Norwegian University of Life Sciences

**Master's Thesis 2020 30 ECTS**
Faculty of Chemistry, Biotechnology and Food Science

# The stabilization effect of DNA polymerase on DNA duplexes

Fredrik Johansen
Chemistry and biotechnology

# Acknowledgment

First and foremost, I would like to thank my supervisors, Pranvera Hiseni, Robert Ch. Wilson, and Knut Rudi, for the guidance, advice, and help throughout this period.

A big thank you to my parents, who have been supportive throughout my entire education. I would also like to thank my grandfather for his continuous encouragement, support, and for taking interest in my studies. Last but not least, I would like to thank my girlfriend for her love and support during these times.

Ås, June 2020

Fredrik Johansen

# Preface

Due to the lockdown caused by COVID-19, the experimental part of this thesis has been slightly limited. Several experiments that should have been done were not possible to do but are presented under “technical issues and improvements of the study” in the discussion section of this thesis. The most pressing of these issues is the lack of larger sample sizes in the results presented.

# Abstract

This study is based upon previous observations of DNA polymerase stabilizing DNA duplexes. With the presence of DNA duplexes and DNA polymerases in many experimental methods in modern biotechnology, it is important to gain knowledge about the interactions between them. In order to acquire this knowledge, several DNA duplexes have been investigated. These DNA duplexes were designed to have different structures, in order to understand what factors might influence the stabilization provided by DNA polymerase. These alterations were made in order to gain information that can further test the hypothesis that DNA polymerase can stabilize DNA duplexes. The duplexes used in these experiments had three different sequences, as well as three alterations made to them. Including mismatches positioned in different locations on the sequence, a length difference between the strands, and replacing the 3'-hydroxyl on the recessed primer strand with a 3'-phosphate group. The DNA duplexes were put through a melting curve analysis in a qPCR machine, using EvaGreen® as an intercalating agent, to find the melting temperature (*T*m) of the different probes, with and without active DNA polymerase. The theoretical *T*m was also calculated utilizing an online prediction tool for comparison.

These experiments were performed with two different DNA polymerases, which provided insights into how their stabilization effect may change under different circumstances. In all the experiments, duplexes incubated with FIREpol® showed a slightly higher *T*m than those with HOT TERMIpol®. If the primer strand was shorter than the template strand, a *T*m shift was observed. The duplexes with a mismatch also had a change in the stability of the duplexes. When comparing the observed and predicted *T*ms of the DNA duplex, it was detected that the further away the mismatch was from the 3' end of the recessed nucleotide, the greater the difference between the theoretical *T*m and the observed *T*m. DNA polymerase did not stabilize duplexes where the 3'-hydroxyl group was replaced by a 3'-phosphate group on the recessed primer strand.

The results achieved in this study demonstrates that DNA polymerase can have a stabilizing effect on DNA duplexes.

# Sammendrag

Denne studien er basert på tidligere observasjoner av DNA duplekser som blir stabilisert av en DNA polymerase. Med tilstedeværelsen av både DNA polymerase og DNA duplekser i mange moderne eksperimentale metoder, er det viktig å ha kunnskap om interaksjonene mellom dem. Hvis disse interaksjonene er ukjent, kan det lede til mistolkning av resultater oppnådd av eksperimentet. For å finne denne informasjonen har det blitt gjort eksperimentelle forsøk på ulike DNA duplekser. Ulike duplekser ble designet for å oppnå kunnskap om faktorer som kan ha betydning for den stabiliserende effekten fra DNA polymerase. DNAet som har blitt forsket på har tre ulike sekvenser, i tillegg til tre modifikasjoner. Modifikasjonene var, mismatcher på DNA duplekset, lengdeforskjell mellom DNA trådene, og noen duplekser hadde en 3'-fosfatgruppe i stedet for en 3'-hydroksylgruppe på primer-tråden. DNAet ble analysert i en smeltekurveanalyse i et qPCR instrument, med EvaGreen® som fargestoff. Dette ble gjort for å finne smeltepunktet for de ulike DNA dupleksene, både med, og uten aktiv DNA polymerase. De forventede smeltepunktene ble også kalkulert ved hjelp av et dataprogram.

Eksperimentene ble utført med to ulike DNA polymeraser, for å teste om den stabiliserende effekten var avhengig av typen polymerase. Disse polymerasene var HOT TERMIpol® DNA polymerase og FIREpol® DNA polymerase, og det viste seg at duplekser med FIREpol® hadde litt høyere smeltepunkt enn dupleksene behandlet med HOT TERMIpol®. DNA duplekser med lengdeforskjell viste et stort skift i smeltepunkt. Det viste seg også at mismatcher påvirket DNA polymerase sin stabilisering av duplekset. Om mismatchen var posisjonert langt unna 3' enden på den korteste tråden, økte den stabiliserende effekten, om den var posisjonert nære enden var den stabiliserende effekten lavere. Når DNA duplekset hadde en 3'-fosfatgruppe i stedet for en 3'-hydroksylgruppe ble det ikke observert noe stabiliserende effekt fra DNA polymerase.

Resultatene presentert i denne studien demonstrerer at DNA polymerase kan ha en stabiliserende effekt på DNA duplekser.

# Table of contents

# Abbreviations

| | |
|---|---|
| ASP | Anti-sense primer |
| Cy3 | Cyanine 3 |
| Cy5 | Cyanine 5 |
| DNA | Deoxyribonucleic acid |
| dNTP | Deoxyribonucleotide triphosphate |
| EG | EvaGreen® |
| FAM | Fluorescein amidite |
| LAD | Liquid array diagnostics |
| NN | Nearest neighbor |
| PCR | Polymerase chain reaction |
| qPCR | Quantitative polymerase chain reaction |
| SP | Sense primer |
| TAMRA | Tetramethylrhodamine |
| *T*m | Melting temperature |

# 1. Introduction

The melting temperature (*T*m) of DNA duplexes is an essential subject for investigation. DNA duplexes have a *T*m unique to its properties, such as base-pairing, base stacking, and mismatches (Aboul-Ela et al., 1985; Yakovchuk et al., 2006). These properties will, in turn, affect the stability of the DNA duplex, which will affect the *T*m of the duplex. Other factors that can influence the stability of the DNA duplex are salt concentration, the type of salt, fluorescent dyes, quenchers, and the pH of the solution (Moreira et al., 2005; Tan & Chen, 2006). The ability to predict the true empirical *T*m of designed probes is central in several methods of research when investigating the qualities of DNA duplexes. Probes are used in many techniques that gather information about DNA; melting curve analysis, HRM, genotyping, and amplification of DNA, are all techniques that require probes. With an understanding of all the factors influencing the stability of DNA duplexes, a true empirical prediction of the *T*m could be possible, this thesis aims to investigate if DNA polymerase is one of these factors.

Previous observations made by Hiseni (2016) show that DNA polymerases appeared to impact the stability of DNA duplexes. Hiseni's research described a novel technique, called liquid array diagnostics (LAD), used for quantifying point mutations in mixed DNA populations. The assay is based on combining sequence-specific labeling of DNA probes with the use of high-resolution melting-curve (HRM) detection. In this way, up to 20-point mutations can be quantified simultaneously. LAD uses a single nucleotide primer extension and an HRM to quantify these point mutations. It can accomplish this using only a qPCR instrument. Therefore it could become an easily adopted diagnostic tool (Hiseni et al., 2019). An additional observation in her study was that the DNA polymerase used might have some DNA duplex stabilizing properties.

This thesis aimed at investigating factors affecting the melting temperatures (*T*m) of DNA duplexes in a melting curve analysis, specifically looking at how DNA polymerase can affect the *T*m of the DNA duplexes. DNA polymerase's effect on the stability of DNA duplexes has been the focus of this study, in order to increase the knowledge of DNA duplex stability and to improve the prediction of melting temperatures in probe design. This study investigated how much stabilizing effect DNA polymerase has on different DNA duplexes.

## 1.1. DNA duplexes

A DNA duplex consists of two strands of DNA intertwined in a double helix formation. The study of these duplexes can be used to find information about many qualities the organism possesses. They contain the code for the organism, how it will grow and function. There are many ways to discover and compare this information with other samples of DNA. One of these methods is a melting curve analysis. A melting curve analysis will visualize the temperature needed for the DNA to denature, and one can find a melting temperature. Among other things, this information can be used to discover what nucleotide is present in a particular position in the DNA, as different nucleotides yield different melting temperatures, it can find point mutations in the same organism by looking at the differences in the dissociation curve. This means that the thermostability of the DNA duplex is essential information. Therefore, accurate and precise prediction and measurement methods are needed. With the advancements in modern biotechnology, this is even more important as probes and primers are often used and designed based on the prediction methods available.

### 1.1.1. The thermodynamics and kinetics of DNA duplexes and duplex formation

DNA duplex formation and stability can be quantified by looking at the free energy change from the formation of two single-stranded DNA strands or the denaturation of a duplex into two separate DNA strands. The free energy is referred to as $\Delta G^{\circ}$. Gibbs free energy is calculated by $\Delta G_T^{\circ} = \Delta H - T\Delta S$. If the change in free energy is negative, it is an exergonic reaction; if it is positive, it is an endergonic reaction. This means that if the formation of the DNA duplex has a negative free energy change, the reaction is exergonic and spontaneous, which in turn means that the disassociation of the DNA duplex is an endergonic reaction. This means that in order to break the DNA duplex, one needs to add energy, with the amount of energy needed depending on how much energy gets released in the formation of DNA.

The formation of a DNA duplex from strands that are not self-complementary gives an equilibrium between the single-stranded DNA and the DNA duplex, $X_{ss} + Y_{ss} \leftrightarrow Z_{ds}$, which can remove the ΔG from our equations and the van't Hoff equation can be used $-lnK = \frac{-\Delta H^{\circ}}{RT} - \frac{\Delta S}{R}$ (Brown & Brown Jr, 2015). Using a modified version of the van't Hoff equation, we can calculate the thermodynamic parameters for non-self-complimentary DNA strands. Where ΔS is entropy changes, ΔH is enthalpy changes, $C_T$ is the concentration of DNA, *T*m is the melting point, and R is the gas constant. We can

calculate these in a temperature-independent thermodynamic analysis using the following van't Hoff analysis (Wu et al., 2002).

Equation 1:

$$T_m^{-1} = \frac{R}{\Delta H} ln \frac{C_T}{4} + \frac{\Delta S}{\Delta H}$$

If the formation of the DNA duplex is from strands that are not self-complementary, the $C_T/4$ is replaced by $C_T$, if the strands are in equal concentration. $C_T/4$ is replaced by $(C_A - C_B/2)$ if the strands are not in equal concentration (SantaLucia, 1998).

Equation 2:

$$T_m^{-1} = \frac{R}{\Delta H} lnC_T + \frac{\Delta S}{\Delta H}$$

The thermodynamics of a temperature-dependent analysis is built upon the heat change capacity of the DNA duplexes. Heat change capacity is decided by the positive contribution of capacity from nonpolar groups and the negative contribution from polar groups. When the double-stranded DNA is denatured into coiled single-stranded strands, there will be a difference in hydration, which will result in a difference in heat capacity. This will be decided by the ratio of polar and nonpolar groups in the DNA duplex. As there is a difference between single-stranded and double-stranded DNA's heat capacity, there will be a temperature-dependent difference in both entropy and enthalpy in the formation of DNA duplexes. In order to calculate the differences in enthalpy and entropy based on the temperature, the following equations have been suggested in (Wu et al., 2002):

Equation 3:

$$\Delta H(T_M) = \Delta H° + \int_{T°}^{T} \Delta C_{p,H} dT = \Delta H° + \Delta C_{p,H}(T_m - T°)$$

Equation 4:

$$\Delta S(T_m) = \Delta S° + \int_{T°}^{T} \Delta C_{p,S} d\ lnT = \Delta S° + \Delta C_{p,S} \ln (T_m/T°)$$

Where the entropy changes are ΔS, and the enthalpy changes are ΔH. The heat capacity change is represented by ΔC, with $\Delta C_{p,H}$ being the heat capacity change regarding enthalpy, and $\Delta C_{p,S}$ is the heat change capacity regarding entropy. These equations should yield the same heat capacity change for both enthalpy and entropy, but there is, in fact, a difference if these equations are used. To be able to accurately use these equations, one needs knowledge about the proportions of the differences between the heat capacity changes for enthalpy and entropy. As there are three unknowns in the equation, this is difficult to calculate. Therefore, it has been suggested to use the average of the heat capacity changes to further calculate the free energy, $\Delta C_p^{ave} = (\Delta C_{p,H} + \Delta C_{p,S})/2$.(Rouzina & Bloomfield, 1999). By using the average value of the heat capacity, it is possible to calculate the temperature-dependent free energy.

Equation 5:

$$\Delta G(T_m) = \Delta H^\circ \left(1 - \frac{T_m}{T}^\circ\right) + C_p^{ave}[T_m - T^\circ - T_m \ln (T_m/T^\circ)]$$

The thermodynamics of duplex formation is one of two parts in the formation of the DNA double helixes, and the other part is the kinetics of the formation. As thermodynamics explains the energy cost of forming or denaturing a DNA duplex over a couple of transition states, it calculates the energy for the DNA sequence; however, it does not explain the mechanism of the formation of the duplex. The formation of DNA duplexes starts with the formation of a nucleus of several base pairs, which then gets expanded through a zip-up model (Gu et al., 2007). This first step is rate-determining, which means it limits the speed of the DNA duplex formation. With a high rate, it will be faster, and with a low rate, it will form slower. This step happens when there is formed a nucleus duplex from two to four bases, which is bonded with hydrogen bonds and via base stacking. Further, there is a transition state which will be made of 15-16 base pairs, which allows the zip up to happen and more bases are paired and stacked until the entire DNA duplex is formed. In this zip-up step, it is hard to distinguish every base by base pairing and stacking, which in turn makes us treat this step as a kinetic cluster of bases. (Rauzan et al., 2013)

### 1.1.2. Stability of DNA duplexes

The stability of DNA duplexes is a widely researched subject. This stability is important in both life and in several experimental methods. As the DNA is the information that codes for life, it is of importance that it can remain stable in the organism. Knowledge about this stability is of great interest to scientists, who can use the denaturation of DNA duplexes to obtain knowledge about DNA in general, as well as DNA specific to an organism. The stability of DNA can be measured quantitively by the free energy change from forming a DNA duplex or reversely by denaturing a DNA duplex (Bren et al., 2010). Knowing what factors can influence the stability and having a method to predict under which conditions the DNA will denature, opens a lot of possible methods for learning about a specific DNA duplex. If one applies heat denaturing to a DNA duplex, it will separate into single-stranded DNA and opens for the possibility for probe hybridization, cDNA expression, and genotyping methods (SantaLucia Jr & Hicks, 2004).

The stability of DNA duplexes is influenced by several factors, with the most important being the DNA sequence itself, with hydrogen bonds, base pairing, and base-stacking (Yakovchuk et al., 2006). The DNA will form hydrogen bonds between nucleotides and water surrounding the DNA, the bonds between the nucleotides are often referred to as base pairing.

Base pairing consists of the bases in the DNA binding together in pairs by hydrogen bonds. This pairing is often referred to as the Watson-Crick pairing, in which purines and pyrimidines bond. This is hydrogen bonding of the respective bases, adenine with thymine, and binding between cytosine and guanine, which forms two polymer chains, that are made by phosphate diesters binding to deoxyribose sugars. This allows the Watson-Crick helical structure to form, where the DNA duplex twists into an antiparallel conformation, with ten nucleotides in each turn (Bansal, 2003). In regard to stability, there is a difference depending on the number of A-T pairs and G-C pairs. While adenine and thymine form two hydrogen bonds that connect them, guanine and cytosine have three hydrogen bonds between them. It has been shown that the G-C bonds have a binding energy of -25.4 kcal/mol$^{-1,}$ and the A-T bonds have a binding energy of -12.4 kcal/mol$^{-1}$ (Mo, 2006). This indicates that high G-C content in the sequence will provide a more stable duplex. A linear relationship between G-C content and thermostability exists was previously discovered, and it is possible to extrapolate G-C content by exposing a DNA duplex to heat and observing the *T*m of the sequence, and by knowing the G-C content, one can derive the A-T content (Schildkraut et al., 1962).

Even though the G-C bonds need nearly twice as much energy to separate, the stabilizing effect they contribute to the DNA duplex is it is not the only determining factor in duplex' stability. Another

important factor is the concept of base-stacking. In short, the concept of base-stacking involves the interactions between the neighboring base pairs in the sequence. Base-stacking includes all the interactions between a base-pair and its neighbor and is essentially a factor relating to duplex structure, sequence, and the properties the sequence can have. This includes stability in the DNA duplex (Hunter, 1993). By having favorable bases stacked next to each other, the energy required to break the DNA duplex will increase, and vice versa, if the bases stacked require less energy to denature, the stability of the DNA duplex will decrease (Petersheim & Turner, 1983). Disputing the fact that base-pairing has a considerable effect on the thermostability of the DNA duplex, there have been findings showing that base-stacking is the determining factor in how thermostable the DNA duplex is. However, they also state that base-pairing will affect the stability of the DNA duplex, by A-T pairings being less stable than G-C pairings, but the stability gained by these pairings is minor compared to the stability provided by base-stacking interactions. In addition to the base pair being more stable, G-C pairs also have more favorable opportunities to increase stabilizing through base-stacking with other base pairs (Yakovchuk et al., 2006). It has also been observed that replacing A-T pairs with G-C pairs is not always a net gain in stability. If the A-T has another A-T as a nearby neighbor, the stability at 25 ˚C will be equal to a G-C replacement of one of the A-T pairs, but at higher temperatures, the stability will change in favor of G-C pairings (Aboul-Ela et al., 1985).

In addition to base-pairing and base-stacking, mismatches in the DNA duplex can influence the thermostability of the DNA duplex. It is shown that duplexes with a mismatch in the sequence have less stability than duplexes without mismatches. The least stable pairings being the pairs with cytosine and the most stable pairings containing guanine. And it has been found that mismatches formed with guanine possess the ability to form hydrogen bonds between the nucleotides which influences the stability. As discussed above, the base stacking will have a greater effect on the stability than the base pairings. The mismatches contribute to the destabilization in this order, with the leftmost being the most stable G-T→G-G→G-A→C-T→A-A = T-T →A-C = C-C (Aboul-Ela et al., 1985).

### 1.1.3. External factors affecting the thermal stability of DNA duplexes

There are several known external factors that can affect the thermostability of DNA duplexes. Pressure, salt, pH, intercalating agents can all influence the thermostability of DNA duplexes, increasing or decreasing the *T*m depending on the factors present. While performing experiments on DNA duplexes, one often uses different reagents to be able to gather the wanted information and to

be able to complete the methodology. Therefore, it is important to know about how these reagents can affect the DNA duplex; in this case, the thermal stability of the DNA duplex is in focus.

Salt is a well-known factor in the stability of DNA duplexes. As the nucleic acid molecules are anions, the forming and stability of DNA duplexes need cations to balance out the charges. The backbone of the DNA strands are phosphate groups, which are negatively charged, and to prevent the phosphates from deterring each other there needs to be a cation to balance the charges. This allows the DNA duplex to make the double-helical formation and for the formation to remain stable (Tan & Chen, 2006). This means that the concentration, size, and charge of the cations are essential to the formation and the stability of the DNA duplex. There are two primary theories on how the ions interact with the DNA. The counterion condensation theory, which classifies the ion that accompanies another ion as the counterion. For example, in NaCl, the sodium ion, which is positively charged, would be the counter ion for chloride. This theory assumes that there is a mean distribution of condensed ions along the polyelectrolyte, i.e., DNA, and that the length of the polyelectrolyte is infinite. Then we have the Poisson-Boltzmann theory, which considers all solute atoms as particles with a low dielectric constant with partial point charges. The solvent that interacts with the solute will often have a much higher dielectric constant, and the theory does not consider rearrangements of polar and charged groups with external electric fields. According to the Boltzmann distribution, a single particle has the average effect of the whole system attributed to it (Fogolari et al., 2002).

It has been found that the free energy needed to form DNA duplexes can be decreased and increased depending on the salt concentration in the solution. With a lower salt concentration, the free energy needed for the initiation a DNA duplex formation is increased (Starikov & Nordén, 2009). Tan and Chen (2006) investigated the effects of salt concentration, size, cation valence, and the chain length would have on the stability of the DNA duplex. They quantified the electrostatic contribution of salt on helix' stability by using the electrostatic folding free energy. Using equation 6, where ΔG is the difference in free energy, they found that a higher ion concentration gave a lower amount of free energy.

