240min irradiation and 45min irradiation (+control). The ds No AQ represents construct with no tethered AQ (negative control), the ss-T=T represented single stranded thymine dimer strand (negative control). The four lanes to the ight is irradiation time-course carried out at 334nm. Note that no significant photoreversal of thymine dimers were seen even after 240min of irradiation. Both proximal and distal GG detectors detected hole transport as usual.**

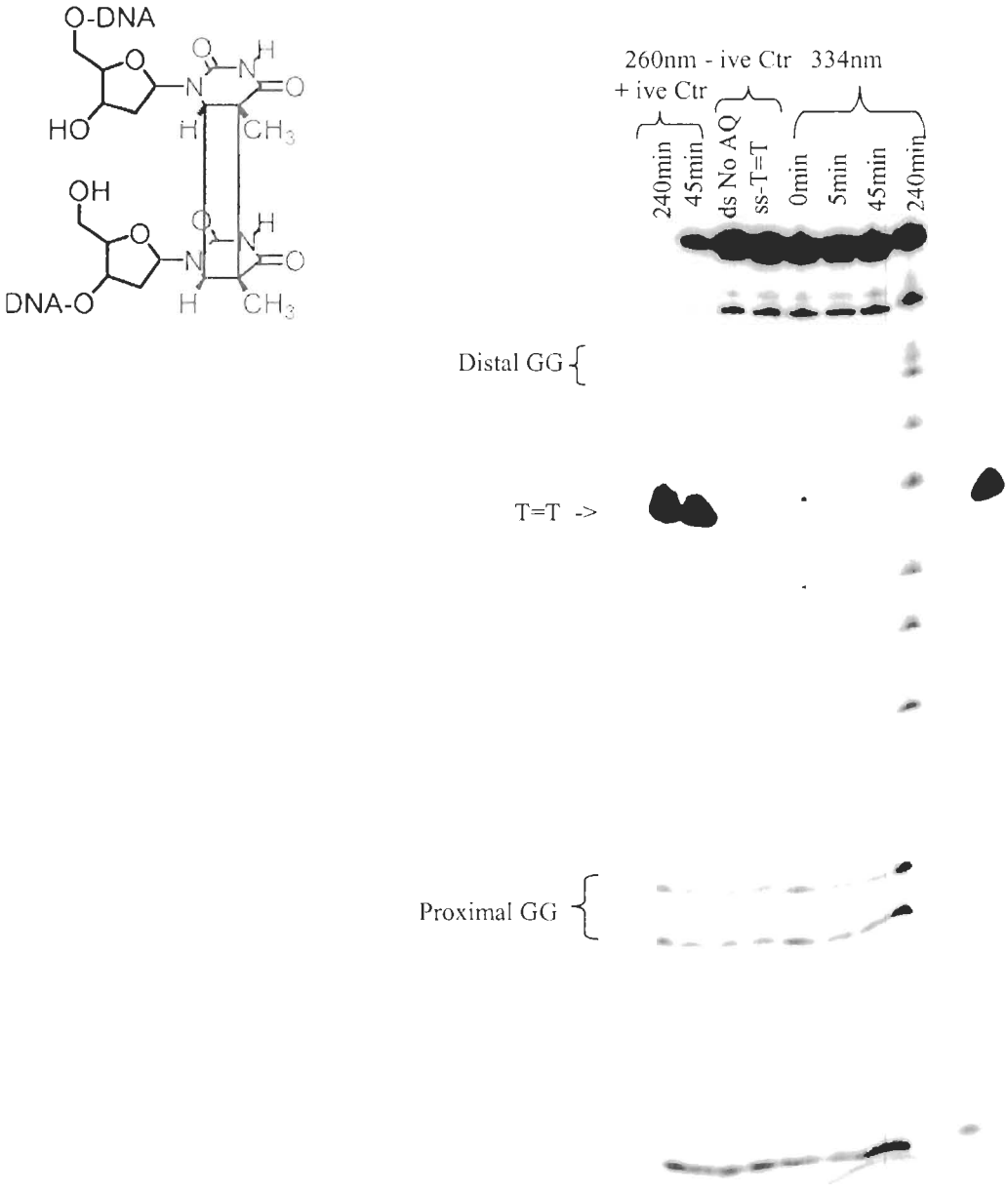


Fig. 5.7 showed the photoirradiated and piperidine treated TD constructs run on a 12% denaturing polyacrylamide gel. The positive controls have been irradiated with 260nm light, known to directly photoreverse dimers. The negative control shows constructs with no AQ and single stranded constructs (ss-T=T). These have not been irradiated. A time-course using 334nm light was used to determine the effectiveness of AQ in photoreversal of the thymine dimers.

Result showed that direct photoreversal at 260nm worked as expected, with the dimers photoreversing almost completely after 45min of irradiation. The proximal and distal GG present within the construct detected the usual damage associated with charge injection by AQ, meaning that charge transport was occurring. However, the excitation of AQ using 334nm light did not seem to photoreverse thymine dimers appreciably. In summary, although the idea is intriguing, thymine dimers were found not to be photoreversed by single-electron oxidation with AQ and so it cannot be used for monitoring charge transport in the selection.

5.4.3 SX Series, Experimental Details

The SX Series *in vitro* selection scheme is illustrated in Fig 5.8. First, the SX selection library oligonucleotides, containing multiple copies of all the possible combinations in the N10 region, (shown in Fig 5.9 Top, the bottom strand) is synthesized and purified. The library strand also contained the GGG detector as well as a 5' radiolabel site. The boxed region is the primer regions to be used for PCR amplification. Irradiation time-courses (not shown) was carried out using the positive control shown in Fig. 5.9 Bottom, and found to give adequate signal at the detector. In Fig. 5.8. (1) The selection pool is then PCR amplified further in the initial round, then exposed to UV light

at 334nm. This causes the excitation of AQ and the subsequent charge injection and charge transport through the N10 region of the SX construct until it is quenched by the GGG detector. (2) A subset of the damaged GGG are prone to piperidine treatment, causing strand cleavage. (3) Members of the selection population able to successfully transport charge to the GGG detector are then isolated by 12% denaturing polyacrylamide gel and (4) selectively amplified for the next round of selection. In order to pass the screening test (step 3), the N10 region must not be too insulating (causing no charge transport to reach GGG), nor too reactive (causing oxidative damage in the N10 region, also not reaching GGG).

Figure 5.8: SX Series construct representing the *in vitro* selection (SELEX) schematic.

