USE OF LUMINESCENCE ENERGY TRANSFER PROBES TO DETECT GENETIC VARIANTS.

Carlos Vaccaro, B.S.

Thesis Prepared for the Degree of

MASTER OF SCIENCE

UNIVERSITY OF NORTH TEXAS

August 2004

APPROVED:

Douglas Root, Major Professor
Rebecca Dickstein, Committee Member
David Tam, Committee Member
Arthur Goven, Chair of the Department of Biological sciences
Sandra L. Terrell, Dean of the Robert B. Toulouse School of Graduate Studies

Vaccaro, Carlos, Use of luminescence energy transfer probes to detect genetic variants. Master of Science (Biochemistry), August 2004, 53 pp., 6 tables, 15 illustrations, references, 36 titles.

The purpose of this research was to study the hybridization of molecular beacons under different conditions and designs. Data collected suggest that the inconsistency found in the emission intensity of several of these probes may be caused by 3 important factors: length of the probe, nucleotide sequence and, the formation of an alternative complex structure such as a dimer. Of all three factors, dimer formation is the most troublesome, since it reduces the emission of the reporter molecules. A new probe design was used to reduce dimer formation. The emission signal of the improved probe was several folds stronger than those probes with the early design.

In this research, dimer formation is detected, furthermore a new probe with a different design was tested. If dimer formation can be reduced molecular beacons can be integrated into more complex hybridization systems providing an important tool in research and diagnosis of genetic disorders.

ACKNOWLEDGEMENTS

My mentor, Dr. Douglas Root, and Zhiling Zhang, for their contributions on hybridization proximity assays. This project is supported by Biosynthesis, Inc., and NIH.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF ILLUSTRATIONS

Page

CHAPTER 1

INTRODUCTION

Nucleic acids in the form of DNA or RNA are complex macromolecules with at least two important and unique properties necessary to sustain life: the capability to direct their own replication for the conservation of the genetic material and also, the ability to direct the synthesis of proteins; the functional units of all living organisms. The required information for both of these functions is mostly present in the linear arrangement of the molecules that compose the DNA or RNA polynucleotide chain.

The precise knowledge of the human DNA genomic sequence is therefore a fundamental piece of information to establish the relationship between genes and diseases. Modifications in the DNA primary sequences as a consequence of mutations can result in numerous cases in detrimental health conditions. The implications of this important fact have led to the creation and improvement of many molecular biology techniques designed to analyze the presence of mutated sequences present in the genome. A core issue is the need to create highly sensitive devices and protocols for mutation detection that can be used in a clinical setting. This research project described here tests the ability of a new set of luminescent probes to detect not only small mutations such as insertions or deletions, but also much larger genetic variants, such as those produced by alternative splicing or other genetic shuffling mechanisms. This research study also describes the formation of complex tertiary structures by these nucleotides, a problem that under certain conditions tends to limit the usefulness of these probes. Several of these experimental techniques use in one way or another the principle of molecular hybridization (Hopman et al., 1995). Molecular hybridization is the

formation of the double helix from two complementary strands. This natural process can be suited to study particular DNA sequences. First described by Murmur and Dogy in 1961 (Southern, 1985), molecular hybridization helped to establish several important factors about the nature of the DNA present in large genome, such as the existence of several types of repetitive and non-repetitive sequences along the overall DNA structure. Molecular hybridization is also the underlying process that gave rise to important studies on DNA, such as the kinetics of its restitution, a complex DNA behavior that is known and described by the C-value paradox and the Cot ½ which is related directly to the amount of DNA in the genome. In short, molecular hybridization is an excellent tool to measure the extent of hybridization between single stranded nucleic acids, and as a result the degree of complementary of the strands can be inferred (Lewin, 1997).

The main purpose of this research is to determine the advantages and disadvantages of luminescent probes referred here as molecular beacons, reporter probes or fluorescent probes to detect genetic variants. To achieve this goal several molecular beacons with different structural designs were tested against synthetic nucleotide sequences. The idea is to collect important information about what experimental factors can alter, decrease or increase, both the specificity of the probes for its target sequence and the emission of the fluorescent probes. Both of these parameters, hybridization specificity and the intensity of the emission signal, are essential conditions that determine the proper detection of the probes.

Experimental data suggest that the specificity of hybridization is dependent mostly on the length of DNA and other identified factors, such as overall base content,

ionic strength and temperature of the system (Britten, 1985). Several types of mutations can arise on DNA. They encompass both bulky changes such as duplication or rearrangement of segments of chromosomes (Lewin, 1997), and point mutations that include mutations that can add or delete a base pair, shifting the reading frame. While genetic variants may include large sections of DNA, more numerous although less identified mutations involving single base pairs have recently drawn renewed attention.

The single nucleotide polymorphism (SNP) is a change in the DNA produced by a single nucleotide (Lewin, 1997). The natural variation at the single nucleotide level determines genetic variations between individual humans. They occur mostly in non-coding fragments of DNA and are present at a high frequency in any normal human DNA. In fact, the presence and uniqueness of DNA polymorphisms makes possible several analyses ranging from forensic sciences to paternity testing and organ transplant.

The typical scenario of the effect of single base mutations is exemplified in sickle cell anemia, a recessive condition resulting from the exchange of a negatively charged glutamic acid by a hydrophobic amino acid, valine, in a critical site of the protein hemoglobin, at position six of the β-globin chain (Hart, 1998). The codons for glutamic acids are either GAA or GAG. A missense mutation can introduce the base uracil in the second position of the codon, thereby, changing GAG to GUA which codes for the amino acid valine. This minute amino acid exchange causes protein aggregation during the deoxy state with devastating results for the proper delivery of oxygen to the tissues.

Initial analysis of the recently drafted human DNA sequence estimates that our genome may contain up to 1.49 million single amino acid substitutions (Francis et al.,

2001) and, perhaps only half the total genes initially predicted a few years ago, a somewhat unexpected figure that adds more complexity to our present outlook of genomic organization (Pollard, 2001). The emerging picture is that although most of the SNPs are clustered on the non-coding section of DNA, known as hot spots. The remaining fraction indeed may contribute to significantoperational changes in certain genes. Point mutations introduced over long periods of time may contribute to gene loss or gene creation through natural selection. As more of the correct human genome is unveiled, the list of disorders associated with SNPs is likely to increase rapidly (Geminis-Sanchez, 2001). Because of the key importance of these genetics variations, probes such as molecular beacons may be useful tools for research.

To test the ability of these probes to hybridize in solution with synthetic targets a DNA region of critical importance was selected: The human leukocyte antigen (HLA) region. The selection of this DNA sequence responds to the fact that HLA regions are deeply implicated in the recognition of self versus non-self antigens. This process plays a critical role in the rejection or acceptance of grafted tissue such as organ transplant. If molecular beacons can hybridize to key sequences of the HLA portions they can also be used to infer small genetic variations.

The synthetic molecular beacon contained sequences similar to the human leukocyte antigen (HLA) coding sequences. The HLA complex is a cluster of closely associated linked genes, each with several different alleles. The complex is located on the short arm of chromosome six. The HLA system is responsible for specific immunological responses of each individual (Owen, 1998).

Three classes of molecules (I, II, and III) are encoded in the HLA complex. All class I and II gene products have related overall structure. At present there are no established functions between class III genes and the class I or class II gene products. Class II molecules are similar to class I molecules in their final conformation. The gene organization of HLA includes several clusters of similar genes named DR, DQ and DP (Owen, 1996). These genes are the DR gene clusters and consist of a single α gene (DRA) and nine β genes (DRB1-DRB9). The proteins from any combination of these genes contain a hypervariable region located in the T cell receptor binding cleft. Presentation of antigen to leukocytes type T by class II polypeptides is critical for immune responses (Hutchinson, 1996). Organ transplant rejections can be reduced by the matching of donor and recipient HLA molecules, especially for HLA-DR class II molecules (Bjorkman et al., 1987). The highly polymorphic HLA-DR cluster of genes can also be used as a marker site for clinical studies, including paternity testing, disease prevention, and tissue transplant.

The luminescent probes synthesized for this research were designed to be complementary to a segment of the HLA-DR1 gene. While the detection of HLA target sequences was entirely performed in heterogeneous solutions molecular beacons can also be used to study tissues by in situ hybridization.

The final part of the research concentrated on the in situ hybridization of troponin T (TnT) mRNA using molecular beacons. Troponin T is a fundamental motor protein (Mader, 1997) found in all three types of muscle tissue, skeletal, smooth and cardiac. Current research on a less known series of clinical conditions generally known as familial hypertropic cardiomyophathy (FHC) have narrowed down the probable cause to

SNPs or other minor genetic variants localized in key elements of the human heart contractile protein complex (Watkins et al., 1995). Clinical observations of this group of autosomal dominant cardiac diseases include enlargement of the heart, arrhythmia, asymmetrical interventricular and left ventricular hypertrophy, myocellular disarray, chronic weakening of the cardiac muscle tissue and in a number of cases sudden heart arrest during exhausting physical activity in young athletes.

The high level of complexity of the motor protein system in cardiac muscle makes our present understanding of it deficient. Heart muscle tissue contains a large number of motor proteins that cooperate intimately with one another and with other regulatory molecules such as ATP and calcium. It is this multiple protein involvement that makes force generation and its modulation possible (Winegrad, 1992).

The first identified cause of FHC was a missense mutation in the myosin heavy chain gene, resulting in an Arg403Gln mutation, and a striking defect in the function of the purified myosin (Geistefer et al., 1990). Recent data (Roopnarine and Leinwand, 1998) indicated that all mutations found to trigger FHC are proteins of the sarcomere and FHC is caused by mutations in at least seven different genes including:

β-myosin heavy chain
troponin T
α-tropomyosin
myosin-binding protein C
ventricular light chain 1 (VLC1)
ventricular light chain 2 (VLC2)
troponin I

The myosin mutations are known to impair function (Rayment et al., 1995), but any functional consequences of the recently recognized HCM-causing mutations in

cardiac troponin T (TnT) remain unknown. This research project concentrated on the design and testing of reporter probes to detect genetic variants of the protein troponin T in rabbit skeletal muscle as a preliminary model before testing on cardiac muscle tissue. The presence of troponin T and tropomyosin on muscle tissue regulates the hydrolysis of ATP by blocking the interactions of myosin heads with actin. (Mader, 2000). Mutations associated with FHC are mostly concentrated at the N–terminal region of TnT (Sweeney et al., 1994). The study of troponin T is complicated because normal expression of TnT involves complex mechanisms of alternative splicing resulting in several modified isoforms of the protein.

Potentially any amino acid substitution introduced into the primary sequence of TnT can produce a conformational modification. The lack of the appropriate three-dimensional structure changes the steric properties of the protein and ultimately results in a deviation from its normal function; the degree of this change can in some cases be manifested not only at the molecular but at the tissue or organ level as well. In the specific case of troponin T, several mutations at the N–terminal seem to affect by 50% the sliding speed in vitro motility assay of actin sliding over myosin. (Sweeney et al., 1994). The effects of many other known or possible mutations on troponin T still are under investigation. Molecular beacons were constructed to detect genetic variations of important sites on troponin T.

An alternative to the more traditional biochemical methods is to investigate patterns of similarity or changes at the DNA or RNA molecular level. A step forward in this direction was undertaken by the human genome project, a remarkable research undertaking that seeks to sequence the entire human euchromatic DNA, the gene rich

region of the genome. The unveiling of the human genome may eventually have profound consequences in the way we study, prevent and fight many forms of diseases in the near future (Bork, 2001). Genomes hold the inventory of genes, and for this same reason it sets a framework of study for all processes that support life. The drafting and proper analysis of the human genome will benefit research on cancer, gene regulation, gene expression, immunology, signal transduction, and many other current biological puzzles (Baltimore, 2001).