Equation 6:

$$\Delta G_{37}^{el} = G_{37}^{el}(helix) - G_{37}^{el}(coil)$$

The higher ion concentration also gave a lower amount of free energy for the double-stranded DNA in a helix formation compared to the single-stranded DNA in a coiled formation. The reasoning behind this is that the double-stranded DNA is more negatively charged than ssDNA and thus has a greater

electrostatic effect, which makes it easier for the ions to bind at higher salt concentrations, in turn meaning that a higher salt concentration gives more stable DNA duplexes (Tan & Chen, 2006).

The use of intercalating agents is quite common in modern methods of DNA analysis. The knowledge of how these intercalating agents can affect the *T*m of the DNA duplexes is important. Studies have shown that fluorescent dyes covalently bound to the DNA may have a destabilizing effect on the DNA duplex with as much as 11 ˚C. It was also reported that the type of agent used could have a differing effect on the stability of the DNA duplex, where fluorescein destabilized the duplex more than tetramethylrhodamine (Moreira et al., 2005). Fluorescein destabilizing the DNA duplex also observed by Mineno et al. (1993). Here it was also shown that the concentration of fluorescein used impacted the *T*m of the duplex. A higher fluorescein concentration gave the DNA duplex a lower melting temperature.

A stabilizing effect was also observed, when there were both fluorescein and rhodamine present on the DNA duplex. These were placed respectively on the 5' and 3' end of the oligomers on the opposite strands of each other. When both fluorophores where present, the DNA duplex was stabilized, with a change of -1.5 kcal ΔG˚. This can also indicate that the positioning of the fluorophores might have an impact on the stabilizing/destabilizing effect of intercalating dyes (Morrison & Stols, 1993). There has also been suggested that some intercalating dyes can have a stabilizing effect on the DNA duplex. This has been observed in the use of some oxazole yellow dyes, YOYO and YO-PRO on self-complementary DNA duplexes. It was observed that the *T*m of the DNA duplex increased proportionally with a higher concentration of the dye. There it was suggested that dyes have three different physical effects on the DNA duplex, which made the duplex more stable. The first effect was that the positively charged dye would shelter the negatively charged molecules in the DNA backbone, the phosphates. Then there is the effect of the structure of the dye, which is aromatic rings, would interact with the base stacking mechanism and further increase the DNA duplexes stability. The third effect is believed to be the unwinding of the DNA duplex made by the intercalating agents. The charge density of the DNA backbone would be smaller as the unwinding of DNA makes the DNA longer (Bjorndal & Fygenson, 2002).

The location of a quencher and a fluorophore on a DNA duplex might be a factor in the stabilization of a DNA duplex. The agents are primarily bound to the DNA. However, there might be an attraction between the two agents, which then can further help to stabilize the DNA. This was observed by Marras et al. (2002), where depending on the dye used, the *T*m of the duplexes containing a fluorophore and a quencher directly opposite of each other increased the *T*m to 51-59 ˚C. Whereas

the *T*m observed without any intercalating agents was determined to be 49 ˚C, using the same DNA duplex.

Further expanding on this topic, Moreira et al. (2005) investigated the effects several fluorophores and quenchers have on the *T*m of a DNA duplex. They tested Cyanine 3 (Cy3), Cyanine 5 (Cy5), TET, Tetramethylrhodamine (TAMRA), Texas Red, fluorescein amidite (FAM) and HEX. While the quenchers used were Black hole 2, Black hole 1, QSY 7, and IOWA black FQ. The fluorophores that showed the highest stabilizing effects on their own were Cy3 and Cy5, with an average of 1.6 ˚C, followed by TAMRA and Texas Red with 0.6 ˚C. TET was found to destabilize the duplexes and reducing the *T*m by 0.5 ˚C, while FAM and HEX have no effect on the stability of the DNA duplex. All the quenchers had a positive effect on the *T*m with the greatest effect being from Black hole 2 and Iowa black RQ, which increased the *T*m by up to 2.6 ˚C. When the fluorophores and quenchers were paired, some of them showed an additive stabilizing effect, while some pairs stabilized less than expected and some more than expected. This again can indicate that the effect fluorophores and quencher have on the stability of a DNA duplex is highly individual to the agent used (Moreira et al., 2005).

## 1.2. DNA polymerases

In order to achieve the best accuracy in DNA research, the modern world of biology requires stable and predictable reactions. DNA polymerases that can endure high temperatures are essential to many of these reactions, one of the most important reactions being the polymerase chain reaction (PCR). The discovery of thermostable DNA polymerases allowed for the wide usage of the modern PCR, where the DNA is subjected to major increases and decreases in temperature. The development of these polymerases has greatly increased the efficiency of running PCR, and a wide range of polymerases for all purposes now exists. There has also been reported that the polymerase has further effects other than elongating the DNA in the PCR. There have been some observations of DNA polymerase that has increased the stability of DNA duplexes (Hiseni, 2016).

Polymerases that are stable under heat have long been under development, with one of the first discoveries of thermostable polymerases being Taq polymerase. Taq polymerase was purified from *Thermus aquaticus* and had an optimal temperature of 80C˚(Chien et al., 1976). Later this enzyme could be used to improve the PCR cycle. Previously researchers had to add polymerase to each step of the PCR, but with the discovery of a thermostable DNA polymerase, one could use the polymerase for the entire run (Saiki et al., 1988). Later in the development of PCR and polymerases, there have been advances in the accuracy and clarity of the results. One of the problems with the thermostable

polymerases was that they retained enzymatic activity at temperatures as low as 20 ˚C. This activity can cause problems like primer dimers and priming on the wrong places on the strand, mispriming. One way of combating this would be to add reagents later in the PCR after the initial heating step. However, this would be another source of possible errors, cross-contamination, and it would take more work. By using a Hot-start polymerase, these issues are easier to resolve, as it works by lowering the activity for the enzyme in lower temperatures while keeping the activity normal at 70C˚ and keeping the thermostable capabilities at even higher temperatures. This was discovered using a cold-sensitive mutant of Taq DNA polymerase, and it was observed that 50% of all PCR runs with a high amount of cycles showed an improved yield if a Hot-start polymerase was used (Kermekchiev et al., 2003).

### 1.2.1. DNA Polymerases mechanisms on DNA duplexes

DNA polymerase is crucial for replication and maintaining a genome. Different polymerases have different functions in the genome. They play different roles, synthesizing DNA, repairing the DNA, with different substrate specificity and different mechanisms for interacting with the DNA. There are six different families of DNA; A, B, C, X, Y, and reverse transcriptases. Even though they have different functions, the main structure of DNA polymerases remains largely the same with three subdomains, palm, fingers, and thumb. The palm has a catalytic core that is responsible for the phosphoryl transfer reaction, which again contains three active site motifs, A, B, and C. A and B are responsible for positioning the catalytic aspartates at the active sites on the polymerase, while the B motif is set up against this structure (Brakmann, 2005). The polymerization mechanism by the hand configuration functions, by the fingers binding to the incoming nucleotides and makes the interaction with the single strand template DNA, the palm has the catalytic residues and will also interact with incoming dNTPs. The thumb is responsible for binding with double-stranded DNA. When the DNA polymerase synthesizes new DNA, the enzyme is translocated in order to add new bases to the DNA. This translocation has a rate-limiting step, suggesting that a conformational transition is needed in order to add nucleotides. It is believed that the polymerase needs to form a pol-DNA-dNTP ternary complex following the pol-DNA binary state in order to function properly (Li et al., 1998). Family A polymerases has members such as DNA pol 1 in prokaryotes, pol γ,θ in eukaryotes, and T3, T5, T7 pol in viruses. Family A polymerases mechanism works by the thumb rotating towards the palm of the polymerase, letting the amino acid residues on the thumb connect to the DNA in the minor groove, resulting in an S form of the DNA (Patel & Loeb, 2001). The way these reactions are catalyzed is by a two-metal ion mechanism. This mechanism is used for the phosphoryl transfer reaction, and the mechanism is

suggested to exist in all polymerases (Steitz, 1999). The two metal ions are bound to three carboxylates and are contained in a domain that is not necessarily homologous to other polymerases. The first metal ion interacts with the three-prime end of the primer strand, with the 3'-hydroxyl group, and it is suggested that it can lower the pKa of the hydroxyl, further allowing the reaction with the alpha-phosphate on the new dNTPs. The second metal ion is responsible for binding to and helping in the leaving of beta and gamma phosphates from the reaction. It is also suggested that the metal ions can have a stabilizing effect on the Penta covalent transition state that happens under the binding of DNA to polymerase (Steitz, 1999). The 3-D structure of Taq-polymerase is shown as an example structure in **Figure 1**.

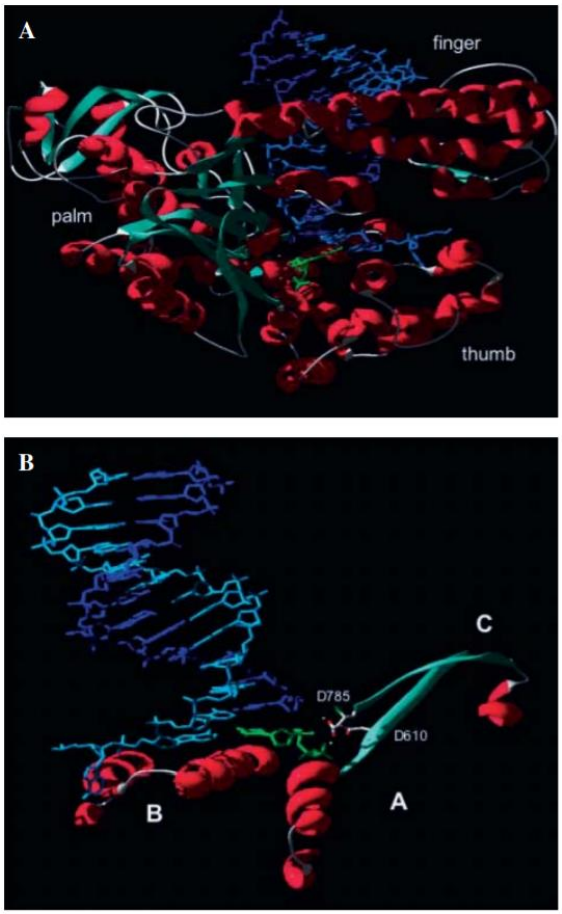



**Figure 1: 3-D structure of Taq-polymerase.** Represented in panel A is the structure of Taq polymerase, which shows the structural configuration of the palm, fingers, and thumb of the polymerase. Panel B shows the A, B, and C motifs, which form a cleft for the incoming nucleotides. Blue colored molecule represents primer-template, while the green molecule represents the incoming nucleoside triphosphate. The Figure was taken from Brakmann (2005)

### 1.2.2. Observed stabilizing effects

DNA duplexes thermostability has been a major focus in this study, and it has been discovered that there might be a stabilizing effect from DNA polymerase on DNA duplexes. This effect was observed by (Hiseni, 2016), where an oligonucleotide duplex, with a sense-primer (SP) of 60 nucleotides and anti-sense primers (ASP) of 16, 20, 30, 40, 50 and 60 nucleotides were analyzed with a high resolution melting curve with two different DNA polymerases, HOT TERMIpol® DNA polymerase, and HOT FIREpol® DNA polymerase. The terminology used when naming the different strands in the DNA duplexes is slightly misleading, as this was not necessarily coding DNA. The experiments run with HOT TERMIpol® DNA polymerase used heparin as an inactivator for the polymerase and 1X buffer C as a substitute in the experiments without an inactivator, with EvaGreen® being used as the dye. The results showed a correlation between the length of the ASP and the stabilizing effect of the polymerase. Where the heparin treated samples that had an ASP with 16 nucleotides showed a 7.5 ˚C higher *T*m when compared to the samples that had 1X buffer C. whereas the ASP containing 20 nucleotides showed a 4 ˚C higher *T*m. The DNA duplex with an ASP of 60 nucleotides showed a *T*m shift of 2 ˚C, which suggests that DNA polymerase can, in fact, stabilize a DNA duplex with varying efficiency depending on the length/variance in length of the DNA duplex (Hiseni, 2016). HOT FIREpol® DNA polymerase also showed some stabilizing effects on the DNA duplex, using SDS as an inactivator. With a shown stabilizing effect of 2-3 ˚C (Hiseni, 2016). This stabilizing effect by DNA polymerase was also anecdotally observed by the PCR-extension of 8-nt-long DNA primers at temperatures far above their theoretical Tm, which strongly indicates a stabilizing effect (Leal et al., 2006).

## 1.3. Theoretical *T*m calculations

To be able to predict the thermostability of DNA duplexes, it is crucial to perform an accurate probe design, which in turn leads to accurate studies. To be able to achieve this, many algorithms and iterations of these algorithms have been developed. There are several factors involved in the thermostability of the DNA duplex, such as the sequence of the DNA, the length of the duplex, salt concentration, what kind of salt used, and if there are mismatches present in the sequence (Aboul-Ela et al., 1985; Tan & Chen, 2006; Yakovchuk et al., 2006).

### 1.3.1. The algorithm used to calculate *T*m

The most common method used for calculating *T*ms is the nearest neighbor (NN) model. This model takes base stacking into account and assumes that the neighbors of the base pairs are a determining factor in DNA duplex stability, by looking at what nucleotides are next to the base-pair and how they are oriented compared to the examined base-pair. By looking at the free energy changes in all possible configurations, one can find the total free energy change in a duplex. Finding the free energy for all combinations of nearest-neighbor interactions has been the focus of many studies. So that one can plug in a sequence, and by looking at the bases and neighboring bases, it is possible to calculate total free energy by looking at table values. One study has tried to unify these studies as best as possible and therefore making a unified view of the nearest neighbor model (SantaLucia, 1998). The equation they used to find the total free energy cost of forming a DNA duplex is based on, the summary of free energy used in forming all the NN-pairs, if there is a terminal A-T or G-C pairing an initial parameter is included in the equation as they require different amount of energy. There is also included an additional parameter, which differs if the strands are self-complementary or non-self-complementary (SantaLucia, 1998). In the formula $n_i$ is the number of times the NN-pair occurs in the sequence, $\Delta G^{\circ}(i)$ is the free energy for the ten possible NN-pairs (in example A-T/T-A), while $\Delta G^{\circ}(sym)$ is the adjustment for non-self-complementary or self-complementary, which is either 0 for non-self-complementary or +0.43Kcal/mol if self-complementary (SantaLucia, 1998). Shown in equation 7, and an example of the calculation is found in **Figure 2**.

Equation 7:

$$\Delta G^{\circ}(total) = \sum_i n_i \Delta G^{\circ}(i) + \Delta G^{\circ}(init\ w/term\ G*C) + \Delta G^{\circ}(init\ w/term\ A*T) + \Delta G^{\circ}(sym)$$

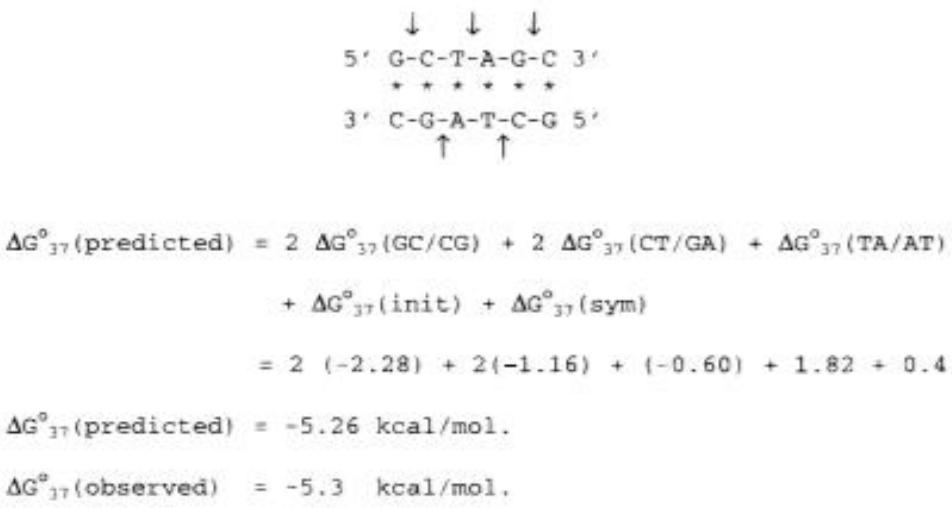



**Figure 2: Example of calculation of the free energy in DNA duplex formation.** The figure shows a calculation using the sequences, GCTAGC/CGATCG, with both predicted and observed ΔG˚ present. Showing the accuracy possible to achieve using the "equation 7". This Figure demonstrates the usage of the NN-model at 37 ˚C and with 1M of NaCl (SantaLucia et al., 1996).

With the total free energy available calculated, one can move on to calculate the predicted *T*m of the DNA duplex. Using equations 1 and 2, depending on the DNA duplex is self-complementary or not. Then there is the effect the salt concentrations will have on the stability of the DNA duplex. There is research suggesting that the sequence of the DNA duplex will not influence the stabilizing properties of salt in the solution. However, the length of the sequence will (Record & Lohman, 1978). The combined effects of the salt concentration and length of the sequence need to be addressed to get an accurate prediction of free energy change if the experiment is performed with different parameters on the NN-model. Therefore, the following formula was developed:

Equation 8:

$$\Delta G°_{37}(oligomer, [Na^+]) = \Delta G°_{37}(Unified\ oligomer, 1M\ NaCl) \pm (stddev\ of\ slope) * N * ln[Na^+]$$

Where $\Delta G°_{37}(Unified\ oligomer, 1M\ NaCl)$ is the predicted free energy change at 1M NaCl, the stddev of the slope is the standard deviation of the slope between N*ln[$Na^+$] and the free energy change of the nearest neighbors, with the intercept through zero. N is the number of phosphates in the DNA duplex divided by 2 (SantaLucia, 1998). In the study from SantaLucia (1998), they used the data from 26 different studies and duplexes and unified them to create a combined agreement on what the free energy change would be for the duplexes.

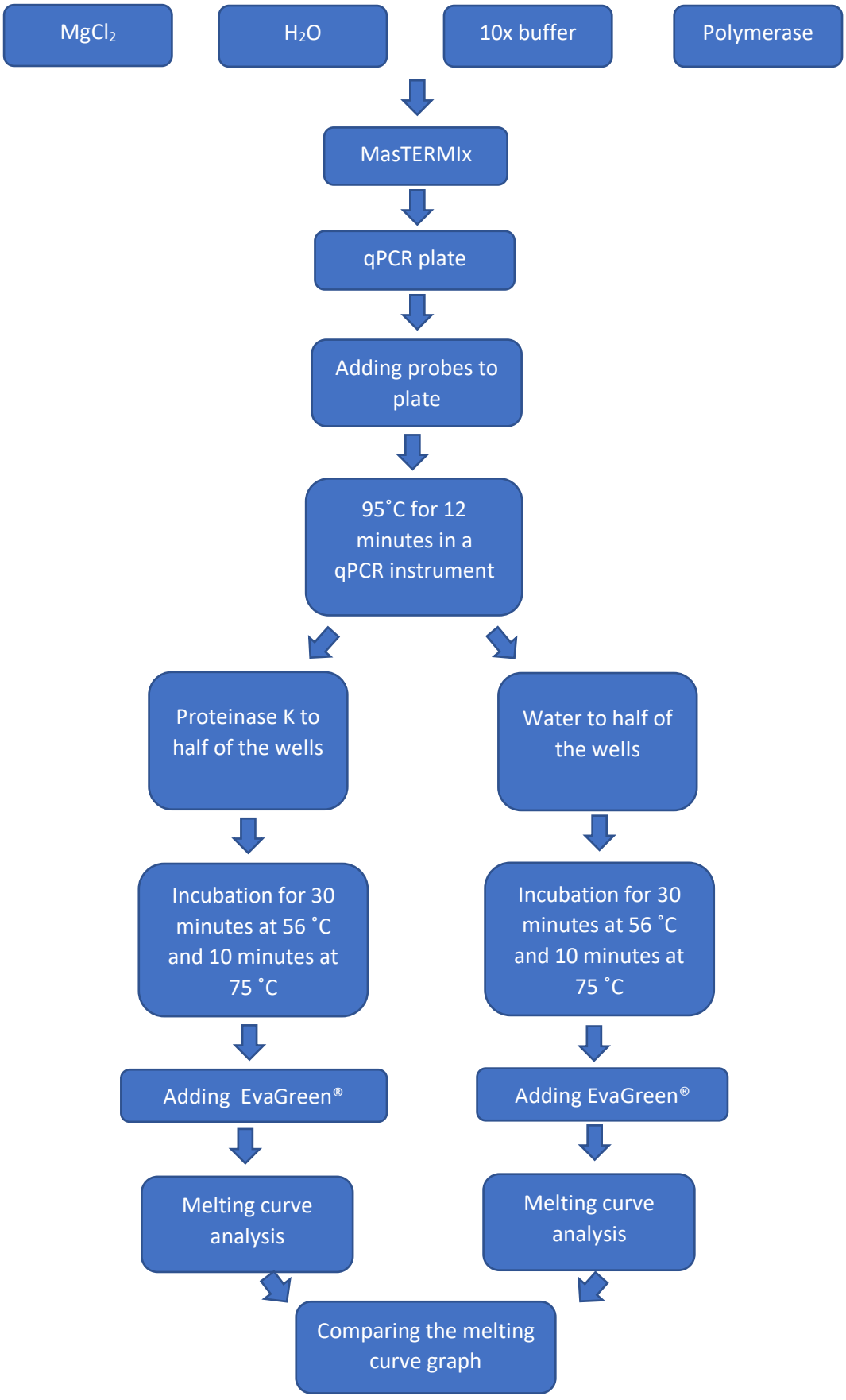
## 1.4. Aim of this study

The aim of this study was to investigate factors that might affect the melting temperature (*T*m) of DNA probes in a melting curve analysis performed utilizing a qPCR machine. The main focus was to investigate if DNA polymerase could have an effect on the *T*m of the DNA duplexes. To establish if the DNA polymerase can cause a *T*m-shift in the melting curve analysis. To be able to explore this, several secondary goals were set:

- Comparing the effect different DNA polymerases have on the *T*m of the DNA probes
- Comparing the difference between the theoretical *T*m value of the probes to their observed *T*m when they are mixed with DNA polymerase.
- Comparing the *T*m of samples containing active DNA polymerase and the *T*m of samples containing inactive polymerase.
- Looking at the effect different mismatches in a DNA sequence can have on the *T*m.