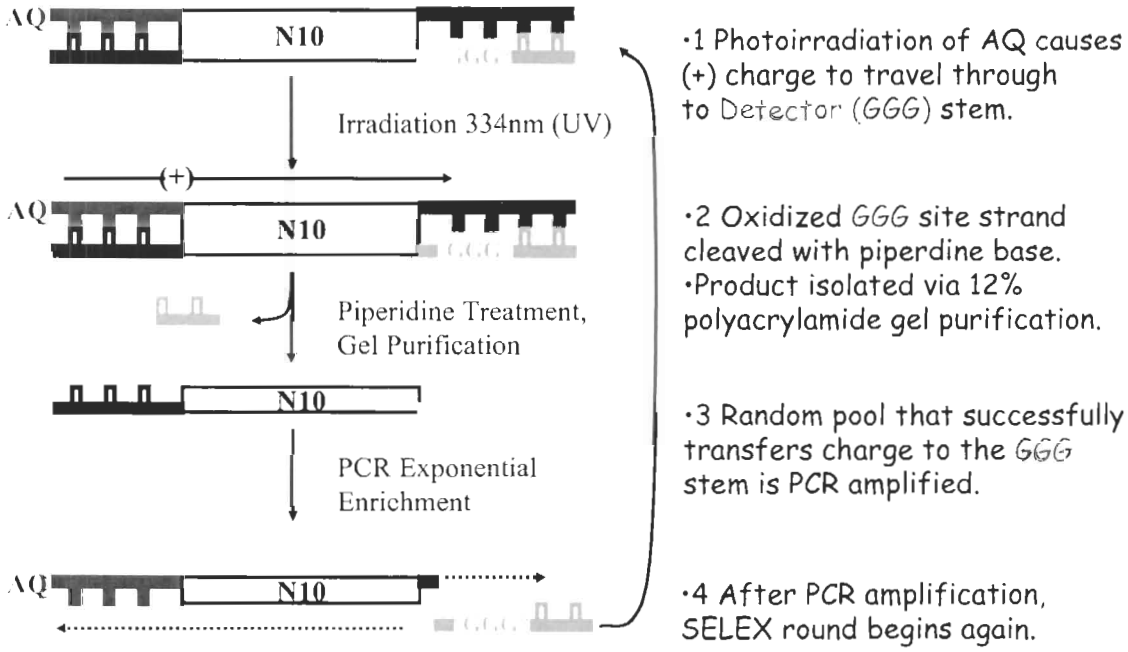


Figure 5.9: SX Selection constructs. At the top is the SX selection construct, features: An N10 random region is in the middle flanked by two primer binding sites. Primer 1 [boxed] contains covalently attached AQ; primer 2 [boxed] contains GGG [bolded] detector. The bottom construct represents the synthetically produced full length oligonucleotides that would serve as the positive control.

Primer 1 (p1AQ-15)

AQ~TTTAG CTCAC tcact XXXXX XXXXX agtgc GAGAC GCTga CCCAT AGTGA 3′
3′ AAATC GAGTG agtga NNNNN NNNNN *tcacg CTCTG CGAct* ***GGG****TA TCACT**5′
N10 Region Primer 2 (p2-25)

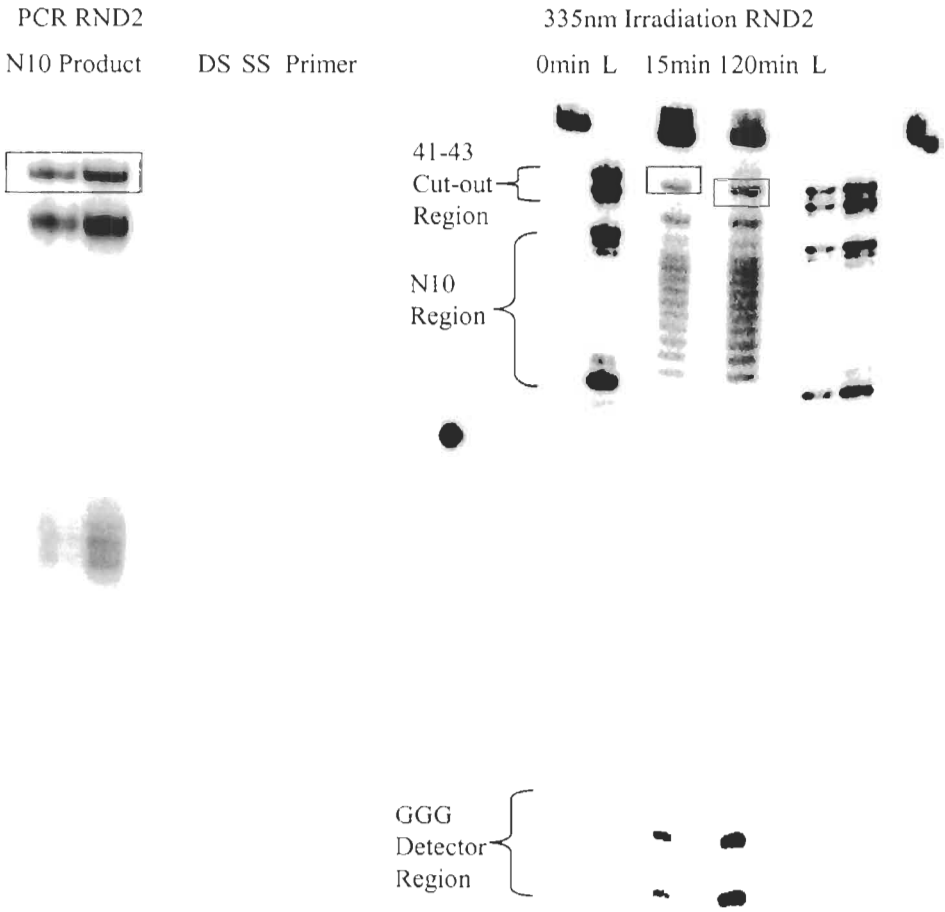
Positive Control

AQ~TTTAG CTCAC tcact *cgagt gagcg* agtgc GAGAC GCTga CCCAT AGTGA 3′
3′ AAATC GAGTG aGtga *gctca ctcgc* tcacg *CTCTG CGAct* ***GGG****TA TCACT**5′

After some initial PCR problems, eventually attributed to one of the primers, the PCR worked properly and the selection experiments started. Rounds 1 through 9 was carried out in two parallel selections. One selection was irradiated at 334nm for 15min, while the other was irradiated for 120min. The rationale behind selecting a 15min pool and a 120min pool was to test if irradiation durations have an effect on types of sequences selected for.

Data from a typical round (Round 2) is shown in Fig. 5.10. On the right hand side, 12% native polyacrylamide gel was used to resolve the products of PCR. Using synthetic full length double strand DNA as a size marker, the correct PCR band corresponding to full length double stranded N10 products (boxed) were isolated and eluted from excess primers and other products. One unexpected product appeared; a band running slightly faster than the double stranded N10 product. This mystery band was not merely a single strand product since it doesn't run the same as single stranded size controls, but is probably caused by secondary structures. And since sufficient amplification was achieved for selection purposes, the band was noted for future reference.