Experimental difficulties to detect SNPs encourages researchers to seek out more sensitive assays. Denaturation and hybridization experiments led to the discovery of several different groups of DNA sequences in the mammalian genome, denaturation and hybridization can be measured by changes in the optical density (OD_{260}) (Zubay, 1996).

In the case of strands of DNA or RNA, hybridization and denaturation are dependent on a number of factors, namely the strand length, base composition, chemical surrounding and DNA or RNA concentration (Dornberger et al., 1997). Long strands of nucleic acids have a large number of hydrogen bonds and therefore they require more energy to separate. In this study the typical length of the synthetic oligonucleotide ranges from 22 to 28 base pairs. A high ratio of GC base content is more difficult to separate and the presence of divalent cations, such as Mg^{2+} ($MgCl_2$) stabilizes the double strand (Tyagi, 1996).

Hybridization between complementary strands is specific; even, slight nucleotide differences between complementary strands greatly diminish the bonding affinity for each other (Lewin, 1997). Therefore, the natural basis for the molecular hybridization of

nucleic acids can be fully exploited to detect large or minute genetic variations on DNA. DNA sequencing offers the ultimate physical map. Sequencing gels have the greatest resolving power since single nucleotide resolution can be obtained, but it is time consuming. Faster automated sequencing methods based on capillary gels are now available (Amhersham Biotech, 2001).

The main drawback characteristic of all these protocols is the present difficulty of their implementation in a large-scale setting such as those required in a clinical environment. In the past, most of these hybridization procedures used radio-nucleotide probes, the main disadvantages of using radioisotopes include not only the important aspect of health and disposal hazard, but also proper personnel training, laboratory certification and regulation compliance. The short half-life of a number of usable isotopes also makes valuable measurements difficult. For all these reasons, the search for perfection of non-isotopic methods derived in a wide range of non-isotopic experimental procedures that are gradually replacing the conventional methods based on iodine-125 or phosphorus-32. These newer assays are based mostly upon two important factors: 1) The overall simplification improved the sensitivity and analytical handiness of the protocol. 2) The new trend in non-isotopic detection is the use of fluorescent probes. Nowadays with the advent of newer and improved spectrophotometers, the limits of luminescence resolution have been increased, and, under certain experimental conditions, even single fluorophores can be successfully detected. Future variants of non-isotopic detection techniques will probably include multi-analyte systems with the capability of a large number of simultaneous assays with a minimal volume of sample, such microarrays systems can contain hundreds of

thousands of probes on a synthetic surface. Microarrays with such capabilities can monitor expression or any genetic change that occur at the DNA or RNA molecular level.

At present the latest technology of these multi-microarrays systems have successfully been designed to detect the presence of specific genes (Amersham Biosciences, 2001). Still, genes are, in most cases, a large polymer of nucleotides, and they can be more readily detected than single nucleotide polymorphisms. Improvements on SNP detection technology may be the logical subsequent step to a technological breakthrough that will incorporate the work of microchip arrays with SNP detection capabilities or perhaps developing tiling technologies (Oleykowski et al., 1998). These technologies may offer a wide range of possibilities not only in research but also in clinical diagnoses (Shoemaker et al., 2001). With microarray systems, containing thousands, or perhaps millions of probes in a small volume, several biological aspects can be identified at once. It may be possible to detect the presence of minute particles of foreign DNA, such as in infectious viral or bacterial diseases with enough precision to even identify the species of the invading microorganism (Lewin, 1999).

Most of the modern non-isotopic technologies are based on the use of fluorescent probes. The use of luminescent probes eliminates most of the hazards associated with the use of radionucleotides. At present, thousands of fluorophores or fluorophore derivatives are commercially available. Fluorescein, rhodamine, coumarin, and bodipy, to name a few, are among the most commonly used luminescent compounds. Minor modifications on the fluorescent molecules can be made to render them suitable for attachment to any specific functional group. At this time many

experimental protocols are intended for single nucleotide polymorphism detection. Usually the probe is another nucleic acid, complementary to the target sequence and contains one or up to several reporter molecules directly or indirectly coupled to it.

Recently, novel sets of fluorescent probes directly linked to short oligonucleotide sequences have been designed (Tyagi, 1996). These reporter probes are constructed to take advantage of two biophysical processes: The natural molecular hybridization of complementary nucleic acid strands and, the transfer of energy between certain chromophores through complex processes such as resonance energy transfer. These luminescent probes, referred to here as molecular beacons, reporter probes, or simply as fluorescent hairpin probes, were designed to analyze nucleic acid hybridization. Because the efficiency of the luminescence emission is a process dependent on the distance between the chromophores, information on several useful parameters can be directly obtained by using these luminescent reporter probes. Resonance energy transfer has been effectively used in several studies, it provides distance information between two sites. Energy transfer can, under proper conditions, effectively resolve distances in the angstrom and nanometer scale.

Molecular beacons are synthetically prepared short single strands of oligonucleotides. Three definite sections compose a typical molecular beacon: the neck region, the inner loop region and the terminal ends of the probe (Fig 1). The neck region is the section that auto-hybridizes. The loop region is the inner sequence of bases that hybridize with the target sequence of interest and, finally, two chromophores are attached at each terminal end or close to the ends of the oligonucleotide.

The key and novel feature of these luminescent probes is their capability to form two different conformations. Because the neck region contains a short sequence of bases perfectly complementary to each other the molecular beacon assumes a closed conformation, similar to a hairpin structure. The strong tendency of the bases to form hydrogen bonds results in a physical closing of the entire strand into a closed structure. Unless the target sequence is present in the solution the open conformation is not thermodynamically feasible due to the base complementarity of the neck region. Fig. 1, shows the two conformations of molecular beacon/hairpin structures. Fig. 1A, shows the closed probe. Observe the base pairing of the short neck region. The loop region is usually larger and complements to the target sequence either RNA or DNA. The presence of the target sequence shifts the thermodynamics of the system to the open conformation, (Fig. 1B). Special care is taken at the time of probe design to favor the natural closing of the probes in solution by selecting the proper bases and the length of the neck region.

Hypothetically, a single mismatch event is sufficient to restrain the probe opening. Previous studies on nucleic acids have identified the dependency of hybridization on external conditions such as sequence length (Bloomers et al., 1998), ionic strength, buffer concentrations and, temperature. Because of the need to create more sensitive and accurate assays, several promising techniques are being currently explored, some of them implement molecular beacons or probes very similar in design. At present several aspects might limit the proper integration of these probes into more complex systems, such as microarrays. Therefore, it is critical to collect preliminary data concerning the actual hybridization behavior of the probe. In other words, what are the

optimal hybridization conditions for such probes and, what are the main drawbacks. The main hypothesis addressed on this research is that it is possible for dimer formation to quench the emission signal, therefore rendering some molecular beacon probes unusable. If this hypothesis is correct then suitable probes can be designed by constructing probes with a nucleotide sequence content that minimizes the formation of this complex dimer structures. Specifically by making either side of the hairpin complementary to the target sequence. Adding an extra complementary section on either side of the hairpin to make hybridization contact more favorable for assays using the principle of energy transfer. Structural studies indicate that several conformations can be adopted by nucleic acids (Ramsing et al., 1989). The structural variety is mostly modulated by the nucleotide sequence composing the polynucleotide chain. To detect single nucleotide polymorphisms such as insertions or deletions, two probes must be used in combination. Since the efficiency of energy transfer is dependent mainly in the distance between probe and acceptor, the information can be used to determine the presence of an extra base (insertion) or, the absence of a base (deletion). Experiments on energy transfer have been previously successful in determining such perturbations of distances between nucleotides (Root, 2000).

My hypothesis is that the formation of dimers can reduce the efficiency of fluorescence emission on this type of probe. This research studies the behavior of the probe under dimer formation, and furthermore shows at the molecular level why the emission is affected. The validity of the hypothesis is tested by designing a nucleotide probe with a reduced thermodynamic tendency to form dimeric structures. The

information gathered by this research, we believe, will lead to the refinement of multi-array assays for early detection of diseases or more complex studies of genomes.

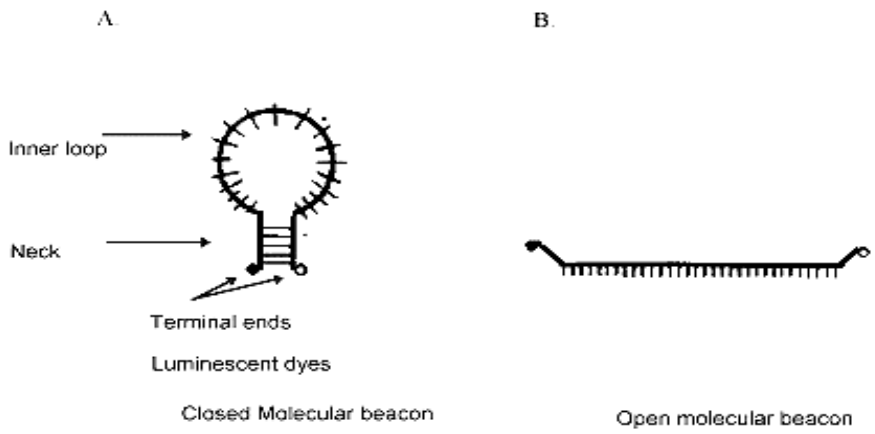

Fig. 1 Closed and open conformations of molecular beacons/hairpin structures.

A. The neck region contains a short sequence of nucleotides perfectly complementary to the other end of the oligonucleotide. Hydrogen bond formation keeps the hairpin in a closed conformation. Close proximity between the luminescent dyes (donor and acceptor) quenches the emission of the terbium chelate (the donor).

B. In the presence of a target sequence, complementary to the inner loop of the molecular beacon, the hairpin structure adopts an open conformation, increasing the distance between the two luminescent dyes attached to it. The increased distance between the luminescent dyes increases the emission intensity of the terbium-chelate.

CHAPTER 2

MATERIALS AND METHODS

Molecular beacons were synthesized at Biosynthesis, Inc. The single stranded nucleic acids composing the molecular beacons were prepared by using an automated phosphoramidite protocol on a DNA synthesizer. In most cases a single end of the probe was labeled with a fluorescent compound, either fluorescein isothiocyanate (FITC) or tetramethyl-rhodamine isothiocyanate (TRITC), during the oligonucleotide synthesis process. Both are fluorescent compounds having different excitation and emission properties (Table 1). The complementary synthetic oligonucleotide target sequences (Table 2) were also synthesized by Biosynthesis, Inc

Preparation of Terbium Chelate

The intrinsic luminescence emission of the lanthanide terbium is relatively weak due to low absorption. To enhance the emission intensity, the lanthanide was bound to a chelate compound with a strongly absorbing ligand attached. By complexing the lanthanide with a chelate, several thousand fold emission intensity increases can be achieved at some excitation wavelengths. The chromophore attached to the chelate, acts as a donor picking up energy in the ultraviolet range and transferring it to the chelated lanthanide element terbium. Two different chelate systems were tested during this research to enhance the emission signal. One contained a cytosine ligand and the other contained a carbostyril-124 compound. The choice of which chelate to use depends on the excitation wavelength selected to excite the compound (Table 1). The cytosine chelate was prepared by making a 120 μM solution of DTPA (diethylenetriaminepentaacetate) and cytosine, both compounds were mixed in a final

volume of 400 µl of the organic solvent solution, DMSO (dimethyl sulfoxide). During the molecular beacon automated synthesis process, a free aliphatic amino or thiol functional group was introduced into the single stranded oligonucleotide to attach the terbium chelate by means of an anhydride reaction to a primary amine or a maleimide reaction with a thiol. A fluorescein derivative, fluorescein isothiocyanate (FITC) is added to the opposite end.