# 2. Materials and methods

## 2.1. Flow chart of the experimental setup



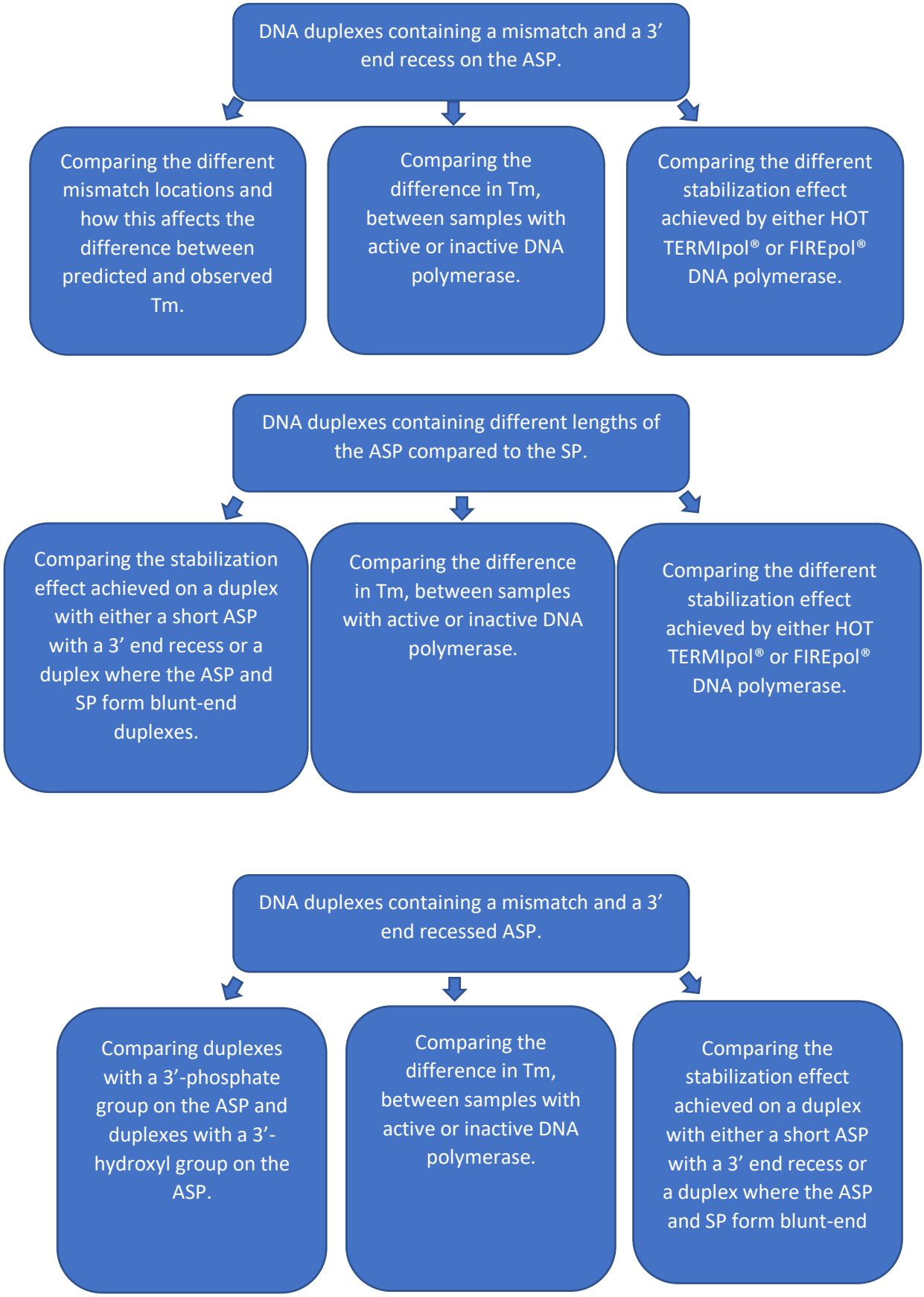
## 2.2. Naming and terminology of the probes

The probes used in this study are named based on either their length, mismatch position if the duplex contains a mismatch or the predicted *T*m of the duplexes. The first part of the name SP or ASP is an abbreviation from sense primer and anti-sense primer, respectively. This terminology might be slightly misleading as the sequences used are not necessarily coding sequences. These oligonucleotides are made with no other purpose than to test at what temperature they will denature, no regards have been given towards the function of the sequence, but the terms ASP and SP will be used in order to categorize the different strands that make the duplex. The letter behind the ASP or SP represents a quality about the probe. If the letter is an L, the number represents the number of nucleotides on the strand. If the ASP or SP has a mismatch in the sequence, it is represented by an M in the name, where the number is the position of the mismatch, calculated from the 3' end of the ASP. With the last letter being the predicted *T*m of the DNA duplex, represented by a T, and the following number is the predicted melting temperature of the DNA duplex in Celsius. The SPs and ASPs that have been used to form duplexes in this experiment are shown in **Table 1**, along with the names of the duplexes. The sequences of the probes can be found in **Table A1** in **Appendix A.**

**Table 1. An overview of probes used in this study, including duplex names, SP and ASP names, what mismatch is present, and the length difference between the ASP and SP in the duplex.**

| Duplex name | SP | ASP | Mismatch | Length difference (nucleotides) |
|---|---|---|---|---|
| SP-M0 | SP-M0 | ASP-M0 | none | 10 |
| ASP-M4 | SP-M0 | ASP-M4 | A-A | 10 |
| ASP-M12 | SP-M0 | ASP-M12 | A-A | 10 |
| SP-M8 | ASP-M0 | SP-M8 | G-G | 10 |
| | | | | |
| ASP-T64 | SPL24 | ASPL24 | None | 0 |
| ASP-T42 | SPL24 | ASPL14 | None | 10 |
| | | | | |
| ASPL61 | SPL61 | ASPL61 | None | 0 |
| ASPL61-3P | SPL61 | ASPL61-3P | None | 0 |
| ASPL25 | SPL61 | ASPL25 | None | 36 |
| ASPL15 | SPL61 | ASPL15 | None | 46 |
| ASPL15-3P | SPL61 | ASPL15-3P | None | 46 |

## 2.3. Flow chart of information gathered from the different probes



## 2.4. Chemicals, buffers and other solutions

**Table 5. The different buffers used in the solution**

| Chemical | Specification | Supplier |
|---|---|---|
| 10X buffer B2 | *0.8 M Tris-HCl, 0.2 M $(NH_4)_2SO_4$* | Solis Biodyne |
| 10X buffer C | *500 mM Tris-HCl pH 9.5 at 25ºC* | Solis Biodyne |

**Table 6. Reagents and concentrations for master mixes**

| | Master mix 1 – HOT TERMIpol® | | Master mix 2 - FIREpol® | |
|---|---|---|---|---|
| **Reagent** | **Volume (µL)** | **End concentration** | **Volume(µL)** | **End concentration** |
| C Buffer | 4/0 | 50mM | 4/0 | 0 |
| B2 Buffer | 4/0 | 0 | 4/0 | 80mM |
| $MgCl_2$ | 1.6 | 1mM | 1.6 | 1mM |
| Polymerase | 1.6 | 0.2 U/µL | 1.6 | 0.2 U/µL |
| $H_2O$ | 22 | | 22 | |

**Table 7. Reagents and concentrations for other solutions**

| Reagent | Volume (µL) | End concentration |
|---|---|---|
| Compensation $H_2O$ | 2/0 | |
| EvaGreen® | 8 | 1.25 µM |
| Probe | 0.8 | 1µM |
| Proteinase K | 0/2 | 0.03 mAU/ µL |

## 2.5. Software and Online Resources

**Table 8. Overview of software and online resources**

| Software and Online Resources | Specifications |
|---|---|
| Oligoanalyzer 3.1 | https://eu.idtdna.com/calc/analyzer |
| Bio-rad CFX maestro | https://www.bio-rad.com/en-no/product/cfx-maestro-software-for-cfx-real-time-pcr-instruments?ID=OKZP7E15 |

## 2.6. Melting curve analysis

Polymerase chain reaction (PCR) is a method for exponential amplification of short DNA samples using a thermostable DNA polymerase. Quantitative PCR (qPCR) uses a fluorescent label, which is proportional to the amount of produced DNA.

The quantitative PCR machine can be used to achieve several goals. Using both a melting curve analysis and amplification of DNA, it can be used to identify and quantify differences in DNA, genes, and nucleic acids. The most common way to perform a melting-curve analysis in the qPCR is by using a fluorescent dye and observing the intensity released from the dye. In this study, we performed a melting curve analysis using the dye EvaGreen®. EvaGreen® is inactive when there is no DNA available and will be activated in the presence of DNA. It does not emit any light until it is bound to DNA. This makes it a good dye for a qPCR melting curve analysis since it will be activated on demand (Mao et al., 2007).

In a melting curve analysis, the DNA is exposed to increasing temperature, and it will start to denature, releasing the dye, and the decrease in intensity can be observed and will correlate with the amount of DNA that has been denatured. As the temperature rises higher, a greater drop in fluorescence can be observed until the temperature is so high that the DNA completely disassociates, and a sharp drop in the fluorescence can be observed. When the DNA has reached 50% disassociation when there is an equal amount of double-stranded and single-stranded DNA, it has reached its melting temperature (*T*m). This *T*m is followed by a further drop in intensity until all the DNA has released its dye; this will happen rapidly. Using the qPCR instrument, it can create a melting curve which shows the intensity of light at different temperatures. This curve can be compared to other samples and will yield information about differences in the *T*m, and such differences between the DNA samples can be detected by looking at the *T*m differences (Reed et al., 2007).

## 2.7. Oligoanalyzer 3.1

Olgioanalyzer 3.1 is a tool made available by Integrated DNA Technologies (IDT). It is an online tool that is able to make predictions about the stability and melting temperatures of DNA duplexes. Oligoanalyzer considers many factors when making a prediction, salt concentration, dNTP concentration, oligo concentration, and mismatches. In order to calculate the *T*m, the tool uses equation 2, presented in the introduction, 1.2.1 The biochemistry of DNA duplexes and duplex formation (Integrated DNA Technologies, 2019). As mentioned in the introduction, $Na^+$ and $Mg^{2+}$ will stabilize the DNA duplex, depending on the concentration of salt. An increase in oligonucleotide

concentration will also require more salts in order to gain the same effect. All the parameters will change the resulting *T*m predicted by the tool. The parameters used in oligoanalyzer to predict the *T*ms of the probes used are shown in **Table 8.** The predicted *T*ms acquired are determined using these parameters and are shown in **Table 9.** In **Table 9,** one can see a discrepancy between the parameters used for probe set 1, and the reagents added into the solution. The parameters used in the prediction contains 1 mM $Na^+$, which is not in the experimental setup. This is because the prediction tool does not allow for the calculation of *T*m in duplexes containing mismatches (probe set 1), without at least 1 mM $Na^+$.

Equation 2:

$$T_m^{-1} = \frac{R}{\Delta H} lnC_T + \frac{\Delta S}{\Delta H}$$

**Table 9. The following parameters were used in Oligoanalyzer for probe set 1, 2, and 3.**

| Target type | DNA (probe set 1) | DNA (probe set 2 & 3) |
|---|---|---|
| Oligonucleotide conc | 1 µM | 1 µM |
| Na | 1 mM | 0 mM |
| Mg | 1 mM | 1 mM |
| dNTP conc | 0 mM | 0 mM |

**Table 10. The predicted *T*ms by oligoanalyzer**

| Probe | Predicted *T*m ( ˚C) |
|---|---|
| ASP-M4 | 37.4 |
| ASP-M12 | 39.9 |
| SP-M0 | 46.9 |
| SP-M8 | 40 |
| ASP-T42 | 42 |
| ASP-T64 | 63.5 |
| ASP-L61-3P | 75.2 |
| ASP-L61 | 75 |
| ASP-L25 | 61.5 |
| ASP-L15-3P | 45.7 |
| ASP-L15 | 45.7 |

## 2.8. DNA polymerases

The DNA polymerases used in this study were FIREpol® DNA polymerase and HOT TERMIpol® DNA polymerases.

FIREpol® DNA polymerase is derived from *E. coli* expressing a recombinant gene encoding a modified version of *Thermus aquaticus* DNA polymerase, which means that it is a DNA polymerase that is highly thermostable. It is a single polypeptide polymerase and has 5'-3' polymerase activity, as well as 5'-3' endonuclease activity, which means that it will add nucleotides from a 5'-3' direction and can cleave phosphor-diester bonds in the same direction. FIREpol® also possesses a non-template-dependent terminal transferase activity. The polymerase does not possess a 3'-5' exonuclease activity (Solis BioDyne, 2020).

Hot TERMIpol® DNA polymerase is also produced through an *E. coli* strain, as a modified version of the *Thermus aquaticus* DNA polymerase. The product of the recombinant *T. aquaticus* gene that is expressed in *E. coli* is called TERMIpol®, which is then chemically modified to produce the final product of HOT TERMIpol®. This means that the polymerase is thermostable and that it needs an activation step, by heating it to 95 ˚C for at least 12 minutes. HOT TERMIpol® is a single polypeptide polymerase and has a 5'-3' polymerase activity, and does not contain, nicking activities, priming activities, non-specific endonuclease or exonuclease activities (Solis BioDyne, 2020)

## 2.9. Method

The aim of this study was to investigate the stabilizing effect DNA polymerase might have on DNA duplexes. In order to do this, a melting curve analysis was performed, with three different probe sets, to see how different DNA duplexes are affected. The designed duplexes had mismatch variations, length differences between the strands, and a difference in length of the sequences, as well as 3' phosphates instead of 3'OH on the recessed ends of the anti-sense primers. Probe set 1 had both length differences between the anti-sense primer (ASP) and the sense primer (SP) in addition to mismatches in different locations in the sequence. Probe set 2 allowed comparisons between ASP and SP, when there was both equal lengths and a length difference between them. Probe set 3, had varying length differences between ASP and SP, as well as some probes with 3' phosphates. In order to test the effects, the different variations have on the stabilizing effect of DNA polymerase, the following experiment was performed.

The first step was to make a master mix that was used for all the probes. The master mix contained 22 µL of nuclease-free water, 0.2 U/µL of either HOT TERMIpol® DNA polymerase or FIREpol® DNA polymerase, 50 mM of 10X C buffer if HOT TERMIpol® was used or 80 mM 10X B2 buffer if FIREpol® was used. Lastly, it contains 1 mM of $MgCl_2$. The master mix was then added to the qPCR plate, and 1 µM of the probes were added, and the plate was run through a 12-minute heating step at 95 ˚C in order to activate the HOT TERMIpol®, samples that used FIREpol® instead was still subjected to this step. After the first activation step, 0.03 mAU/µL of proteinase K was added to half the samples, and 2 µL of nuclease free $H_2O$ was added to the other half in order to compensate for the prot.K volume. The plate was then run through an additional heating step in the qPCR machine for 30 minutes at 56 ˚C to activate the proteinase K, then for 10 minutes at 70 ˚C in order to inactivate the proteinase K again. The last step is then to add 1.25 µM of EvaGreen® and to run the melting curve analysis with a total volume of 40 µL. The melting curve analysis was run from 31-85 ˚C with a 0.5 ˚C increment over 10 seconds. Overview of chemicals used and suppliers can be found in **Appendix A, Table A2.**

# 4. Results

The results of this study have been divided into three sections. In each section, the DNA duplex has a different sequence, as well as one additional quality that was investigated. The sections are categorized as probe set 1, probe set 2, and probe set 3. Probe set 1 includes duplexes with a mismatch and an anti-sense primer, which is ten nucleotides shorter than the sense primer. While probe set 2 has duplexes with a different sequence and compares duplexes with varying lengths of the ASP. Probe set 3 contains duplexes, where the 3'end of the ASP is blocked by a 3'-phosphate group, and the set has different lengths of the ASP.

## 4.1. Probe set 1 – comparing the effect of mismatches and length differences

The effect polymerase has on the *T*m of DNA has been investigated by using both FIREpol® DNA polymerase and HOT TERMIpol® DNA polymerase, using proteinase K as an inactivator of the polymerase. The position of the mismatch in these duplexes, is specified in the probe name, in example, ASP-M4 will have a mismatch in position 4 from the 3' end on the anti-sense strand and SP-M8 will have a mismatch in position 8 counted from the 3' end of the anti-sense strand, the specific mismatch will be supplied in the Figure text. The sequences of the probes can be found in **Figure A1** in **Appendix A**.

Samples containing HOT TERMIpol® DNA polymerase are presented in **Figure 3A** and the samples containing FIREpol® DNA polymerase in **Figure 3B**. The theoretical melting point of ASP-M4 is 37.4 ˚C. **Figure 3A** shows an average melting temperature of 47.3 ˚C in the samples containing active HOT TERMIpol® DNA polymerase. This is a 9.9 ˚C increase in *T*m compared to the theoretical value. The average melting temperature in samples treated with proteinase K was not observable as there are no clear peaks in the graphs. However, a steady decline in intensity at approximately 35 ˚C is present in all samples containing proteinase K. **Figure 3B** shows an average melting point in samples containing active FIREpol® DNA polymerase of 49.8 ˚C. The samples containing proteinase K do not have a clear peak, and it is not possible to determine the *T*m. However, there is a steady decline in intensity after 34 ˚C has been reached in the melting curve analysis. The sample labeled ASP-M4-1 shows a much higher intensity than the other samples and reaches 640 -d(RFU)/dT.