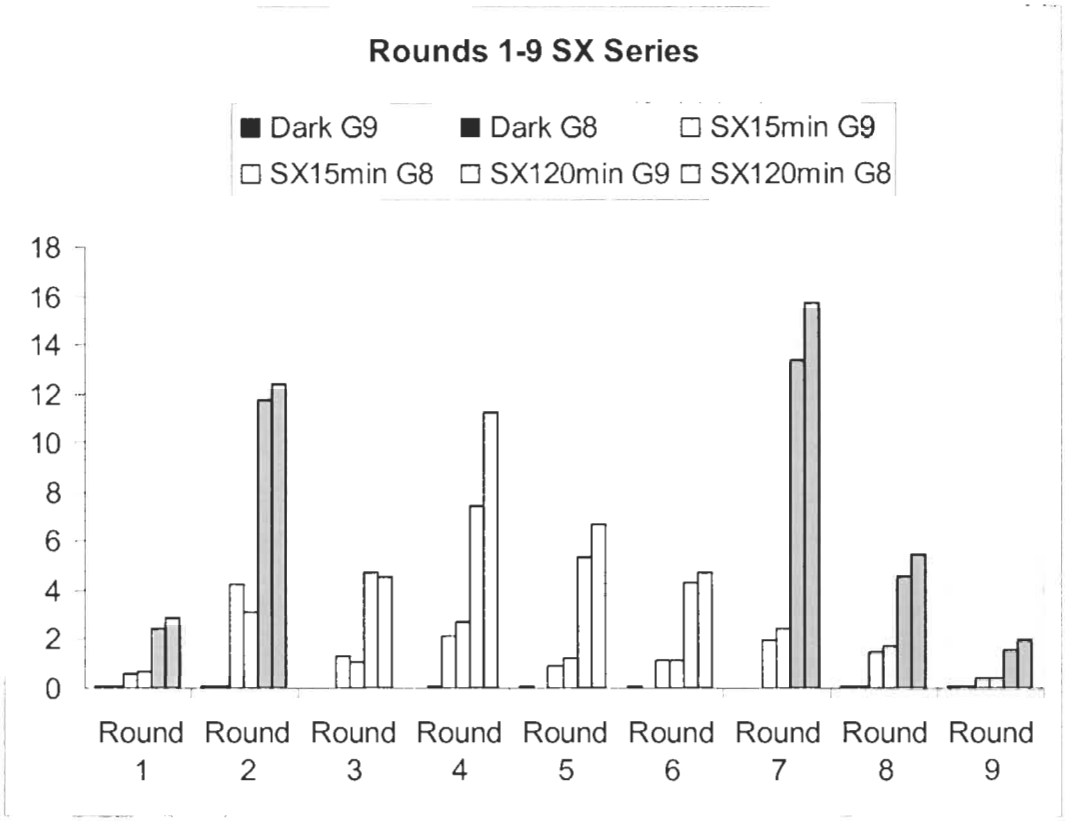
Figure 5.10: **SX *in vitro* selection Round 2. Results: Left is a 12% native gel used to purify PCR products (note the "unknown band" below the primary double stranded band). Lanes (right to left) includes P2 primer, single stranded size control, double stranded size control as well as the round 2 selection PCR. Top band was the double stranded product to be cut and eluted. After the selection construct was purified by native PAGE, irradiated, piperidine treated to induce strand cleavage, the lyophilized results were loaded and ran on a 12% denaturing PAGE and shown on the right. Lanes from left to right is 0 min irradiation (dark) controls, Size ladders to represent size of successfully cleaved candidate sequences, 15min irradiation pool, 120min irradiation pool, and another ladder lane. Boxed area represents successfully cleaved candidate sequence to be cut, eluted, and amplified for subsequent selection rounds.**



After the PCR amplified double stranded N10 were isolated from the native gel, it was eluted and irradiated at 334nm for 15min or 120min, depending on the selection pool, and treated with piperidine. The denaturing gel on the right side of Fig. 5.10 showed the post-irradiation, post-piperidine treatment samples. The 41-43 cutout region represents the size of the fragment that successfully passed the selection test. Since strand cleavage also cleaves off the 5' radioactive end-labeling, the selection fragment is cold, and size markers (L) were needed to visualize it.

Fig. 5.11 showed the densitometry analysis monitors the GGG detector over a period of nine rounds (expressed in % cleavage compared with the overall signal). If % cleavage of GGG detector increases as the rounds progress, then this signal increase may represent the selective amplification of "good" charge transport sequences. In general, the 15min irradiation pool exhibited less damage than the 120min pool, an understandable outcome considering the damage is proportional to the irradiation time. The signal peaked during round 7, but with the large amount of fluctuation seen between the rounds, it is difficult to access whether the selection succeeded or not. Because the number of unique sequences in this selection was small (N=10 or $1x10^6$ different possible combinations) and the copy number very high (~$2x10^7$ copies of every unique sequence in an average round), one would expect that effective conductors would be amplified in relatively few rounds. Unfortunately, after 9 rounds, there has not been a significant increase in signal. Instead, round 9 has even less signal then starting sequence. Round 9 products were cloned but not sequenced.

Figure 5.11: SX Series Round 1 -9. Densitometry data expressed as % cleavage at G9-G8 site of GGG detector. Bar graph is listed in the same order as legend (i.e. first bar from the left represents Dark G8 % cleavage from round 1, second bar from the left represents Dark G9 % cleavage from round 1) White represents 15min irradiation pool. Grey represents 120min irradiation pool. Black represents dark (round 3 and 7 dark unavailable).



Two main shortcomings perceived in the SX series was the short lengths of the primer regions (15bp only) and the short length of the random region (10bp only). Primers generally are 17 nucleotides or longer to allow for proper annealing. 10bp random region was also not ideal, considering the evidence suggesting that charge transport polarons may be up to 6bp in length. Originally the design was constrained by

the concern that a longer construct would not conduct charges efficiently to the detector. But after evaluating the experimental evidence detailing high 5% range cleavage in the 15min pool, a longer construct was justified. The longer SY selection was designed to overcome these shortcomings.