The cytosine chelate used for binding terbium was produced by using an organic solvent DMSO and equimolar concentrations of cytosine and DTPA andhydride that contains a total of five carboxyl groups. Three of the carboxyl ends are linked to the tri-valent lanthanide terbium (Tb^{3+}), one end attaches to the amine group of cytosine and the remaining carboxyl end attaches to the single stranded oligonucleotide via an amino and/or thiol functional group. A 2-aminoethylmaleimide linked the anhydride to the thiol group when present.

Table 1 Spectral Properties of Luminescent Compounds

Reactive Chromophores	Excitation Peak	Emission Peak
FITC	490nm	520nm
TRITC	541nm	572nm
Cy5	650nm	670nm
Terbium-cytosine chelate	280nm	547nm
Terbium-carbistyril-124	340nm	547nm

Molar Extinction Coefficients for Acceptor (M^{-1} cm^{-1})	
$FITC_{(abs\ 494)}$	77000
$TRITC_{(abs\ 544)}$	84000
$Cy5_{(abs\ 678)}$	250000

Table 2. Oligonucleotides Sequences of Molecular Beacons and Target Sequences

Probes designed for HLA typing experiments

Serial number	Nucleotide Sequence 5'-3'
F468-1	CCT GCX TCC CCC ACG TCG CTG TCG AAY GCA GG X=FITC, Y=Tb
A026-3 A977-3	FITC-TTC GAC AGC TCC CCC ACG TCG CTG TCG AA-Tb

Target sequences for HLA typing experiments

Serial number	Nucleotide Sequence 3'-5'
A026-5, A977	TTC GAC AGC GAC GTG GGG GA
F468-3	G TAC GTG CGC TTC GAC AGC GAC GTG GGG GAG TAC

Probes designed for in situ hybridization experiments

Serial number	Nucleotide Sequence 5'-3'
A977-8	Tb-AGG CTC TAT TTT CCA GCG CCC GAG AGC CT-FITC

Target sequence for in situ hybridization

mRNA contained in 20 μm section of rat skeletal muscle tissue (striated muscle). Tb= terbium chelate, FITC= fluorescein isothiocyanate, TRITC= tetramethyl rhodamine isothiocyanate, CY5= cyanide5. Accession number for human NM_000364. Accession number for rat NM_012676

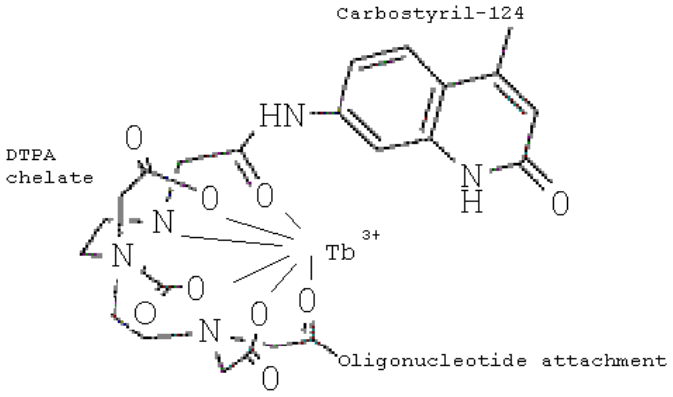

Fig. 2 Chemical structure of carbostyril-124-terbium DTPA chelate. DTPA contains a total of five carboxyl groups, three are covalently linked to the lanthanide element terbium (Tb^{3+}). One attaches to the carbostyril-124 while the remaining carboxyl end acts as a linker to one end of the oligonuecleotide via amino or thiol functional group.

Oligonucleotide Preparation Outline

Drying by centrifugation

↓

Determination of unlabelled oligonucleotide concentration

↓

Terbium-chelate synthesis

↓

Terbium-chelate attachment to oligonucleotide

↓

Isolation and purification of labeled oligonucleotides

↓

Determination of labeled oligonucleotide concentration

↓

Gel shift assay

↓

Hybridization Assay

Fig. 3 Outline of the preparation protocol for labeling and testing of molecular beacons.

As seen in figure 3, the first step is the removal of the buffer by drying with centrifugation. Most of the aqueous buffer is evaporated during this process. Absorbance measurements at 260 nm determine the concentration of oligonucleotides,

the extinction coefficients of each individual sample must be obtained from the manufacturer, since they vary depending on the nucleotide length.

Terbium is attached to the EDTA and later attached to the oligonucleotide in a non-aqueous medium such as DMSO (dimethyl sulfoxide). The labeled oligonucleotides are isolated by size exclusion chromatography using a sephadex G-25 column. Absorbance measurements at 260 nm (oligonucleotides) and 492 nm (FITC) determine the ratio of labeled oligonucleotides. The hybridization of molecular beacons with the target sequences are tested in a 1% acrylamide gel and in aqueous solutions (homogeneous solution).

Determination of Unlabelled Oligonucleotide Concentration

Prior to the labeling of the synthetic oligonucleotides, their initial concentration must be determined. By analyzing the ratio of labeled versus unlabelled oligonucleotide fractions one can roughly determine how efficiently the labeling reaction proceeded. The concentration can be determined by measuring the absorbance at 260 nm and the use of Beer's law:

$$C=A\ 1/e$$

where

C = molar concentration
A = absorbance (260)
1 = pathlength in cm
e = molar extinction coefficient

The aromatic bases of the oligonucleotides (adenine, cytosine, guanine and thymine or uracil) absorb strongly at 260 nm, in addition, the absorbance of a present label, such as FITC or other chromophore can also be used to determine more precisely the final oligonucleotide concentration. The molar extinction coefficient was determined

by the manufacturer. Each oligonucleotide batch has a unique extinction coefficient, since each sample is different from one another in base and length composition.

Labeling of Molecular Beacons

Most of the beacons prepared contained a free amino or a sulfhydril group at one single end of the probe. The functional group was introduced at the time of synthesis with the purpose to have a functional group where the luminescent terbium complex can be efficiently linked. Oligonucleotides were manufactured either as a solid powder or as a solution. If they were in dry-form, 250 µg of the oligonucleotides were mixed with 100 µl of the cytosine-DTPA-terbium solution and the reaction proceeded for 3 hours. If oligonucleotides were in liquid form, they can be mixed directly to the same solution. Another more efficient alternative included letting the oligonucleotide solution to slowly evaporate by the use of a centrifugal rotor, and then using the remaining solid pellet.

Purification of Oligonucleotides

Several products are present after the labeling reaction is completed:

- labeled oligonucleotides
- unlabeled oligonucleotides
- free chelate
- free terbium
- any other by-products of the phosphoramidite reaction
- salts from buffers
- free FITC or any other label used during the oligonucleotide synthesis process.

The isolation of the molecular beacons was carried out by size-exclusion chromatography, using a Sephadex G-25 (Amersham Biosciences) column. The beads of this column are made by cross-linking dextran. The fractionation range or molecular weight range is 1000-5000 for globular proteins. The average diameter of the bead is 20

to 50 μm, the bed volume is approximately 4-6 mL per gram. The sample was run in a FPLC (fast protein liquid chromatography) composed of a 25 mL column. The sample containing the sample was injected by using a 100 μL plastic loop in a running buffer composed of 2mM $MgCl_2$, 10 mM imidazole at pH 7.0. To isolate the fractions containing most of the labeled oligonucleotides, overlapping chromatographs of absorbance and phosphorescence were compared. Only the fractions containing the highest peak of labeled oligonucleotides were used for the remaining experiments. The flow was set at 0.5 mL per minute and the fraction collector was set at 0.5 mL per fraction.

The chromatograph in fig 4 shows the elution of the different compounds present in the reaction vial. The numbers on the x-coordinate represents the fraction numbers while the y-coordinate represents the absorbance at 254 nm. The first peak is the elution of the labeled oligonucleotides; because they have the highest molecular weight of all compounds present in the reaction vial they elute first. The small differences in molecular weight between the labeled and unlabeled oligonucleotides can not be resolved by this type of chromatography. As a result the fractions collected may contain a small fraction of unlabeled oligonucleotides. Overlapping chromatographs of absorbance and phosphorescence were compared (Fig. 5). The intensities are not set to scale.

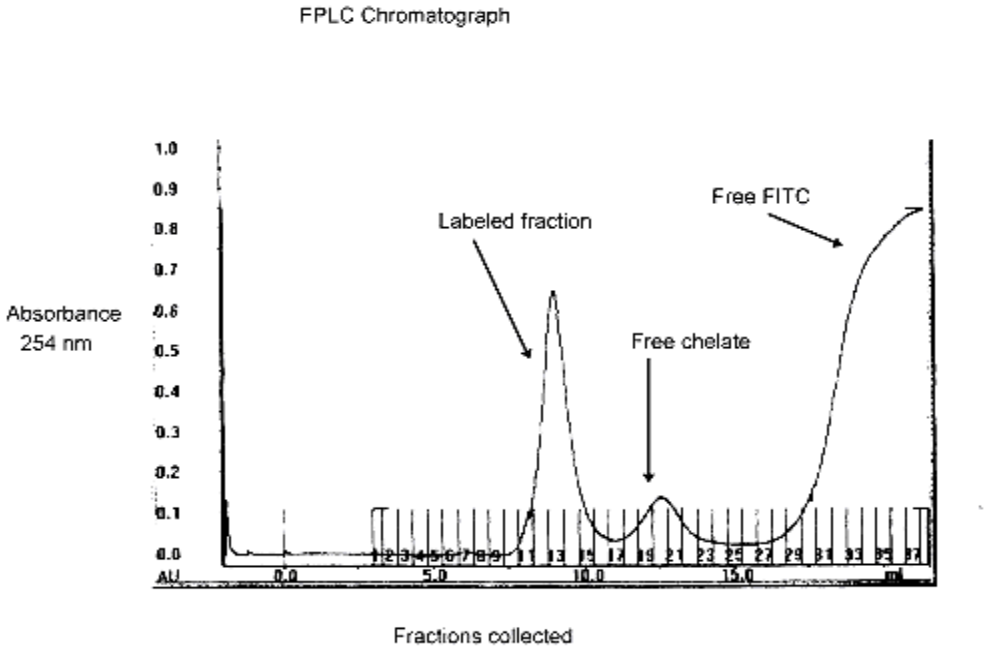

Fig. 4 Typical chromatograph of labeled oligonucleotide purification by size-exclusion. Peaks of absorbance at 254 nm show the elution of the different substances present in the column. Elution from this type of column is a function of the molecular weight of the substance. Since labeled oligonucleotides have the highest molecular weight they elute first.

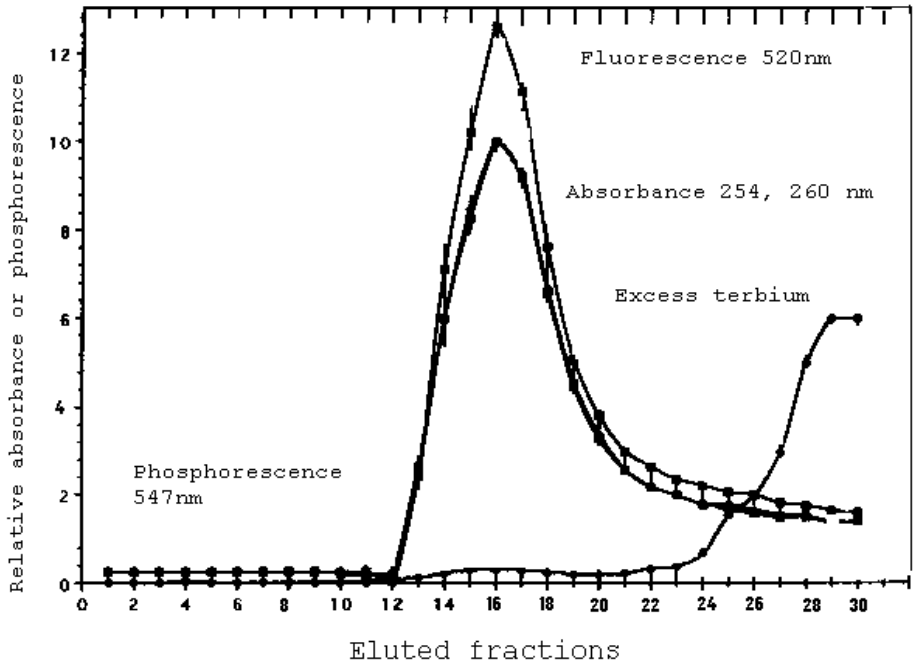

Fig. 5 Overlapping measurements of absorbance at 254 nm, 260 nm (bases of oligonucleotides), fluorescence at 520 nm (FITC), and phosphorescence (terbium chelate) at 547 nm. To identify the fractions containing the labeled oligonucleotides absorbance and fluorescence measurements are taken. The fractions containing the highest absorbance at 260 nm should overlap with a similar peak of fluorescence at 520 nm and phosphorescence at 547 nm. The units on the Y-axis are not set to scale.

Principles of Fluorescence

Luminescence is the emission of light from a substance and involves processes that are associated with the removal of energy from an electronically excited molecule. The term luminescence can be further divided into two main categories; fluorescence and, phosphorescence also known as delayed fluorescence. Because of the intrinsic electric field associated with a light wave, an electron or any charged particle placed in the path of the light wave will suffer a force and will be capable of absorbing energy from the electric field of the light wave. If the electron absorbs energy from the electric field it will be moved to an unoccupied orbital and it will be moved from one location into another in the molecule.

The absorption of energy by a molecule and the subsequent promotion of an electron to a higher excited state is called the electronic transition or dipole electronic transition. Only specific light frequencies can be absorbed by molecular electrons. Excited electrons eventually fall back to the ground state and several alternative processes can take place at this stage, typically, emission of light, heat or both. In a fluorophore molecule, emission of visible or ultraviolet light is the predominant mechanism.

Fluorescence and phosphorescence takes place usually in chromophores having a restricted vibrational freedom, such as aromatic compounds. The absorption process is rapid, about 10^{-15}s but, the chain of events that return the excited molecule to its initial ground state is a much slower process, ranging from 10^{-13}s to several seconds. In the case of free terbium, the lanthanide used for my experiments, its lifetime is approximately 0.44 ms and, when attached to the DTPA chelate is even longer

approximating 0.95 ms. Fluorescent emission has several important characteristics: the Stokes' shift, fluorescence lifetime and quantum yield. The energy of the emission is generally less than that of absorption, a phenomenon first observed by Sir G.G. Stokes in 1852 in Cambridge University. Fluorescent emission takes place at lower energies or longer wavelengths due to energy losses mainly to thermal release of energy, interactions with the solvents, excited state reactions and/or energy transfer between nearby molecules.

The fluorescent quantum yield (Q) is the ratio of the number of photons emitted to the number absorbed. The quantum yield is always less than unity because of natural losses. The quantum yield for elemental terbium and terbium attached to the DTPA chelate is shown in Table 3. Lifetime is approximately the average time the molecule spends in the excited state before returning to the ground state. The lifetime of the fluorescent molecule in the absence of any non-radiative process is called the intrinsic or natural lifetime for terbium its intrinsic lifetime is approximately 4.4 ms. Terbium-DTPA has slightly different intrinsic lifetime of about 4.75 ms.

Table 3 Intrinsic Lifetime of Terbium

Sample	Free terbium[a]	DTPA + terbium[a]
Quantum yield	0.10	0.20
Lifetime	0.44 ms	0.95 ms
Intrinsic lifetime	4.4 ms	4.75 ms

[a]Xu and Root, 1999

The intensity of fluorescence can be decreased by several processes (Lakowicz, 1999). This process is called quenching and can take place by: collision of the excited state fluorophore with some other molecule in the solution, attenuation of the absorbed

light, photobleaching, and also energy transfer between molecules. Forming non-fluorescent complexes with other chromophores can also produce fading of the signal. The instrumental settings for fluorescence detection of fluorescein used during my experiments are shown on Table 4.

Table 4 Luminescence spectrophotometer setting for fluorescence detection.

Emission Wavelength		Monochromators Excitation	
Lower limit	505.0 nm	Wavelength	494.0 nm
Upper limit	650.0 nm	Bandpass	4.0 nm
Scan rate	1.00 nm/sec		
Repetitions	1.0		

Time-Resolved Energy Transfer

Luminescent measurements can be put into two main categories (Lakowickz, 1999), steady-state and time resolved. In steady state measurements, the sample is continuously illuminated by the incident beam of light and the intensity of the emission spectrum is directly recorded. Because of the short lifetime of fluorescence, steady state is reached immediately.

Phosphorescence, also known as delayed fluorescence, is the result of a triplet excited state, in which the spin of the excited electron has the same spin orientation as in the electron on the ground state. The transition to the ground state is therefore forbidden resulting in slow emission rates (decay). Phosphorescence lifetimes are much longer than their fluorescent counterparts, usually ranging from microseconds to seconds.

Time-resolved measurements are used to determine lifetimes and provide an excellent way to improve the signal to noise ratio. In time-resolved measurements, the sample is exposed to a short pulse of excitation light by using the specific wavelength that results in greatest emission intensity. The pulse of light is much shorter than the lifetime of the sample. In time-resolved fluorescence a variable time window can be created by setting up the detection system with specific parameters. A graphic description of time-resolved gating can be seen in Fig. 6.

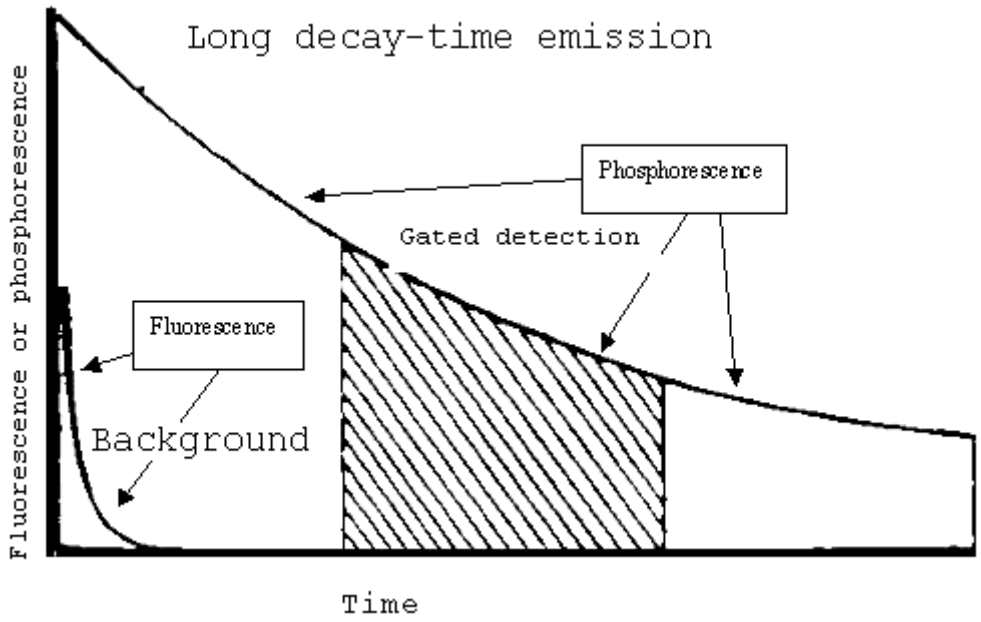

Fig. 6 Time-resolved fluorescence. Representation of fluorescence and delayed fluorescence of luminescent compounds. Time-resolved measurements use a timed-window to discriminate the autofluorescence background. (Modified from Hemmila et al., 1995).

The use of luminescent lanthanides such as terbium or europium allows the implementation of time-resolved measurements. To minimize the intrinsic background of the samples used for this research project, the detection system was preset to begin detecting the emission intensity signal 200 μ-seconds after the pulse of light was triggered. During that interval most of the intrinsic autofluorescence has rapidly decayed

to a negligible point while the lanthanide emission remains relatively close to its maximum and decays at a much slower rate. Settings for time-resolved measurements used for all of my experimental testing are shown in Table 5 and 6.

Table 5 Luminescence spectrophotometer setting for time-resolved phosphorescence detection.

Monochromators Excitation		Monochromators Emission	
Lower limit	200.0 nm	Wavelength	547.0 nm
Upper limit	400.0 nm	Bandpass	16.0 nm
Bandpass	16.0 nm		

Table 6 Luminescence spectrophotometer setting for time-resolved detection (phosphorescence detection)

Flash type	Decay curve
Lower limit	200 µ-seconds
Upper limit	5000 µ-seconds
Resolution	100 µ-seconds

The selection of wavelengths, bandpass, as well as the settings and recording is all electronically controlled by software tailored to the spectrometer specifications, an Aminco-Bowman, series 2 luminescence spectrometer. Proper settings for the time window include emission detection only after 200 µ-seconds following the flash pulse. This setting of the 200 to 4500 µ-seconds window allows ample time for the auto fluorescence to decay to negligible levels before acquiring luminescence; therefore, no significant contribution from the sample can be detected. Most of the signal at this point onward is the result of the terbium chelate emission only.

The data obtained from each single experiment consist of an excitation and emission reading before the sample is mixed with the target sequence. The excitation spectrum determines the peaks of maximum absorption for the molecular beacons. Peaks of strong absorption are analyzed to select the best excitation wavelength. The emission spectrum is independent of the excitation wavelength but the intensity depends on the excitation source. Excitation and emission scans are taken also after the hybridization reaction is completed. The excitation scan identifies natural peaks and valleys of spectral absorbance. The excitation measurement is taken by the setting emission detection at 547 nm, the wavelength of maximum emission of all four main peaks of the terbium chelate. The emission spectrum registers the emission intensity at increasing wavelengths. The emission spectrum is recorded by setting the excitation pulses at 280 or 340 nm depending on the type of chelate used. It is also important to determine the baseline or background emission, which was often done by taking a time resolved measurement of the closed molecular beacons in a 50 μL plastic cuvette without the presence of the target sequence.

Time-resolved measurement of the hybridization reaction proceeds as follows: After 60 or 100 seconds of recording the initial emission of the beacons the target sequence is added. If hybridization occurs an increase above the baseline on the intensity of the terbium emission can be observed. If there is no hybridization the baseline remains unmodified.

Resonance Energy Transfer

An important process that can place between certain chromophores (light energy absorbing compounds) is resonance energy transfer (RET). Migration of energy can

occur when the emission spectrum of a luminescent molecule, called the donor, overlaps significantly with the excitation spectrum of another molecule called the acceptor. There are several additional parameters influencing the efficiency of resonance energy transfer: quantum yield, and intermolecular distance of the probes and molecular orientation of the probe.

In RET the light emitted by the donor is not involved in the transferring of energy, rather, the process takes place because of the dipole-dipole interactions between the two different chromophores. In RET, the energetic exchanges between donor and acceptor are primarily dependent on the distance between them. If the proximity between the probes is extremely close the phosphorescence of the donor chelate complex will be almost quenched because most of the energy collected is immediately transferred to the acceptor molecule. In contrast, if the molecular beacon is in the open conformation the greater distance between the fluorescent dyes contributes to the several fold augmentation of the donor emission intensity. Hence, in more simple terms if hybridization takes place the molecular beacons are in the open conformation and the terbium phosphorescence intensity increases. On the other hand, if the probe is in the closed conformation, indicative of no hybridization, the phosphorescence intensity of terbium remains near the baseline.

The thermodynamics of the structure determines the extent and specificity of hybridization; consequently molecular beacons may be well suited to detect single mismatch base pairing. Optimal conditions for hybridization must be created to increase the specificity of the hybridization reaction. For instance, non-specific hybridization can take place under certain conditions, such as, low temperatures and very large

oligonucleotides sequences. Therefore, in these types of hybridization reactions it is essential that the experimental conditions favor specific hybridizations only. The critical factors playing a dominant role are the length of the loop sequence and the design of the neck region.

Luminescent probes with hairpin/molecular beacon design may have a number of advantages over other conventional fluorescent probes. The most significant improvement is the use of time-resolved measurements. The combination of a conventional chromophore and a long lived phosphorescent compound offers the ability to single out the signal from the reaction while rejecting the fluorescent background noise from the sample. Also, these probes can be used in solution or in tissue samples as well.

CHAPTER 3

RESULTS

The optimal conditions for molecular beacon hybridization were found at low ionic strength, 1-2 mM $MgCl_2$, 10 mM imidazole pH 7.0 (Root, 2001). Optimal condition refers to the maximum emission obtained during hybridization of target and probes sequences. Chaotropic agents such as TMA (tetramethyl ammonium acetate) and SSC were found to inhibit the emission signal at concentration (data not shown).

Two different chelate complexes were tested with the lanthanide terbium to find the one with the strongest intensity signal. One complex used a cysteine-DTPA chelator and the other was composed of carbostyril-124. The central roles of the chelate are to amplify the luminescent intensity of terbium emission and affix the lanthanide on the oligonucleotide.

Molecular beacons constructed for this research were tested under two different experimental conditions: One set of experiments was carried out in solution. Specific concentrations of target and probe were both mixed in a buffer and the fluorescence emission intensity was recorded. Other testing was performed in tissue. The target sequences used for the homogeneous (in solution) part of experiments consisted of sequences similar to those targeted for HLA-DR typing experiments (sequences described in Table 1). Molecular beacons for the in situ experiments were constructed against troponin T mRNA present in a 20 μm section of rat skeletal muscle tissue.

Fig. 7 shows the luminescent increase of the emission at 547 nm after addition of the probe to the target sequence in solution. The phosphorescence intensity was recorded as a function of time (seconds) generally referred to as a time trace

measurement. The molecular beacons (F468-1) and the target sequence (probe sequence A026-5) were mixed in a 50 µL of a buffer consisting of 2 mM $MgCl_2$, 10 mM imidazole at pH 7.0. The concentration of the probe sequence was 7.5 X 10^{-7} M while the target concentration was 2.0x10^{-8} M. The chelate was constructed with DTPA which has an excitation peak at 280-310 nm.

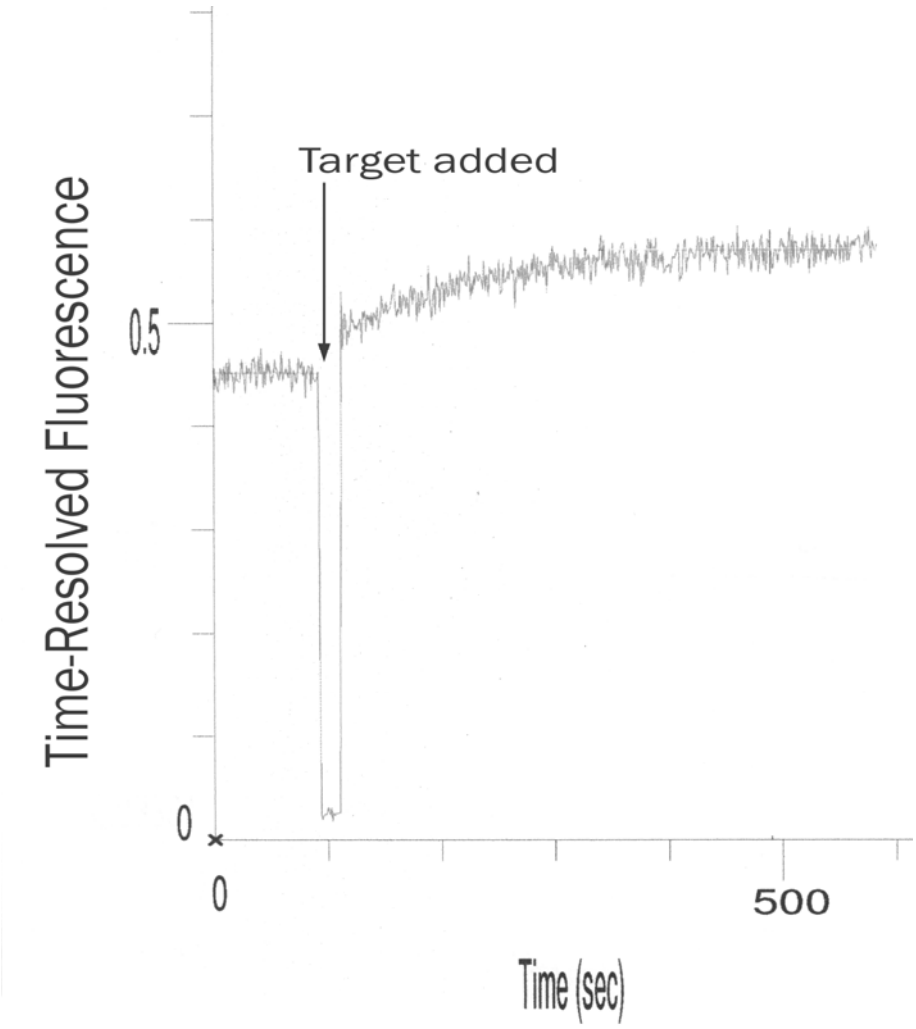

Fig. 7 Time trace of molecular beacons after addition of complementary sequence F468-1 with target sequence F468-3. The increase in emission intensity is due to the separation of the luminescent probes FITC and terbium chelate. The basal emission intensity ranging from times 0 to approximately 80 seconds is before the addition of the target sequence.

The increase in emission after the addition of the probes is due to the opening of the hairpin structure of the beacon. The distance between the acceptor and donor increases as the oligonucleotide probe adopts a helical linear structure. As a result, the terbium is no longer quenched, and its phosphorescence emission intensity is increased several fold. The response shown in Fig. 7 is the typical response of most molecular beacons used for this research. Note also, the baseline level before and after addition of the probe. Approximately a 3.4 fold increase is obtained after 1000 seconds of measurements. Each single experiment carried out consisted of luminescence measurements of excitation, emission and time trace measurements. Emission and excitation scans were performed before and after target-probe were mixed together. The purpose was to monitor the increase or decrease of the emission intensity at certain excitation wavelengths. Time trace measurements describe the kinetics of the reaction probe-target.

Fig. 8 shows the increase of the uncorrected excitation spectrum before and after the addition of probe sequences in a similar experiment and under the same experimental conditions as Fig. 7. Fig. 9 shows the uncorrected emission increase before and after the addition of the same probe. Previous to the addition the probes were mixed with a 50 µL aliquot of a buffer consisting of 2 mM $MgCl_2$, 10 mM imidazole at pH 7.0. The excitation scan reveals sites of weak or strong absorbance of the reporter probes. In fluorescence measurement, the peak of strongest excitation must be used to excite the sample to obtain the maximum emission intensity of the sample. Fig. 8 shows that for terbium-DTPA that maximum absorption is around 280 nm.

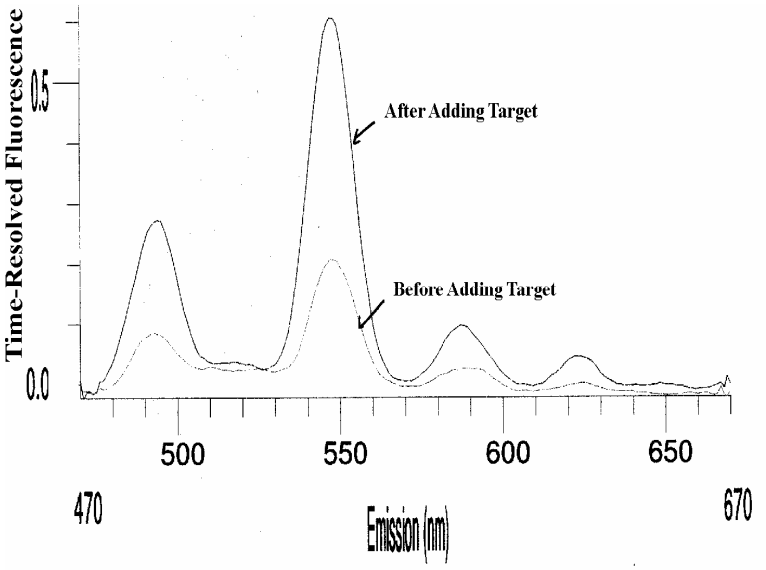

Fig. 8 Excitation spectra of labeled molecular beacon F468-1 before and after hybridization reaction. Excitation scans are produced by recording the emission wavelength at 547 nm while the excitation wavelength is increased from 200 nm to 400 nm. Excitation scans reveals peaks of strong or weak absorbance of a fluorescent compound. Typically, the peak with the strongest absorbance can be used to maximize the emission intensity. The graph shows the excitation spectrum of terbium-cytosine chelate.

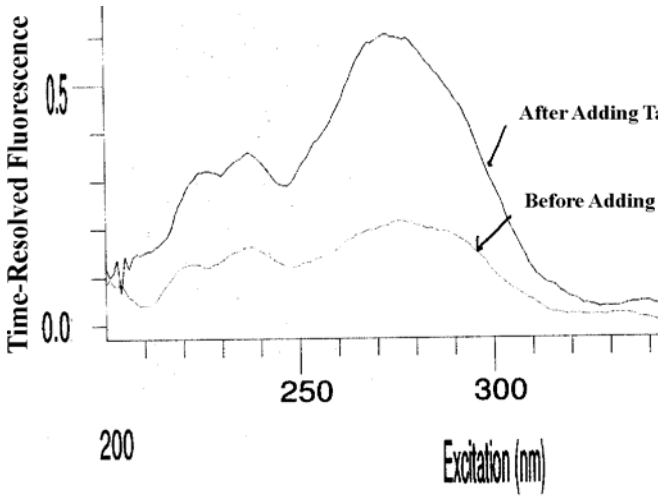

Fig. 9 Emission spectra of labeled molecular beacon F468-1 before and after hybridization reaction. The maximum emission wavelength of the complex terbium chelate is registered at 547 nm. Excitation wavelengths are increased from 400 to 600 nm one wavelength increase per second.

Fig. 10 shows the time trace measurement after the fluorescent probe (A977-8) was added to a rat tissue sample section (in situ experiment). The initial baseline during seconds 0 to 200 indicates the emission registered at 546 nm when the molecular beacon is absent. After the addition of the probe (A977-8) the increase of the phosphorescence indicates that hybridization between the probe and the target sequence is taking place.