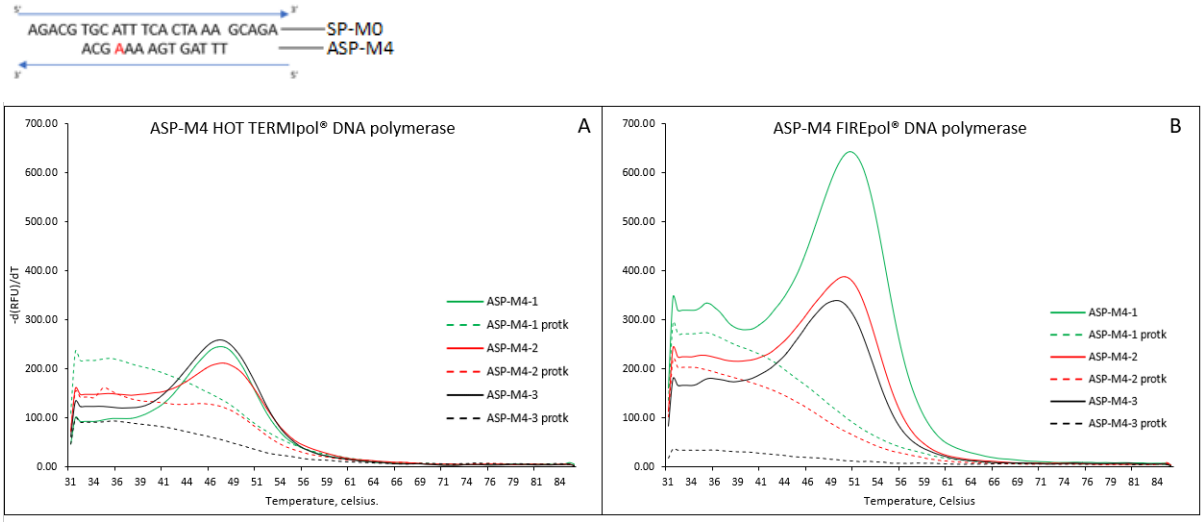



**Figure 3: EvaGreen® based melting curve analysis of probe duplexes containing a sense primer and an anti-sense primer with an A-A mismatch on position 4 in the ASP.** A) the probes either contained active HOT TERMIpol® DNA polymerase or inactive polymerase, inactivated by proteinase K. The experiment was done in a triplicate indicated by "ASP-M4-1", "ASP-M4-2" and "ASP-M4-3" if containing active polymerase, or "ASP-M4-1 ProtK", "ASP-M4-2 ProtK" and "ASP-M4-3 ProtK" if containing inactive polymerase. Panel B has the same setup as panel A, but HOT TERMIpol® was replaced by FIREpol®. The Figure shows the influence of DNA polymerase on the *T*m of the duplexes. Reactions in panel a were performed in a solution of 50 mM Tris-HCl pH 9.5, 1 mM $MgCl_2$, 0.2 U/µL HOT TERMIpol® DNA polymerase, 1.25 µM EvaGreen®, and 1 µM oligonucleotides. Stippled lines contain 0.03 mAU/µL Proteinase K, while regular lines contain an equal volume of water instead. Panel B consisted of the same reagents except the buffer and polymerase, the buffer used for panel B is 80mM Tris-HCl with 0.2M $(NH_4)_2SO_4$ and the polymerase used is 0.2 U/µL FIREpol® DNA polymerase. The triplicate with active polymerase in panel A, shows an average *T*m of 47.3 ˚C, which is 15 ˚C higher than the theoretical *T*m of 37.4 ˚C. Panel B shows the triplicate containing samples with active fire polymerase; these yielded an average *T*m of 49.8 ˚C, which is 12.4 ˚C higher than the theoretical *T*m. No definite *T*m-shift, as proteinase K treated samples in both A) and B), shows a steady decline of derived intensity from approximately 35 ˚C with no peaks.

The results of the melting curve analysis of ASP-M12, which contained an A-A mismatch in position 12 on the ASP, is presented in **Figure 4.** Panel A represents samples treated with HOT TERMIpol® DNA polymerase, and panel B represents samples treated with FIREpol® DNA polymerase. The theoretical melting point of ASP-M12 was found to be 39.9 ˚C. Probes treated with HOT TERMIpol® showed an average *T*m of 56.3 ˚C, which was an increase in *T*m of 16.4 ˚C from the theoretical value. The proteinase K treated samples showed a decrease in intensity from 37 ˚C. However, the sample "ASP-M12-1 ProtK" had a small peak at 54 ˚C. **Figure 4B** had a greater variance in the samples containing active polymerase and had a *T*m averaging 57.7 ˚C. This was a 17.8 ˚C difference in *T*m from the theoretical value. Proteinase K treated samples showed a lower intensity than the samples with active polymerase. "ASP-M12-1 ProtK" produced a clear peak at 57 ˚C but also shows a slight increase in intensity between 39-45 ˚C. "ASP-M12-2 ProtK" had a wide peak at 43 ˚C which was 15.5 ˚C lower than the corresponding peak containing active polymerase. "ASP-M12-3 ProtK" had an increasing intensity from 34 ˚C, but the intensity declined from 40 ˚C, resulting in a wide peak. In general, this sample had a low intensity with 150 -d(RFU)/dT at its highest.

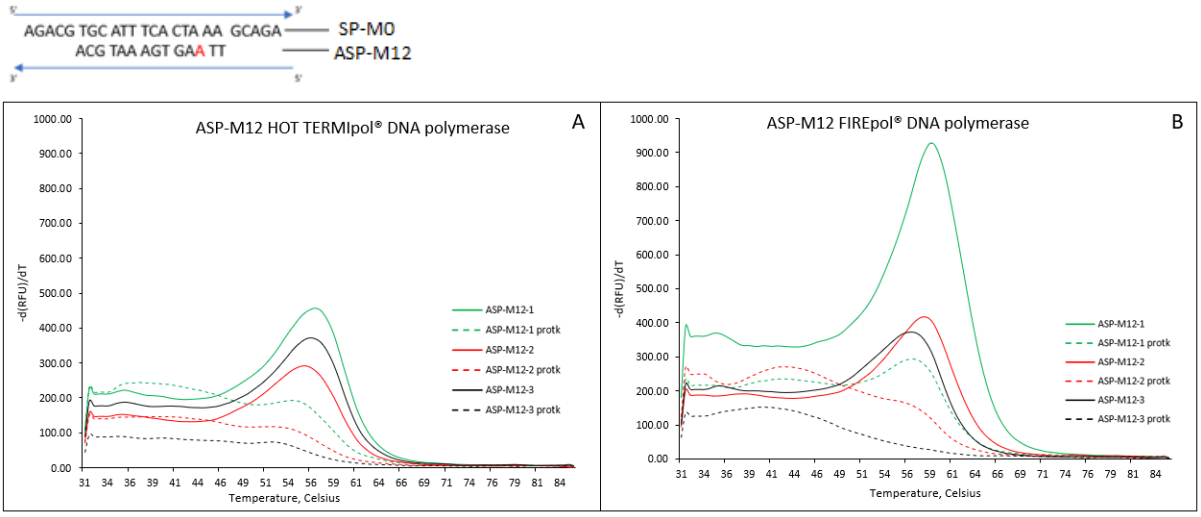



**Figure 4: EvaGreen® based melting curve analysis of probe duplexes containing a sense primer and an anti-sense primer with an A-A mismatch on position 12 in the ASP.** A) the probes either contained active HOT TERMIpol® DNA polymerase or inactive polymerase, inactivated by proteinase K. The experiment was done in a triplicate indicated by "ASP-M12-1", "ASP-M12-2" and "ASP-M12-3" if containing active polymerase or "ASP-M12-1 ProtK", "ASP-M12-2 ProtK" and "ASP-M12-3 ProtK" if containing inactive polymerase. Panel B has the same setup as A), but HOT TERMIpol® was replaced by FIREpol®. The experimental setup was the same as shown in **Figure 3**. The triplicate with active polymerase in panel A shows an average *T*m of 56.3 ˚C, which is 16.4 ˚C higher than the theoretical *T*m of 39.9 ˚C. Panel B shows the triplicate containing samples with active fire polymerase; these yields an average *T*m of 57.7 ˚C, which is 17.8 ˚C higher than the theoretical *T*m.

**Figure 5** represents the melting curve analysis of SP-M0. This duplex did not contain any mismatches and was the duplex in its original form. As with the previous figures, the primers treated with HOT TERMIpol® DNA polymerase are represented in panel A, and the samples treated with FIREpol® DNA polymerase is represented in panel B**.** The triplicate treated with HOT TERMIpol® DNA polymerase had an average *T*m of 54.7 ˚C, with "SP-M0-2" having a higher *T*m than the two others at 56 ˚C. The theoretical *T*m was found to be 46.9 ˚C and was 7.8 ˚C lower than the average *T*m of the samples that contained active polymerase. The proteinase K treated samples showed a large variance in -d(RFU)/dT and displayed peaks at a higher *T*m than previously recorded with other probes. Most notable is "SP-M0-3 ProtK" and "SP-M0-2 ProtK" which had *T*ms of 57 ˚C and 56˚, correspondingly 2.3 ˚C and 1.3 ˚C higher than the average *T*m in samples that contained active polymerase. With "SP-M0-1 ProtK" having a small peak at 54 ˚C, which was 0.7 ˚C lower than the average *T*m of samples with the active polymerase. They also had higher intensities than their counterparts, which had not been observed before. **Figure 5B** has no observed signal for "SP-M0-3" and a large variance between "SP-M0-1" and "SP-M0-2". The average *T*m would be 58.7 ˚C, which was 11.8 ˚C higher than the theoretical *T*m. "SP-M0-3 ProtK" showed a wide peak at 43 ˚C, but its corresponding sample showed no signal, while the other samples with inactive polymerase had no clear peaks. "SP-M0-1" had a very high intensity with a -d(RFU)/dT at 1120.

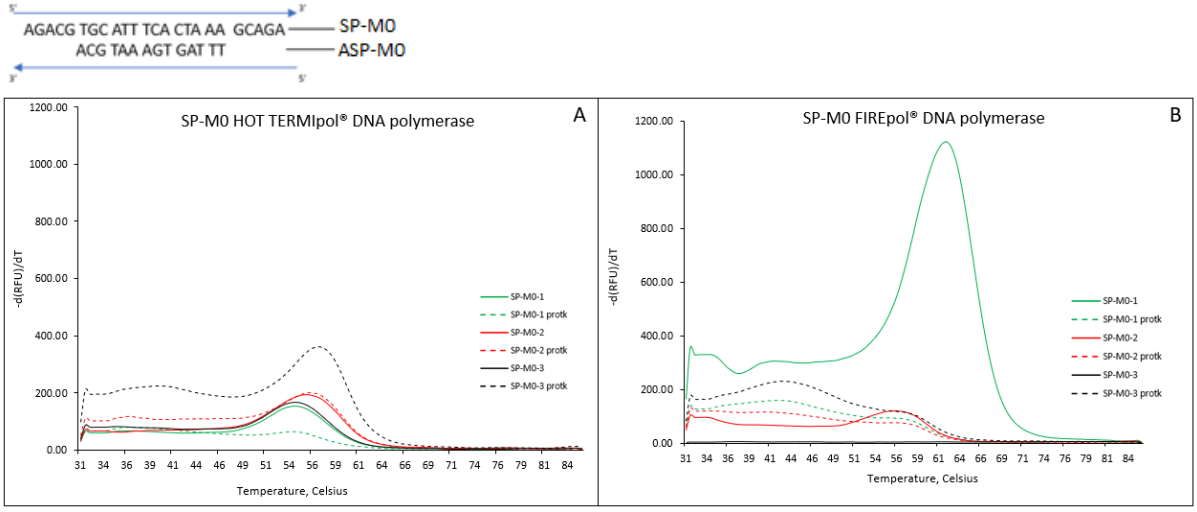



**Figure 5**: **EvaGreen® based melting curve analysis of probe duplexes containing a sense primer and an anti-sense primer with no mismatches in the duplex.** A) the probes either contained active HOT TERMIpol® DNA polymerase or inactive polymerase, inactivated by proteinase K. The experiment was done in a triplicate indicated by "SP-M0-1", "SP-M0-2" and "SP-M0-3" if containing active polymerase or "SP-M0-1 ProtK", "SP-M0-2 ProtK" and "SP-M0-3 ProtK" if containing inactive polymerase. Panel B had the same setup as panel A, but HOT TERMIpol® was replaced by FIREpol®. The experimental setup was the same, as shown in **Figure 3**. The triplicate with active polymerase in panel A shows an average *T*m of 54.7 ˚C, which is 7.8 ˚C higher than the theoretical *T*m of 46.9 ˚C. Panel B shows the triplicate containing samples with active fire polymerase; these yielded an average *T*m of 58.7 ˚C, which was 11.8 ˚C higher than the theoretical *T*m. SP-M0-3 in panel B showed no signal, SP-M0-1 showed a six times higher intensity than SP-M0-2 and a 6 ˚C higher melting point.

SP-M8 had a G-G mismatch on position 8 in the SP, and its theoretical *T*m was found to be 40 ˚C. **Figure 6A** shows the data of the samples treated with HOT TERMIpol® DNA polymerase. With an average observed *T*m of 49.5 ˚C in the samples with active polymerase, which was 9.5 ˚C higher than the theoretical *T*m. The proteinase K treated samples showed an increase in intensity starting at 35 ˚C with a decline that started at 42-44 ˚C. **Figure 6B** shows an average *T*m of 51.7 ˚C in samples containing active polymerase, which is 11.7 ˚C higher than the theoretical value. "SP-M8-2" showed a rapid drop in intensity at 37 ˚C and increased at 40 ˚C. The proteinase K treated samples had peaks, although wide, at respectively, 44 ˚C, 45 ˚C, and 42 ˚C, with the average *T*m of 43.7 ˚C. This average *T*m was 3.7 ˚C higher than the theoretical *T*m and 8 ˚C lower than the samples with active FIREpol® DNA polymerase.

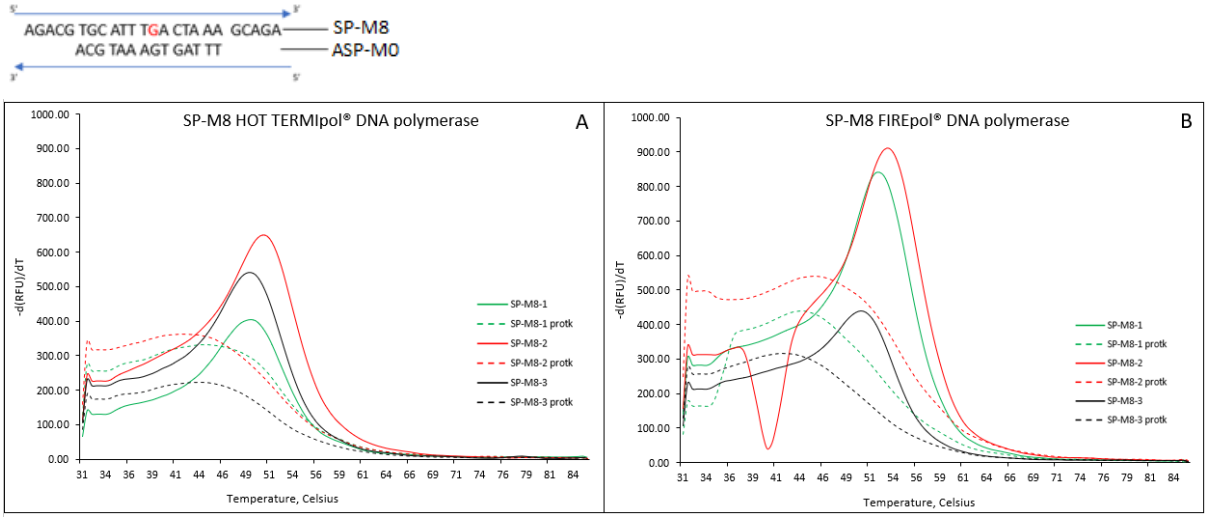



**Figure 6**: **EvaGreen® based melting curve analysis of probe duplexes containing a sense primer and an anti-sense primer with a G-G mismatch in the 8th position on the SP.** A) the probes either contained active HOT TERMIpol® DNA polymerase or inactive polymerase, inactivated by proteinase K. The experiment was done in a triplicate indicated by "SP-M8-1", "SP-M8-2" and "SP-M8-3" if containing active polymerase or "SP-M8-1 ProtK", "SP-M8-2 ProtK" and "SP-M8-3 ProtK" if containing inactive polymerase. Panel B has the same setup as A, but HOT TERMIpol® was replaced by FIREpol®. The experimental setup is the same as shown in **Figure 3**. The triplicate with active polymerase in panel A, shows an average *T*m of 49.5 ˚C, which was 9.5 ˚C higher than the theoretical *T*m of 40 ˚C. Panel B shows the triplicate containing samples with active fire polymerase; these yielded an average *T*m of 51.7 ˚C, which is 11.7 ˚C higher than the theoretical *T*m.

With the goal of getting a better understanding of how the position of the mismatch can influence the stability of the DNA duplex, **Figure 7** presents a comparison between all the probes in probe set 1, with the focus on comparing the average observed *T*m and the predicted *T*m. The results from samples treated with HOT TERMIpol® DNA polymerase is presented in **Table 13,** and the samples treated with FIREpol® DNA polymerase are presented in **Table 14**. These tables represent the results from the entire probe set as well as the results presented in **Figures 3, 4, 5 & 6**. In **Figure 7** the mismatched positions are presented in different colors depending on the relative positioning of the mismatch. The results presented in **Figure 7** are only from samples treated with HOT TERMIpol® DNA polymerase. The average *T*ms are based on the average of the triplicates in the experiment. A mismatch in position 2-8 had an average *T*m of 48.6 ˚C. While positions 9-10 had the average *T*m of 51.3 ˚C. Positions 11-14 had an average *T*m of 55.7 ˚C. The sample with no mismatches in the sequence showed an observed *T*m of 54.7 ˚C. The probe with a mismatch in the first position showed a *T*m of 38.9 ˚C.

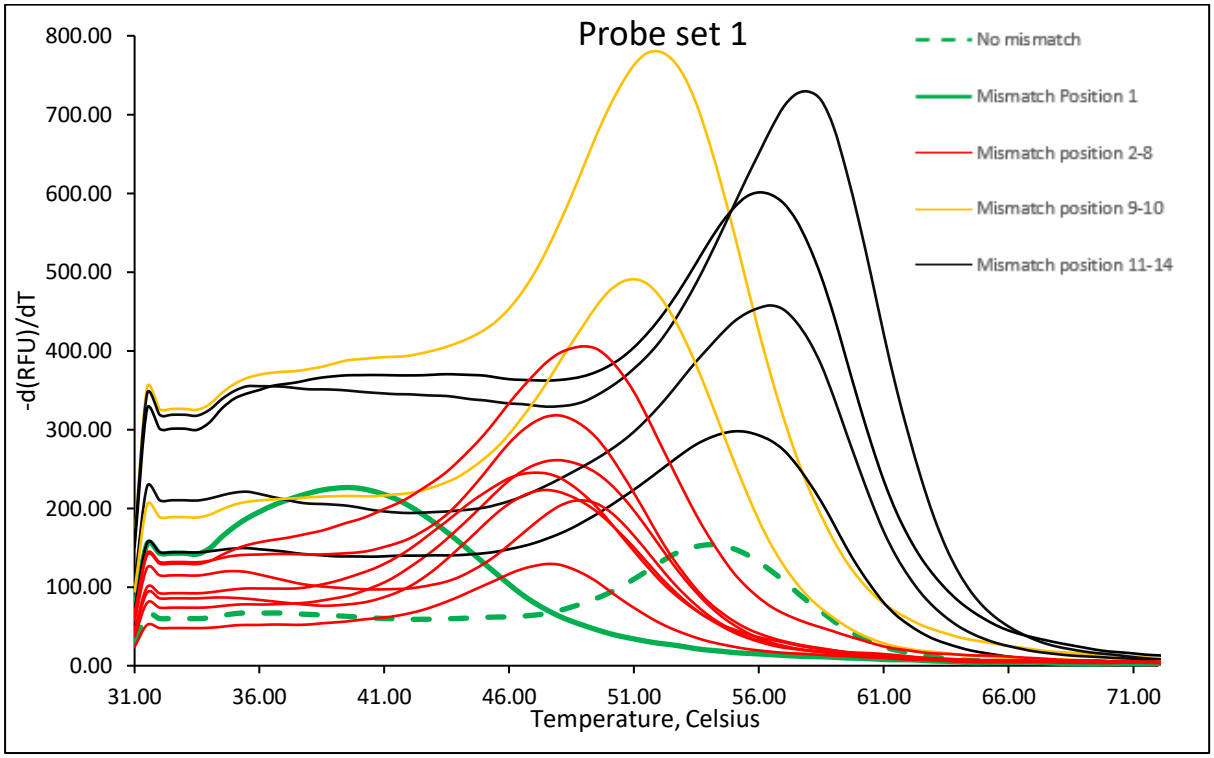



**Figure 7: Overview of the effect of a mismatch in a DNA duplex**. The position of the mismatch was counted from the 3' end on the ASP. The experimental setup is the same as shown in **Figure 3**, with HOT TERMIpol® DNA polymerase used. This Figure is demonstrating the first series of the triplicate. With the green dotted line being the probe without mismatches, the solid green line is the probe with a mismatch in position 1. The red line is mismatches from positions 2 to 8, orange represents the positions 9-10, and the black graphs are the positions 11-14.