5.4.4 SY Series, Additional Experimental Details.

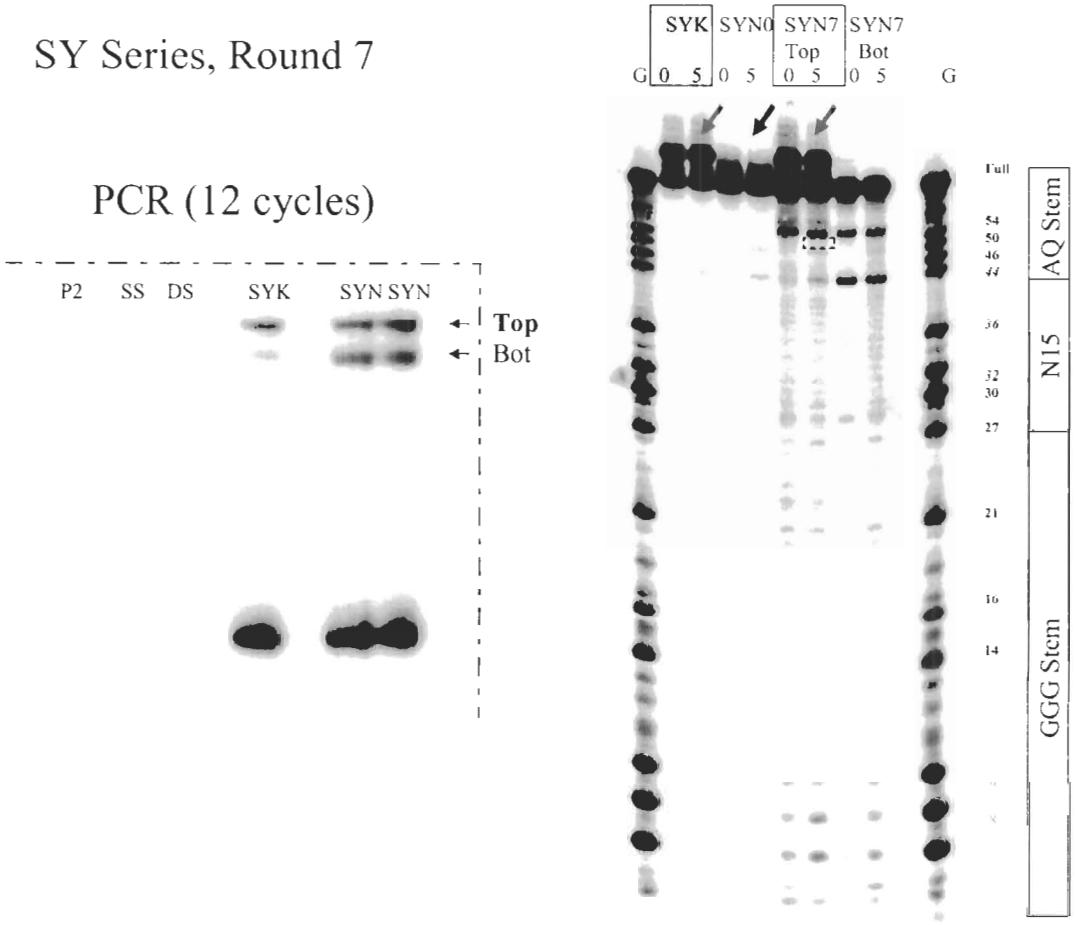
5.4.4.1 SY Selection Round 1

See Section 3.3.1.

5.4.4.2 SY Selection Round 7

Fig. 5.14, left showed the round 7 PCR results. PCR amplification still produced the regular selection product (called SYN7 Top), which had the same mobility as the double stranded size control, and an unknown product (called SYN7 Bot or "Bottom"). SYN7 Bot product ran faster than ds and was probably a misfolded or even single stranded species. In the next section, DEPC probing were carried out to figure out the identities of the PCR bands in Section 3.3.1. We were curious, so SYN7 Bot was irradiated and treated with piperidine just like the selection construct SYN7 Top.

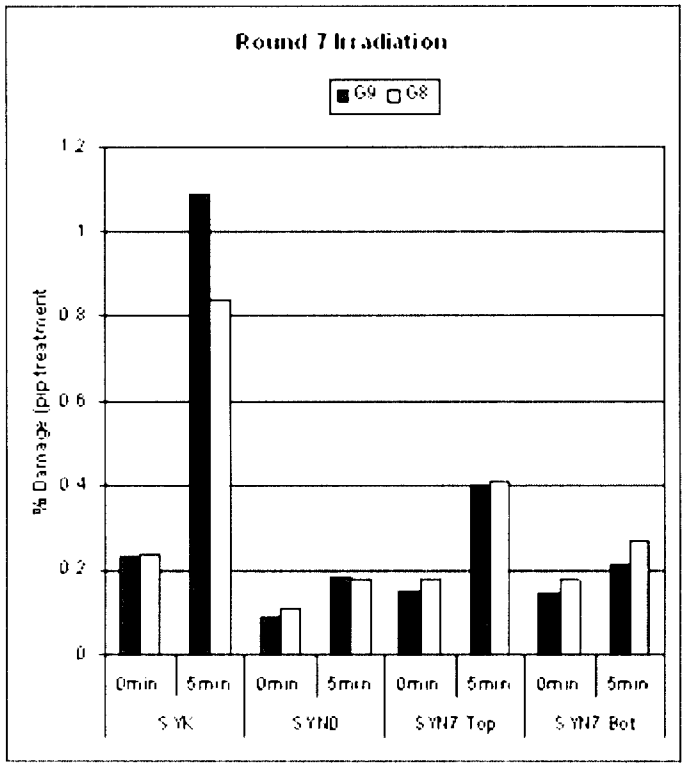
Figure 5.12: SY *in vitro* selection, Round 7 Results: Left is a 12% native gel used to purify PCR products. Lanes (left to right) P2 primer, single stranded size control, double stranded size control, SYK known sequence, SYN selection sequence. Top band was the double stranded product to be cut and eluted. After the selection construct was purified by native PAGE, irradiated, piperidine treated to induce strand cleavage, the lyophilized results were loaded and ran on a 12% denaturing PAGE and shown on the right. Lanes from left to right is: G-ladder, SYK 0min irradiation (dark), SYK 5min, SYN0 initial round control 0min, SYN0 5min, SYN7 current selection Top band from PCR 0min, SYN7 5min, SYN7 Bot band from PCR 0min, SYB7 Bot 5min, G-ladder.



On Fig. 5.14, left and Fig 5.15 showed the 12% denaturing gel results and the densitometry data. Each construct were either not irradiated (labeled 0min) or irradiated for 5 min. SYK was similar to round 1 complete with the unknown "slow SYK band". Densitometry data showed the signal at the GGG detector to be ~1% cleavage, quite low compared with round 1. SYN0 is a sample of the initial library for charge conducting comparisons. Densitometry data showed ~0.2% also a lower signal than in round 1. The round 7 selection pool was labeled SYN7 Top. Densitometry data showed it cleaved slightly better than SYN0 at ~0.4%. The unknown PCR product was labeled SYN7 Bot. It also had low cleavage %. An important observation is the fact that SYN7 Top acquired an "unknown slow band" shown by the arrows in Fig. 5.14, left similar to the one on SYK. We thought it may be some kind of double stranded construct that survived the "denaturing" conditions. This was tested in section 2.3.5.4.

