Studies on oxidative damage of proteins by metallodesferals and designer oligonucleotides in DNA diagnostics and restriction enzyme recognition.

THESIS

SUBMITTED TO

UNIVERSITY OF POONA

FOR

THE DEGREE OF DOCTOR OF PHILOSOPHY

IN BIOTECHNOLOGY

BY

VASANT R. JADHAV

NATIONAL CHEMICAL LABORATORY

PUNE - 411 OO8

MARCH 1998

Dedicated to my parents.....for sharing their genes with me!

CERTIFICATE

Certified that the work incorporated in the thesis entitled "**Studies on oxidative damage of proteins by metallodesferals and designer oligonucleotides in DNA diagnostics and restriction enzyme recognition**" submitted by **Mr. Vasant R. Jadhav** was carried out by the candidate at National Chemical Laboratory, Pune, under my supervision. Such materials as obtained from other sources have been duly acknowledged in the thesis.

(K. N. GANESH)

Research Guide

Head, Division of Organic Chemistry (Synthesis)

National Chemical laboratory

Pune-411 008.

March 1998

CANDIDATE'S DECLARATION

I hereby declare that the thesis entitled "**Studies on oxidative damage of proteins by metallodesferals and designer oligonucleotides in DNA diagnostics and restriction enzyme recognition**" submitted for the degree of Doctor of Philosophy in Biotechnology to the University of Poona has not been submitted by me for a degree to any other university or institution. This work was carried out at the National Chemical Laboratory, Pune, India.

Jadhav

(**Vasant R. Jadhav**)

National Chemical Laboratory
Pune- 411 008.
March 1998

ACKNOWLEDGMENTS

It gives me great pleasure to express my deep sense of gratitude to Dr. K. N. Ganesh for introducing me to the interdisciplinary science of chemistry and biology of nucleic acids. His receptive attitude, untiring enthusiasm and positive approach will always remain a source of inspiration.

I am indebted to my parents who have been a source of encouragement and I am grateful for all their sacrifice.

I am thankful to Prof. D. N. Deobagkar and all my teachers for their encouragement in my initial years. I take this opportunity to thank the Director, National Chemical Laboratory, for allowing me to work in this laboratory and providing necessary facilities.

I was fortunate to meet three wonderful people - Rajeev, Gopal and Jayashree, during my stay here. It is due to their help, this thesis has taken present shape.

I take this opportunity to express my sincere thanks to Drs. A. A. Natu, Mrs. V. A. Kumar, Mangesh Bankar and Dinesh Barawkar for their valuable scientific discussion and help during the course of this work..

My sincere thanks are due to Prof. Anil Lala, IIT, Mumbai, Profs. S. Maitra & Doraiswamy, TIFR, Mumbai, for allowing me to use their facilities. I thank Prof. D. Balasubramanian, CCMB, for help in amino acid analysis and Prof. Arthur Broom, University of Utah, for providing mass spectra of oligonucleotides.

I am thankful to Mrs. A. Gunjal and Mrs. S. Kunte for their help in oligonucleotides synthesis.

I am thankful to Ramesh Babu for his expert help in amino acid analysis and Dr. Ameeta Ravikumar for help in Protein work.

My thanks are also due to Mrs. M. V. Mane for HPLC analysis and the library staff for excellent facilities. I gratefully acknowledge and thank Mr. A. G. Samuel, Dr. G. J. Sanjayan and Mr. Sathe for their help in recording NMR spectra and Mr. Parag for help in photography.

I am indebted to my friends Prakash Lohar, Ramesh, Rajendra, Raju, Mohan, Namdev and Murari for their help and encouragement.

I take this opportunity to thank my colleagues, Vipul, Vallabh, Gangamani, Meena, Pradeep, Moneesha, Anand, Leena, Tanya, Sanjib, Praveen, Nagamani and Dr. Dimpan.

Tomal, Vijay, Bharat Gawande, Gyanendra Tripathi, Hegde, Sreedhara, Patrick, Lakshmikant and friends in plant molecular biology unit have helped me in various ways for which I shall remain indebted.

I am grateful to Drs. R. R Joshi, T. P. Prakash, M. Renil, Krishna Kumar and K. Sakthivel for their encouragement.

I am thankful to Anuj, Tushar, Radha, Partha, Prasad, Saumen, Rajesh, Chengal, Uday, Santhosh, Deepak, Karthikeyan, Shirish, Selvam and Ravi for providing me a cheerful hostel atmosphere.

I thank Prakash, Khandekar, Makhar, Sunil and Bhumkar for their help and cooperation.

Finally I acknowledge CSIR (INDIA) for valuable support in the form of fellowship.

VASANT R JADHAV

CONTENTS

Synthesis and application for identifying restriction cleavage sites

LIST OF PUBLICATIONS:

1. PCR based test for the detection of Bovine Herpes Virus in Indian isolates
   Indian Scientific and Industrial Research (1994), **2**(2), 148-150.
   Jadhav, V. R., Natu, A. A. & Ganesh, K. N.

2. Synthesis, Characterisation and DNA interaction studies of Cr(III) products isolated from Cr(IV) reduction with -SH containing molecules.
   Polyhedron (1996) **15**(5-6), 765-774.
   Kaiwar. S. P., Rao, C. P., **Jadhav, V. R.** & Ganesh, K. N.

3. 5-amido-(carboxyfluorescein)-2'-dU oligonucleotides: Novel primers for fluorescent detection of PCR amplified DNA.
   Nucleosides & Nucleotides (1997) **16**(1&2), 107-114.
   Jadhav, V. R., Natu, A. A. & Ganesh, K. N.

4. Triplex formation at physiological pH: comparative studies on DNA triplexes containing 5-Me-dC tethered at N^4 with spermine and tetraethyleneoxyamine
   Nucleic Acids Res. (1997) **25**(21), 4187-4193.
   Rajeev, K. G., **Jadhav, V. R.** & Ganesh, K. N.

5. Synthesis of chiral hydroxamic acid derivatives and the study of interaction of their metal complexes with DNA
   Chem. Soc. Perkin Trans I (1998) In press.
   Chittari, P., **Jadhav, V. R.,** Ganesh, K. N. & Rajappa, S.

6. Oxidative damage of proteins by Cu(II)-desferal complex
   J. Biochem. Mol. Bio. & Biophys (1998) In press.
   Jadhav, V. R. & Ganesh, K. N.

7. Design of a combinatorial oligonucleotide library containing all possible hexamer palindromes: PCR synthesis and application for identifying restriction cleavage sites
   Biochem. Biophys. Res. Commun. (1998) In press.
   Jadhav, V. R. & Ganesh, K. N.

Part of work in chapter 2 and 4 is under preparation for publication.

ABSTRACT

CHAPTER 1: Applications of synthetic oligonucleotides as research, diagnostic and therapeutic agents

The development of chemical synthesis of oligonucleotides and the total automization of the process has revolutionized the field of molecular biology. It has chiefly contributed to the generation of a vast amount of information in molecular genetics. The property of specific recognition by oligonucleotides of its complementary nucleic acid sequences is exploited in their use as hybridization probes for detection and sequencing of a number of genes. Apart from the classical applications, synthetic oligonucleotides are now viewed as drugs, diagnostic and therapeutic agents. This chapter overviews the classical applications of oligonucleotides while the newer applications in improving drugs through oligonucleotide combinatorial chemistry, drug development by antisense/antigene therapy, DNA diagnostics through use as primers in polymerase chain reaction (PCR) are described in detail. The other novel applications like sequencing by hybridization using oligonucleotide arrays, their role in targeted cleavage of DNA and uses in nanotechnology are also described.

CHAPTER 2: Fluorescent oligonucleotides: Synthesis, evaluation of fluorescence properties, effect of spacer chain and PCR applications.

Fluorescence detection of nucleic acids for application in DNA diagnostics and molecular biology is attracting wide attention due to its safe and easy handling. The lower sensitivity of this technique compared to conventional radiolabels can be compensated by either multiple labeling or signal amplification of target through PCR technique. The detection procedures employing fluorescent oligonucleotides are of significant importance in automated DNA sequencing, nucleic acid hybridization, DNA-protein interactions and PCR. The nucleobase labeling strategies reported so far involve the use of spacer arm between fluorophore and nucleobase.

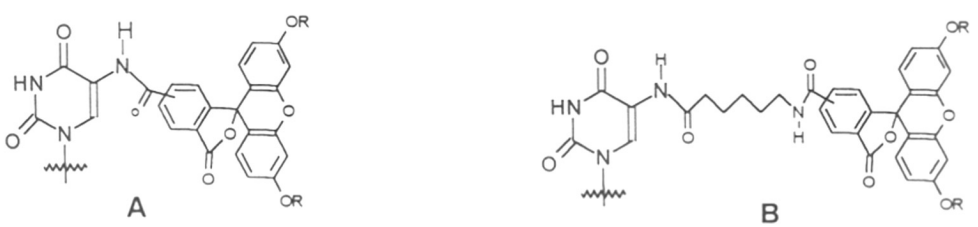


Figure 1. The fluorescent deoxyuridine nucleobase. **A**) Attachment of fluorophore through short and rigid amido spacer, **B**) Attachment of fluorophore through long and flexible hexanoyl spacer.

This chapter describes the synthesis and study of fluorescence properties of oligonucleotides in which the fluorescent dye 5/6 carboxyfluorescein is linked to 5-NH_2-deoxyuridine either with short and rigid amido spacer or with long and flexible hexanoyl spacer

chain (Figure 1). The quenching of fluorescence was found to be more when fluorophore is directly attached to base. These oligonucleotides showed changes in the fluorescence intensities upon duplex formation, a property which was utilised to monitor the duplex formation. The duplex formation resulted in the increase of anisotropy values of fluorophore. The decreased mobility of DNA strand upon duplex formation was transferred more efficiently to directly linked fluorophore than the one attached via spacer. These fluorescent oligonucleotides were used to monitor the binding of netropsin in the minor groove of DNA by induced changes in fluorescence. The binding constant of netropsin with DNA was found to be in the range of 10^6 M^{-1} which is close to the reported values in literature. These oligonucleotides were also used as PCR primers for the amplification of 861 bp globin gene region in human genomic DNA. The oligonucleotides having fluorophore attached through spacer were found to be better as they were efficiently utilised by *Taq* polymerase and showed higher fluorescence sensitivity. Overall, the work presented in this chapter aims to evaluate the consequences of spacer chain on various fluorescent observables.

CHAPTER 3: Design of a combinatorial oligonucleotide library containing all possible hexamer palindromes: PCR synthesis and application for identifying restriction cleavage sites

In the last two decades, a number of Type II restriction endonucleases which play a major role in gene cloning technology have been reported in literature and their number continues to increase. Many of these enzymes recognize tetra or hexamer palindromic sequences of duplex DNA as specific sites of cleavage. The identification of the specific DNA cleavage sites of these enzymes currently involves cloning, DNA sequencing, and biocomputational techniques. This chapter proposes and explores a new improved method towards this end. It involves design and synthesis of a set of eight oligonucleotides (33-37 nt), containing all possible hexamer palindromic sequences present as unique sites. In any hexamer palindrome, the first three bases decide the sequence of next three bases. The four bases of DNA permit a total of 64 different possible triplets and hence only 64 possible corresponding hexamer palindromes. The maximum length of DNA sequence containing all these sites will be 64 x 6 = 384 bases. This length of DNA could be reduced by tandem design strategy by grouping hexamers in such a way that last two bases of one hexamer palindrome site correspond to first two bases of next hexamer palindrome. Addition of one of the four bases at each time at third position will complete the first three bases of palindrome which automatically fixes the next three bases of hexamer palindrome. This algorithm leads to a combination set in which each oligonucleotide represent all possible hexamer palindromes starting with two complementary doublets.. The 6 base pair complementarity at the 3' end of these oligonucleotides has been exploited to make them double stranded by PCR using primer-dimer approach. These duplex PCR products screened as substrates for known restriction enzymes showed the cleavage at appropriate sites. As these kind of oligonucleotides represent all possible hexamer palindromes, they have potential application for characterization of new, unknown restriction enzymes that

recognize hexamer palindromic sequences, thus simplifying the existing methods. The double stranded PCR products of these oligonucleotides are self complementary sequences containing two inverted repeats which exists in two forms. The intrastrand association results in hairpins while interstrand association leads to linear duplexes. The experimental parameters to control and evaluate the equilibration of formation of these two forms have been studied.

CHAPTER 4: 5-amido-deoxyuridine containing oligonucleotides: Effect on *Eco*RV activity and DNA major groove polarity

The specific interactions between proteins and nucleic acids play extremely important role in biological processes. The restriction endonucleases are a class of enzymes which cleave DNA at a particular sequence with very high specificity. The molecular interactions involved in this specific interaction are generally studied by using modified DNA substrates in which the functional groups of particular bases are modified. Using various modified bases, it has been postulated that the 5-methyl group of thymine residues present in recognition sequence of *Eco*RV (GATATC) is involved in hydrophobic interaction with enzyme. This chapter describes the studies on the replacement of hydrophobic 5-methyl group of thymine with 5-NH_2 group (Figure 2) and its incorporation into the oligonucleotides containing *Eco*RV recognition sequence.

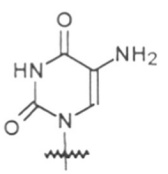


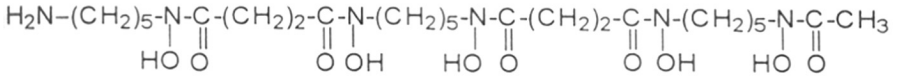
Figure 2. Structure of 5-NH_2-deoxyuridine.

The results showed that replacement of thymine in recognition site with 5-NH_2-dU inhibits the cleavage by enzyme. This may be due to the loss of hydrophobic interaction between modified base and enzyme. The effect of the 5-NH_2 replacement on the polarity of major groove of DNA was studied using covalently conjugated dansyl fluorophore which suggested an increased polarity of major groove of DNA upon incorporation with 5-NH_2-dU. It is concluded that the hydrophobic interaction between 5-methyl group of thymine base and *Eco*RV play important role in its enzyme activity.

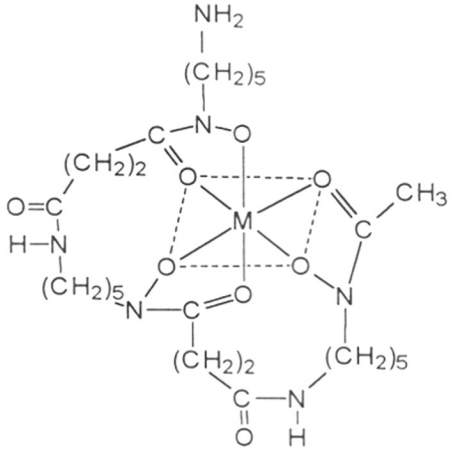
CHAPTER 5: Oxidative damage to proteins by Cu(II)Desferal complex

Biological damage induced by reactive oxygen species is an important factor in disease and aging. There are various systems which generate oxygen free radicals that catalyze the oxidative damage of proteins, DNA and lipids. This chapter reports protein damage activity of Cu(II)Desferal complex that was previously shown to posses DNAse activity. Desferal (Figure 3), a naturally occurring siderophore produced by actinomycetes, is used for iron transport by microbial systems. Due to its high affinity for Fe(III) and the redox inactivity of ferrioxamine, it is

widely used as a drug for the treatment of iron overload diseases despite the associated cerebral and ocular toxicity. Although Fe(III)Desferal complex is unable to produce highly reactive hydroxyl radicals (.OH), desferal therapy administered with vitamin C (ascorbate) shows toxic side effects.



1



2 M = Fe, Cu

Figure 3. Structure of desferal **1** and its metal complex **2**

Owing to its importance as a drug and related side effects, studies have been carried out on the effect of CuDFO on proteins such as lysozyme and BSA. This chapter reports the interesting protease activity of CuDFO in presence of H_2O_2 (Figure 3, M = Cu) on lysozyme and BSA analysed through gel electrophoresis, enzyme activity, protein fluorescence, amino acid analysis and peptide sequencing. Exposure of CuDFO/H_2O_2 to lysozyme resulted in fragmentation and aggregation of proteins accompanied by the loss of enzyme activity. The tryptophan fluorescence of proteins was found to decrease upon CuDFO/H_2O_2 treatment indicating the damage of tryptophan. The specific protein fragments of lysozyme resulting from CuDFO/H_2O_2 treatment were analyzed by N- terminal sequencing. The amino acids methionine, tyrosine and proline were prone to modification by Cu(II)Desferal attack as analysed by amino acid composition.

CHAPTER 1

Applications of synthetic oligonucleotides as research, diagnostic and therapeutic agents

1.1 INTRODUCTION

Synthetic oligonucleotides are short streches of single stranded DNA fragments which are not isolated from natural sources but assembled from mononucleotide units by chemical synthesis according to a defined sequence. After the pioneering work of Michelson & Todd (1955) on phosphotriester synthesis, Khorana and coworkers (1978, 1979) were able to develop phosphodiester-chemistry to an impressive height, culminating into the chemical synthesis of tRNA gene. This opened the field of chemical synthesis of oligonucleotides which all that time did not have much biological applications, but was rather challenging to the organic chemist. Over the past twenty years this situation has dramatically changed. There is hardly a field in biology in which synthetic DNA has not been used or does not have the potential to be used. The main reason for such a turnabout is the simultaneous developments of methods for the rapid and efficient synthesis of oligonucleotides and the advent of molecular cloning techniques. It is not surprising that these two technologies had an impact on each other, since they both dealt with the creation of new combinations of nucleotide sequences (Itakura et al., 1984).

The increasing availability of synthetic DNA sequences has led to a revolution in the way molecular biologists approach problems. Synthetic oligonucleotides make possible the creation of man-made genes, the creation of site-specific mutations (precise point mutations, insertions and deletions) and the diagnosis of mutations responsible for human genetic diseases. Thus biology and chemistry have symbiotically developed in this as nowhere (Engels & Uhlmann, 1988).

The aim of this chapter is to give brief overview of the classical applications of synthetic oligonucleotides with an emphasis on the new roles of oligonucleotides as diagnostic, therapeutic agents and other novel applications.

1.2 CLASSICAL APPLICATIONS

Since the availability of synthetic oligonucleotides in 70's and 80's, they are primarily being used for following classical applications.

1.2.1 Selection of Genes

It is not always possible to identify directly the gene corresponding to the messenger RNA coding for a protein of interest. If however, the amino acid sequence of protein in part or whole is available, it is possible to predict the corresponding gene sequence. Oligonucleotides synthesised according to a predicted sequence can then be used to locate the required component from a mixture of DNA or RNA species. Because of the "wobble" in the genetic code, prediction of a unique oligonucleotide sequence is sometimes difficult. Mixtures of oligonucleotides covering all possible sequence variants can be synthesised and used in such a case. Synthetic oligonucleotides are routinely used to select sequences from both mRNA and genomic DNA (Messing et al., 1981).

1.2.2 Site Directed Mutagenesis

Synthetic DNA mediated mutagenesis provides one of the most powerful methods of studying proteins and DNA-protein interactions. The systematic changes in the amino acid sequence of protein can be obtained by changing the corresponding gene sequence (Zoller & Smith, 1983). Mutagenesis at specific sites can be carried out by use of a mismatched oligonucleotide primer hybridised to the wild type single stranded DNA template. This can be achieved by cloning in M13 bacteriophage or by partial removal of one strand of plasmid duplex (Figure 1). This strategy is routinely used to produce deletion and transversion mutants (Morinaga et al., 1984).

1.2.3 Sequencing Primers

All known DNA polymerases require a short segment of DNA as a primer to initiate DNA synthesis. This has been exploited in Sanger's method of enzymatic DNA sequencing (Sanger et al.,1977) which requires a specific primer for the synthesis of complementary DNA from a single strand template using DNA polymerase. These oligonucleotides can be radiolabeled to detect the sequencing products.

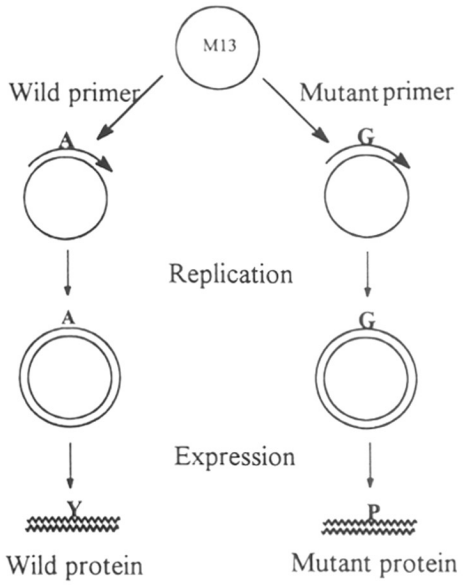


Figure 1. Schematic representation of site directed mutagenesis. The single stranded template (usually M13 phage) is annealed separately with wild and mutant primers containing desired changes in bases which after replication results in the duplex carrying expected changes. Expression of this DNA gives the protein with designed changes in amino acid sequence. In above hypothetical case, the change of base A→G in DNA results in the change of amino acid Y to P in protein.

1.2.4 Linkers and Adaptors

During genetic manipulation, it is often useful to create known restriction enzyme cleavage sites at specified points in DNA molecule. In particular, during cloning it is desirable to have known restriction sites at the points of insertion of foreign DNA into the vector molecule (Davis & Thorner, 1987). This allows the predictable recovery of cloned insert sequences from the recombinant by the use of oligonucleotide linkers and adapters (Scheller et al., 1977). Linkers are self complementary single stranded oligonucleotides which can self anneal to form duplex containing restriction enzyme site. Such molecules are joined to the DNA fragment of interest which can be further cloned after digestion with restriction enzyme specific for a site present in linker. Adaptors are single stranded oligonucleotides containing multiple restriction sites. Single adptors may be used to join otherwise incompatible cohesive ends of DNA. Pairs of partially

complementary adaptors may be used to join non-identical restriction ends of DNA whilst creating a restriction site for a further enzyme at the point of joining.

1.2.5 Synthetic Genes

The first methodology for assembling synthetic oligomers into extended duplexes was developed by Khorana (1979) using ligase-mediated linking of relatively short oligonucleotides (10-20 mers) carrying overlapped complementary sequences. Further developments in oligonucleotide synthesis made it possible to prepare 50-100 nucleotide-long polynucleotides. This paved way for the synthesis of duplexes involving DNA polymerase-mediated repair synthesis of double-stranded polynucleotide from a partial duplex formed by the 3'-termini of two long polynucleotides (Figure 2). This approach was used for preparation of human interferon gene fragments (Itakura, 1982).

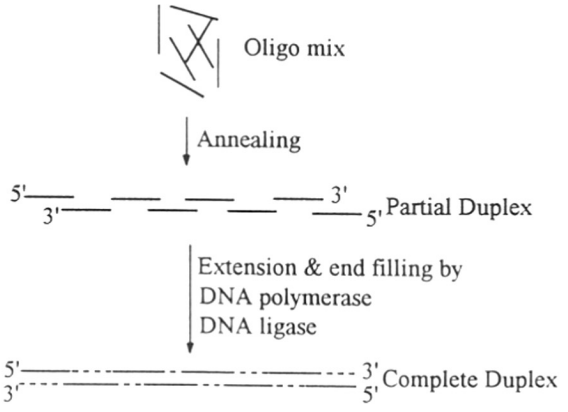


Figure 2. Schematic representation of gene synthesis. The pool of oligos having unique 3'- end complementary sequences are mixed and under appropriate hybridisation conditions it results in to partial duplex formation. Such partial duplexes are extended by DNA polymerase and the ends are filled by DNA ligase to obtain a complete duplex.

1.2.6 Probes for Genetic Diseases

Based on the ingenious blotting method developed by Southern (1975), oligonucleotide probes are used to detect single base pair changes within human genome. For example, sickle cell anemia caused by a transversion (A→T) within β-globin gene can be analysed with labeled oligonucleotides. Under stringent hybridisation conditions , it is possible to analyse whether one or two of family's

alleles contain the mutant gene. This has also been successfully used for bacterial systems (Trevors, 1985).

1.2.7 DNA-Ligand Interaction

Nucleic acids interact with a broad range of chemical species that include water, metal ions and their complexes, small organic molecules and proteins (Wilson, 1990). The prerequisite of studying such interactions is the availability of pure sequence specific nucleic acids in reasonably good amounts. Synthetic oligonucleotides are widely used in such applications by employing techniques such as crystallography, NMR etc. (Kopka et al., 1987). Use of well defined oligonucleotide sequences has permitted identification and investigations of structural polymorphism in DNA, (A-, B- and Z-DNA) with great detail (Saenger, 1984). The higher order structures such as triplex (Dervan,1992), tetraplex (Sen & Gilbert, 1988) and cruciform (Wells, 1988) are systematically elucidated by using synthetic oligonucleotides. The interaction of proteins with DNA which is of paramount in gene expression is widely studied using synthetic oligonucleotides as substrates (Brennan & Matthews, 1989; Rajendrakumar et al., 1990). The basic knowledge obtained using synthetic oligonucleotides about variations of DNA structure and its interaction with different molecules has resulted in development of many newer and novel applications of oligonucleotides. One of the most important aspect of synthetic oligonucleotides is their ability for selective modification or conjugation with different molecules such as reporter groups at different sites on DNA (5' or 3'-ends, nucleobases, sugar and phosphate backbone). A variety of conjugated reporter groups - fluorescent labels, biotin etc. are widely useful in the detection of nucleic acids (Kricka, 1992).

The last decade has witnessed new role of oligonucleotides as diagnostics and therapeutic agents. The following section focusses on recent applications of synthetic oligonucleotides in Polymerase Chain Reaction (PCR) for DNA based diagnostics, in combinatorial methods for drug discovery and in antisense-antigene technology for therapeutics. It also describes the novel role

of oligonucleotides in sequencing by hybridisation, genome mapping and nanotechnology.

1.3 DIAGNOSTICS APPLICATIONS

Conventional diagnostics based on enzymatic and immunological procedures involve the detection of specific gene product associated with the disease under condition. The applicability and success of this technique depends on the detectable expression of these products. In several viral and genetic diseases, a defective gene, although present, may not express itself, thus preventing its detection by the above methods. The route, DNA→RNA→proteins, which is the central dogma of information flow during gene expression is the basis of new DNA-based diagnostics (Persing et al., 1993). Detection at DNA level rather than gene product (protein) level offers attractive possibilities for the development of new diagnostic systems. Although the widespread use of DNA probes for the diagnosis of infectious diseases has been discussed for many years (Moseley et al., 1980), the last decade witnessed a near explosive growth in the number and variety of specific applications. Among these, one of the most outstanding discovery of 80's is polymerase chain reaction (PCR) technique (Mullis, 1990) in which synthetic oligonucleotides play a very important role as primers for polymerase-directed nucleic acid synthesis. This section describes the principle, methodology and applications of PCR.

1.3.1 Background

Only a decade ago, the prospect of producing billions of amplified copies of specific nucleic acid sequences by performing successive rounds of *in vitro* nucleic acid replication would have been considered science fiction. Interestingly, the basic ingredients of an *in vitro* nucleic acid amplification method were described by Kleppe et al., (1971) in which extensive synthesis of tRNA gene by primer-directed DNA repair was postulated. However, this work did not result in an exponential amplification process (the hallmark of PCR), and the concept remained largely unrecognised in literature for many years.

In 1983, in a moonlit drive through mountains of northern California, Cetus scientist Kary Mullis envisioned a process of *in vitro* nucleic acid amplification procedure that would eventually become known as PCR (Mullis, 1990). Buoyed by several technologies that matured in 1980s, including automated oligonucleotide synthesis, PCR became a reality in a relatively shorter time, leading to the publication by Saiki et al.,(1985) of its first practical application and since then it has developed as a mainstay technique in many molecular biology laboratories.

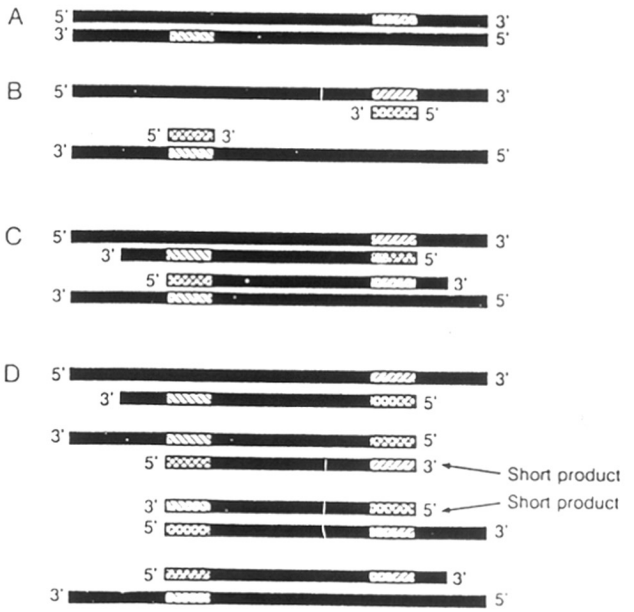
1.3.2 PCR Principle

The apparently simple principle of PCR is outlined in Box 1. The method uses repeated cycles of oligonucleotide-directed DNA synthesis to perform *in vitro* replication of target nucleic acid sequences. The synthetic oligonucleotide primers whose sequences are determined by the target nucleic acid are synthesised as complementary to their annealing sites within two different strands of a target DNA sequence.

1.3.3 Experimental Methodology

In its simplest form, each cycle of PCR consists of three steps: (i) a denaturation step, in which the target DNA is incubated at high temperatures so that the target strands are separated apart and thus made accessible to hybridisation by specific oligonucleotide primers; (ii) an annealing step, in which the reaction mixture is cooled to allow the primers to anneal on their complementary target sequences; and (iii) chain extension reaction, usually done at an intermediate temperature, in which the primers are extended on the DNA template by a DNA polymerase. These three steps are linked to what is referred to as a thermal cycle. A typical PCR protocol comprises 25 to 40 thermal cycles. After completion of a single thermal cycle, there is theoretical doubling of the target sequence. Thus, repeating the thermal cycle results in a geometric accumulation of amplified target sequences.

Box 1



PCR principle. In first PCR cycle, a double stranded DNA target sequence is used as template, with the primer binding sites indicated by hatched lines (A). These two strands are separated by heat denaturation and two synthetic oligonucleotide primers (cross hatched lines) anneal to their respective recognition sequences in the 5'-to-3' orientation (B) facing each other. *Taq* DNA polymerase initiates the synthesis at 3' ends of each primer (C). Extension of primer via DNA synthesis results in new primer binding sites. The net result after one round of synthesis is two "ragged" copies of the original target DNA molecule. In the second cycle, each of the four DNA strands shown in panel C anneal to the primers (present in excess) to initiate a new round of DNA synthesis (D). Of the eight single stranded products, two are of a length defined by distance between and including the primer annealing sites which accumulate exponentially in subsequent cycles. Thus, using defined primers, PCR can amplify any particular DNA sequence whose length can be from few base pairs to kilobase pairs (Persing, 1993).

1.3.4 PCR Product Detection

The simplest method of analysing amplified DNA involves gel electrophoresis of products according to their molecular weight and detection of these products by ethidium bromide staining. Infact, PCR can be done in

presence of ethidium bromide and the increase in fluorescence can be detected online without opening the reaction tube (Higuchi et al., 1992).

The next level of complexity in detection system involves primers or probes which are derivatised with ligands or signal moieties. Chehab and Kan (1989) labeled each primer with a different fluorescent dye; after separation of the unincorporated primers by selective filtration, they detected the amplified DNA by gel electrophoresis and UV light. Alternatively primers can be labeled with biotin and the product is then detected with standard avidin-peroxidase reagent. The incorporation of digoxigenin-11-dUTP or biotin-dUTP during PCR results in labeled product. After electrophoresis, the gel is blotted on nylon membrane and the labeled DNA is detected with anti-digoxigenin antibody:alkaline phosphatase conjugate and a chemiluminescent substrate (Ou et al., 1990). The gel based methods are still relatively cumbersome and offer mainly an increased sensitivity over the original method of simply detecting a band of expected size on an agarose gel, but are prone to PCR carryover and false positive results (Kwok and Higuchi, 1989).

A novel method of detection that makes use of an exonuclease activity of some thermostable DNA polymerases was recently described (Holland et al., 1992). The 5'-to-3' exonuclease activity of *Taq* DNA polymerase was used to generate a sequence specific probe for signal generation during PCR. In each cycle of amplification, the enzyme degrades the probe and releases a smaller fluorescent fragment from the quencher component. This results in an increase in the fluorescence which can be monitored in the reaction tube itself. It is to be noted that as fluorescent probe is specific for the desired amplification product, the increase in fluorescence occurs only during specific amplification. This method has potential for giving a very broad dynamic range for quantitative assays and requires little or no handling of the sample after amplification. Such approaches may reduce the chance of contamination and combining several differentially labelled probes in a single reaction, permits multiplex analysis of several agents.

1.3.5 PCR in Diagnostics

Although PCR has found widespread applications in molecular biology, its singlemost important application is in diagnosis of genetic and infectious diseases. Indeed, published reports of new applications for the diagnosis of genetic and infectious diseases, in terms of technology that they exploit, are accumulating at an exponential rate. New organisms, many of which have never been cultivated on artificial medium, are being identified, and their associations with human diseases are being elucidated. In only few years after introduction of PCR, diagnostic molecular microbiology, has evolved from a mere technologic curiosity into a field of vast breadth and complexity, with the potential to revolutionise practice of medicine (Persing et al., 1993).

The key advantage in using PCR is its ability of specific signal amplification. The signal present in the form of very minute quantities of nucleic acids is amplified to detectable limits. The selective amplification of target sequence that are present in low abundance in a background of complex genomic DNA is achieved. This feature makes it potentially useful for the diagnosis of pathogens present in small numbers whose DNA or RNA copurifies with genomic DNA. One of the best studied example of this is detection of proviral sequences for human immunodeficiency virus type 1 (HIV-1), in which the low prevalence in human mononuclear cells of virus specific sequences precludes the use of conventional hybridisation technique (Ou et al., 1988). The list of application of PCR for the sensitive and specific detection of bacterial, fungal, and parasitic pathogens has grown immensely. The current applications of PCR to the detection of clinically relevant infectious microorganism are recently reviewed (Persing et al., 1993).

PCR has achieved widespread use in the analysis of genetic diseases. The target region is usually amplified from genomic DNA and cDNA, and examined for mutations and polymorphisms by sequencing, hybridization with allele-specific oligonucleotides, restriction analysis, or enzymatic or chemical cleavage (Ehrlich, 1989). Deletions in genomic DNA can be identified by

determining if a segment of DNA can be amplified (Chamberlain et al., 1989). The single base mutations are identified using ARMS-PCR (Amplification Refractory Mutation System) in which the difference between wild type and muant primer is the single base difference. The primer hybridisation/annealing conditions are used in such a way that even single base mismatch is not tolerated (Newton et al., 1989).

The application of PCR technology in research and diagnostics has proceeded at an unprecedented rate. Because of its extreme sensitivity and versatility, the long-term implication of this technology in all areas of biological research and medicine are potentially tremendous. Although conceptually, PCR method is simple and is based on the long-understood DNA replication process in life forms, in reality the power and applicability of this method are phenomenal.

1.4 OLIGONUCLEOTIDES IN DRUG DISCOVERY

Many factors must favorably coincide in order for a compound to be useful as a drug. Although the compound must obviously have inherent therapeutic activity, several other pharmacological criteria (e.g., lack of toxicity, delivery to the site of action and bioavailability) must also be satisfied. Several characteristics steps are involved in the evolution of a drug. Typically, at first a lead compound is discovered that has a desired activity profile in an assay. The compound is then tested for safety, pharmacokinetics, and pharmacodynamics in animals and ultimately in patients, prior to being tested for therapeutic activity. Classical approaches to identification of a lead compound are based on repeated assays of interaction of small organic molecules with selected therapeutic targets. The source of these molecules has generally been large in-house libraries of diverse compounds each synthesised for potential activity against a particular target. In view of the multiple criteria that a successful drug must satisfy through several stages of testing, it is hardly surprising that generating drugs has historically been a time-consuming, labour-intensive, and expensive undertaking with a relatively low probability of success. Needless to

say, any strategy that efficiently reduced time or uncertainty in the drug discovery process would have significant value (Trotta et al., 1995).

1.4.1 Combinatorial Chemistry

Combinatorial chemistry is now being widely used for drug discovery in an effort to maximise probability of success while minimising the time required to reach a lead product (Hogan Jr, 1996). In the combinatorial chemistry approach, very large libraries of compounds are synthesised, often simultaneously. Entire libraries or large portions thereof are assayed simultaneously as well.

The challenges and limitations inherent to combinatorial chemistry involve ability to generate libraries in a labour-efficient manner and detect the active compounds among a very larger population of chemically related inactive ones. Combinatorial library synthesis and selection strategies focus on these factors since they influence the probability of identifying active compounds if the procedure includes a random screening approach (Broach & Thorner, 1996).

1.4.2 Combinatorial Chemistry with Oligonucleotides

Since oligonucleotides can be synthesised efficiently and easy methods are available for their identification, these compounds would, in theory, be appropriate for combinatorial chemistry approaches (Eaton and Pieken, 1995). Indeed, oligonucleotides were perhaps the first class of compounds used in combinatorial chemistry (Tuerk and Gold, 1990; Ellington and Szostak, 1990). Oligonucleotides (either DNA or RNA) can be synthesised in an automated fashion in a nucleic acid synthesiser. Numerous procedures are available that allow sequencing of oligonucleotides. Further, the availability of PCR and cloning technologies present a unique opportunity (relative to other types of oligomers or polymers that might be used in combinatorial chemistry approaches) to amplify minute amounts of nucleic acid material to generate concentrations that are amenable to sequencing.

1.4.3 SELEX

SELEX (Systematic Evolution of Ligands by Exponential enrichment. (Tuerk and Gold, 1990) is a combinatorial chemistry methodology in which vast

numbers of oligonucleotides (DNA, RNA, or unnatural compounds) are screened rapidly for sequences that have appropriate binding affinities and specificities toward any target.

Strategy: The selection concept of oligonucleotide combinatorial chemistry libraries is to expose the target to a population of oligonucleotides, isolate the target molecule and elute the bound oligonucleotide (Figure 3). Collectively, these three steps are referred to as selection cycle. The sequences of the eluted oligonucleotides can be determined at that point, but more commonly the selection cycle is repeated one or more times prior to sequencing so as to isolate the sequences that bind most tightly to the target molecule.

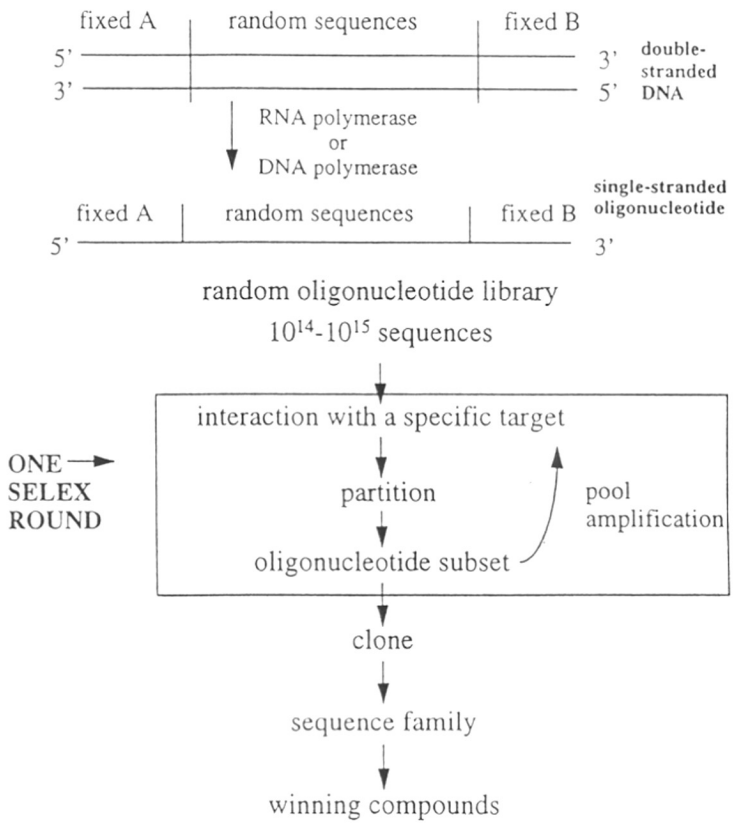


Figure 3. Schematic representation of SELEX procedure. Fixed A and B refer to defined sequences present on each member of the library that flank the random region. Fixed sequences permit amplification and transcription. The steps in the SELEX process are further described in text. Source: Gold et al., 1995

The bound oligonucleotides can be amplified by PCR following some or all of the selection cycles. Since it is common for more than one oligonucleotide sequence to be selected from among a complex library, these oligonucleotides are commonly identified by cloning and sequencing of a number of individual clones (Bartel and Szostak, 1993).

***Library size*:** Nucleic acids are synthetically randomized by condensing a mixtures of activated monomers over some specified number of positions (n), where the sequence diversity for a nucleic acid with four different bases is n^4. The availability of novel, additional base pairs could allow even more random sequences over the same length (Switzer et al., 1993). For SELEX experiments reported in literature, the number of molecules in the starting library has been 10^{14} -10^{15}, but if necessary this number could be higher. Large lengths of randomness quickly provide sequence diversity potential that dwarfs even the number 10^{20}. In the SELEX experiments reviewed, the winning ligands were most often found to represent between 1 in 10^9 to 1 in 10^{13} of the starting library (Gold et al., 1995).

***Length of random library*:** Random regions of roughly 30 nucleotides are used in many of the first experiments. The known motifs in single stranded oligonucleotide such as hairpins, bulges within helices, pseudoknots, and G-quartets, can be built from 30 nucleotides (Gold et al., 1995). Thus, the major structural framework for single-stranded oligonucleotides are reached easily with randomization over lengths that can be searched thoroughly. The same targets have been used for a few SELEX experiments in which the library had different lengths of randomized sequences; while the same motifs and sequences emerged, it seems likely that large and very precise structural motifs will emerge idiosyncratically if long randomized domains are used.

***RNA, DNA, or Non-natural Single-stranded Oligonucleotides*:** The first SELEX experiments were done with RNA (Ellington and Szostak, 1990). Since then, numerous experiments have been reported using RNA, DNA, and modified nucleotides (Gold et al., 1995). Binding affinity and specificity do seem

correlated with the number of SELEX rounds performed as observed from the small number of experiments done in which weakly bound ligands were identified after a small number of SELEX rounds. For many purposes, the winning ligands should be nuclease resistant, and modifications toward that end can be incorporated into the entire library (Irvine et al., 1991; Lin et al., 1994). High affinity, high specificity ligands have been found no matter what monomers have been used to make the library.

1.4.4 Uses of Molecules Derived From SELEX

SELEX-derived oligonucleotides may be an alternative to antibodies for research purposes. The common step of preparing antibodies to a new protein for intracellular localization experiments could be complemented or replaced with SELEX-derived reagents. Moreover one could modify oligonucleotides with visualisation-enhancing adducts and reporters. Similarly, SELEX-derived reagents could be used for affinity purification of proteins. One major effort since the invention of selex has been aimed at direct use of oligonucleotides as drugs. For example, a target for oncology has been angiogenesis, the process by which most tumors recruit new vasculature for the required supply of nutrients. SELEX has been used to make antagonists that block angiogenesis.

A number of successful selections of RNA or DNA oligonucleotides have been reported against targets ranging from low molecular weight organic molecules to proteins and oligonucleotides. Many of the oligomers identified through combinatorial approach have high affinity for their target, usually in nanomolar range (Pei et al., 1991).

***Small Molecules*:** In an early success with *in vitro* selection, Ellington and Szostak (1990) utilised multiple cycles of amplification and selection to isolate RNAs that were highly specific for various low molecular weight organic dyes immobilised on an appropriate substrate. It is notable that although average length of randomised sequence was 100 nt, the region that interacted with dyes appeared to be related to 20-30 nt folding into the predicted patterns of

secondary structure. This suggests that significant reduction in size could be achieved with a retention of affinity and specificity.

***Nucleic Acid-Binding Proteins*:** Oligonucleotide selections have been used to define recognition sites for known nucleic acid binding proteins. Even when individual binding sites are already defined, selection procedures might help in determining the bases critical for interaction. Examples of successful application of this technique include identification of common consensus sequence recognised by the helix-loop-helix proteins myoD, E2A (Blackwell & Weintraub, 1990; Blackwell et al., 1990); and elucidation of the role of the *E-coli*-35 promoter sequence motif in transcription initiation (Horwitz & Loeb, 1988).

***Antibodies*:** Some oligonucleotide selection experiments have provided valuable basic information towards therapeutic opportunities. For example, Tsai et al. (1991) utilised an antiserum against 13 amino acid peptide to screen a pool of RNAs chosen to represent an auto-recognition site in U1 snRNA. The oligonucleotide was designed to consist of ten-nucleotide loop within a constant stem region. A specific RNA sequence was identified that could compete with peptide for the antigen combining site. Such oligonucleotides could have therapeutic utility in sequestering autoimmune antibodies.

***Thrombin*:** Due to its pivotal role in blood coagulation, inhibitors of thrombin have potential applications as anti-coagulants. Using single stranded DNA selection technology, a 15 mer sequence has been identified which binds to thrombin and inhibit thrombin catalysed blood clot formation (Bock et al., 1992). These results were extended to an *ex vivo* whole artery angioplasty in rabbit model where the oligonucleotide was shown to suppress significantly the formation of fibrinopeptide A and reduce platelet deposition (Li et al., 1994).

***HIV Inhibitors*:** RNA selection has been used to identify a psuedoknot inhibitor of HIV reverse transcriptase. One of the sequence demonstrated binding to target at nanomolar concentrations and exhibited specificity for HIV-RT over reverse transcriptase from other sources (Tuerk et al., 1992). Beutel et al.,

(1994) have identified RNA oligonucleotide that bind tightly to HIV protease and inhibit enzyme activity.

***Triplex Forming Oligonucleotides*:** A combinatorial approach has been used to identify individual RNA molecules from a large population of sequences that bind 16-base pair homopurine-homopyrimidine DNA sequence through triple-helix formation (Pei et al., 1991). The identified molecules were highly homologus to target sequence and in addition contained hairpin loops, interior loops and nonstandard base triplets at various positions. Systematic variation in both target DNA sequence and buffer components should provide increased insight into the molecular interactions required for triple-helix mediated recognition of natural DNA.

***Taq Polymerase Inhibitors*:** Dang & Jayasena et. al. (1996) have identified the oligonucleotide sequences which bind specifically to *Taq* polymerase. These oligonucleotides are utilised as reagents for "hot start" PCR. At low temperature during PCR, these oligonucleotides bind to enzyme and thus prevent the nonspecific extension by polymerase. The same oligonucleotides dissociate from enzyme at higher temperatures where specific amplification of target can occur.

1.4.5 Prospects

A successful combinatorial chemistry methodology should rapidly provide: (i) enough compounds with very high affinity individual molecules for any target, (ii) high specificity of the emergent ligands towards their cognate targets and weak binding to either related targets or to other potentially interactive molecules in the environment of the intended use, (iii) cost-appropriate, generic methods to use the ligands for the intended environment; (iv) formulation steps that retain ligand activity with high bioavailability and low toxicity and (vi) immediate testing in appropriate animal disease models if the intended use is therapeutic.

A fully successful combinatorial chemistry methodology for therapeutics yields drug candidates that require extensive further experimentation and reiterative medicinal chemistry accompanied by tedious functional and toxicity assays. That is, combinatorial chemist should not seek lead compounds, but

"hits" for structural analysis and further development. Combinatorial chemistry technology may provide, compounds for direct animal testing in animals, and finally in humans, the clinical outcome is all that matters. SELEX may meet most, if not all, of the criteria for a robust, generic drug discovery technology (Gold, 1995).

1.5 OLIGONUCLEOTIDES AS THERAPEUTICS AGENTS

Many organisms function through transfer of genetic information from DNA to RNA to protein. In the case of retroviruses, the information is stored in viral RNA. Novel approaches to therapeutics, involving intervention in these processes using oligonucleotides, have been developed in past few years (Figure 4). These have the capability to replace the standard approach to the development of pharmaceuticals. In effect, this approach represents a completely novel strategy, that is in some ways intermediate between classical drug paradigm of small molecule (lock and key) inhibitors, and gene therapy that involves the introduction of whole genes into affected individuals.

1.5.1 Antisense Strategy

The most widely applied approach to interrupt the flow of genetic information is to use oligonucelotides as inhibitors of translation. The complementary or 'antisense' base sequence is targeted to a specific 'sense' sequence in the mRNA. Thus, expression of specific protein can be regulated or inhibited. It should be noted that antisense mRNAs occur naturally as a regulatory mechanism (Inouye, 1988; Krystal, 1990). Mechanisms of antisense inhibition includes interference with ribozyme binding and processing of mRNA (Liebhaber, et al., 1991), interference with mRNA conformation (Vickers, et al., 1991) or mRNA splicing (Kole et al., 19891), and RNase-H activation of mRNA digestion (Walder and Walder, 1988).

The preferred target of antisense inhibition is the 5'-initiation codon. In general, an oligonucleotide of ~15-20 bases in used for antisense inhibition, since this should approximate to only one unique target sequence in the human

genome. In addition, an oligomer of this length hybridizes well with its complementary target mRNA.