**Table 13. Probes treated with HOT TERMIpol® DNA polymerase. Comparing the average observed *T*m to the predicted *T*m of the duplexes.**

| Probe | Mismatch position from the 3' end on the ASP | Observed *T*m (˚C) | Predicted *T*m (˚C) | Difference in *T*m (˚C) |
|---|---|---|---|---|
| SP-M0 | 0 | 54.7 | 46.9 | 7.8 |
| SP-M1 | 1 | 38.9 | 45.2/38.5[1] | -6.3/0.4[1] |
| ASP-M2 | 2 | 48.8 | 38.5 | 10.3 |
| SP-M3 | 3 | 48.33 | 37.1 | 11.2 |
| ASP-M4 | 4 | 47.3 | 37.4 | 9.9 |
| SP-M5 | 5 | 49.6 | 39.5 | 10.1 |
| SP-M6 | 6 | 49.8 | 37.6 | 12.2 |
| SP-M7 | 7 | 47 | 37.1 | 9.9 |
| SP-M8 | 8 | 49.5 | 40 | 9.5 |
| ASP-M9 | 9 | 51 | 36.9 | 14.1 |
| SP-M10 | 10 | 51.5 | 36.3 | 15.2 |
| SP-M11 | 11 | 54.3 | 38.3 | 16 |
| ASP-M12 | 12 | 56.3 | 39.9 | 16.4 |
| ASP-M13 | 13 | 55.7 | 38 | 17.7 |
| ASP-M14 | 14 | 56.3 | 45.8/42.4[1] | 10.5/13.9[1] |

[1] Due to a believed overestimation of the prediction tool, two values have been presented. Where the leftmost value is the original predicted Tm, and the rightmost value is the predicted Tm without the mismatched base pair included in the calculation.

**Table 14. Probes treated with FIREpol® DNA polymerase. Comparing the average observed *T*m to the predicted *T*m of the duplexes.**

| Probe | Mismatch position from the 3' end on the ASP | Observed *T*m (˚C) | Predicted *T*m (˚C) | Difference in *T*m (˚C) |
|---|---|---|---|---|
| SP-M0 | 0 | 58.7 | 46.9 | 11.8 |
| SP-M1 | 1 | 39.7 | 45.2/38.5[2] | -5.5/1.2[2] |
| ASP-M2 | 2 | 49.7 | 38.5 | 11.2 |
| SP-M3 | 3 | 51.3 | 37.1 | 14.2 |
| ASP-M4 | 4 | 49.8 | 37.4 | 12.4 |
| SP-M5 | 5 | 50.8 | 39.5 | 11.3 |
| SP-M6 | 6 | 50.5 | 37.6 | 12.9 |
| SP-M7 | 7 | 49.3 | 37.1 | 12.2 |
| SP-M8 | 8 | 51.7 | 40 | 11.7 |
| ASP-M9 | 9 | 50.8 | 36.9 | 13.9 |
| SP-M10 | 10 | 55.8 | 36.3 | 19.5 |
| SP-M11 | 11 | 57.7 | 38.3 | 19.4 |
| ASP-M12 | 12 | 57.7 | 39.9 | 17.8 |
| ASP-M13 | 13 | 56.7 | 38 | 18.7 |
| ASP-M14 | 14 | 58.2 | 45.8/42.4[2] | 12.4/15.8[2] |

[2] Due to a believed overestimation of the prediction tool, two values have been presented. Where the leftmost value is the original predicted Tm, and the rightmost value is the predicted Tm without the mismatched base pair included in the calculation.

## 4.2. Probe set 2 – comparing length differences

In order to see what effects a length difference between the SP and ASP might have on the stabilization properties of DNA polymerase, an experiment was run with two duplexes. One had an ASP with 23 nucleotides, and the other duplex had an ASP with 14 nucleotides, while the length of the SP was 24 nucleotides in both duplexes. The names of the duplexes are ASP-T42 (ASP has 14 nucleotides), and ASP-T64 (ASP has 23 nucleotides), with the T representing the predicted *T*m of the duplexes. The experiments in probe set 2 have been run on separate plates at separate times. The reagents and program for the melting curve analysis were the same as in probe set 1, one experiment with FIREpol® DNA polymerase and one experiment with HOT TERMIpol® DNA polymerase. The second set of probes did not have any mismatches in the duplexes. The results from these probes are presented as an overview in **Table 15**, as well as in **Figures 8 & 9**.

**Table 15: Summary of the average results for probe set 2.**

| **Probe** | **Length** | **Observed $_{\text{no prot. K}}$ *T*m (˚C)** | | **Observed$_{\text{+ prot. K}}$ *T*m (˚C)** | | **Predicted *T*m (˚C)** |
|---|---|---|---|---|---|---|
| | | **Hot TERMIpol® (˚C)** | **FIREpol® (˚C)** | **Hot TERMIpol® (˚C)** | **FIREpol® (˚C)** | |
| ASP-T42 | 14 | 61.5 | 63 | 37.5 | 38.5 | 42 |
| ASP-T64 | 23 | 61.5 | 62.5 | 61 | 62.5 | 63.5 |

ASP-T42 is represented in **Figure 8,** showing the results of the melting curve analysis on the duplex, unlike probe set 1, this probe set has been tested as duplicates. Panel A in **Figure 8** shows the HOT TERMIpol® treated duplexes and has two clear peaks in the samples with the inactive polymerase. The average *T*m of these peaks was 37.5 ˚C. This was the same as the theoretical *T*m of 37.5 ˚C and was 24 ˚C lower than the average *T*m of the samples with the active polymerase. The samples with active polymerase had an average *T*m of 61.5 ˚C. In panel B, the average *T*m of the samples with active polymerase had an average *T*m of 63 ˚C and is 25.5 ˚C higher than the theoretical *T*m. The proteinase K treated samples showed an average *T*m of 38.5 ˚C.

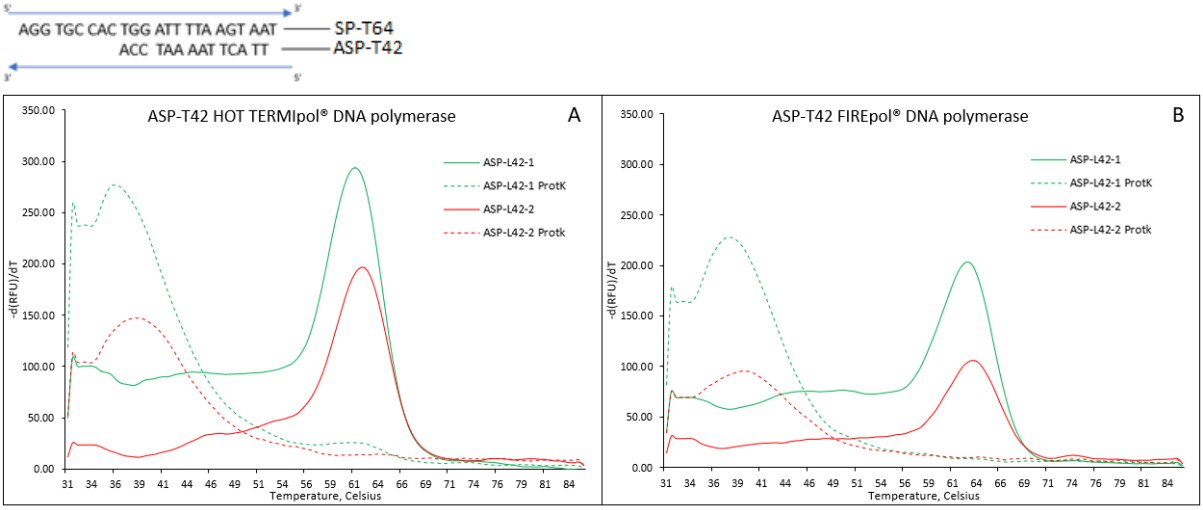



**Figure 8**: **EvaGreen® based melting curve analysis of probe duplexes containing a sense primer and an anti-sense primer with a 10-nucleotide difference in length of the primers.** A) the probes are either containing active HOT TERMIpol® DNA polymerase or inactive polymerase, inactivated by proteinase K. The experiment was done in duplicates, indicated by "ASP-T42-1" and "ASP-T42-2" if containing active polymerase or "ASP-T42-1 ProtK" and "ASP-T42-2 ProtK" if containing inactive polymerase. Panel B had the same setup as panel A, but HOT TERMIpol® was replaced by FIREpol®. The experimental setup was the same as shown in **Figure 3**. The duplicate with active polymerase in panel A showed an average *T*m of 61.5 ˚C, which was 24 ˚C higher than the theoretical *T*m of 37.5 ˚C. Panel B shows the duplicate containing samples with active FIREpol® DNA polymerase, and these yielded an average *T*m of 63 ˚C, which was 24.5 ˚C higher than the theoretical *T*m.

In **Figure 9**, a difference in *T*m between samples containing active or inactive DNA polymerase was not observed. In panel A, the average melting point in samples with inactive polymerase was observed to be 61 ˚C. While the *T*m of the samples containing active polymerase showed an average of 61.5 ˚C, which was 2 ˚C lower than the theoretical *T*m of 63.5 ˚C. Panel B represents the samples treated with FIREpol® DNA polymerase and they had an average *T*m of 62.5 ˚C in the samples with the active polymerase. This is 1 ˚C lower than the theoretical *T*m, while the samples with inactive polymerase had a *T*m of 62.5 ˚C, which also is 1 ˚C lower than the predicted *T*m.

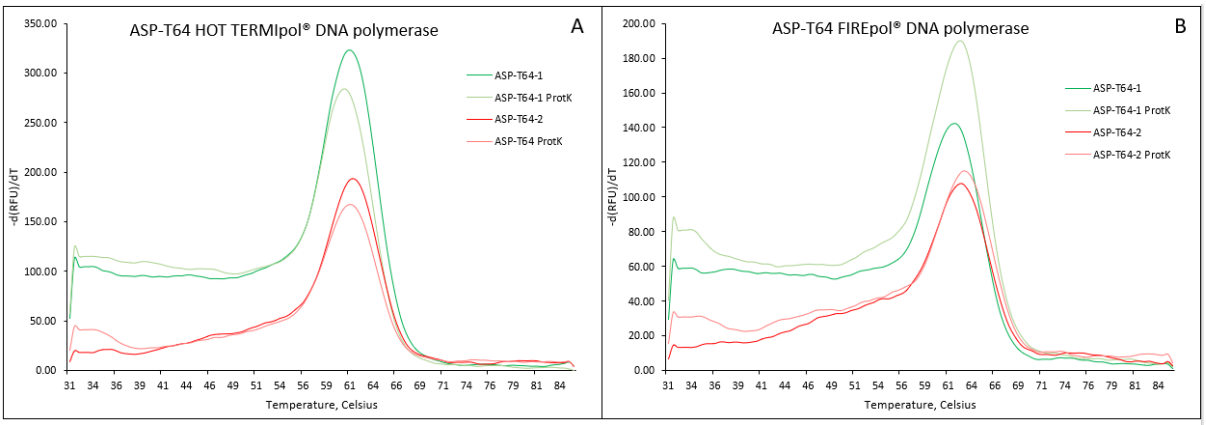




**Figure 9**: **Melting curve analysis of DNA duplexes with one nucleotide difference in the length between the ASP and SP.** A) the duplexes contained either active HOT TERMIpol® DNA polymerase or inactive polymerase, inactivated by proteinase K. The experiment was done in duplicates, indicated by "ASP-T64-1" and "ASP-T64-2" if containing active polymerase or "ASP-T64-1 ProtK" and "ASP-T64-2 ProtK" if containing inactive polymerase. Panel B had the same setup as panel A, but HOT TERMIpol® was replaced by FIREpol®. The experimental setup was the same as shown in **Figure 3**. The duplicate with active polymerase in panel A shows an average *T*m of 61.5 ˚C, which was 2 ˚C lower than the theoretical *T*m of 63.5 ˚C. Panel B shows the duplexes containing samples with active FIREpol®; these yielded an average *T*m of 62.5 ˚C, which was 1 ˚C lower than the theoretical *T*m.

## 4.3. Probe set 3 – the effect of a 3'-phosphate group on the ASP

In order to see further changes in the DNA duplex stability with different lengths in the ASP compared to the SP, and to see the effects of changing the 3'-hydroxyl group to 3'-phosphate in the ASP, the same experimental setup as probe set 1 were performed on probe set 3, only using HOT TERMIpol® DNA polymerase. Some of the duplexes in probe set 3 has been designed with the intent of mimicking a standard qPCR run, with a long SP, which acts as a template and short ASPs that acts as a primer. This was done to see if a qPCR run could experience *T*m-shifts caused by DNA polymerase. The experiments in probe set 3 was run as a triplicate. The third probe set did not contain any mismatches, and all the probes had the same sequence. There were three different lengths of the ASP in probe set 3 with, 15 nucleotides in ASP-L15, 25 nucleotides in ASP-L25, and 61 nucleotides in ASP-L61. While the SP always had 61 nucleotides. ASP-L61-3P and ASP-L15-3P had an end blocking phosphate on the 3' end of the ASP.

**Table 16. Represents the values of the average observed *T*m with polymerase, the average *T*m with proteinase K, and the predicted *T*ms for the duplexes.**

| Probe | Length | Obs.$_{\text{no prot. K}}$ *T*m (˚C) | Obs.$_{\text{+ prot. K}}$ *T*m (˚C) | Pred. *T*m (˚C) |
|---|---|---|---|---|
| ASP-L61 | 61 | 74 | 74 | 75.2 |
| ASP-L61-3P | 61 | 74.7 | 74.7 | 75.2 |
| ASP-L25 | 25 | 73 | 62.7/73[3] | 61.5 |
| ASP-L15 | 15 | 73 | 43.5/75[3] | 45.7 |
| ASP-L15-3P | 15 | 46 | 46 | 45.7 |

[3] Rightmost numbers represent secondary peaks in the melting curve analysis.

**Figure 10,** represents ASP-L61-3P, where the samples containing active polymerase had an average *T*m of 74.7 ˚C. With the theoretical *T*m being 75.2 ˚C, there was a 0.5 ˚C difference. The triplicate with inactivated polymerase also had an average *T*m of 74.7 ˚C. Showing that there was no difference between the samples with inactive polymerase and the samples with active polymerase.

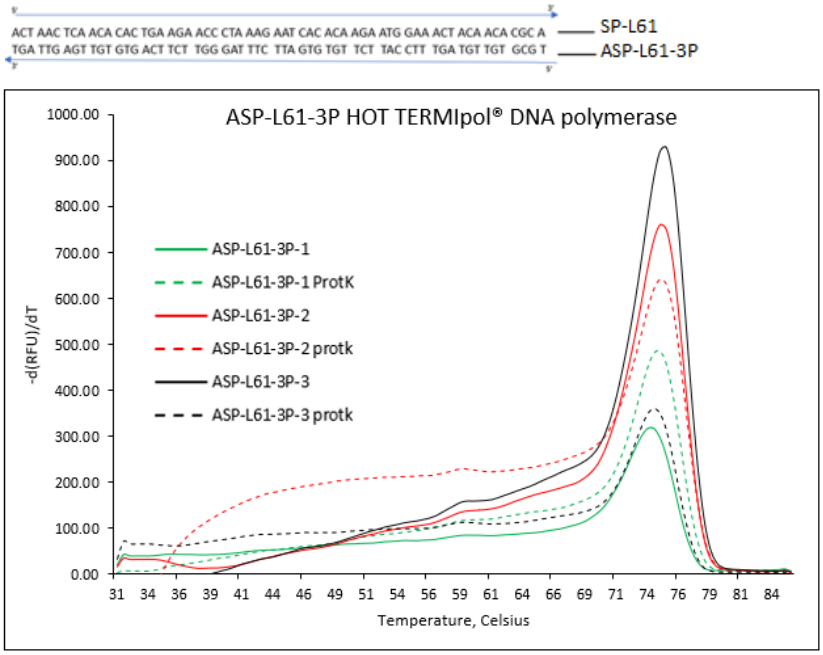



**Figure 10**: **Melting curve analysis of duplexes containing a 3'-phosphate on the ASP and without a length difference.** The polymerase used for these duplexes was HOT TERMIpol®. The polymerase was in half the samples inactivated by proteinase K. The experiment was done in a triplicate indicated by "ASP-L61-3P-1", "ASP-L61-3P-2," and "ASP-L61-3P-3" if containing active polymerase or "ASP-L61-3P-1 ProtK" and "ASP-L61-3P-2 ProtK" if containing inactive polymerase. The experimental setup was the same as in **Figure 3**. The triplicate with active polymerase in Figure 10 showed an average *T*m of 74.7 ˚C, which was 0.5 ˚C lower than the theoretical *T*m of 75.2 ˚C. The samples containing inactive polymerase also had an average *T*m of 74.7 ˚C. The anti-sense primer had a 3'-phosphate group instead of a 3'-hydroxyl group.

**Figure 11** represents the results from the ASP-L61 duplex. The results showed that the proteinase K treated samples and the samples with active polymerase were again similar when there was no difference in length between the ASP and SP. With the average *T*m in samples with active polymerase being 74 ˚C, and duplexes with inactive polymerase also had an average *T*m of 74 ˚C. The theoretical *T*m is 75.2 ˚C and is 1.2 ˚C higher than the observed *T*ms.

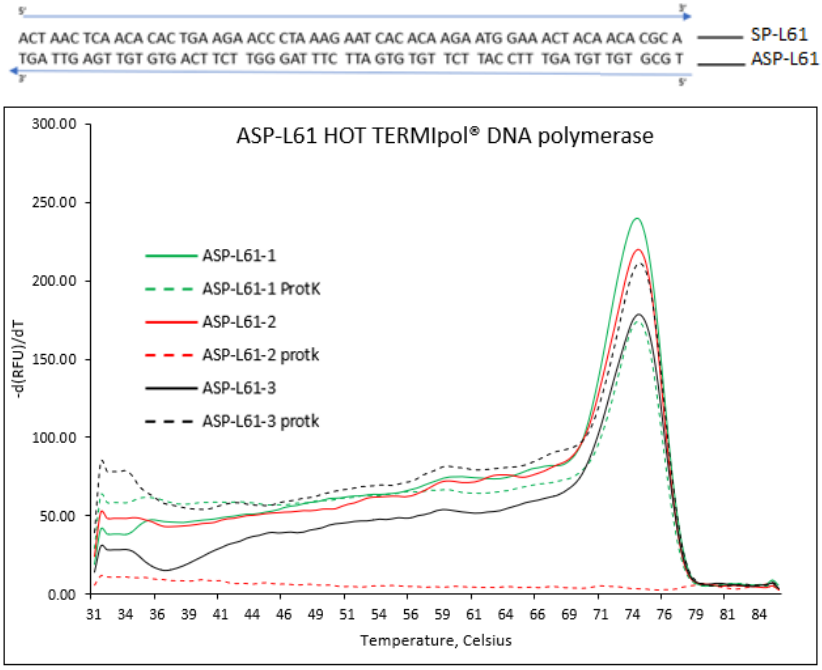



**Figure 11**: **Melting curve analysis of duplexes with no length difference between the SP and ASP.** The polymerase used for these duplexes was HOT TERMIpol®. Half the samples were inactivated by proteinase K. The experiment was done in a triplicate indicated by "ASP-L61-1", "ASP-L61-2," and "ASP-L61-3" if containing active polymerase, or "ASP-L61-1 ProtK", "ASP-L61-2 ProtK" and "ASP-L61-3 ProtK" if containing inactive polymerase. The duplexes with active polymerase showed an average *T*m of 74 ˚C, which was 1.2 ˚C lower than the theoretical *T*m of 75.2 ˚C. The duplexes with inactive polymerase had an average of *T*m 74 ˚C, with "ASP-L61-2 ProtK" having no signal.

**Figure 12** shows the results of the triplicates using ASP-L25, which had a length difference of 36 nucleotides between the ASP and SP. The triplicate with active polymerase had an average *T*m of 73 ˚C, while the theoretical *T*m is 61.5 ˚C, which was a difference of 11.5 ˚C. The samples with inactive polymerase had one set of peaks at an average of 62.7 ˚C and a small peak at 73 ˚C; however, the intensity of the second peak was half of the first peak.