By the end of round 7, there were two unknown bands to test. One band was called "Bottom or SYN7 Bot" arising from PCR. The other was seen in 12% denaturing experiments in both SYK and SYN7 constructs; we dubbed it "Slow SYK band".

Figure 5.13: Densitometry data for Round 7 expressed as % cleavage. SYK is positive control, SYN0 is initial library, SYN7 is round 7 selection. SYN7 Top was the PCR



		G9	G8
SYK	0min	0.238	0.238
	5min	1.09	0.841
SYN0	0min	0.09	0.108
	5min	0.184	0.183
SYN7 Top	0min	0.151	0.182
	5min	0.395	0.412
SYN7 Bot	0min	0.147	0.177
	5min	0.211	0.269

5.4.4.3 Solving the SYN7 Bot Mystery, DEPC Probing Experiments

DEPC experiments, using as the active reagent Diethyl Pyrocarbonate, was used as a way of testing for single stranded DNA. DEPC preferentially causes the carboxyethlation of adenosine and to a lesser extent guanine, and cause cleavage of strand when subjected to piperidine treatment. More importantly, it *preferentially damages single stranded* but not double stranded DNA, a useful test for probing to what extent an unknown bands are single stranded.

Figure 5.14: **DEPC Experiment Results. Left is a 12% native PAGE used to purify PCR products. Lanes (left to right) single stranded size control, double stranded size control, SYK known sequence. Various sequences were treated with 1% DEPC at room temperature followed by piperidine-mediated strand cleavage. Samples were run on a 12% denaturing PAGE and shown on the right side. Lanes (from left to right) is as follows: G-ladder, 2x single stranded control (expect to see cleavage at G and T sites), PCR product Top band (expected double stranded product), PCR product Bot band, and 2x double-stranded control. Note that Top lane did not show much DEPC cleavage compared with single strand and Bot lane. Top lane did exhibit a band running at single stranded mobility, this is due to the denaturing conditions of the gel. Double stranded controls shows that double stranded DNA can exist in denaturing gel.**

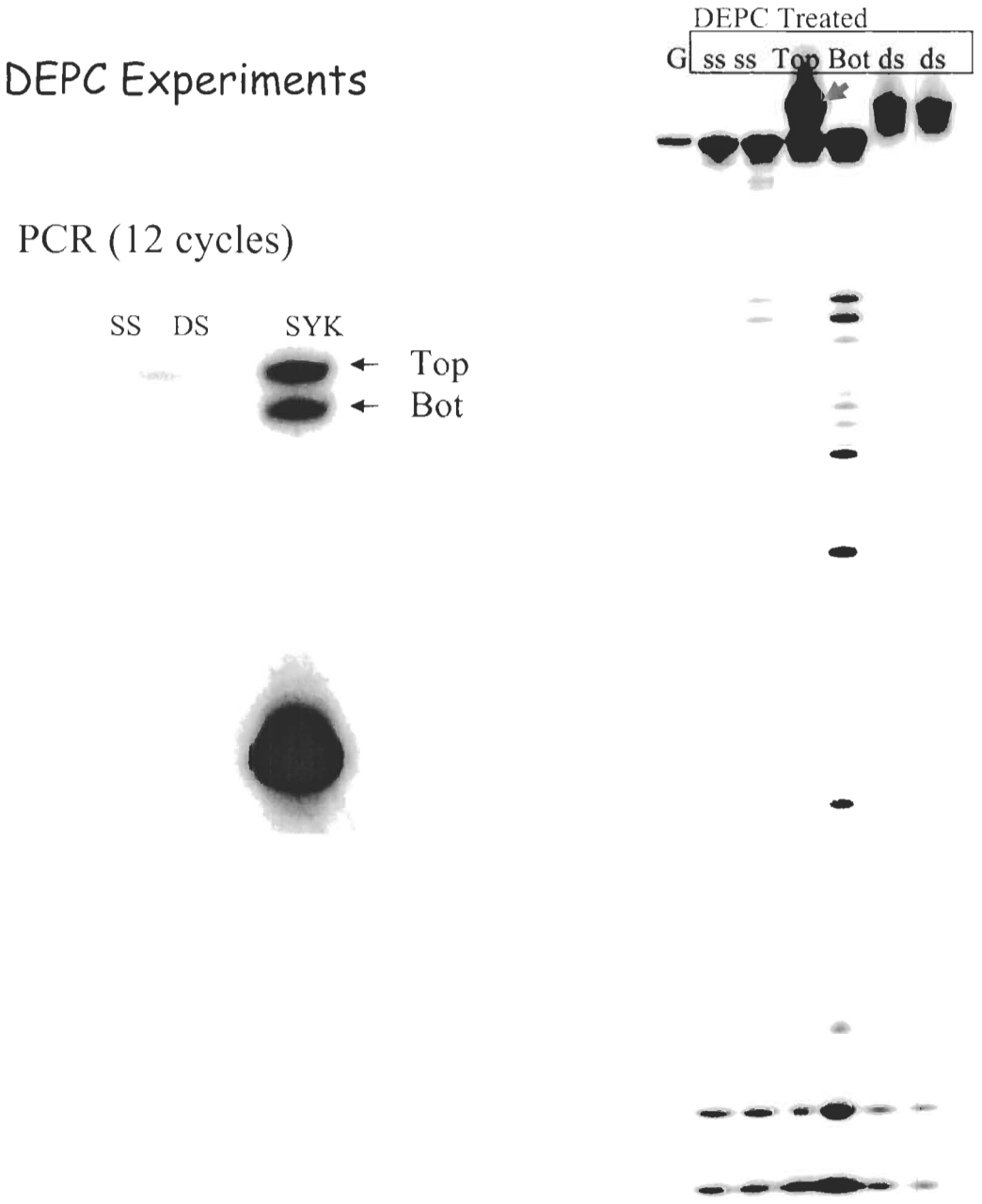


Fig. 5.16, right showed the 12 rounds of PCR in 12% native gel, where the Top and Bottom products were taken out and subjected to DEPC experiment (using SYK as template). Top was the correct double stranded product while Bottom was unknown. Top and Bottom, in addition to ss (single stranded) and ds (double stranded) controls were subjected to 1% DEPC for 30min at 21°C. The samples were then subjected to standard piperidine treatment (30min at 90°C).