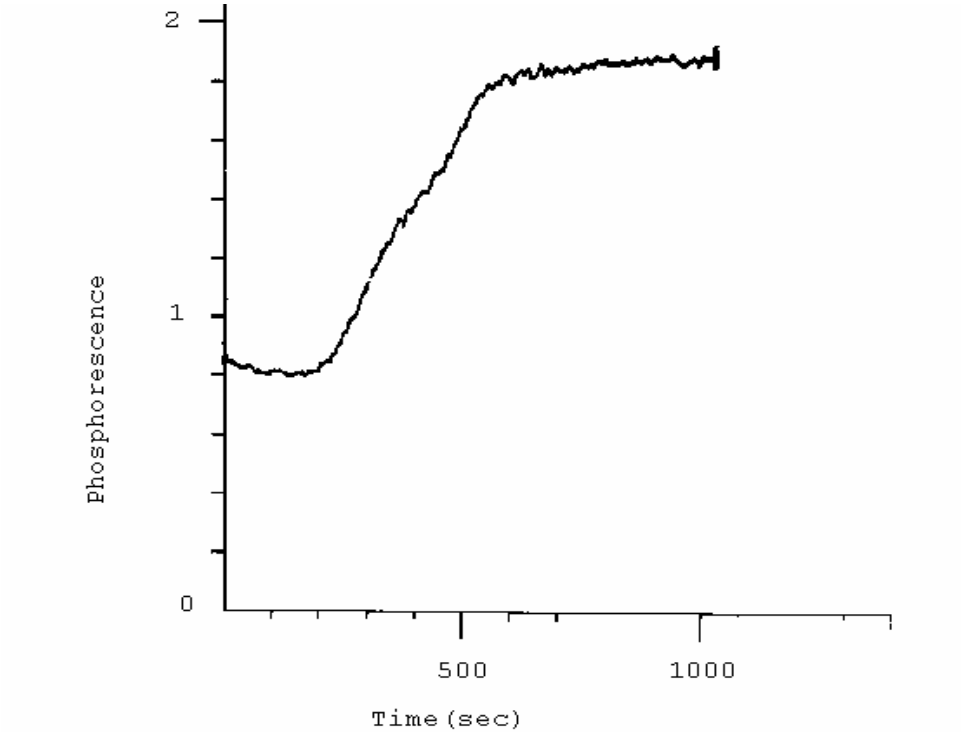

Fig. 10 Time trace measurement of the in situ hybridization reaction with A977-8 on rat skeletal muscle. The increase in intensity is due to the hybridization between the molecular beacon and the target sequence. Settings for the phosphorescent detection are described on Tables 4, 5 and 6. The tissue sample is held in a glass slide introduced into the luminescent spectrometer's cuvette holder. Molecular beacons were labeled with FITC and contained in a 2mM $MgCl_2$, 10 mM imidazole pH 7.0 buffer.

Fig. 11 shows a steady state emission increase at different time intervals. The increase in phosphorescence intensity reaches a plateau after 600 seconds. The shape of the emission spectra of the terbium is independent of the excitation/absorption wavelength used. Although each chelate either cytosine or carbostyril, can be excited

with different wavelengths (cytosine=280 nm, carbostyril-124=340 nm), the emission shape of the terbium chelate remains identical in both scenarios. The key difference is the intensity of the signal collected by the detection system. Fig. 12 shows the detection of dimer formation in a solution containing a mixture of labeled oligonucleotides F468-1 with a fraction of unlabeled oligonucleotides containing exactly the same sequence. The several fold increase in the emission registered by this experiment suggests duplexes of probes were present in most of my experiments and perhaps the variability of the intensity of the signal is connected to the presence of these duplex structures. The higher than normal intensity of emission indicates that in the presence of unlabeled probes containing identical sequences to the molecular beacons, a population of dimer is created because the arms of the probes can hybridize with their unlabeled counterpart therefore minimizing the quenching of terbium.

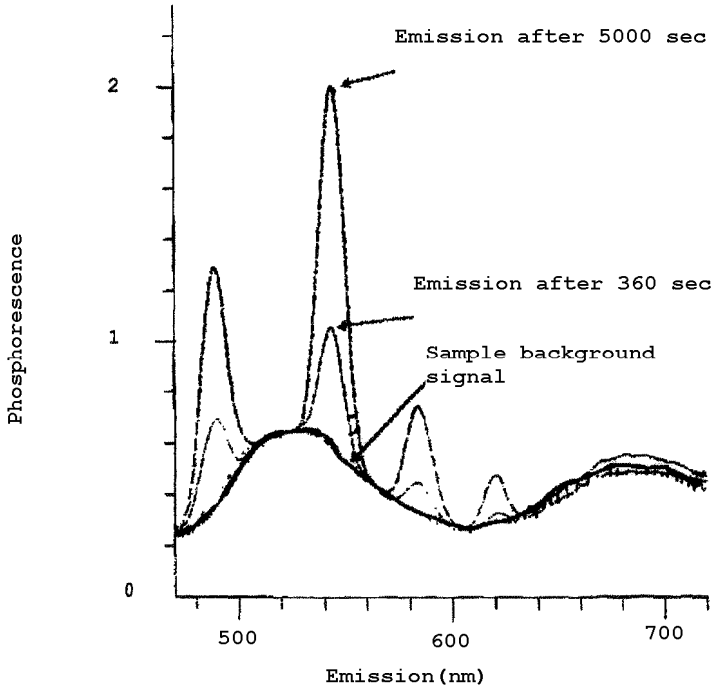

Fig. 11 Overlapping spectrum of A977-8 with rat skeletal muscle at different time intervals: approximately at 0 seconds or baseline, 360 and 5000 seconds.

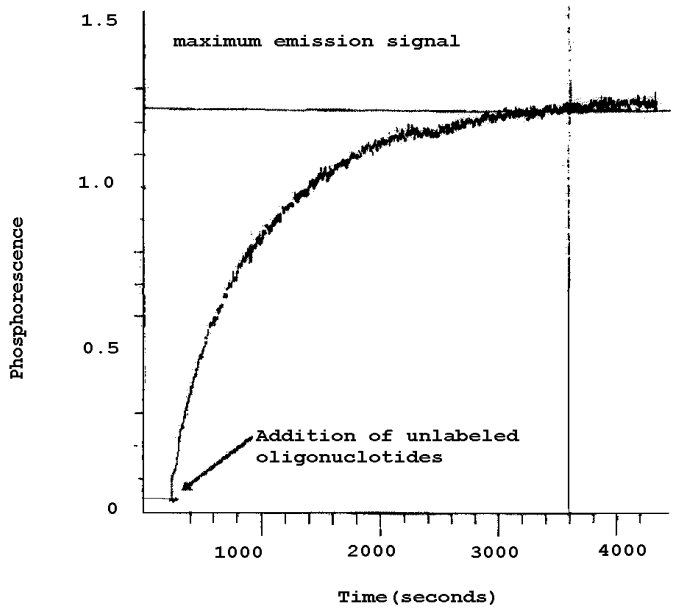

Fig. 12 Reaction of labeled F468-1 with unlabeled F468-1 produces a several fold enhancement of the emission signal. The formation of dimers between labeled molecular beacons reduces the emission intensity due to the proximity between donor and acceptor. In this experiment a sample of unlabeled molecular beacons were added. Since the unlabeled fraction does not contain luminescent probes attached to the ends, the intensity of the emission is not diminished.

When DNA duplexes such as those detected on this specific experiment are formed, the enhancement of the signal is much greater than the intensities observed with other reactions, with exceptions of those improved hairpins where the neck is made to hybridize to the target sequence. There are significant effects of alternative conformations of molecular beacon hairpin structure on the emission intensity. The closed conformation quenches the terbium emission, and there is no enhancement of the signal. Presence of the target sequence opens the probe increasing the distance between the chromophores. The emission of terbium increases several fold in magnitude. Duplex formation by complementarity of the neck region produces a quenching effect due the close proximity of donor-acceptor probes of different strands. In contrast, the presence of unlabeled molecular beacons with the exact sequence will

produce a strong enhancement of the intensity of terbium, because the terbium will no longer be positioned near a quenching chromophore. Therefore, the presence of DNA duplex can be detected by matching labeled hairpins with a concentration of unlabeled oligonucleotides.

After the DNA duplex presence was detected in the experiments, the neck region of the molecular beacons was improved. The key feature was to redesign the neck region to hybridize also with the target sequence. Figs. 13 and 14 show the enhancement of the signal after the probes were re-designed (A026-3). Notice the before and after hybridization emission. The newly improved probe sequence A026-3 gave a 200-fold increase under the described conditions.

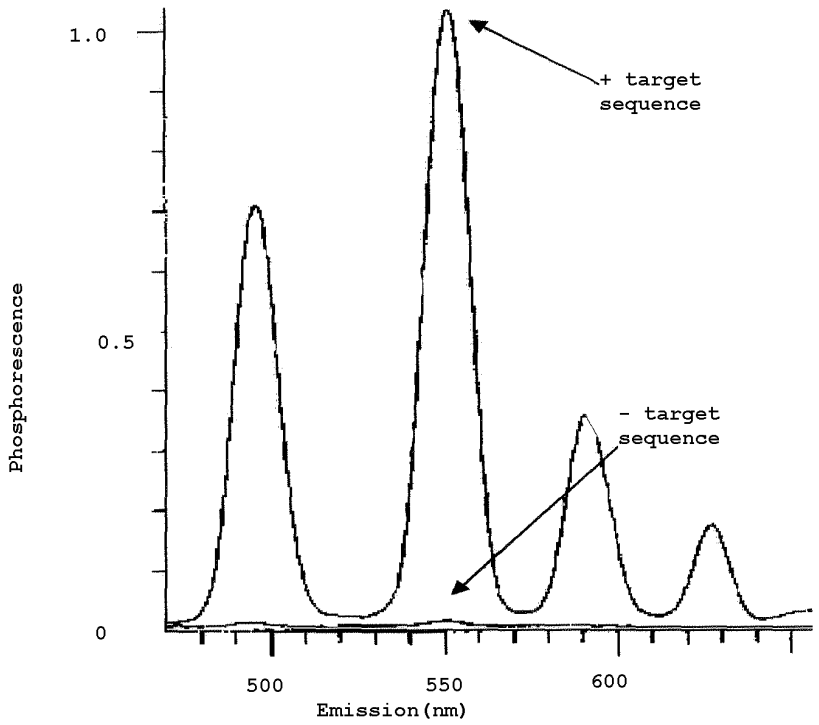

Fig. 13 Improved molecular beacon hybridization reaction, probe A026-3. The neck region has been modified to be part complementary to the target sequence. A several fold enhancement of the signal suggest dimer formation is reduced. Addition of the target sequence increases the intensity of emission several times.

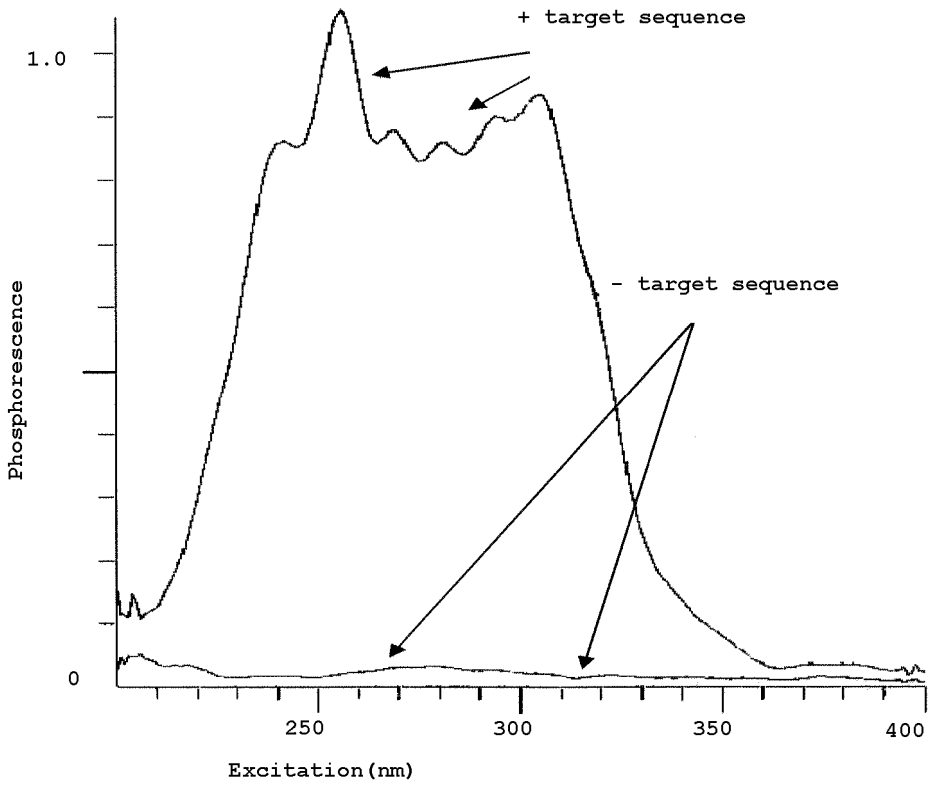

Fig. 14 Excitation scan of improved molecular beacon, probe A026-3. Modification of the neck region reduces duplex formation and therefore the quenching effect is also reduced. Addition of the target sequence increases its excitation intensity.

Another important aspect considered during the experiments was the need to enhance the emission of the chromophores by reducing the background signal. To address the problem, the hairpin probes were modified to accommodate a combination of a lanthanide-chelate complex with a fluorophore, instead of the conventional fluorophore-fluorophore system.

The main idea behind molecular beacons is to produce a reporter probe that can be used to detect single nucleotide polymorphisms, larger types of mutations and also have enough flexibility to be used for instance in microarray technology or a clinical setting. Repeated testing with rat skeletal muscle tissue samples substantiate the high efficiency of molecular beacons to bind to their mRNA counterparts. Parallel to these studies, the probes were tested in a homogeneous medium (in solution) also with

positive results. Data obtained from these experiments confirms the validity of molecular hairpins as reporter probes. The main drawback was that these types of hairpin probes can also form other specialized tertiary structures by dimerization referred to here as DNA duplexes. Duplex formation significantly reduces the emission intensity of the chromophore. A significant result that confirms not only the intricacy of nucleic acids structure, but also limits the capabilities of hairpin probes such as molecular beacons to be integrated into more complex detection systems, such as multi-analyte, micro-array systems.

To minimize the formation of DNA duplex or other structures, an improved molecular beacon was designed (Figure 15).

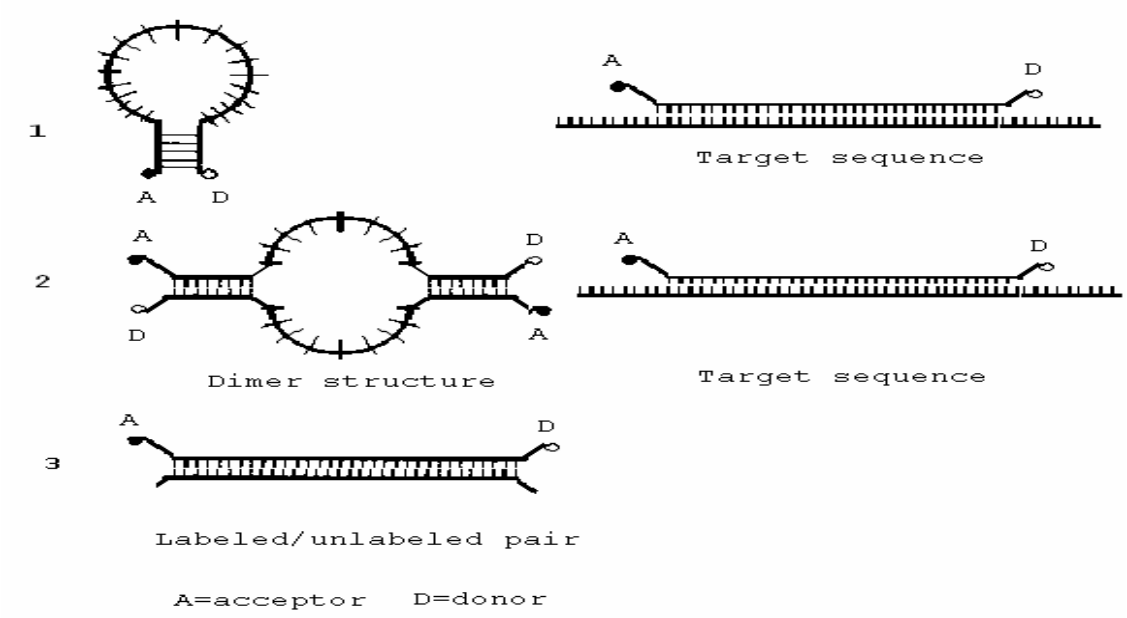

Fig. 15 Alternative conformations of molecular beacons. Closed molecular beacons hybridize with the complementary target sequences. The opening of the probe increases the distances between the terbium chelate and quencher resulting in an increase in the intensity of terbium emission (1). Molecular beacons can also form duplex structures or dimers due to hybridization between the neck regions of identical beacons. In this state the fluorescence of terbium chelate is drastically quenched because of the donor-acceptor proximity to each other (2). Unlabeled molecular beacons identical in sequence to labeled molecular beacons opens the probe and the intensity of the signal is several fold increased due to the absence of the acceptor molecule (3).

The neck region was modified to also base pair with the target sequence of interest. The new improved probe reported a signal enhancement of approximately 200 fold. Probes containing the early construct, with a single exception, showed a poor signal enhancement, averaging 2 to 8 fold increases.

CHAPTER 4

DISCUSSION

Clinical and experimental research data on DNA collected in recent years emphasizes the importance of mutations as the main cause of several diseases. Examples of mutations with adverse phenotypic changes are sickle cell anemia, familial hypertrophic cardiomyopathy and many tumor-related diseases. In order to enhance our current approach to detect mutations, several assays are presently being developed. A new set of luminescent probes, called molecular beacons were first designed by Tyagi and Kramer in 1996 to detect DNA hybridization. These fluorescent probes have a hairpin structure which allows hybridization with complementary target sequences. Along the same line of research, other recent protocols have incorporated similar hairpin structures into more complex studies such as the hybridization proximity assay (HYPA) that uses resonance energy transfer to measure the distances between nucleotide bases. In addition, fluorescent oligonucleotides of similar construct are being incorporated into large-scale assays known generally under the name of microarray systems.

DNA microarray technology has been created as a large scale means to study genomic activity. Large numbers of DNA samples are immobilized on solid surfaces, such as glass. The DNA or the target sequences are tagged with fluorescent labels. The arrays are then scanned to detect the signals produced by the hybridization of the DNA-target complex. Such approach has been successful to study gene expression in prokaryote and eukaryotes as well (Dugan, 1999)

Several hybridization array techniques are currently possible: Sequencing by hybridization arrays (SBH) uses sequence specific hybridization of oligonucleotides to determine DNA structure. In another approach high density polynucleotide probe arrays are being produced in large quantities by light directed synthesis and subsequent addition of a fluorochrome at free hydroxyl ends (Lipshutz et al., 1999). This is having a great impact on molecular biology and clinical diagnostics since large-scale hybridization assays based on array technology are characterizing gene expression patterns (McGall, 2001).

The main purpose of this study was to determine the usefulness of luminescence reporter probes such as molecular beacons to detect the presence of genetic variants along a section of DNA or RNA. In addition, since alternative conformations, such as dimers, can be adopted by these luminescent probes, we propose a simple protocol to detect those higher structural shapes. Because luminescent probes such as molecular beacons or similar hairpin structures are rapidly becoming an integral part of most array detection systems, the implication of having structures of complex orders that may interfere with correct hybridization become extremely relevant.

In this study I detected formation of the mentioned dimers, and furthermore we have re-designed and tested molecular beacons with improved efficiency for hybridization in solution and in situ as well. For the experimental testing of the molecular beacons, two separate sets of experiments were carried out; one designed to hybridize the probes in an homogeneous solution against synthetic target sequences and the other to hybridize the probes in a thin section of sample tissue. These experiments were

performed to validate the usefulness of these probes under conditions that are relevant for clinical diagnosis.

The synthetic target sequences were designed to be similar to the HLA-DR complex sequences. The knowledge of the base sequence of this region of DNA is of critical importance for organ transplant, since they encodes molecules which recognize self versus non-self in an immunological context. If molecular beacons can successfully hybridize to the correct sequence and discriminate other variants with similar base content, it may be possible to incorporate these types of probes into clinical protocols for organ matching. The probes were also tested in situ against a section at the C–terminus of one of the sarcomeric proteins from rat skeletal muscle: troponin T (TnT).

Several factors play important roles during hybridization. Previous studies on nucleic acids have concluded that hybridization reactions of synthetic oligonucleotides are dependent on the length of the oligonucleotide (Southern, 2001), base composition and external agents such as pH, ionic strength and temperature. In these experiments several different variables were tested to find the optimal hybridization conditions. Previous studies have found that large oligonucleotide sequences, for instance, increase the hybridization time, in some cases to the extent of hours. Shorter sequences between 5-10 bp, in contrast, were found to have much faster hybridization rates, on the order of seconds or minutes (Zlatanova, 2001).

Although, the base composition of the inner loop has also a peculiar effect on the hybridization rate, such effects were less clearly identified in our experiments. In previous studies attempts have been made to predict the stability of nucleic acids based on the relevant data from calorimetric studies (Bresslauer, 1986). The thermodynamic

results can be used to predict or characterize the behavior of short duplex sections of DNA from information of its base sequence. In future experiments, we may be able to incorporate this information to predict the stability of the gene-probe complex, and the effect of length and sequence on the final conformation of the reporter oligonucleotide probe

Structural studies on nucleic acids are accumulating evidence regarding the many alternative conformations DNA can achieve based on all those factors previously discussed. For instance guanosine rich deoxynucleotide strands in the presence of MgCl2 can form triple helix structures which seem to regulate cytokine gene expression (Helene, 1992). Other oligonucleotide complex structures include hairpin-dimers, triple helix (Helene, 1999), and four stranded complexes. The structural complexity of nucleic acids emphasizes the need for further detailed characterization of all of these conformational states.

In this study it was attempted to obtain the melting temperature of each batch of fully labeled molecular beacons. Thermal denaturation analysis would have perhaps allowed us help to establish a more clear relationship between base length and composition with the kinetics of hybridization. Calculation of the melting temperature of the single-stranded labeled oligonucleotides by measuring the fluorescence intensity while increasing the temperature was not possible. Further testing on the behavior of terbium-chelate under different temperatures is necessary to reach any conclusive interpretation, but it was proposed that the terbium chelate complex might become unstable above certain temperatures (data not shown) and therefore quenching any fluorescence signal.

The optimal ionic strength, after several experiments was found at a low ion concentration of the divalent cation Mg^{2+}, ($MgCl_2$) in the 1-2 mM range which is consistent with early studies thought to stabilize the neck of hairpin structures.

Molecular beacons as reporter probes have been successfully used in several different types of research such as detection of foreign viral and microbial DNA (Peter Schofield, 1996). In situ hybridization of molecular beacons requires a preliminary DNA search for sequences similar to those present in the target sequence to ensure the molecular beacons are not complementary to any other gene. Also, non-specific hybridization with very similar sequences remains always an open possibility. Caution should be taken to ensure optimal hybridization conditions. Several public domain DNA databases can be searched for sequence matching purposes, such as GenBank or the Protein Data Bank (PDB) which are both available on the Internet.

The approach to detect mis-sense mutation can be significantly different from the approach to detect insertions and deletions. Mis-sense mutations do not necessarily change the distance between the donor and acceptor probes, since it is a replacement of a nucleotide by another. Therefore it can only be detected as a function of the hybridization affinity with the strand. On the other hand, insertions and deletions increase or decrease the distances between the probes proportional to the amount of missing or added bases (Root, 2001).

Reports on single point mutations located in the middle of the target sequence, a mis-sense mutation, shows a decrease up to 80% in the intensity of the signal indicating that molecular beacons can indeed discriminate single nucleotide polymorphisms. Phosphorescence was used as a means to detect the conformational status of the

molecular beacon: closed or open. Recent experiments took a step further and used resonance energy transfer to detect distance information between the probes. The efficiency of energy transfer can be used to detect spatial changes in the angstrom scale. For instance a single insertion will be registered as an increase of approximately 3.6 Å, 2 insertions 7.2 Å, and so on. The opposite is true for missing bases.

The essential pre-assumption made in this model is that molecular beacons and their target sequence will adopt a helical double stranded conformation, as described by the Watson and Crick model of double stranded DNA. It is interesting to notice that some deviations from the typical structures may play a regulatory role in DNA replication and protein expression.

Although these experiments confirm that molecular beacons can be used potentially to detect genetic variants, a significant limiting factor for efficient detection was found: structural dimerization of the reporter probes. Duplex formation causes a significant decrease in the emission signal and complicates the integration of these types of luminescent probes into more complex detection systems, such as those currently under development like DNA arrays.

Previous studies on hairpin probes have predicted the possibility of forming dimers or even more complex structures (Dornberg, 1997). A conclusive evidence of dimer (duplex) formation was obtained in our laboratory when the synthetic molecular beacon was mixed with the unlabeled molecular beacon containing the same sequence. There was a several fold increase in the intensity of emission as a result of the reaction. The reason is that during dimerization between two labeled beacons, the terbium chelate is close enough to the fluorescent signal of the opposite strand and therefore

maximum quenching takes place. Dimerization between a labeled and unlabeled beacon removes this quenching.

Data obtained from this research confirms the presence of these dimers (duplex) structures, furthermore a modified hairpin probe was designed and successfully tested to minimize dimer formation and therefore obtain an enhanced emission intensity signal. In addition, a simple experimental protocol to detect the presence of dimer structures has been developed.

While dimer formation can potentially be reduced by changes in the design of the hairpin/molecular beacon, the fluorescence of most biological samples presents another particular problem: a high background signal. The intrinsic background, commonly referred as background noise is another factor that limits the use of fluorescent probes to detect minor base changes such as insertions and deletions.

To successfully use resonance energy transfer as a means to detect minor distance changes, the sensitivity of the assay had to be improved. In some cases the auto fluorescence signal from the sample is high enough to limit the efficiency of the assay.

Most of the current available protocols use the conventional approach of attaching two fluorescent compounds with short fluorescent lifetimes. The disadvantage of using two such fluorescent dyes is that the auto fluorescence is not significantly eliminated which increases the experimental error. In order to reduce the unwanted background, either in solution or in situ, a lanthanide element such as terbium was tested. The use of terbium in combination with time resolved measurement allowed us

to obtain a more sensitive detection of the signal by an effective reduction of the background.

Terbium, an element with long fluorescence decay, has been previously used in many other types of assays such as the DELFIA assay. In most experiments the element terbium is not directly coupled to the analyte. Our efforts were concentrated to the direct coupling of one end of the molecular beacon with terbium, so critical distance information can be obtained in future experiments.

These results clearly indicate that time-resolved measurements might be effectively incorporated to detect minor genetic changes. The use of terbium permits accurate distance measurements leading to enhanced use of beacons for hybridization proximity assays.

The benefits of using a lanthanide include not only the effective elimination of the fluorescent background signal, but also terbium-chelates unlike many other luminescent compounds used in resonance energy transfer has no fluorescence polarization. This lack of anisotropy is significant, since the orientation factor, also known as κ^2, is a factor that greatly contributes to the efficiency of energy transfer. In the original fluorophore-fluorophore molecular beacon design, the orientation of the donor-acceptor must be experimentally determined before distance information can be achieved. In practice, the orientation factor can contribute to significant sources of errors.

The forces that dictate the three dimensional conformation patterns of oligonucleotides still remain unclear. Natural examples confirming the existence of complex structure of nucleic acids are found in tRNA, ribosomal RNA and hairpin structures found along DNA and RNA (Lewin, 1995).

The labeled oligonucleotides used on this study were tested against synthetic target sequences with identical sequences of the Human leukocyte antigen complex (HLA complex) and, mRNA troponin T for several reasons. The most relevant was to study the suitability of these molecular reporter probes to hybridize to synthetic HLA sequence. The results of these experiments might provide significant information as to whether the probes can be used to detect genetic variants and furthermore whether or not they can comply with the requirements needed in a clinical setting such as those required for organ transplant.

Current testing of luminescent probes include: testing with complementary sequences containing base insertion and deletions. Data on this recent work is not shown on this research and is only properly referenced.

Molecular beacons' ability to detect variants may not only be restricted to SNP. Any other larger variation can be detected, such as mutations involving large sections of chromosomes or genetic variants triggered by tissue specific changes such as alternative splicing. Molecular beacons/hairpin probes based on resonance energy transfer protocols can be used to detect the presence of complementary target sequences, in solution and in situ. Finally, results of the experimental data presented here can be used to design more efficient hybridization proximity assays (HYPA) based on RET. Among the considerations that are critical for such improvements:

1) The design of the molecular beacon/hairpin probe must taken into consideration the formation of tertiary structures including dimers or more complex structures, each different molecular beacons designed must be fully characterized and tested with the complementary target sequence.

2) The combination of a lanthanide-fluorophore to allow time-resolved measurements instead of conventional fluorescence measurements decreased sample background interference.

3) New designs are proposed in which the nucleotide sequence of neck region matches part of the complementary target sequence of interest to minimize dimer formation.

4) Optimal hybridization conditions were found to include $MgCl_2$ as a stabilizer divalent cation and buffer solution with pH close to physiological.

It is expected that hybridization proximity assays that use probes similar to molecular beacons/hairpin can be integrated into more complex systems such as oligonucleotide arrays. At present, the combination cannot be fully exploited, since many of the thermodynamic factors governing small hybridization systems are not completely understood. Further research and testing is critical for the refinement of these newer micro-technologies. This development is not only limited to the probes, fluorescent detection system may also have to be improved to detect even smaller quantities of analyte. Oligonucleotide arrays represent both a challenge, and a next step in molecular biology.

REFERENCES

Aelen, J.M.A., Van der Boom, M. and Hilbers, C.W. (1989) *Biochemistry* 28: 7491.

Baltimore, D. (2001) *Nature* 409: 814.

Bjorkman, P., Saper, M. Samaouri, B., Bennet, W., Strominger, J. and Wiley D. (1987) *Nature* 329: 506.

Bloomers, M.J.J., Walters, J.A.L.I., Hasnoots, C.A.G., Bork, P. and Copley, R. (2001) *Nature* 409: 818.

Breslauer, K., Frank, R., Blocker, H. and Marky, L. (1986) *Proc. Natl. Acad. Sci. USA* 83: 3746.

Britten, R. and Davidson, E. (1991) *Nucleic Acid Hybridization*. Oxford: IRL Press, pp. 3-15.

Dornberger, U., Behlke, J., Birch-Hirschfeld, E. and Frietzsche, H. (1997) *Nucleic Acids Research* 25: 822.

Duggan, D., Bittner, M., Chen, Y., Meltzer, P. and Trent, J. (1999). *Nat. Genetics* 21: 10.

Hart, D and Jones, E. (1998) *Genetics: Principles and Analysis, 4th ed*. Sudbury: Jones and Bartlett, pp. 16-18.

Helene, C., Thong, N. and Harel-Bellan, A. (1992) *Annual NY Academy of Sciences* 660: 27.

Hopman, A. H. N., Speel, E. J., Voorter, C. and Ramaekers, F.C. (1995) *Non-Isotopic Methods in Molecular Biology*. Oxford: IRL Press pp. 1-24.

Hutchinson, I. (1996) *Immunology*, 4th ed. London: Times Mirror International Publishers, pp. 26.1-26.8.

Jordanka, Z. and M. Andrei (2001) *DNA Arrays*. Totowa, New Jersey: Humana Press Inc., pp. 11-13.

Lakowicz, J. (1999) *Principles of Fluorescence Spectroscopy*, 2nd ed. New York: Plenum Publishers, pp. 17-19.

Lewin, B. (2000) *Genes VII*. New York: Oxford University Press Inc., pp. 876-911.

Lewin, S.R., Vesanen, M., Kotrikis L., Hurley A., Zhang L., Ho, D.D. and Markowitz, M. (1999) *Journal of Virology* 73: 6099.

Lipshutz, R., Fodor, S.P.A., Gingeras, T. R. and Lockhart, D.J. (1999) *Nature Genetics* 21: 20.

Mader, S. (2000) *Inquiry into Life*, 9th ed. Boston: McGraw Hill, pp. 385-389.

McGall G. and, Fidanza J. (2001) *DNA Arrays*. Totowa, New Jersey: Humana Press, Inc., pp. 71-99.

Morrison E. (1995), *Nonisotopic Probing, Blotting, and Sequencing*, 2nd ed. San Diego: American Press, Inc., pp. 433-456.

Oleykowsky, C.A., Bronson Mullins, C.R., Godwin, A.K. and Yeung, A.T (1998) *Nucleic Acids Res.* 26: 4597

Owen, M. (1996) *Immunology*, 4th ed. London: Times Mirror international Publishers, pp. 5.1-5.9.

Pollard, T. D. (2001) *Nature* 409: 842.

Ramsing, B., Rippe, K., and Jovin, T. (1989) *Biochemistry* 28: 9528.

Roopnarine, O. and Leinwand, L. A. (1998) *Biophys. J*. 75: 3023.

Root, D. D., Shangguan, X., Xu, J. and McAllister, M. (1999) *Journal of Structural Biology* 127: 22.

Root, D. D. (1997) *Proc. Natl. Acad. Sci. USA* 94: 5685.

Shoemaker, D., Schadt, E., Armour, C., Garret-Engele, Y., McDonagh, P. and Loerch, P. (2001) *Nature* 409: 922.

Southern, E. (1985) *Nucleic Acid Hybridisation*. Oxford: IRL Press, pp. 1-2.

Southern E. (2001) *DNA Arrays*. Totowa, New Jersey: Humana Press, Inc., pp. 11-13.

Sweeney, H. L., Straceski, A. J., Leinwand, L. A., Timo, L. and Antii, I. (1995) *Nonisotopic Probing, Blotting, and Sequencing*, 2nd ed. San Diego: American Press, Inc., pp. 331-376.

Tikunov, A. and Faust, L. (1994) *J. Biol. Chem*. 269: 603.

Tyagi, S. and Kramer, F.R. (1996). *Nature Biotechnol*. 14: 303.

Watkins, H., McKenna, W. J., Thierfelder, L., Suk, H.J., Anan, R., O'Donoghue, A., Spirito, P., Matsumori, A., Moravec, C.S., Seidman, J.G., and Seidman, C.E. (1995) *N. Engl. J. Med*. 332:1058.

Winegrad, S., (1992) *Muscular Contraction*. Cambridge: Cambridge University Press, pp. 107-116.

Zubay, G. L. (1996) Biochemistry, Fourth edition., Dulbuque, Iowa., Wm. C.Publishers Brown., pp. 751.