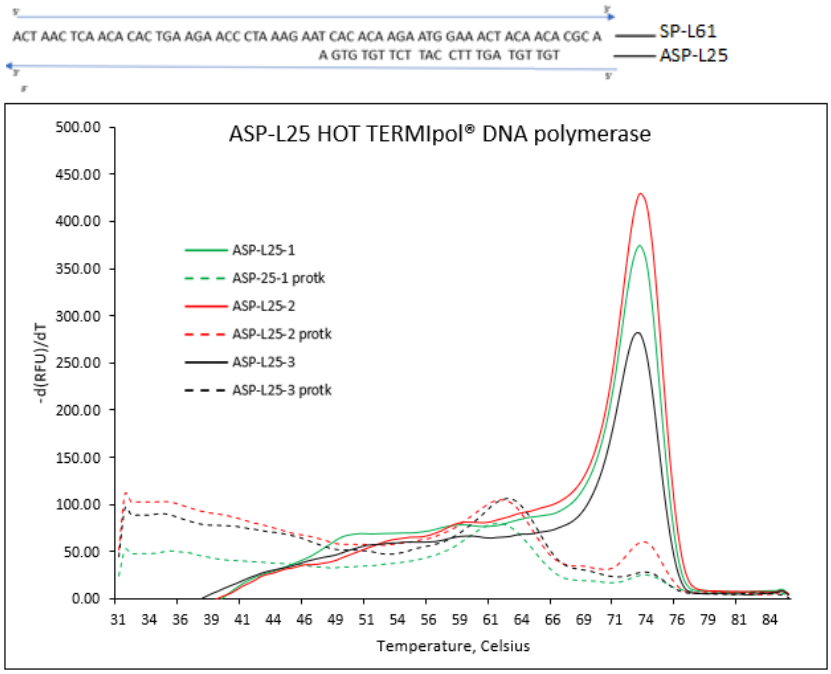



**Figure 12**: **Melting curve analysis of duplexes with a length difference of 36 nucleotides between the ASP and SP.** The polymerase used for these duplexes was HOT TERMIpol®. Half the samples were inactivated by proteinase K. The experiment was done in a triplicate indicated by "ASP-L25-1", "ASP-L25-2," and "ASP-L25-3" if containing active polymerase or "ASP-L25-1 ProtK", "ASP-L25-2 ProtK" and "ASP-L25-3 ProtK" if containing inactive polymerase. The triplicate with active polymerase showed an average *T*m of 73 ˚C, which was 11.5 ˚C higher than the theoretical *T*m of 61.5 ˚C. Samples containing Inactivated polymerase had an average *T*m of 62.7 ˚C.

**Figure 13** shows the analysis of ASP-L15. The theoretical *T*m was found to be 45.7 ˚C and was 27.3 ˚C lower than the average *T*m of 73 ˚C in the samples with the active polymerase. "ASP-L15-2 ProtK" had a high *T*m of 75 ˚C, while "ASP-L15-1 ProtK" and "ASP-L15-3 ProtK" had a *T*m of 44 ˚C and 43 ˚C, which was 29 ˚C and 30 ˚C lower than the samples with the active polymerase.

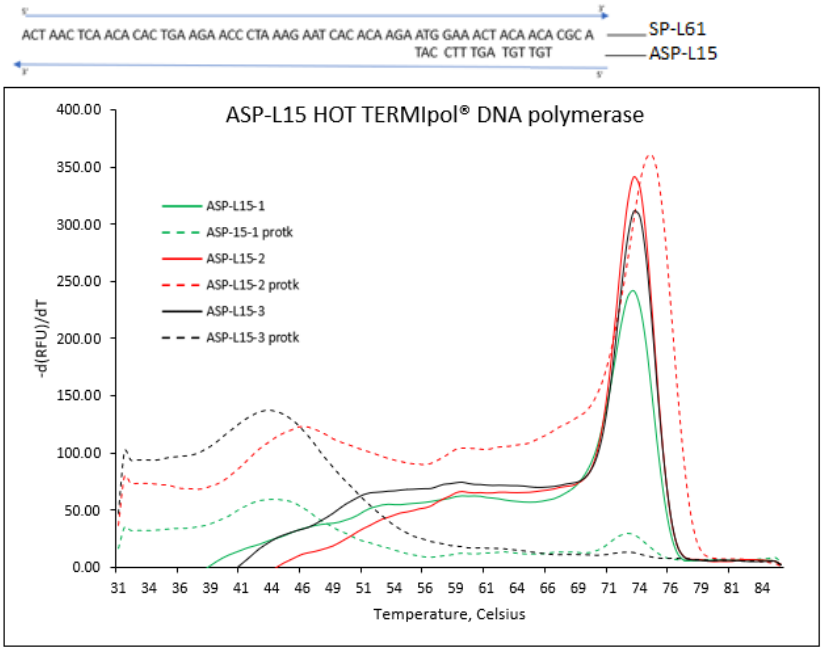



**Figure 13**: **Melting curve analysis of probe duplexes a length difference of 45 nucleotides between the ASP and the SP.** The polymerase used for these duplexes was HOT TERMIpol®. Half the samples were inactivated by proteinase K. The experiment was done in a triplicate indicated by "ASP-L15-1", "ASP-L15-2," and "ASP-L15-3" if containing active polymerase or "ASP-L15-1 ProtK", "ASP-L15-2 ProtK" and "ASP-L15-3 ProtK" if containing inactive polymerase. The duplexes with active polymerase showed an average *T*m of 73 ˚C, which was 27.3 ˚C higher than the theoretical *T*m of 45.7 ˚C. The samples containing inactivated polymerase had a large variance, where ASP-L15-1 had a *T*m of 44 ˚C, ASP15-2 had a *T*m of 75 ˚C, and ASP-L15-3 had a *T*m of 43 ˚C.

**Figure 14** represents the results of ASP-L15-3P. The duplexes with an active polymerase had the same *T*m as the samples with the inactive polymerase, at an average *T*m of 46 ˚C. While the Theoretical *T*m was 45.7 ˚C, which was a difference of 0.3 ˚C. With no difference between the *T*m of active and inactive polymerases in the samples.

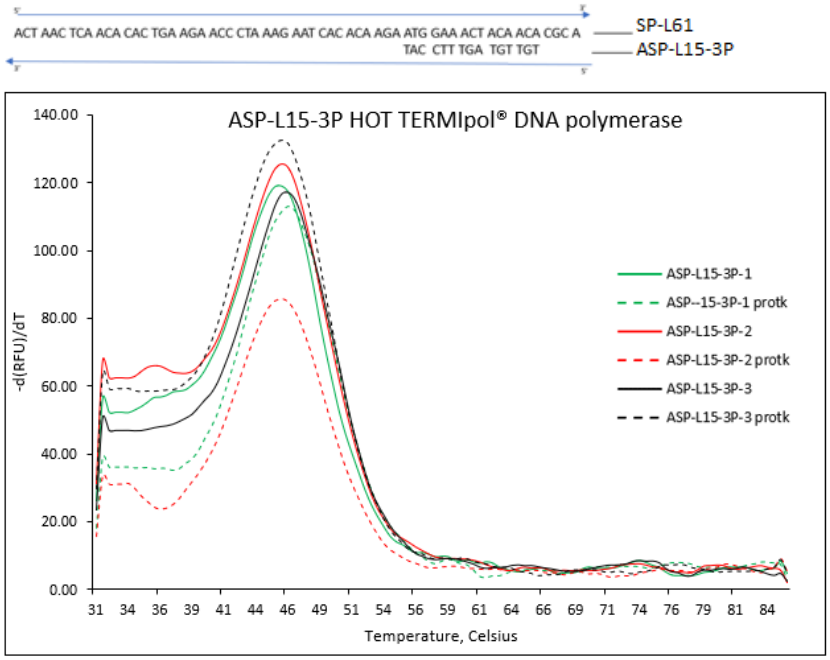



**Figure 14**: **Melting curve analysis of duplexes containing a 3'-phosphate on the ASP and with a length difference of 46 nucleotides between the ASP and SP.** The polymerase used for these probe sets was HOT TERMIpol®. Half the samples were inactivated by proteinase K. The experiment was done in a triplicate indicated by "ASP-L15-3P-1", "ASP-L15-3P-2," and "ASP-L15-3P-3" if containing active polymerase, or "ASP-L15-3P-1 ProtK", "ASP-L15.3P-2 ProtK" and "ASP-L15-3P-3 ProtK" if containing inactive polymerase. The triplicate with active polymerase showed an average *T*m of 46 ˚C, which was 0.3 ˚C higher than the theoretical *T*m of 45.7 ˚C. The samples inactivated by proteinase K had an average *T*m of 46 ˚C. The ASP had a 3'-phosphate group instead of a 3'-hydroxyl group.

**Figure 15,** investigates the effect of replacing the 3'-hydroxyl group on the ASP with a 3'-phosphate group, comparing the differences between duplexes with and without phosphate, namely ASP-L15 and ASP-L61. These duplexes were treated with HOT TERMIpol®, and the results shown in **Figure 15** do not contain any samples treated with proteinase K. The dotted lines represent the samples with 3'OH, while the solid graphs represent the probes with 3'P. Observed *T*ms can be found in **Table 16**.

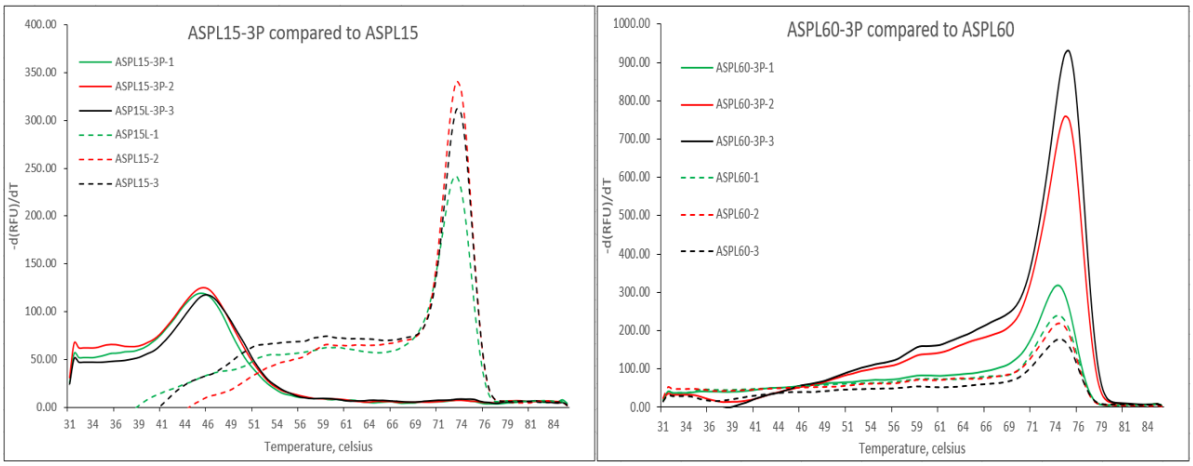



**Figure 15: A comparison between 3'OH and 3'-phosphate on the 3' end of the anti-sense primer**. The Figure shows four oligonucleotide DNA duplexes. ASP-L15-3P, ASP-L15, ASP-L61, and ASP-L61-3P. Both panels are demonstrating the difference in *T*m achieved when the ASP has a 3'-phosphate instead of a 3'-hydroxyl group. With panel A having a length of fifteen nucleotides on the ASP, while panel B has 61 nucleotides on the ASP. These graphs are gathered from **Figures 14,13,12 &11**; only the samples with active polymerase are presented.

## 4.4. Summary of the results

**Figure 16** shows a summary of the results in probe set 1. The values presented are the averages of the triplicates used in probe set 1. All samples that had an active polymerase showed a higher observed *T*m than predicted *T*m. The samples with an active polymerase, except SP-M0 with HOT TERMIpol®, showed a higher observed *T*m than the samples that had inactive polymerase.

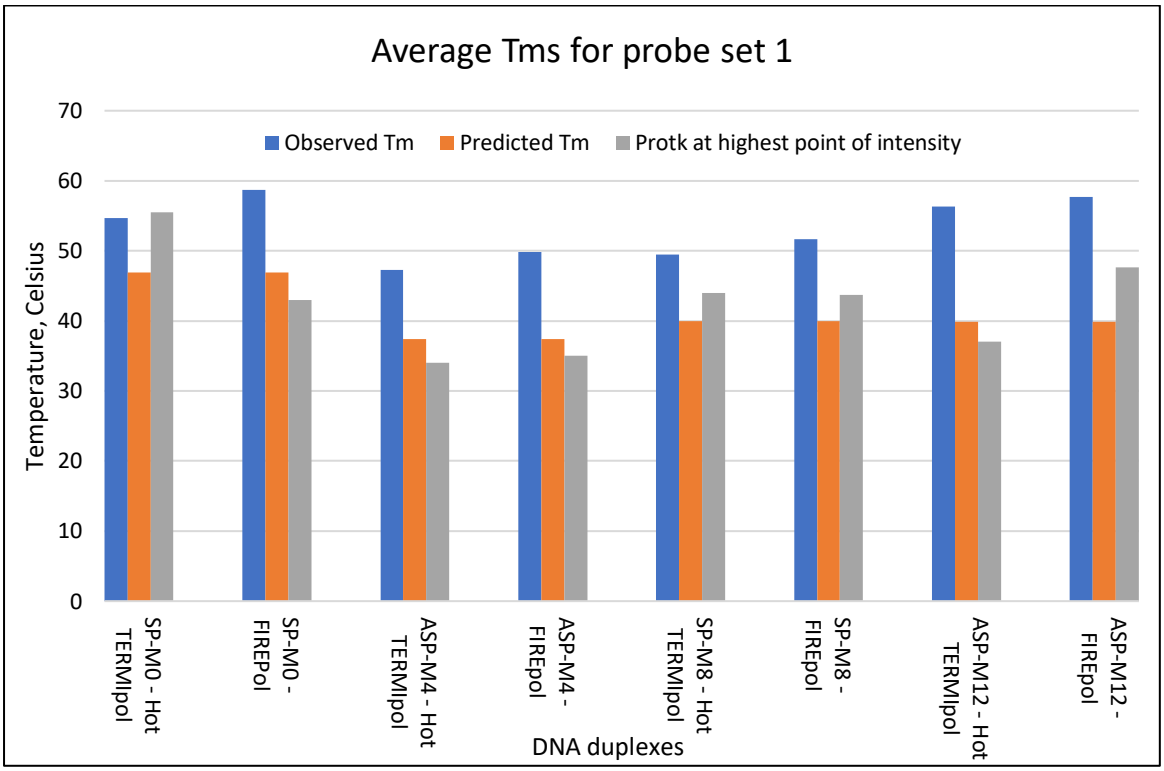



**Figure 16: Summary of average observed and predicted *T*ms in probe set 1.** The Figure shows SP-M0, SP-M8, ASP-M4 and ASP-M12, treated with HOT TERMIpol® and FIREpol® DNA polymerase.

**Figure 17** shows a summary of the results in probe set 2. The values presented are the averages of the triplicates used in probe set 2. The DNA duplexes that had a one nucleotide length difference between the SP and ASP (ASP-T64) showed no differences in the *T*ms. While the DNA duplexes with a ten-nucleotide length difference between the SP and ASP (ASP-T42) showed a higher observed *T*m in the samples with active DNA polymerase, compared to samples with inactive polymerase, it is also shown that the predicted *T*m is lower than the observed *T*m.

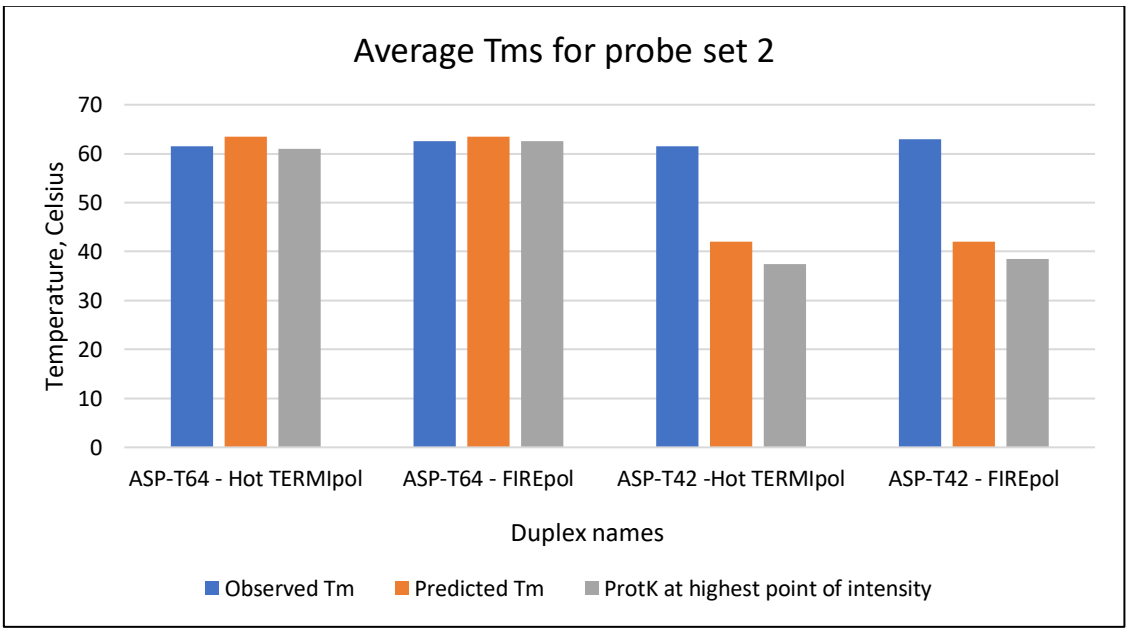



**Figure 17: Comparing duplexes with different lengths on the ASP.** The Figure shows ASP-T64 and ASP-T42 treated with either HOT TERMIpol® or FIREpol® DNA polymerase. This Figure is a summary of the results observed in **Figure 8 & 9**.

**Figure 18** shows a summary of the results in probe set 3. The values presented are the averages of the triplicates used in probe set 3. The duplexes with a 3'-phosphate group on the ASP, shows no changes in *T*m between the series. There was only observed a *T*m-shift when the duplexes had an ASP that was shorter than the SP.

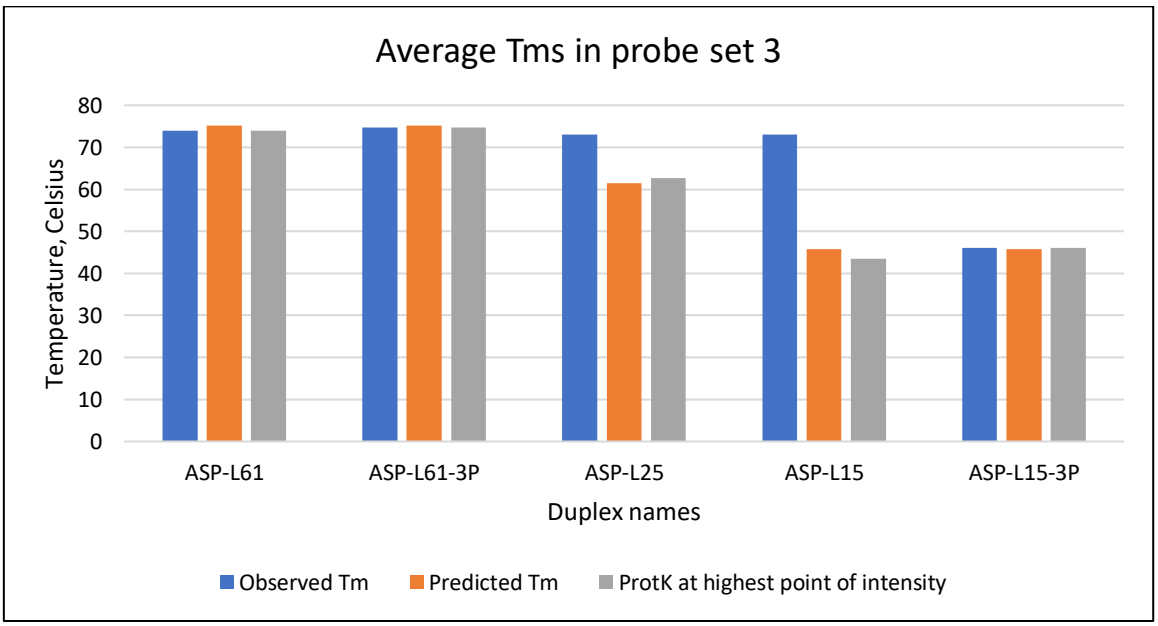



**Figure 18: Comparison of duplexes with a 3'-hydroxyl group and duplexes with a 3'-phosphate group on the ASP.** The Figure shows ASP-L61, ASP-L-61-3P, ASP-L25, ASP-L15 and ASP-L15-3P, treated with HOT TERMIpol® DNA polymerase. Figure 18 is a summary of **Figures 10, 11, 12, 13 & 14**.

# 5. Discussion

The first goal of the study was to establish that DNA polymerase could influence the thermostability of DNA duplexes. Then it compared the stabilizing effects between different DNA polymerases. The last goal was to investigate if the various states of the DNA duplexes could affect the stabilization ability of DNA polymerase. There were three main differences in the probes used in this study, represented in the three probe sets. All the probe sets were subjected to experiments with and without active polymerase; this was done to compare the *T*m of a DNA duplex without active polymerase and the same DNA duplex with an active polymerase to gain baseline information about the possible stabilizing effect of DNA polymerase.