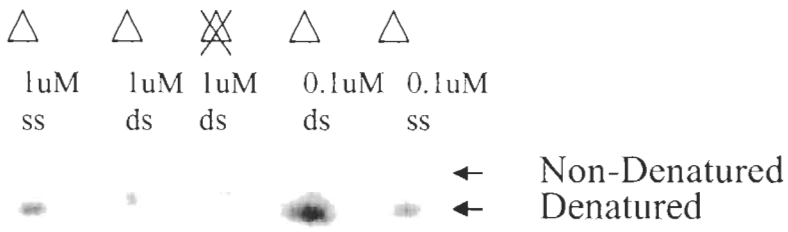
The results were run on a 12% denaturing polyacrylamide gel on Fig. 5.16, left. Going from left to right, was: G-ladder, ss or synthetic single stranded full length, the Top and Bottom PCR products, and ds or double stranded synthetic construct. Double stranded control (ds) offered little reaction when exposed to DEPC, as expected. Single stranded control (ss) on the other hand showed clear bands of cleavage at every A and G site. Much like ds controls, top band showed little reaction with DEPC, consistent our assignment of Top band as the double stranded product of PCR. Bottom band showed significant reaction to DEPC, consistent with being single stranded. Based on this evidence, Bottom unknown band is probably a single stranded oligonucleotide. The presence of the bottom band does not interfere with the selection.

5.4.4.4 Solving the "Slow" SYK band Mystery, Denaturation experiments.

As mentioned before, SYK consistently showed a band that was slower in mobility than full length on the 12% denatured polyacrylamide gel (G-ladder represented full length). Dubbed "slow SYK band" since it was first noticed on SYK denaturing gel in round 1. We initially hypothesized if salt wasn't properly removed, double stranded constructs may not be denatured even under the denaturing conditions of 9M Urea in denaturing polyacrylamide. The constructs also contained a lot of guanines, which would

increase the melting temperature and make it harder to dissociate. This would lead to a band slightly slower in mobility than full length single stranded band. Using OligoAnalyzer 3.0 from IDTDNA, http://scitools.idtdna.com/scitools/Applications/OligoAnalyzer/default.aspx the SYK oligo with a concentration of 0.25uM and 50mM salt has a predicted melting temperature (Tm) of more than 70°C, and much approaching 85°C when salt concentration reaches 300mM (Salt should only be a big issue if there are ethanol precipitation problems).

Figure 5.15: SYK denaturation experiment in 12% denaturing polyacrylamide gel. Triangle symbol represents heating the sample to 90°C before loading on gel. Crossed triangle means no pre-heating before loading. SS means single stranded SYK, DS means double stranded SYK. Samples also listed based on concentration.



In the denaturation experiment shown in Fig. 5.17, SYK construct, made of two complementary synthesized 59mer oligonucleotides, were produced in two concentrations, 1μM or 0.1 μM. These constructs were pre-heated in boiling water, then added in denaturing loading buffer and loaded into the denaturing gel. Single stranded

controls were also included for size reference. Double stranded controls were constructs at 1μM concentration added directly to the loading buffer and loaded on gel without any heating.

Results show that at 1μM concentrations, pure single stranded control (ss) had different mobility than the non-heated double stranded control (no heat ds). This showed that double stranded DNA could still exist in 9M Urea conditions of the denaturing gel. Heated 1μM ds showed two bands, one had the same mobility as single stranded control, another had the same mobility as the double stranded control. This result is consistent with previously seen "slow SYK band". A lower concentration of each complimentary DNA strand to (0.1μM) seems to promote disassociation, which made sense because duplex DNA dissociates easier at lower concentrations.

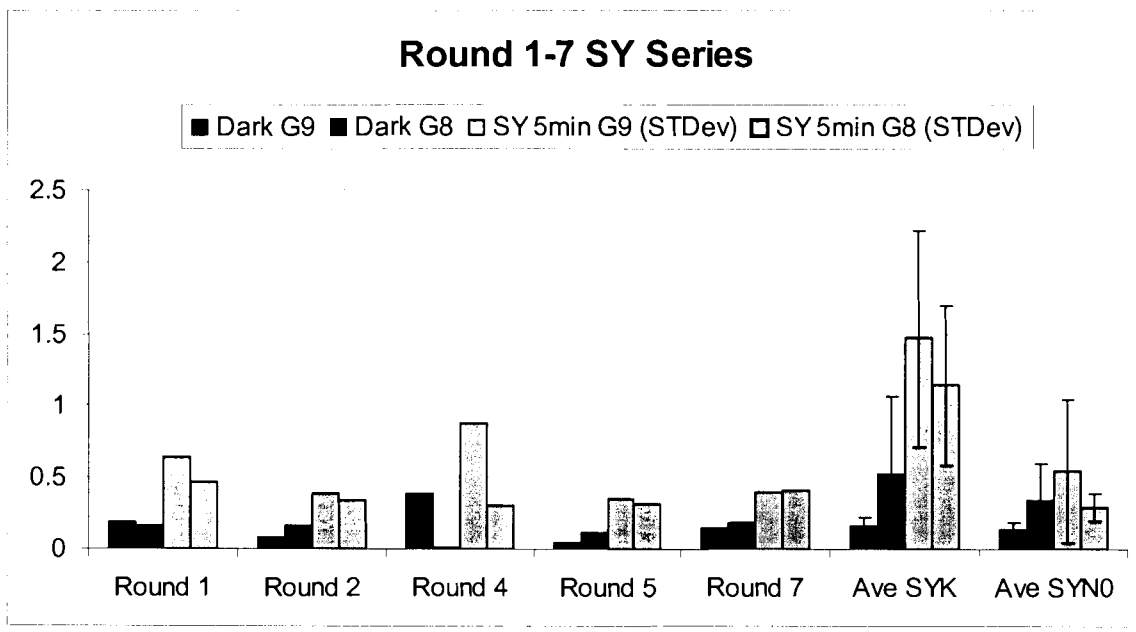
The denaturation experiment showed that the "slow SYK band" anomaly was probably nothing more than double stranded DNA that failed to denature properly. Heating with boiling water alleviated this problem. While this decreased the yield of harvested template for selection, it has little effect on the validity of the selection itself.

5.4.4.5 SY Selection, 7 Rounds of Selection Summary.

7 Rounds of selection were completed. The round by round results was tabulated in Fig. 5.18 as cleavage % of the two 5'most guanines in the GGG detector. Based on average SYK positive control, and average SYN0 initial pool controls, it appeared that rounds 1-7 exhibited little change in terms of increased damage to the detector. Rounds 3 and 6 were excluded from the data analysis because the guanine detector was not properly resolved on the 12% denaturing polyacrylamide gel. The unknown PCR band was determined to be single stranded construct by DEPC experiments. The unknown

"slow SYK band" on the denaturing gel was determined to be undenatured double stranded constructs. Both should have minimal effect on the selection process. However, it was interesting to note that by round 7, the selection pool seemed to exhibit the same "slow SYK band" seen on the positive controls, perhaps signifying a homogenization of the selection population. This was one of the reasons why by round 7 the selection was cloned and sequenced.

Figure 5.16: SY Series Round 1 -7 Summary expressed as % cleavage at G9-G8 of GGG detector. Black represents the dark reaction, grey represents 5min irradiation. Rounds 1,2,4,5,7 are included. SYK is average of all positive control values, STDev error bars. SYN0 is average of all initial pool values, STDev error bars.



REFERENCE LIST

Anderson, R. & Wright, G. (1999) Energetics and rate of electron transfer in DNA from base radical anions to electron-affinic intercalators in aqueous solution. *Phys. Chem. Chem. Phys.* **1**, 4827-31.

Barton, J.K., Wan, C., Fiebig, T., Kelley, S.O., Treadway, C.R. & Zewail, A.H. (2000) Femtosecond dynamics of DNA-mediated electron transfer *Proc. Natl. Acad. Sci.* **97**, 12759.

Barnett RN, Cleveland CL, Joy A, Landman U, Schuster GB. Charge migration in DNA: ion-gated transport. (2001) *Science*, **294**(5542):567-71.

Behrens C, Carell T. (2003) Excess electron transfer in flavin-capped, thymine dimer-containing DNA hairpins. *Chem Commun (Camb)*, **21**(14), 1632-3.

Besaratinia A, Synold TW, Xi B, Pfeifer GP. (2003) G-to-T transversions and small tandem base deletions are the hallmark of mutations induced by ultraviolet a radiation in mammalian cells. *Biochemistry*, **43**(25), 8169-77.

Bixon, M., Giese, B., Wessely, S., Langenbacher, T., Michel-Beyerle, M.E., & Jortner, J. (1999) Long-range charge hopping in DNA. *Proc. Natl. Acad. Sci.* **96**, 11713-16.

Borer, P.N., Pante, S.R., Kumar, A., Zantta, N., Martin, A., Hakkinen, A., & Levy, G.C. (1994) C-13-NMR relaxation in 3 DNA oligonucleotide duplexes – Model-free analysis of internal and overall motion. *Biochemistry* **33**, 2441-2450.

Braun, E. Eichen, Y. Sivan, U. and Ben-Yoseph, G., (1998) *Nature* **391**, 775

Burrows, C., Muller, J.G., (1998) Oxidative Nucleobase Modifications Leading to Strand Scission. *Chem Rev.* **98**(3), 1135-1152.

Conwell, E.M. & Bloch, S.M. (2006) Base Sequence Effects on Transport in DNA. *J. Phys. Chem. B.* **110**, 5801-5806

Debije, M.G., Milano, M.T. & Bernhard, W.A. (1999) DNA responds to ionizing radiation as an insulator, not as a 'molecular wire'. *Angew. Chem. Int. Ed.* **38**, 2752-2756.

Dohno, C.; Ogawa, A.; Nakatani, K.; Saito, I. (2003) Hole Trapping at N6-Cyclopropyldeoxyadenosine Suggests a Direct Contribution of Adenine Bases to Hole Transport through DNA. *J. Am. Chem. Soc.*, **125**, 10154

Duarte, V., Muller, J.G., & Burrows, C.J. (1999) Insertion of dGMP and dAMP during *in vitro* DNA synthesis opposite an oxidized form of 7,8-dihydro-8-oxoguanine. *Nucleic Acids Research* **27**(2), 496-502.

Eley, D.D. & Spivey, D.I. (1962) Nucleoic acid in the dry state. *Trans. Faraday Soc.* **58**, 411-415.

Ellington, A.D. & Szostak, J.W. (1990) *In Vitro* Selection of RNA Molecules that Bind Specific Ligands. *Nature*, **346**, 818-822.

Fahlman, R.P. (2001) DNA Nanotech: Expanding the repertoire of DNA for the assembly of nanoscale objects and electrical devices. Ph.D. Thesis.

Fahlman, R.P., Sharma, R.D., Sen, D. (2002) The charge conduction properties of DNA holliday junctions depend critically on the identity of the tethered photooxidant. *J. A. Chem. Soc.* **124**, 12477.

Fahlman, R. P. & Sen, D. (2002) DNA conformational switches as sensitive electronic sensors of analytes. *J. Am. Chem. Soc* **124**, 4610-4616.

Famulok, M.; Szostak, J. W., (1992) *In vitro* Selection of Specific Ligand Binding Nucleic Acids. *Angew. Chem.* **31**, 979-988.

Famulok, M.& Szostak, J. W., (1993) Selection of Functional RNA and DNA Molecules from Randomized Sequences. *Nucleic Acids and Molecular Biology* **7**, 271.

Friedberg, E.C., Walker, G.C., & Seide, W. (1995) DNA Repair and Mutagenesis, *ASM Press*.

Gasper, S.M. & Schuster, G.B. (1997) Intramolecular Photoinduced Electron Transfer to Anthraquinones Linked to Duplex DNA: The Effect of Gaps and Traps on Long-Range Radical Cation Migration. *J. Am. Chem. Soc.*, **119**, 12762-12771.

Giese, B. (2002) Long-distance electron transfer through DNA. *Annu. Rev. Biochem.* **71**, 51-70.

Görner, H. (2002) Photoreduction of 9,10-Anthraquinone Derivatives: Transient Spectroscopy and Effects of Alcohols and Amines on Reactivity in Solution. *Photochemistry and Photobiology* **77**(2), 171–179.

Herbert, A., Rich, A., (1999), Left-handed Z-DNA, structure and function. *Genetica* **106,** 37-47.

Hermon, Z., Caspi, S. & Ben-Jacob, E., (1998) *Europhys. Lett.* **43**, 482.

Hickerson, R. P., Burrows, C. (1999) Sequence and Stacking Dependence of 8-Oxoguanine Oxidation: Comparison of One-Electron vs Singlet Oxygen Mechanisms. *J. Am. Chem. Soc.*, **121** (40), 9423 -9428.

James, A.M., Cocheme, H.M. & Murphy, M.P. (2005), Mitochondria-targeted redox probes as tools in the study of oxidative damage and ageing. *Mechanisms of Ageing and Development,* **126**, 982–986.

Joyce, G.G. (1989). Amplification, Mutation and Selection of Catalytic RNA. *Gene*, **82**, 83-87.

Kawai, K., Takada, T., Tojo, S., & Majima, T., (2003) Kinetics of Weak Distance-Dependent Hole Transfer in DNA by Adenine-Hopping Mechanism, *J. Am. Chem. Soc.* **125**, 6842-6843.

Kanvah, S. & Schuster, G.B. (2005) The sacrificial role of easily oxidizable sites in the protection of DNA from damage. *Nucleic Acids Research*, **33**, 5133.

Klug, S. & Famulok, M., (1994) All you wanted to know about SELEX. *Mol. Biol. Reports* **20**, 97-107.

Kornyushyna, O.; Berges, A. M.; Muller, J. G.; Burrows, C. J. (2002) *In vitro* Nucleotide Misinsertion Opposite the Oxidized Guanosine Lesions Spiroiminodihydantoin and Guanidinohydantoin and DNA Synthesis Past the Lesions Using Escherichia coli DNA Polymerase I (Klenow Fragment). *Biochemistry* **41**, 15304-15314.

Krebs, F.C., Laursen, B.W., Johannsen, I. (1998) *Acta Crystallogr., Sect. B: Struct. Sci.* **55**, 410.

Lewis, F.D. Wu, T., Zhang, Y., Letsinger, R.L., Greenfield, S.R. & Wasielewski, M.r. (1997) Distance-dependent electron transfer in DNA hairpins. *Science* **277**, 673-676.

Liu, C.-S.; Hernandez, R.; Schuster, G. B. (2004) Mechanism for Radical Cation Transport in Duplex DNA Oligonucleotides. *J. Am. Chem. Soc.* **126**(9), 2877-2884.

Liu, C.S. & Schuster, G.B. (2003) Base Sequence Effects in Radical Cation Migration in Duplex DNA: Support for the Polaron-Like Hopping Model. *Am. Chem. Soc.*, **125** (20), 6098.

Maniatis, T., Sambrook, J., and Fritsch, E.F. (1989) Molecular Cloning: A Laboratory Manual. *Cold Spring Harbor Laboratory Press, NY,* **1, 2, 3**.

Muller, J. G.; Hickerson, R. P.; Perez, R. J.; Burrows, C. J. J. Am. Chem. Soc. 1997, 119, 1501-1506.

Murphy, C.J., Arkin, M.R., Jenkins, Y., Ghatlia, N.D., Bossmann, S.H., Turro, N.J. & Barton, J.K. (1993) Long-Range Photoinduced Electron Transfer Through a DNA Helix. *Science*, **262**, 1025.

Nakamura, T.; Sugiyama, H.; Fujisawa, K.; Dohno, C.; Nakatani, K.; Saito, I. Nucleic Acids Symp. Ser. 1996, 35, 89-90.

Nakatani, P. D. & C. Saito, I. (2000) Modulation of DNA-mediated hole transport efficiency by changing superexchange electronic interactions. *J. Am. Chem. Soc.* **122**, 5893-5894.

Núñez, M.E. & Barton, J.K. (2000) Probing DNA charge transport with metallointercalators. *Curr. Opin. Chem. Biol.*, **4**, 199–206.

Nunez, M.E., Hall, D.B., & Barton, J.K. (1999) Long-range oxidative damage to DNA: effects of distance and sequence. *Chem. Biol.* **6**, 85-97.

Núñez, M.E., Holmquist, G.P., & Barton, J.K., (2001) Evidence for DNA Charge Transport in the Nucleus. *Biochemistry*, **40**(42), 12465 -12471

Piette,J. (1991) Mechanism of DNA cleavage mediated by photoexcited nonsteroidal antiinflammatory drugs. *J. Photochem. Photobiol. B*, **11**, 241–260.

Rakitin, A., Aich, P., Papadopoulos, C., Kobzar, Y., Vedeneev, A. S., Lee, J. S., and Xu, J. M. (2000) Metallic Conduction through Engineered DNA: DNA Nanoelectronic Building Blocks. *Phys. Rev. Lett.* **86**, 3670–3673.

Reynisson J, Schuster GB, Howerton SB, Williams LD, Barnett RN, Cleveland CL, Landman U, Harrit N, Chaires JB. (2003) Intercalation of trioxatriangulenium ion in DNA: binding, electron transfer, x-ray crystallography, and electronic structure. *J Am Chem Soc.* 2003 **125**(8), 2072-83.

Sanii, L. & Schuster, G.B. (2000) Long-Distance Charge Transport in DNA: Sequence-Dependent Radical Cation Injection Efficiency. *J. Am. Chem. Soc.*, **122**(46), 11545-11546.

Sankar, C.G. & Sen, D. (2004) DNA helix-stack switching as the basis for the design of versatile deoxyribosensors. *J Mol Biol.* **340**(3),459-67.

Schoonveld, W.A., Vrijmoeth, J., Klapwijk, T.M. (1998) *Appl. Phys. Lett.* **102**, 9625-9634.

Schuster, G.B. (2004) Long-Range Charge Transfer in DNA I and II. *Springer-Verlag, Heidelberg.*

Schuster, G.B. (2000) Long-Range Charge Transfer in DNA: Transient Structural Distortions Control the Distance Dependence. *Acc. Chem. Res.*, **33**, 253.

Schwogler, A., Burgdorf, L. (2000) Self-reparing DNA based on Reductive Electron Transfer through base stack. *Acc. Chem. Res.* **33**(21), 3918-20.

Seidel, C.A.M., Schultz, A. & Sauer, M.H.M. (1996) Nucleobase-specific quenching of fluorescent dyes. Nucleobase one-electron redox potentials and their correlation with static and dynamic quenching efficiencies. *J. Phys. Chem.* **100**, 5541-5553.

Silverman, SK. (2005) *In vitro* selection, characterization, and application of deoxyribozymes that cleave RNA. *Nucleic Acids Res.* **33**(19), 6151-63.

Spassky, A. & Angelov, D. (1997) Influence of the Local Helical Conformation on the Guanine Modifications Generated from One-Electron DNA Oxidation. *Biochemistry*, **36**(22), 6571-6576.

Steenken, S., & Jovanovic, S.C. (1997) How easily oxidizable is DNA? One-electron reduction potentials of Adenosine and Guanosine radicals in aqueous solution. *J. Am. Chem. Soc.* **119**, 617-18.

Takada, T., Kawai, K., Cai, X., Sugimoto, A., Fujitsuka, M., & Majima, T. (2004) Charge Separation in DNA via Consecutive Adenine Hopping. *J. Am. Chem. Soc.* **126**, 12843-12846.

Tombelli, S., Minunni, M., & Mascini, M. (2005) Analytical applications of aptamers. *Biosens Bioelectron*, **20**(12), 2424-34.

Tuerk, L., & Gold, L. (1990). Systematic Evolution of Ligands by Exponential Enrichment: RNA Ligands to Bacteriophate T4 DNA Polymerase. *Science,* **249**, 505-600.

Turro, C.J. & Barton, J.K. (1998) Paradigms, supermolecules, electron transfer and chemistry at a distance. What's the problem? The science or the paradigm. *J. Biol. Inorg. Hem.* **3**, 201-209.

Ventra, M.D. and Zwolak, M., (2004) Encyclopedia of Nanoscience and Nanotechnology. *American Scientific Publishers*, **2**, 475.

Wang, S.Y., ed. (1976) Photochemistry & Photobiology of Nucleic Acids: Chemistry *Academic*, **1**.

Watson, J.D. & Crick, F.H.C. (1953) A structure for deoxyribose nucleic acid. *Nature* **171**, 737-738

Zumdahl, S. (1995) Chemical Principles, *D.C. Heath & Co*, **2**, 749-751.