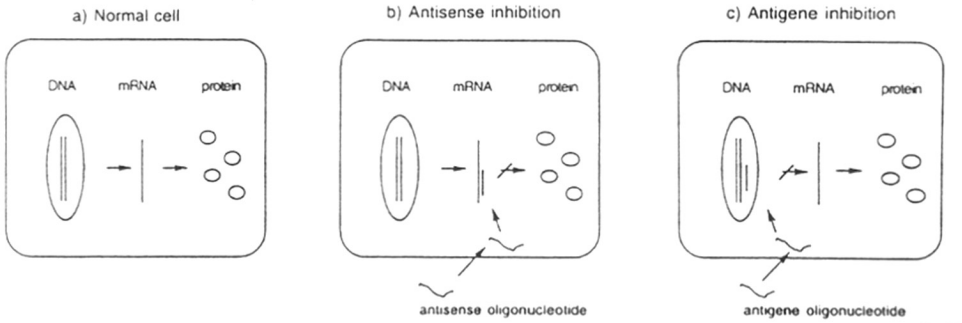


Figure 4. Schematic illustration of (a) Normal gene expression. DNA is transcribed into mRNA followed by translation to give multiple copies of product; (b) Antisense inhibition. An antisense oligonucleotide binds specifically to mRNA via Watson-Crick hydrogen bonding wherby it inhibits translation of mRNA into protein; (c) Antigene inhibition. Transcription is inhibited by the binding of an antigene oligonucleotide to DNA by triplex formation or strand displacement.

1.5.1a Requirements for Therapeutic Use of Antisense Oligonucleotides

The use of antisense oligos as therapeutic agent presupposes that six criteria can be satisfied (Stein & Cheng, 1993).

***The oligos can be synthesised easily and in bulk*:** The development of phospharamidite chemistry (Beaucage & Caruthers, 1981) and its elaboration into an automated technology have greatly enhanced the ease with which oligos are synthesised and consequently their availability. Methods for large-scale oligo synthesis are being commercially pursued. Although the cost has been dramatically reduced, the final cost to the consumer of a "treatment" with oligo has not been determined.

***The oligos must be stable in vivo*:** This precludes the use of phosphodiester oligos as therapeutics because serum and intracellular nucleases will degrade them. The significant effort has been extended by synthetic chemists to develop nuclease resistant oligos. Perhaps the greatest success in this direction have been achieved with phosphorothioate and methylphosphonate oligos, which can be synthesised with relative ease.

***The oligos must be able to enter the target cell*:** The ability of oligos to penetrate the cell membrane and the mechanisms of entrance are critical

considerations in developing these agents as therapeutics. Unlike many other antineoplastic agents, which are of low molecular weight and hydrophilic, oligos (with the exception of methyl phosphonate types) are polyanionic and cannot passively diffuse across cell membranes. Despite this, phosphorothioate oligos are internalised within cells and many examples of sequence specific mRNA translation exist (Stein & Cheng, 1993). The recent work on covalent linkage of spermine to the nucleobase renders extremely good triplex forming ability of the oligos and has potential for cellular uptake due to the interaction DNA-spermine conjugate with polyamine receptors and the neutralisation of the net charge on the oligos (Barawkar et al., 1996; Rajeev et al., 1997).

***The oligos must be retained by the target cell*:** Oligos undergo exocytosis (Yakubov et al., 1989) from H9 cells (Crooke, 1991), and HL60 cells (Stein and Cheng, 1993). Oligo chain extension has been seen in oligos extracted from the kidneys and intestines of mice after administration of PS oligos (Agarwal et al., 1991). The origins and effects of this process on the antisense inhibition are unknown. The question of oligo exocytosis from the target cell must also be considered in the context of the rate of hybridization of the oligo to its target mRNA. If these rates become comparable, then the ability to produces the desired antisense effect may be compromised.

Modification of an oligonucleotide by prolyl-L-lysine, transferrin-polylysine, or a cholesteryl (chol)-moiety conjugated at the 5' terminus can increase its net intracellular internalization or retention or both (Leonetti et al., 1990; Citro, et al., 1992; Letsinger et al., 1989; Boutourin et al., 1989). The poly-L-lysine modification masks the negative charge of the oligo, preventing it from interacting with nonspecific cellular binding proteins. This may help to destabilise endosomal membranes and thus permit transit of oligo into the cytoplasm. On the other hand, 5'-cholesterol oligos can bind to low density lipoprotein (LDL) and may be internalized, at least partially, by means of the LDL receptor (Kreig et al., 1993). Intracellular internalization by adsorptive endocytosis may also be increased by a 5' hydrophobic modification. This may lead to improved antisense

efficacy: the 5' cholesterol modified (as opposed to unmodified) phosphorothioate oligos are effective antisense inhibitors of the expression of the MCF-7 murine endogenous retrovirus envelope protein in murine spleen cells (Kreig et al., 1993). However, the cholesterol modification tends to increase cellular cytotoxicity, perhaps by promoting extracellular calcium influx (Saxon et al., 1992).

***The oligos must be able to interact with their cellular targets*:** Intracellular targets for oligos (such as mRNA, pre-mRNA, and genomic DNA) are invariably protein bound, and many sites are probably not accessible for Watson-Crick base pairing. The location of the best targets for the antisense mRNA approach has been well discussed by Goodchild (1989). One approach, used to inhibit globin translation in the reticulocyte lysate system, has been to target the 5' cap or initiation codon (AUG) regions. However, in HL60 cells the 5' cap region of the c-*myc* mRNA is a significantly better target than the AUG region and about as useful as the first splice junction site (Bacon and Wickstrom, 1991). The most useful target, as judged by the ability of the antisense oligo to decrease p21 Hras protein production, was the 5' cap of the mRNA. Contradictory results from different cellular systems exemplify the intersystem variability in oligo technology and the need to avoid overgeneralization.

***The oligos should not interact in a non-sequence-specific manner with other macromolecules*:** Although this statement is self evident, the fact that charged oligos are polyanions, is overlooked. Naturally occurring sulfated polyanions, such as glycosaminoglycans, heparan, dermatan and chondroitin sulfates, play several vital physiological roles. For example, these polyanions can bind to and sequester tumour (heparin-binding) growth factors. Glycosaminoglycans are also potent anti-angiogenesis agents. Several proteins also bind to both charged oligos and sulfated polyanions. However, alteration of the polyanionic backbone of a charged oligo can alter the binding constants to protein, as occurs on phosphorothioate substitution and produce non-sequence-specific effects.

1.5.1b Modified Oligonucleotides: Better Antisense Agents

Early on, it was anticipated that there would be significant limitations with the use of natural DNA as antisense therapeutics. As a result, a significant number of modifications of the oligonucleotide structure have been made (Sanghvi and Cook, 1994). Figure 5 shows a partial structure with examples of the positions in which modifications have been synthesized for testing their effects on hybridization and nuclease resistance, and in some cases, pharmacological activity.

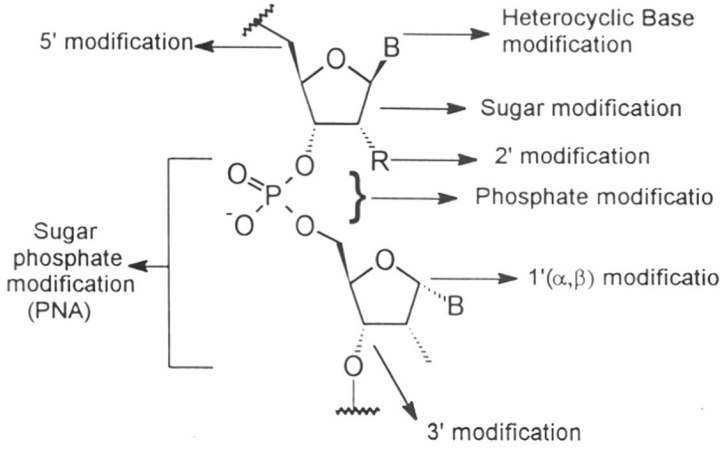


Figure 5: Position and types of oligonucleotide modifications.

Many modifications have been performed on the nucleobase. Among the more promising are substitutions at C5 position of pyrimidines, resulting in increased hybridization affinity and maintaining recognition of the heteroduplex by RNase H. In particular, 5-propynyl C and T substitutions increase the T_m of an oligonucleotides by as much as 2°C per modification and still support RNase H activity (Sagi, et al., 1993). The more pyrimidine substitutions that are made in a given oligonucleotide, the more benefit will be derived. The 5-propynyl substitutions do not provide any significant nuclease resistance; therefore, the internucleosidic linkage must still be protected by modifications such as phosphorothioate substitutions. A number of modifications on the 2' position of the sugar have resulted in increased affinity for RNA and in some cases increased nuclease resistance (Martin, 1995). However, oligonucleotides

modified with current 2' substitutions such as O-methyl, fluoro, O-propyl, O-allyl and methoxyethoxy do not support RNase H activity. Therefore, either non-RNase H mechanisms must be exploited to fully capitalize on the benefits of such oligonucleotides, or alternatively, a chimeric strategy must be used to retain some RNase H activity (Wagner, et al., 1993). Modifications with phosphate backbone have been made with the aim of increasing nuclease resistance, increase affinity, and possibly enhancing the pharmacokinetic properties of the oligonucleotides. A number of modifications on either the phosphate backbone or substitution of the phosphate groups have been made and are providing promising biophysical and biological activity. Methylphosphonate oligonucleotides eliminate the charge on the backbone but introduce a chiral center. Although much has been written about the benefits of methyl phosphonate substitutions (Miller, 1991), they have not achieved widespread use. This may be due in part to not supporting RNase H activity and the low affinity of chirally impure oligonucleotides. Recently, chirally pure dimers of methyl phophonates have demonstrated marked increased affinity when multiple dimers are incorporated into oligonulceotides (Arnold, 1995).

A number of modifications have been made in which the phosphate has been replaced with achiral linkages. In particular, the methyleneimino,and thioformacetal linkages are very promising (Jones, et al., 1993). In each case, the linkages support hybridization and may even increase affinity and provide nuclease resistance; they are also compatible with existing phosphoramidite chemistries. Currently these substitutions are prepared as dimers and incorporated into oligonucleotides with a phosphate linkage between each dimer. Like the 2' modifications, replacement of the phosphate linkage results in loss of RNAse H activity; therefore, similar strategies will need to be employed to take full advantage of them.

More radical substitutions have been made in which the sugar and phosphate residues were replaced with an alkylamide linkage (Nielsen, et al., 1991). Surprisingly, this modification, peptide nucleic acid (PNA), increased

affinity for both RNA and DNA and preserved specificity. Furthermore, it is not susceptible to either nuclease or peptidase cleavage. Similarly morpholino nucleoside oligomers with a carbamate internucleosidic linkage have been prepared that preserve hybridization. Finally, various functional groups that have been covalently attached to oligonucleotides to change their physical properties, provide ligands for interacting with specific receptors, and provide additional nuclease resistance.

Table 1. Properties of selected phosphodiester backbone analogs

backbone analogue	activation *	resistance #
phosphodiester	yes	[a]
phosphorothioate	yes	+
methylphosphonate	no	++
phosphoramidate	no	+
peptide nucleic acids	no	++
3'thioformacetal	no	++
methylene(methylimino)	no	++
3'-N-carbamate	no	++

* activation by RNAse H., # nuclease susceptibility, [a] -, rapidly degraded by nucleases; +, resistant to nucleases; ++, no nuclease degradation. Source: Milligan et al., (1993)

One of the most efficient mechanism of antisense oligonucleotides require specific hybridisation of the oligonucleotide to its complementary RNA sequence followed by cleavage of the mRNA strand of RNA/DNA duplex by RNase H. As unmodified DNA oligomers lack stability in the presence of serum nucleases, investigations have been focused on designing more stable oligonucleotides (See Table 1). Although many of these analogs are resistant to nucleases and interact with mRNA, only natural phosphodiesters and phosphorothioates induce RNase H cleavage. In order to overcome this problem, chemically or photochemically activable reagents are developed which irreversibly modify the mRNA target and hence inhibit translation even in absence of RNAse H activity (Godard et al., 1994). Although each of these modifications have improved the properties of the current first-generation phosphorothioate oligonucleotides, it is unlikely that any single chemical modification will solve all the issues. Therefore,

a combination of approaches will probably have to be undertaken to provide the optimum benefit chemical modifications have to offer.

1.5.1c Long-Term Therapeutic Potential

The long-term therapeutic potential of antisense oligonucleotides is not defined yet, and it is still possible that with a more sophisticated understanding of this technology, researchers will conclude that the technology has an as yet unidentified fatal flaw. Nevertheless, the progress to date provides justification for continued exploration and optimization. It will be several years before the therapeutic value of phosphorothioate oligonucleotides is understood and hopefully, many years before advances in medicinal chemistry of oligonucleotides no longer result in new classes of antisense drugs with improved therapeutic properties and utility.

1.5.2 Antigene Strategy: Transcription Inhibition

By targeting DNA, transcription arrest by an oligonucleotide would result in direct inhibition of gene function (Figure 4b). However, the DNA duplex is very stable compared with single-stranded mRNA which, although extensively folded, is readily accessible. Use is therefore made of the fact that a third strand can be accommodated in the major groove of B-form DNA duplex to form a triplex structure.

The triple helix formation requires a polypyrimidine strand (third) which forms Hoogsteen base pairing with polypurine strand of Watson-Crick duplex (Figure 6). Two important features of this motif are that (i) while A:T base pairs are recognised by neutral T, G:C Watson:Crick pair requires protonated C for triad formation with full complementation and (ii) a purine must occupy the central position of triad. These limitations have led to exploration of base pair recognition diversity in nucleic acids to accomodate other Watson:Crick base pairs.

From the therapeutical point of view, the modifications are being made in oligonucleotides which can form stable and specific triplexes under physiological

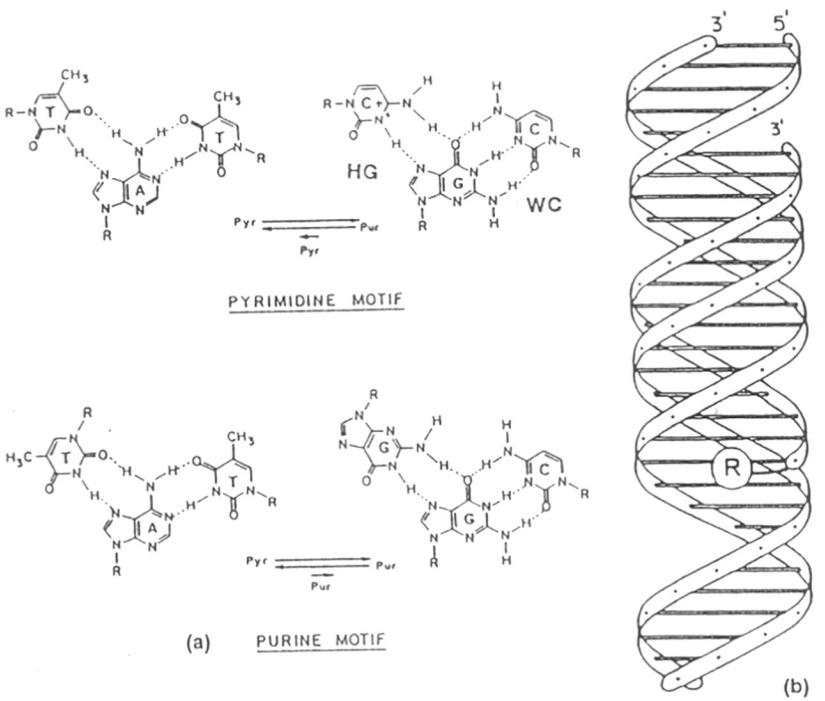


Figure 6. Hydrogen bonding in triple helix structures. Third strand is parellel to purine strand in pyrimidine motif and antiparellel in purine motif. (b) Triplex structure showing the hybridisation of third strand. Third strand can be conjugated with different moieties ® which can cleave DNA in close vicinity.

conditions with duplexes of wide sequence variety (Barawkar et al., 1996; Best & Dervan, 1995). The other criterions like cell permeability, nuclease resistance etc. are simillar to those for antisense reagents which are described earlier. Another approach to transcription arrest is to sequester transcription factors that recognize specific base sequences by addition of the recognition sequence as an oligomer. It is anticipated that the oligomer will avidly bound and saturate the transcription factors, preventing its normal function. An example of the inhibition of transcription using a short duplex phosphorothioate oligomer has been reported (Bielinska et al., 1990).

1.5.3 Ribozymes

A further area where oligonucleotides are emerging as potential drugs is that of ribozymes. Ribozymes cleave their target using either transesterification or hydrolysis mechanism (Cech et al., 1981). By suitable chemical or molecular manipulation, ribozymes can be engineered to either bind specifically to external desired sequence targets and cleave them, to inhibit a gene function, or to ligate new pieces of RNA onto the target by trans-splicing to create a new gene function. Therein lies their therapeutic potential (Christoffersen & Marr, 1994).

While ribozymes have received a great deal of attention, including an application to HIV, few parameters need to be satisfied for its success. As ribo-oligomers, they are more difficult to synthesize and are susceptible to degradation by RNAses. This is a significant challenge to establish ribozymes as human therapeutics. They can be chemically modified (e.g., with phosphorothiotes or with 2'-O-methyl groups) but the effect of these modifications on their catalytic activity is uncertain (Yang et al., 1992). To overcome these problems the delivery of ribozymes through retrovirus based vectors is becoming popular. These vectors transfect and express therapeutic genes in disease susceptible cells (Samulski, 1993).

Therapeutic applications of ribozymes are potentially quite broad but have so far been applied to situations involving inhibition of overexpression of gene. The gene target may be foreign, as in a viral infection, or may be a normal gene which has undergone mutation such as an activated protooncogene. (Christoffersen & Marr, 1995).

1.6 OLIGONUCLEOTIDE ARRAYS

Hybridization is a hydrogen-bonding interaction between two nucleic acid strands that obey Watson-Crick complementarity rules. All other base pairs which do not follow this rule are mismatches and destabilize hybrids. Since a single mismatch decreases the melting temperature (T_m) of the hybrid by upto 10 $^{\circ}C$ (Wallace et al., 1979), conditions can be found in which only perfect hybrids survive. Hybridization comprises contacting the strands, one of which is usually

immobilized on a solid support and the other usually bears a radioactive or fluorescent label. After hybridisation, the unreacted labelled strands are separated from hybridised ones by washing the support. Hybrids are recognized by detecting the label bound to the surface of the support (Bains & Southern, 1988; Khrapko et al., 1989).

Oligonucleotide hybridization is widely used to determine the presence in a nucleic acid of a sequence that is complementary to the oligonucleotide probe. In many cases, this provides a simple, fast, and inexpensive alternative to conventional sequencing methods (Sanger et al., 1977; Maxam and Gilbert, 1977). Hybridization does not require nucleic acid cloning and purification, chemical reactions, or tedious elctrophoretic separations. Hybridization of oligonucleotide probes has been successfully used for various purposes, such as the analysis of genetic polymorphisms (Saiki et al., 1989), diagnosis of genetic diseases (Conner et al, 1983), cancer diagnosis, detection of viral and microbial pathogens, screening of clones, genome mapping, and the ordering of fragment libraries (Chetverin and Kramer, 1994).

An oligonucleotide array is comprised of a number of individual oligonucleotide species tethered to the surface of a solid support in a regular pattern, each one in a different area, so that the location of each oligonucleotide is known. These arrays can be prepared by parellel synthesis of all oligonucleotides, directly on the support, in combination with site-directed masks (Southern, 1989). Such masks direct a particular nucleotide monomer (A, T, C or G) to react with a predetermined exposed area on the surface of the support. Four masks with non-overlapping windows and four coupling reactions are required to increase the length of the tethered oligonucleotides by one. In each subsequent round of synthesis a different set of four masks is used, and this determines the unique sequence of the oligonucleotides synthesised in each particulate area.

Thus, a miniature array can contain a large number of oligonucleotide probes, and all of them can be used simultaneously for hybridization to a nucleic

acid sample in one experiment. This greatly reduces the time required for analysis and eliminates the need for the costly synthesis of individual oligonucleotides. An array may contain a chosen collection of oligonucleotides, e.g., probes for all known clinically important pathogens or specific probes for all known sequence markers of genetic diseases (Maskos and Southern, 1993). Such an array can satisfy the needs of a diagnostic laboratory. Alternately, an array may contain all possible oligonucleotides of a given length *n*. Hybridization of a nucleic acid with such a "comprehensive" array results in a list of all its constituent *n*-mers, which can be used for unambiguous gene identification (eg., in forensic studies), for determination of unknown gene variants and mutations, for finding overlapping clones, and for checking sequences determined by conventional methods. Finally, surveying the *n*-mer by hybridization to a comprehensive array can provide sufficient information to determine the sequence of a totally unknown nucleic acid.

1.6.1 Sequencing by Hybridization

Surveying the *n*-mers in a nucleic acid is analogous to listing the words contained in a text. This would not make much sense unless we know how the words are connected. Fortunately, unlike common words, the *n*-mers in a nucleic acid strand overlap one another so that each non-terminal *n*-mer includes the last *n*-1 nucleotides of the preceding *n*-mer and the first *n*-1 mer nucleotides of the next *n*-mer. This allows the surveyed *n*-mers to be assembled by overlapping them at their (*n*-1)-long subsequences [(*n*-1)-mers] in order to reconstruct the sequence of the analyzed nucleic acid (Figure 7). This strategy for sequence determination is called "sequencing by hybridization" (SBH). SBH can surpass the conventional sequencing procedures in a number of a parameters, including speed, cost, quality of the results, and ease of automation. Test sequencing of ≈100 nucleotide-long DNA strands by hybridization with octamers has demonstrated that the method is feasible and tolerant of occasional hybridization errors (Drmanac et al,. 1993).

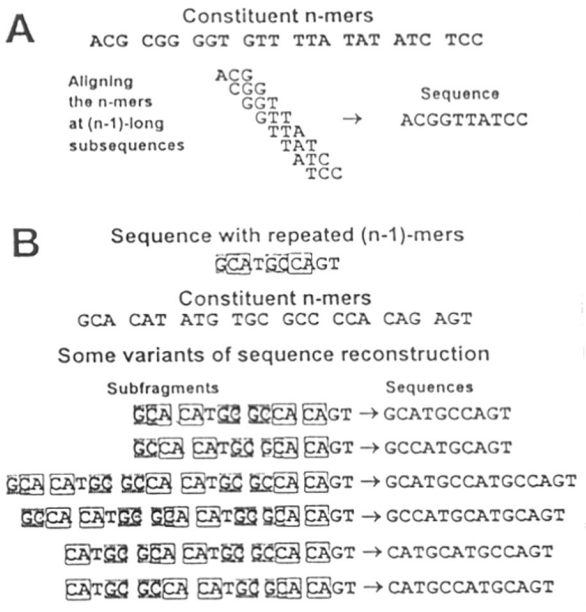


Figure 7. Principle of sequencing by hybridisation. (A) Reconstruction of a nucleic acid sequence from the list of its constituent of n-mers identified by its hybridisation to a comprehensive set of oligonucleotide probes. (B) If repeated (n-1)-mers are present in a sequence, it cannot be reconstructed unambiguously. Assembly of the n-mers result in multiple subfragments that can be permuted and/or repeated in different ways. In these examples, n = 3 in order to simplify the illustration. Crosshatches and boxes indicate repeated (n-1)-mers.

However, there is an inherent flaw in the SBH method that undermines its advantages. SBH relies exclusively on short range information provided by the sequences of the surveyed *n*-mers, and success in assembling the *n*-mers is absolutely dependent on whether or not their (*n*-1)-long overlaps are unique. success is dependent on whether or not there are repeated (*n*-1)-mers in the nucleic acid being analysed. Strand reconstruction terminates at non-unique (*n*-1)-mers and the resulting subfragments can be permuted and/or repeated in many different ways without conflicting with the hybridization data. When $n = 8$, only 94%, 32% and 0.9% of random sequences of 50, 200 and 400 nucleotides in length, respectively can be reconstructed unambiguously (the remaining sequences contain repeated heptamers). The situation is even worse with natural nucleic acids, since they usually contain more repeats than do random sequences. Utilising longer probes would reduce the ambiguities, but would

result in an exponential increase in cost. For example, if *n* is increased from 8 to 12, then the length of random strands that can be sequenced with 95% success increases from 47 to 666 bases (14 fold), whereas the number of probes required increases 256 fold.

1.7 MAPPING OF GENOMIC DNA

The primary goal of the Genome Program is to make a series of descriptive maps of each chromosome at increasingly finer resolutions. A single chromosome is cut by some rare-cutting agent into large pieces, which are then ordered and subdivided. The smaller pieces are then mapped further. Site-specific binding of triplex-forming oligonucleotides may form the basis of rare cutting of chromosomal DNA (Dervan, 1992).

1.7.1 Triplex Assisted Genome Mapping Strategies Using Endonucleases

Fragmentation of long chromosomal DNA for mapping and, in particular, sequencing purposes by naturally occurring restriction enzymes is sometimes inappropriate because of their too frequent cleavage sites. This is due to the relatively short recognition sites for most restriction nucleases. Therefore, alongside the discovery and application of new rare-cutting restriction endonucleases such as *Not*I and *Mlu*I (Hanish and Mcleland, 1990; Billings et al., 1991), efforts are being made to make some restriction sites more rare and to develop different kinds of synthetic endonucleases (Sigman, 1990; Dervan, 1992).

Transient site-specific protection from methylation by a site-specific protein leaves only rare recognition sites unmethylated for the restriction enzymes (Koob et al., 1988). When the protecting protein dissociates, only such unmethylated "Achilles heel" sites are susceptible to the action of corresponding endonucleases, thereby providing much rarer cleavage of DNA. The other recognition sites are not cleaved, due to methylation. In the oligonucleotide equivalent of the "Achilles heel" methodology, protection of restriction sites against methylation is accomplished by the triplex-forming oligonucleotide (Strobel and Dervan, 1992). Subsequent triple-helix disruption and cleavage by

a restriction enzyme yields larger DNA fragments (Figure 8). By preventing *Eco*RI methylase reaction at a specific site with a triplex-forming oligonucleotide, it was possible to cleave the yeast chromosome III (340-kb size) with *Eco*RI endonucleases at a single site.

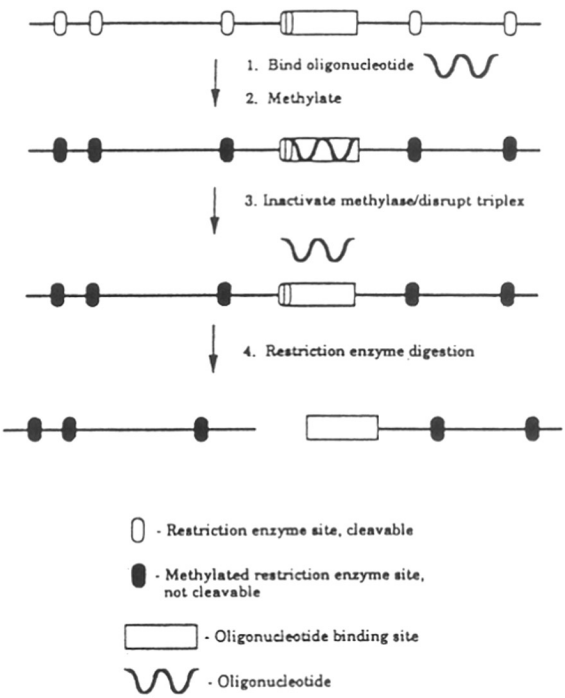


Figure 8.A generalised scheme for single-site restriction cleavage of genomic DNA by oligonucleotide directed triplex formation. The triplex between chromosomal DNA and oligo is formed in a methylase-compatible buffer. *Eco*RI methylase methylates all adenines in DNA except the partially triplex protected site 5'-GAATTC-3'. After triplex disruption, all recognition sites for *Eco*RI are resistant to cleavage, with only the temporarily triplex-protected site being susceptible. (Strobel & Dervan, 1992).

As an extension of this methodology, Nehls et al. (1994) suggested introduction of triplex-forming tracts, which overlap *Eco*RI restriction site, into mammalian genome via homologous recombination. In such a way, large chromosomal domains may be created for cloning, sequencing and functional analysis.

One of the promising directions in the development of artificial endonucleases is the use of triplex-forming oligonucleotides carrying DNA-cleaving groups. The ability to vary the length of the recognition site and control over the stability of complex stability can provide targeting oligonucleotides at

distant sites and produce DNA fragments of any desired size. Even a simple analysis of randomly chosen sequences of 25 human genes shows that oligonucleotides of 9 bases and longer can be used to produce chromosome fragments of several dozen kilobases in length. For example, an arbitrary 9-*mer* consisting of the bases of four types is statistically expected to occur once in a sequence 32Kb in length. This is the average size of fragments that can be produced with triplex-binding 9-*mer* bearing the DNA-cleaving group. Because of the nonrandom base distribution in the DNA sequences, for two-thirds of 9-*mers* found, the frequency of repeats was less than one in 68 kb, so they could be used to produce fragments longer than 68 kb.

Another method makes use of the specific feature of triple helical complex of DNA with an oligomer peptide nucleic acid (PNA). Formation of the PNA/DNA triplex results in a displacement of the one strand which is recognized and digested by a single-strand specific enzyme (e.g., S1 or mung bean nucleases) (Demidov et al., 1993). Although the yields of enzymatic reactions are greater than 75%, they do not result in point cleavage and the lengths of the affected sites are 5 to 10 bp. This may be due to excessive freedom of the targeted staphyloccocal nuclease (Pei et al., 1990) or the presence of a long looped part of the DNA (Demidov et al., 1993).

1.7.2 Triplex Mediated Genome Mapping Strategies using Endonuclease

These type of strategies involve covallenet attachment of DNA cleaving moeity to the third strand of DNA. The third strand places the nuclease moeity at specific site after which site specific cleavage is achieved (Figure 6). Depending upon the type of group which is attached to third strand for nuclease action, these strategies can be separated mainly into three groups.

(i) In the first (enzymatic) group, the oligonucleotides-directed recognition of DNA is coupled with the enzymatic cleavage of the latter. This can be accomplished via an attachment of the DNA-cleaving enzyme (e.g., staphylococcal nuclease) to the oligonucleotide (Pei et al., 1990). Specific

double-stranded cleavage occurred at AT-rich sites at the 5' and of the P_yP_u tract for both 5'-mono- and 5'-3'-diderivatized oligomers.

In the second (chemical) group, DNA-binding molecules are combined with reactive functionalities that result in the oxidation of the deoxyribose (Strobel and Dervan, 1990) or the electrophilic modification of the bases (Povsic et al., 1992). Among the best known reagents are those capable of generating reactive species, such as the hydroxyl radical. For instance, metal complexes (e.g., Fe-EDTA and Cu-phenanthroline) are used to generate diffusable HO· radicals which attack the deoxyriboses within one of the double helix grooves. The DNA molecule may be cleaved either in the course of the oxidative reaction or during additional chemical treatment. DNA at oxidized sites can also be cleaved enzymatically. As a result of sufficient mobility of the cleaving moieties coupled to the oligonucleotide ends and a spatial diffusion of hydroxyl radicals, oxidative chemical cleavages occur at the lengths of several base-pairs. Reaction yields are relatively low, 25% or less (Strobel and Dervan, 1990). For Cu-phenanthroline, however, 70% cleavage efficiency was reported (Francois et al., 1989).

The third (photochemical) group includes DNA-binding molecules that carry photoactivable moieties in order to produce either direct or indirect strand breaks (that is, appearing after chemical or enzymatic treatment (Dobrikov et al., 1992). The photochemical reaction is initiated by long-wavelength (λ>300 nm) irradiation, which does not affect nucleic acids themselves. A number of photoactivable groups were tested. The photocrosslinks produced by azidoproflavine, azidophenacyl, and psoralen (Le Doan et al., 1990, Praseuth et al., 1988, Takasugi et al., 1991) can be converted into strand breaks under alkaline conditions. The yield of the photoinduced crosslinking reaction can be quite high - about 80% (Takasugi et al., 1991; Dobrikov et al., 1992). Ellipticine derivatives covalently attached to oligonucleotides were used for photoinduced cleavage of the two strands of a target P_yP_u sequence (Perroualt et al., 1990; Le Doan et al., 1991). The drawbacks of such photoinduced cleavages are that they

are not 100% complete, and due to "diffusable" active groups, photodamaged sites spanned several base pairs. In addition, some of the above-mentioned agents produce multiple effects: for instance, porphyrins induced crosslinking reactions and oxidation of guanine bases, whereas ellipticine induces both crosslinks and strand scissions (Le Doan et al., 1990, 1991).

Synthetic oligonucleotides that contain photosensitizing nucleobase analogs (e.g., halogen- or thiopurines and pyrimidines) were suggested for modifying DNA at specific sites. Since halogenated pyrimidines have the tails of absorption spectra spanning much further into long-wavelength UV region than common bases, and thiopurines and pyrimidines have absorption maxima at 330 to 430 nm, such base analogs can be activated by long-wavelength ($\lambda > 300$ nm) light (Rahn, 1992). Once incorporated into the triplex forming oligonucleotides, under UV irradiation these analogs will induce photomodifications at specific DNA sites. Among these modifications are photodimers, strand breaks, and crosslinks with the target (Rahn, 1992). This approach seems to have some important advantages over other photosensitizers. First, nucleobase analogs contribute to complex stability, as they form hydrogen bonds with the DNA target. Moreover bromination of pyrimidines enhanced the affinities of the oligonucleotides to their double-stranded target (Povsic and Dervan, 1989). Second, one or several preferable photoactive base analogs can be introduced into desirable sites of the same oligonucleotide. Third, a nondiffusable photoactive moiety is expected to produce a highly localized effect (Sugiyama et al., 1990).

A recent report (Panyutin and Nuemann, 1994) describes the formation of double-stranded breaks (DSBs) induced by ^{25}I-labeled oligonucleotides targeted to a P_uP_y tract in the *nef* gene of the human immnodeficiency virus. The DSBs found only under triplex-forming conditions were produced with a high efficiency (0.8 DSB per decay) and were distributed within 10 base pairs of a maximum, located just opposite the position of iodinated cytosine in the oligonucleotide.

1.8 DNA AS MATERIAL FOR NANOTECHNOLOGY

Miniaturization is the thrust area in the field of electronics and there is a constant search for materials that can scale down the size of each component (Niemeyer, 1997). A lot of recent research is being devoted to generate nanometer-scaled structural and functional elements. Biological macromolecules such as proteins and nucleic acids may be used as components for construction of nanostructured systems for which DNA seems particularly suitable. This molecule, a unique information carrier, not only occurs in a large variety of structures but also displays comparably high physicochemical stability and rigidity (Stryer, 1988; Hagerman, 1988).

As an example, DNA might be used in the construction of chips to attain circuit sizes below 100 nm. DiMauro and Hollenberg formulated the concept of construction of DNA networks on suitable solid supports using oligonucleotides such as 1 (Figure 9) as initiation points for enzymatic DNA synthesis or hybridization. These networks may then be used as scaffolds for deposition of conducting materials such as gallium arsenide or indium phosphite by chemical vapor desorption (CVD) procedures. The calculable advantage of DNA technology is based on the precision of DNA biosynthesis and hybridization in the process of matrix generation. The accuracy of photolithographic techniques may therefore be improved by about two orders of magnitude. Further, since the DNA double helix is only 2 nm in diameter, the lower size limit of the structures should be substantially reducible.

The concept of using DNA as a framework for the precise spatial arrangement of molecular devices has recently been applied to the assembly of supramolecular aggregates from DNA-derivatized gold colloids (Robinson and Seeman, 1987; Alivisatos et al., 1996). Such defined arrangements of nanocrystalline metal clusters (quantum dots) are being investigated for new material properties and may have actual applications in the field of laser technology (Weller, 1996).

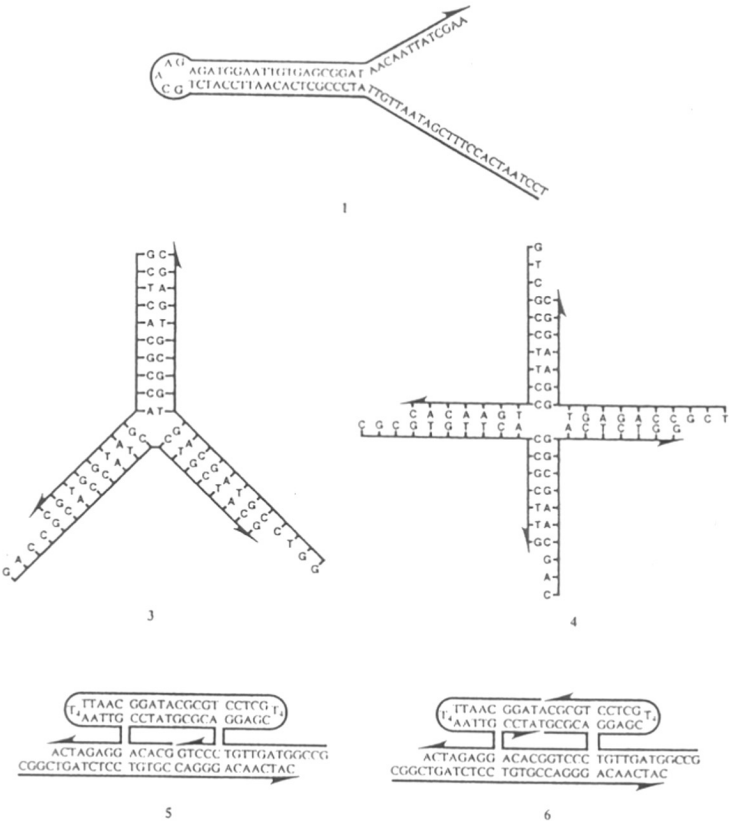


Figure 9. Synthetic DNA motifs for the construction of DNA frameworks: Oligonucleotide **1** with hairpin loop motif used as an initiation point for enzymatic DNA synthesis; three (**3**) and four (**4**) arm junction. DNA crossover molecules **5** & **6**. The arrowheads indicate 3'-end. Source: Niemeyer, 1997.

In the last 15 years Seeman and coworkers (1982, 1987, 1997) have been engaged in the rational construction of DNA frameworks. They have used branched DNA molecules containing three, four and more double helical arms (termed "DNA junctions", **3**, **4**, in Figure 9 and Figure 10). These molecules are similar in structure to the replicational junction and Holliday junctions of genetic recombination (Stryer, 1988). They can be prepared easily by briefly heating stoichiometric mixtures of the oligonucleotide components and then stepwise cooling. The synthesis of periodic matter from DNA was initially attempted with three-arm junctions. Connection of the cohesive end of **3** using DNA ligase yielded a mixture of linear and cyclic oligomers starting with trimer. Since double

helical arms 5-7 nm in length behave like rigid rods, the proportion of cyclized products can lead to conclusions about the variation of the valance angle between the ligation arms of the molecules. Subsequent experiments with four-, five-, and six-arm junctions confirmed the high flexibility of the molecules. Thus these "simple" motifs are not suitable for the assembly of large repetitive constructs.

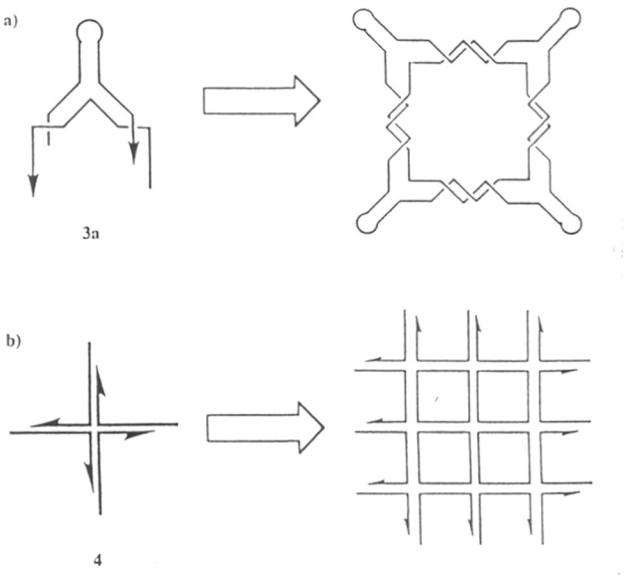


Figure 10. Construction of geometric objects and periodic frameworks from DNA. a) synthesis of a macrocyclic molecule from bivalent tree-arm DNA junctions (**3a**) having two cohesive ends. b) construction of two-dimensional DNA lattices from tetravalent molecules. Linear double helical regions are represented by parellel lines.

The utilization of DNA in biotechnological application for electrical, optical, catalytic, and numeric devices may have additional advantages: DNA can be processed in aqueous media and is suitable for convenient (e.g. biological) decomposition. In addition, the reversibility of the Watson-Crick basepairing is supplemented by the possibility of fixating the structures with selective crosslinking (Woo, et al., 1993). Furthermore, the prospect of high economic profits in nanotechnology justify the current financial expenditure and the recently reported electron transfer in DNA promise further exciting developments in DNA technology (Stemp, et al., 1995).

1.9 PRESENT WORK

The above review describes the widespread applications of synthetic oligonucleotides which span into various fields of science. The work presented in this thesis reports few novel applications of synthetic oligonucleotides in (i) fluorescent DNA detection (ii) combinatorial library for screening restriction enzymes (iii) DNA-restriction enzyme recognition. It also describes the artificial protease activity of Cu(II)desferal complex, a reagent which is useful in triple helix mediated DNA cleavage.

Chapter 1: This chapter gives an overview of the applications of synthetic oligonucleotides. The applications of oligonucleotides in drug discovery, antisense/antigene based therapeutics, sequencing by hybridisation and their role in genome mapping as artificial endonucleases are highlighted.

Chapter 2: This chapter describes the synthesis of fluorescent oligonucleotides by conjugating carboxyfluorescein directly at C-5 position of deoxyuridine through short and rigid amido spacer as well as through long and flexible hexanoyl spacer. The comparative fluorescent properties of these oligonucleotides are studied. The applications of these oligonucleotides are shown in monitoring duplex hybridisation, netropsin-DNA interaction and as primers for the nonradioactive detection of PCR products.

Chapter 3: This chapter describes an algorithm for designing a combinatorial library comprehensively representing all hexamer palindrome sequences at uniquely defined sites. The expected size for such a library of 64 possible hexamer palindromes is 384 bases, which is reduced to 266 bases spread over 8 oligonucleotides through a linear overlap of rationally selected hexamer palindromes. The utility of these duplex oligomers for identifying cleavage sites of restriction enzymes recognizing hexamer palindromes has been demonstrated using some representative enzymes. The sets of oligonucleotides with high information content, though designed for direct and unambiguous characterization of cleavage sites of isolated restriction enzymes, has potential

applications as templates for characterizing sequence selective binding and interaction of small molecules with nucleic acid.

Chapter 4: This chapter describes the synthesis of oligonucleotides containing modified base, 5-NH_2-deoxyuridine to study the interaction of restriction enzyme *Eco*RV with its substrate DNA. The oligonucleotides containing modified base in the recognition sequence of *Eco*RV were found to be refractory for cleavage by enzyme. The effect of such modifications on major groove polarity of DNA was studied using covalently conjugated dansyl probe by systematic replacement of thymine base by 5-NH_2-deoxyuridine. The results suggested that such modification increases the polarity of major groove of DNA which may be the reason for the loss of enzyme activity.

Chapter 5: This chapter describes the artificial protease activity of Cu(II) complex of a widely used drug desferal. By using a combination of gel electrophoresis, fluorescence spectroscopy and amino acid analysis/sequencing it is shown that Cu(II)-DFO complex causes oxidative damage to proteins in presence of the oxidizing agent H_2O_2. The exposure of Cu(II)-DFO/H_2O_2 to proteins resulted in protein fragmentation and amino acid modifications. Lysozyme exhibited a specific fragmentation pattern accompanied by loss of enzyme activity. Since the oxidative damage of proteins is reported to play an important role in aging, the results are of importance due to the utilization of desferal as a drug and related side effects of desferal therapy.

1.10 REFERENCES

Agarwal, S., et al., (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 7595.

Alivisatos, A. P. et al., (1996) *Nature* **382**, 609.

Arnold, L.J., (1995) *Antisense Res. Dev.* **5**, 108.

Bacon, T. & Wickstrom, E. (1991) *Oncogene Res.* **6**, 13.

Bains, M. W. & Southern, E. M. (1988) *J. Ther. Biol.* **135**, 303.

Barawkar, D. A., Rajeev, K. G., Kumar, V. A. & Ganesh, K. N. (1996) *Nucleic Acids Res.* **24**, 1229.

Bartel, R. R., and Szostak, J. W. (1993) *Science* **261**, 1411.

Beaucage, S., & Caruthers, M. (1981) *Tetrahedron Lett.* **37**, 3557.

Bej, A. K., Steffan, R. J. & Atlas, R. M. (1991) *Crit. Rev. Biochem. Mol. Biol.* **26**, 301.

Best, G. C. & Dervan, P. B. (1995) *J. Am. Chem. Soc.* **117**, 1187.

Beutel, B.A., Sherman, M.I., Coppola, G.R. (1994) *PCT Patent, W094/29479.*

Bielinska, A., Shivadarsani, R.A., Zhang, L., Babel, G.J. (1990) *Science* **250**, 997.

Billings, P. R., Smith, C. L., and Cantor, C. R. (1991) *FASEB* J. **5**, 28.

Blackwell, T.K. & Weintraub, H. (1990) *Science* **249**, 505.

Blackwell, T.K., Kretzner, L., Blackworth E.M, Eisenman R.N., Wientraub, H. (1990) *Science* **250**, 1149.

Bock, L.C., Griffin, L.C., Latham, J.A., Vermaas, E.H., Toole, J.J. (1992) *Nature* **355**, 564.

Boutourin, A., et al., (1989) *FEBS Lett.* **254**, 129.

Brennan, R. G. & Mattews, B. W. (1989) *TIBS* **14**, 286.

Broach, J. R. & Thorner, J. (1996) *Nature* **384**, 14.

Cech, T. R., Zaug, A. J. & Gabrowski, P. J. (1981) *Cell* **29**, 487.

Chamberlain et al., (1989) *UCLA Symp. Cell. Mol. Biol. New. Ser.,* **93**, 951.

Chehab, F. F. & Kan, Y. W. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9178.

Chen, J. -H., Kallenbach, N. R., Seeman, N. C. (1989) *J. Amer. Chem. Soc.* **111**, 6402.

Chetverin, A. B., and Kramer, F. R. (1994) *Bio/Technology* **12**, 1093.

Christoffersen, R. E. & Marr, J. J. (1995) *J. Med. Chem.* **38**, 2023.

Citro, G., et al., (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 7031.

Conner, B. J., Reyes, A. A., Morin, C., Itakura, K, Teplitz, R. L. and Wallace, R. B. (1983) *Proc. Natl. Acad. Sci. USA.* **80**, 278.

Crooke, R. (1991) *Anti Cancer Drug Des.* **6**, 609.

Dang, C. & Jayasena, S. (1996) *J. Mol. Biol.* **264**, 268.

Davis, T. N. & Thorner, J. (1987) *Methods Enzymol.* **139**, 248.

Demidov, V. V., Frank-Kamenetskii, M. D., Egholm, M., Buchardt, O., and Nielsen, P. E. (1993) *Nucleic Acids Res.* **21**, 2103.

Dervan, P. B. (1992) *Nature* **359**, 87.

Dobrikov, M. I., et al., (1992) *Bioorg. Khim.* **18**, 1190.

Drmanac, R., Drmanac, S., Strezoska, Z., Paunesku, T., Labat, I, Zeremski, M., Snoddy, J., Funkhouser, W. K., Koop, B, Hood, L. and Crkvenjakov, R. (1993) *Science* **260**, 1649.

Du, S. M., Wang, H., Tes-Dinh, Y. -C., Jing, Q., Seeman, N. C. (1995) *Biochemistry* **118**, 6131.

Eaton, B. E., and Pieken, W. A. (1995) *Annu. Rev. Biochem.* **64**, 837.

Ehrlich, H. A., Ed., (1989) *PCR Technology:Principles and Applications for DNA Amplification*, Stockton press, New York.

Ellington, A. D., and Szostak, J. W. (1990) *Nature* **346**, 818.

Engels, J. & Uhlmann, E. (1988) *Adv. Biochem. Engi./Biotech.* **37**, 73.

Francois, J. C. et al., (1989) *Proc. Natl. Acad. Sci.* USA. **86**, 9702.

Godard et al., (1994) *Nucleic Acids Res.* **22**, 4789.

Gold, L. (1995) *J. Biol. Chem.* **270**, 13581.

Gold, L., Polisky, B., Uhlenbeck, O., and Yarus, M. (1995) *Annu. Rev. Biochem.* **64**, 763.

Goodchild, J. (1989) In "*Oligodeoxynucleotides, Antisense Inhibitors of Gene Expression*", J. Cohen, Ed. (Macmillan, London) pp. 53.

Hagerman, P. J. (1988) *Annu. Rev. Biophys. Chem.* **17**, 265.

Hanish, J., and McClelland. M. (1990) *Nucleic Acids Res.* **18**, 3827.

Higuchi, R., Dollinger, G., Walsh, P. S. & Griffith, R. (1992) *Bio/Technology* **10**, 413.

Hogan Jr, J. C. (1996) *Nature* **384**, 17.

Holland, P. M., Abramson, R. D., Watson, R. Will, S., Saiki, R. K. & Gelfand, D. H. (1992) *Clin. Chem.* **38**, 462.

Horwitz, M.Z. & Loeb, L.A. (1988) *J. Biol. Chem.* **263**, 14724.

Inouye, M. (1988) *Gene* **72**, 25.

Irvine, D., Tuerk, C., and Gold, L. (1991) *J. Mol. Biol.* **222**, 739.

Itakura, K, Rossi, J. J., and Wallace, R. B. (1984) *Annu. Rev. Biochem.* **53**, 323.

Itakura, K. (1982) *TIBS* **7**, 442.

Jellinek, D., Green, L. S., BellC., and Janjie, N. (1994) *Bochemistry* **33**, 10450.

Jones, R.J., Lin, K-Y., Milligan, J.F., Wadwani, S., Matteucci, M.D. (1993) *J. Org. Chem.* **58**, 2983.

Khorana, H. G. (1978) *Bioorg. Chem.* **1**, 351.

Khorana, H. G. (1979) *Science* **203**, 614.

Khrapko, K. R., Lysov, Yu. P., Khorlin, A. A. Shik, V. V., Florentiev, V. L. & Mirzabekov, A. D. (1989) *FEBS Lett.* **256**, 118.

Kleppe, K., Ohtsuka, E., Kleppe, R., Molineux, R. & Khorana, H. G. (1971) *J. Mol. Biol.* **56**, 341.

Koob, M., Grimes, E., and Szybalski, W. (1988) *Science* **241**, 1084.

Kopka, M. L., Pjura, P. E., Goodsell, D. S. & Dickerson, R. E. (1987) In "*Nucleic Acids and Molecular Biology*" ed. F. Eckstein and D. M. J. Lilley,Springer, Berlin **1**, 1.

Kreig, A., et al., (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1048.

kricca, L. J. (1992) In "*Non-isotopic DNA Probe Techniques* " Academic Press, San Diego.

Krystal, G.W., Armstron, B.C., Battey, J.F. (1990) *Mol. Cell. Biol.* **10**, 4180.

Kwok, S. & R. Higuchi. (1989) *Nature* **339**, 237.

Landegren, U., Kaiser, R., Sanders, J. and Hood, L. (1988) *Science* **241**, 1077.

Le Doan, T. D. et al., (1991) *Antisense Res. Dev.* **1**, 43.

Le Doan, T., Praseuth, L., Perrouault, M., Chassignol, N. T., Thuong, & Helene, C. (1990) *Bioconjugate Chem.* **1**, 108.

Leonetti, J., Degols, G., LeBleu, B. (1990) *Bioconjugate chem.* **1**, 149.

Letsinger, R., et al., (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6553.

Li, W.-X., Kaplan, A.V., Grant, G.W., Toole, J.J. (1994) *Blood* **83**, 677.

Li, X., Xiaoping, Y., Jing, Q., and Seeman, N. C. (1996) *J. Am. Chem. Soc.* **118**, 6131.

Liebhaber, S.A., Russel, J.E., Cash, F., Eshelman, S.S. (1991) *J. Cell. Biochem.* SD:CD 007.

Lin, Y., Qui, Q., Gill, S. C., and Jayasena, S. D. (1994) *Nucleic Acids Res.* **22**, 5229.

Martin, P. (1995) *Helv. Chim. Acta* **78**, 486.

Maskos, U. and Southern, E. M. (1993) *Nucleic Acids Res.* **21**, 2269.

Maxam, A. M. and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA.* **74**, 560.

Messing, J., Crea, R. & Seeburg P. H. (1981) *Nucleic Acids Res.* **9**, 309.

Michelson, A. M. & Todd A. R. (1955) *J. Chem. Soc.* 2632.

Miller, P.S. (1991) *BioTechnology* **9**, 358.

Milligan, J. F., Matteucci, M. D. & Martin, J. C. (1993) *J. Med. Chem.* **36**, 1923.

Morinaga, Y., Franceschini, T., Inonye, S. & Inonye, M. (1984) *BioTechnology* **7**, 636.

Moseley, S. L., Huq. I., Alim., A. R. M. A., So, M., Samabadpour-Motalebi, M. & Falkow, S. (1980) *J. Infect. Dis.* **142**, 892.

Mullis, K. B. (1990) *Sci. Am.* **262**, 56.

Nehls, M. C., Krause, S. and Boehm, T. (1994) *Mammalian Genome* **5**, 183.

Newton et al., (1989) *Nucleic Acids Res.* **17**, 2503.

Nielsen, P.E., Egholm, M., Berg, R.H., Buchardt, O. (1991) *Science* **254**, 1497.

Niemeyer, C. M. (1997) *Angew. Chem. Int. Ed. Engl.* **36**, 585.

Ou, et a., (1990) *AIDS Res. Hum. Retroviruses* **6**, 1323.

Panyutin, I. G. and Neumann, R. D. (1994) *Nucleic Acids Res.* **22**, 4979.

Pei, D., Corey, D. R. and Shultz, P. G. (1990) *Proc. Natl. Acad. Sci.* USA. **87**, 9858.

Pei, D., Ulrich, H. D. and Shultz, P. G. (1991) *Science* **253**, 1408.

Perrouault, L. (1990) *Nature* **344**, 358.

Persing, D. H. (1991) *J. Clin. Microbiol.* **29**, 1281.

Persing, D. H., Smith, T. F., Tenover, F. C. & White, T. J. (1993) *Diagnostic Molecular Biology-Priniciples & Applications* American society for microbiology, Washington. D. C.

Povsic, T. J. and Dervan, P. B. (1989) *J. Amer. Chem. Soc.* **112**, 9428.

Povsic, T. J., Strobel, S. A and Dervan, P. B. (1992) *J. Amer. Chem. Soc.* **114**, 5934.

Praseuth, D. et al., (1988) *Proc. Natl. Acad. Sci.* USA. **85**, 1349.

Rahn, R. O. (1992) *Photochem. Photobiol.* **56**, 9.

Rajeev, K. G., Jadhav, V. R. & Ganesh, K. N. (1997) *Nucleic Acids Res.* **25**, 4187.

Rajendrakumar, G., Ganesh, K. N. & Chatterji, D. (1990) *J. Biol. Chem.* **265**, 22300.

Robinson, B. H., and Seeman, N. C. (1987) *Protein Eng.* **1**, 295.

Saenger, W. (1984) In "*Principles of Nucleic Acid Structure* ", Springer Verlag, New York.

Sagi, J., et al., (1993) *Tetrahedron Lett.* **34**, 2191.

Saiki et al., (1985) *Science* **230**, 1350.

Saiki, R. K., Walsh, P. S., Levenson, C. H. and Erlich, H. A. (1989) *Proc. Natl. Acad. Sci. USA.* **86**, 6230.

Samulski, R. J. (1993) *Curr. Opin. Genet. Dev.* **3**, 74.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA.* **74**, 5463.

Sangvi, Y.S. & Cook, P.D. (1994) In "*Carbohydrate Modifications in Antisense Research", A.C.S. Symposium Series No. 580*, Ed. Sanghvi, Y.S., Crook, P.D. (American Chemical Society, Washington-DC) pp. 1.

Saxon, M., et al., (1992) *Antisense Res. Dev.* **2**, 243.

Scheller, R. H., Dickerson, R. E., Boyer, H. W., Riggs, A. D. & Itakura, K. (1977) *Science* **196**, 180.

Schneider, T. D., Stormo, G. D., Gold, L., and Ehrenfeucht, A. *J. Mol. Biol.* **188**, 415, (1986).

Seeman, N. C. (1982) *J. Theor. Biol.* **99**, 237.

Seeman, N. C. (1997) *Acc. Chem. Res.* **30**, 357.

Sen, D. & Gilbert, W. (1988) *Nature* **334**, 364.

Sigman, D. S (1990) *Biochemistry* **29**, 9097.

Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503.

Southern, E. M. In "Analysing polynucleotide sequences", PCT Application WO 89/10977.

Stein, C. A. & Cheng, Y. C. (1993) *Science* **261**, 1004.

Stemp, E. D. A., Arkin, M. R. & Barton, J. K. (1995) *J. Am. Chem. Soc.* **117**, 2325.

Strobel, S. A and Dervan, P. B. (1990) *Science* **249**, 73.

Strobel, S. A and Dervan, P. B. (1992) *J. Amer. Chem. Soc.***111**, 7286.

Stryer, L. In "Biochemistry", 3rd Ed, (1988) W. H. Freeman, New York, p 677

Sugiyama, H., Tsutsumi, Y., and Saito, I. (1990) *J. Amer. Chem. Soc.* **112**, 6720.

Switzer, C. Y., Moroney, S. E. and Benner, S. A. (1993) *Biochemistry* **32**, 10489.

Takasugi, M. et al., (1991) *Proc. Natl. Acad. Sci.* USA. **88**, 5602.

Trevors, J. T. (1985) *Trends in Biotechnology* **3**, 291.

Trotta, P. P., Beutel, B. A. & Sherman, M. I. (1995) *Med. Res. Rev.* **15**, 277.

Tsai, D.E., Harper, D.S., Keene, J.D. (1991) *Nucleic. Acid. Res.* **19**, 4931.

Tuerk, C. & Gold, L. (1990) *Science* **249**, 505.

Tuerk, C., MacDougal, S., Gold, L. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 6988.

Vickers, T., et al., (1991) *Nucleic Acid Res.* **19**, 3359.

Wagner, R.W., Matteucci, M.D., Lewis, J.G., Gutierrez, A.J., Moulds, C., Forehler, B.C. (1993) *Science* **260**, 1510.

Walder, R.Y. & Walder, J.A.(1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5011.

Wallace, R. B., Shaffer, J., Murphy, R. F., Bonner, J., Hirose, T., and Itakura, K. (1979) *Nucleic Acids Res.* **6**, 3543.

Weller, H., (1996) *Angew. Chem. Int. Ed. Engl.* **35**, 1079.

Wells, R. D. (1988) *J. Biol. Chem.* **263**, 1095.

Wilson, W. D. (1990) In "*Nucleic Acids in Chemistry & Biology*" Ed. Blackburn, G. M. & Gait, M. J. IRL press, Oxford.

Wu, D. Y., Ugozzoli, L., Pal, B. K. & Wallace, R. B. (1989) *Proc. Natl. Acad. Sci. USA.* **86**, 2757.

Yakubov, L., et al., (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86,** 6454.

Yang, J-H., Usman, N. Chartrand, P. & Cedergren, R. J. (1992) *Biocemistry* **31**, 5005.

Zarytova, V. F., Levina, A. S., Khalimskaya, L. M. & Sergeeva, Z. A. (1992) *Bioorg. Khim.* **18**, 640.

Zoller, M. J. & Smith, M. (1983) *Methods Enzymol.* **100**, 468.

CHAPTER2

Fluorescent oligonucleotides: Synthesis, evaluation of fluorescence properties, effect of spacer chain and PCR applications

2.1 INTRODUCTION

The specific molecular recognition (enzyme-substrate, DNA-protein, ligand-receptor etc.) is a key feature underlying many biological processes. The discovery of structure of DNA double helix with specific molecular recognition by complementary strands of nucleic acids by Watson and Crick in 1953, decoded the nature's secret of storage and transmission of the genetic information. Since then, the most important principle of molecular recognition of nucleic acids used by nature for preservation of life is being exploited by the scientific community in understanding molecular genetics (Southern, 1975; Viscidi et al., 1987; Smith, 1993) and finding newer applications like DNA diagnostics (Saiki et al., 1985) and therapeutics (Crooke & Bennet, 1996). With the advent of chemical synthesis of oligonucleotides and later the automization of the process, synthetic oligonucleotides have found widespread applications in molecular biology due to their property of specific hybridization to complementary nucleic acid sequences (Kricca, 1992). The procedures for detection of nucleic acids in these studies involve the use of radioisotopes. Although radiolabelling is far more sensitive than any other detection methods, it suffers from the problems associated with stability, handling, storage and disposal. Owing to such problems, a number of studies have been directed towards nonradioactive labeling of nucleic acids which involve biotinylated (Telser et al., 1989), digoxigenin (Rihn et al., 1995) or fluorescent (Haralambidis et al., 1990) labeled probes. The widely used biotin and digoxigenin probes are part of a heterogeneous system, in which the label is detected through a cascade of reactions, which are prone to artifactual results while the fluorescent labeled probes provide direct, reliable detection.

Fluorescent labeling of nucleic acids has attracted wide attention for probing nucleic acid hybridization (Gilham, 1987; Kingsbury, 1987; Keller & Manak, 1989; Kricca, 1992.), DNA sequencing (Ansorage et al., 1987; Prober et al., 1987; Brumbagh et al., 1988; Landegren et al., 1988), nucleic acid-protein interactions (Allen et al., 1989; Patel et al., 1990; Guest et al., 1991; Hagmar et al., 1995) and for applications in DNA diagnostics (Landegren et al., 1988;

Chehab & Kan, 1989). The lower sensitivity of this technique compared to conventional labels can be overcome either by multiple labeling (Tang & Agarwal, 1990; Conway & McLaughlin, 1991; Tong et al., 1993) or signal amplification through PCR technique (Saiki et al., 1985). The reported fluorescent labeling methods frequently involve conjugation of a fluorescent dye to 5'/3' terminus of oligonucleotide (Asseline et al., 1984; Chollet & Kawashima, 1985; Kempe et al., 1985; Smith, 1985; Agarwal et al., 1986; Wachter et al., 1986; Connolly, 1987; Thiesen et al., 1992; Guisti & Adriano, 1993; Markiewicz et al., 1997) phosphate backbone (Conway & McLaughlin., 1991; Murakami et al., 1991; Tong et al., 1993), sugar (Manoharan et al., 1991) or nucleoside bases using spacer arms (Draper, 1984; Ruth, 1984; Haralambidis et. al, 1987; Singh et al., 1990). Nucleobase labeling has advantages of incorporation at internal sites without affecting the backbone geometry and it leaves 5' or 3' ends free for enzymatic reactions like polynucleotide kinase and DNA polymerase. So far, most of the strategies for labeling involve synthesis of nucleic acids with modified pyrimidines carrying nucleophilic handles such as amines or alkyl sulphydryls, followed by labeling reaction (Asseline et al., 1984; Chollet & Kawashima, 1985; Kempe et al., 1985; Smith, 1985; Agarwal et al., 1986; Wachter et al., 1986; Connolly, 1987; Haralambidis et al., 1987; Urdea et al., 1988; Thiesen et al., 1992; Guisti and Adriano, 1993). The regiospecificity and degree of labeling in such a reaction is governed by the specificity of fluorophoric reagent used and the reaction conditions. Though this method is satisfactory for monolabeling, incorporation of multiple labels leads to an inseparable mixture of oligonucleotides with components that differ in number of labels per oligonucleotide chain. The alternative method of prelabeling before the oligonucleotide assembly by phosphoramidite gives control over uniform distribution of labels and has advantage of multiple labeling to increase the sensitivity. Such modified monomers, carrying fluorescent groups should be protected in such a way that the label is stable during assembly and purification of oligonucleotides .

Many chemical studies are directed towards development of new and simpler procedures to obtain such derivatives (Goodchild, 1989; Beaucage & Iyer, 1993). Singh et al., (1990) have reported the synthesis and fluorescent properties of oligonucleotides containing adenines covalently linked at C-8 with dansyl fluorophore using variable polymethylene spacer chain. The C-5 position of pyrimidines is the preferred site for attachment of label, since this would interfere least with hybridization, an important criterion to be met for internally labeled oligonucleotides. The reported methods of attaching fluorescent group at C-5 (Figure 1) involve the synthesis of oligonucleotides containing modified base with suitable linker at C-5 position followed by postsynthetic fluorophore labeling (Ruth, 1984; Haralambidis et al., 1987; Cook et al., 1988; Telser et al., 1989).

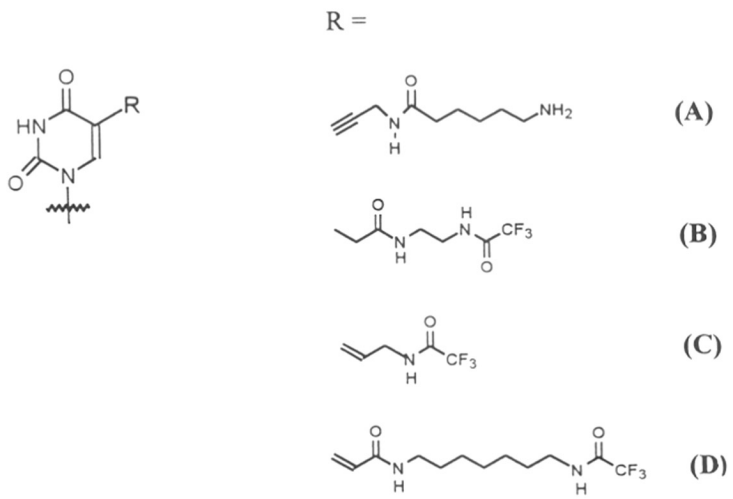


Figure 1. The different linkers attached at C-5 site of deoxyuridine reported in literature. **A**: Haralambidis et al., (1987); **B**: Telser et al., (1989) **C**: Cook et al., (1988); **D**: Jablonski et al., (1986).

2.2 Present Work

This chapter describes a strategy for chemical synthesis of fluorescent labels and their incorporation during oligonucleotide synthesis to enable complete control over the position and number of labels in each DNA molecule

and a systematic study of the consequent fluorescence properties and their applications. This is achieved by synthesis of protected phospharamidite monomers which contain 5/6-carboxyfluorescein fluorophore attached to 2'-deoxyuridine base residue at C-5 position through a short and rigid amido spacer as well as long and flexible 6-aminohexanoyl spacer. The comparative fluorescent properties of these oligonucleotides containing fluorophores linked with short and long spacers are described. The hybridisation of these fluorescent oligonucleotides with complementary DNA is studied by monitoring changes in fluorescence intensity and anisotropy. Their utility in studying DNA-drug interactions, and as primers in PCR for fluorescent detection of amplified products is reported.

2.3 EXPERIMENTAL

The chemicals used were of laboratory or analytical grades. All the solvents used were purified according to the literature procedures (Perrin & Armarego, 1989). Usual work-up implies sequential washing of the organic extract with water, brine, water followed by drying over anhydrous sodium sulfate and evaporation of solvent under vacuum. GeneAmp PCR reagent kit from Perkin Elmer was used for preparing PCR amplification reactions which were performed on Perkin-Elmer DNA thermal cycler. Netropsin was obtained from Boehringer Mannheim. Agarose and ethidium bromide were obtained from Amersham.

2.3.1 Chromatography

Column chromatography was carried out for purification of compounds on Loba silica gel (100-200 mesh). The TLC was performed on silica gel GF_{254} precoated aluminium sheets (Merck 5554). Spots were visualized by spraying with 60% perchloric acid-ethanol solutions when DMTr compounds show orange spots, followed by heating when dark spots show up. The amino compounds were detected by ninhydrin spray.

2.3.2 Spectroscopy

^{1}H (200 MHz) and ^{13}C (50 MHz) NMR spectra were recorded on a Bruker ACF 200 spectrometer fitted with an Aspect 3000 computer and all the chemical shifts are refered to internal TMS for ^{1}H, chloroform-*d* for ^{13}C and 85% H_3PO_4 for ^{31}P. The chemical shifts are quoted in δ scale (ppm). Qualitative UV-visible spectra and thermal denaturation studies were recorded on Perkin-Elmer Lambda 15 UV/VIS spectrophotometer fitted with Julabo water circulator.

2.3.3 Protocol for Synthesis of Fluorescent Amidites

3',5'-Di-O-acetyl-5-bromo-2'-deoxyuridine, 2

A suspension of 2'-deoxyuridine **1** (5 gm, 21.9 mmol) in 30 ml of acetic anhydride was heated until dissolution occurred. A solution of bromine (3.8 gm, 23.8 mmol) in 3 ml glacial acetic acid was added, with cooling, to maintain a temperature of 25 °C. After keeping overnight in cold (4 °C), the solution was evaporated to a thick syrup under reduced pressure. On addition of about 30 ml of ethanol, the diacetate **2** crystallized slowly. (7.2 gm, yield 84%), m.p. 153°C. Rf (10% MeOH in DCM) = 0.8.

^{1}H NMR ($CDCl_3$) δ 7.8 (s, 1H, H6), 6.3 (t, 1H, H1'), 5.2 (m, 1H, H3'), 4.4 to 4. 15 (m, 3H, H4', H5' and H5"), 2.55 (m, 2H, H2' and H2"), 2.15 and 2.0 (s, 6H, 2x$OCO\underline{C}H_3$) . **^{13}C NMR** ($CDCl_3$) δ 170 and 169.8 ($O\underline{C}OCH_3$), 158.5 (C4), 149.4 (C2), 138.4 (C6), 97.2 (C5), 85.6 (C4') 82.5 (C1'), 73.7 (C3'), 63.5 (C5'), 38 (C2'), 20.56 ($OCO\underline{C}H_3$).

5-Bromo-2'-deoxyuridine, 3

Compound **2** (3.9 gm, 10 mmol) was dissolved in 85 ml of anhydrous methanolic ammonia (10%) and the solution was kept at room temperature for 24 hours. The solvent was evaporated and the residue was dissolved in a minimum volume of hot, anhydrous ethanol. On cooling, crystals of **3** were obtained (3.3 gm, yield 85%), m. p. 193 °C. Rf (10% MeOH/DCM) = 0.4.

^{1}H NMR (DMSO-d_6) δ, 8.4:(s, 1H, H6), 6.1 (t, 1H, H1'), 5.4-5.1 (brd, 2H, 3'OH and 5' OH), 4.25 (m, 1H, H3'), 3.8 (m, 1H, H4'), 3.6 (m, 2H, H5', H5"), 2.1 (m, 2H,

H2', H2"). **^{13}C NMR** (DMSO-d_6) δ 159.2 (C4), 149.8 (C2), 140.4 (C6), 95.8 (C5), 87.7 (C4'), 85 (C1'), 70.1 (C3'), 60.9 (C5'), 40.2 (C2').

5-Bromo-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridine, 4

Compound **3** (3.0 gm, 10 mmol) was dried by co-evaporation with dry pyridine and redissolved in dry pyridine (15 ml). To this 4,4'-dimethoxytrityl chloride (3.7 gm, 1.1 eq) was added in two portions at 3 hour. The stirring was continued at room temperature for further 3 hr. The progress of reaction as monitored by TLC was complete in 6 hr. It was quenched by addition of MeOH (1 ml), concentrated to dryness, dissolved in CH_2Cl_2 and washed with water, purified by column chromatography. (3.65 gm, yield 60%) Rf (4% MeOH/DCM) = 0.4

^{1}H NMR ($CDCl_3$) δ 8.15 (s, 1H, H6), 7.25 to 7.75 (m, 9H, ArH DMT), 6.85 (m, 4H, ArH DMT), 6.35 (q, 1H, H1'), 4.6 (M, 1H, H3'), 4.15 (m, 1H, H4'), 3.8 (s, 6H, 2xOCH_3 DMT), 3.4 (M, 2H, H5', H5"), 2.6 and 2.35 (M, 2H, H2', H2"). **^{13}C NMR** ($CDCl_3$) δ 159.4 (C4), 149.1 (C2), 139.3 (C6), 97.1 (C5), 86.6 (C4'), 85.7 (C1'), 72.1 (C3'), 63.4 (C5'), 55.1 (OCH_3), 41.2 (C2').

5-Amino-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridine, 5

Compound **4** (1.5 gm, 2.5 mmol) was taken in a glass tube placed in a stainless steel bomb to which 20 ml of cooled liquid ammonia was added. The reaction mixture was kept at 60°C for 30 hr and after slow removal of ammonia, the residue was taken in CH_2Cl_2 and washed with water to remove NH_4Br formed during reaction. The product gave a single spot positive for DMTr and ninhydrin on TLC which was used without any further purification. (yield 1.3 gm, 97%). Rf (4%MeOH/DCM) = 0.35

^{1}H NMR ($CDCl_3$) δ 7.7 to 7.2 (m, 10H, H6 and ArH DMT), 6.85 (m, 4H, ArH DMT), 6.5 (t, J = 6.5 Hz, 1H, H1'), 4.65 (m, 1H, H3'), 4.1 (m, 1H, H4'), 3.75 (s, 6H, 2xOCH_3 DMT), 3.45 (m, 2H, H5', H5"), 2.4 (m, 2H, H2', H2"). **^{13}C NMR** ($CDCl_3$) δ 160.7 (C4), 150 (C2), 127 (C6), 116.6 (C5), 86.1 (C4'), 84.6 (C1'), 72 (C3'), 63.6 (C5'), 55.1 (OCH_3), 40.4 (C2').

***Synthesis of N-benzyloxycarbonyl-ω-amino hexanoic acid PFP ester,* 7**

ω-aminohexanoic acid (1.3 gm, 10 mM) was taken in 10% Na_2CO_3 solution to which benzyl chloroformate (0.8 ml, 11 mM) was added at room temperature under constant stirring. The pH was adjusted to 8.0 with Na_2CO_3 solution. After 1 hr, reaction mixture was concentrated and pH was adjusted to 2-3 with dilute HCl and extracted with ethyl acetate to obtain product **6**. This compound was used for the next step without further purification.

The compound **6** (1.82 gm, 5 mM) was taken in 5 ml of THF to which PFP (1.0 gm, 5.5 mM) and DCC (1.1 gm, 5.5 mM) were added. The reaction was carried out for 3 hours and checked by TLC. After the usual work up, the product was separated on a silica gel column (25 gm) using 50% DCM/Pet ether solvent. Yield (1.94 gm, 90%), Rf (75% DCM/Pet ether)= 0.62.

^{1}H NMR ($CDCl_3$) δ 7.4 (s, 10H, C_6H_5), 5.1 (s, 2H, CH_2), 4.85 (bs, 1H, NH), 3.2 (q, 2H, OC-CH_2), 1.5 (m, 6H, $(CH_2)_3$)

***Synthesis of 5'-O-(4,4'-dimethoxytrityl)-5-N-ω-benzyloxycarbonyl-amino hexanoyl-5-amino-2'-deoxyuridine,* 8**

The compound 5-amino-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridine **5** (400 mg, 0.75 mM) was taken in 3 ml of dry pyridine to which **7** (390 mg, 0.90 mM) and HOBT (100 mg, 0.74 mM) were added. The reaction was kept under stirring for 24 hours and the product formation was checked by TLC. The reaction mixture was dried under reduced pressure and after usual work up, the product was separated by silica gel chromatography. Yield (440 mg, 75%), Rf (4% MeOH/DCM) = 0.41.

^{1}H NMR, (CDCl3) δ10.08 (bs, 1H, exchangeable with D_2O, NH), 8.62 (s, 1H, H6), 7.45-7.20 (m, 18H, aromatic H's), 6.34-6.27 (t, 1H, j = 6.5 Hz, H1'), 5.05 (s, 2H, benzylic CH_2), 4.94 (bd, 1H, exchangeable with D_2O, NH), 4.44-4.36 (m, 1H, H4'), 4.03-3.97 (m, 1H, H3'), 3.74 (s, 6H, 2 x OCH_3 of DMT), 3.50-3.38 (m, 2H, H5', H5''), 3.17-3.10 (q, 2H, N- CH_2), 2.44-2.20 (bm, 4H, H2', H2'' and CO-CH_2 of the side chain) and 1.72-1.27(m, 6H, 3 x CH_2 of the side chain)

^{13}C NMR, ($CDCl_3$) δ 171.0, 160.8, 158.2, 156.6, 149.1, 148.8, 139.1, 135.5, 130.0, 128.2, 128.0, 127.8, 127.6, 123.7, 114.8, 112.9, 86.5, 85.8, 85.4, 71.6, 66.7, 64.0, 54.9, 41.3, 40.2, 36.2, 29.9, 26.0 and 24.8

***Synthesis of 5'-O-(4,4'-dimethoxytrityl)-5-N-ω-aminohexanoyl-5-amino-2'-deoxyuridine*, 9**

The compound **8** (400 mg, 0.5 mM) was taken in 5 ml of MeOH to which palladium charcoal (20% w/v) and 3 to 4 equivalent excess ammonium formate were added. This reaction of catalytic transfer hydrogenation for removal of benzyloxycarbonyl group was complete in 10 minutes. The reaction was monitored by TLC in which the product showed ninhydrin positive spot. After removal of MeOH under reduced pressure, the product was directly purified by silica gel chromatography and eluted with 15% MeOH/DCM solvent. Yield (80%, 275 mg), Rf (15% MeOH/DCM) = 0.25.

***Spiro[isobenzofuran-1-(3H)-9'-[9H]xanthene]-5(and-6)carboxylic acid, 3'.6'-(2,2-dimethylpropanoyloxy)-3oxo pentafluorophenol*, 11**

Trimellitic anhydride (4.8 gm, 25 mmol) and resorcinol (5.5 gm, 55 mmol) were mixed thoroughly and placed in a 190 °C oil bath for 1 hr. The temperature was then increased to 210 °C and was maintained for 5 hr by which time the melt gets solidified to a dark red solid. It was then cooled down and dissolved in hot ethanol followed by precipitation from water. The precipitate was filtered out and dried to yield a fluorescent compound . This was dissolved in 30 ml of dry pyridine and kept on ice. To this, trimethylacetyl chloride (9.2 ml, 0.15 mol) was added dropwise, allowed to stir for 24 h, filtered, and the filtrate was concentrated to thick syrup. The resultant was redissolved in 400 ml of ethyl acetate, washed with 1M H_2SO_4 (2 x 300 ml) and water (1 x 300 ml), dried over anhydrous sodium sulfate and reconcentrated to syrup which was chromatographed over silica gel column to give 6.2 gm of crude non-fluorescent lactone **10**.

To a mixture of **10** (1.2 gm, 2.2 mmol), pentafluorophenol (0.5 gm, 2.7 mmol) and DCC (0.56 gm, 2.7 mmol), 20 ml of dry THF was added and allowed

to stir for 24 h. The precipitated dicyclohexyl urea was filtered out and the filtrate was concentrated to dryness followed by purification over silica gel chromatography. (yield 1.0 gm, 64%). Rf (10%MeOH/DCM) 0.71 and 0.68, corresponding to 5 and 6 isomers.

^{1}H NMR ($CDCl_3$) δ 8.20 (m, 1H, H-4 and H-5 of 6-isomer), 8.87 (s, 0.5H, H-4 of 5-isomer), 7.96 (s, 0.5H, H_7 of 6-isomer), 7.38 (d, J = 8 Hz, 0.5H, H_7 of 5-isomer), 7.11 (s, 2H, H4'), 6.83(s, 4H, H1' and H2'), 1.4 (s, 18 H, 2 x $(CH_3)_3$).

***Synthesis of 5'-O-(4,4'-dimethoxytrityl) -5-N-carboxyfluoresceinyl -5-amino-2'-deoxyuridine,* 12**

To compound **5** (270 mg, 0.5 mmole) in 5 ml dry pyridine, **11** (420 mg, 0.6 mmole) and HOBT (60 mg, 0.5 mmole) were added and stirred overnight at room temperature. The reaction was monitored by TLC, evaporated to dryness, and redissolved in CH_2Cl_2 (30 ml) and the solution was washed with aqueous $NaHCO_3$ (2x10 ml) concentrated and followed by product purification over silica gel. Yield (63%, 330 mg), Rf (4% MeOH/DCM) = 0.42.

^{1}H NMR ($CDCl_3$) δ 8.9-7.5 (m, 4H, H6 & ArH), 7.5-6.7 (m, 18H, DMT & ArH), 6.4 (t, 1H, H1'), 4.5 (m, 1H, H3'), 4.1 (m, 1H, H4'), 3.7 (s, 6H, -OCH_3), 3.5 (m, 2H, H5', H5"), 2.55-2.3 (m, 2H, H2', H2"), 1.4 (S, 18H, 2X$C(CH_3)_3$).

***Synthesis of 5'-O-(4,4'-dimethoxytrityl)-5-N-ω-carboxyfluoresceinyl-amino hexanoyl-5-amino-2'-deoxyuridine,* 13**

The compound **9** (250 mg, 0.37 mM) was taken in 5 ml pyridine to which **11** (290 mg, 0.4 mM) and HOBT (50 mg, 0.37 mM) were added and kept for stirring for 24 hours. The reaction was monitored by TLC and the pyridine was removed under reduced pressure. After usual work up, product was purified by silica gel chromatography using 4% MeOH/DCM solvent system. Yield (65%, 295 mg), Rf (4% MeOH/DCM) = 0.4.

^{1}H NMR ($CDCl_3$) δ 8.61-8.56 (m, 3H, ArH) 8.47 (s, 1H, H6), 8.21- 8.04 (m, 2H, ArH), 7.75-7.66 (m, 2H, ArH), 7.44-6.95 (m, 15H, ArH), 6.34-6.27 (t, 1H, j = 6.5 Hz, H1'), 4.39-4.36 (m, 1H, H4'), 4.00-3.97 (m, 1H, H3'), 3.75-3.71 (d, 6H, 2 x

OCH_3 of DMT), 3.45-3.24 (m, 4H, H5', H5" and N-CH_2 of the side chain), 2.28-2.17 (bm, 4H, H2', H2" and CO-CH_2 of the side chain), 1.66-1.56 (bm, 4H, 2 x CH_2 of the side chain) and 1.36-1.24 (bs, 20H, 1 x CH_2 of the side chain and 2 x $C(CH_3)_3$).

^{13}C NMR, ($CDCl_3$) δ 176.7, 176.5, 171.6, 171.4, 168.7, 168.4, 165.7, 160.3, 160.2, 158.5, 155.2, 152.8, 152.7, 151.6, 151.5, 149.4, 149.0, 144.9, 141.5, 137.0, 136.4, 130.2, 129.8, 128.8, 128.2, 127.8, 127.3, 126.8, 126.3, 126.6, 124.4, 118.0, 114.8, 115.7, 114.8, 114.7, 113.2, 110.5, 86.6, 85.9, 85.5, 82.2, 81.9, 71.7, 64.1, 55.2, 40.2, 40.0, 39.2, 36.4, 29.1, 27.1, 26.4 and 24.8.

***General Procedure for Phosphitylation of 12 & 13 to 3'-O-(N,N-Diiso propylamino)(β-cyanoethoxy) phosphines*, 14 & 15**

Compound **12** & **13** (0.2 mmol) were separately dried by coevaporation with tetrazole (14 mg, 0.5 mmol) in dichloroethane and suspended in 1 ml of dichloroethane. To each, 2-cyanoethyl N,N,N',N'-tetraisopropyl phosphoramidite (0.1 ml, 0.31 mmol) was added with stirring under dry conditions. After stirring for 3 hr, the solutions were diluted with dry CH_2Cl_2 and washed with 5% aqueous $NaHCO_3$, concentrated to dryness. These phosphoramidites were dried overnight over P_2O_5 and KOH in a dessicator before using on DNA synthesiser. TLC, RF (EtOAc:CH_2Cl_2, 1:1 with 0.5% TEA) shows two close moving spots for two distereoisomers, **14** = 0.6 & 0.42, **15** = 0.65 & 0.45

^{31}P NMR ($CDCl_3$) **14** = 149.1 & 149.3 ppm, **15** = 150.6 & 149.6 ppm.

2.3.4 Oligonucleotide Synthesis

Base protected standard nucleoside phospharamidites (A, T, G and C) and nucleoside derivatized controlled pore glass supports were purchased from Cruachem, UK. The DNA synthesis was carried out on Pharmacia LKB-Gene Assembler Plus. Dry solvents were used for DNA synthesis. The commercially available amidites (0.1 M) were dissolved in dry acetonitrile while 0.15 M solutions were prepared for modified amidites and 4A molecular sieves were added to it to remove traces of moisture. Acetonitrile was distilled twice over

P_2O_5 and finally over CaH_2 immediately before use. Dichloroethane was dried by distilling twice over P_2O_5. For oxidation, after each coupling, 0.01 M iodine in collidine, water and acetonitrile, while for capping 20% acetic anhydride in acetonitrile were used. The solid phase synthesis protocol is summarized in Figure 2. The modified amidites were incorporated at desired sites of ODN by using automated Pharmacia GA Plus DNA synthesizer.

2.3.5 Oligonucleotide purification

The crude oligonucleotides were cleaved from support, deprotected using aqueous ammonia at 55 °C for 16 hr and desalted using NAP-10 gel filtration columns. The desalted oligonucleotides were purified by gel electrophoresis under denaturing conditions. The oligonucleotides (10 O.D.) were dissolved in 200 µl of 95% formamide, heated at 80 °C for 5 min and loaded on 20% polyacrylamide gel containing 7 M urea. The gels were run at constant voltage (200 V) till the dye reached the bottom of gel (Effcavitch, 1990). The fluorescent oligonucleotides were visible on gel as coloured bands while unmodified oligonucleotides were detected by UV shadowing. The gel slices containing full length oligonucleotides were cut and crushed in a dry tube. To this, 5 ml of sterile water was added and the tubes were kept at 50 °C overnight. The solution was filtered through Whatman paper to remove polyacrylamide particles. The filtrate was concentrated, dissolved in minimum amount of water (~500 µl) and purified twice over NAP-10 gel filtration column (Pharmacia) to remove dissolved acrylamide and urea. The purity of oligonucleotide was rechecked by reverse phase HPLC using gradient of 10% to 30% acetonitrile in 0.1 M TEAA (pH 7.0) buffer for 20 min. The absorption spectrum of each peak was scanned in the range of 200-600 nm using diode array detector.

2.3.6 DNA Melting Experiments

All melting experiments were done by hybridising 1 µM concentrations of each strand of duplex in 10 mM Tris-HCl (pH 7.0) buffer containing 10 mM $MgCl_2$, 100 mM NaCl. The samples were annealed by heating at 85 °C for 3 minutes, allowed to come to room temperature followed by storage at 4 °C

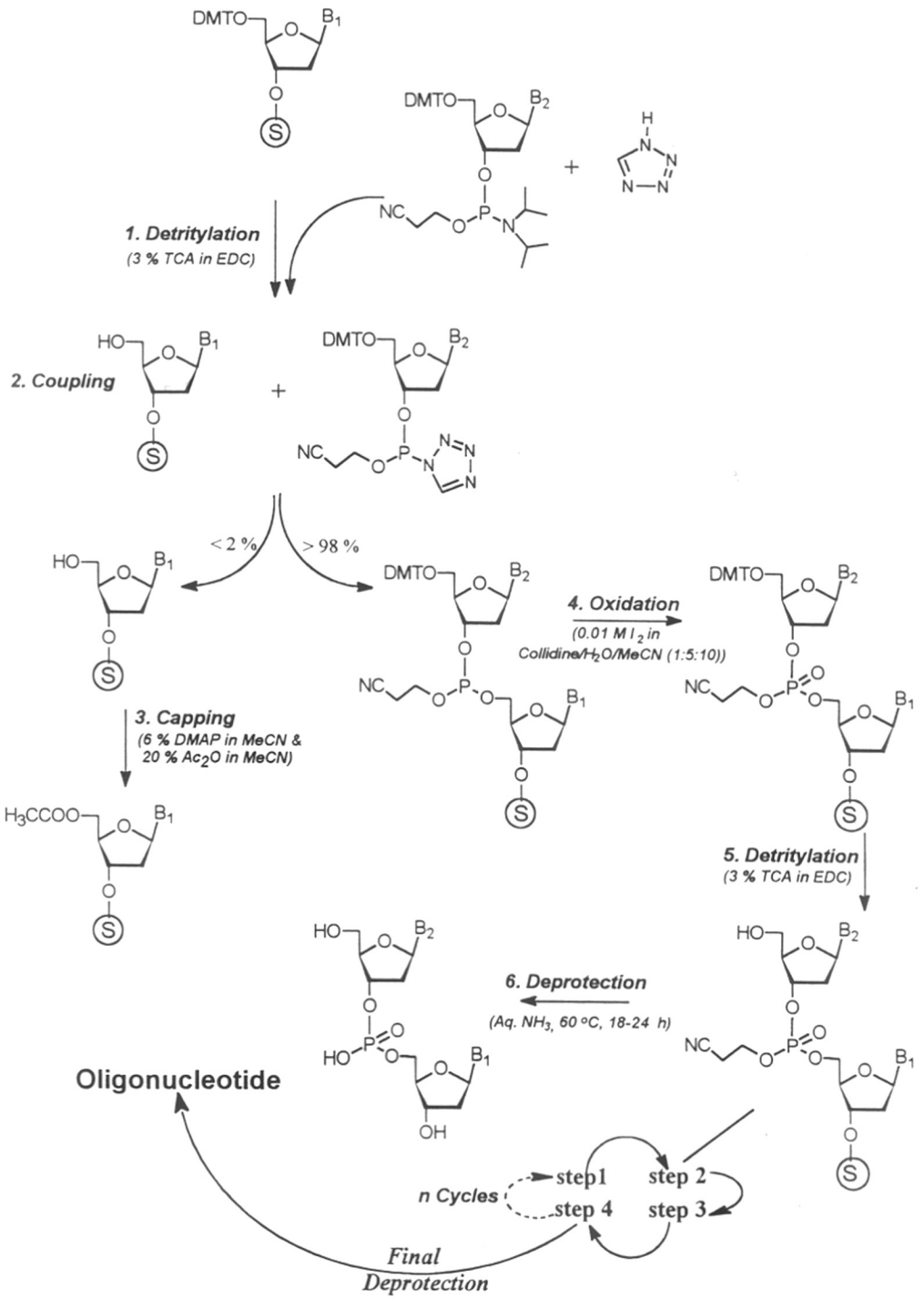


Figure 2: Solid phase chemical synthesis of oligodeoxynucleotide

overnight. The temperature of the sample compartment of the spectrofluorimeter was maintained using a Julabo water circulator in the range of 20 °C-85 °C. For melting, the annealed duplexes were heated at 0.5 °C per minute and absorption recorded every minute. In a plot of relative change of absorption vs temperature, the temperature corresponding to 50% rise in absorption was taken as the melting temperature (T_m) of the DNA.

2.3.7 Fluorescence Spectroscopy

Fluorescence measurements were performed on a Perkin-Elmer Model LS 50B spectrofluorimeter. Oligomer concentrations were determined from the absorption at 260 nm and using molar extinction coefficients for A=15.4, C=7.3, G=11.7 and T=8.8 $cm^2/\mu M$ at 260 nm. The absorbances of the labeled oligomers were corrected for fluorescein contribution at 260 nm by substraction of one-fifth of the fluorescein absorbance at 480 nm. Unless otherwise stated the concentration of oligomers in the sample was 75 nM. The samples were prepared in 10 mM Tris-HCl (pH 7.0) buffer containing 10 mM $MgCl_2$, 100 mM NaCl. The temperature of the sample compartment of the spectrofluorimeter was maintained using a Julabo water circulator at 20 °C for all experiments but varied in the range of 20 °C-45 °C for studies at different temperature.

The fluorescence spectra of oligonucleotides were recorded by excitation at 495 nm and fluorescence emission monitored from 500-560 nm using the excitation slit width of 10 nm and emission slit width of 5 nm. KI was used as a fluorescence quencher in the concentration range of 0 to 0.6 M.

2.3.8 Steady-state Anisotropy Measurements

Steady-state anisotropy measurements were performed using in the L-format configuration. The samples were excited at 495 nm and emission was monitored at 520 nm using excitation slit width of 10 nm and emission slit width of 5 nm. Fluorescence anisotropy (A) was calculated using the measurements of the emission intensity, I, according to

$$A = I_{vv} - I_{vh}G / I_{vv} + 2I_{vh}G$$

where the subscripts v and h refer to the orientation (vertical or horizontal) of the polarizers in the excitation beam (first subscript) and the emission beam (second subscript) for that intensity measurement. The instrumental correction factor, *G*, is the ratio of I_{hv} to I_{hh}.

The changes in anisotropy upon addition of restriction enzyme *Hinf*I to oligo duplexes were performed using an enzyme concentration of 1000 units/ml in a buffer containing 10 mM Tris-HCl (pH 7.5); 100 mM NaCl; 7 mM 2-mercaptoethanol, 10 mM KCl; 0.01 mM EDTA; 0.1 mM DTT; 0.015% Triton X-100; 0.001% BSA and 5% glycerol. The KCl, glycerol, BSA, Triton X-100, DTT and β-mercaptoethanol were contributed by the stock *Hinf*I solution. Experiments lacking *Hinf*I also contained these reagents. The *Hinf*I cleavage of DNA upon hybridization is inhibited in presence of EDTA, which chelates the cofactor Mg^{2+} essential for enzyme activity.

2.3.9 Netropsin Binding Assay

Netropsin binding assay was performed by successive addition of 10 nM of netropsin to the fluorescent oligonucleotide duplexes and following changes in fluorescence emission intensity at 520 nm until it reached near saturation (excitation at 495 nm). The binding affinity of netropsin was calculated from a plot of normalized fluorescence intensity at λ_{em} maximum (at each titration point) against log of netropsin concentration (Chaires et al., 1993). In such a plot, the antilog of the point on X-axis corresponding to the midpoint of the curve gave the binding constant (K_a).

2.3.10 Fluorescence Lifetime Experiments

Time-resolved fluorescence decay were observed by employing a CW mode-locked frequency-doubled Nd-YAG laser-driven dye (Rhodamine 6G) laser operating at a repetition rate of 800 kHz with pulse width of the order of 4-10 ps and tunability in the region of 570-640 nm. Fluorescence decay curves were obtained by using a time-correlated single-photon-counting setup coupled to a multichannel plate photomultiplier (Model 2809U; Hamamatsu Corporation). The instrument response function (IRF) was obtained at 230 nm using a second

harmonic output of the dye laser from an angle-tuned ADP crystal and the fluorescence emission was collected through a 450 nm cut-off filter followed by a monochromator. The cut-off filter was used to prevent scattering of the excitation beam from the sample. The peak counts obtained in control experiments were comparable to the background level. In all life-time measurements, the emission was monitored at the magic angle of 54.7° to eliminate the contribution from the decay of anisotropy.

The fluorescence decay curves were analyzed by deconvoluting the observed decay with instrument response factor to obtain the intensity dacay function represented as a sum of three exponentials:

$$I(t) = \sum \alpha_i exp\ (-t/\tau_i) \qquad i = 1\text{-}3$$

where I(t) is the fluorescence intensity at time t and α_i is the amplitude of the lifetime τ_i such that $\sum_i \alpha_i = 1$.

The system was tested by measuring the fluorescence lifetimes of two reference compounds at room temperature, N- acetyl tryptophanamide (2 ns) and ethidium bromide in 10 mM Tris buffer (pH 7.2), containing 100 mM NaCl and 20 mM $MgCl_2$, free and DNA bound form (6 ns and 24 ns) respectively.

2.3.11 Human Genomic DNA Purification

The Human genomic DNA used as target DNA to amplify the 861 bp region of β-globin gene was purified essentially as described (John et. al, 1991). To 5 ml of blood, solution A (5 ml, 10 mM Tris pH 7.6; 10 mM KCl; 10 mM $MgCl_2$) was added. To this 120 µl of Nonidet P40 was added, mixed well and centrifuged at 3000 rpm for 10 min. The supernatant was discarded and pellet was resuspended in 700 µl of solution B (10 mM Tris pH 7.6; 10 mM KCl; 10 mM MgCl2; 0.5 M NaCl; 0.5% SDS; 2 mM EDTA) and kept for 15 min. This was transferred to 1.5 ml microcentrifuge tube containing 400 µl of buffered phenol, mixed well, centrifuged at 12000 rpm for 5 min. The aqueous phase was collected and the procedure was repeated with chloroform:isoamylalcohol. The aqueous phase was collected, to which 1 ml of chilled ethanol was added to

precipitate the DNA. The DNA precipitate was washed with 70% chilled ethanol, dissolved in sterile water for use in PCR. The concentration of DNA was calculated in terms of μg/μl using the relation : 1 O. D. = 50 μg/ ml.

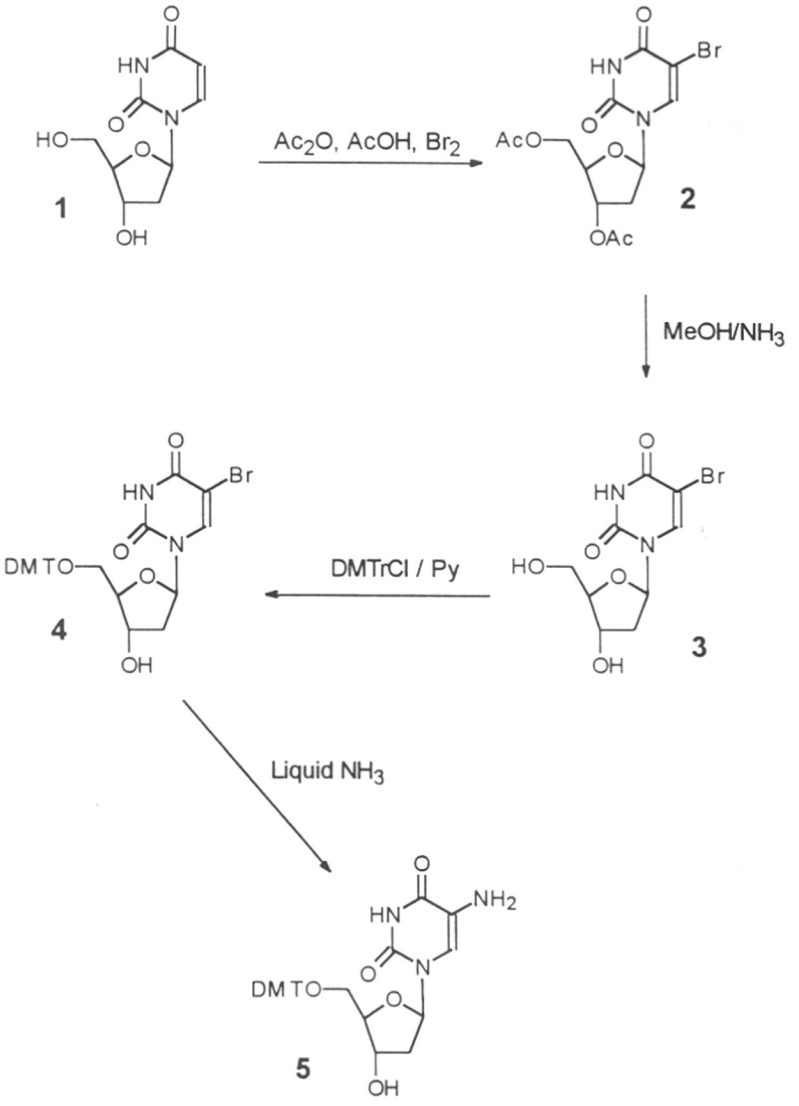
2.3.12 Polymerase Chain Reaction

PCR amplification reactions were done using different pair of primers to amplify the 861 bp globin gene fragment of Human genomic DNA. The PCR mixture (50 μl) contained 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% gelatin, 100 μM dNTPs, 100 pmoles of each primer, 0.1 μg of target DNA, and 2.5 Units of *Taq* polymerase. The reaction mixtures, overlaid with mineral oil to prevent evaporation, were subjected to 35 thermal cycles: 94 °C for 1 min, 62 °C for 1 min, and 72 °C for 2 min., followed by further keeping at 72 °C for 10 minutes and stored at 4 °C before analysis. The PCR products were analyzed by loading 20 μl of product on 1.2% agarose gel till the bromophenol blue dye reached 3/4 th of gel. The gels were visualised over UV transilluminator with and without ethidium bromide (0.5 μg/ml) staining and photographed. The densitometric scanning of gels was carried out using the UVP gel documentation system, UK.

2.4 RESULTS AND DISCUSSION

2.4.1 Synthesis of 5'-O-DMT Derivative of 5-NH_2-2'-deoxyuridine

The target compound **5** was synthesized (Scheme 1) according to the literature procedure (Visser, 1968; Barawkar & Ganesh, 1993). The 2'-dU **1** was first treated with acetic anhydride to yield the corresponding 3',5'-di-O-acetyl derivative, to which bromine in glacial acetic acid was added to give the 3'-5'-di-O-acetyl-5-bromo-2'-dU **2**. The diester **2** was hydrolysed with anhydrous methanolic ammonia to obtain 5-bromo-2'dU **3**. The formation of bromo derivative is shown by the disappearance of signal due to H5 and downfield shift of H6 in ^1H NMR. Compound **3** was converted to its corresponding 5'-O-DMT derivative **4**. This was treated with liquid ammonia in a steel bomb and the


HN
O
N
HO
O
OH
1
Ac_2O, AcOH, Br_2
Br
AcO
OAc
2
$MeOH/NH_3$
DMTrCl / Py
DM TO
4
3
Liquid NH_3
NH_2
5

SCHEME 1

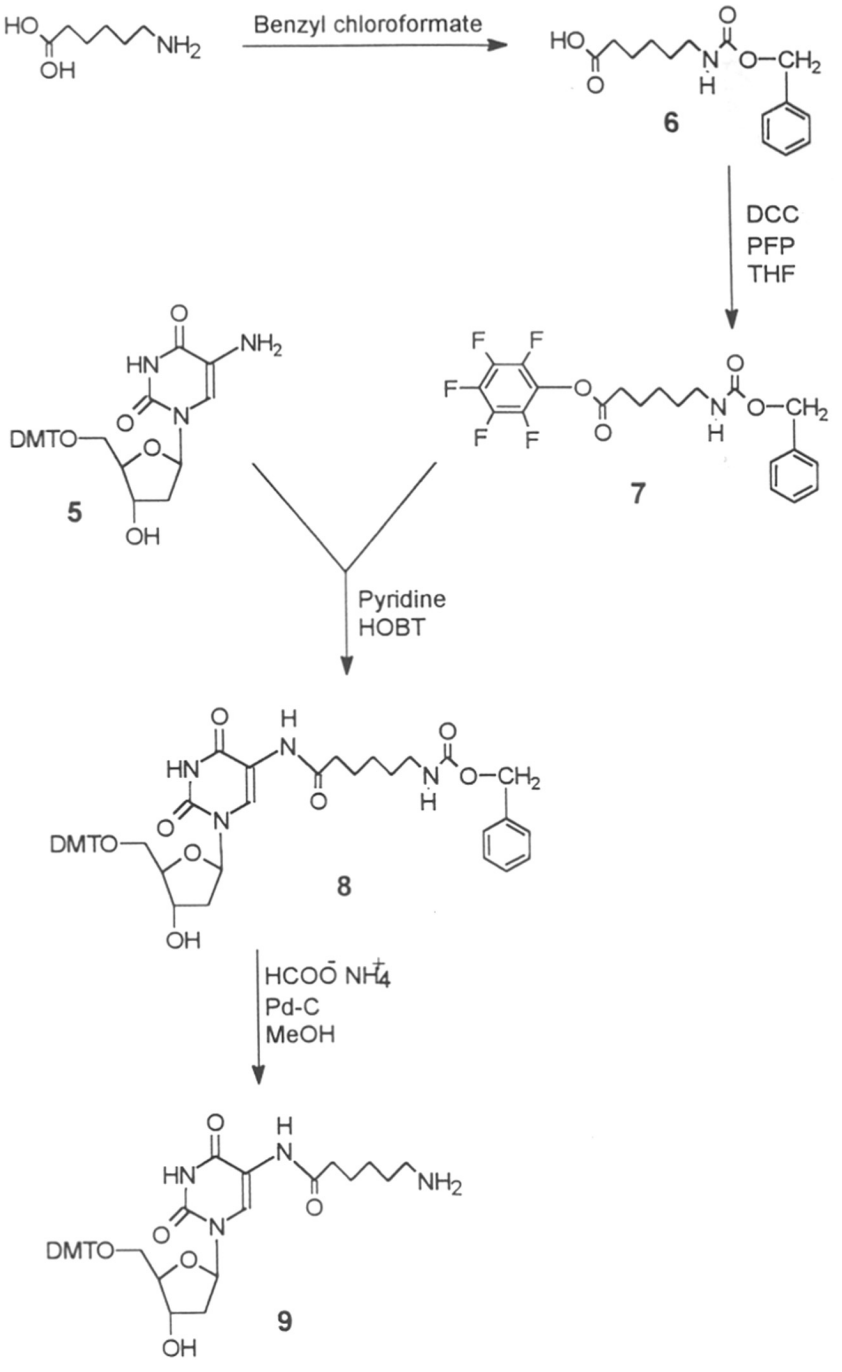
removal of excess NH_3 gave compound **5**. The displacement of 5-bromo by amino group was indicated by an upfield shift of H6 in ^1H NMR.

2.4.2 Synthesis of 5'-O-(DMT) derivative of 5-N-ω-aminohexanoyl-5-amino-2'-deoxyuridine, 9

The amino hexanoic acid was N-protected with benzyloxycarbonyl to give **6** and its PFP ester was prepared using DCC and PFP to yield **7**. This was reacted with 5-NH_2-deoxyuridine **5** in presence of HOBT to yield **8**. Figure 3 shows its ^1H NMR and Figure 4 shows ^{13}C NMR. The deprotection of benzyloxycarbonyl group was done by catalytic transfer hydrogenation using ammonium formate and Pd/C (10%) without any detritylation. The disappearance of benzylic methylene group in ^1H NMR indicated successful deprotection (Scheme 2). Attempts to protect amino group of spacer with Fmoc reagent were not successful as the detritylation was observed during the removal of Fmoc group.

2.4.3 Synthesis of Protected Carboxyfluorescein, 11

Fluorescein is used in many applications in view of its high fluorescence quantum yield (Chen et al., 1979; Munkholm et al., 1990). The synthesis of protected carboxyfluorescein **11** was carried out (Scheme 3) using literature procedure (Haralambidis et al., 1990) by heating trimellitic anhydride and recorcinol together in 1:2 ratio, which gave rise to a mixture of two products, isomers 5 and 6-carboxyfluorescein. The phenolic hydroxyl groups of the mixture were protected with trimethyl acetyl chloride to give the nonfluorescent lactone **10**. The pentafluorophenol ester **11** was prepared from **10** using DCC and PFP. The advantage of using nonfluorescent compound is that it becomes fluorescent only after deprotection of phenolic hydroxyls during ammonia treatment and hence there is less likelihood of photochemical bleaching of its fluorescence during the synthetic procedure (Haralambidis et al., 1987).


Benzyl chloroformate
6
DCC
PFP
THF
5
7
Pyridine
HOBT
8
HCOO⁻ NH₄⁺
Pd-C
MeOH
9

SCHEME 2

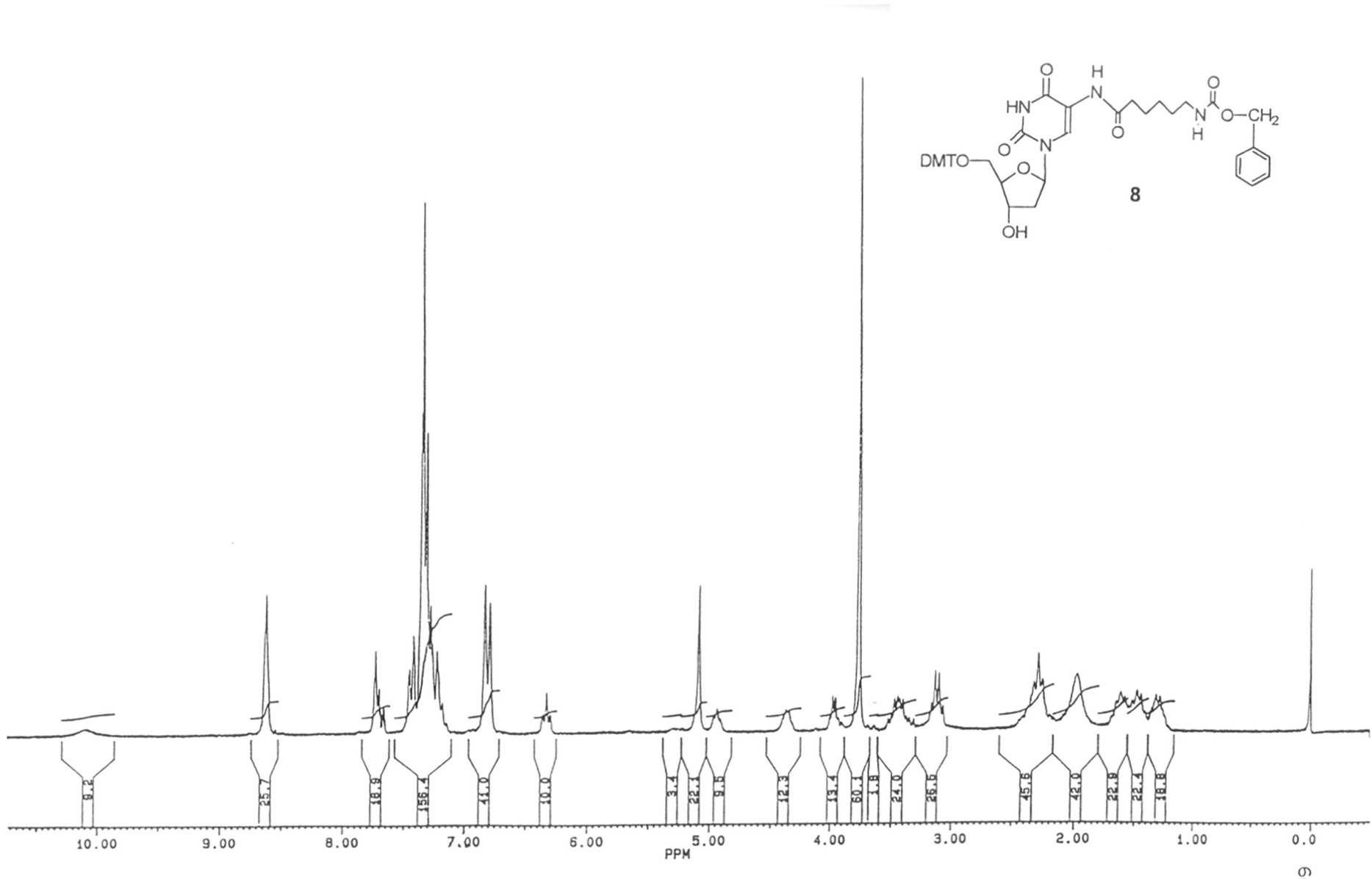


Figure 3. ^{1}H NMR spectrum of compound **8** in $CDCl_3$

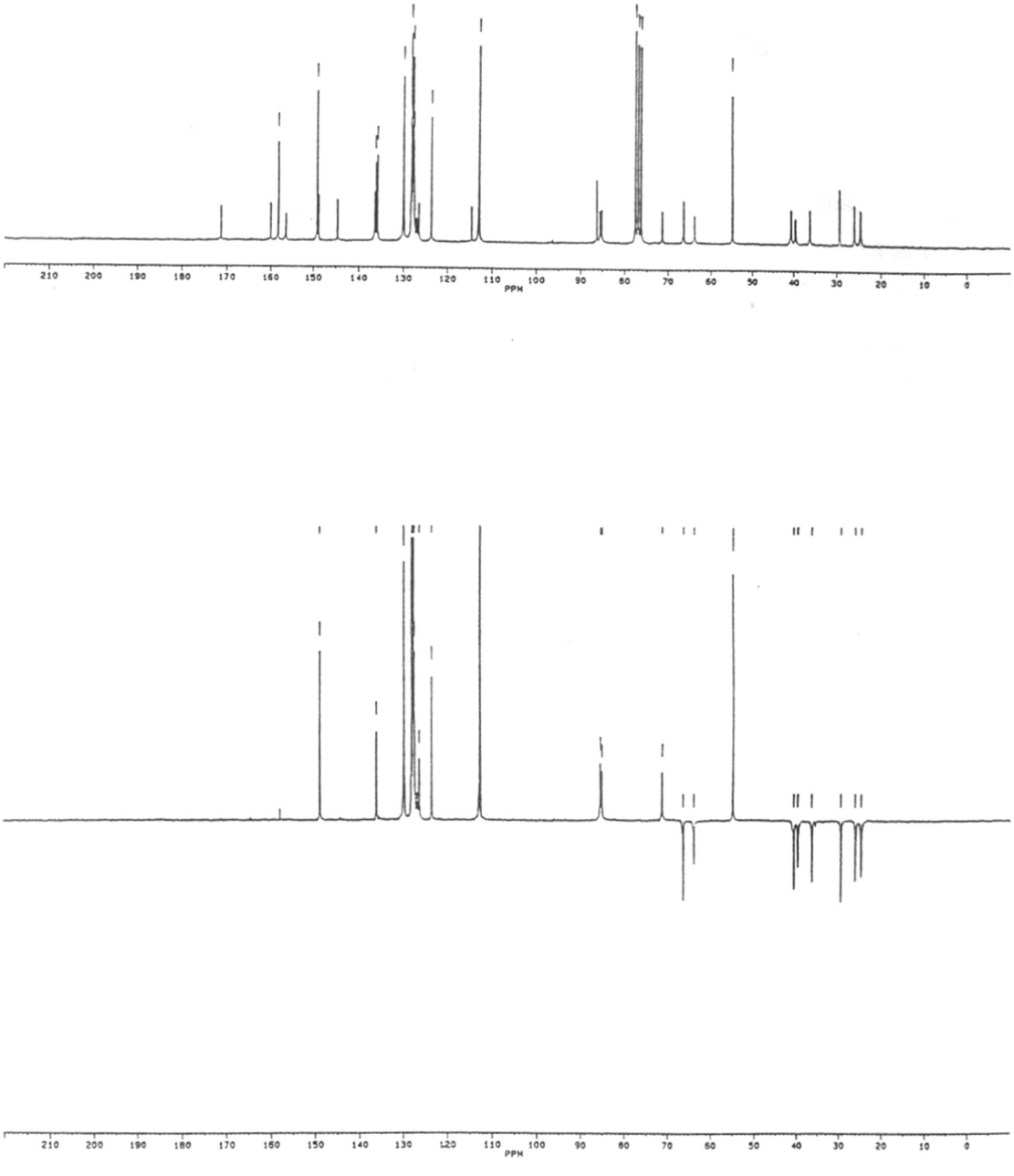


Figure 4. ^{13}C NMR spectrum of compound **8** in $CDCl_3$

2.4.4 Synthesis of Fluorescent Nucleosides With and Without Spacer

The PFP active ester of 5/6 carboxyfluorescein **11** was condensed separately with 5-NH_2-deoxyuridine **5** and 2'-dU nucleoside bearing hexanoyl spacer **9** in presence of activating agent HOBT in pyridine to yield corresponding fluorescein conjugated nucleosides **12** and **13** (Scheme 3). Figures 5 and 6 show the ^1H NMR and ^{13}C NMR of **13** respectively. These were converted into required β-cyanoethyl phspharamidite **14** & **15** . (Sinha et al., 1984).

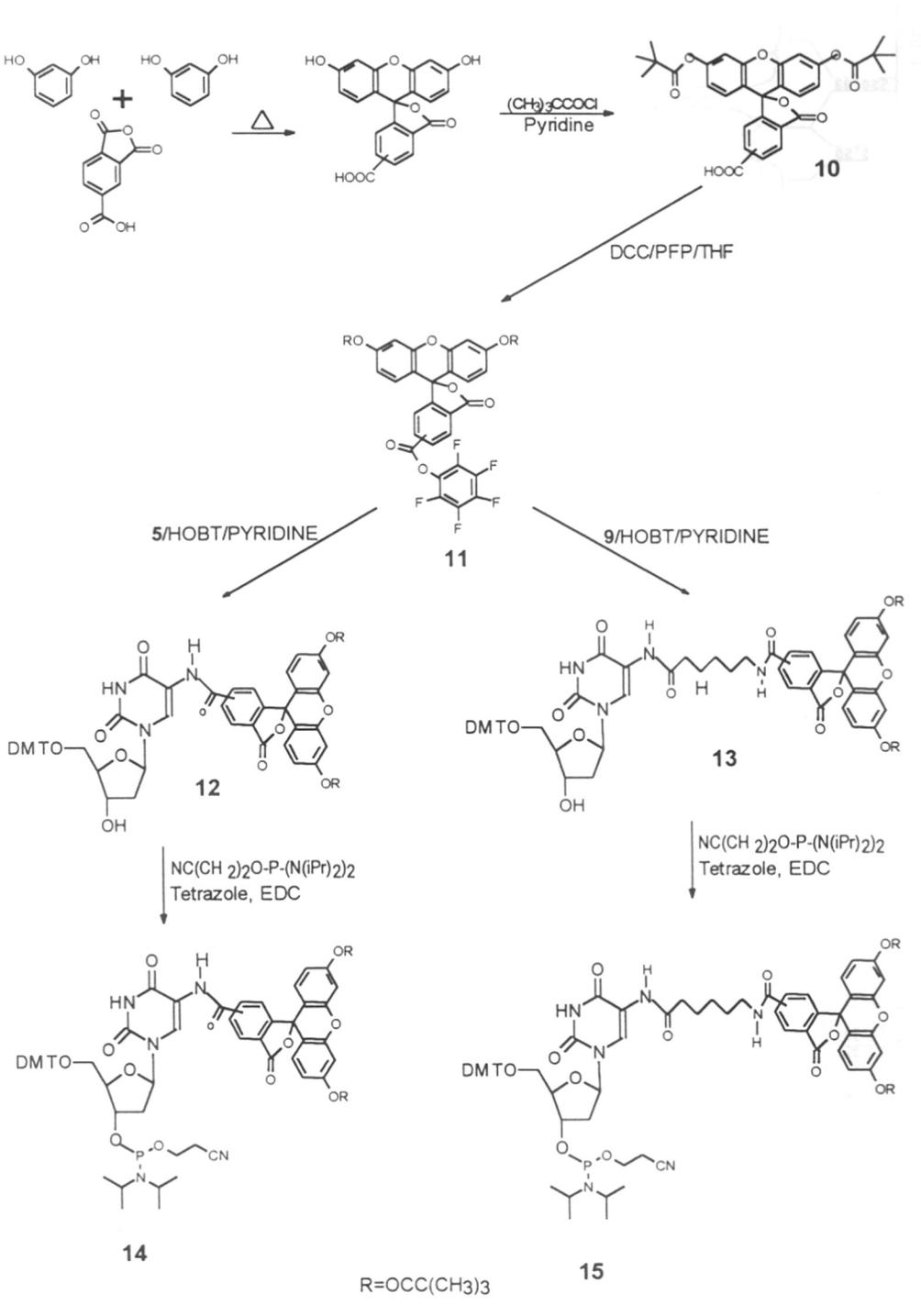
2.4.5 Synthesis and Characterization of Fluorescent Oligonucleotides.

The oligonucleotides incorporating the fluorescent base analogues **14** and **15** were synthesised using standard solid phase phosphoramidite chemistry and these sequences are given in Table 1. The parent ODN **16** & **22** are the primer sequences used for the amplification of 861 bp region of Human globin gene (Varawalla et al., 1991). The ODNs **21** and **27** are complementary sequences of **16** and **22** respectively. The coupling time for modified monomers was increased by 5 minutes and the coupling efficiency for their incorporation was found to be similar to that of unmodified monomers.

The nomenclature of ODNs synthesised and used in this study is illustrated as below. The ODNs in which fluorescein is attached directly to 5-NH_2 of dU without spacer, **L**, is absent.

e. g. **19-LF$_1$**
where
19 is ODN entry number
L indicates attachment of label through hexanoyl spacer
F indicates carboxyfluorescein attachment
1 indicates number of fluorescein molecules attached

The resin bound fluorescent oligonucleotides were deprotected by standard ammonia treatment to which 5-amido linkage is stable (Barawkar & Ganesh, 1993) with no detectable loss of conjugated fluorophore. A visible yellow-green color was observed upon ammonia treatment due to the deprotection of O-pivaloyl on carboxyfluorescein. The deprotected oligonucleotides were purified by preparative gel electrophoresis on which its direct visualization as a fluorescent band was possible.


(CH3)3CCOCl
Pyridine
10
DCC/PFP/THF
11
5/HOBT/PYRIDINE
9/HOBT/PYRIDINE
12
13
NC(CH 2)2O-P-(N(iPr)2)2
Tetrazole, EDC
NC(CH 2)2O-P-(N(iPr)2)2
Tetrazole, EDC
14
15
R=OCC(CH3)3

SCHEME 3

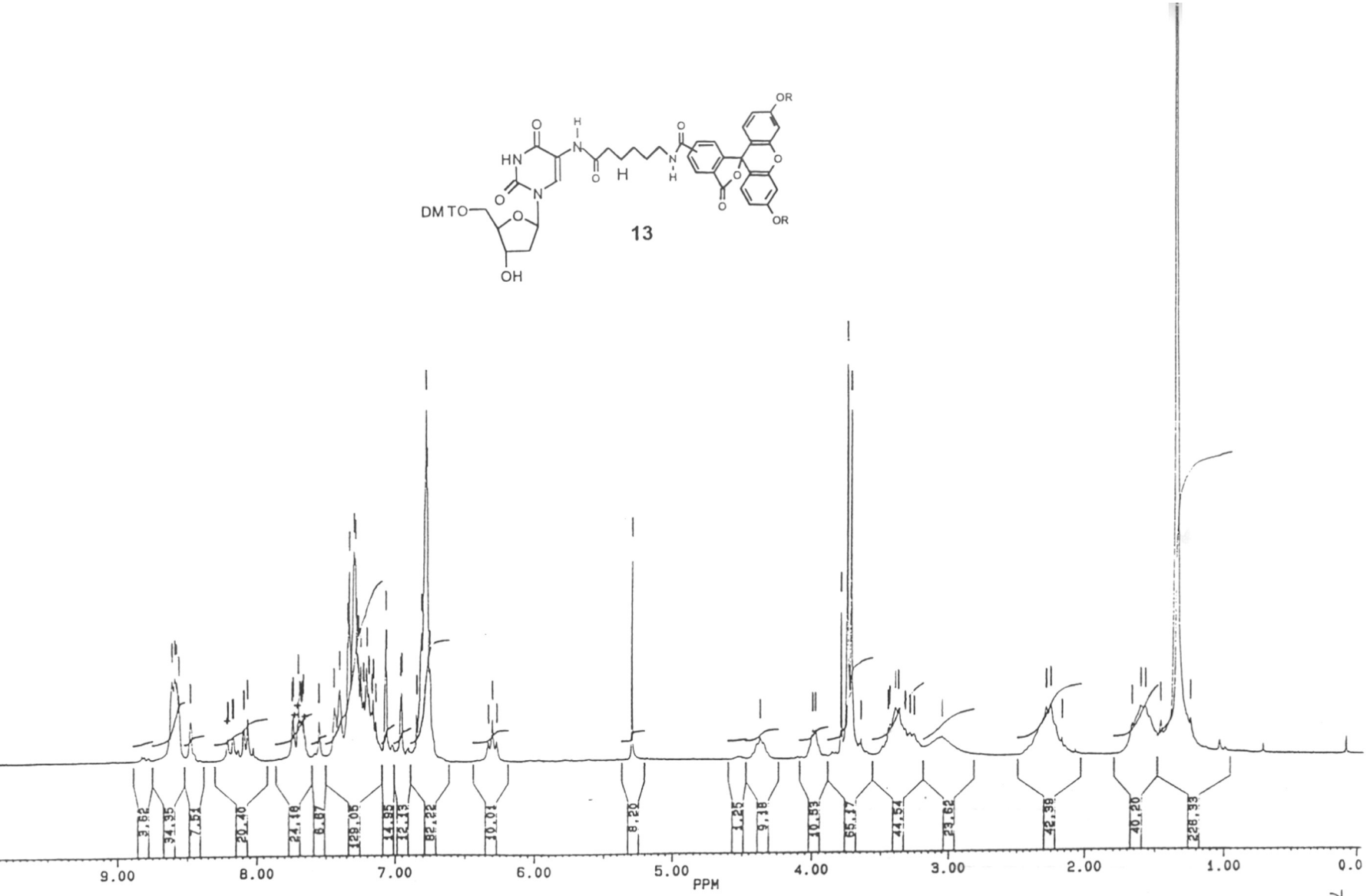


Figure 5. ^{1}H NMR spectrum of compound **13** in $CDCl_3$

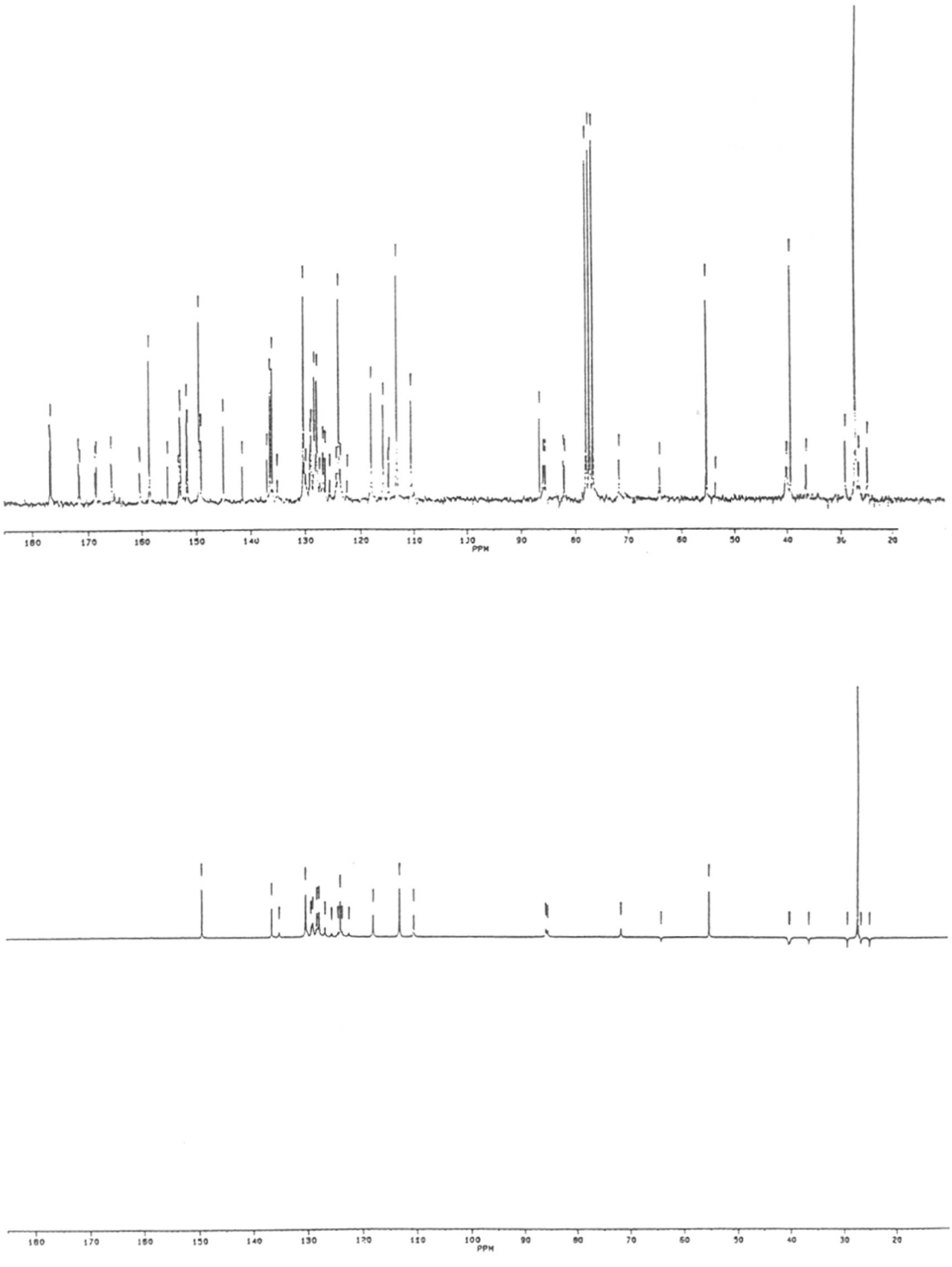


Figure 6. ^{13}C NMR spectrum of compound **13** in $CDCl_3$

The retention of fluorescein conjugates on oligonucleotides was confirmed by its absorption spectrum in 220 - 600 nm region showing the characteristic band in the region of 430-480 nm for fluorescein. Figure 7 shows the representative UV-VIS absorption spectrum of **19-LF$_1$** and **20-LF$_2$**. The ratio of absorption intensity at 480/260 nm was found to be 0.05 for monolabelled oligonucleotides and was exactly doubled to 0.1 confirming the complete retention of fluorophores. It is to be noted that the UV-VIS absorption spectrum for fluorescein labeled with and without spacer were exactly same.

The laser desorption mass spectral measurement of purified oligonuleotides **17-F$_1$** and **19-LF$_1$** (Figure 8) gave the expected mass of 7907 and 8019.4 respectively, thus confirming the retention of fluorophore in purified oligonucleotides. Moreover, the difference in the observed mass of these two oligonucleotides was found to be 112.4, which is close to the expected value of 113 , corresponding to that of spacer chain.

Table 1 Sequences of oligonucleotides

ODN									
16	5'- T	CAA	TGT	ATC	ATG	CCT	CTT	TGC	ACC
17-F$_1$	5'- **X**	CAA	TGT	ATC	ATG	CCT	CTT	TGC	ACC
18-F$_2$	5'- **X**	CAA	TG**X**	ATC	ATG	CCT	CTT	TGC	ACC
19-LF$_1$	5'- **Y**	CAA	TGT	ATC	ATG	CCT	CTT	TGC	ACC
20-LF$_2$	5'- **Y**	CAA	TG**Y**	ATC	ATG	CCT	CTT	TGC	ACC
21	5'- G	GTG	CAA	AGA	GGC	ATG	ATA	CAT	TGA
22	5'- T	GAG	TCA	AGG	CTG	AGA	GAT	GCA	GGA
23-F$_1$	5'- **X**	GAG	TCA	AGG	CTG	AGA	GAT	GCA	GGA
24-F$_2$	5'- **X**	GAG	**X**CA	AGG	CTG	AGA	GAT	GCA	GGA
25-LF$_1$	5'- **Y**	GAG	TCA	AGG	CTG	AGA	GAT	GCA	GGA
26-LF$_2$	5'- **Y**	GAG	**Y**CA	AGG	CTG	AGA	GAT	GCA	GGA
27	5'- T	CCT	GCA	TCT	CTC	AGC	CTT	GAC	TCA

X = 14 and **Y = 15**

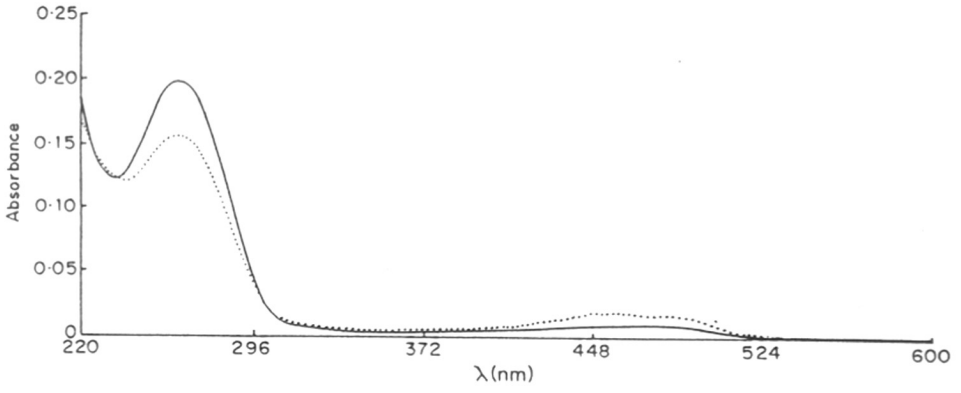


Figure 7. UV-VIS spectra of **19-LF$_{10}$** (—) and **20-LF$_2$** (.....)

The HPLC analysis of modified monolabelled oligonucleotides showed two well resolved peaks of equal intensity, each of which can be assigned to ODN carrying one of the isomers of 5/6-carboxyfluorescein (Figure 9a). Similarly, dilabelled ODN should have four isomeric oligonucleotides each representing one of the four different possible combinations of carboxyfluorescein isomers (5,5; 5,6; 6,6; and 6,5). Indeed, as shown in Figure 9b, all these four peaks were seen by HPLC analysis with **18-F$_2$** which were of almost equal intensity. However, in **20-LF$_2$**, only three peaks were resolved and one of the peaks had double the intensity of the other two (Figure 9c), indicating the coelution of two peaks. The retention times of oligonucleotides are given in Table 2. All these peaks showed absorption maxima at 260 nm as well as 500 nm, thus confirming the retention of fluorophore on oligonucleotides. The presence of such 5/6-carboxyfluorescein isomers in ODN does not affect the fluorescence detection and can be used without any separation of isomers.

Table 2. Elution times of different ODNs on HPLC #

ODN	R_t time (min)
17-F1	9.80, 10.62
18-F2	9.95, 10.53, 10.94, 11.37
19-LF1	11.12, 11.99
20-LF2	12.12, 12.79, 13.37

HPLC gradient: 10-30% acetonitrile in 0.1 M TEAA (pH 7.0) for 20 min.

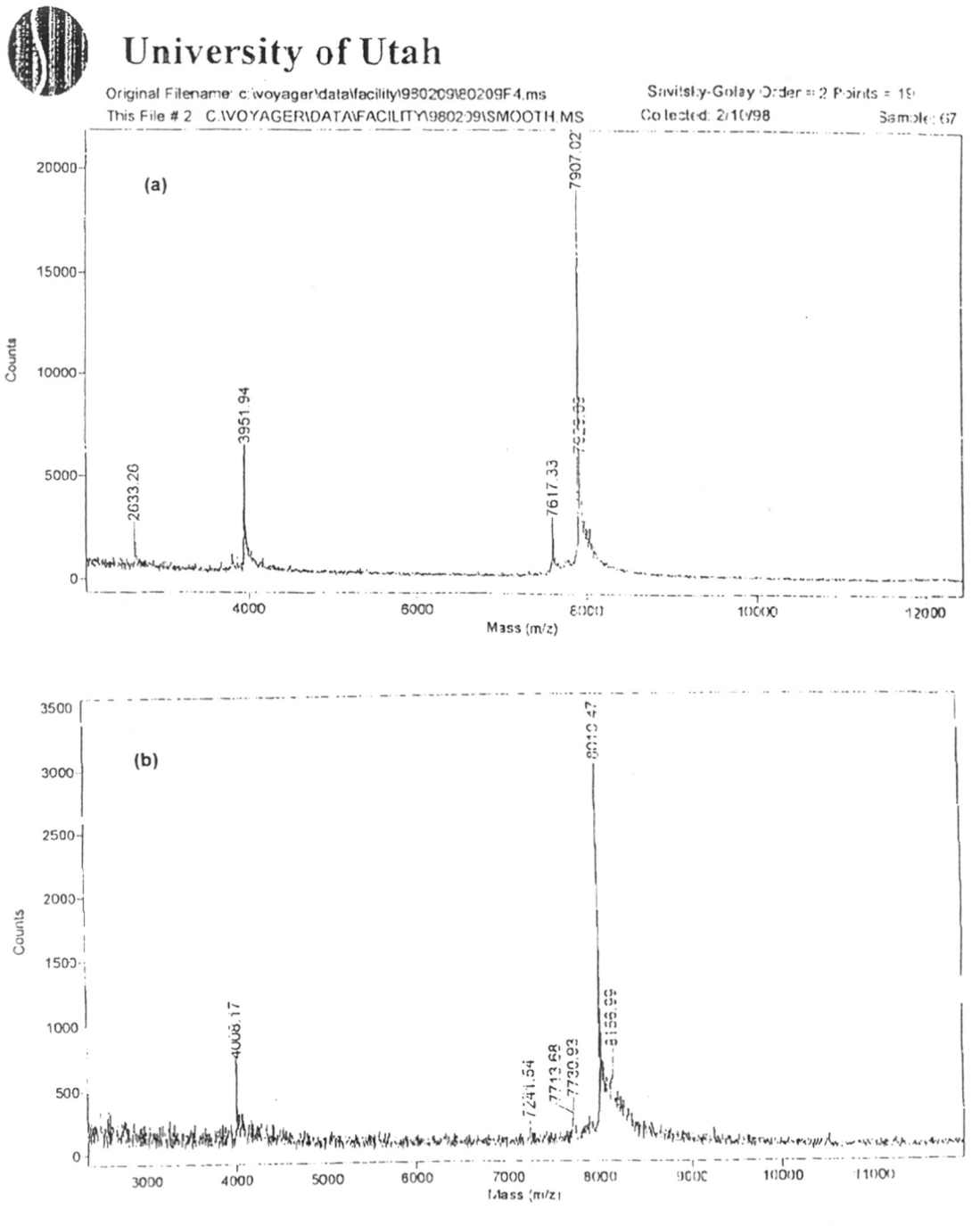


Figure 8. Laser desorption (MALDI) mass spectra of (a) **17-F$_1$** and (b) **19-LF$_1$**

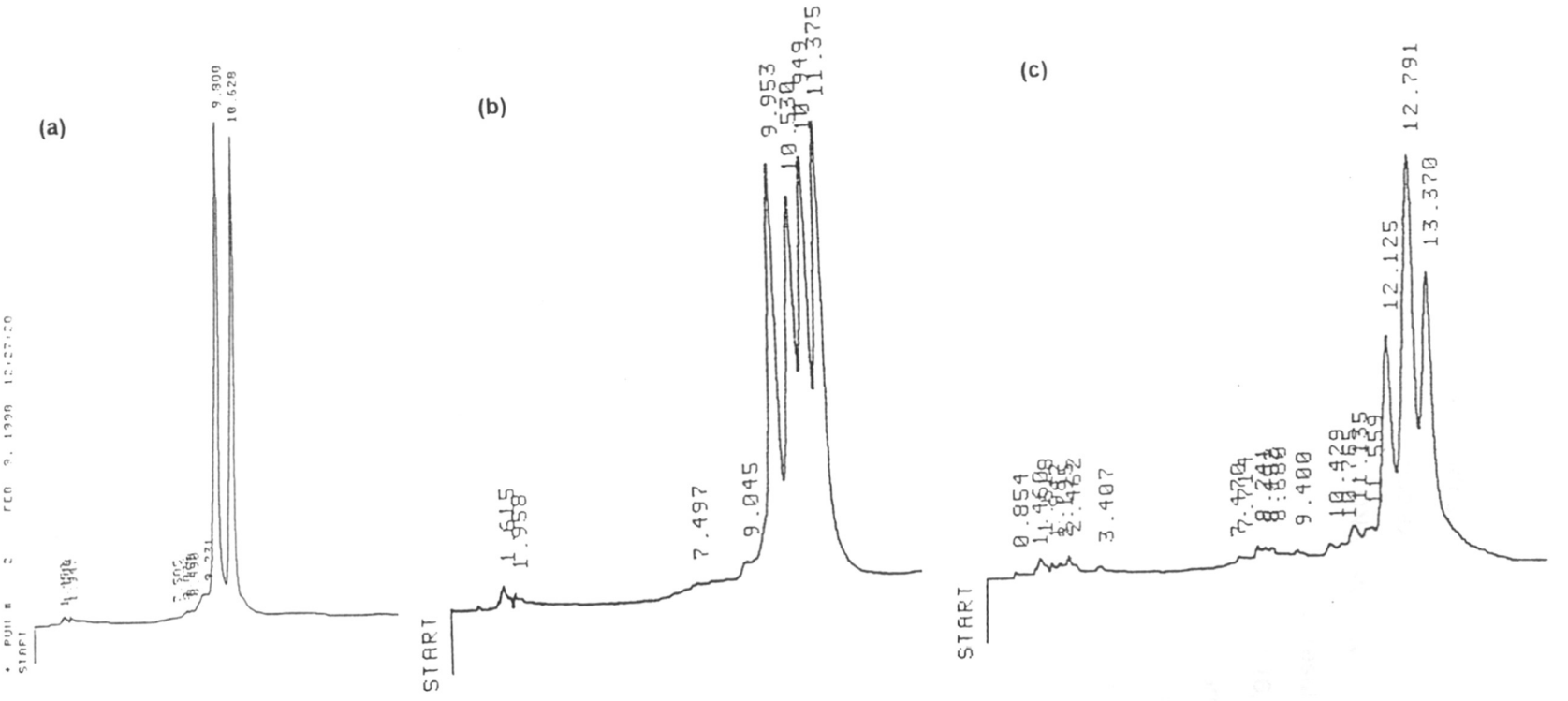


Figure 9. HPLC profiles of (a) **17-F$_1$** (b) **18-F$_2$** and (c) **20-LF$_2$**.

2.4.6 Duplex Stability of Modified Oligonucleotides

The DNA modifications on base affect the duplex stability of DNA. To see the effect of fluorescein labelling on duplex DNA stability, the thermal denaturation experiments of ODN duplexes were done. The results of these studies in the form of melting profiles of the duplexes derived from **16**, **17-F_1** , **18-F_2** , **19-LF_1** and **20-LF_2** with common complementary strand **21** are shown in Figure 10. The duplex T_m's of the oligonucleotides used in this study are given in Table 3. The results showed that monolabelled duplexes with direct attachment of label are destabilised by 3 °C (entry 2,7) and those with spacer by 1 °C (entry 4,9) compared to unmodified one (entry 1). The dilabelled duplexes with direct attachment of label are destabilised by 5-6 °C (entry 3, 8) and with spacer by 3 °C (entry 5, 10) compared to unmodified one.

Table 3 . Melting temperatures (T_m) of duplexes

Entry	Duplex	T_m °C	Δ T_m °C	Entry	Duplex	T_m °C	Δ T_m °C
1	**21:16**	69		6	**27:22**	72	
2	**21:17-F_1**	66	(-3)	7	**27:23-F_1**	69	(-3)
3	**21:18-F_2**	64	(-5)	8	**27:24-F_2**	66	(-6)
4	**21:19-LF_1**	68	(-1)	9	**27:25-LF_1**	71	(-1)
5	**21:20-LF_2**	66	(-3)	10	**27:26-LF_2**	69	(-2)

T_m = ± 0.5 ^{0}C, Δ T_m compared to the unmodified control duplex

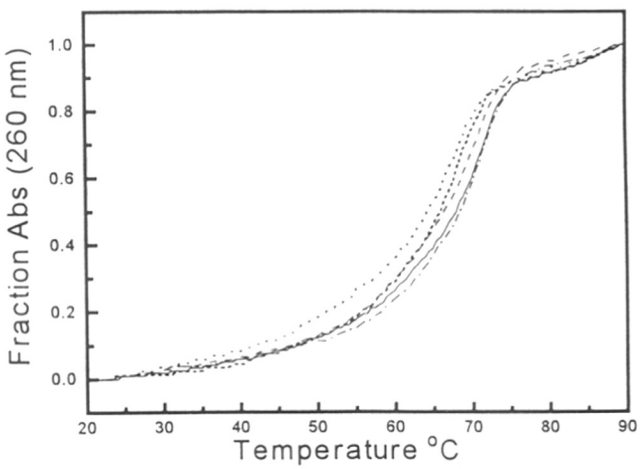


Figure 10. Melting profile of DNA duplexes **21:16** (-.-.-.-); **21:17-F_1** (——), **21:18-F_2** (....); **21:19-LF_1** (—); and **21:20-LF_2** (----).

These results indicate that by direct attachment of fluorophore caused more destabilisation compared to those in which the fluorophore is linked through spacer. The direct rigid attachment of the bulky fluorophore to the base in the major groove of DNA may cause slight distortion of the helix leading to destabilization of duplex. In contrast, the introduction of a spacer arm keeps the bulky fluorophore away from the helix, with a decreased effect on the stability of duplex. In any case, considering the size of fluorescein label, the observed destabilization with spacer fluorophore is not significant even with internal labeling since the site of attachment i.e. C-5 of deoxyuridine is not involved in Watson-Crick hydrogen bonding during hybridization. In any DNA probe design, the attachment of reporter group should have minimal effect on hybridisation properties of duplex. The probes described here are suitable for such studies as they can be multiply labeled to enhance the detection sensitivity without affecting the duplex stability significantly. Among the different sequences, those bearing fluorophore through spacer were better in terms of duplex stability.

2.4.7 Fluorescence Properties of Oligonucleotides

The fluorescence spectra of the various labeled oligomers were qualitatively similar, with an excitation maximum at 495 nm and the emission maximum at 520 nm. The relative fluorescence intensities of various derivatives showed some differences (Figure 11A-D). The Table 4 shows the relative fluorescence intensities of labeled oligonucleotides compared to free carboxyfluorescein under same concentration.

Table 4. Relative fluorescence intensities of oligos#

	Relative fluorescence/CF
Carboxyfluorescein (CF)	1
17-F_1	0.61
18-F_2	0.42
19-LF_1	0.85
20-LF_2	1.4

All under same conditions and concentration

The dilabelled **18-F_2** ODN showed much less fluorescence compared to monolabelled ones. The ODN containing fluorescein moiety attached via spacer

showed 85% fluorescence compared to free carboxyfluorescein which is in accord with the results described by Haralambidis et al., (1987). However, ODN

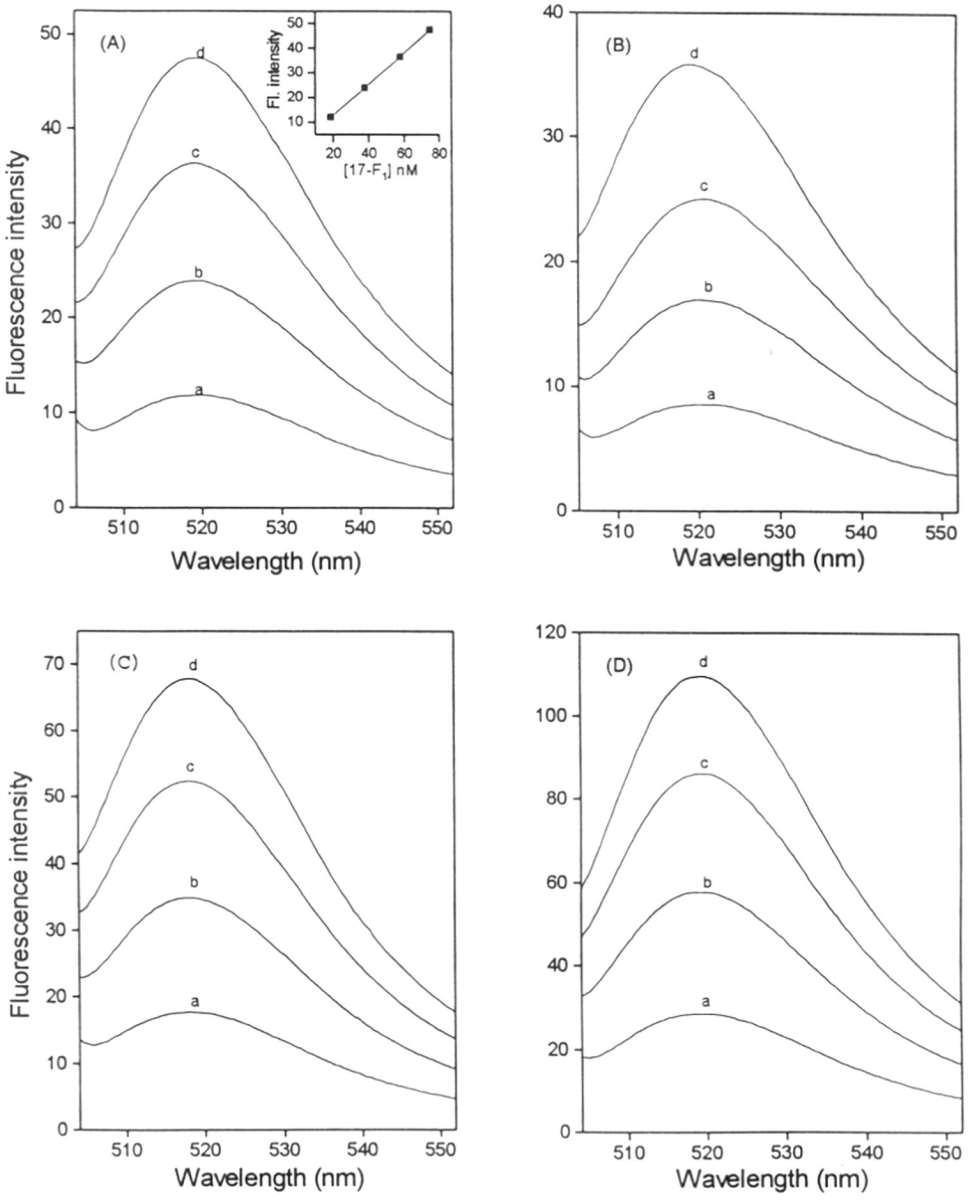


Figure 11. The concentration dependent fluorescence emission spectra of (A) **17-F$_1$**, (B) **18-F$_2$**, (C) **19-LF$_1$** and (D) **20-LF$_2$**. In each spectra the a, b, c, and d corresponds to ODN concentrations 19 nm, 38 nm, 58 nm and 75 nm respectively. The inset in (A) shows the plot of oligo conc. vs. fluorescence intensity derived from the spectral data of (A).

having two fluorescein molecules attached via spacer showed 69% fluorescence intensity per label compared to free CF. In contrast, the monolabelled ODN having fluorophore attached to base without spacer had only 61% of fluorescence compared to free CF and the dilabelled ODN showed only 21% of fluorescence per label compared to free CF.

The covalent attachment of fluorescent label to DNA (containing numerous functional groups), can provide many pathways for nonradiative relaxation of the fluorophore's electronic excited state (Telser et al., 1989; Fick et al., 1982). The multiple labeling density of fluorophore has been shown not to necessarily increase the sensitivity of probe in a linear way (Wallner et al., 1993).

In principle, two types of molecular associations may influence these processes, (I) fluorophore-fluorophore and (ii) fluorophore DNA interactions; these may occur both at intermolecular and intramolecular levels, resulting in a net loss of fluorescence in labeled oligomers compared to the free label alone. In order to delineate the relative effects of these interactions and their dependence on the position and distance between fluorophore and DNA, the fluorescence of ODNs **17-F_1**, **18-F_2**, **19-LF_1** and **20-LF_2** were recorded under different concentrations. As shown in Figure 11A (inset), fluorescence intensity increased linearly with the increase in ODN concentration. This suggests that in the concentration ranges used in these studies, intermolecular (interduplex) DNA-fluorophore and fluorophore-fluorophore associations do not contribute significantly to the quenching process.

The overall results indicate that the extent of fluorescence quenching is much higher in ODNs having fluorophore attached without spacer. The most likely mechanism for this quenching of the fluorescence is non-radiative energy transfer- through space via dipole-dipole interaction (Morawetz, 1988). This can occur since there is a considerable overlap between the excitation and emission spectra of fluorescein, leading to an interaction between an excited molecule and a nearby chromophore in the ground state. The present results on fluorescence

quenching of fluorescein attached to base with and without spacer indicate that non-radiative energy loss is more with fluorescein moiety attached to base without spacer perhaps due to restricted freedom of rotation. The possibility of energy transfer occurring between DNA bases and fluorophore cannot be ruled out as the quenching is more with fluorophores which are directly linked to bases.

2.4.8 Quenching Studies

Iodide is known to be a dynamic, or collisional fluorescence quencher. The degree of quenching of fluorescence by KI gives information about the accessibility of the fluorophore. The quenching experiments were performed with 150 nM of fluorescent duplexes by successive addition of KI in the concentration range of 0-0.6 M. The quenching constant obtained using stern-Volmer plots are given in Table 5. The quenching constants were higher for ODNs labelled through spacer. The results indicate direct attachment of label to the ODN in the middle of sequence (entry 3 & 4) reduces its accessibililty to quencher compared to the one with spacer. The difference between the quenching constants of monolabeled oligos with and without spacer was marginal (entry 1,2) which may be due to the DNA breathing at the ends of DNA which renders fluorophore exposed to quencher irrespective of mode of attachment.

Table 5. Quenching constants for duplexes obtained from stern-Volmer plot

Entry no.	ODN duplex	K_q (M^{-1})
1	**21:17-F_1**	0.746
2	**21:19-LF_1**	0.778
3	**21:18-F_2**	0.892
4	**21:20-LF_2**	1.075

2.4.9 Time Dependent Duplex Formation Monitored by Change in Fluorescence Intensity

The duplex formation by fluorescent ODNs with their complementary sequence was studied at 20 °C using time drive experiments. Upon excitation at 495 nm, the fluorescence emission intensity at 520 nm was monitored for fifteen minutes immediately after the addition of complementary strand. As shown in Figure 12 , the increase in fluorescence intensity observed using oligos with

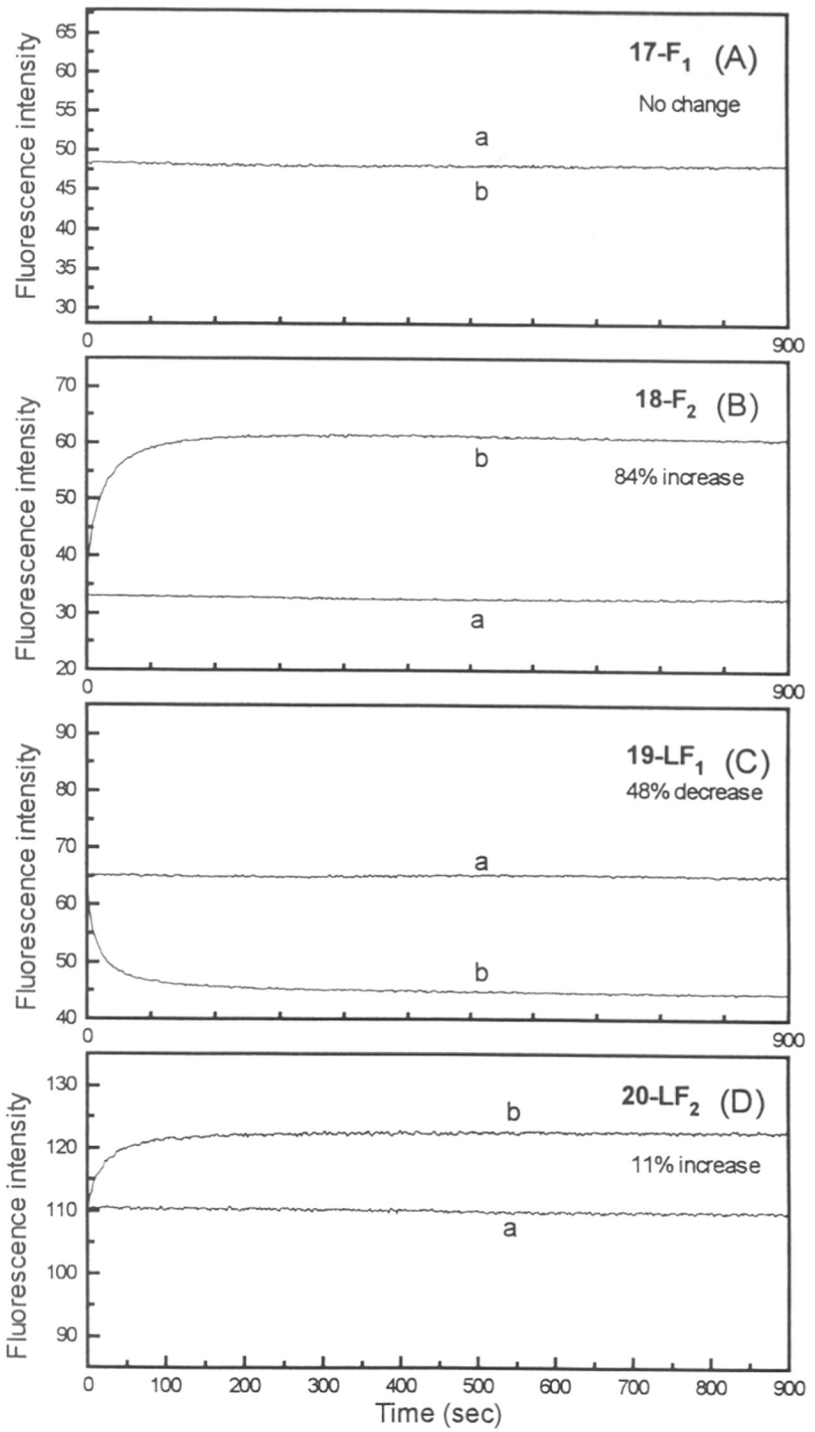


Figure 12. Time drive spectra of duplex formation monitored by change in fluorescence emission intensity at 520 nm upon excitation at 495 nm. (A) **17-F$_1$**; (B) **18-F$_2$**; (C) **19-LF$_1$** and (D) **20-LF$_2$** In each spectra, 'a' corresponds to single strand fluorescence emission while 'b' corresponds to fluorescence monitored immediately after addition of complementary strand **21**.

directly attached label was insignificant with monolabeled ones (Figure 12A), while an increase of 84% was observed using dilabelled oligo upon hybridisation (Figure 12B). This increase may primarily arise from a change in local environment sensed when the fluorescein moiety is present in the middle of sequence, but not when present at the end. Upon hybridisation, the fluorescein moiety attached directly to the base at 5'-terminal end in the sequence is not likely to sense same duplex environment experienced by the label present in the interior of sequences. In contrast, the oligo **19-LF$_1$**, which is monolabelled through spacer, showed 48% decrease in fluorescence upon duplex formation (Figure 12 C). It is likely that the fluorophore attached to the terminal base with spacer may adopt a specific orientation induced by duplex formation, such as stacking with terminal base pair (Kumke et al., 1995). However, the dilabelled ODN with fluorescein attached through spacer showed an increase in fluorescence intensity by 11%. Although the observed increase is only 11%, the actual increase in intensity by the fluorescein moiety present in the middle of sequence may be much higher as the terminal fluorophore causes a decrease in fluorescence upon hybridisation. There were no significant changes in fluorescence upon addition of either unlabelled ODN with same or noncomplementary sequence, indicating that changes observed are specific to duplex formation between Watson-Crick complementary strands.

It is known that fluorescein does not intercalate with DNA (Murakami et al., 1991) and its fluorescence is significantly quenched by acidification (pH 5-6) with an increased quantum yield at higher pH (8.0) (Aguirre et al., 1986). Such changes in fluorescence intensity have been used in understanding the protein-protein recognition (Patel et al., 1990). In order to understand the effect of pH on fluorescence intensity, the fluorescence spectra were recorded in the buffers with pH range of 5-9. Figure 13 shows the change in fluorescence intensity of **17-F$_1$**& **19-LF$_1$** under different pH conditions. Tris buffer was used in the pH range 7-9 while MES buffer was used in the pH 5-6 range. Both ODNs showed a

significant quenching in the acidic region with almost nondetectable fluorescence at pH 5 while the optimum fluorescence intensity was observed in basic medium at pH 8.0. Further increase in pH to 9.0 decreased the fluorescence intensity.

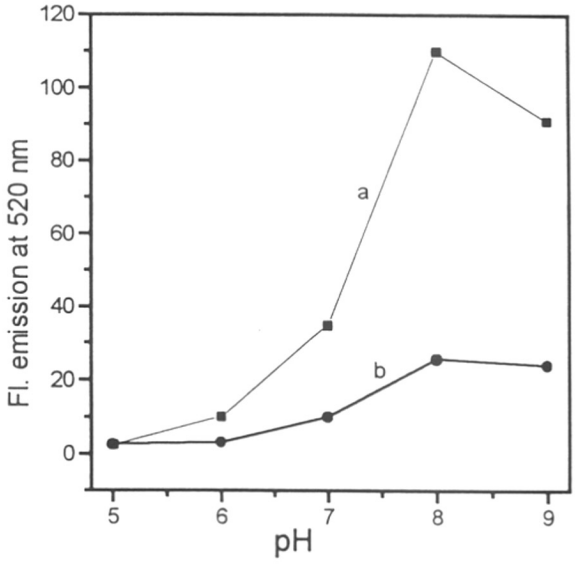


Figure 13. The pH dependent fluorescence emission intensity of (a) **19-LF$_1$** and (b) **17-F$_1$** in different pH buffers. The samples were excited at 495 nm and fluorescence emission was monitored at 520 nm.

In the present work, the attachment of fluorophores at C-5 position of deoxyuridine places fluorophores in the major groove of duplex DNA. An increase in fluorescence intensity upon hybridisation, may indicate an enhanced basicity in the microenvironment around the fluorophore. It may be noted that in time dependent duplex formation study, the observed percentage increase in fluorescence intensity was much higher in the duplex having fluorophore directly attached to base (Figure 12B), where the fluorophore is firmly placed in major groove and thus could be more sensitive to the local environment in the major groove compared to those in which fluorophore is attached through flexible spacer arm.

2.4.10 Steady State Anisotropy

An increase in fluorescence anisotropy value is associated with a decrease in motional freedom (restricted motion) of the probe in solvent

(Lakowicz, 1983). The data on anisotropy given in Table 6 indicate that anisotropy values were higher in ssODN as well as duplexes in which fluorescein is attached without spacer (entry 1,2) compared to the one with spacer chain (entry 3,4). The direct attachment significantly restricts the range of conformations available for fluorophore when it is attached to DNA while attachment via spacer will have more freedom of rotation and this may explain the observed differences in the anisotropy of same ODN when fluorophore is attached with and without spacer. The anisotropy values are higher in duplexes compared to single strand. The duplex formation in DNA enhances the rigiditiy of the probe thereby restricting its rotation compared to that in single stranded DNA. This leads to an increase in anisotropy values.

Table 6. Anisotrpy values for ODNs

Entry	ODN	ss	ds
1	**17-F_1**	0.060	0.090
2	**18-F_2**	0.071	0.081
3	**19-LF_1**	0.047	0.075
4	**20-LF_2**	0.050	0.061

The reaction conditions are given in experimental section. Oligo **21** was used as complementary strand for duplex formation.

The anisotropy values for both ss and ds for the ODNs having same sequence were higher when the fluorophore is directly attached compared to the one in which fluorophore is attached through spacer. The increased rigidity of DNA strand upon hybridisation is transfered more efficiently to the directly linked fluorophore than the one attached with spacer. The rigidity of direct attachment of fluorophore minimises its contribution to anisotropy values while with flexible spacer the anisotropy values are the net effect of mobility of DNA as well as fluorophore.

The fluorescence anisotropy values were also measured with duplex DNA in the temperature range of 20-45 °C. The results shown in Figure 14 indicate that the anisotropy decreases with increase in temperature, suggesting an increased diffussional mobility upon denaturation. The decrease in anisotropy was more in

the ODNs having fluorophore attached via spacer (Figure 14: c, d) which may be due to the increased rotation of fluorophore itself.

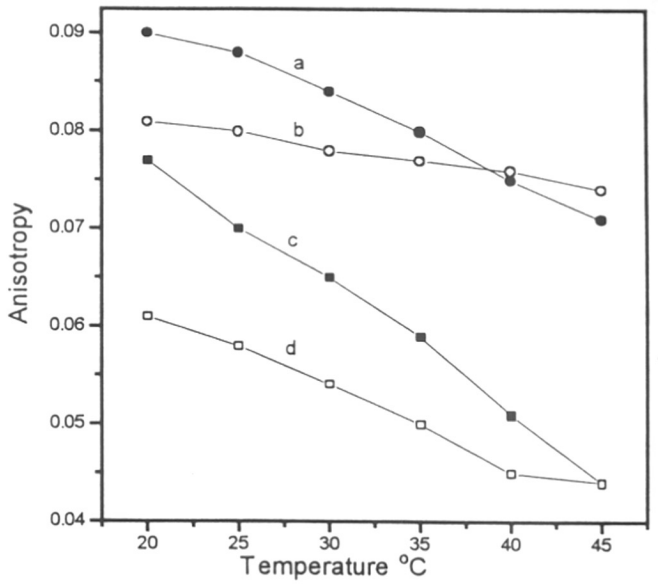


Figure 14. The effect of temperature on anisotropy values of various labelled duplexes. a:**21:17-F_1**; b:**21:18-F_2**; c:**21:19-LF_1**; d:**21:20-LF_2**

The detection of DNA hybridisation *in situ* without separation of unhybridised single stranded species is reported using fluorescence anisotropy (Kumke et al., 1995; 1997) The further enhancement in detection of hybridisation using anisotropy values specific to double stranded DNA were achieved by binding of *Eco*RI to the duplex containing its recognition sequence. The oligonucleotides (**22** and its analogs) used in present study, contain *Hinf*I recognition site (GAGTC) towards 5'-end of duplex.

22- 5'-T**GAGTC**AAGGCTGAGAGATGCAGGA - 3'

The ODNs **23-F_1** and **25-LF_1** have fluorophore on 5' terminal base, adjacent to *Hinf*I recognition sequence while the ODNs **24-F_2** and **26-LF_2** have two fluorophores, one on the T base which is part of *Hinf*I recognition and other on 5' terminal base (See table 1). These ODNs were thus useful to see the effect of *Hinf*I binding to duplex by studying anisotropy values. The effect of *Hinf*I addition

to various duplexes was followed by change in anisotropy values and the results are shown in Table 7. Only the duplex **27:23-F_1** which is monolabelled outside the recognition sequence without spacer showed significant increase in anisotropy indicating that the binding of enzyme to its recognition site leading to further restriction of the mobility of probe in duplex. All other duplexes showed insignificant changes in anisotropy values upon addition of enzyme. The enzyme does not bind to ODN duplexes (**24-F_2** and **26-LF_2**) in which the fluorophore is present in recognition site of enzyme while the ODN duplex (**27:25-LF_1**) having a fluorophore on terminal base linked via spacer is less sensitive to such events as seen by fluorescence anisotropy data. This is due to a larger inherent freedom of rotation of label linked through a spacer (Netzel et al.,1995).

Table 7. Steady state Fluorescence Anisotropy#

ODN	ss	ds	ds+ *Hin*fI
23-F_1	0.055	0.068	0.081
24-F_2	0.061	0.077	0.077
25-LF_1	0.037	0.060	0.063
26-LF_2	0.038	0.046	0.047

The anisotropy readings for ss and duplex were taken (at pH 7.5 and 20 °C) in same buffer as used for enzyme binding. Oligo **27** was used as complementary strand.

2.4.11 Netropsin Binding to Fluorescent ODNs

Netropsin is a DNA binding peptide antibiotic with base specificity towards A/T regions of double stranded DNA, with a non-intercalative, minor groove binding mode (Berman et al., 1979; Kopka et al., 1985). Barawkar & Ganesh, (1994;1995) have studied the binding event of netropsin in minor groove using the dansyl fluorophore located in major groove by direct attachment of fluorophore to C-5 position of deoxyuridine. The oligonucleotides used in present work, contain A/T rich stretches which are potential binding site for netropsin. The studies were carried out to check the sensitivity of carboxyfluorescein (CF) located in major groove of DNA to monitor the effect of netropsin binding in minor groove.

The addition of netropsin to ODN duplexes (**21:18-F_2** and **21:20-LF_2**) resulted in a decrease of carboxyfluorescein fluorescence. No change in

fluorescence was observed upon addition of netropsin to free carboxyfluorescin. As shown in Figure 15, the decrease in fluorescence intensity was observed upto

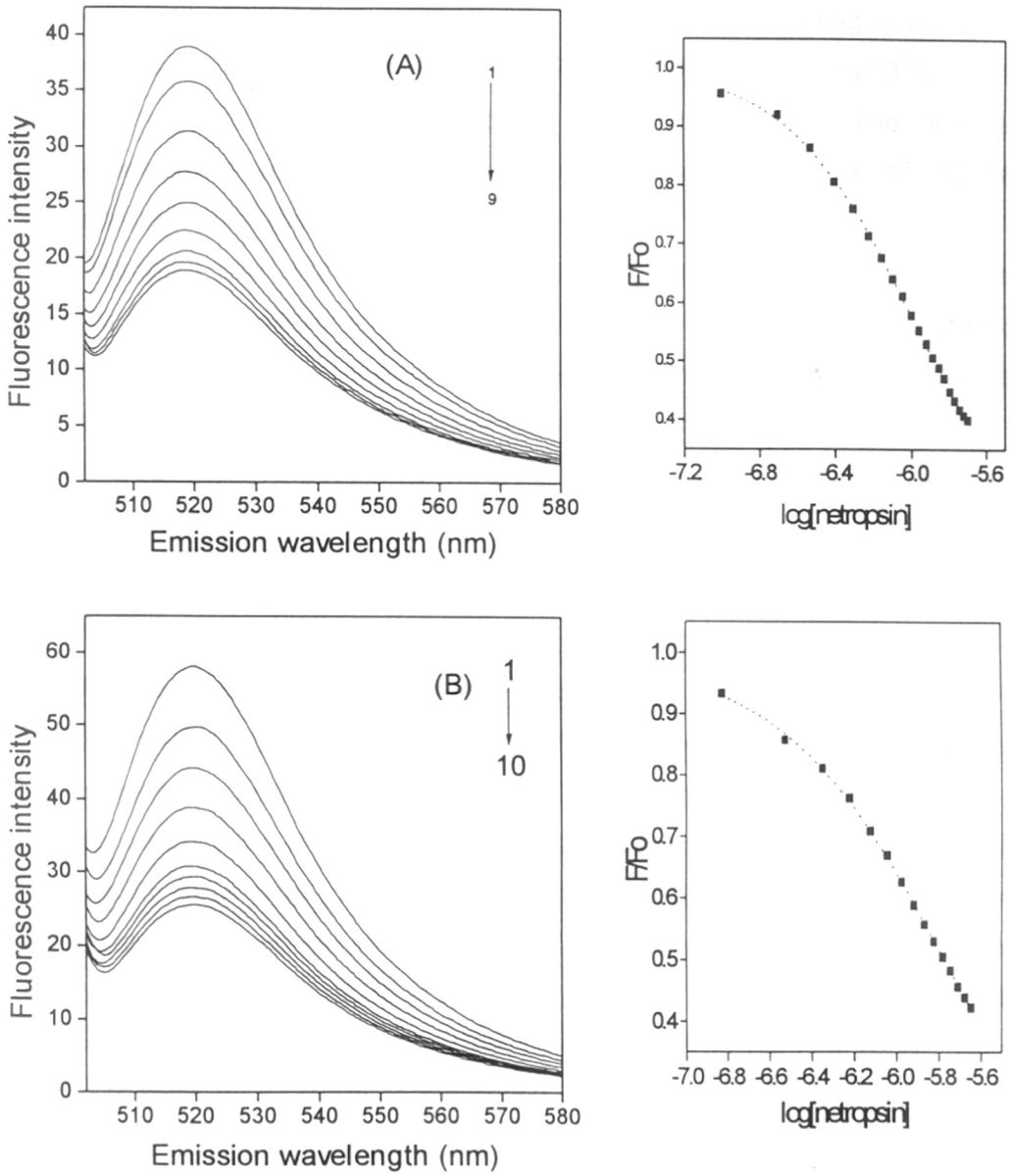


Figure 15. Changes in the fluorescence spectra of upon addition of increasing concentrations of netropsin to (A) **21:18-F_2** and (B) **21:20-LF_2**. The plots on the right show the corresponding binding isotherms from the data in (A) and (B).

the addition of three equivalents of netropsin. However in both these cases, no complete saturation of drop in fluorescence was observed. The binding affinity of

netropsin was calculated from a plot of normalised fluorescence intensity at λ_{em} (at each titration point) against log of netropsin concentration (Figure 15). In such a plot, the antilog of the point on X-axis corresponding to the midpoint of the slope gives the binding constant (K_a). The observed binding constant of netropsin with **21:18-F_2** was $1x10^6\ M^{-1}$ while with **21:20-LF_2**, it was $5x10^5\ M^{-1}$. The stoichiometry of binding was found close to 1 in both cases. The reported binding constant of netropsin for its cognate sites are in the range of 10^6 - $10^8\ M^{-1}$ while in present work they are ~ $10^5\ M^{-1}$.

Netropsin has high affinity to its binding sites corresponding to AT seqences with a strech of four bases. However, the oligonucleotides used in present work contain AT stretches extended to only three bases in length which may be the reason for the observed lower binding constants owing to a decreased specificity. Netropsin binding constant values obtained with duplexes having direct attachment of fluorophore were closer to the reported values. Thus, the drug binding event was more efficiently sensed by directly attached fluorophore.

2.4.12 Fluorescence Lifetime Analysis (FLA)

FLA allows a quantitative resolution of discrete excited state lifetimes from the fluorescence decay curve, which reflects probable different modes of molecular interactions of the probe within a target macromolecule. Using FLA as technique for DNA-drug interaction, studies with ethidium bromide, DAPI and metal complexes have led to identification of binding modes of these ligands (Kumar et al., 1985; Barcellona et al., 1990; Heller et al., 1994). Thus, FLA has extensive utility not only to monitor DNA binding ability of the fluorescent ligand but also to examine its mode of interaction with the double helix. The results of fluorescence lifetime studies of ODNs used in present work (single strand and duplex form) are given Table 8.

As seen from Table 8, under present conditions, free carboxyfluorescein itself shows two lifetimes (3.66 ns with larger amplitude & 1.41 ns with shorter amplitude (entry 9)). It is interesting to note that covalently attached fluorescein

with or without spacer showed fluorescence lifetimes not very different to that of free form. In single stranded form, the oligos with direct attachment of label have both lifetime components with equal amplitude (entry 1, 5). In oligos with label attached via spacer the longer lifetime components have a higher amplitude (entry 3, 7). This trend is same with mono as well as dilabelled oligos.

Table 8. Fluorescence lifetimes of labeled ODNs (nanoseconds)*

Entry	ODN	τ_1 (a_1)	τ_2 (a_2)	χ^2
1	**17-F_1**	4.05 (*0.58*)	1.71 (*0.42*)	1.09
2	**21:17-F_1**	3.98 (*0.55*)	1.70 (*0.45*)	0.98
3	**19-LF_1**	4.22 (*0.81*)	**1.73 (*0.19*)**	1.02
4	**21:19-LF_1**	4.22 (*0.81*)	**1.28 (*0.19*)**	1.0
5	**18-F_2**	3.63 (*0.44*)	0.91 (*0.58*)	1.5
6	**21:18-F_2**	3.62 (*0.58*)	1.17 (*0.42*)	1.12
7	**20-LF_2**	4.17 (*0.71*)	1.26 (*0.29*)	1.06
8	**21:20-LF_2**	4.20 (*0.68*)	1.39 (*0.32*)	1.06
9	**CF**	3.66 (0.88)	1.41 (0.12)	1.07

*τ is the excited state lifetime of the fluorophore, a is the weightage of the particular component in the total fluorescence. χ^2 value close to one indicate a good statistical fit of the decay curves. The fluorescence measurements of oligos (300 nm) were done at 520 nm upon exciting the sample at 310 nm. The excitation and emission slitwidth was 4 nm each.

Upon duplex formation, the longer lifetime component remains unchanged irrespective of whether the probe is attached with or without spacer. However, a marked decrease in the shorter lifetime values are observed with **19-LF_1** upon duplex formation without any change in amplitude (entry 3,4). The results are in agreement with time dependent duplex formation study in which 48% decrease in fluorescence intensity of oligo **19-LF_1** was observed upon duplex formation.

In general, it is seen for the case of probe directly attached to DNA that the long and short lifetime components contribute equally both in single strand and duplex forms (entry 1,2,5 and 6) whereas when the probe is attached via spacer the longer lifetime has distinctly higher amplitude component (entry 3,4,7 and 8). This suggests two different probable modes of association of probe with DNA depending on whether the linkage is rigid or flexible. The results were similar even when two fluorophores are attached per DNA molecule.

The intercalative binding mode of fluorophores is associated with large

increase in fluorescence lifetime values. In present study, only a marginal difference in the fluorescence lifetime of free and conjugated fluorophore were observed, ruling out an intercalative mode of binding. The decrease in shorter lifetime component of **19-LF$_1$** upon duplex formation indicates that the change in conformation of fluorophore attached via spacer occurs upon hybridisation. The interpretation of fluorescence lifetime data requires extensive analysis which is not straightforward.

2.4.13 Fluorescent Oligos as Primers in PCR

The polymerase chain reaction is one of the most widely used technique in molecular biology and medicinal genetics (Saiki et al., 1988). The extraordinary power of PCR to amplify the specific nucleic acid sequences of definite length from a complex mixture of genomic DNA is exploited in diagnosis of pathogenic and genetic diseases (Persing et al., 1993). The analysis of PCR products is generally done by electrophoresis in which the products are separated according to their molecular weights. The detection of products on gel is generally done by either ethidium bromide staining or radioactive labeling. Both these methods are quite popular but have certain disadvantages. The ethidium bromide is a known carcinogen which requires careful handling and disposal. Although the radiolabels are highly sensitive, they are hazardous, have short shelf life and waste disposal is serious concern. Thus, the development of methods of detection which are safe, fast and sensitive with ease of waste disposal will greatly facilitate the routine use of PCR for diagnosis. The fluorescent labeling of oligonucleotides provides a convenient methodology which fulfills above criterion. Fluorescent labeling may also avoid the use of electrophoresis for PCR analysis which is of paramount importance on line screening of large amount of samples by PCR (Lo et al., 1990).

In an earlier work (Jadhav et al., 1996), the synthesis and PCR application of oligonucleotides containing nucleobase fluorescent derivative 5-NH_2-(carboxy fluorescein)-2'-deoxyuridine, in which the chromophore fluorescein is directly linked to nucleobase of 2'dU at C-5 via an amido function was

described. PCR was employed to amplify 285 bp region of globin gene of Human genomic target DNA using different combinations of forward and reverse primers (Table 9).

The primer sets **29-F_1:32-F_1**, **30-F_2:33-F_2** and **30-F_2:34-F_3** were chosen to yield product DNA containing two, four and five fluorescent groups respectively per amplified duplex. The gel photograph shown in Figure 16 indicated a successful and efficient amplification of the 285 bp fragment with both labeled (lanes 1-3) and unlabeled (lane 4) sets of primers.

Table 9. Sequences of oligonucleotides primers used in PCR

Forward primer	
28	5'-ACC TCA CCC TGT GGA GCC AC-3'
29-F_1	5'-ACC **X**CA CCC TGT GGA GCC AC-3'
30-F_2	5'-ACC **X**CA CCC **X**GT GGA GCC AC-3'
Reverse primer	
31	5'-CTC CTT AAA CCT GTC TTG TAA CCT TGT TAG-3'
32-F_1	5'-CTC CT**X** AAA CCT GTC TTG TAA CCT TGT TAG-3'
33-F_2	5'-CTC CT**X** AAA CC**X** GTC TTG TAA CCT TGT TAG-3'
34-F_3	-5'-CTC CT**X** AAA CC**X** GTC TTG **X**AA CCT TGT TAG-3'

X = 14 (fluorescent nucleobase).

Under identical PCR conditions, fluorescence intensity of amplified products as seen on agarose gel was found to be enhanced with an increasing number of fluorophores. Optimal and unambiguous detection of amplified product on gel was possible with 250 ng of DNA containing four fluorophores per amplified duplex, obtained from primer set (**30-F_2:33-F_2**). Increasing the number of fluorescein tags beyond this level resulted in a non-specific amplification of PCR product giving additional bands (Figure 16, lane 3).

In the present work, comparative effects on PCR utility of oligonucleotide primers containing 5/6-carboxyfluorescein attached to 5-NH_2-deoxyuridine with and without the spacer chain are studied. The modified bases containing carboxy fluorescein labeled to 5-NH_2-deoxyuridine directly and via long and flexible spacer chain were synthesised. The modified bases were incorporated at the desired positions in oligonucleotides during chemical synthesis of oligonucleotides.

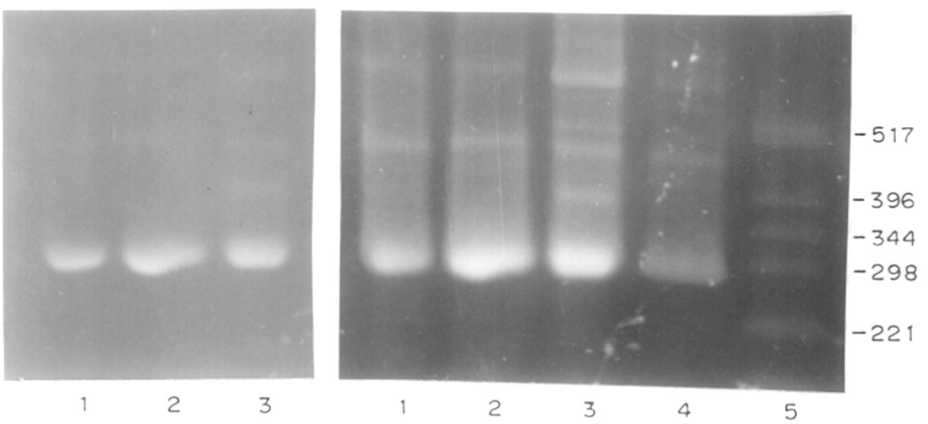


Figure 16. Gel photograph of PCR amplified product before (A) and after (B) ethidium bromide staining. The lanes show the PCR product with following set of primers; lane 1, **29-F_1** and **32-F_1**; lane 2, **30-F_2** and **33-F_2**; lane 3, **30-F_2** and **34-F_3**; lane 4, **28** and **31**; lane 5, Molecular weight marker pBR322 Hinf I digest.

The oligonucleotide sequences (Table 1) for amplifying the 861 bp region of β-globin region of Human genomic DNA were synthesized for use in PCR as primers (Figure 17).

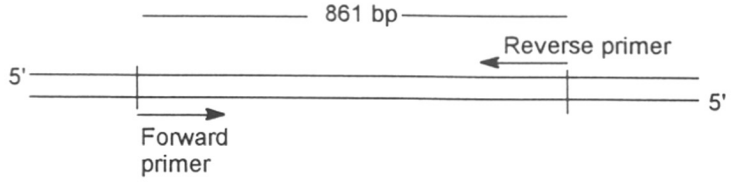


Figure 17. The figure shows schematic representation of β-globin gene fragment of Human genomic DNA. The position of PCR primers is indicated which are used for amplifying 861 bp region. The forward and reverse primer pairs used in this study are shown in Table 10.

Figure 18 shows the successful amplification of 861 bp PCR product using modified as well as unmodified primers. The annealing temperature of the reaction was kept at 62 °C, below the duplex T_m values of primer sequences. Although, there is no apparent complementarity at 3' end of primer pair, at annealing temperature below 55 °C, only primer-dimer formation was observed, which is in agreement with the previously reported results (Watson, 1990; Ferrie et al., 1992). The high concentrations of primers used in PCR may show weak inter/intra primer interactions which can be extended by *Taq* polymerase resulting in primer-dimer formation.

Table 10. The primer pairs used in PCR #

Entry	Forward : Reverse
1	**16 : 22**
2	**17-F_1 : 23-F_1**
3	**18-F_2 : 24-F_2**
4	**19-LF_1 : 25-LF_1**
5	**20-LF_2 : 26-LF_2**

The primer sequences are given in Table 1

The efficiency of incorporation of modified primers with monolabel on terminal base (entry 2,4) was same as that of unmodified primers for ODNs (entry 1). The dilabelled primers with fluorescein attached through spacer (entry 5) also showed satisfactory incorporation. The efficiency of PCR product formation was poor with dilabelled primers (52% of PCR product formation with unlabelled primers, entry 3) carrying fluorescein which is directly attached to the

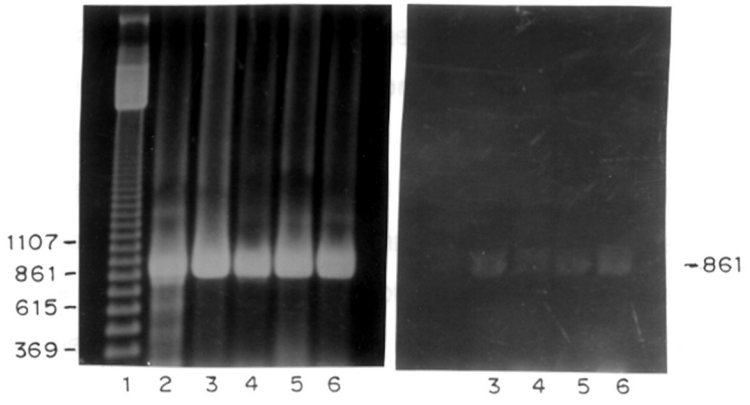


Figure 18. Gel photograph of PCR amplified product after (A) and before (B) ethidium bromide staining. The lanes show the PCR product with following set of primers; lane 1, 123 bp molecular weight marker lane 2, **16** and **22**; lane 3, **17-F_1** and **23-F_1**; lane 4, **18-F_2** and **24-F_2**; lane 5, **19-LF_1** and **25-LF_1**; lane 6, **20-LF_2** and **26-LF_2**.

base. The results indicate that when fluorescein is attached through spacer, the interference with *Taq* polymerase activity is minimal while the direct attachment of fluorophore interferes in the enzyme recognition causing a lower efficiency of PCR amplification. The results are supported with the observation that direct attachment of fluorophore results in destabilisation of duplex stability by 3-4 °C per modification while destabilisation by laelling via spacer is 1-2 °C per modification.

Under identical reaction conditions, the fluorescent intensities of the PCR products differed with degree of substitution and the mode of attachment of fluorophore. The products obtained with dilabelled primers having direct attachment of fluorophore showed only 60% of fluorescence intensity compared to their counterparts carrying fluorophore through spacer. These results are in agreement with the fluorescent properties of oligonucleotides described earlier which show the quenching of fluorescence using direct attachment of probes. Overall results indicate that the primers carrying fluorophores attached through spacer are better candidates for PCR detection as they are efficiently incorporated, show higher fluorescence intensity and thus enhancing the sensitivity of PCR product detection.

2.5 CONCLUSIONS

This study describes the synthesis, characterisation and various applications of fluorescent oligonucleotides. The synthetic strategy for attaching spacer at 5-NH_2 position of deoxyuridine base has been developed. Such monomers with long & flexible spacer having functional -NH_2 group at the end is useful to attach various reporter groups. In the present work, such monomers were used to attach carboxyfluorescein molecule which were incorporated into oligonucleotides. The fluorescence properties of these oligomers were compared with oligonucleotides in which carboxyfluorescein attachment was done directly to base at C-5 position of deoxyuridine through short and rigid amido spacer.

The thermal stability experiments suggested that inclusion of spacer between fluorophore and base does minimal effect on duplex stability (1- 1.5 °C

per modification) while direct attachment of label results in destabilisation by 3- 4 °C per modification. The duplex formation was efficiently monitored with these oligonucleotides by monitoring change in fluorescence intensity upon hybridisation. In such studies, the oligos with direct attachment of fluorophore showed up to 84% increase in fluorescence whereas with oligos having fluorophore attached through spacer, it was only 11%. The steady state anisotropy values were found to be higher with oligos having direct attachment of label suggesting its higher sensitivity in monitoring DNA hybridisation.

These oligos were also used to monitor netropsin binding to DNA. A drop in fluorescence intensity was observed with addition of netropsin. The observed binding constants of netropsin with DNA were found to be close to reported values using oligos having direct attachment of label and hence making them better candidates to study drug interaction.

However, the relative fluorescence intensities of oligonucleotides having fluorophore attached through long, flexible spacer were much higher than the ones in which fluorophores are directly to the base. The quenching of fluorescence became higher upon double labelling with directly attached fluorophore. In PCR study, the oligos with fluorophore attached through spacer were more efficiently utilised by *Taq* polymerase and showed higher sensitivity of detection compared to oligos with direct attachment of fluorophore.

The overall results suggest that, the selection of attachment of fluorophore with or without spacer will depend on the application for which they are required. The oligos with direct labelling of fluorophores to base are suitable for studying DNA hybridisation, drug-DNA interaction etc. as such changes are sensed efficiently them. The direct and rigid attachment of label places fluorophore in major groove of DNA and hence it is more suitable to monitor changes occurring in the major groove. On the other hand, the fluorophore labelling through spacer reduces fluorescence quenching and such property is important in applications like PCR where higher sensitivity of detection is desired.

2.6 REFERENCES

Agarwal, S., Christoulou, C. & Gait, M. J. (1986) *Nucleic Acids Res.* **14**, 6227.

Aguirre, R., Gonsoulin, F. & Cheung, F. C. (1986) *Biochemistry* **25**, 6827.

Allen, D. J., Darke, P. I. & Benkovic, S. J. (1989) *Biochemistry* **28**, 4601.

Ansorage, B., Sproat, B. S., Stegman, J., Schwager, G. & Zenke, M. (1987) *Nucleic Acids Res.* **15**, 4593.

Asseline, U., Toulme, F., Thoung, N. T., Delarue, M., Therese, M. G. & Helene, C. *EMBO. J.* **3**, 795 (1984).

Barawkar, D. A. & Ganesh, K. N. (1993) *BioMed. Chem. Lett.* **3**, 347.

Barawkar, D. A. & Ganesh, K. N. (1995) *Nucleic Acids Res.* **23**, 159.

Barawkar, D. A., & Ganesh, K. N. (1994) *Biochem. Biophys. Res. Comm.* **203**, 53.

Barcellona, M. L., Cardiel, G. & Gratton, E. (1990) *Nucleic Acids Res.* **20**, 27.

Beaucage, S. L. & Iyer, R. P. (1993) *Tetrahedron* **49**, 1925.

Berman, H. M., Neidle, S., Zimmer, C. & Thrum, H. (1979) *Biochem. Biophys. Acta* **561**, 124.

Brumbaugh, J. A., Middendorf, L. R., Grone, D. L. & Ruth, J. L. (1988) *Proc. Natl. Acad. Sci. USA.* **85**, 5610.

Chaires, J. B., Priebe, W., Graves, D. E. & Burke, T. G. (1993) *J. Am. Chem. Soc.* **115**. 5360.

Chehab, F. F. & Kan, Y. W. (1989) *Proc. Natl. Acad. Sci. USA.* **86**, 9178.

Chen, S. C., Nakamura, H. & Tamura, Z. (1979) *Chem. Pharm. Bull.* **27**, 475.

Chollet, A. & Kawashima, E. H. (1985) *Nucleic Acids Res.* **13**, 1529.

Connolly, B. A. (1987) *Nucleic Acids Res.* **15**, 3131.

Conway, N. E. & McLaughlin, L. W. (1991) *Bioconjugate Chem.* **2**, 452.

Cook, A. F., Vuocolo, E. & Brakel, C. L. (1988) *Nucleic Acids Res.* **16**, 4077.

Crooke, S. T. & Bennet, C. F. (1996) *Ann. Rev. Pharmacol. Toxicol.* **36**, 107.

Draper, D. E. (1984) *Nucleic Acids Res.* **12**, 989.

Effcavitch, J. M. (1990) in *Gel Electrophoresis of Nucleic Acids - A Practical Approach;* Rickwood, D. Hames, B. D. Ed. Oxford University Press; Oxford.

Ferrie et al., (1992) *Am. J. Hum. Genet.* **51**, 251.

Fick, J., Lawakzeck, R. & Schneider, F. W. (1982) *Eur. J. Biochem.* **126**, 367.

Gilham, I. C. (1987) *Trends in Biotech.* **5**, 332.

Goodchild, J. (1989) *Bioconjugate Chem.* **1**, 165.

Guest, C. R., Hochsrasser, R. A., Dupuy, C. G., Allen, D. J., Benkovic, S. J. & Millar, D. P. (1991) *Biochemistry*, **30**, 8759.

Guisti, W. G. and Adriano, T. (1993) *PCR Methods and Applications* **2**, 223.

Hagmar, P., Bailey, M., Tong, G., Haralambidis, J., Sawyer, W. and Davidson, B. E. (1995) *Biochim. Biopys. Acta* **1244**, 259.

Haralambidis, J., Angus, K., Pownal, S., Duncan, L., Chai, M. & Tregear, G. W. (1990) *Nucleic Acids Res.* **18**, 501.

Haralambidis, J., Chai, M. & Tregear, G. W. (1987) *Nucleic Acids Res.* **15**, 4857.

Heller, D. O. & Greenstock, C. L. (1994) *Biophys. Chem.* **50**, 305.

Jablonski, E., Moomaw, E. W., Tullis, R. H. & Ruth, J. L. (1986) *Nucleic Acids Res.* **14**, 6115.

Jadhav, V. R., Barawkar, D. A., Natu, A. A. & Ganesh, K. N. (1997) *Nucleosides & Nucleotides* **16**, 107.

John, S. W. M., Weitzner, G., Rozen, R. & Scriver, C. R. (1991) *Nucleic Acids Res.* **19**, 408.

Keller, G. H. & Manak, M. M. (1989) in *DNA Probes:* Stockholm Press: NY.

Kempe, T., Wesley, I., Sundgeist, C. F. & Hu, S. I. (1985) *Nucleic Acids Res.* **13**, 45.

Kingsbury, D. T. (1987) *Trends in Biotech.* **5**, 107.

Kopka, M. L., Yoon, C., Godsell, D., Pjura, P. & Dickerson, R. E. (1985) *Proc. Natl. Acad. Sci.* USA **82**, 1376.

Kricca, L. J. (1992) in Non-isotopic DNA Probe Techniques: Academic Press Inc. San Diego.

Kumar, C. V., Barton, J. K. & Turro, N. J. (1985) *J. Am. Chem. Soc.* **107**, 5518.

Kumke, M. U., G, Li., McGown, L. B., Walker, G. T. & Linn, C. P. (1995) *Anal. Chem.* **67**, 3945.

Kumke, M. U., Shu, L., McGown, L. B., Walker, G. T., Pitner, J. B. & Linn, C. P. (1997) *Anal. Chem.* **69**, 500.

Lakowicz, J. R., (1983) In "*Principles of Fluorescence Spectroscopy*" Plenum press, New York

Landegren, U., Kaiser, R., Caskey, C. T. & Hood, L. R. (1988) *Science* **242**, 229.

Lo, Y., Mehal, W. Z. and Fleming, K. A. (1990) *Lancet* **i**:679.

Manoharan, M., Guinosso, C. J. & Cook, P. D. (1991) *Tetrahedron Lett.* **32**, 7171.

Markiewicz, W. T., Groger, G., Rosch, R., Zebrowska, A., Markiewicz, M., Klotz, M., Hinz, M., Godzina, P. & Seliger, H. (1997) *Nucleic Acids Res.* **25**, 3672.

Morawetz, H. (1988) *Science* **240**, 172.

Munkholm, C., Parkinson, Don-R. & Walt, D. R. (1990) *.J. Am. Che.Soc.***112**, 2608.

Murakami, A., Nakamura, M., Nakatsuji, Y., Nagahara, S., Tran-Cong, Q. & Makino, K. (1991) *Nucleic Acids Res.* **19**, 4097.

Netzel, T. L., Zaho, M., Nafisi, K., Headrick, J., Sigman, M. S. & Eaton, B. E. (1995) *J. Am. Chem. Soc.* **117**, 9119.

Patel, L., Abate, C. & Curran, T. (1990) *Nature (London)* **347**, 572.

Perrin D. D. and Armarego, W. L. F. (1989) *Purification of Laboratory Chemicals* third edition Pergamon Press.

Persing, D. H., Smith, T. F., Tenover, F. C. & White, T. J. (1993) In "*Diagnostic Molecular Microbiology-Principles and Applications*", American Society for Microbiology, Washington

Prober, J. M., Trainor, G. L., Dam, R. J., Hobbs, F. W., Robertson, C. W., Zagursky, R. J., Cocuzza, A. J., Jensen, M. A. & Baumeisten, K. (1987) *Science* **238**, 336.

Rihn, B., Bottin, M. C., Coulais, C. & Martinet, N. *J.* (1995) *Biochem. Biophys. Methods* **30**, 91.

Ruth, J. L. (1984) *DNA*, **3**, 123.

Saiki et al., (1988) *Science* **239**, 487.

Saiki, R. K., Scharf, S. Faloona, F., Mullis, K., Horn, G. T., Erlich, H. A. & Arnheim, N. (1985) *Science*, **230**, 1350.

Singh, D., Kumar, V. A. & Ganesh, K. N. (1990) *Nucleic Acids Res.* **18**, 3339.

Sinha, N. D., Biernat, J., McManus, J. & Koster, H. (1984) *Nucleic Acids Res.* **12**, 4539.

Smith, L. M., Fung, S., Hunkapiller, M. W. & Hood, L. E. (1985) *Nucleic Acids Res.* **13**, 2399.

Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connel, C. R., Heiner, C., Kent, B. H. & Hood, E. (1986) *Nature (London)* **321**, 674.

Smith, M. (1994) *Angew. Chem. Int. Ed. Engl.* **33**, 1214.

Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503.

Tang, J. Y. & Agarwal, S. (1990) *Nucleic Acids Res.* **18**, 6461.

Telser, J., Cruikshank, K. A., Morrison, L. E. & Netzel, T. L. (1989) *J. Am. Chem. Soc.* **111**, 6966.

Thiesen, P. McCollum, C. & Andrus, A. (1992) *Nucleic Acids Symposium Series,* **19**, 99.

Tong, G., Lawlor, J. M., Tregear, G. W. & Haralambids, J. (1993) *J. Org. Chem.* **58**, 2223.

Urdea, M. S., Warner, B. D., Running, J. A., Stempien, M., Clyne, J. & Horn, T. (1988) *Nucleic Acids Res.* **16**, 4937.

Varawalla, N. Y., Old, J. M. Sarkar, R., Venkatesan, R. & Weatherall, D. J. (1991) *Brit. J. Haematology* **78**, 242.

Viscidi, R. P. & Volken, R. G. (1987) *Molecular and Cellular Probes* **1**, 3.

Visser, D. W. (1968) *"Synthetic Procedures in Nucleic Acid Chemistry".* John-Wiley, NY, **1**, 401

Wachter, L., Jablonski, J. A. & Ramchandran, K. L. (1986) *Nucleic Acids Res.* **14**, 7985.

Wallner, G., Amann, R. & Beisker, W. (1993) *Cytometry* **14**, 136.

Watson, J. D. & Crick, F. H. C. (1953) *Nature (London)* **171**, 737.

Watson, R. (1989) *Amplifications*, 5.

CHAPTER 3

Design of a combinatorial oligonucleotide library containing all possible hexamer palindromes: PCR synthesis and application for identifying restriction cleavage sites

3.1 INTRODUCTION

Ever since the finding of restriction enzymes in the late sixties (Lin and Arber, 1968; Messelson and Yuan, 1968; Smith and Wilcox, 1970), the introduction of gel electrophoresis for analysis of restriction cleavage fragments (Danna and Nathans, 1971) and the development of DNA sequencing procedures (Maxam and Gilbert, 1977; Sanger et al., 1977), the search for newer restriction enzymes has continued unabated due to their overwhelming importance in recombinant DNA technology. The restriction endonucleases are the enzymes produced by microorganisms which cleave DNA at specific site. These are the part of host modification and restriction system in which sequences of foreign DNA are cut by restriction enzymes while the host DNA is protected by modification using methylases. In any gene analysis, the specific fragment of DNA carrying the particular gene needs to be separated from the rest of genomic DNA. Restriction enzymes are widely used in such studies by cutting the genomic DNA in manageble size which can be manipulated for further analysis. A most recent compilation of known restriction enzymes has 2821 entries and the high frequency of finding new enzymes has necessitated the database to be updated almost daily (Roberts and Macelis, 1997). Based on the corresponding gene and enzyme structure, cofactor dependence and specificity of binding as well as cleavage, the restriction enzymes are divided into three classes - I, II and III (See Box 1). Among these, the subtype - IIP endonucleases are the most common class-II enzymes, that recognize tetra-, hexa- and octanucleotide palindrome sequences (Kesseler and Manta, 1990).

Class II restriction endonucleases are powerful tools for the dissection of DNA and they are essential for the construction of recombinant DNA molecules and for DNA sequence determination (Roberts 1976). While all the known class II restriction enzymes cleave both the strands of DNA, the particular phosphodiester bonds hydrolysed vary from enzyme to enzyme. Thus, in some cases DNA fragments are produced with flush ends and others produce termini

with either 5' or 3' extensions. Clearly, characterisation of the cleavage site remains an important aspect of restriction enzyme analysis.

BOX I

Nomenclature and Classification of Restriction Enzymes (Kesseler and Manta, 1990)

Nomenclature: All restriction enzymes are named by *at least* a capital and two small letters in italics, followed in each case by a Roman numeral. The letters reflect the initials of bacterial genus and species classification. The Roman numeral refers to the number of enzymes discovered in the respective microorganisms. If additional information is needed for discrimination (e.g., extended strain specification of related species), these abbreviations are added (capital letter or arabic numerals) between the initial three letters in italics and final Roman numeral (e.g., *Asp*HI from *Achromobacter* species is an isoschizomer of *Hgi*AI; *Ecl*XI from *Enterobacter clocae* is an isoschizomer of *Xma*III).

Classification: The restriction and modification enzymes are classified in three types based on their gene and protein structure, cofactor dependence, as well as specificity of binding and cleavage reactions.

Class I enzymes exhibit both restriction as well as DNA modification activities located on different subunits of multifunctional enzyme complex. They require Mg^{2+} ions, ATP, and AdoMet as cofactors. These enzymes cleave DNA at unspecific sites, usually 100 to 1000 bp downstream of their recognition sequences.

Class II enzymes act as separate proteins. Class II restriction enzymes carry out site specific hydrolytic endonuclease activity in presence of Mg^{2+} ions on particular phosphodister bonds within both strands of dsDNA. These enzymes are further classified by the ds recognition sequence, the cut position within both strands of DNA, and their methylation sensitivity. The subclass IIP restriction enzymes are the most common ones which recognise tetra-, hexa- and octanucleotide perfect palindrome sequences.

Class III enzymes like those of class I, also combine restriction and modification activities on single enzyme complex composed of different subunits. However, these enzymes are lacking both the ATPase activity of class I and absolute requirement for AdoMet. Class III enzymes recognise specific sequences, and in the presence of Mg^{2+} ions, cleave DNA at a rather fixed position, 25 to 27 bp distal of their recognition sequence.

Different strategies have been developed for characterisation of the cleavage sites which are in general based on ^{32}P end labeling of mixture of fragments produced by a given enzyme, followed by the determination of nucleotide sequence at the labeled ends (Brown and Smith, 1980). There are generally three types of methods as given below.

(i) The earliest method of characterization of restriction enzyme sites used for *Hind*II involved an analysis of fragments released after DNAse I digestion of the 3'- or 5'- labeled termini of the restriction enzyme cleaved fragments. The products are compared with standard di- or trinucleotide fragments to deduce the recognition site (Kelly and Smith, 1970).

(ii) The method of primed synthesis on restriction cleavage product coupled with DNA sequencing is extensively used to analyse the cleavage site of restriction enzymes (Brown et al., 1980). However, the unambiguous identification of specific sequence for recognition-cleavage by new isolates of restriction enzyme is not always straightforward.

(iii) Another simple method for identifying palindromic sequences recognised by restriction endonucleases has been described by Fuchs et al., (1978). The method involves computer search based tabulation of frequencies, distances between position of all possible tetra-, penta- and hexanucleotide palindromes occurring in published sequences of various viral DNAs. This information is utilised for identifying the sequence recognised by restriction endonuclease by comparing the experimental determination of the number and appropriate sizes of the fragments obtained by digestion of ϕx174 RF or SV40 DNAs with particular enzyme. The palindroimic sequence which gives the same pattern of location on particular DNA obtained using computer search, as that obtained from enzyme digest is the recognition sequence of enzyme. Although the method is simple, it determines only the recognition sequence and not the actual cutting/cleavage site within the sequence. Moreover, using agarose or polyacrylamide gel electrophoresis, the number of bands observed can be less than the number of DNA fragments produced in a cleavage reaction due to similar or very small size of fragments. Hence, the number and sizes of the restriction fragments and the locations of the cleavage sites need to be determined with accuracy.

Combinatorial approaches are being presently used increasingly in drug discovery to maximize the possibility of success while minimizing the time required to reach lead compound (Trotta et al., 1995). Since oligonucleotides of defined sequences can be synthesized with ease and methods are readily available for their identification, these approaches hold a great potential for newer applications when coupled with combinatorial methods (Bock et al., 1992; Chetverin et al., 1994). One of the most important aspects in combinatorial

library of oligonucleotides is the separation of oligonucleotide - ligand complex from the large pool of unbound oligonucleotides. For some applications, it may not be necessary to use the combinatorial library of all possible sequences of oligonucleotides of specific length. In such cases, the logical reduction of the size of library will greatly facilitate the separation of DNA bound ligand from the unbound ones.

3.2 PRESENT WORK

This chapter describes, a combinatorial strategy for design and synthesis of an oligonucleotide library which contains all possible hexamer palindromes and demonstrate its utility in characterisation of restriction enzyme cleavage sites (*RE*-oligos: Restriction enzyme site recognising oligonucleotides). The library consisting of a set of 8 oligonucleotides is derived through a linear overlap of rationally arranged hexamers to minimize the final size and accommodates all possible hexamer palindromes at uniquely defined sites. Further, it also contains all possible tetramer palindromes, located at least at four sites within a library. The set of oligonucleotides described here, when screened against new restriction endonucleases, has the potential for a direct and unambiguous determination of the tetra/hexamer palindromic sequences specifically cleaved by the enzyme.

3.3 MATERIALS AND METHODS

All reagents and chemicals used were of reagent quality or better grade. The base protected nucleoside phosphoramidites and nucleoside derivatised controlled pore glass supports (CPG) were purchased from Cruachem UK. Ampli *Taq* DNA polymerase used for PCR reactions was from Perkin Elmer, T4 polynucleotide kinase from USB Corporation and the restriction enzymes from Promega. The buffers used for PCR and restriction digestion were same as that obtained from enzyme source. [γ-^{32}P] dATP was obtained from Bhabha Atomic Research Centre, Bombay.

3.3.1 Oligonucleotide Synthesis, Purification and Labeling.

DNA oligonucleotides were synthesized on 0.2 μM scale using the solid phase phosphoramidite chemistry on a Pharmacia LKB - Gene Assembler Plus followed by deprotection with aqueous ammonia (55 °C, 17 hours). The oligonucleotides were desalted by gel filtration using Pharmacia NAP-10 columns and further purified by preparative electrophoresis on 20% denaturing polyacrylamide gel using "crush and soak" method as described in chapter 2 section 2.3.5 and their purity was rechecked by HPLC. The concentration of oligonucleotides was calculated by considering 1 O.D. at 260 nm corresponds to 30 μg of DNA. The oligonucleotides were labeled radioactively at 5'-end using [γ-^{32}P]-dATP and T4 polynucleotide kinase according to standard procedures (Sambrook et al., 1989).

3.3.2 PCR Assembly of *RE*-Oligonucleotides

All PCR experiments were performed on a Perkin Elmer DNA Thermal Cycler over 30 cycles employing the following temperature cycle: (a) denaturation, 92 °C for 20 sec (b) annealing, 15 °C for 20 sec and (c) chain extension, 72 °C for 20 sec. A typical PCR mixture contained specific oligonucleotide (500 ng), each dNTP (200 μM), $MgCl_2$ (5 mM) and Ampli *Taq* polymerase (2.5 Units) in 1x polymerase buffer (50 μl). The unlabeled PCR products were analyzed on 2% agarose gel, stained with ethidium bromide and visualized under UV light. The labeled PCR products were analyzed on 15% native as well as denaturing PAGE containing 7 M urea followed by autoradiography. The PCR products were purified by phenol:chloroform extraction followed by ethanol precipitation in presence of 0.15 M NaCl. The pellets were washed with 70% ethanol, dried and redissolved in deionised water.

3.3.3 Restriction Enzyme Digestion of PCR Products and Analysis

A typical reaction mixture for restriction enzyme cleavage analysis contained 2 μl of the substrate oligonucleotide (500 cpm/μl), sonicated calf thymus DNA (0.1 mg/ml) and restriction enzyme (1.0 Unit) in appropriate buffer supplied by manufacturer. The reactions were done for 30 min at 37 °C. The A +

G specific Maxam-Gilbert (1977) chemical sequencing reactions of the oligonucleotides were performed using piperidine formate as base modifying agent as described (Sambrook et al., 1989). All reaction products were analyzed by gel electrophoresis on 8% denaturing sequencing PAGE containing 7 M urea at 1200V till the bromophenol blue reached 3/4th of gel. The gels were exposed to X-ray film for 24 hours and later developed.

3.4 RESULTS AND DISCUSSION

3.4.1 Design of Minimal Length Sequences for All Possible Hexamer Palindromes

In any hexamer palindrome, the first three base sequence determines the next three base sequence due to invert complementarity and two fold rotational symmetry. A total of 64 base triplets arise from all possible combinations of the four DNA bases, resulting in an equal number of hexamer palindrome sequences. If all these hexamer palindromes are arranged tandemly as a single fragment, the length of such a DNA would be 64 x 6 = 384 bases. This section describes design of a set of 8 oligonucleotides (total of 266 bases) that encompass all the possible hexamer palindromes through a linear overlap method in which the last two bases of one hexamer palindrome becomes the first two bases of next hexamer palindrome. For e.g., starting from a single hexamer palindrome 1 having first two bases as AA and the last two bases as TT, the next palindrome 2 will have the first two bases as TT and the last two bases necessarily as AA. Thus, the first

$$\overset{1}{\overline{\text{AAAT}TT\text{ATAA}}}\underset{2}{} \qquad \overset{1}{\overline{\text{AAAT}TT}}\text{AT}\overset{3}{\overline{AATA}}\text{TT}\underset{2}{}$$

two bases of palindrome 1 is regenerated as the last two bases of palindrome 2. The extension of palindrome 2 on this principle would only repeat this palindrome duet. To avoid this in further sequence build up, a combinatorial variation of base at third position was done. Thus addition of T as the third base at the end of last AA doublet of 2, instead of the existing A as third base in

palindrome 1, gives rise to the third hexamer palindrome 3. Similarly, the terminal TT doublet of palindrome 3 can be extended by the addition of a third base different from that in TT palindrome 2 and complete the next hexamer palindrome. The doublet at the end of each last hexamer can be extended by a systematic addition of all four bases at third position one at a time to obtain an oligonucleotide which finally comprises of all feasible hexamer palindromes starting with two doublets AA and TT.

1. AA***A***TTT
2. TT***A***TAA
3. AA***T***ATT
4. TT***T***AAA
5. AA***G***CTT
6. TT***G***CAA
7. AA***C***GTT
8. TT***C***GAA

AA***A***TTT***A***TAA***T***ATT***T***AAA***G***CTT***G***CAA***C***GTT***C***GAA

This strategy leads to a polynucleotide of 34 bases containing eight hexamer palindromes, each starting and terminating with either AA or TT and reducing the oligomer length by 14 bases (expected 8 x 6 = 48). The individual members of the oligonucleotide library were assembled in a similar way starting from other doublets.

There are a total of 16 possible doublets, which can be classified in two groups: group I containing the 12 non-self complementary doublets and group II consisting of 4 self-complementary doublets. The group I doublets can be further divided into to 6 complementary pairs: (i) AA and TT (ii) CA and TG, (iii) CT and AG, (iv) GA and TC, (v) GG and CC and (vi) GT and AC, which must occur as either the first or last two bases in a hexamer palindrome. Using the iterative approach described above, 6 polynucleotide sequences of 34 base length can be designed for the 6 group I doublets pairs, each sequence representing all the possible hexamer palindromes starting with one of the group I doublet pairs. In contrast, due to the self-complementary nature of group II doublets, the present method leads to the formation of only 18 mer oligonucleotides for each one of

the doublets. The oligonucleotides representing group II palindromes from AT (4) and TA (5), have identical pentamer sequence at one of their ends (***TATAT***) and this feature was utilized to overlap-link them into a single sequence of 31 bases containing all hexamer palindromes starting with AT as well as TA.

4 ATTAATGCATCGA***TATAT***
5 ***TATAT***AATTAGCTACGTA

ATTAATGCATCGA***TATAT***AATTAGCTACGTA

Similarly, a construct of oligonucleotide of 31 bases representing all palindromes starting with GC and CG can be deduced.

6 GCATGCTAGCCGG***CGCGC***
7 ***CGCGC***GATCGTACGGCCG

GCATGCTAGCCGG***CGCGC***GATCGTACGGCC

Using the above described logic, a set of 8 oligonucleotide sequences (**1-8, Table 1**) representing all possible hexamer palindromes at uniquely defined sites were designed.

Table 1. Oligonucleotide sequences **1 - 8**

Oligomers*		1	4	7	10	13	16	19	22	25	28	31	34	37
1 *RE*AA-TT	d	GCA	AAT	TTA	TAA	TAT	TTA	AAG	CTT	GCA	ACG	TTC	GAA	C
2 *RE*CA-TG		GCC	AAT	TGA	TCA	TAT	GTA	CAG	CTG	GCC	ACG	TGC	GCA	
3 *RE*CT-AG		GCC	TAT	AGA	TCT	TAA	GTA	CTG	CAG	GCC	TCG	AGC	GCT	
4 *RE*TC-GA		GCT	CAT	GAA	TTC	TAG	ATA	TCG	CGA	GCT	CCG	GAC	GTC	
5 *RE*CC-GG		GCC	CAT	GGA	TCC	TAG	GTA	CCG	CGG	GCC	CCG	GGC	GCC	
6 *RE*GT-AC		GCG	TAT	ACA	TGT	TAA	CTA	GTG	CAC	GCG	TCG	ACC	GGT	
7 *RE*AT-TA		GCA	TTA	ATG	CAT	CGA	TAT	ATA	ATT	AGC	TAC	GTA		
8 *RE*GC-CG		GCG	CAT	GCT	AGC	CGG	CGC	GCG	ATC	GTA	CGG	CCG		

**RE*NN-NN (N=A,T,G,C) represents Restriction Enzyme recognizing oligonucleotides comprising all hexamer palindromes starting with a pair of group I or II NN doublet. All sequences have an additional GC clamp at 5'-end.

This approach not only decreases the library size but also can be used for creating different sequences containing the same information as far as tetra-hexamer palindromes are concerned, by changing the sequence in which third base is added in sequence build up. For e.g.,

AA***A***TTT***A***TAA***T***ATT***T***AAA***G***CTT***G***CAA***C***GTT***C***GAA

can be designed in following way by changing the sequence in which third base is added during sequence build up.

AA***T***ATT***T***AAA***G***CTT***G***CAA***C***GTT***C***GAA***A***TTT***A***TAA

Both these oligos have the hexamer palindromes starting with AA and TT with only difference being the positions at which they are located.

While this work was in progress, a related report has recently appeared in literature (Hecker and Rill, 1997; Rill and Hecker, 1996) describing the design of polynucleotide templates representing all base quartets in a minimal length sequence (144 nt) that was cloned into plasmids. The minimal length of oligonucleotides were designed by maximum overlap approach between two quartets.Though no specific application was demonstrated, it has potential for characterization of sequence-selective binding of drugs and other small molecules to nucleic acids. However, the design strategy employed in this work for building hextet palindromes is based on a logic different from that used for quartets.

3.4.2 Synthesis and Purification of Oligonucleotides

The oligonucleotides **1-8** were synthesised by standard phosphoramidite chemistry on solid support using an automated DNA synthesiser. Following deprotection and cleavage from support, the oligonucleotides were purified by denaturing gel electrophoresis. Figure 1 shows the purity of oligonucleotides (Table 1) by appearance of single band as observed by denaturing gel electrophoresis. Slight differences were seen in the mobility of oligonucleotides due to different sequence composition. Except *RE*GT-AC (36-mer; lane 1) which showed more retardation, all other 36-mers showed simillar mobility. As expected, the 33-mers *RE*GC-CG (Lane 2) and *RE*AT-TA (Lane 3) showed faster moving bands compared to higher molecular weight oligomers.

3.4.3 Double Stranded Dimer Formation Through PCR

An immediate practical application of these set of oligonucleotides would be to determine the nucleotide sequence at cleavage sites of various type II endonucleases recognizing hexamer palindromes. All restriction enzymes

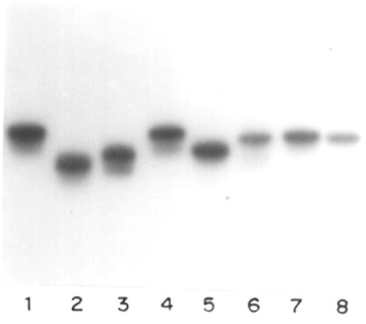


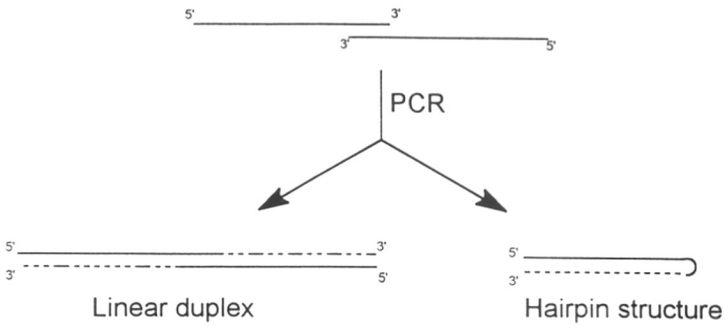
Figure 1: Autoradiograph of 15% denaturing PAGE showing the purity of *RE*-oligonucleotides. lane 1: *RE*GT-AC (**6**), lane 2: *RE*GC-CG (**8**), lane 3: *RE*AT-TA (**7**), lane 4: *RE*AA-TT (**1**), lane 5: *RE*CC-GG (**5**), lane 6: *RE*CT-AG (**3**), lane 7: *RE*TC-GA (**4**), lane 8: *RE*CA-TG (**2**). Inside parenthesis are oligonucleotide number as given in Table 1.

require DNA substrate in double stranded form for their reaction, which was achieved using PCR strategy. The GC clamp was added to the 5' end of each oligo so that the first hexamer palindrome can also be used for cleavage by appropriate restriction enzyme. The double stranded DNA needed for this purpose was prepared by PCR, taking advantage of the 6 bp complementarity at the 3'-end of these oligonucleotides. The sequences of PCR products are given in Table 2. This is identical to the normal primer-dimer formation (Saiki, 1989) observed in PCR where partial duplex formed between the primers due to the 3'-end complementarity, is extended by DNA polymerase. In the present study, the partial duplex (a) is formed between two molecules of same primer due to 3' end complementarity. Although, these oligos have the potential to form another partial duplex (b) by using 5' end complementarity, such duplexes are not extended by DNA polymerase. To increase the stability of (a) form over that of (b) in all oligonucleotides, the third and fourth base of 3' end palindromes were always C or G while A or T were selected as third and fourth bases of 5' end palindrome. Moreover, the addition of GC clamp at 5' end interupts the 5' end complementarity. Such kind of sequence arrangements are expected to increase the extent of form (a) which acts as substrate for PCR.

(a) (b)

Extension by PCR No extension by PCR

The partial duplex formed using 3' end complementarity is extended by DNA polymerase resulting in to a PCR product which itself is a self-complementary sequence containing two inverted repeats. Due to its palindromic nature, the PCR product can exist in two structures: the intermolecular association resulting in a linear double strand and intramolecular association leading to hairpin structure.



Although primer-dimer formation is well known as an artifact in PCR (Saiki, 1989; Rychlik, 1995), the objective in present study was to obtain a complete dimerisation of the starting oligonucleotides. The formation of partial duplex is mandatory for extension by DNA polymerase and in self complementary primers, the proportion of such duplex formed during annealing step is low and hence the efficiency of primer-dimerization per PCR cycle is only modest. Moreover, it has been reported that for primer-dimer formation, inter-primer extensions are better substrates than intra-pimer interactions (Brownie et al., 1997). To increase the partial-duplex form which is the substrate for extension by *Taq* polymerase in PCR, the $MgCl_2$ concentration and annealing temperature were varied. The efficiency of reaction was found to be maximum at a $MgCl_2$ concentration of 5 mM and a low annealing temperature of 15 °C. Figure 2 shows the agarose gel electrophoretic analysis of *RE*GA-TC PCR product after different PCR cycles. There was gradual increase in PCR product formation and to achieve complete dimerisation and obtain maximal yields of all primer-dimers at least 30 cycles of PCR were needed. The single stranded oligonucleotide at the start of reaction was not visible by ethidium bromide staining while the PCR products showed intense staining with ethidium bromide indicating the duplex nature of products. Under these conditions, in case of *RE*AT-TA **1**, the results were poor due to a weaker binding of its AT rich complementary sequence at 3'end. This was improved by the addition of a single base C at 3'-end of *RE*AA-

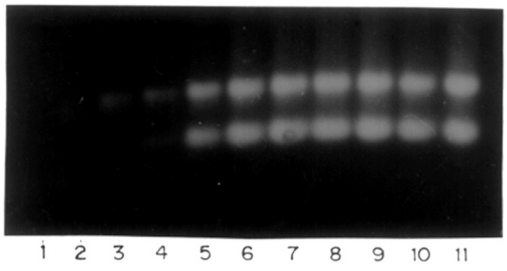


Figure 2: Cycle dependent PCR product analysis of *RE-GATC* on 2% agarose gel. Lane 1 and 2 show *RE*GA-TC PCR mixture in absence and presence of *Ampli Taq* polymerase respectively (as they contain single stranded oligo, they are not stained by ethidium bromide). The other lanes show the *RE*GA-TC PCR product after increasing number of cycles. lane 3: 5 ; lane 4: 10; lane 5: 15; lane 6: 20; lane 7: 25; lane 8: 30; lane 9: 35; lane 10: 40; lane 11: 45 cycles. In each lane the faster moving band is hairpin duplex while the slower moving one is linear duplex PCR product.

TT to generate a 8 bp complementarity that was better in partial duplex formation and gave a higher yield of the dimer under above PCR conditions.

```
5'----------GTTCGAA                5'----------GTTCGAAC
           AAGCTTG----------5'                 CAAGCTTG----------5'
```

Poor PCR substrate Better PCR Substrate

Under identical conditions of PCR as described above, all *RE*-oligos were subjected to PCR for 30 thermal cycles (each thermal cycle consisted of 92 °C denaturation, 15 °C for annealing and 72 °C for extension: each step for 20 sec). All the single stranded starting *RE*-oligonucleotides showed two bands on 2% agarose gel (Figure 3), one faster moving band slightly ahead of the expected 32 bp duplex, suggesting it to be the hairpin duplex while the slower moving band around 60-65 bp mobility arising from linear duplex. Both these bands were purified from low melting agarose gel, reloaded separately on agarose gel to find that both reverted to the same equilibrium composition of linear duplex and hairpin structure as in the original PCR product. The formation of these products confirmed that no other hybridisation in oligonucleotides occurs which can give PCR product.

Table 2. Oligonucleotide sequence of PCR products **1a-8a**[#]

		1	11	21	31	41	51	61
1a	*d*	GCAAATTTAT	AATATTTAAA	GCTTGCAACG	TTCGAACGTT	GCAAGCTTTA	AATATTATAA	ATTTGC
2a		GCCAATTGAT	CATATGTACA	GCTGGCCACG	TGCGCACGTG	GCCAGCTGTA	CATATGATCA	ATTGGC
3a		GCCTATAGAT	CTTAAGTACT	GCAGGCCTCG	AGCGCTCGAG	GCCTGCAGTA	CTTAAGATCT	ATAGGC
4a		GCTCATGAAT	TCTAGATATC	GCGAGCTCCG	GACGTCCGGA	GCTCGCGATA	TCTAGAATTC	ATGAGC
5a		GCCCATGGAT	CCTAGGTACC	GCGGGCCCCG	GGCGCCCGGG	GCCCGCGGTA	CCTAGGTACC	ATGGGC
6a		GCGTATACAT	GTTAACTAGT	GCACGCGTCG	ACCGGTCGAC	GCGTGCACTA	GTTAACATGT	ATACGC
7a		GCATTAATGC	ATCGATATAT	AATTAGCTAC	GTAGCTAATT	ATATATCGAT	GCATTAATGC	
8a		GCGCATGCTA	GCCGGCGCGC	GATCGTACGG	CCGTACGATC	GCGCGCCGGC	TAGCATGCGC	

[#] The underlined bases indicate presence of GNA triplet in loop region of potential hairpin structure formed by PCR products.

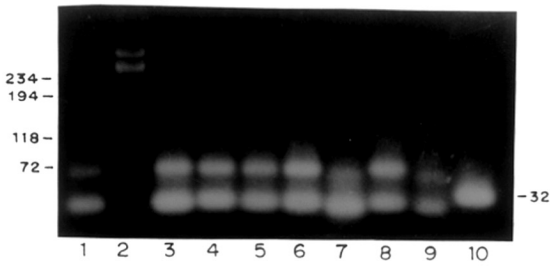


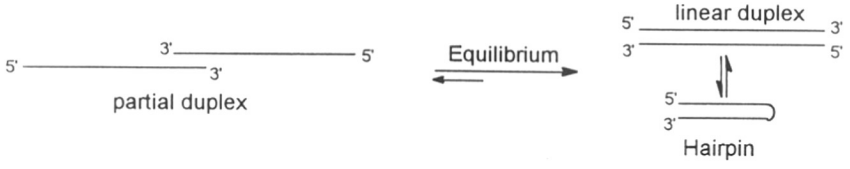
Figure 3: Analysis of PCR products obtained from *RE*-oligonucleotides on 2% agarose gel. lane 1: *RE*AT-TA (**7a**), lane 2: ϕX174-*Hae*III digest (shown partially), lane 3: *RE*CA-TG (**2a**), lane 4: *RE*TC-GA (**4a**), lane 5: *RE*CT-AG (**3a**), lane 6: *RE*CC-GG (**5a**), lane 7: *RE*GC-CG (**8a**), lane 8: *RE*GT-AC (**6a**), lane 9: *RE*AA-TT (**1a**), lane 10: 32 bp duplex.

3.4. 4 Hairpin ⇔ Duplex Equilibrium of PCR Products

The palindrome DNA duplex have sequence that is same when one strand is read left to right or the other is read in right to left direction. This selfcomplementary property of the sequences allows palindromes to form duplex as well as hairpin structure. It is well known in literature that for palindromic sequences, the relative proportion of hairpin and linear duplex formation depends on salt concentration and temperature (Scheffler et al., 1970; Cantor and Schimmel, 1980). The existence of considerable hairpin secondary structure may affect the efficiency of some restriction enzymes as they are specific to double stranded DNA. To assess the hairpin ⇔ duplex equilibrium, the purified double stranded PCR products of *RE*-oligos were subjected to the effect of various concentrations of NaCl and $MgCl_2$. The results were consistent with earlier findings (Scheffler et al., 1970) that increasing salt concentrations favored the linear duplexes, while at lower concentration hairpin formation was favored. However, under these identical conditions used in PCR, the relative ratio of hairpin structure to linear duplex differed from oligo to oligo (Figure 3). The *RE*AT-TA (lane 9), *RE*GC-CG (lane 7) and *RE*AA-TT (lane 1) PCR products showed ~75% hairpin structure while other PCR products showed ~60% hairpin structure. The results indicate that not only the salt concentration but sequence of DNA also contributes in the formation of hairpin structure. Indeed, it has been shown that the presence of GNA triplet in the loop region greatly facilitates the highly stable hairpin structure formation (Yoshizawa et al., 1997). In present work, as shown in Table 2, three oligos i.e. (i) *RE*AT-TA **7a**, (ii) *RE*AA-TT **1a** and (iii) *RE*CA-TG **2a** will have GNA triplet in the hairpin loop region considering the size of loop to be at least six bases (GNA triplets are underlined in Table 2). As shown in Figure 3, in accordance with the results of Yoshizawa et al., (1997), two oligos, (i) *RE*AT-TA **7a** (lane 1) and (ii) *RE*AA-TT **1a** (lane 9), showed higher hairpin formation compared to other oligos. However, *RE*CA-TG **2a** (lane 3) showed same hairpin to duplex ratio as observed in oligos devoid of potential hairpin stabilising sequence GNA in the loop region. In contrast, *RE*GC-CG **8a**

(lane 7) also showed higher hairpin formation (~75%), although it does not contain GNA in its probable loop forming region. The results suggest that, factors in addition to presence of GNA in loop region play important role in hairpin structure stabilisation. The salt dependent effects further proved that the two bands in PCR product are arising from the same DNA which exists in interconvertible hairpin and linear duplex forms.

As the starting oligonucleotides (without having additional GC clamp towards 5' end) have 6 bp complementarity at 5' as well as 3'-ends, the dimers resulting from PCR regain the complementarity at 3' end. Based on this, the dimers can also form partial duplexes and get extended by DNA polymerase. The repetition of this in each PCR cycle should give higher molecular weight products, but such products were not detected.



This may be due to the fact that in single stranded palindrome DNA, a rapid renaturation favors only a hairpin or full duplex structure with negligible partial duplex formation that is crucial for extension by DNA polymerase (Hemat and McEntee, 1994; Palmer et al., 1995).

3.4.5 Application of *RE*-oligonucletides in Identification of Type II Restriction Enzyme Sites.

To demonstrate the utility of these oligonucleotides for finding the cleavage site of type-IIP endonucleases, the commercially available restriction enzymes were chosen, at least one for each of the PCR products. The restriction enzymes, oligonucleotides used and the expected length of the digested fragment are shown in Table 3. The experimental data in Figure 4 indicates that the restriction enzymes cleaved the appropriate *RE*-duplex products efficiently at specific sites. The sequence identification of the generated fragments were done

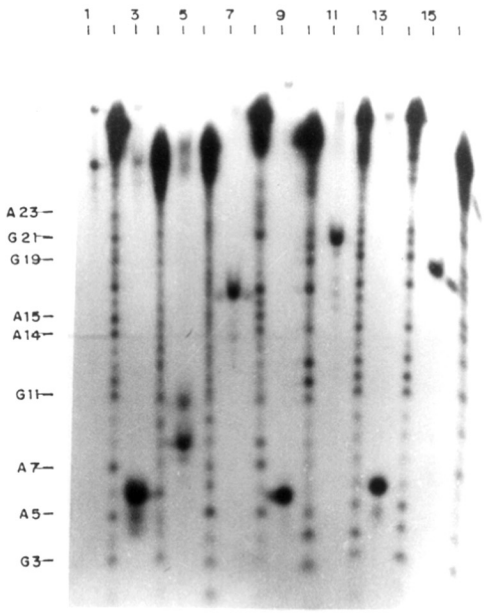


Figure 4: Autoradiograph of restriction enzyme digest of PCR products (odd numbered lanes) and A+G Maxam-Gilbert sequencing (even lanes) of *RE*-oligonucleotides on 8% denaturing PAGE. lane 1: **2a** + *Pvu*II, lane 2: **2,** lane 3: **4a** + *EcoR*I, lane 4: **4** lane 5: **3a** + *Pst*I, lane 6: **3**, lane 7: **5a** + *BamH*I, lane 8: **5**, lane 9: **1a** + *Hind*III, lane 10: **1**, lane 11: **7a** + *Cla*I, lane 12: **7**, lane 13: **8a** + *Nhe*I, lane 14: **8**, lane 15: **6a** + *Sal*I, lane 13: **6**. Numbers on left indicate the A+G sequencing of **2** (lane 2). The lengths of enzyme digest fragments are shown in Table 3.

with A+G chemical sequencing of corresponding *RE*-oligos which proved that the cleavage occurred at the expected site. Inspite of a significant hairpin ⇔ duplex equilibrium in the substrate, the enzyme digestion was neat and complete, proving that both hairpin as well as linear duplexes are efficient substrates for the enzymes. It is possible that for restriction sites in and around the hairpin loop, the cleavage may be more efficient in the duplex DNA whose consumption shifts the equilibrium from hairpin form towards the duplex DNA.

Table 3. PCR primer-dimer and restriction cleavage sites

RE-Oligo	No of bases	PCR product (No of bp)	Restriction enzyme	Cleavage site *	Location on oligomer #
1	37	**1a** (66)	*Hind*III	A*AGCTT	19
2	36	**2a** (66)	*Pvu*II	CAG*CTG	21
3	36	**3a** (66)	*Pst*I	CTGCA*G	23
4	36	**4a** (66)	*Eco*RI	G*AATTC	7
5	36	**5a** (66)	*Bam*HI	G*GATTC	7
6	36	**6a** (66)	*Sal*I	G*TCGAC	27
7	33	**7a** (60)	*Cla*I	AT*CGAT	12
8	33	**8a** (60)	*Nhe*I	G*CTAGC	7
1	37	**1a** (66)	*Taq*I	TC*GA	33
1	37	**1a** (66)	*Alu*I	AG*CT	21
1	37	**1a** (66)	*Dra*I	TTT*AAA	17

* the position of cut in the recognition site of enzyme.
\# expected length of 5' end fragment after enzyme digestion

It is seen from the Figure 4, lane 5, that the restriction enzyme *Cla*I generated a fragment corresponding to 10 bases in addition to the expected 12 base fragment. This may arise from a relaxed specificity for this enzyme. Such effects have been observed (Polisky et al., 1975; Fuchs and Blakesley, 1983) for some restriction endonucleases under non-standard conditions where they are capable of cleaving sequences that are similar but not identical to their defined recognition sequence ("star" activity). The oligonucleotide library described here are composed of closely related sequences and hence has potential for identifying star sites. As they also contain all tetramer palindromes, each represented as unique sites in 4 of the RE-oligomers [e.g. AATT in oligomer **1** (nt 4-7), **2** (nt 4-7), **4** (nt 8-11) and **7** (nt 21-24)], the application can be extended to identify restriction sites of enzymes with tetramer specificity. This was proved by reaction of the duplex **1a** with tetramer recognizing restriction enzymes *Taq*I

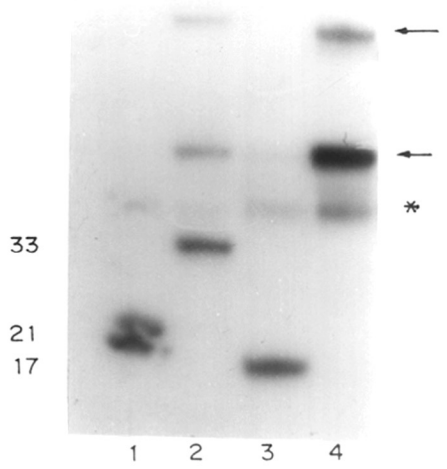


Figure 5: Non-denaturing gel photograph showing digestion of 5' end labeled *RE*AA-TT (**1a**) with different restriction enzymes. lane 1: *Alu*I; lane 2: *Taq* I; lane 3: *Dra*I; lane 4: Only PCR product *RE*AA-TT (**1a**). In addition to the two PCR products (←), small amount of single stranded *RE*AA-TT (**1**) is also seen as faster moving band (•). The numbers on the left indicate sizes of corresponding enzyme digestion fragments.

(T*CGA) and *Alu*I (AG*CT) which gave the expected cleavage fragments (Figure 5). This can also be used to distinguish enzymes which cleave outside the tetramer palindrome recognition site.

3.5 CONCLUSIONS

This chapter describes an algorithm for design of a combinatorial library which represents all possible hexamer palindrome sequences at uniquely defined sites. A linear overlap of rationally arranged hexamer palindromes results in reducing the expected size of 384 bases for 64 hexamer palindromes to 266 bases. The chemically synthesized single stranded oligonucleotides were converted into duplex dimers using PCR. The utility of these duplex oligomers for identifying the cleavage sites of restriction enzymes recognizing hexamer palindromes has been demonstrated for some representative enzymes. The library is also useful for screening restriction enzymes with tetramer cleavage sites. The methods utilizing naturally occurring DNAs as endonuclease substrates determines only the recognition sequence and not the position of cuts within the sequence (Fuchs et al., 1978). Further, when identical cleavage sites are present either too close or overlapping, the resulting fragments will be too small for detection. The present polynucleotide library overcomes this problem since all possible 64 hexamer palindromes are present as unique defined sites allowing a direct determination of both the recognition sequence and the position of cuts. The length of all polynucleotides in the present library is more than 60 bp with a high duplex T_m and these may therefore be suitable substrates as for thermophilic restriction enzymes as well. The templates described here for palindromic sextet recognition and that reported recently for quartet recognition (Hecker and Rill, 1997) have a high information content for novel applications in characterization of sequence-selective binding of enzymes, proteins, drugs and other small ligands to various forms of nucleic acids. They also overcome limitations of natural DNA sequences in terms of unequal representation of all base combinates.

The restriction enzyme recognition site may not be the one which is the strongest binding site of enzyme. In nature, the interactions with intermediate strength (optimal energy) are prefered over those with very strong or very weak interaction energy. In codon:anticodon recognition, out of all possible pairs for a single amino acid, the pair which is neither too sticky or too loose are prefered (Grosjean and Fiers 1982). By providing closely related sequences on single oligomer, *RE*- oligonucleotides provide the comprehensive array of sites which can be used to check above hypothesis.

3.4 REFERENCES

Bock, L. C., Griffin, L. C., Latham, J. A., Vermaas, E. H. & Toole, J. J. (1992) *Nature* **355**, 564.

Brown, N. L. & Smith, M. (1980) *Met. Enzyml.* **65**, 391.

Brown, N. L., Hutchison, C. A. & Smith, M. (1980) *J. Mol. Biol.* **140**, 143.

Brownie et al., (1997) *Nucleic Acids Res.* **25**, 3235.

Cantor, C. R. & Schimmel, P. R. (1980) *Biophysical Chemistry, Part III*, W. H. Freeman & Co. NY.

Chetverin, A. B. & Kramer, F. R. (1994) *Bio/Technology*, **12**, 1093.

Danna, K. & Nathans, D. (1971) *Proc. Natl. Acad. Sci. USA.* **68**, 2913.

Fuchs, C., Rosenvold, E. C., Honigman, A. & Szybalski, W. (1978) *Gene* **4**, 1.

Fuchs, R. and Blakesley, R. (1983) In Grossman, L. & Moldave, K. (eds), *Methods in Enzymol*ogy, Academic Press, INC., California, Vol 100, 30-32.

Grosjean, H. & Fiers, W. (1982) *Gene* **18**, 199.

Hecker, K. H. & Rill, R. L. (1997) *Anal. Biochem.* **244**, 67.

Hemat, F. & McEntee, K. (1994) *Biochem. Biophy. Res. Com.* **205**, 475.

Kelly, Jr., T. J. & Smith, H. O. (1970) *J. Mol. Biol.* **51**, 393.

Kesseler, C. & Manta, V. (1990) *Gene* **92**, 1.

Lin, S. & Arber, W. (1968) *Proc. Natl. Acad. Sci. USA.* **59**, 1300.

Maxam, A. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA.* **74**, 560.

Messelson, M. & Yuan, R. (1968) *Nature (London)* **217**, 1110.

Palmer, E. L., Gewiess, A., Harp, J. M., York, M. H. & Bunick, G. J.(1995) *Anal. Biochem.* **231**, 109.

Polisky, B., Greene, P., Garfin, D. E., McCarthy, B. J., Goodman, H. M. & Boyer, H. W. (1975) *Proc. Natl. Acad. Sci. USA.* **72**, 3310.

Rill, R. L. & Hecker, K. H. (1996) *Biochemistry* **35**, 3525.

Roberts, R. J. & Macelis, D. (1997) *Nucl. Acids. Res.* **25**, 248.

Roberts, R. J. (1976) *Crit. Rev. Biochem.* **4**, 123.

Rychlik, W. (1995) *Mol. Biotechnol.* **3**, 129.

Saiki, R. K.(1989) in Erlich. H.(Ed), *PCR Technology*, 1st Ed. Stockton Press, NY, pp. 7.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) In *"Molecular Cloning: a Laboratory Manual"*, Cold Spring Harbor Press.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA.* **74**, 5463.

Scheffler, I. E., Elson, E. I. & Baldwin, R. L. (1970) *J. Mol. Biol.* **48**, 145.

Smith, H. O. & Wilcox, K.W. (1970) *J. Mol. Biol.* **51**, 379.

Trotta, P. P., Beutel, B. A. & Sherman, M. I.(1995) *Med. Res. Rev.* **4**, 277.

Yoshizawa, S., Kawai, G., Watanabe, K., Miura, K. & Hirao, I. (1997) *Biochemistry* **36**, 4761.

CHAPTER 4

5-amido-deoxyuridine containing oligonucleotides: Effect on *Eco*RV activity and DNA major groove polarity

4.1 INTRODUCTION

Many proteins and enzymes show a very high specificity of interaction with particular DNA sequences. There are two classes of proteins that are extensively studied to understand the mechanism by which this high specificity occurs. The first class includes repressor proteins which bind to operator DNA sequences ranging between 14 and 17 base pairs in length (Ohlendorf & Matthews, 1983; Pabo & Sauer, 1984; Brennan & Matthews, 1989) and the other class includes type II restriction enzymes that recognise and cleave tetra- or haxamer palindrome duplexes (Modrich, 1982; Bennet & Halford, 1989). The studies with these groups have led to general models to explain how specific protein - DNA interaction occur. In the "direct readout" model , the proteins interact directly with bases in DNA (Otwinowski et al., 1988; Matthews, 1988; Brennan & Matthews, 1989). At its simplest, this model assumes that the overall B-DNA geometry does not significantly change with base sequence and that specificity arises from a binding of protein to an unique set of functional groups that may only be presented by its cognate sequence. By using X-ray crystallography, the direct interaction of proteins with bases has been established for several repressor proteins (Anderson et al., 1987; Jordan & Pabo, 1988; Aggarwal et al., 1988; Wohlberger et al., 1988) and the *Eco*RI restriction endonuclease (McClarin et al., 1986). In the alternative "indirect readout" model, where particular sequence of bases change the conformation of DNA, for example, by altering the relative dispositions of the phosphate groups, the protein then makes specific interactions with the altered phosphate group array. For the direct readout mechanism, the groups on the bases of double stranded B-DNA that are accessible to proteins have been determined (Seeman et al., 1976). Figure 1 shows these groups for dA/T base pair [similar features are found in other base pairs; Seeman et al., (1976)]. Most of the interactions between bases and proteins occur via specific hydrogen bonds made either in the major or the minor groove of B-DNA. The only exception to these hydrogen bonding rule is the 5-CH_3 group of T capable of van der Walls interaction with

proteins in the major groove. The direct readout mechanism can be probed by base analogues, and Figure 1 gives the sites within the T and dA bases that are altered to examine their interaction with protein. The examples include the replacement of 6-NH_2 group of dA or 5-CH_3 group of T with hydrogen.

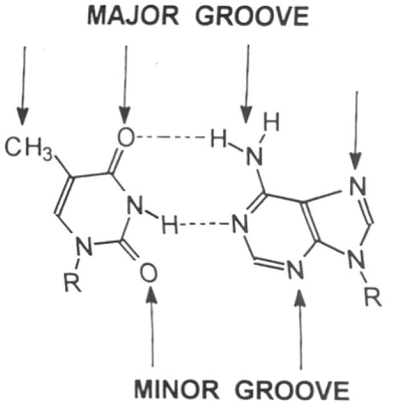


Figure 1. Structure of dA/T base pair in double stranded DNA. The arrows indicate the part of each base that are capable of interacting with proteins in either major or minor grooves. Most interactions are hydrogen bonds with the exception of potential van der Waals contact between proteins and the 5-CH_3 group of T (Seeman et al., 1976).

As restriction endonucleases and repressor proteins show activity with short oligonucleotides, the easiest way is to prepare oligonucleotides containing appropriate base analogues. The structural determinants on DNA responsible for specific interaction between restriction endonucleases and DNA are studied by analysing the effect of nucleotide substitutions in defined positions of the recognition sequence. This approach has been extensively used with restriction endonucleases (Dwyer-Hallquist et al., 1982; Ono et al., 1984; Yolov et al., 1985; Fliess et al., 1986,1988; Seela & Driller, 1986; Jiricny et al., 1986; Brennan et al., 1986; McLaughlin et al., 1987; Seela & Kehne, 1987; Ono & Udea, 1987; Hayakawa et al., 1988; Mazzarelli et al., 1989). Usually the most difficult part of this approach is the preparation of suitable base analogues and their incorporation into oligonucleotides. The success of this approach depends on the availability of a large number of suitably modified DNA substrates which

carry the modification only in one position of recognition sequence. The *Eco*RV interaction with its substrate has been studied using different base analogues (Figure 2) by Newman et al., (1990); Waters & Connolly (1994), and Fliess et al., (1988). For T analogues, these studies involved replacement of 5-CH_3 by either hydrogen which inhibited the enzyme activity or bromine which did not affect enzyme activity. The analogous studies using a hydrophilic substituent such as NH_2 at C-5 position of thymine were not reported possibly due to the unavailability of synthetic procedures describing incorporation of 5-NH_2-deoxyuridine in oligonucleotides. Barawkar and Ganesh (1993) reported the first solid phase synthesis of DNA containing 5-NH_2-deoxyuridine.

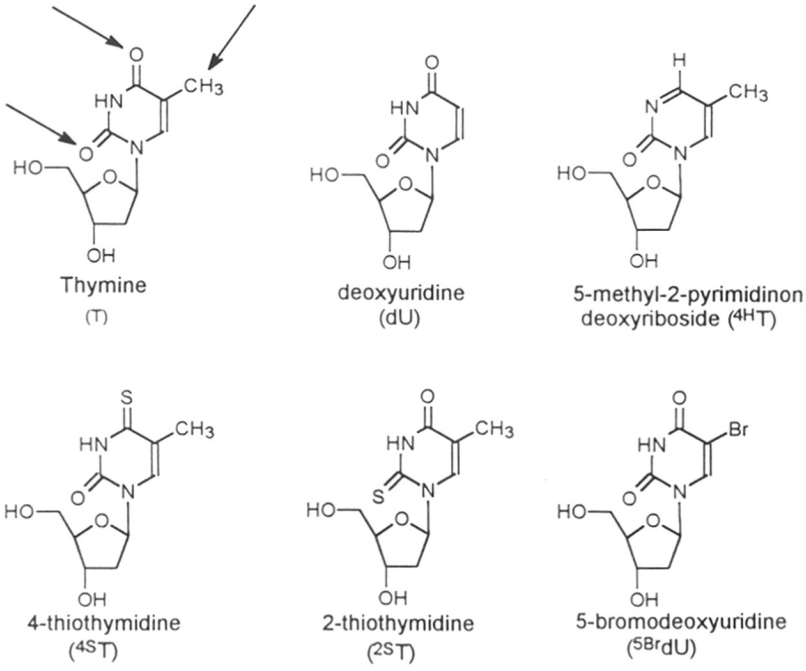


Figure 2. Structure of T. The arrows indicate the part of base capable of interacting with proteins when they are incorporated in double-helical B-DNA. The structures of different analogues of T, reported in literature to probe the these potential contacts are also shown.

4.2 Present Work

This chapter describes studies on consequences of replacement of hydrophobic 5-CH_3 group of thymine by hydrophilic substituent NH_2 in the form of 5-NH_2-deoxyuridine in the recognition sequence of *Eco*RV to further understand specificity of interaction between enzyme and its recognition sequence. In addition, the accompanying changes in polarity of microenvironment in DNA major groove by presence of hydrophilic 5-NH_2 group has been probed by use of 5-amido-dansyl-dU containing DNA. The different oligonucleotides were designed which contain dansyl fluorophore flanked by thymine bases which were systematically replaced by 5-NH_2-dU. The effects of such substitution on the fluorescence properties of dansyl fluoroprobe were monitored to characterise the major groove polarity.

4.3 MATERIALS AND METHODS

All chemicals used were of reagent or better grades. All the solvents used were purified according to literature procedures (Perrin and Armarego, 1989). Usual work-up implies sequential washing of the organic extract with water, brine, water followed by drying over anhydrous sodium sulphate and evaporation of solvent under vacuum. Dansyl chloride was obtained from Aldrich, USA. Restriction enzyme EcoRV was obtained from Amersham. The column chromatography and NMR data was obtained as described in chapter 2, section 2.3.1 and 2.3.2.

4.3.1 Synthetic Procedures to Obtain 4 & 5.

***5'-O-(4,4'-dimethoxytrityl) 5-N-trifluroacetyl-5-amino-2'-deoxyuridine*, 2**

Compound 5-amino-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridine **1** (500 mg, 1mmol; see chapter 2, section 2.3.3 for synthesis of **1**) was taken in dry MeOH (12 ml) to which dry TEA (0.69 ml, 10 mmol) was added and stirred for two minutes. To this, CF_3COOEt (1.2 ml, 10 mmol) was added and the reaction mixture was kept for stirring for 18 hours. The reaction was monitored by TLC on which product showed slightly faster moving ninhydrin negative spot. The solvent was removed under reduced pressure followed by usual work up and purification

by silica gel chromatography (yield 0.5 gm, 78%), Rf (4% MeOH in CH_2Cl_2) = 0.4 ninhydrin negative spot, DMTr positive spot.

H^1 NMR ($CDCl_3$) δ 8.6 (s, 1H, H6), 7.75-7.2 (m, 9H, ArH, DMT), 6.8 (m, 4H, ArH, DMT), 6.35 (t, j=6. Hz, 1H, H1'), 4.4 (m, 1H, H3'), 4.1 (m, 1H, H4'), 3.8 (s, 6H, 2xOCH_3 DMT), 4.45-3.35 (m, 2H, H5', H5") 2.45-2.2 (m, 2H, H2', H2")

***5-N-dansyl-5'-O-(4,4'-dimethoxytrityl)-5-amino-2'-deoxyuridine*, 3**

A solution of dansyl chloride (300 mg, 1.1 mmole) in 2 ml of DMF was added slowly to 5-amino-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridine **1** (540 mg, 1mmole) in 5 ml DMF containing 1.5 ml of saturated $NaHCO_3$ solution. After 30 minutes, 1 ml of 20% of ammonia solution was added and the reaction was continued 15 minutes. After this, reaction mixture was evaporated to dryness and precipitated from water, followed by usual work up and purified by silica gel column. The required compound eluted with gradient of methanol in CH_2Cl_2 containing 0.5% pyridine (yield 0.57 gm, 74%). Rf (4% Methanol in CH_2Cl_2) = 0.46, ninhydrin negative, UV positive fluorescent spot.

H^1 NMR ($CDCl_3$) δ 8.65-7.9 (3xd, 3H, Dansyl ArH), 7.8-7.0 (m, 13H, H6, DMT & dannsyl ArH), 6.85 (d, 4H, DMT & dannsyl ArH), 6.16 (t, J=6.3 Hz, 1H, H1'), 4.36 (m, 1H, H3'), 4.1 (m, 1H, H4'), 3.8 (s, 6H, 2xOCH_3 DMT), 3.53-3.35 (m, 2H, H5', H5") 2.8 (s, 6H, 2xNCH_3), 2.45-2.0 (m, H2', H2").

***3'-O-(N,N-Diisopropylamino)(β-cyanoethoxy)phosphines* 4 & 5**.

Compound **2** or **3** (0.2 mmol) and tetrazole (14 mg, 0.2 mmol) were dried by coevaporation with dichloroethane and suspended in dry dichloroethane (1 ml). To this 2-cyanoethyl N,N,N',N'-tetraisopropylphosphorodiamidite (0.1 ml, 0.31 mmol) was added with stirring for 3 hours. The reaction mixture was diluted with CH_2Cl_2 and washed with 5% aqueous $NaHCO_3$, concentrated to dryness and dried over P_2O_5 and KOH in desiccator before using on DNA synthesizer. Rf (EtOAc: CH_2Cl_2, 1:1 and 0.5% TEA) two close moving spots for two distereoisomers **4**=0.62 ; **5**=0.65

^{31}P NMR ($CDCl_3$) **4** 149.6 & 149.3 ; **5** 149.3 and 149.1 ppm,

4.3.2 Oligonucleotide Synthesis and Purification

The base protected nucleoside phosphoramidites and the nucleoside derivatised controlled pore glass supports were purchased from Cruachem, UK. The DNA synthesis was carried out on Pharmacia LKB-Gene Assembler Plus. The oligonucleotide synthesis and purification was carried out as described in Chapter 2, section 2.3.4.

4.3.3 UV Melting Studies

DNA melting studies were performed on a Perkin-Elmer Lambda 15 spectrophotometer. UV absorption was monitored at 260 nm with a 2 nm slit and 0.5 sec response time. The duplex annealing was done by mixing 1 μM of each of strand in 10 mM Tris buffer (pH 7.0) containing 100 mM NaCl and 10 mM $MgCl_2$, heated to 80 °C for 5 minutes and allowed to cool to room temperature. The samples were thereafter stored at 4 °C overnight. For melting, the annealed duplexes were heated at 0.5 °C per minute and absorption recorded every minute. In a plot of relative change of absorption vs temperature, the temperature corresponding to 50% rise in absorption was taken as the melting temperature (T_m) of the DNA.

4.3.4 Circular Dichroism

Circular dichroism spectra were recorded on a JASCO J600 spectrometer. The samples were scanned with a scan speed of 10 nm/min, slit width of 1 nm and a time constant of 1 sec, in 195 - 240 nm range. Each spectrum was plotted as an average of 5 scans using a 0.5 mm cell. The spectra were recorded using 1 μM of duplex in 10 mM Tris buffer (pH 7.0), containing 100 mM NaCl and 10 mM $MgCl_2$. The buffer alone did not show any background signals in this region.

4.3.5 Cleavage of Synthetic Oligonucleotides by *Eco*RV

The parent oligonucleotide used in this study was a self complementary 13-mer sequence d(GACGATATCGTCA) containing recognition site of *Eco*RV i.e. GATATC. It forms a 12 bp duplex with the overhang of single base A at its 3'

end. This addition of single base A at 3' was done to improve the HPLC resolution of products obtained after enzyme digestion (Figure 3).

The endonuclease assay was performed at 25 °C in 50 µl volume of 10 mM Tris-HCl (pH 7.5), 7 mM $MgCl_2$, 150 mM NaCl, 7 mM 2-mercaptoethanol, and 0.01% BSA. Cleavage experiments were carried out with 40 µM of DNA and 30 Units of enzyme. The progress of reaction was monitored by HPLC using a C18 column and a linear gradient composed of 5 - 30% acetonitrile containing 0.1 M TEAA (pH 7.0) for 20 minutes at a flow rate of 1 ml/min. Under these conditions the two products and substrate oligonucleotide were well resolved. The oligonucleotides which were refractory to cleavage under these conditions were analysed for their ability to inhibit competitively the *Eco*RV catalysed cleavage of unmodified oligonucleotides.

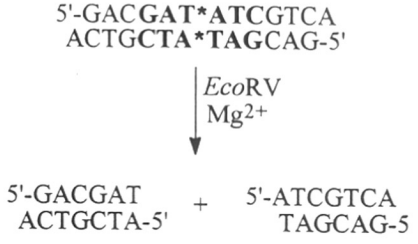


Figure 3. The oligo duplex containing *Eco*RV cleavage site (indicated by *). The cleavage products are identical and in single stranded form they are 6 and 5 nt long.

4.3.6 DNA Major Groove Polarity Measurements

The effect of 5-NH_2-dU substitution in DNA on major groove polarity was characterised by covalently conjugated dansyl fluoroprobe. The fluorescence experiments were carried out on a Perkin-Elmer LS-50B luminescence spectrometer at room temperature. The 2 µM of dansylated oligonucleotides in single strand as well as duplex form were dissolved in 2 ml of Tris-HCl (pH 7.5) buffer containing 10 mM $MgCl_2$ and 100 mM NaCl and the fluorescence measurements were done using excitation and emission slit width of 5 nm each. The Stokes shift was calculated according to the following equation

$$\text{Stokes shift} = (1/\lambda_{ex,max} - 1/\lambda_{em,max}) \times 10^{-7}\ cm^{-1}$$

From excitation spectrum (300 to 400 nm, emission monitored at 500 nm) the wavelength with maximum intensity was taken to be the excitation maximum ($\lambda_{ex,max}$). The emission maximum ($\lambda_{em,max}$) was obtained by scanning emission spectra (range 400 to 600 nm) upon excitation at respective excitation maximum ($\lambda_{ex,\ max}$) of the sample.

4.4 RESULTS

4.4.1 Synthesis of Oligonucleotides Containing 5-NH_2-deoxyuridine

The modified base 5-NH_2-deoxyuridine **1** was synthesised and its 5-NH_2 function was protected with trifluoroacetyl group (Scheme 1) (Barawkar and Ganesh, 1993). This modified nucleoside **2** was converted into required β-cyanoethyl phosphoramidite **4** by usual procedure and incorporated into the oligonucleotides containing recognition sequence of *Eco*RV (5'-GATATC-3'). All oligonucleotides were purified by denaturing gel electrophoresis and their purity was rechecked by HPLC. The oligonucleotide sequences are given in Table 1 indicating the position of modified base. The oligonucleotides **7** & **8** contain 5-NH_2-dU at unique positions within the recognition sequence while **9** has 5-NH_2-dU one base away from recognition sequence towards 3' end.

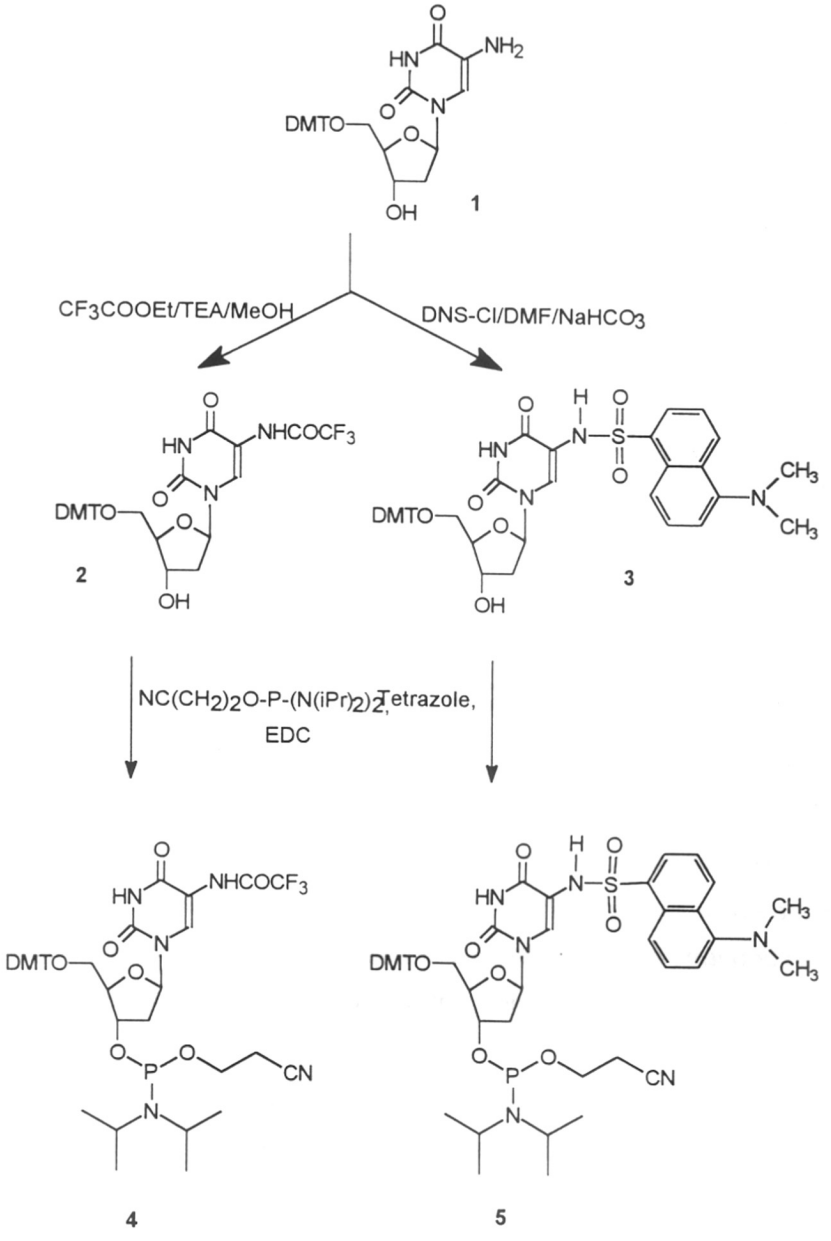
Table 1. Oligonucleotide sequences and their duplex melting temperatures (T_m)#

Oligo no		Sequence	T_m °C
6	d-	GAC**GATATC**GTCA	52
7		GAC**GAU*ATC**GTCA	49
8		GAC**GATAU*C**GTCA	49
9		GAC**GATATCGU***CA	50
10		GACGAT	

U* = 5-NH_2-dU. # T_m values are ± 0.5 °C range. The sequences in bold letters indicate *Eco*RV recognition site.

4.4.2 Spectroscopic Studies of Oligonucleotides

In order to check the general structural properties of the modified oligonucleotides their duplex stability and CD spectra were recorded. As shown in Table 1, modified duplexes were slightly destabilised by 2-3 °C compared to unmodified one but all duplexes had T_m above 45 °C.


DMTO
NH2
HN
O
N
OH
1
CF3COOEt/TEA/MeOH
DNS-Cl/DMF/NaHCO3
NHCOCF3
2
3
CH3
CH3
NC(CH2)2O-P-(N(iPr)2)2, Tetrazole,
EDC
CN
4
5

SCHEME 1

Within the limits of error, all oligonucleotides showed very similar characteristic B-DNA CD spectral pattern, with a positive band at 280 nm, a negative band at 247 nm and a crossover at 265 nm, as shown in Figure 4. The results indicated that, the 5-NH_2-dU substitution in duplex oligonucleotides does not lead to any major distortion of the B-helix. It is an essential prerequisite for such studies that modifications do not change the gross structural features of the modified oligonucleotides compared with unmodified oligonucleotides.

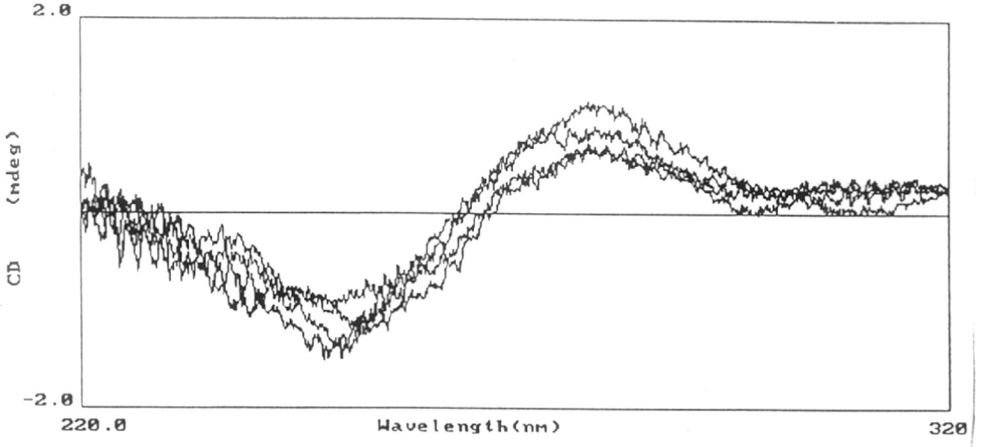


Figure 4. CD spectra of oligos **6**, **7**, **8** and **9** in Tris Buffer (pH 7.0) containing 100 mM NaCl and 10 mM $MgCl_2$ showing minute changes.

4.4.3 Interaction of the Oligonucleotides with *Eco*RV Endonuclease

The cleavage experiments were done at 25 °C, well below the duplex T_m of the oligonucleotides used in this study to ensure the double stranded nature of oligonucleotides. As shown in Figure 5, the unmodified oligonucleotide **6** was efficiently cleaved to give two expected products i.e. 5'-GACGAT-3' **10** & 5'-ATCGTCA-3' **11**, in 12 hours of reaction. The identity of products was proved by coinjecting authentic 5'-GACGAT-3' with reaction mixture which resulted in coelution of these fragments at Rt time (12.4 min). On the other hand, under identical conditions, oligonucleotides **7** and **8** which contain modified base within the recognition sequence were not cleaved by enzyme while **9** which has

Figure 5. HPLC profiles of (a) **10**, (b) *Eco*RV digest of **6**, (c) *Eco*RV digest of **9**,(d) *Eco*RV digest of **7**,(e) *Eco*RV digest of **8**

modified base one base away from recognition site was cleaved as efficiently as unmodified one. The retention times of the oligonucleotides are given in Table 2.

Table 2. HPLC Retention time (R_t) of oligos*

Entry	oligo	R_t (min)
1	**6**	14.5
2	**7**	14.1
3	**8**	13.8
4	**9**	14.0
5	**10**	12.4

* HPLC gradient: 5-30% acetonitrile in 0.1 M TEAA (pH 7.0) for 20 min at a flow rate of 1 ml/min.

The cleavage reactions at 40 μM conc. of unmodified duplex **6** by *Eco*RV (30 units) were done in presence of increasing concentrations of modified duplex **7** for 6 hours and the reactions were analysed by HPLC. As shown in Figure 6, the increase in inhibition of reaction was observed with increasing concentrations of **7** and by 1.25 ratio of [**7**]/[**6**], 32% inhibition of cleavage was observed.

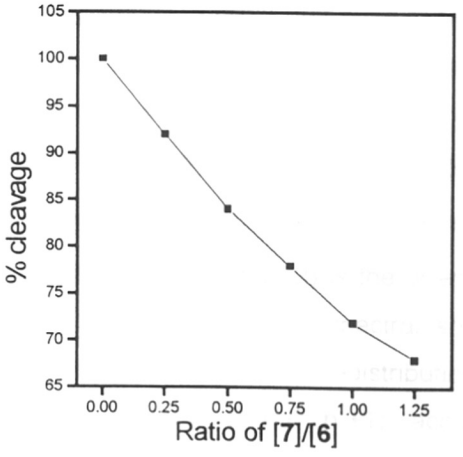


Figure 6. The cleavage of unmodified duplex **6** by *Eco*RV in presence of varying amounts of modified duplex **7**. The cleavage efficiency was calculated by checking the intensity of product **10** formed during reaction. The cleavage efficiency was taken 100% in reaction containing only unmodified duplex **6**.

4.5 POLARITY MEASUREMENTS

Principle: The Lippert equation (1) (Lakowicz, 1983) relates the Stokes shift of given fluorophore with the polarity/dielectric constant of the medium:

$$\Delta\nu = (2\Delta\mu^2/hca^3) \times f \qquad (1)$$

where $\Delta\nu$ is the Stokes shift (the difference between the excitation and emission wavelength, expressed as energy difference in cm^{-1}), $\Delta\mu$ is the difference in dipole moment between the excited and the ground states of the fluorophore, h is the Planck's constant, c is the speed of light and *a* is the radius of the cavity in which the fluorophore resides. The sensitivity of the Stokes shift to orientation polarity (*f*) can be used to estimate the change in dipole moment resulting upon the interaction of the molecule, due to the change in its immediate environment. From a calibration curve for variation of Stokes shift of the free probe with respect to orientational polarity in various aqui-organic binary media (which is usually linear for general solvent effects), the orientational polarity or the solvent environment of the complexed fluorophore can be estimated from its Stokes shift in the bound form.

The interactions between the solvent and fluorophore molecule affect the energy difference between the ground and the excited states. To a first approximation this energy difference (in cm^{-1}) is a property of the refractive index (*n*) and dielectric constant (ε) of the solvent as described by Lippert equation:

$$\upsilon_a - \upsilon_f = 2/hc\ [(\varepsilon-1)/(2\varepsilon+1) - (n^2-1)/2n^2+1)] \times [(\mu - \mu^*)/a^3] + \text{const.} \qquad (2)$$

The term in the first square brackets in eq. (2) is the orientational polarizability (Δf). The term $(\varepsilon-1)/(2\varepsilon+1)$ accounts for the spectral shifts due to both the reorientation of the solvent dipoles and to the redistribution of the electrons in the solvent molecules. The term $(n^2-1)/(2n^2+1)$ accounts for only the redistribution of electrons. The difference of these two terms accounts for the spectral shifts due to reorientation of the solvent molecules.

4.5.1 Use of Dansyl Probe to Characterise Major Groove Polarity

Barawkar and Ganesh (1995) have utilised dansyl conjugated oligonucleotides to determine the polarity of major groove of DNA. Dansyl was chosen as fluorophore because it has large stokes shift and further its high sensitivity to perturbations in local environments such as changes in solvation, ligand binding, etc., by undergoing spectral shifts. The dansyl probe was covalently linked to 5-NH_2-dU, thus placing the fluorophore in major groove (Figure 7). The fluorophore is regiospecifically and rigidly conjugated to DNA by a sulphonamide bond, without much freedom for flexibility averaging over different environments and hence this method was used for groove polarity measurement. Moreover, the Stokes shift of dansyl fluorophore was found to be independent of ionic strength (0.1 M to 5 M) and pH (3 to 5).

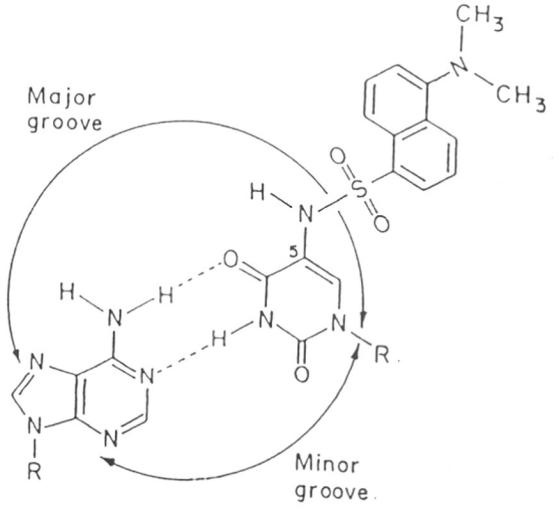


Figure 7. 5-amidodansyl-dU:dA base pairing showing the position of dansyl fluorophore in major groove.

4.5.2 Effect of 5-NH_2-dU Substitution in DNA on Major Groove Polarity

In order to understand the effect of 5-NH_2-dU substitution in oligonucleotides on the polarity of major groove of DNA, the oligonucleotides containing modified bases 5-NH_2-dU and DNS-5NH-dU were synthesised (Table 3). The oligonucleotides were designed in such a way that 5NH-DNS-dU was

flanked by two T in oligo **12**, in which each T was systematically replaced with 5-NH_2-dU to give oligos **13**, **14**, & **15**. The oligonucleotides were designed to be noncomplementary to delineate the effect between single stranded and duplex structures. The sequence of these 13 mer oligonucleotides are given in Table 3.

Table 3. Oligonucleotide sequences used for polarity measurements and their duplex melting temperatures (T_m)

Oligo		Sequence	T_m °C
12	d-	GGTGAT**U^D** TAAGCG	42
13		GGTGAT**U^D** **U$^\#$**AAGCG	39
14		GGTGA **U$^\#$U^D** **U$^\#$** AAGCG	37
15		GG **U^D**GATT **U$^\#$** AAGCG	41
16		CGCTTAAATCACC	

U^D = DNS-5NH-dU and **U$^\#$** = 5-NH_2-dU. The sequence **16** is complementary strand for **12-15**.

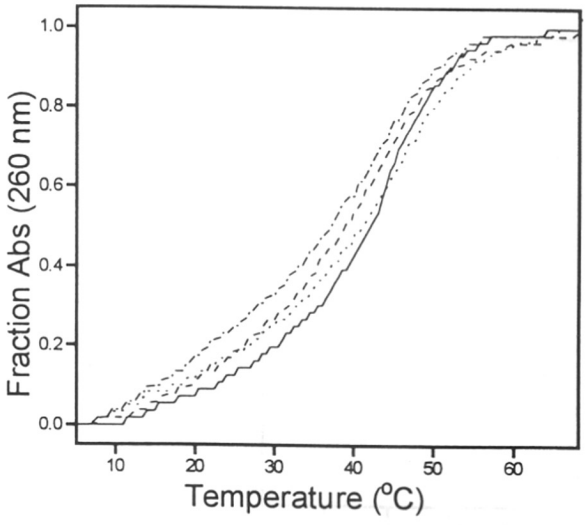


Figure 8. Melting profile of DNA duplexes **12:16** (—);**13:16** (----); **14:16** (-.-.-.-); **15:16** (....)

In order to see the effect of these modifications on duplex stability, the melting behaviour of appropriately annealed duplexes was studied in the range of 5- 85 °C in the same buffer as used for fluorescence studies. As shown in

Figure 8, T_m values for all these duplexes were in the range of 37 to 42 °C and destabilisation of 2-3 °C per 5-NH_2-dU substitution was observed compared to the duplex containing only 5NH-dansyl-dU, **12** (Table 3).The fluorescence properties of differentially labeled oligonucleotides (**12-15)** in single strand as well as duplex form were measured to obtain $\lambda_{ex,max}$ and $1/\lambda_{em,max}$ (Figure 9) to calculate the Stokes shift. These values were compared with the Stoke's shift values of 5-NH_2-dU nucleoside measured in different concentrations of water:Dioxane mixture to calculate the dielectric constant of the oligos (Table 4) as described by Barawkar & Ganesh (1995). The dielectric constant of major groove for **10:14** duplex was found to be ~40D, while for **11:14** and **12:14** it was ~55D and ~62D respectively. The dielectric constant of the major groove of **13:14** duplex was found to be ~52D. The dielectric constant values for all single stranded oligonucleotides (**10-14**) were close to 80D.

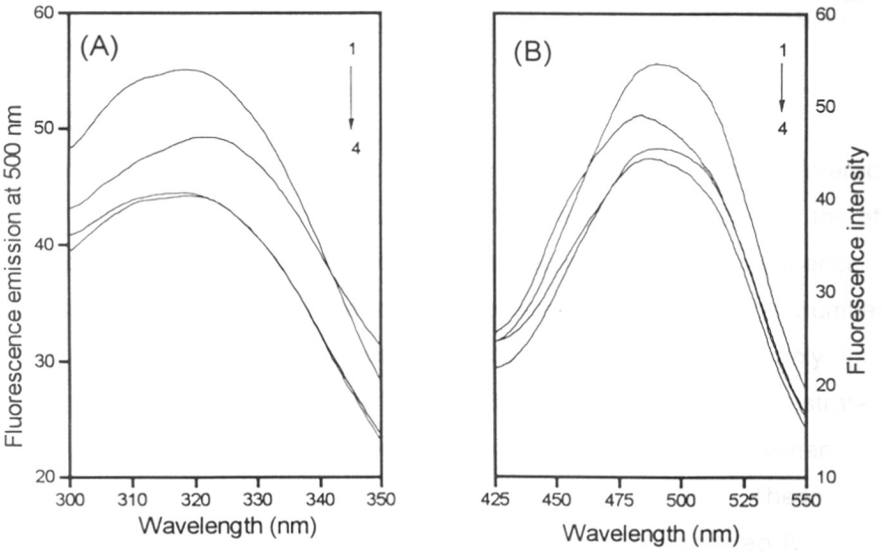


Figure 9 . (A) Excitation spectra and (B) Emission spectra of duplexes. In both spectra 1:**13:16**; 2: **12:16**; 3:**14:16**; 4:**15:16**

Table 4. Stokes shift of 5-aminodansyl-2'-dU in media of different dielectric constants generated from varying ratio of dioxane:water.#

Oligo	%1,4-Dioxane	ε	*n*	*f*	λex (nm)	λem (nm)	Δν cm^{-1}
	0	78.5	1.333	0.320	330	523	11180
	5	72.5	1.338	0.317	328	523	11367
	10	67.0	1.343	0.314	330	531	11470
	15	63.3	1.349	0.312	332	530	11253
	20	58.2	1.354	0.309	334	530	11080
	25	54.2	1.359	0.306	336	530	10900
	30	50.4	1.364	0.303	338	529	10680
	35	45.8	1.369	0.300	338	528	10641
	40	41.3	1.374	0.296	340	528	10481
	45	37.3	1.379	0.292	342	527	10250
	50	32.7	1.383	0.288	342	526	10220
	55	28.2	1.389	0.283	343	524	10060
	60	24.0	1.392	0.277	343	523	10030
	65	20.0	1.397	0.270	343	522	10000
12:16		***40***			***322***	***484***	***10394***
13:16		***55***			***319***	***490***	***10939***
14:16		***62***			***317***	***491.5***	***11200***
15:16		***52***			***319***	***487***	***10814***

The table is adapted from Barawkar and Ganesh (1995). The values in bold indicate the results obtained in present work.

4.6 DISCUSSION

The structural determinants on DNA responsible for specific interaction between restriction endonucleases and DNA are studied by analysing the effect of nucleotide substitutions in defined positions of the recognition sequence. The success of this approach depends on the availability of a large number of suitably modified DNA substrates which carry the modification only in one position of recognition sequence. The *Eco*RV interaction with its substrate has been studied using different base analogues (Figure 2) by Newman et al., (1990); Waters & Connolly (1994), and Fliess et al., (1988). These studies involved replacement of 5-CH_3 by either hydrogen which inhibited the enzyme activity or bromine which did not affect enzyme activity. The studies using a hydrophilic substituent NH_2 at C-5 position of thymine (instead of hydrophobic 5-CH_3) were not possible due to the unavailability of synthetic procedures for incorporation of 5-NH_2-deoxyuridine in oligonucleotides. Barawkar and Ganesh

(1993) reported the first solid phase synthesis of DNA containing 5-NH_2-deoxyuridine using trifluoroacetyl protecting group for 5-NH_2 function.

The restriction enzyme *Eco*RV recognizes the hexamer palindrome d(GATATC) in duplex DNA with nuclease cleaving between central T and A residues (Schildkraut et al., 1984; D'Acry et al., 1985). The high resolution crystal structures of both free enzyme and its complex with oligonucleotides containing its recognition sequence GATATC have been reported (Winkler, 1992; Winkler et al., 1993). The protein is dimeric with a deep cleft between the two subunits that accommodates the DNA-binding site. The enzyme-specific DNA complex shows that the DNA binds with its minor groove facing toward the bottom of cleft and the major groove oriented towards the top.

The enzyme makes most of its contact with DNA via two polypeptide loops. The first of these, centered on Asn 70, approaches the minor groove and contacts phosphate groups. The second loop is centered on Asn185 and penetrates into the major groove of the DNA. Twelve hydrogen bonds (six from each subunit) are made between this loop and the GATATC sequences as shown in Table 5. The exact role of these hydrogen bonds (Table 5) can be probed by using base analogues.

Table 5: Direct contacts observed by crystallography between the *Eco*R V enzyme and its GATATC cognate sequence.

#	Base	Loci on base	Protein contact	Interaction
1	G	6-O 7-N	Asn 185 (backbone) Gly 184 (backbone)	H- bond H- bond
2	A	6-NH_2 7-N	Asn 185 (backbone) Asn 185 (backbone)	H- bond H- bond
3	T[a]			
4	A[a]			
5	T	4-O 5-CH_3	Thr 186 (side chain) Asn 185/Thr 186(side chain)	H- bond vdW
6	C	4-NH_2	Gly 182 (backbone)	H- bond

The bases are numbered from 5'-end according to their position in recognition sequence [a] No direct contacts between oligonucleotide and protein because of extreme DNA distortion.

The studies described here suggest that incorporation of modified base 5-NH_2-dU within recognition sequence of *Eco*RV (T-3/T-5) inhibit the enzyme activity. Such modified oligonucleotides inhibit the cleavage of the unmodified substrates in a competitive manner suggesting that the inactivity of these modified oligonucleotides in acting as a substrate for *Eco*RV is not due to interference in binding. Moreover, the oligonucleotide carrying 5-NH_2 modification just one base outside the recognition site was as good a substrate as the unmodified one. This suggests that the modifications within the recognition site perhaps affects the cleavage mechanism by interfering with formation of an enzyme stabilised transition state.

The reported crystal structure results indicate that, the third T in recognition sequence GATATC does not directly contact the enzyme while fifth T is involved in van der Waals interaction between protein and DNA. The studies by Newman et al., (1990) included deoxyuridine as base analogue for T to check the effect of replacement of 5-CH_3 by hydrogen. Such a replacement resulted in a loss of enzyme activity which may be due to the large difference in size of hydrogen and methyl group, thus suggesting the importance of 5-CH_3 group. However, the replacement of 5-CH_3 by bromine does not affect the cleavage efficiency of *Eco*RV enzyme (Fliess et al., 1986; 1988). It was reasoned that as bromine is only slightly larger than the methyl group and is of similar hydrophobicity (Petruska & Horn, 1983), 5-bromo-dU can replace T in recognition site of *Eco*RV without affecting its activity. Thus *Eco*RV endonuclease seems to interact with its recognition sequence d(GATATC) by making contacts both to the purine and pyrimidine bases via hydrogen bonds and through hydrophobic interactions with the thymine methyl groups (Fliess et al., 1986; 1988). It is therefore very interesting to note from present work that replacement of hydrophobic methyl group by hydrophilic amino group, unlike that of 5-Br replacement completely inhibits the enzyme activity.

These results strengthen the view that hydrophobic interaction between thymine methyl residue and enzyme is of importance during cleavage. Bromine

atom, though more electronegative has a van der Waal radius similar to that of methyl group. It may therefore compensate for the loss of 5-methyl group by forming a direct contact with the protein or by strengthening a nearby interaction. The delineation of contributions from the steric and hydrophobic parameters (Hansch et al., 1973) for 5-substituents of uracil derivative are shown in Table 6. The small size and low hydrophobic parameter for hydrogen atom explains the observed inhibition of hydrolysis of dU substituted oligonucleotides by *Eco*RV. Conversely , the increased size and positive hydrophobic parameter of the bromine atom like that of methyl group restores the substrate activity.

Table 6. Comparison of selected substituents[#]

Substituent	Steric parameter	Hydrophobic parameter	Enzyme activity
$-CH_3$	5.63	0.56	Yes
-H	1.03	0	No
-Br	8.83	0.86	Yes
$-NH_2$	5.42	-1.23	No*

[#] Taken from Hansch et al., (1976)
* Present work

When compared with the above substituents, the amino group has steric parameter similar to that of methyl group, but has a more negative hydrophobicity (hydrophilic) than other substituents. The small difference in steric parameters of methyl group and amino group may not directly affect the hydrolysis by *Eco*RV, since bromine atom was tolerated for hydrolysis under similar conditions. The observed inhibition of cleavage by *Eco*RV of 5-NH_2-dU containing oligonucleotides may be attributed to the enzyme sensitivity to hydrophobic environment which is reversed by amino group substitution. Such functional group replacements may also lead to altered microenvironment in DNA grooves due to changes induced in local *dielectric constant*. The importance of such effects is reflected in the fact that while the major groove of DNA (ε=54.2) is non polar compared to bulk water (ε=80), but considerably more polar than the minor groove (ε~20) (Barawkar & Ganesh, 1995). This may to

some extent explain the predominant interaction of small molecules like drugs in the minor groove while macromolecules such as proteins bind DNA in the major groove. The innumerable crystal structure studies of oligonucleotides have also pointed out significant sequence dependent structural variations which have been correlated with sequence dependent rate of nuclease digestion (Lomonossoff et al., 1981). In light of these results, it is possible that DNA also exhibits a sequence dependent local dielectric changes (microenvironment changes) leading to subtle changes in micropolarity of the groove as a function of sequence. The strategy of changing major groove environment by replacing hydrophobic methyl group by hydrophilic amino group may therefore offer suitable opportunity to study such sequence dependent polarity effects. Hence, this was explored in the present work by determining the Stokes shifts of 5-amido-dansyl-dU incorporated into sequences designed to change the micropolarity. The polarity measurement experiments using dansyl probe showed that the polarity of major groove of DNA significantly changes according to the sequence of bases. The oligonucleotide **12** having thymine (hydrophobic) on both sides of dansyl probe showed a major groove dielectric constant of ~40 while the substitution of one of the thymines by 5-NH_2-dU (hydrophilic) as in **13** increased the dielectric constant to 55 and substitution of both T's by 5-NH_2-dU **14** increased the dielectric constant further to 62. These results indicate a successive increase in hydrophlicity of major groove of DNA by 5-NH_2-dU substitution. The DNA **15** in which 5-amidodansyl-dU is far separated from 5-NH_2-dU, the observed ε is 52, similar to that reported earlier for Dickerson's dodecamer (Barawkar & Ganesh, 1995). Thus, a sequence dependent micorpolarity variation is certainly feasible in DNA, which may considerably modulate its molecular recognition features. It is possible that, the presence of an amino group in the major groove may enhance local hydration network through hydrogen bonding with water, which is absent with hydrophobic methyl group. Thus, the restriction enzyme which bind in the major groove of DNA is placed in much more hydrophlic environment by the 5-NH_2-dU substitution

compared to the hydrophobic environment provided by thymine. The results indicate the importance of thymine methyl groups and support the notion (Ivarie, 1987) that thymine methyl groups are as important in the recognition of specific DNA sequences by protein (Knight & Sauer, 1989; Rajendrakumar et al., 1990) as the more widely recognised hydrogen bonding sites of bases in the major groove (Seeman et al., 1976).

4.7 CONCLUSIONS

This chapter reports the sequence dependent microenvironmental effects in DNA major groove caused by substitution of the hydrophilic 5-amino-dU in place of the standard T (5-methyl-dU) which is hydrophobic. Such a substitution within the recognition site of *Eco*RV makes DNA refractory to enzyme while substitutions just outside the recognition site (even 1 nt away) does not affect the enzyme activity. The enzyme does bind to modified substrate as seen from a competitive binding assay, but cannot cleave it. The dielectric constant as monitored from the fluorescence of incorporated 5-amidodansyl-dU suggests an increase in local polarity in the vicinity of 5-amino-dU. The studies clearly demonstrate the potential of 5-amino-dU and 5-amidodansyl-dU as novel biophysical probes in investigating the role of hydrophobic forces in DNA interaction with other molecules such as drugs and proteins.

4.8 REFERENCES

Aggarwal, A. K., Rogers, D., Drotter, M. Ptashne, M. & Harrison, S. C. (1988) *Science*, **242**, 899.

Anderson, J. E., Ptashne, M. & Harrison, S. C. (1987) *Nature*, **326**, 846.

Barawkar, D. A. & Ganesh, K. N. (1993) *BioMed. Chem. Lett.* **3**, 347.

Barawkar, D. A. & Ganesh, K. N. (1995) *Nucleic Acids Res.* **23**, 159.

Bennet, S. P. & Halford, S. E. (1989) *Curr. Top. Cell. Regul.* **30**, 57.

Brennan, C. A., VanCleve, M. D. & Gumport, R. I. (1986) *J. Biol. Chem.* **261**, 846.

Brennan, R. G., & Matthews, B. W. (1989) *J. Biol. Chem.* **264**, 1903.

D'Arcy, A., Brown, R. S., Zabeau, M., Van Resandt, R. W. & Winkler, F. K. (1985) *J. Biol. Chem.* **260**, 1987.

Dwyer-Hallquist, P., Kezdy, F. J. & Agarwal, K. L. (1982) *Biochemistry* **21**, 4693.

Fliess et al., (1986) *Nucleic Acids Res.* **14**, 3463.

Fliess, A., Wolfes, H., Seela, F. & Pingoud, A. (1988) *Nucleic Acids Res.* **16**, 11781.

Hansch, C., Albert, L., Unger, S. H., Kim, K. H., Nikaitani, D. & Lien, E. J. (1973) *J. Med. Chem.* **16**, 1207.

Hayakawa, T., Ono, A. & Ueda, T. (1988) *Nucleic Acids Res.* **16**, 4761.

Ivarie, R. (1987) *Nucleic Acids Res.* **15**, 9975.

Jiricny, J., Wood, S. G., Martin, D., & Ubasawa, A. (1986) *Nucleic Acids Res.* **14**, 6579.

Jordan, S. R. & Pabo, C. O. (1988) *Science* **242**, 893.

Knight, K. L. & Sauer, R. T. (1989) *J. Mol. Biol.* **264**, 13706.

Lakowicz, J. R. (1983) In "Principles of Fluorescence Spectroscopy" Plenum Press, New York, USA.

Lomonossoff, G. P., Butler, P. J. G. (1981) *J. Mol. Biol.* **149**, 745.

Matthews, B. W. (1988) *Nature* **335**, 294.

Mazzarelli, J., Scholtissek, S. & McLaughlin, L., W. (1989) *Biochemistry* **28**, 4616.

McClarin et al., (1987) *Science* **234**, 1526.

Modrich, P. (1982) *Crit. Rev. Biochem.* **13**, 287.

Newman, P. C., Nwosu, V. U., Williams, D. M., Cosstick, R., Seela, F. & Connolly, B. A. (1990) *Biochemistry*, **29**, 9891.

Ohlendorf, D. H. & Matthews, B. W. (1983) *Annu. Rev. Biophys. Bioeng.* **12**, 259.

Ono, A., Sato, M., Ohtani, Y. & Ueda, T. (1984) *Nucleic Acids Res.***12**, 8939.

Otwinowski et al., (1988) *Nature* **335**, 321.

Pabo, C. O. & Sauer, R. T. (1984) *Annu. Rev. Biochem.* **53**, 293.

Perrin, D. D. & Armarego, W. L. F. (1989) In "*Purification of Laboratory Chemicals*", third ed. Pergamon Press.

Petruska, J. & Horn, D. (1983) *Nucleic Acids Res.***11**, 2495.

Rajendrakumar, G., Ganesh, K. N. & Chatterji, D. (1990) *J. Biol. Chem.* **265**, 22300.

Schildkraut, I., Banner, C. D., Rhodes, C. S. & Parekh, S. (1984) *Gene* **27**, 327.

Seela, F. & Driller, H. (1986) *Nucleic Acids Res.* **14**, 2391.

Seela, F. & Kehne, A. (1987) *Biochemistry* **26**, 2232.

Seeman, N. C., Rosenberg, J. M. & Rich, A. (1976) *Proc. Natl. Acad. Sci.* USA. **73**, 804.

Waters, T. R., & Connoly, B. A. (1994) **33**, 1812.

Winkler et al., (1995) *EMBO J.* **12**, 1781.

Winkler, F. K. (1992) *Curr. Opin. Struct. Biol.* **2**, 93.

Wohlberger, C., Dang, Y., Ptashne, M. & Harrison, S. C. (1988) *Nature* **35**, 789.

Yolov et al., (1985) *Nucleic Acids Res.* **13**, 8983.

CHAPTER 5

Oxidative damage to proteins by Cu(II)Desferal complex

5.1 INTRODUCTION

Radical mediated protein damage has been recently recognized to be an important process during aging (Stadman, 1992). This has led to an increasing number of studies on the chemistry, biochemistry and pathology of radical mediated protein oxidation initiated by metal ion dependent reactions, ionizing radiations, electron leakage and autooxidation of lipids and sugars (Dean, et. al, 1997). *In vivo*, oxidative modifications of proteins mark them for degradation by cytosolic proteases and protein oxidation contributes to the pool of damaged enzymes which increases in size during aging and various pathological stages (Stadman, 1992; Davies & Delsignore, 1987). The primary free radical in most oxygenated biological systems is the superoxide radical ($O_2^{-}\cdot$) in equilibrium with hydroperoxyl radical ($HO_2\cdot$) (Halliwell and Gutteridge, 1989). Biological systems convert these into more reactive species such as peroxyl radical (ROO·), alkoxyl (RO·) and hydroxyl radical (HO·). In the cycled Fenton reaction, the metal ion reduction is effected by $O_2^{-}\cdot$ while oxidation of H_2O_2 leads to a continuous production of reactive hydroxyl radicals (Dean, et. al, 1997).

$$Fe(II) + H_2O_2 \text{ ------} Fe(III) + OH^- + OH\cdot$$

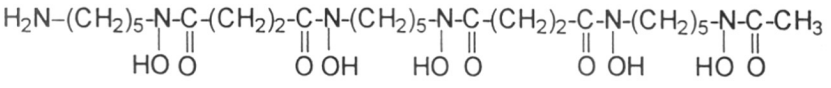
Iron and copper are biologically important transition-metal ions whose reduced forms rapidly cleave organic hydro(lipids) peroxides that initiate the damaging chain reactions. Many complexes of ferrous iron can undergo Fenton reaction. The version of this reaction developed by Udenfriend (1953) uses anionic iron chelate $[Fe(EDTA)]^{2-}$.

$$[Fe(EDTA)]^{2-} + H_2O_2 \text{ ------> } [Fe(EDTA)]^{1-} + OH^- + OH\cdot$$

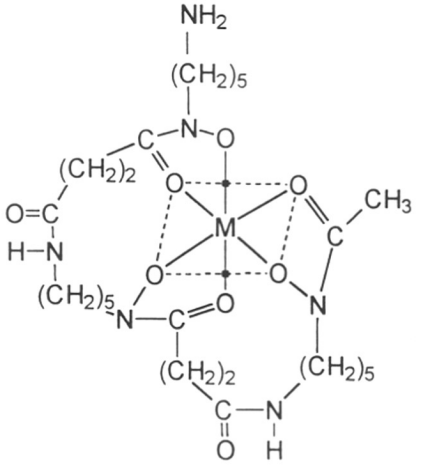
Figure 1 shows the structure of Desferal (DFO) **1**, a naturally occurring siderophore secreted by *Actinomycetes*. In the microbes it occurs as the Fe(III)complex (Ferrioxamine B). Fe(III) can be removed by chemical means to give desferrioxamine B as a colorless crystal. Structurally, it is composed of one molecule of acetic acid, two molecules of succinic acid and three molecules of 1-amino-5-hydroxylaminopentane. These organic units are interlinked via amide

bonds to form a chain containing three hydroxamic acid groups (Halliwell and Gutteridge, 1989).

Many microbes synthesise and secrete the ferric ion selective chelating agents called siderophores like ferrioxamine, which are used to transport iron which is otherwise extremely insoluble. Desferal has an extremely high affinity for ferric (K_a association constant, 10^{31}) while the stability constant for other metals are lower by at least 10^{17} times ($Cu^{2+} > Co^{2+} \sim Zn^{2+} \sim Fe^{2+}$). Owing to its high affinity for ferric ion, it is a highly effective drug used in chelation therapy of iron overload diseases, such as β-thalassemia (Dobbin & Hilder, 1991). It is also used in treatment of Alzheimer's disease to chelate and remove the excess aluminium (Martin, 1991).



1



2 M = Fe, Cu

Figure 1. Struture of desferal **1** and its metal complex **2**

Desferal forms a stable octahedral coordination Fe(III) complex **2** (M = Fe) (Ferrioxamine B), with a metal affinity constant 10^6 times higher than EDTA (Halliwell and Gutteridge, 1986). In contrast to Fe-EDTA which generates hydroxyl radical (HO·) in the presence of a reducing agent such as ascorbate, FeDFO does not undergo a red-ox cycling, thus being ineffective in iron catalyzed hydroxyl radical formation (Graf, et al, 1984). However, desferal administration, has been shown to lead to neurotoxic effects (Arden et al., 1984; Blake et al., 1985; Olivieri et al, 1986; Freedman et al., 1988;) and to enhance the toxicity of paraquat (Osherfoff et al., 1985) and alloxan (Grankvist & Marklund, 1983). Long tern desferal therapy for Alzheimer's disease is also predicted to have side effects (Kruck et al., 1993). Desferal doubles the oxidative inactivation of alkaline phosphatase caused by ascorbate system (Mordente et al., 1990). Earlier it has been shown that while FeDFO is passive in DNA scission, desferal when complexed with other metal ions such as Cu, Co and Ni damages DNA via hydroxyl radical production (Joshi and Ganesh, 1992a, 1992b). This property of Cu (II)-Desferal complex (CuDFO, **2**, M = Cu) has been utilised in triplex mediated targeted cleavage of DNA by attaching metal complex to third strand of DNA (Joshi and Ganesh, 1994). In view of such observations, work was undertaken to examine the effect of metallodesferals on proteins and this chapter describes the oxidative damage to proteins by CuDFO in presence of H_2O_2, a finding that may have important clinical relevance for drug administration. The artificial protease activity (hereafter refered only as protease acivity) of CuDFO is studied by using gel electrophoresis, fluorescence spectroscopy, amino acid modifications and peptide sequencing techniques.

5.2 MATERIALS AND METHODS

Desferal was a kind gift of Hindustan Ciba - Geigy. The highly purified protein substrates were obtained from Sigma USA. Chicken egg white lysozyme (L 7651), Bovine serum albumin (A 0281), Catalase (C 40), Myoglobin (M 0630), Bovine hemoglobin (H 2500), Trypsin (T 8253), Cytochrome C (C 2506) and

Hexokinase (H 5250). The numbers in brackets indicate the catalog number. Lyophilized *Micrococcus lysodeiktius* was obtained from Sigma, USA. Sephadex G-10 was obtained from Pharmacia, Sweden. All reaction solutions were prepared in highly purified water (resistivity = 18 Mohm cm) devoid of metal ion contamination obtained through Milli-Q water purification system (Millipore). All UV-VIS spectra were taken on a Perkin Elmer λ15 spectrophotometer. All other reagents used were of highest purity grade available.

5.2.1 Synthesis and Purification of CuDFO Complex

The CuDFO was prepared and purified as described earlier (Joshi and Ganesh, 1992a). $CuCl_2$ (410 mg, 6.1 mmole) and desferal (210 mg, 0.72 mmole) were mixed in 50 ml water at pH 3.0 and stirred at room temperature for two days. The reaction was monitored by TLC, in which CuDFO showed faster moving spot (Solvent: 65% methanol and 35% of 0.035 M formic acid, Rf = 0.7, green coloured spot, ninhydrin positive). The reaction mixture was lyophilized to obtain green powder. The compound was redissolved in minimum amount on deionised water (1 ml) and was chromatographed on Sephadex G-10 column (150 ml, 10" x 1"). The compound was eluted from the column using deionised water as eluent. Fractions (2 ml each) were collected and scanned on a UV-VIS spectrophotometer from 200 - 600 nm. The fractions containing the peak at 343 nm were then pooled and lyophilised to obtain a green powder. The purity of complex was checked by HPLC using reverse phase C18 Column (solvent system: 30% acetonitrile in triethylammonium acetate, 0.5 M, pH 7.0). Retention time : 5.95 min

5.2.2 Protein Degradation Reaction by CuDFO

The standard reaction mixture contained 100 μg of various proteins were treated separately with 500 μM CuDFO in the presence of 5 mM H_2O_2, in a total volume of 100 μl deionised water. The effect of reducing agent on reaction was checked by addition of β-mercaptoethanol (0.5 mM). The deionised water was used to make all of aqueous reaction ingredients since buffers tend to affect the HO· mediated reactions (Davies et al., 1987b). The reactions were carried out at

37 ^{0}C for different time intervals and terminated by freeze drying. The reactions of lysozyme with CuDFO were also carried out in the presence of various HO· scavengers such as sodium azide (0.1 mM), thiourea (0.1 mM), thymine (10 μM), butanol (2 M), glycerol (2.0 M), mannitol (0.1 M) & catalase (0.5 mg/ml) at 37 ^{0}C for 24 hours in deionised water and analysed by SDS-PAGE.

5.2.3 SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

The CuDFO treated protein reactions were analyzed by SDS-PAGE using a high molarity Tris buffer system as described (Fling and Gregerson, 1988). The Table 1 shows the solutions required and the concentrations used for preparation of SDS-PAGE.

Table 1. Solutions and buffers used for SDS-PAGE

Solutions / Buffers	Stock concentration	Final concentration
A. Acrylamide / Bis (30:1)	40% (w/v)	4 - 15%
B. Resolving and spacer gel buffer		
Tris (pH 8.85)	3.0 M	0.75 M
SDS	0.4% (w/v)	0.1%
C. Stacking gel buffer		
Tris (pH6.8)	0.5 M	0.125 M
SDS	0.4%	0.1%
D. Sample buffer		
Tris (pH6.8)	0.063 M	0.055 M
SDS	2.3%	2.0%
Glycerol	8%	7.0%
2-Mercaptoethanol	5%	4.3%
E. Running buffer		
Tris (pH 8.5)	-	0.05 M
Glycine	-	0.19 M
SDS	-	0.1%

The procedure was modified to obtain better resolution by introducing an additional gel layer, "spacer gel", (containing higher concentration of poly acrylamide than in resolving gel) between stacking gel and resolving gel. The gels were 1 mm in thickness and 2, 2 and 14 cm in length for stacking, spacer and resolving gels respectively. Gel solutions were degassed prior to polymerisation with 0.02% ammonium persulfate followed by addition of 0.04%

TEMED. The resolving gel was polymerised first and was overlaid with water to get even boundary of polymerised gel at top. A similar protocol was followed for spacer gel and stacking gel. The polymerised gels were kept for overnight before use. The CuDFO-protein reaction products (10 μg) were dissolved in 10 μl of sample buffer (Table 1), kept for 5 minutes in boiling water and chilled immediately by keeping ice cold water. These samples were loaded on gels which were run at constant voltage (150 V) until the bromophenol blue dye reached the 3/4th of the gel. The gels with different concentrations of polyacrylamide in stacking, spacer and resolving gels were used for separation for different proteins. Table 2 shows concentrations of polyacrylamide used for making gels to separate different protein-CuDFO reaction products.

Table 2. The % of acrylamide:bisacrylamide (29:1) used in preparing gels for separating different proteins

Protein	Stacking gel %	Spacer gel %	Resolving gel %
Lysozyme, Cytochrome C, Hexokinase, Trypsin, Catalase	4	20	15
BSA, Myoglobin, Hemoglobin,	4	12	8

5.2.4 Silver Staining of Protein Gels

Silver staining is one of the most sensitive technique for the detection of proteins on gel. In present work, to detect the cleavage products with high sensitivity, the gels were developed by silver staining for visualization (Blum, et. al, 1987). The flowchart of the steps followed are summarised in Table 3. All the reagent solutions were prepared fresh just before use. After staining, the gels were dried under vacuum at 80 ^{0}C for one hour by placing the gel between the filter papers in a gel dryer. The densitometric scanning of gels was done using Gel Documentation System from UV Products, UK.

5.2.5 Protein Damage by γ-Radiolysis

The damage to proteins by free radicals generated using water radiolysis is reported by various groups (Davies, 1987; Franzini et al., 1993). To compare the results of CuDFO/H_2O_2 damage to proteins with such systems, the proteins, lysozyme (6.9 μM) and BSA (5.0 μM) in either 10 mM phosphate buffer (pH 7.4) or deionised water at room temperature(25 ^{0}C) were exposed to a steady state irradiation in saturated N_2O solutions for 160 sec using a ^{60}Co irradiator with a source intensity of 0.33 Gy/s.

Table 3. Flowchart of the steps followed in silver staining of SDS-PAGE

STEPS	SOLUTION CONC	WORKING SOLUTIONS	TREATMENT TIME
1. Fixation	50% Methanol, 12% acetic acid	100 ml methanol, 25 ml acetic acid, 75 ml dH_2O	Atleast 1 hr or overnight
2. Wash	50% Ethanol	150 ml ethanol, 150 ml dH_2O	3 x 20 minutes
3. Pretreat	0.02% $Na_2S_2O_3.5H_2O$	0.04 gms in 200 ml dH_2O	Exactly 1 minute
4. Rinse	deionised H_2O	300 ml	3 x 20 sec.
5. Impregnate	0.2% $AgNO_3$, 0.075% of HCOH (37%)	0.4 gms $AgNO_3$, 150 μl HCOH (37%) in 200 ml dH_2O	20 minutes in dark
6. Rinse	deionised H_2O	300 ml	3 x 20 sec.
7. Developer	6% Na_2CO_3, 0.05% HCOH (37%), 0.0004% $Na_2S_2O_3.5H_2O$	12 gm Na_2CO_3, 100 μl HCOH (37%), 200 μl of 4 mg/ml $Na_2S_2O_3.5H_2O$ in 200 dH_2O.	Until the desired intensities of bands appear (around two min.)
8. Wash	deionised H_2O	200 ml	2 x 2 min.
9.Termination of development	50% Methanol, 12% acetic acid	100 ml methanol, 25 ml acetic acid, 75 ml dH_2O	Gels can be stored for a day
10. Gel drying	3% glycerol	6 ml glycerol in 194 ml dH_2O.	30 min

5.2.6 Fluorescence Measurements

The fluorescence measurements were carried out using a Perkin Elmer LS50B spectrophotometer. For tryptophan fluorescence, the protein samples were dissolved in 6M Guanidine hydrochloride prepared in 10 mM phosphate buffer (pH 7.0) and the fluorescence emission was monitored over range of 330-

345 nm upon excitation at 295 nm using excitation and emission slit width of 5 nm each (Capeillere-Blandin et al., 1991). The guanidine hydrochloride used for this study was crystallised from ethanol:benzene mixture (Nozaki, 1972). The time dependent fluorescence loss from BSA and lysozyme during CuDFO reaction was compared by taking the fluorescence from the protein sample at 0 min reaction as 100%. For bityrosine detection, proteins were denatured in 6M urea prepared in 10 mM phosphate buffer (pH 7.0) and the fluorescence emission was monitored in the range of 300-500 nm upon excitation at 315 nm using excitation and emission slit width of 10 nm each (Franzini et al., 1993).

5.2.7 Lysozyme Enzyme Activity

The activity of the CuDFO treated lysozyme was measured according to the method of Shugar (1952) using a Perkin Elmer Lambda 15 UV/VIS Spectrophotometer. The bacterial suspension of *Micrococcus lysodeiktius* (0.7 O.D. at 450 nm) in 2 ml of 10 mM phosphate buffer (pH 7.2) was incubated with CuDFO treated lysozyme samples and decrease in absorbance at 450 nm was recorded after one minute of reaction. The standard curve was obtained by checking the change in absorbance (450 nm) at different time intervals with varying concentrations of lysozyme.

5.2.8 Amino Acid Analysis

Amino acid analysis of CuDFO treated BSA and lysozyme was carried out using a Hewlett Packard Ti series 1050 amino acid analyzer with an auto sampler and 1046A programmable fluorescence detector. Precolumn derivatisation with o-pthalaldehyde (OPA) and fluorenylmethylchloroformate (Fmoc) was employed. All the derivatisation reactions were performed by the autosampler with a programmable injector. The machine was controlled by an interfaced computer using HP Chemstation software. For amino acid analysis, samples were prepared as follows: The CuDFO treated BSA or lysozyme samples (2 mg protein/ml) were hydrolyzed with 6N HCl (1.5 ml) at 120 °C for 24 hours. The hydrolysates were concentrated, dissolved in 300 µl of 0.4 N Borate buffer, pH-10.2 and 1 µl of it was taken for analyses. The sample was mixed with

a 1 µl solution of OPA and mercaptopropionic acid, which resulted in the formation of fluorescent derivatives of all the resulting amino acids from hydrolysis except proline. The imino acid proline was derivatised by addition of 1 µl of Fmoc solution. This mixture was injected on to a C-18 RP column and was analysed with the following buffer system: Buffer A- water with 30 mM sodium acetate, 0.1 mM EDTA and 0.7% THF. Buffer B- 20% 100 mM sodium acetate buffer with 0.1 mM EDTA and 80% acetonitrile. Gradient of A to B was (0 to 9 min - 30% B, at 11 min - 50% B and at 14 min - 50% B) used as given in HP AminoQuant operators manual. The degassing and maintenance of pH was achieved by constant bubbling of buffer solutions with helium gas. These derivatives were monitored with λ_{ex} = 230 nm and λ_{em} = 455 nm (for OPA derivatives); with λ_{ex} = 266 nm and λ_{em} = 305 nm (for Fmoc derivatives). Calibration runs with amino acid standards (100, 50, 25 pmol each) were performed at regular intervals.

5.2.9 N-terminal Sequence Analysis

The amino terminal sequence analysis of lysozyme fragments produced by treatment with CuDFO and H_2O_2 was performed on a Shimadzu PPSQ-10 gas phase protein sequencer equipped with an Edman reaction unit, an on-line phenylthiohydantoin analyzer and a CR-7A data processor. The lysozyme fragments resulting from CuDFO/H_2O_2 treatment were separated on 15% SDS-PAGE. After the electrophoresis, the gel and PVDF membrane were sandwiched between Whatman paper and placed in the blotting cassette. The tank was filled with 10 mM CAPS buffer, pH 11.0, containing 10% methanol and the electrotransfer was carried out under a constant current of 250 mA for 40 min. The PVDF membrane was then washed several times with Milli-Q water and stained with Coomassie Brilliant Blue R-250. The N-terminal amino acid sequence of the first six residues of the protein fragments were determined by subjecting the blot to Edman degradation on an automated protein sequencer as described by LeGendre et al., (1993).

5.2.10 Susceptibility Of Damaged Proteins To Protease

Lysozyme, both native and that reacted with CuDFO/H_2O_2 upto 2 hours were examined for their susceptibility to trypsin by UV spectrometric assay (absorbance at 280 nm) of the shorter protein fragments produced by trypsin digestion. The treatment of CuDFO/H_2O_2 for 24 hours to lysozyme shows 80% of intact protein and only 20% of fragmentation as analysed by SDS-PAGE. Hence, the studies of proteolytic susceptibility were done with protein modified by CuDFO upto 2 hours of treatment. The samples were prepared by addition of 5% TCA to the protease reaction mixture, kept for 15 minutes to allow precipitation of intact protein. The reaction mixture was centrifuged for 15 min at 10,000 rpm and the supernatant was used for checking its absorbance at 280 nm.

5.3 RESULTS

5.3.1 SDS-PAGE Analysis of Protein-CuDFO Reactions

The effect of CuDFO treatment on various proteins was studied using SDS-PAGE and the results are summarised in Table 4. Figures 2, 3, 4 and 5 show SDS-PAGE results of the effect of CuDFO treatment on different proteins under various conditions. The treatment of lysozyme with CuDFO for 24 hours resulted in specific fragmentation of protein in the presence of H_2O_2 giving rise to the three lower molecular weight fragments of size 10, 7 and 4.5 kDa (Figure 2, lane 4). In the absence of either H_2O_2 (lane 2) or CuDFO (lane 3), no cleavage of protein was observed indicating the requirement for both reagents. As shown in Figure 3, In case of myoglobin, CuDFO treatment for 5 hours resulted in specific fragmentation as well as aggregation in presence of H_2O_2 (lane 3) while addition of CuDFO alone was ineffective (lane 2). Bovine hemoglobin (Figure 3, lane 6), Cytochrome C (Figure 4, lane 3) showed extensive aggregation while no fragmentation was observed upon CuDFO/H_2O_2 treatment for 7 hours. Hexokinase (Figure 4) and Trypsin (Figure 5) showed multiple bands in the control itself and CuDFO/H_2O_2 treatment for 7 and 5 hours respectively, resulted in loss of intensities of these bands while no low molecular weight fragments were observed. As shown in figure 5, extensive aggregates of catalase were

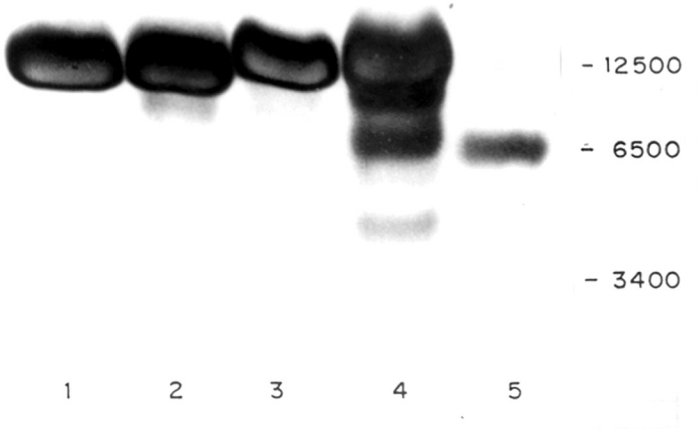


Figure 2. SDS-PAGE analysis of the 24 hour reaction of lysozyme (6.9 μM) with CuDFO as visualized by silver staining. Lanes 1-4 contain lysozyme with the following additions. lane 1: none; lane 2: H_2O_2 (5 mM); lane 3: CuDFO (500 μM); lane 4: CuDFO (500 μM) + H_2O_2 (5 mM); lane 5: Low molecular weight protein marker.

observed in presence of CuDFO alone within 5 hours of reaction while further addition of addition of H_2O_2 enhanced the aggregation. Unlike the Fe-EDTA system, no reducing agent (e.g. 2-mercaptoethanol) was required in the CuDFO mediated reaction. A similar set of reactions on BSA (Figure 6) with CuDFO/H_2O_2 resulted in the major fragmentation of the protein within an hour. In contrast, lysozyme showed only a partial fragmentation even after 24 hour of reaction (Figure 2). Ionizing radiation is well known to damage proteins and for comparison, the pattern of fragmentation by γ-radiolysis of BSA is shown in Figure 6 (lane 9). The degradation pattern of BSA observed in radiolysis was different compared to that induced by CuDFO. Protein damage reactions were analysed by native and denaturing gel in presence of 7 M urea, also showed results similar to that obtained with SDS-PAGE indicating the covalent nature of aggregates.

Table 4. The effect of CuDFO/H_2O_2 treatment on various proteins as analysed by SDS-PAGE

Protein	**Reaction time (Hours)**	**Effect of CuDFO/H_2O_2 treatment**
Lysozyme	24	Fragmentation
BSA	2	Fragmentation
Bovine hemoglobin	5	Aggregation
Catalase	5	Aggregation
Cytochrome C	7	Aggregation
Myoglobin	5	Aggregation / fragmentation
Trypsin	5	Loss of intensities of protein bands
Hexokinase	7	Loss of intensities of protein bands

After establishing the protease activity of CuDFO/H_2O_2 on various proteins, further studies were concentrated on BSA and lysozyme as they are structurally very well characterised proteins in literature (Brown, 1977; Blake et al., 1967). Several reports of modification by oxygen radicals of these proteins are available for comparison (Davies, 1987; Franzini et al., 1993). Moreover, BSA represents the class of serum proteins which are easily accessible during desferal therapy.

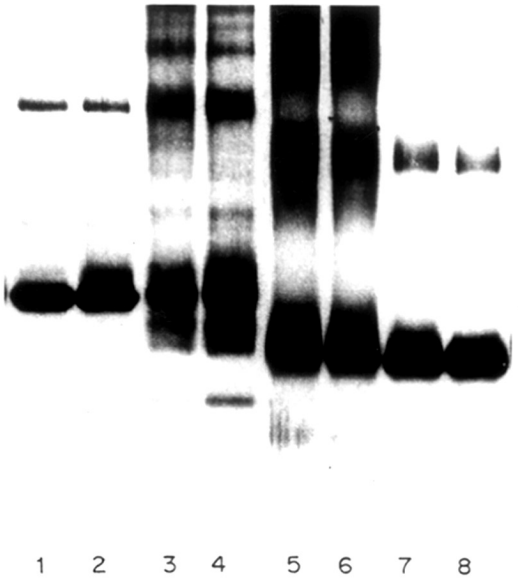


Figure 3. SDS-PAGE analysis of the 5 hour reaction of myoglobin (lane 1-4) and hemoglobin (lane 5-8) with CuDFO as visualized by silver staining. Lanes 1-4 contain myoglobin with the following additions. lane 1: none; lane 2: CuDFO (500 μM); lane 3: CuDFO (500 μM) + H_2O_2 (5 mM); lane 4: CuDFO (500 μM) + H_2O_2 (5 mM) + ME (0.5 mM). Lanes 5-8 contain bovine hemoglobin with following additions. lane 5: CuDFO (500 μM) + H_2O_2 (5 mM) + ME (0.5 mM): lane 6: CuDFO (500 μM) + H_2O_2 (5 mM); lane 7: CuDFO (500 μM); lane 8: none.

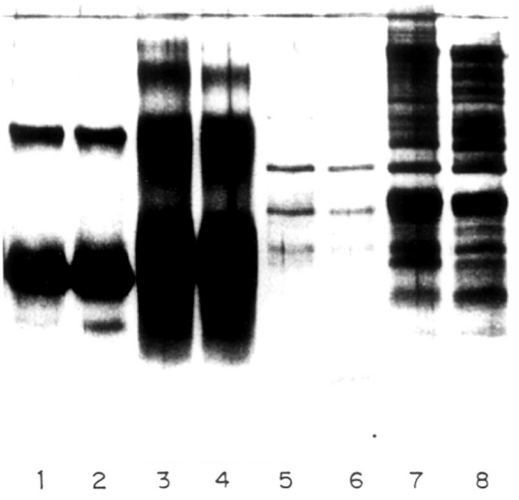


Figure 4. SDS-PAGE analysis of the 5 hour reaction of cytochrome C (lane 1-4) and Hexokinase (lane 5-8) with CuDFO as visualized by silver staining. Lanes 1-4 contain cytochrome C with the following additions. lane 1: none; lane 2: CuDFO (500 μM); lane 3: CuDFO (500 μM) + H_2O_2 (5 mM); lane 4: CuDFO (500 μM) + H_2O_2 (5 mM) + ME (0.5 mM). Lanes 5-8 contain hexokinase with following additions. lane 5: CuDFO (500 μM) + H_2O_2 (5 mM) + ME (0.5 mM): lane 6: CuDFO (500 μM) + H_2O_2 (5 mM); lane 7: CuDFO (500 μM); lane 8: none.

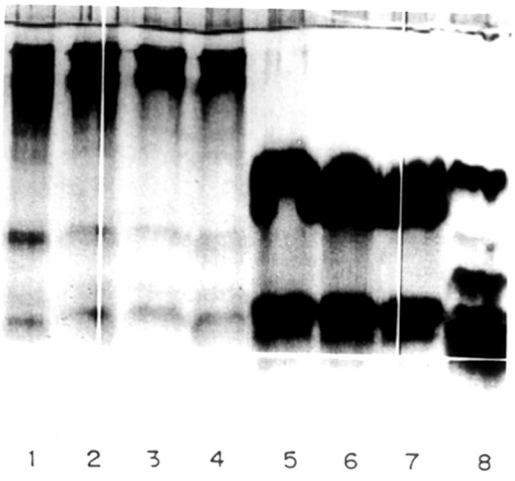


Figure 5. SDS-PAGE analysis of the 5 hour reaction of trypsin (lane 1-4) and catalase (lane 5-8) with CuDFO as visualized by silver staining. Lanes 1-4 contain trypsin with the following additions. lane 1: none; lane 2: CuDFO (500 μM); lane 3: CuDFO (500 μM) + H_2O_2 (5 mM); lane 4: CuDFO (500 μM) + H_2O_2 (5 mM) + ME (0.5 mM). Lanes 5-8 contain catalase with following additions. lane 5: CuDFO (500 μM) + H_2O_2 (5 mM) + ME (0.5 mM): lane 6: CuDFO (500 μM) + H_2O_2 (5 mM); lane 7: CuDFO (500 μM); lane 8: none.

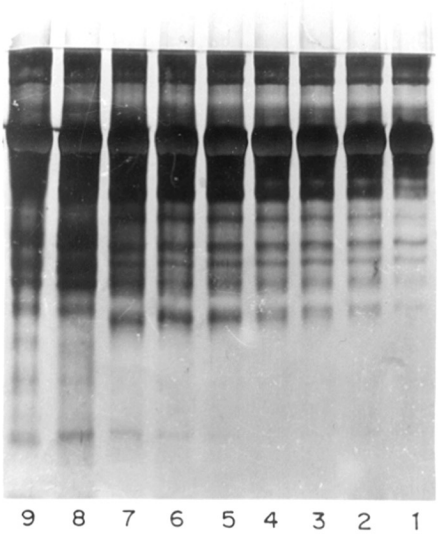


Figure 6. SDS-PAGE analysis of BSA (5 μM) treated with CuDFO in presence of H_2O_2 for different time intervals. lane 1: 0 min; lane 2: 15 min; lane 3: 30 min; lane 4: 45 min; lane 5: 60 min; lane 6: 90 min; lane 7: 120 min; lane 8: 150 min, lane 9: BSA (5 μM) ionized by γ- radiolysis (55 Gy).

5.3.2 Effect of HO· Scavangers

It is certian that some active free radical species generated by CuDFO/H_2O_2 causes the structural damage to proteins. Various scavangers of free radicals are useful in determining the nature of damaging species and its mode of action. The effect of HO· scavangers on lysozyme fragmentation by CuDFO/H_2O_2 treatment for 24 hours was studied (Figure 7) by SDS-PAGE analysis. The reaction was completely inhibited by free radical scavengers such as sodium azide, thiourea and catalase while boiled catalase, glycerol, butanol, mannitol and thymine were found to be ineffective.

5.3.3 Effect on Protein Fluorescence

The aromatic side chain amino acids tryptophan and tyrosine are most susceptible to oxidative damage (Davies et al., 1987a) and the extent of their damage in CuDFO/H_2O_2 treated proteins was monitored by fluorescence studies. Although BSA contains several tyrosine and phenylalanine residues and only 2 tryptophans, more than 99% of the fluorescence observed at 330-345 nm region (295 nm excitation) can be attributed to tryptophan fluorescence (Davies et al., 1987a).

Upon exposure to CuDFO/H_2O_2, a gradual decrease in tryptophan fluorescence emission within the 332-345 nm region on excitation at 295 nm was seen for BSA in time dependent manner. By 8 hours, 68% of loss in fluorescence intensity was observed in addition to the slight blue shift in the emission maximum (≈11 nm) (Figure 8, curves a-e). The crosslinking (both inter and intramolecular) of tyrosine molecules in proteins results in the formation of fluorescent bityrosine product which gives rise to covalent aggregates (Boguta & Dancewics, 1983). The formation of fluorescence active bityrosine in the region of 410-430 nm upon excitation at 315 nm was detected following radiolysis of BSA (Figure 8, curve g). However, no such characteristic fluorescence emission was seen with BSA exposed to CuDFO/H_2O_2 for 8 hours. (Figure 8, curve f)

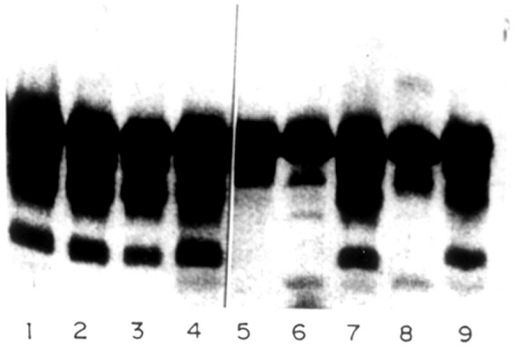


Figure 7. SDS-PAGE analysis of the 24 h reaction of lysozyme (6.9 μM) with CuDFO/ H_2O_2 in presence of various scavangers as visualized by silver staining. lame 1. No scavenger addition; lane 2: thymine (10 μM); lane 3: glycerol (2.0 M); lane 4: mannitol (0.1 M); lane 5: sodium azide (0.1 mM); lane 6: thiourea (0.1 mM); lane 7: butanol (2 M), lane 8: catalase (0.5 mg/ml) and lane 9: catalase (0.5 mg/ml)

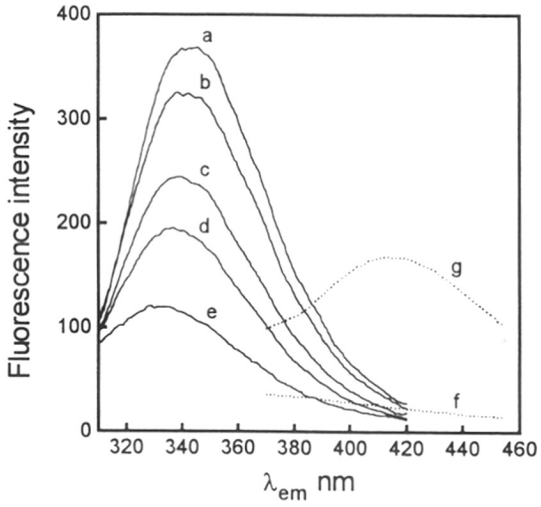


Figure 8. The fluorescence emission spectrum of BSA treated with CuDFO/H_2O_2 (a-f) and BSA exposed to radiation dose of 55 Gy (g). a-e recorded at different time intervals of reaction (tryptophan ex 295 nm, spectral bandwidth 5 nm). a: 0 h, b: 0.5 h, c: 2 h, d: 4 h, e: 8 h. The emission spectra f (CuDFO/H_2O_2, 8 h) and g were recorded by excitation at 315 nm (bityrosine) and spectral bandwidth of 10 nm.

5.3.4 Effect on Amino Acid Composition of Proteins

The amino acid composition analysis of BSA and lysozyme exposed to CuDFO/H_2O_2 were carried out to assess the oxidative susceptibility of the constituent amino acids of BSA and lysozyme. Figure 9 shows the representative chromatographs of amino acid compositions of (A) standard mixtures containing 100 pmol of each amino acid, (B) lysozyme and (C) lysozyme treated with CuDFO/H_2O_2 for 24 hours. By comparing with standard, % molar ratio of each amino acid present per mole of protein was calculated. The results of such analysis for BSA and lysozyme reactions are given in Table 5 and 6 respectively. As shown in Table 5, in case of BSA, the levels of the four amino acids - methionine, histidine, lysine, and proline were considerably decreased during the

oxidation while in lysozyme histidine, tyrosine and methionine were found to be affected (Table 6).

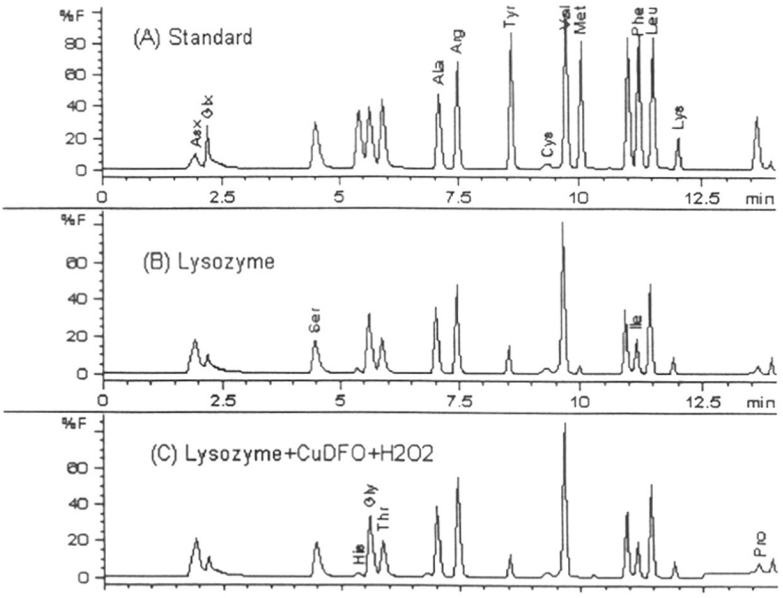


Figure 9. The chromatographs showing the amino acid composition of (A) Standard mixture 100 pmol of each amino acid, (B) lysozyme and (C) lysozyme treated with CuDFO/H_2O_2 for 24 hours.

Table 5. Amino acid composition of lysozyme and BSA upon treatment with CuDFO/H_2O_2.*

Amino acid	% amino acid # at 0 min	% amino acid # at 24 hours	% change
His	3.70	2.59	-30
Met	0.63	0.29	-54
Lys	11.43	9.96	-13.1
Pro	5.87	4.84	-18.6

* % change in other amino acids was ± 5

Table 6. Amino acid composition of lysozyme upon treatment with CuDFO/H_2O_2.*

Amino acid	% amino acid # at 0 min	% amino acid # at 24 hours	% change
His	1.14	0.99	-14
Tyr	2.15	1.69	-21
Met	1.11	0.74	-34

* % change in other amino acids was ± 5

5.3.5 N-terminal Sequencing of Lysozyme Fragments from CuDFO Reaction

The treatment of CuDFO/H_2O_2 to lysozyme results into the specific fragmentation of protein (Figure 2). In order to determine the site of cleavage and hence specificity, the products of lysozyme reaction with CuDFO/H_2O_2 were subjected to amino terminal sequence analysis. The three products corresponding to 10, 7 and 4.5 kDa fragments resolvable on electrophoretic gel (Figure 2) were electroblotted on PVDF membranes which were then subjected to N-terminal sequencing. The results showed that while the 10 and 7 kDa fragments had the amino terminal sequence of the native protein itself while the 4.5 kDa fragment could not be sequenced possibly due to modification at the N-terminal end during the oxidative cleavage.

5.3.6 Effect on Lysozyme Activity

The impact of the damage on lysozyme caused by CuDFO was also assessed by the effect on enzyme activity. As shown in Figure 10, lysozyme exposed to CuDFO/H_2O_2 is slowly deactivated and by 24 hours, 55% of its activity is lost.

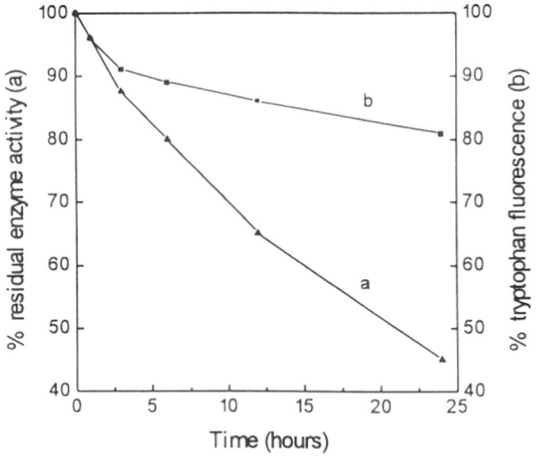


Figure 10. The plot of lysozyme activity (curve a) and tryptophan fluorescence (curve b) as a function of time after reaction with CuDFO/ H_2O_2.

5.3.7 Proteolytic Susceptibility

As shown in Figure 11, CuDFO/H_2O_2 damaged lysozyme showed increased susceptibility to trypsin digestion in time dependent manner. The time dependent increase in the TCA soluble material was observed upon treatment of lysozyme with CuDFO/H_2O_2. The 2 hour exposure of protein increased the acid soluble protein fraction by three times indicating the increased susceptibility of damaged protein.

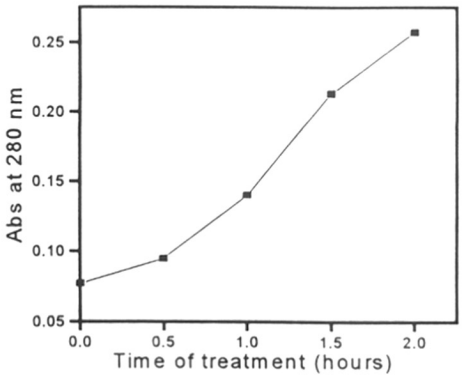


Figure 11. The plot shows increase in absorbance at 280 nm of TCA soluble material upon trypsin digestion of lysozyme treated with CuDFO/ for different time intervals (0 - 2 hours).

5.4 DISCUSSION

The gel electrophoretic analysis of the reaction products of various proteins with the CuDFO/H_2O_2 systems indicated two types of modifications: fragmentation and covalent aggregation. The fragmentation may occur at preferred sites in some proteins (e.g. lysozyme) leading to specific products of low molecular weight, whereas in other cases (e.g. BSA), the smear observed on gel indicates a non-specific fragmentation. The formation of covalent aggregates in some proteins was shown by low mobility, high molecular weight bands on the gel. This is in agreement with the identical behavior of proteins upon treatment with other radical producing systems such as Fe-EDTA/ascorbate (Uchida, et. al, 1989) or water radiolysis (Davies, 1987). Such aggregations are known to occur

through intermolecular processes mediated by radicals, predominantly by crosslinking of side chains through S-S and Tyr-Tyr bonds (Stadman, 1993).

The inhibition of protein damage reactions by radical quenchers such as sodium azide, thiourea and catalase confirms the mediation of hydroxyl radicals in the oxidative process. Boiled catalase did not inhibit the reaction confirming the necessity of H_2O_2 in reaction. However, the metal ion catalysis involves no prior reduction to lower oxidation state since the reaction does not require a reductant such as DTT or ascorbate. The H_2O_2 reacts directly with CuDFO leading to hydroxyl radical production accompanied by reduction of the metal to a lower oxidation state. The inhibition was not observed with glycerol, mannitol, butanol and thymine possibly due to a restricted access of these trapping agents to the hydroxyl radicals. Such reactions may be viewed as caged processes in which the active oxygen species is not released in the surrounding medium but preferentially reacts with functional groups of amino acid residues at CuDFO binding site. This is supported by the selective nature of the CuDFO catalyzed oxidation observed with lysozyme (Figure 2) as a consequence of localization of the complex at particular sites on protein. The appearance of three distinct fragments by exposure of lysozyme to CuDFO suggests that the cleavage occurs at least two specific sites. It is known that radiolytic cleavage of BSA, hemoglobin and myoglobin lead to extensive fragmentation in particular, at proline sites. A comparison of data with the metal ion peroxide system shows that the fragmentation patterns are distinctly different (Puchala and Schyessler, 1993). The difference in the size of fragments of BSA on exposure to CuDFO/ H_2O_2 observed and ionising radiation implying a difference in mechanism and/ or kinetics of fragmentation (Figure 6).

The time required for the damage of proteins by the CuDFO treatment suggests that large proteins such as BSA, are more susceptible to degradation than small compact proteins such as lysozyme. The high susceptibility of the large proteins like BSA compared to small compact proteins like lysozyme is in agreement with a view that although many amino acids are susceptible to

modifications by $HO^{\cdot}$ radical action, the protein structure could greatly influence the reactivity of oxygen radicals (Phelps et al., 1961)

Although, all amino acids are susceptible to $HO^{\cdot}$ attack, the primary, secondary and tertiary structure of a protein greatly influences the reactivity of each amino acid (Davies et al., 1987a; Stadman, 1993). It is known that, in BSA certain amino acids (tyrosine, phenylalanine, tryptophan, histidine, methionine, cystine) are the preferred targets for attack by hydroxyl radicals (Davies, et al, 1987a). The data of Table 5 on amino acid analysis of BSA reacted with CuDFO, suggests that in the case of BSA significant oxidative damage occurs with histidine, lysine, proline and methionine. In case of lysozyme, methionine, tyrosine and histidine were more susceptible to attack by CuDFO/H_2O_2. This difference in susceptibility of these two proteins could be due to the difference in three dimensional protein structure. While histidine and lysine are hot spots for complexation with Cu and hence prone to metal ion generated hydroxyl radical attack, methionine is most susceptible to oxidation. The loss in tryptophan fluorescence intensity of BSA accompanied by a slight blue shift in emission maxima following its treatment with CuDFO/H_2O_2 indicates the damage of tryptophan as well as an alteration in its environment effected by modification and/or protein unfolding. Unlike BSA radiolysis, which leads to bityrosine production through interchain/intermolecular coupling of tyrosyl radicals, CuDFO-BSA reaction did not result in bityrosine formation (Figure 8). This is also supported by the fact that no covalent aggregates were observed on SDS-PAGE for BSA or lysozyme treated with CuDFO. However, amino acid analysis results of lysozyme treated with CuDFO showed 21% loss of tyrosine while no bityrosine formation was observed using fluorescence spectroscopy (Figure 10). The results suggest that the oxidation of tyrosine by CuDFO/H_2O_2 may lead to products other than bityrosine which is not involved in crosslinking.

Two tryptophans have been shown to be part of the active site of lysozyme (Blake et al., 1967) and their damage by radicals may lead to enzyme inactivation. Although only 19% of the tryptophan fluorescence was lost during

the 24 h treatment of lysozyme, corresponding loss of enzyme activity was 55% (Figure 10). This indicates that not only tryptophan but other amino acids like methionine, tyrosine, and histidine damaged by CuDFO treatment as evidenced by amino acid analysis cause the structural changes in protein leading to a loss of enzyme activity. These results indicate that the oxidative damage of enzymes not only increases its proteolytic susceptibility but also reduces its activity and thus may play an important role in aging.

In the lysozyme-CuDFO/H_2O_2 reaction, no aggregation was observed but the production of protein fragments of definite size indicate either a specific binding of metallodesferal with the protein or the presence of "weak" sites on proteins, susceptible to cleavage. The oxidative fragmentation of protein which does not involve the formation of a new N-terminus poses difficulties in N-terminal sequencing (Dean et al., 1997). A well established modification is the peroxy radical mediated α-amidation to produce a carboxy terminal amide and an NH_2 blocked by α-ketoacyl group (Garrison, 1987). The N-terminal sequence analysis of the 10 and 7 kDa fragments of lysozyme/CuDFO reaction showed them to have the identical sequence of the native protein over a stretch of six amino acids. The third fragment produced in the reaction could not be sequenced satisfactorily possibly due to a modification at N-terminal end.

The synthetic metallopeptidases carrying out specific fragmentation of proteins are generating interest due to their potential use in elucidating the solution structure of proteins as well as sequencing of large proteins (Hegg and Burstyn, 1995; Rana and Meares, 1992). The specific fragmentation of proteins caused by Cu-DFO may be useful in such applications.

Oxidative modifications of proteins by $HO^{\cdot}$ generated using ionising radiation are known to cause an increased proteolytic susceptibility of the damaged proteins and this phenomenon is taken as a sensitive biological index of protein modification (Davies & Goldberg, 1987; Davies et al.,1987b; Stadman, 1992). This may be primarily due to the denaturation of proteins leading to its loss of native structure since heat denatured proteins also show an increased

proteolytic susceptibility (Davies & Goldberg, 1987). Protein oxidation contributes to the pool of damaged enzymes, which increase in size during aging and various pathological states. Stadman (1992) has postulated that the age related increase in amounts of oxidised protein may reflect the age-dependent accumulation of unrepaired DNA damage. This, in a random manner, affects the concentrations or activities of numerous factors that govern the rates of protein oxidation and degradation of oxidised protein.

5.5 CONCLUSIONS

Free radical damage to cellular functions through oxidative modifications of proteins, nucleic acids and lipids is associated with age-related diseases such as artherosclerosis, arthritis, muscular dystrophy and cancer (Stadman, 1992). It is demonstrated that the copper complex of desferal, a chelation drug used in therapy of β-thalassemia, causes oxidative modification of amino acids in proteins, leading to aggregation, fragmentation and loss of enzyme activity. The key amino acids susceptible to CuDFO induced oxidative damage are tryptophan, lysine, methionine, cysteine and histidine and may differ among proteins, depending on the binding site of CuDFO on protein. The damage seen in many proteins of the circulation system, such as BSA, haemoglobin etc. may occur through their reactivity with hydroxyl radicals. These findings hold importance as clinical trials have shown that increased toxicity of desferal is observed when supplemented with ascorbate (Davies et al., 1983; Nieuhaus, 1981). In normal blood plasma containing micromolar levels of Cu(II), ascorbate is loosely bound to the albumin (Oelrichs et al., 1984; Marx and Chevion, 1985). The oxidative damage of proteins in presence of reducing agents such as ascorbate and the accompanying toxicity may originate from complexes of desferal with metals like Cu(II). In view of the findings that desferal, generally administered intramuscularly, exhibits acute toxicity (Davies et al., 1983; Nieuhaus, 1981) the present findings of protease activity of CuDFO and previously known oxidative DNAse activity (Joshi & Ganesh, 1992a), may have importance in understanding the toxic side effects during desferal therapy.

5.6 REFERENCES

Arden, G. B., Wonke, B., Kennedy, C. & Huehns, E. R. (1984) *Brit. J. Ophthalmol.* **68**, 873.

Blake, D. R., wyngard, P., Lunec, J., Williams, A., Good, P. A., Crewes, S. J., Gutteridge, J. M. C., Rowely, D., Halliwell, B., Cornish, A. & Hider, R. C. (1985) *Q. J. Med.* **56**, 345.

Blake. C. C. F., Johnson, L. N., Mair, G. A., North, A. C. T., Phillips, D. C. and Sarma, V. R. (1967) *Proc. Roy. Soc. Lon. Ser. B.* **167**, 378.

Blum, H., Beier H. & Gross, H. J. (1987) *Electrophoresis* **8**, 93.

Boguta, G. & Dancewics, A. M. (1983) *Int. J. Radiat. Biol.* **43**, 249.

Brown, J. R. (1977) In *Albumin Structure, Biosynthesis, Function* Eds. Peters, T. and Sjoholm, I. Permagon press, New York, **50**, 1.

Capeillere-Blandin, C., Delaveau, T. & Descamps-Latscha, B. *Biochem.* (1991) *J.* **277**, 175.

Davies, K. J. A. & Delsignore, M. E. (1987) *J. Biol. Chem.* **262**, 9908.

Davies, K. J. A. & Goldberg, A. L. (1987) *J. Biol. Chem.* **262**, 8220.

Davies, K. J. A. (1987) *J. Biol. Chem.* **262**, 9895.

Davies, K. J. A., Delsignore, M. E. & Lin, S. W. (1987a) *J. Biol. Chem.* **262**, 9902.

Davies, K. J. A., Lin, S. W. & Pacifici, R. E. (1987b) *J. Biol. Chem.* **262**, 9914.

Davies, S. C., Hungarford, J. L., Arden, G. B., Marcus, R. E., Miller, M. H. & Huchins, E. R. (1983) *Lancet* **1**, 181.

Dean, R. T., Fu, S., Stocker, R. & Davies, M. J. (1997) *Biochem J.* **324**, 1.

Dobbin, P. S. & Hilder, R. C. (1991) *Chem. Brit.* 565.

Fling, S. P. & Gregerson, D. S. (1988) *Anal. Biochem.* **155**, 83.

Franzini, E., Sellak, H., Hakim, J. & Pasquier, C. (1993) *Biochim. Biophy. Acta* **1203**, 11.

Freedman, M. H., Boyden, M., Taylor, M. & Skarf, B. (1988)*Toxicology* **49**, 283.

Garrison, W. M. (1987) *Chem. Rev.* **87**, 381.

Graf, E., Mahoney, J. R., Bryant, R. G. & Eaton, J. W. (1984) *J. Biol. Chem.* **259**, 3620.

Grankvist, K. & Marklund, S. L. (1983) *Life Sci.* **33**, 2535.

Halliwell, B. & Gutteridge, J. M. C. (1986) *Arch. Biochem. Biophys.* **246**, 501.

Halliwell, B. & Gutteridge, J. M. C. (1989) Free Radicals in Biology and Medicine, Clavendon, Oxford

Hegg, E. L. & Burstyn, J. N. (1995) *J. Am. Chem. Soc.* **117**, 7015.

Joshi, R. R. & Ganesh, K. N. (1992a) *Biochem. Biophys. Res. Commun.* **182**, 588.

Joshi, R. R. & Ganesh, K. N. (1992b) *FEBS Lett.* **313**, 303.

Joshi, R. R. & Ganesh, K. N. (1994) *Biochim. Biophys. Acta.* **1201**, 454.

Kruck et al., (1993) *Adv. Biosci.* **87**, 257.

LeGendre, N., Mansfield, M., Weiss, A. & Matsudaira, P. In *A practical guide to protein and peptide purification for microsequencing.* Matsudaira P, ed. 2nd ed. Academic press, San Diego 71 (1993).

Martin, D. *Chem. Brit.* (1991) 689.

Marx, G. & Chevion, M. (1985) *Biochem. J.* **236**, 397.

Mordente, A., Meucci, E., Miggiano, G. A. D. & Martorana, G. E. (1990) *Arch. Biochem. Biophys.* **277**, 234.

Nieuhaus, A. W. (1981) *N. Eng. J. Med.* **304**, 170

Nozaki, Y. (1972) *Meth. Enzy.* **26**, 48.

Oelrichs, B. A., Kratzing, C. C. & Winzor, D. J. (1984) *J. Vitam. Nutr. Res.* **54**, 61.

Olivieri, N. F., Buncic, J. R., Chew, E., Gallant, T., Harrison, R. V., Keenan, N., Logan, W., Mitchel, D., Ricci, G., Skarf, B., Taylor, M. & Freedman, M. H. (1986) *N. Engl. J. Med.* **314**, 869.

Osherfoff, M. R., Schaich, K. M., Drew, R. T. & Borg, D. C. (1985) *J. Free Radicals Biol. Med.* **1**, 71.

Phelps, R. A., Neet, K. E., Lynn, L. T. & Putman, F. W. (1961) *J. Biol. Chem.* **236**, 96.

Puchala, M. & Schuessler, H. (1993) *Int. J. Radat. Biol.* **64**, 149.

Rana, T. M. & Meares, C. F. (1992) *J. Am. Chem. Soc.* **112**, 2457.

Shugar D. (1952) *Biochim. Biophys. Acta.* **8**, 302.

Stadman, E. R. (1992) *Science* **257**, 1220.

Stadman, E. R. (1993) *Ann. Rev. Biochem.* **62**, 797.

Uchida, K., Enomoto, N., Itakura, K. & Kawakishi, S. (1989) *Agri. Biol. Chem.* **53**, 3285

Udenfriend, S., Clark, C. T., Axelrod, J. & Bordie, B. B. (1953) *Federation Proc.* **11**, 731.