The first probe set had a mismatch in different locations in the DNA duplex. This was done to get information on how mismatches could influence the stabilization effect provided by DNA polymerase and compare the stabilization achieved when the mismatches were in different positions on the duplex. Probe set 2 revolved more around the different length combinations that are possible in a DNA duplex, and how these combinations could affect the stabilization achieved. This probe set compared duplexes with an equal length between the ASP and SP, and duplexes where the ASP was significantly shorter than the SP. This was done to gain information on how the length of the different strands in the DNA duplex impacted the stabilizing effect of DNA polymerase. Probe set 3 was designed to look into length differences, just as probe set 2. Its secondary goal was to investigate how an end blocking 3'-phosphate on the ASP, would affect the possible stabilizing ability of a DNA polymerase. The duplexes with a 3'-phosphate group could then be compared to the duplexes with a hydroxyl group on the 3' end, effectively emulating the experiments with the proteinase K, with the advantage of not adding additional reagents in the solution. These probe sets were used to gain information that could confirm or refute the possible stabilization effect that DNA polymerase might have on a DNA duplex.

## 5.1. Comparing the effect of mismatches, length, and different polymerases

The ASP in probe set 1 was paired five nucleotides away from both the 3' end and the 5' end of the SP. This left a 5-nucleotide overhang from both ends, which leaves room for discussion of the effect these overhangs might have on the stabilizing effect from DNA polymerase on DNA duplexes. All the probes that were used in this section contained a mismatch in a unique position in the sequence, except for SP-M0, which did not contain a mismatch. The probes were tested with both active and

inactive polymerase to investigate the difference in *T*m of the DNA duplex, and the polymerase was inactivated by proteinase K. A summary of the results for probe set 1 is shown in **Figure 16,** including the observed *T*m and the predicted *T*m. The figure also shows the results acquired from the proteinase K treated samples. Due to the sporadic nature of these results, the temperature at the highest point of intensity is shown in the figure.

The result section for probe set 1 shows the melting curve analysis on ASP-M4, ASP-M12, SP-M0, and SP-M8. The probes had different mismatched base pairs, A-A, and G-G on different positions on the duplexes. They were individually treated with both HOT TERMIpol® DNA polymerase and FIREpol® DNA polymerase. When the ASP-M4 probe was combined with HOT TERMIpol® DNA polymerase in the melting curve analysis, the average *T*m was observed to be 47.3 ˚C, which was a Tm shift of 9.9 ˚C, when compared to the predicted *T*m (**Figure 3**). The proteinase K treated samples showed no clear peaks of intensity and started at a high intensity with a slight rise before it declined steadily until it reached the baseline. This can be interpreted as a steadier denaturing process, where the DNA duplexes denature at the same rate during the entire analysis until there are no more DNA duplexes. Compared to the samples with active polymerase where the intensity spikes and falls rapidly. Suggesting that the active polymerase might keep more of the DNA duplex stable until it reaches its melting point. It can also be observed in **Figure 16** that the highest point of intensity for proteinase K treated samples was a lot lower than for the samples with active polymerase, indicating that the *T*m of the proteinase K treated samples was a lot lower.

ASP-M12 showed the same tendencies and had a 16.4 ˚C lower predicted *T*m than observed *T*m. With the proteinase K treated samples being slightly different from ASP-M4, where they had a relatively flat graph from 31 ˚C and tendencies of peaks could be observed around 52-56 ˚C before a steady decline in intensity (**Figure 4**).

**Figure 5** shows the results for SP-M0 treated with HOT TERMIpol® DNA polymerase, this sequence did not contain any mismatches and had an average observed *T*m 7.8 ˚C higher than the predicted *T*m. Contrary to the previous tendencies, the proteinase K treated samples of SP-M0 showed an average *T*m of 55.66 ˚C, which was a higher *T*m than the samples with an active polymerase. There can be some explanations for these results, the first being that polymerase has not been able to stabilize these duplexes as much as the other samples in the probe set, with the major difference between these duplexes and the other probes, was that SP-M0 had no mismatches. However, since both probe set 2 and 3 did not contain mismatches and still showed stabilization, this is unlikely. There is also the possibility of a human error, which will be discussed in "technical issues and suggested improvements to the study." As the proteinase K treated samples had a higher *T*m than the predicted Tm, it is possible

that the presence of proteinase K or EvaGreen® had a stabilizing effect on the DNA duplex. Even though the proteinase K is inactivated, it is possible that its presence had a stabilizing effect on the DNA duplex. Intercalating dyes have been shown to be able to increase the *T*m of a DNA duplex (Mao et al., 2007). However, whether it is stabilizing or destabilizing depends on the agent, its position and if there are several agents interacting with the duplex, as described in the introduction (Marras et al., 2002; Moreira et al., 2005).

EvaGreen® has been observed to stabilize DNA duplexes, depending on the concentration of the dye. Firstly, if there is too little dye to saturate all the DNA duplexes, the EG from the melted DNA duplexes might interact with other DNA duplexes, which lacks the dye. Therefore, increasing the observed *T*m as they will emit more fluorescence when binding to a new duplex and make a change in the observed fluorescence drop, which can result in broad peaks in a melting curve analysis. It was reported that the *T*m shift caused by EG was at its smallest at 0.5X while otherwise increasing in order, from 0.25X, 1X, 2X, 5X, with a 2-3 ˚C difference between 0.5X and 5X. Further looking into the concentration of dye, it was found that 0.5X EG should be enough for 1 µM of a 15bp DNA duplex, in order to avoid unsaturated dye (Wang et al., 2016). It has been observed that at a high concentration of EvaGreen®, the *T*m of a DNA duplex can shift up to 10.7 ˚C. This was observed with a concentration of 10 µM of EvaGreen®, demonstrating that the effects EG has on the *T*m can be quite large (Radvanszky et al., 2015).

For probe set 1, there was used 1X EG and 1 µM combined with duplexes that have 14 base pairs, so following the research of Wang et al., there should be enough EG to reach saturation. As there was used 1X EG in this thesis, we should expect some stabilization of the DNA duplex from the EG. Considering the difference between 1X EG and 5X EG was 2-3 ˚C in (Wang et al., 2016), it is assumed, that we have not achieved stabilization at the level of a 10 ˚C *T*m shift in this study. Because of this stabilizing effect from EG, the comparison of *T*ms between samples with active polymerase and samples with inactive polymerase becomes even more important. Both these samples contained an equal amount of EG. The only difference between these samples was if they contained water (active polymerase) or proteinase K (inactive polymerase). In turn, meaning that the difference in *T*m, between samples with proteinase K and samples without, could be caused by a stabilizing effect from DNA polymerase.

As for SP-M8, the triplicates containing HOT TERMIpol® DNA polymerase showed an average observed *T*m 9.5 ˚C higher than the predicted Tm. The proteinase K treated samples had very broad peaks at the average *T*m of 45 ˚C (**Figure 6**). As mentioned above, broad peaks can be the result of a lack of dye, and they can also be a result of A-T areas melting before G-C areas, which can lead to EG binding

to the G-C-areas that are not yet denatured. As the EG dye binds to G-C, it would then again begin emitting fluorescence (Mao et al., 2007).

In general, the experiments with FIREpol® DNA polymerase have shown slightly higher observed *T*ms compared to the samples treated with HOT TERMIpol®. ASP-M4 treated with FIREpol® DNA polymerase showed a 12.4 ˚C higher average observed *T*m than the predicted *T*m (**Figure 3B**). The proteinase K treated samples had the highest intensity close to the beginning of the melting curve analysis at approximately 35 ˚C. There were no concrete peaks, and the analysis showed a steady decline until complete disassociation. Compared to its HOT TERMIpol® counterpart, there was a 2.7 ˚C difference in observed average temperature in samples with polymerase. While the proteinase K treated samples showed a similar trend in both **Figure 3A & B**. This can suggest that proteinase K affects both HOT TERMIpol® and FIREpol® DNA polymerase in the same manner.

For ASP-M12 treated with active FIREpol® DNA polymerase, it had a 17.8 ˚C difference between the average observed *T*m and predicted *T*m (**Figure 4B**). The proteinase K treated samples showed a large variance in the triplicate, with an average of 46.7 ˚C. However, one of the triplicates had a high *T*m compared to the others. The *T*m of this sample was approximately the same as in the samples with an active polymerase. This might be due to a human error, where the proteinase K was not properly added, variance in the qPCR machine or the proteinase K has not been properly activated. Even so, the average *T*m for the proteinase K treated samples were 11.3 ˚C lower than the average of samples with active FIRE pol DNA polymerase, as mentioned above, this can be indicative of DNA polymerase having a stabilizing effect on the DNA duplex. The difference between the predicted *T*m and the average *T*m of samples without active polymerase was also quite large at 6.8 ˚C, which might represent a stabilizing effect from EG.

SP-M0, when treated with FIREpol® DNA polymerase, showed a very large variance in both *T*m and intensity. With SP-M0-3 showing no intensity, it is very likely that an error in the experiment has occurred. It might be that EvaGreen® was not added or that no DNA duplexes were added to the solution. As shown in **Figure 5B,** there is a large variance in the intensity of the remaining peaks, with one of the triplicates having an intensity nearly ten times higher than the other. Even though there was a large variance, the approach taken was to treat these results as the rest of the samples and use the average *T*m of 59 ˚C. The average observed temperature for the proteinase K treated samples was 42.5 ˚C and were 16.5 ˚C lower than the observed *T*m in samples with active FIREpol®. The *T*m for proteinase K treated samples were lower than the predicted *T*m of 46.9 ˚C, but due to the broad peaks and linear behavior of the graphs, it was not easy to determine if 42.5 ˚C was the true melting point of the sample. If it was, these results would suggest that EG does not affect the stability of the DNA

duplex positively in all cases, as it has been observed to both increases the *T*m in ASP-M12 and decrease the *T*m in SP-M0. As the proteinase K treated samples had a lower *T*m than the samples with active FIREpol®, this is further proof that FIREpol® DNA polymerase might have a stabilizing effect on the DNA duplex.

ASP-M4, ASP-M12, SP-M0, and SP-M8 (**Figure 16**), all have shown to have a higher observed *T*m when treated with FIREpol® compared to HOT TERMIpol® DNA polymerase, in the range of 1.7-4.3 ˚C (**Figures 3,4,5 & 6**), which was a small, but noticeable increase in *T*m. With SP-M8 having an 11.7 ˚C higher *T*m than the predicted *T*m, it was shown that all the duplexes represented in the results had a higher observed *T*m than the predicted *T*m in a range of approximately 8-18 ˚C. As shown in **Figure 6B,** the SP-M8 probe also had the clearest results for proteinase K treated samples, the melting curve analysis still shows broad peaks, but they were much more similar in shape compared to its counterpart with active DNA polymerase. The average observed Tm for proteinase K treated samples were 43.7 ˚C, and with a smaller variance in both intensity and *T*m, it was easier to compare with the predicted *T*m of 40 ˚C. The perfect scenario in this study is that proteinase K treated Tm and predicted *T*m is the same, in order to verify that it is, in fact, DNA polymerase that is responsible for the *T*m-shifts.

For most of the experiments in probe set 1, there can be seen a difference between the predicted *T*m and the observed *T*m. There are differences between proteinase K treated samples and samples with active polymerase, both in intensity readings and in melting temperatures. The different probes have varied results, SP-M8 showed a smaller variance in all triplicates, in both DNA polymerase treated samples and proteinase K treated samples (**Figure 6**). While proteinase K treated ASP-M4, ASP-M12 and SP-M0 showed smaller intensity peaks than their counterparts; the peaks were often very broad if to be considered peaks at all. The lesser amount of intensity observed might be an interaction between proteinase K and EG. It might be the result of incomplete inactivation of proteinase K, causing the proteinase K to degrade the EG, or it could be that the presence of the inactivated proteinase K in the solution has a quenching effect on the EG. However, all the probes except SP-M0 show a higher tangible *T*m in the samples with active polymerase, opposed to the samples with the inactive polymerase (**Figure 5**). This provides credibility to the hypothesis of DNA polymerase having a stabilizing effect on DNA duplexes, as both the samples with active and the samples with inactive polymerase had the same reagents, except proteinase K, which should be inactivated. However, SP-M0 showed a slightly higher *T*m in the samples without active polymerase. This can be interpreted as several things, such as the proteinase K was not properly added, the inactivated proteinase K might influence the solution, or it could be a well to well variation in the qPCR, due to unequal distribution of temperature on the plate. Due to the limited sample size, it is difficult to conclude what was the

reason for the irregularity. The predicted *T*ms were always lower than the observed *T*ms in samples with active polymerase, but they are hard to compare to the samples treated with proteinase K in probe set 1, as the behavior of proteinase K treated samples in the melting curve analysis is very variable.

As mentioned above, all the samples with active polymerase showed a higher *T*m than predicted, even though the $Mg^{2+}$ concentration and the duplex' sequence were the same in the prediction tool. However, the prediction tool does not allow *T*m calculations for duplexes containing mismatches, without at least 1 mM $Na^{+}$ as a parameter, meaning that there is a difference in the calculation of the theoretical *T*m and the experimental setup. However, higher salt concentration results in higher DNA duplex stability (Tan & Chen, 2006), which leads to a higher *T*m. This means that the predicted *T*m should, in theory, be a bit lower than presented. Suggesting that there should be an even larger difference between the observed *T*m and the predicted Tm. The difference in predicted *T*m and observed *T*m shows that there are other factors influencing the stability of the DNA duplex other than oligonucleotide concentration, sequence, and salt concentration, which are the parameters in the prediction tool. Indicating that some of the reagents in the experimental setup is not considered in the prediction. Additional reagents added to the experiment that is not in the prediction tool is the intercalating agent EG, and DNA polymerase. As mentioned above, both EG and DNA polymerase has shown stabilizing effects on DNA duplexes in previous studies.

The mismatches in probe set 1 has provided some interesting results. As shown in **Table 13** and **Table 14,** there is some correlation between the positioning of the mismatches in the sequences and the difference in the observed and predicted *T*ms. As mentioned above, samples with FIREpol® show a slightly higher observed *T*m than samples with HOT TERMIpol®. Nevertheless, all samples show the same trend when comparing the positions of the mismatches. For clarity, only results from samples treated with HOT TERMIpol® will be further discussed; all results for FIREpol® treated samples are found in **Table 14**. **Figure 7** shows that the mismatch positions in the DNA duplexes are clustered in three clusters, with SP-M1 and SP-M0 represented as individual graphs. They are excluded from the clustering because SP-M0 contains zero mismatches and works as a baseline for the duplex, while SP-M1 has a much lower observed *T*m than the rest of the duplexes. For SP-M1, the predicted *T*m was 45.2 ˚C which was 6.2 ˚C higher than the observed *T*m. The predicted *T*m might be an overestimation from oligoanalyzer, as all the triplicates showed an average observed *T*m of 38.9 ˚C. If the mismatched base pair is removed from the calculation in oligoanalyzer, the predicted temperature is 38.5 ˚C, much closer to the observed value. Further confirmed by having a similar observed value in the samples containing FIREpol®.

The first cluster observed in **Figure 7** contains duplexes with a mismatch in position 2-8, from the 3' end of the ASP. Here it was observed an average difference between the observed and predicted *T*m of 10.4 ˚C. While the second cluster was the mismatch positions 9 and 10, and had an average difference of 14.7 ˚C. The third cluster contained a mismatch in the positions 11-13 and had an average difference of 16.7 ˚C (**Figure 7**). These results showed a correlation between the placement of the mismatch and the difference between predicted and observed *T*m, which means that the stabilization effect from DNA polymerase is affected by mismatches in different locations on the DNA strand. DNA polymerase binds at the 3'OH on the primer strand, which would be the ASP in this probe set, as the SP has a 5-nucleotide overhang on both sides of the ASP. This means that the SP functions as a template strand, which is also required for DNA polymerase to function (Steitz, 1999). This might be a reason for the increase in *T*m; the further away the mismatch is located from the 3' end. With SP-M1 showing no difference in observed and predicted *T*m, if the mismatch is omitted from the calculation (**Figure 7**). Suggesting that the DNA polymerase is unable to achieve a stabilizing effect on the duplex if there is a mismatch located on the 3' end of the ASP. Even though the *T*m difference is increasing the further away the mismatch is from the 3' end, ASP-M14 showed a much lower difference in *T*m compared to positions 9-13 (**Figure 7**). However, as this was at the end of the duplex, the prediction tool has been shown to overestimate the predicted Tm, with the base pair removed from the calculation the predicted *T*m is 42.4 ˚C. Making the difference much closer to the average value of the cluster 11-13, with a *T*m difference of 15.8 ˚C.

These results show that DNA duplexes treated with active polymerase has in general, a higher *T*m than the duplexes treated with an inactive polymerase. The location of a mismatch on a duplex will also affect the stabilization achieved from DNA polymerase, when the mismatch is located further away from the 3' end, the DNA duplex will be more stable.

## 5.2. Comparing the effect of length differences between the SP and ASP with different polymerases

Probe set 2 had the same experimental setup as probe set 1, the probes did not have mismatches, but compared the *T*m between duplexes that had ASPs with different lengths. This section compares two duplexes, ASP-T42 and ASP-T64. ASP-T42 had an ASP with 14 nucleotides and an SP with 24 nucleotides, and ASP-T64 had an ASP with 23 nucleotides, and an SP that had 24 nucleotides (**Figures 8 & 9**). Both duplexes were treated with HOT TERMIpol® and FIREpol® DNA polymerase. Where the most interesting observation was that the stabilizing effect was dependent on whether the ASP had a similar length as the SP or not, this is summarized in **Figure 17.**

The ASP-T42 duplex had an average *T*m of 61.5 ˚C with HOT TERMIpol® DNA polymerase and an average *T*m of 63 ˚C with FIREpol® DNA polymerase (**Figure 8**). Matching the observations in probe set 1, with FIREpol® DNA polymerase having a slightly higher *T*m than samples treated with HOT TERMIpol® DNA polymerase (**Figure 17**). The proteinase K treated samples were showing peaks at 37.5 ˚C for HOT TERMIpol® DNA polymerase and 38.5 ˚C for FIREpol® DNA polymerase, which was lower than the predicted melting points for the probe, at 42 ˚C. Again showing that the stabilization effect from EG was not consistent for all the duplexes, and it might be destabilizing as well as stabilizing. However, to be sure of this, one would need to have an experiment where the *T*m of the duplexes was tested with and without EG, which makes it necessary to use a different method than a qPCR melting curve analysis. One example is using a UV absorbance measurement method. This was also investigated in probe set 3, where duplexes with an end blocking 3'-phosphate group were compared to duplexes with a 3'-hydroxyl group on the ASP.

The difference between proteinase K treated samples and samples with active polymerase is very indicative of the DNA polymerase having a stabilizing effect on the DNA duplex (**Figure 17**), even without mismatches present. Suggesting that DNA polymerase can have an impact on the *T*m of the DNA duplex, when there is a length difference between the ASP and SP, with the ASP having an overhang over the 3' end.

While in ASP-T64, both proteinase K treated samples and polymerase treated samples showed nearly the same Tm's in both the duplicates. The predicted *T*m was also approximately the same as the observed temperatures (**Figure 17**). The difference between ASP-T64 and the previous duplexes was that the ASP and SP have nearly the same length, and that that the one nucleotide overhang was over the 5' end of the ASP. This suggests that the length of the SP and ASP is very relevant for the stabilization effect of DNA polymerase. In order for the DNA polymerase to add dNTPs, the DNA polymerase requires a primer with a free 3' end that is paired with the template strand in order to work (Stryer et al., 2002). The correlation between a shorter ASP and a larger *T*m shift was also observed by Hiseni (2016), where an SP of 61 nucleotides was used, which were added to ASPs with lengths ranging from 17 to 61 nucleotides. Where the shortest ASP, with 17 nucleotides, showed a *T*m shift of 7.5 ˚C and the longest ASP of 61 nucleotides had a *T*m shift of 2 ˚C (Hiseni, 2016).

As already shown in probe set 1, when there is an overhang over the 3' end of the ASP, a stabilization effect is observed, which can mean that the length difference between the strands in the duplex is a factor that can allow the DNA polymerase to stabilize the DNA duplex. Compared to the duplexes in probe set 1, the proteinase K treated samples in probe set 2 (**Figures 8 & 9**) had much more defined peaks in the melting curve analysis. There were three differences between these probe sets. With the

first being that probe set 1s ASP was bound to the SP exactly in the middle, with a 5 nucleotide overhang on both ends, while ASP-T42 was bound to its SP on the 3' end of the SP, with a 1 nucleotide overhang on the 3' end of the SP and a 9 nucleotide overhang on the 5' end. The second difference was the sequence of the probes, as mentioned in the introduction, base-stacking and base-pairing influences the *T*m of the duplex, so the melting curve differences between the probes could be the result of different thermostability caused by the mismatches. There were also the mismatches present in the first probe set, which might have influenced the behavior of the duplexes in the melting curve analysis.

## 5.3. The effect of a 3'-phosphate group and length differences

Probe set 3 was treated with only HOT TERMIpol® DNA polymerase and showed some promising results. The 3P on the end of the probe names symbolizes a 3'-phosphate group on the end of the ASP. This phosphate blocks the DNA polymerase from binding to the DNA duplex, thus making it possible to compare the same probes with and without polymerase without the need of proteinase K. The probes have still been treated with proteinase K, in order to get more data. A summary of the results from probe set 3 is displayed in **Figure 18**.

As mentioned in the discussion about probe set 2, the length of the ASP compared to SP will influence the stabilization effect seen from DNA polymerase. ASP-L61 and ASP-L61-3P are represented in **Figures 11 & 10**, respectively, showing that samples with active polymerase did not have any difference in *T*m compared to the samples without active polymerase, demonstrating that DNA polymerase did not stabilize the DNA duplex. Further confirmed by comparing the ASP-L61 duplex and the ASP-L61-3P, both with an active polymerase. If there were stabilization, it should be present in the ASP-L61 duplex, but not in the ASP-L61-3P duplex.

On the contrary, this can be observed by looking at ASP-L15 and ASP-L15-3P, where they had a *T*m of 73 ˚C and 46 ˚C, respectively, when treated with an active polymerase (**Figure 15**). By using a DNA duplex with a 3'-phosphate group on the ASP, DNA polymerase should be unable to bind to the duplex. The first step in the polymerization catalysis is the deprotonation of the 3'-hydroxyl group, allowing it to attack the phosphate of the incoming dNTPs. It was also observed in **Figure 13** that the proteinase K treated samples of ASP-L15 had a *T*m of 43.5 ˚C, except for ASP-L15-2-ProtK, that had a *T*m of 75 ˚C. This is most likely caused by a human error, which would be solved by having a larger sample size, which will be discussed in "technical issues and suggested improvements." This *T*m of 43.5 ˚C was 2.5

˚C lower than the average *T*m observed in ASP-L15-3P with an active polymerase. This makes a good comparison point for the proteinase K treated samples, showing that the *T*m of these samples can be quite accurate and can be used to investigate the stabilizing effects of DNA polymerase.

## 5.4. Technical issues and suggested improvements

The study of the stabilizing effects of DNA polymerase has shown some exciting results. Though, there have been some suspected technical issues, mainly human errors in some of the melting curve analyses and some problems with the prediction of the *T*m of certain duplexes. This was observed in probe set 1, regarding SP-M0 (**Figure 5**) in samples with both, active and inactivated polymerase. As well as the prediction of the *T*m in the duplexes that had a mismatch on the ends of the duplex and the prediction tool not allowing the use of the exact parameters that were used in the experiment. In order to mitigate some of these issues, some modifications have been made in the prediction tool. There are also several suggested improvements that should be implemented if the research is taken further. This is to add a higher degree of credibility and accuracy in the study. These improvements will also have to be done in order to quantify the possible stabilizing effect of DNA polymerase.

### 5.4.1. Technical issues

The technical issues of this study revolve mainly around the first probe set. When SP-M0 was treated with HOT TERMIpol® DNA polymerase, the proteinase K treated samples showed a *T*m that was 7.8 ˚C higher than the predicted temperature, and it was almost the same as the *T*m observed with an active polymerase (**Figure 5**). In addition to the suggestions mentioned above, the reason for this irregular result might be a human error, in which the proteinase K has not been added to the samples. The same reasoning can be applied to the lack of signal in FIREpol® treated SP-M0-3, where the EvaGreen® might not have been added properly.

Then there are the issues with the prediction tool for the duplexes with mismatches. As the prediction tool does not allow calculation without a minimum of 1 mM $Na^+$, the predicted *T*ms might be slightly overestimated. This can result in a higher predicted *T*m and would result in an underestimation of the stabilizing effects of the DNA polymerase when comparing the observed *T*m with the predicted Tm.

The *T*m prediction of DNA duplexes with a mismatch on either end of the ASP is very different from what is observed in the rest of the samples. With SP-M1 showing a much higher predicted *T*m than the observed Tm. While ASP-L14 is showing a much lower difference between the predicted and

observed Tm, than the trend of the other duplexes, where the difference has been shown to increase the further away the mismatch is from the 3' end of the ASP. Removing the mismatched bases from the calculation yields a predicted *T*m much closer to the observed result. As this gave a result closer to reality, there predicted *T*m for these duplexes were calculated without the mismatched bases.

### 5.4.2. Suggested improvements of the study

There are several different suggestions that could improve this study. Many of the technical issues in probe set 1 could be solved by having a larger sample size. This would help to rule out human errors, as the irregular results could be compared with several triplicates of the same probes. A benefit of a larger sample size would also be the possibility of quantifying the stabilization effects of DNA polymerase. This would enable quantifying the effect by finding the true average stabilization effect and the variance of this effect, depending on different factors in a DNA duplex, such as mismatches, length differences and so on. Effectively finding the specific effect the polymerase would have on the DNA duplex, maybe even allowing a prediction of *T*m with different polymerases as a parameter in the future.

As seen in the discussion, this study has a lot of factors that might interfere with the *T*m of the DNA duplex. The first of these factors is proteinase K, where the role of proteinase K is to inactivate the polymerase, the proteinase K is then inactivated by its own step in the experimental setup. However, an experiment with only proteinase K should have been done in order to see the effect that it would have on the stability of the DNA duplex, without DNA polymerase present in the solution. Allowing the comparison of samples with proteinase K in the presence or absence of polymerase, which would show if the proteinase K influences the stability of the DNA duplex independent of the polymerase.

As well as running an experiment with neither polymerase nor proteinase K, in order to see the accuracy of the predicted *T*ms, as many of the results and the conclusions that can be drawn from these results are dependent on its accuracy. This is also the case for EvaGreen®. As mentioned previously in the discussion, EG can influence the stability of DNA duplexes. This is somewhat explored in probe set 3, by comparing the probes with a 3'-phosphate and the probes with a 3'-hydroxyl group, which shows the difference in *T*m when both samples have EvaGreen®. For probe set 1 and 2, there is no such comparison, therefore running a sample with only EvaGreen® would be beneficial to see what stabilizing effects it might represent, when compared to the predicted Tm. For probe set 1, it would also be beneficial to add 1mM of $Na^+$ to this experiment, in order to have the same parameters as used to predict the Tm.

As mentioned earlier in the discussion, samples with FIREpol® DNA polymerase showed slightly higher *T*ms than samples treated with HOT TERMIpol®. This suggests that DNA polymerases have different stabilizing effects. However, the samples with HOT TERMIpol® used C buffer, and the samples treated with FIREpol® used B2 buffer, which might influence the results. Testing different polymerases, which requires the same buffer would allow the comparison of *T*m without the buffers being a potential source of error. As well as testing several polymerases would be necessary if the goal of increasing the quality of probe design should be met.

# 6. Conclusion and future perspectives

The results show that DNA polymerase has a definite effect on the DNA duplex. In all probe sets, there was a difference between samples that had active polymerase and samples that had inactive polymerase, both the *T*m of the duplex and the behavior of the duplex in the melting process was affected. As discussed above, the effects proteinase K and EvaGreen® have on the DNA duplex stability has not been investigated. However, the experiment with ASP-L15 in probe set 3 with and without an end blocking phosphate showed a *T*m-shift. Confirming that there is a stabilizing effect received from DNA polymerase, as there is no proteinase K involved and both samples have equal amounts of EG.

The difference between the DNA polymerases was noticeable in nearly all samples. The FIREpol® DNA polymerase has a slightly higher average observed *T*m, compared to the samples with HOT TERMIpol® DNA polymerase. Showing that the stability provided by polymerases is variable depending on the polymerase used if the use of different buffers is disregarded.

The suggestion that DNA polymerase can increase the stability of DNA duplexes is further demonstrated by the comparison of the theoretical *T*m and the observed *T*m where certain conditions are met. In all samples where the ASP was shorter than the SP, the predicted *T*m was lower than the observed *T*m. Except for SP-M1, which had a mismatch on the 3' end of the ASP. Which leads to the discussion of what impact mismatches have on the *T*m of the DNA duplex. Where the position of the mismatch has been shown to influence the stabilization effect from DNA polymerase. The further the mismatch is from the 3' end of the ASP, the difference between predicted *T*m and observed *T*m increases. Except for the mismatch located at the 5' end of the ASP, which might be due to the prediction tool, but without testing, it is not possible to conclude that it should have the highest difference.

All these results lean toward a conclusion that FIREpol®- and HOT TERMIpol® DNA polymerase has a stabilizing effect on the DNA duplexes. However, certain conditions need to be met for the polymerase to exhibit such an effect. The ASP needs to be shorter than the SP, and according to these results, the shorter it is compared to the SP, the greater the *T*m shift. As well as mismatches at the 3' end of the ASP will inhibit the stabilization effect, and the duplex cannot contain an end blocking 3' phosphate at the 3' end of the ASP if stabilization by the DNA polymerase is wanted.

# 7. References


Aboul-Ela, F., Koh, D., Tinoco Jr, I. & Martin, F. H. (1985). Base-base mismatches. Thermodynamics of double helix formation for dCA3XA3G+ dCT3YT3G (X, Y= A, C, G, D. *Nucleic acids research*, 13 (13): 4811-4824.

Bansal, M. (2003). DNA structure: Revisiting the Watson–Crick. *Current Science*, 85 (11): 557.

Bjorndal, M. T. & Fygenson, D. K. (2002). DNA melting in the presence of fluorescent intercalating oxazole yellow dyes measured with a gel-based assay. *Biopolymers: Original Research on Biomolecules*, 65 (1): 40-44.

Brakmann, S. (2005). Directed evolution as a tool for understanding and optimizing nucleic acid polymerase function. *Cellular and molecular life sciences*, 62 (22): 2634.

Bren, U., Lah, J., Bren, M., Martínek, V. & Florián, J. (2010). DNA duplex stability: the role of preorganized electrostatics. *The Journal of Physical Chemistry B*, 114 (8): 2876-2885.

Brown, T. & Brown Jr, T. (2015). DNA Duplex Stability. *Nucleic Acids Book. Np: np, nd N. pag. ATDBio. ATDBio Ltd. Web*, 5.

Chien, A., Edgar, D. B. & Trela, J. M. (1976). Deoxyribonucleic acid polymerase from the extreme thermophile Thermus aquaticus. *Journal of bacteriology*, 127 (3): 1550-1557.

Fogolari, F., Brigo, A. & Molinari, H. (2002). The Poisson–Boltzmann equation for biomolecular electrostatics: a tool for structural biology. *Journal of Molecular Recognition*, 15 (6): 377-392.

Gu, X.-B., Nakano, S.-i. & Sugimoto, N. (2007). Consecutive GC base pairs determine the energy barrier of DNA duplex formation under molecularly crowded conditions. *Chemical communications* (26): 2750-2752.

Hiseni, P. (2016). *Liquid Array Diagnostics (LAD)-based multiplexed genotyping tests for double-muscling mutations in beef cattle* Hedmark university.

Hiseni, P., Wilson, R. C., Storrø, O., Johnsen, R., Øien, T. & Rudi, K. (2019). Liquid array diagnostics: a novel method for rapid detection of microbial communities in single-tube multiplex reactions. *BioTechniques*, 66 (3): 143-149.

Hunter, C. A. (1993). Sequence-dependent DNA structure: the role of base stacking interactions. *Journal of molecular biology*, 230 (3): 1025-1054.

Integrated DNA Technologies. (2019). Available at: https://eu.idtdna.com/calc/Analyzer/Home/definitions.

Kermekchiev, M. B., Tzekov, A. & Barnes, W. M. (2003). Cold-sensitive mutants of Taq DNA polymerase provide a hot start for PCR. *Nucleic acids research*, 31 (21): 6139-6147.

Leal, N. A., Sukeda, M. & Benner, S. A. (2006). Dynamic assembly of primers on nucleic acid templates. *Nucleic acids research*, 34 (17): 4702-4710.

Li, Y., Korolev, S. & Waksman, G. (1998). Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of Thermus aquaticus DNA polymerase I: structural basis for nucleotide incorporation. *The EMBO journal*, 17 (24): 7514-7525.

Mao, F., Leung, W.-Y. & Xin, X. (2007). Characterization of EvaGreen and the implication of its physicochemical properties for qPCR applications. *BMC biotechnology*, 7 (1): 76.

Marras, S. A., Kramer, F. R. & Tyagi, S. (2002). Efficiencies of fluorescence resonance energy transfer and contact-mediated quenching in oligonucleotide probes. *Nucleic acids research*, 30 (21): e122-e122.

Mineno, J., Ishino, Y., Ohminami, T. & Kato, I. (1993). Fluorescent labeling of a DNA sequencing primer. *DNA Sequence*, 4 (3): 135-141.

Mo, Y. (2006). Probing the nature of hydrogen bonds in DNA base pairs. *Journal of molecular modeling*, 12 (5): 665-672.

Moreira, B. G., You, Y., Behlke, M. A. & Owczarzy, R. (2005). Effects of fluorescent dyes, quenchers, and dangling ends on DNA duplex stability. *Biochemical and biophysical research communications*, 327 (2): 473-484.

Morrison, L. E. & Stols, L. M. (1993). Sensitive fluorescence-based thermodynamic and kinetic measurements of DNA hybridization in solution. *Biochemistry*, 32 (12): 3095-3104.
Patel, P. H. & Loeb, L. A. (2001). Getting a grip on how DNA polymerases function. *Nature structural biology*, 8 (8): 656-659.
Petersheim, M. & Turner, D. H. (1983). Base-stacking and base-pairing contributions to helix stability: thermodynamics of double-helix formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp. *Biochemistry*, 22 (2): 256-263.
Radvanszky, J., Surovy, M., Nagyova, E., Minarik, G. & Kadasi, L. (2015). Comparison of different DNA binding fluorescent dyes for applications of high-resolution melting analysis. *Clinical biochemistry*, 48 (9): 609-616.
Rauzan, B., McMichael, E., Cave, R., Sevcik, L. R., Ostrosky, K., Whitman, E., Stegemann, R., Sinclair, A. L., Serra, M. J. & Deckert, A. A. (2013). Kinetics and thermodynamics of DNA, RNA, and hybrid duplex formation. *Biochemistry*, 52 (5): 765-772.
Record, M. T. & Lohman, T. M. (1978). A semiempirical extension of polyelectrolyte theory to the treatment of oligoelectrolytes: Application to oligonucleotide helix-coil transitions. *Biopolymers: Original Research on Biomolecules*, 17 (1): 159-166.
Reed, G. H., Kent, J. O. & Wittwer, C. T. (2007). High-resolution DNA melting analysis for simple and efficient molecular diagnostics.
Rouzina, I. & Bloomfield, V. A. (1999). Heat capacity effects on the melting of DNA. 1. General aspects. *Biophysical journal*, 77 (6): 3242-3251.
Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, 239 (4839): 487-491.
SantaLucia, J., Allawi, H. T. & Seneviratne, P. A. (1996). Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry*, 35 (11): 3555-3562.
SantaLucia, J. (1998). A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proceedings of the National Academy of Sciences*, 95 (4): 1460-1465.
SantaLucia Jr, J. & Hicks, D. (2004). The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.*, 33: 415-440.
Schildkraut, C. L., Marmur, J. & Doty, P. (1962). Determination of the base composition of deoxyribonucleic acid from its buoyant density in CsCl. *Journal of molecular biology*, 4 (6): 430-443.
Solis BioDyne. (2020). *FIREPol® DNA Polymerase*. Available at: https://www.solisbiodyne.com/EN/product/name=FIREpol-DNA-Polymerase
Starikov, E. B. & Nordén, B. (2009). DNA Duplex Length and Salt Concentration Dependence of Enthalpy− Entropy Compensation Parameters for DNA Melting. *The Journal of Physical Chemistry B*, 113 (33): 11375-11377.
Steitz, T. A. (1999). DNA polymerases: structural diversity and common mechanisms. *Journal of Biological Chemistry*, 274 (25): 17395-17398.
Stryer, L., Berg, J. & Tymoczko, J. (2002). DNA polymerases require a template and a primer. In: WH Freeman.
Tan, Z.-J. & Chen, S.-J. (2006). Nucleic acid helix stability: effects of salt concentration, cation valence and size, and chain length. *Biophysical journal*, 90 (4): 1175-1190.
Wang, J., Pan, X. & Liang, X. (2016). Assessment for melting temperature measurement of nucleic acid by HRM. *Journal of analytical methods in chemistry*, 2016.
Wu, P., Nakano, S. i. & Sugimoto, N. (2002). Temperature dependence of thermodynamic properties for DNA/DNA and RNA/DNA duplex formation. *European journal of biochemistry*, 269 (12): 2821-2830.
Yakovchuk, P., Protozanova, E. & Frank-Kamenetskii, M. D. (2006). Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic acids research*, 34 (2): 564-574.

# 8. Appendices

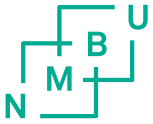
## Appendix A

**Table A1. Overview of all probes used in the study with the corresponding sequences of the probes and the abbreviation of the probe name. The probes are arranged in three different probe sets, with probe set 1,2, and 3.**

| Probe set | Probe name | Abbreviation | Sequence 5'-3' |
|---|---|---|---|
| **1** | Sense probe 0 | SP-M0 | AGACG TGC ATT TCA CTA AA GCAGA |
| **1** | Sense probe 1 | SP-M1 | AGACG AGC ATT TCA CTA AA GCAGA |
| **1** | Sense probe 3 | SP-M3 | AGACG TGG ATT TCA CTA AA GCAGA |
| **1** | Sense probe 5 | SP-M5 | AGACG TGC AAT TCA CTA AA GCAGA |
| **1** | Sense probe 6 | SP-M6 | AGACG TGC ATA TCA CTA AA GCAGA |
| **1** | Sense probe 7 | SP-M7 | AGACG TGC ATT ACA CTA AA GCAGA |
| **1** | Sense probe 8 | SP-M8 | AGACG TGC ATT TGA CTA AA GCAGA |
| **1** | Sense probe 10 | SP-M10 | AGACG TGC ATT TCA GTA AA GCAGA |
| **1** | Sense probe 11 | SP-M11 | AGACG TGC ATT TCA CAA AA GCAGA |
| **1** | Anti-sense probe 0 | ASP-M0 | TTT AGT GAA ATG CA |
| **1** | Anti-sense probe 2 | ASP-M2 | TTT AGT GAA ATG GA |
| **1** | Anti-sense probe 4 | ASP-M4 | TTT AGT GAA AAG CA |
| **1** | Anti-sense probe 9 | ASP-M9 | TTT AGA GAA ATG CA |
| **1** | Anti-sense probe 12 | ASP-M12 | TTA AGT GAA ATG CA |
| **1** | Anti-sense probe 13 | ASP-M13 | TAT AGT GAA ATG CA |
| **1** | Anti-sense probe 14 | ASP-M14 | ATT AGT GAA ATG CA |
| | | | |
| **2** | Sense probe | SP-T64 | AGG TGC CAC TGG ATT TTA AGT AAT |
| **2** | Anti-sense probe 60.3 | ASP-T64 | TTA CTT AAA ATC CAG TGG CAC CTt |
| **2** | Anti-sense probe 30.3 | ASP-T42 | TTA CTT AAA ATC CA |
| | | | |
| **3** | Sense probe 60 | SP-L61 | ACT AAC TCA ACA CAC TGA AGA ACC CTA AAG AAT CAC ACA AGA ATG GAA ACT ACA ACA CGC A |
| **3** | Anti-sense probe 60 | ASP-L61 | TGC GTG TTG TAG TTT CCA TTC TTG TGT GAT TCT TTA GGG TTC TTC AGT GTG TTG AGT TAG T |
| **3** | Anti-sense probe 60-3P | ASP-L61-3P | TGC GTG TTG TAG TTT CCA TTC TTG TGT GAT TCT TTA GGG TTC TTC AGT GTG TTG AGT TAG T |
| **3** | Anti-sense probe 25 | ASP-L25 | TG TTG TAG TTT CCA TTC TTG TGT GA |
| **3** | Anti-sense probe 15 | ASP-L15 | TG TTG TAG TTT CCA T |
| **3** | Anti-sense probe 15-3P | ASP-L15-3P | TG TTG TAG TTT CCA T |

**Table A2. Overview of chemicals used in this thesis, including formulas and suppliers**

| **Chemical** | **Supplier** |
|---|---|
| EvaGreen® | |
| FIREpol® DNA polymerase | Solis Biodyne |
| Hot TERMIpol® DNA polymerase | Solis Biodyne |
| Magnesium chloride($MgCl_2$) | Solis Biodyne |
| MilliQ $H_2O$ ($dH_2O$) | MilliQ |
| Probes | Biomers |
| Proteinase K | Qiagen |




**Norges miljø- og biovitenskapelige universitet**
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway