Differential Scanning Fluorimetry on RNA

Final Research Paper submitted to Austrian Marshall Plan Foundation

submitted by

Lisa Katharina John

Marshall Plan Scholarship Recipient 2020

FLORIDA STATE UNIVERSITY

UNIVERSITY OF APPLIED SCIENCES UPPER AUSTRIA

supervised by

DI Dr. Thomas Eidenberger	**University of Applied Sciences Upper Austria, Wels**
Dr. Robert Silvers	**Florida State University, Tallahassee**

Abstract

As is generally known, macromolecules such as proteins are extremely important for a healthy and functioning body, since they have control over processes like nerve transmission, transport of oxygen in blood, energy metabolism, inheritance and reproduction. But even though they are so important for our continued existence, we do not know enough about their behavior during folding, assembly into larger structures, their stability and dynamics. However, it is extremely important to understand them, especially the way they behave within certain circumstances, in order to discover ways to support them with their tasks.

The goal of this project is briefly described: to test the method of Differential Scanning Fluorimetry (DSF) in order to identify its efficiency in examining the melting temperature T_m of RNA and thus the stability of the RNA. In order to have reliable comparison data, measurements are also carried out with a well-tried method, for which UV-Vis spectrometry is used. The measurements are performed under different environmental conditions to see which factors have an influence on the melting temperature, e. g. stabilize or destabilize the RNA. Additionally, different fluorescent dyes are used for the measurements to see if the different mechanisms of operation of these dyes also have an influence on the temperature. The following dyes were used in this project: Quant-iT™ RiboGreen®, SYBR® Green II, Quant-iT™ OliGreen®, UltraPure™ Ethidium Bromide, Acridine Orange, Invitrogen™ TOTO™-1 Iodide.

The raw data was extracted and smoothed and in addition to that the nearest neighbor method with 1000 intervals was used. To make the evaluation of the melting curve easier, the curve was flipped around the x-axis which was obtained by multiplication of -1 and giving the derivative. Since this way of evaluation is rather imprecise, the data was evaluated via Matlab too. The melting curves obtained by UV-Vis spectroscopy and Differential Scanning Fluorimetry were fitted via Matlab and a fitting function was developed. Therefore, no routine contained in the program was used, but an own function was created. It is based on a non-weighted least square procedure, which is available in Matlab.

It was possible to find a suitable fit for the UV-Vis method. An approach was tested for the DSF method, but there are still some differences between the measured and fitted data. Furthermore, a possible influence of DMSO, the RNA concentration and the different reaction modes of the fluorescent dyes could be determined. It was also shown that the sample volume has an influence on the behavior of the melting temperature T_m.

However, there is still a lot of research to be done on this topic to get more significant results.

Acknowledgement

I would like to thank DI Dr. Thomas Eidenberger for officially supervising this project on the side of the University of Applied Sciences Upper Austria. His support and feedback on my matters has been very helpful.

I also want to thank Dr. Robert Silvers for giving me the opportunity to work on such an interesting topic and working in cooperation with the FSU.

Then I would like to thank Angelique Stevens who was my contact person at the CGE and was always trying to help me with my issues.

Next, I want to thank the International Office at the FH Wels which provided a lot of support before and during this work. Thanks to them it was also possible to quickly change the mode of the project and return to Austria to continue working there.

Thanks to Nolan Blackford who performed all the measurements for my project, as it was not possible for me to do them myself.

I want to thank everybody who supported me during the time of my internship and was not mentioned above.

Finally, I want to thank the Austrian Marshall Plan Foundation for granting me the Marshall Plan Scholarship and generously allowed to amend my proposal when I had to return to Austria.

TABLE OF CONTENTS

ABSTRACT ..... **II**

ACKNOWLEDGEMENT ..... **III**

1. INTRODUCTION **1**

1.1 Reasons for this Project 1

1.2 Goals of this Project 3

1.3 Background Knowledge 4

1.3.1 DNA 4

1.3.2 RNA 5

1.3.3 UV-Vis Spectrometry 5

1.3.4 Differential Scanning Fluorimetry 6

1.3.5 Fluorescence 8

2. MATERIAL AND METHODS **10**

2.1 RNA constructs 10

2.2 Fluorescent Dyes 10

2.3 Measurements 10

2.3.1 Differential Scanning Fluorimetry 10

2.3.2 UV-Vis Spectroscopy 11

2.4 Data treatment and analysis of the data 12

2.5 Melting curve fitting 14

2.5.1 Fitting Method and Options 14

3. FITTING RESULTS **15**

3.1 UV-Vis Spectroscopy 15

3.2 Differential Scanning Fluorimetry 18

3.3 Baselines 21

3.4 Melting Temperature 25

4. DISCUSSION **26**

4.1 DMSO Concentration 26

4.2 RNA Concentration 27

5. CONCLUSION AND OUTLOOK 30

6. BIBLIOGRAPHY 31

1. INTRODUCTION

Organic macromolecules such as proteins are the main actors in all essential chemical processes in the human body. In general, the body consists entirely of chemical or biochemical reactions. The gas exchange, which is the absorption of oxygen and the release of carbon dioxide, is a chemical reaction that keeps us alive. Without acetylcholine, one of the most important neurotransmitters in the human body, the communication between brain, nerves and muscles would not be possible. DNA contains the entire genetic material and thus the most important information, while RNA is responsible for the conversion of genetic information into proteins, for which protein biosynthesis, i.e. transcription and translation, is used. It is therefore obvious that those macromolecules are essential for a healthy and functioning body, as they have control over processes like nerve transmission, transport of oxygen in blood, energy metabolism, inheritance, and reproduction. The problem is that not enough is known about their behavior during folding, assembly into larger structures, stability and dynamics even though it is extremely important to understand them, especially the way they behave within certain circumstances and how they can be supported in their tasks [1].

1.1 Reasons for this Project

As already mentioned, proteins are indispensable in most of the essential reactions in the human body. But how are these important macromolecules formed?

As is generally known, DNA contains all genetic information, including the ones required to produce proteins. For this information to be used, first transcription must take place, in the course of which a transcript is created in form of mRNA. Once this has been done, the tRNA translates the existing information into an amino acid sequence, i.e. into a protein, during the translation step [2].

The following figure (1) schematically illustrates this process.

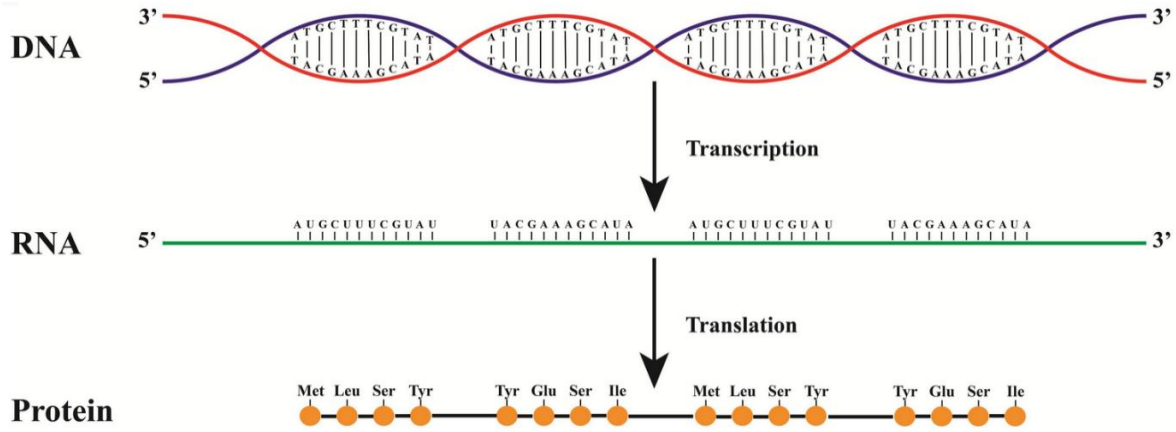

Figure 1. Steps of the protein biosynthesis [3].

In order to perform properly, however, the protein folding must be carried out correctly in addition to the proper execution of the protein biosynthesis. If there is just a single change or deviation in the linear sequence of amino acids, diseases can be caused because the higher levels of organization of proteins are dependent on the primary structure [2].

In the following figure (2) the different structures are portrayed.

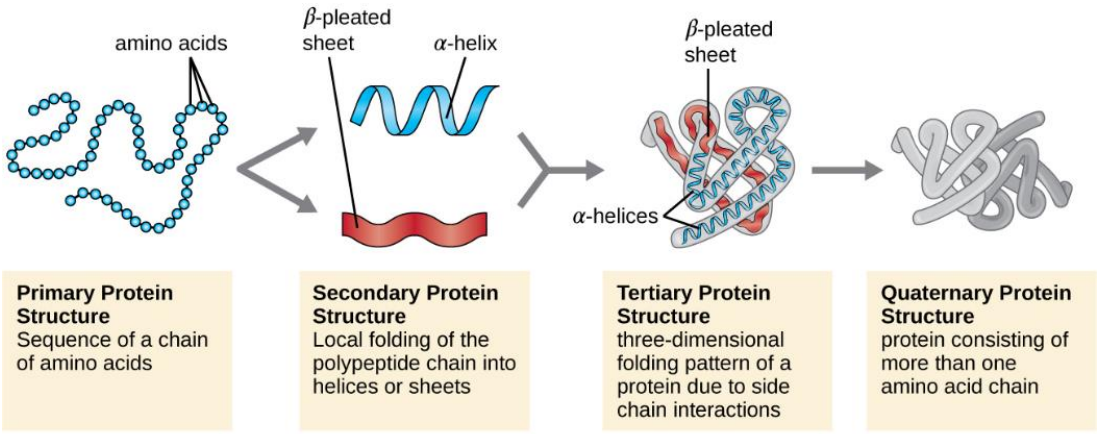

Figure 2: Different protein structures [4].

Proteins mostly work in their native confirmation, which is the structure of energetically most favorable state of the specific protein. This state corresponds to the confirmation with the lowest free energy. Even if everything in the development of a functional protein has worked, it is still possible that the protein denatures. Possible reasons for this to happen are extreme high temperatures, high concentrations of salt or a drastic change in term of the pH concentration [5].

In recent biochemical research the folding of proteins still is a crucial topic. Too little is known about the way proteins fold, the influence different environmental factors can have and the relation between the different proteins. It is important though to

expose more of their nature, because of the critical impact abnormalities can have on the human body.

1.2 Goals of this Project

RNA can fold itself into complex tertiary structures, which leads to a variety of possible RNA structures. They are essential for Hybridization, Catalysis, and protein binding. The problem is that a change of RNA stability, which may occur by change of environmental factors, has a great effect on the function of RNA.

Therefore, it is important to analyze the stability and thus also the melting behavior of RNA under different environmental conditions, like for example the pH-value or the presence of salts and ligands, in order to find out under which conditions the sample is influenced. It is also of great importance to be able to detect the heat stability of RNA correctly, which is only possible with functioning and suitable reporter dyes.

The thermal shift assay, also known as differential scanning fluorimetry, is a method to measure the thermal stability of proteins or, in this case, RNA. Using this method, it is possible to perform measurements under varying environmental conditions (fluctuating pH-value, different salt concentrations, addition of ligands), and different protein reporter dyes to investigate whether the results will be influenced in any way. The main task of this research will be to take a closer look at such measurements and the method itself in order to find out whether the measurements of RNA thermostability can be influenced by various environmental factors and whether the reporter dyes may also have an influence on the measurement result or the behavior of the RNA during the experiment.

The content of this research can be divided into a theoretical and an evaluation section.

The theoretical part consists of literature research on the method and already known stabilizing and destabilizing effects on RNA. This will also be very helpful to understand what exactly happens during the Thermal Shift Assay and why the melting curve looks the way it does in the end. By understanding the process, it becomes easier to determine the reasons for deviations of the maximum melting temperature and the progression of the melting curve during different runs of the same procedure. Since there are several projects ongoing at Silvers Lab, and some of them are related to this project, previously measured data can be used for this project as well. Therefore, another part of the theoretical section of the project is to deal with the evaluation and interpretation of such data. The evaluation will be performed using Matlab. A fitting of the data and thus the determination of the maximum melting

temperature T_m, as well as the progression of the specific melting curve of the different runs, will be done.

Due to the corona situation it was not possible to perform own measurements. Since it is not foreseeable when it will be possible to travel to the US again, this will not be possible in the nearest future. The evaluated measurements were performed by the laboratory assistant Nolan Blackford, who is a student at Florida State University and a member of the Silver's Lab. For the measurements UV-Vis and RTD-PCR respectively Differential Scanning Fluorimetry were used.

1.3 Background Knowledge

1.3.1 DNA

The deoxyribonucleic acid (DNA) consists of polymeric molecules of nucleotide building blocks, which contain 2'-deoxyribose, the bases Adenine, Thymine, Guanine and Cytosine and phosphate groups. DNA can take on different conformations, the most common form being the so-called B-DNA. The DNA consists of two opposing polydeoxynucleotide strands, which have the shape of a right-handed double helix. The backbone of the DNA consists of deoxyribose and phosphate residues, which are linked together by phosphorus diester bonds. Between these backbones there is the large and small groove in which the more accessible bases and DNA binding factors are located, so the interactions are more likely to be bound in this place. A distinction is made between specific bonds such as the hydrogen bridge bonds and the non-specific bonds such as the Van-der-Waals forces and the electrostatic interaction. The base pairs are located inside the double helix, which leads to the fact

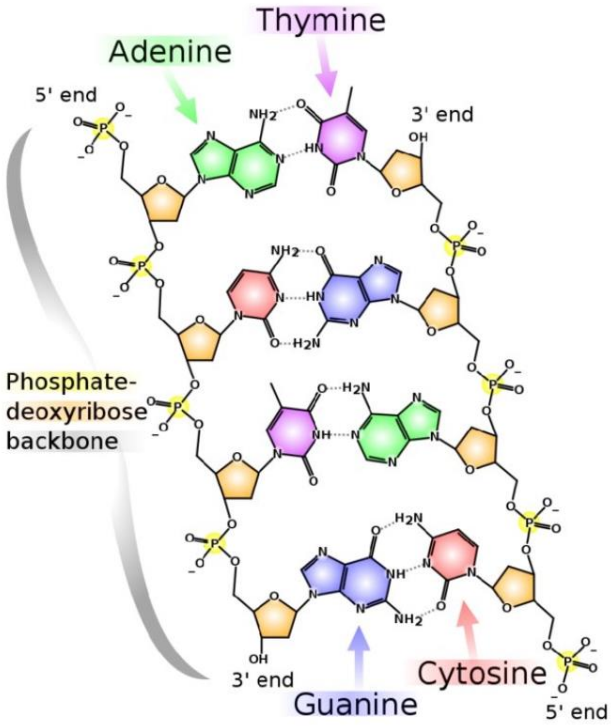

Figure 3: Structure of DNA; red is Cytosine, blue is Guanine, green is Thymine, violet is Adenine; Backbone consists of 2'-deoxyribose and phosphate groups [6].

that the DNA is non-polar in these areas. The surface, where the backbone is located, is polar and negatively charged due to the sugar and phosphate groups [1].

Figure (3) shows the structure of the DNA.

1.3.2 **RNA**

Ribonucleic Acids (RNA) are extremely diverse molecules that, even though they are quite similar to the Desoxyribonucleic acid (DNA), can take on many more complex secondary structures. This is due to the fact that Uracil is less selective in its base pairing than Thymine and because of the additional hydroxyl group on the ribose ring. RNA has significant differences to the DNA; it cannot form double helices and is mostly single-stranded. It contains the bases Cytosine, Guanine, Adenine and Uracil and the backbone is formed of ribose and phosphate groups. There are different types of RNA that perform different tasks. A distinction is made between tRNA (Transfer RNA), mRNA (Messenger RNA), rRNA (Ribosomal RNA) and snRNA (small nuclear RNA) [1].

Figure (4) shows the structure of the RNA graphically.

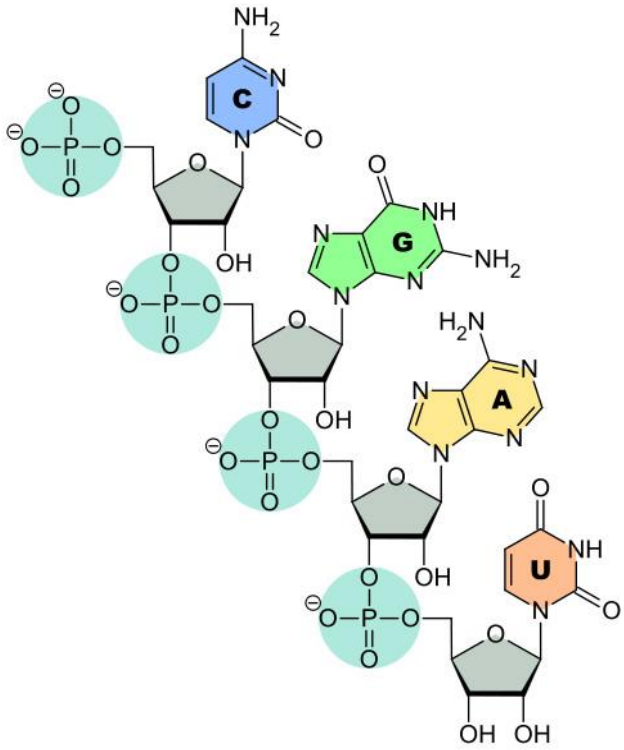

Figure 4: Structure of RNA; blue is Cytosine, green is Guanine, yellow is Adenine, orange is Uracil; Backbone consist of ribose and phosphate groups [3].

1.3.3 **UV-Vis Spectrometry**

When molecules are stimulated by electromagnetic radiation (light waves) with wavelengths in the ultraviolet to visible range (100 - 800nm), they absorb energy at certain wavelengths. This stimulates electrons in the outer orbital region of the molecule depending on the molecular structure. The negative decadic - logarithmic relationship of the intensity difference from input to output is referred to as extinction. The extinction is dependent on the wavelength being measured and is displayed as UV/Vis spectrum by a detector. It is capable of measuring the extinction at a certain wavelength and determining the concentration of a substance in a solution. The equation is based on Lambert-Beer's law, which is represented by the formula (1).

$$E_\lambda = \varepsilon_\lambda \cdot c \cdot d \ (1)$$

E_λ Extinction

E_λ molar extinction coefficient

c concentration in $molL^{-1}$

d layer thickness in cm

In the following figure (5) the process of a measurement via UV-Vis spectroscopy is shown graphically.

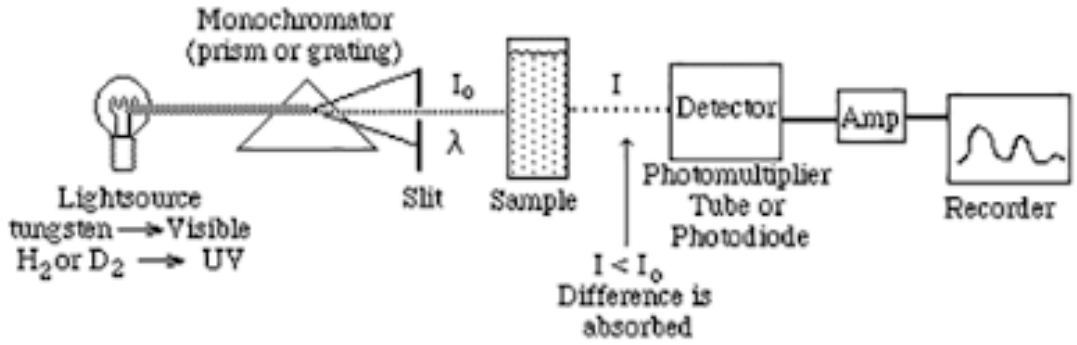

Figure 5: Simplified scheme of a UV-Vis spectrometer [7]

Because RNA absorbs ultraviolet light at a wavelength of 260 nm, a UV-Vis spectrometer with a programable temperature gradient is often used for melting temperature measurements. The absorbance will increase as the RNA unfolds. While UV-Vis spectrometers can be used to test multiple samples at once, they are usually limited to six or twelve at a time, each requiring an expensive quartz cuvette with low head space to minimize evaporation. While the technique is indeed very sensitive, it can struggle with very low concentrations of RNA and, unfortunately, usually requires around 200 µL of sample volume per cuvette. Anyone working with small quantities of expensive RNAs is already acutely aware of these limitations. The large sample volume required and the limiting number of samples that can be tested simultaneously hinder the high throughput screening of RNA stability [8,9,10].

1.3.4 **Differential Scanning Fluorimetry**

Differential Scanning Fluorimetry (DSF) is a powerful technique which is used to quickly test the stability of proteins in a variety of conditions simultaneously while using relatively small amounts of sample. As an example, only 50 µl per well are needed for the project described in this report. All Differential Scanning Fluorimetry measurements were performed via a Real Time Detection Polymerase Chain Reaction (RT-PCR).

The principle behind the real-time PCR, also known as qPCR, is that it monitors the accumulation of PCR amplicons in real time. This is possible as the device measures the change in emission of fluorescence from the fluorescent RNA binding dyes [11]. In the following figure (6) some examples of what the results of a DSF measurement could look like are shown graphically.

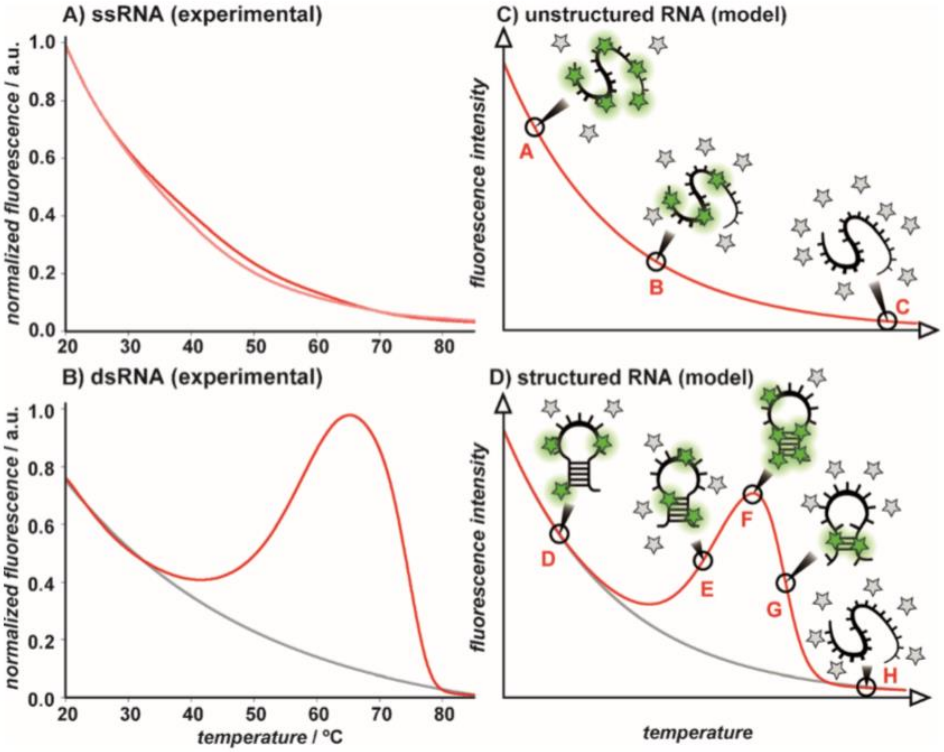

Figure 6: Differential Scanning Fluorimetry (DSF) of a single-stranded RNA (ssRNA) (A) and a double-stranded RNA (dsRNA) (B) with RiboGreen dye and proposed model for fluorescence behavior of DSF signals (C for ssRNA, D for dsRNA) [12].

Bound fluorescent reporter molecules are shown as green stars while the non-fluorescent reporter molecules are shown as gray stars.

For the ssRNA, there is a continuous exponential drop of fluorescence intensity depending on the rise of temperature. In the case of the dsRNA, on the other hand, a deviation from the model can be observed. Looking at graph B) the fluorescence begins to rise at about 40 °C until it reaches its peak at about 65 °C. This result can be explained as follows:

Firstly, the dye binds to the unstructured and bondable parts in the RNA. This produces the initial fluorescence, which decays due to temperature-dependent dissociation and thus decreases exponentially (Figure (6)D; D). Now that the fluorescence has decreased, however, it increases again at about 40 °C because more compact RNA structures destabilize, which in turn enables the dye to be bound. This allows the point of maximum fluorescence to be reached (Figure 6D; E - F). However, the melting process of RNA continues and the binding affinity of the

fluorescent dye decreases. As a result, the fluorescence decreases until it completely disappears when the RNA is completely unfolded (Figure 6D; G – H).

The following figure (7) schematically describes how the RNA and the fluorescent dye SYBR Green are reacting during a DSF.

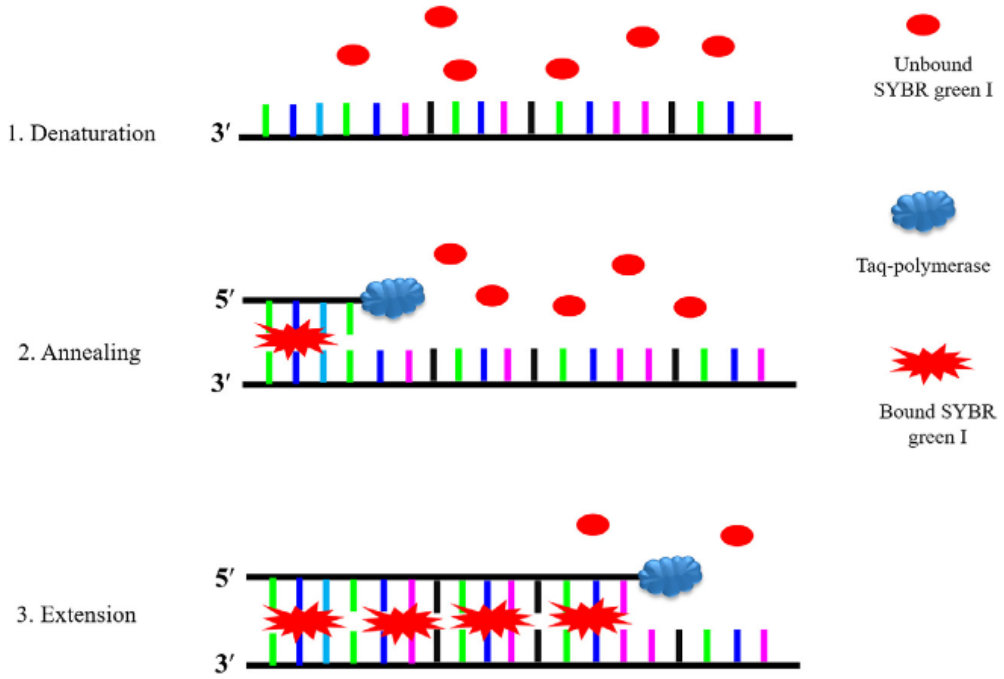

Figure 7: Description of the process occurring during the usage of a real-time PCR; [13].

1.3.5 **Fluorescence**

What is Fluorescence?

Fluorescence is a phenomenon that can be explained as follows: When light of a certain excitation wavelength hits a molecule, photons are absorbed and the electrons of the molecule in concern are lifted to an energetically higher level (which is what happens by the excitation). However, this state of increased energy is not maintained for long. At some point the molecules fall back to their original energy level and the released energy is set free as photons or light and heat, which basically is Fluorescence.

It should be noted though that the emitted fluorescent light is always lower in energy than the incident light since part of the energy is converted into heat. The difference between excitation and radiation wavelengths is about 20 - 50 nm, this phenomenon is also known as Stokes Shift [14].

Fluorescent Dyes

In the case of this project it is a so-called secondary fluorescence, also known as fluorescent staining. This means that synthetically produced fluorescent dyes are used to visualize certain structures, like here the slowly denaturing RNA.

An example of this procedure is following: When a protein melts or unfolds due to heat, the inner hydrophobic regions will be exposed to the aqueous environment. In Differential Scanning Fluorimetry, the temperature is increased in the presence of a fluorescent dye such as RiboGreen which binds to these hydrophobic regions. When the dye binds to the protein, fluorescence will be emitted and the intensity of the light is measured at various temperatures to form a melting curve. Unlike proteins, RNAs form secondary structures without the help of hydrophobic interactions. While RNA DSF would work on the same principles as protein DSF, the dyes used in RNA DSF must work via a fundamentally different mechanism of binding and show specificity for ether single or double stranded RNA so that there will be a change in fluorescence due to melting. Most common nucleic acid stains like Ethidium Bromide are intercalating, which means they bind by fitting between the two strands of double stranded DNA or RNA. This is useful as they show specificity for double stranded RNA, but this also means that they drastically affect the stability of the RNA. This makes the results of an RNA DSF experiment using these dyes difficult to analyze. Therefore, also measurements with a non-intercalating dye, e.g. Quant-iT RiboGreen should be performed to see if there are any differences in the results.

In the course of this project the following fluorescent dyes were tested for their compatibility with the RNA:

- Ethidium Bromide
- Acridine Orange
- RiboGreen
- OLI Green
- SYBR Green
- TOTO-1 Iodide

These fluorescent dyes were used in various concentrations respectively dilutions to find out whether the amount of the dye influences the stability of the used RNA [11; 14].

2. Material and Methods

2.1 RNA constructs

Both single stranded RNA (ssRNA) and double stranded RNA (dsRNA) constructs were used in the course of this project. dsRNA was reconstituted from equimolar concentrations of ssRNA1 and ssRNA2 [15]. Two complementary strands of ten nucleotides with no known secondary structures as the initial standard were used (10-mer A and B).

2.2 Fluorescent Dyes

Six different Fluorescent Dyes were used. In this process, Quant-iT™ RiboGreen® (RIBO, Life Technologies, Ex_{max}=~500 nm, Em_{max}=~525 nm), SYBR® Green II (SYBR, Life Technologies, Ex_{max}=~497 nm, Em_{max}=~520 nm), Quant-iT™ OliGreen® (OLI, Life Technologies, Ex_{max}=~500 nm, Em_{max}=~525 nm), UltraPure™ Ethidium Bromide (EtBr, Sigma, Ex_{max}=~510 nm, Em_{max}=~590 nm), Acridine Orange (AO, ThermoFisher Scientific, Ex_{max}=~502 nm, Em_{max}=~525 nm), Invitrogen™ TOTO™-1 Iodide (TOTO, ThermoFisher Scientific, Ex_{max}=~514 nm, Em_{max}=~533 nm) were used. All fluorescent dyes were used according to the manufacturer's instructions. The concentrations of the dyes in the final sample were fluctuating since different concentrations were tested to see if the dyes have any effect on the stability of the RNA. Stocks for all dyes were 10000x in DMSO and were diluted in water prior to addition to the samples.

2.3 Measurements

2.3.1 Differential Scanning Fluorimetry

All DSF measurements were performed in 96-well plates (MicroAmp® Fast Optical 96-Well Reaction Plate; Life Technologies). The sample volume was 50 µL per well. In order to measure the Fluorescence intensities / absorbance a StepOnePlus RT-PCR (Life Technology) was used whereby its melting curve routine was applied.

The samples were incubated at 20 °C for 3 minutes, after that a heating rate of 0.5 °C per minute up to 95 °C was applied. In order to record the fluorescence intensities continuously, the setting for SYBR® Green reagents was used.

For the DSF measurements 25.2 nanograms of single stranded, respectively double stranded RNA was used. After some testing in order to determine the best conditions

for the planned experiment, the environmental conditions were settled on 50 mM phosphate buffer, pH 7, 1mM EDTA, 5mM NaCL and 2 % DMSO by volume.

To ensure that the measurements were statistically valid, which means that they are reproduceable, and in order to improve the quality of the data, several runs were performed and replicated multiple times. The data was averaged, and the standard deviation was used for error determination.

In some measurements the concentration of RNA was also varied to see if this could have an effect on the resulting melting curve / temperature.

2.3.2 **UV-Vis Spectroscopy**

The RNA samples for UV absorbance measurements were unfolded prior to measurements at 85°C for 3 minutes and then slowly cooled to room temperature (20 °C). All UV-Vis absorbance measurements were performed in a 1.2 mL quartz cuvette (path length: 10 mm) harboring a temperature probe. The sample volume was 200 µL per cuvette. For that an initial absorption of ~0.4 at 260 nm was measured. For the measurements, a Thermo Scientific Evolution 300 UV/Vis spectrometer with its programmed routines was used. The absorbance was monitored at 260 nm. The sample preparation is described below:

The RNA was held at 20 °C for 10 minutes, then the sample was heated to 80 °C with a rate of 0.5 °C per minute. This was followed by 3 minutes of incubation at 80 °C. Finally, the sample was cooled back to 20 °C using a rate of – 0.5 °C per minute.

For the UV-Vis measurements 1.26 micrograms of single stranded, respectively double stranded RNA was used. After some testing in order to determine the best conditions for the planned experiment, the environmental conditions were settled on 50 mM phosphate buffer, pH 7, 1mM EDTA, 5mM NaCL and 2 % DMSO by volume.

No significant hysteresis could be detected. UV absorbance data traces were subsequently analyzed by data smoothing and derivation.

In some measurements the concentration of RNA was also varied to see if this could have an effect on the resulting melting curve / temperature.

2.4 Data treatment and analysis of the data

In general, the treatment and analysis of the obtained data can be separated into 3 steps, which the following figure (8) illustrates graphically.

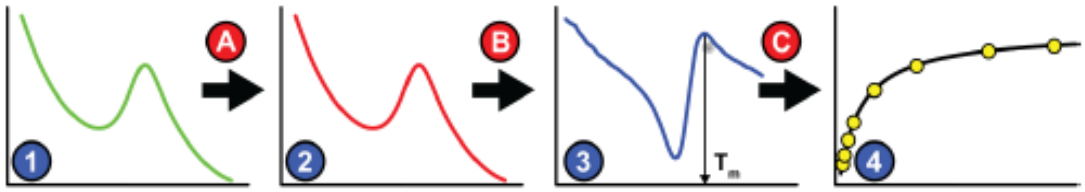

Figure 8: Different steps of Treatment and Analysis of obtained data, the shape of the curves in the figure simulates the shape of DSF data with x = temperature and y = fluorescence intensity [16].

In all the Differential Scanning Fluorimetry measurements fluorescence was recorded ranging from 20 °C to 95 °C whereas the temperature during the UV-Vis Spectrometry measurements fluctuated between 20 °C and 80 °C. For the DSF data the device collected 407 data points, which equals a data pitch of ~0.19 °C. For the UV-Vis data the device typically collected 569 data points, which equals a data pitch of ~0.10 °C. Also, it should be mentioned that the devices recorded the temperature for each well (DSF) and every run of the UV-Vis spectrometer. Concerning the DSF data it has to be pointed out that the standard deviation typically was below 0.05 which is the reason why the average of all 96 temperature values measured in the 96 wells was used for each data point.

Before fitting the data with Matlab, it was prepared and controlled as can be obtained in Figure (8). The following description refers to the numbering of this figure. First, the raw data (1) was extracted and smoothed (A), whereas the smoothing was performed using SigmaPlot 11.0 and its implemented smoothing routine. A bisquare smoother was used as well, with a sampling proportion of 0.1 and a polynomial degree of 1 or 2, depending on the performance of the smoothing routine. In addition to that, the nearest neighbor method with 1000 intervals was used. This resulted in a smoothed data curve with 1001 data points which at first glance could not be distinguished from the original curve. This means that the method did not change the data in any unwanted way (2). The data also was normalized to its highest initial intensity by dividing all individual data values by the highest initial value.

To make the evaluation of the melting curve easier, the curve was flipped around the x-axis which was obtained by multiplication of -1 and taking the derivative (3). All of this leads to the goal of easily extracting and analyzing the melting temperature T_m, which is the maximum of the final curves (C).

If needed, the slope of a data point (i) could be calculated by the following formula (2) [16]:

$$slope(i) = \frac{[y(i+2) - y(i-2)]}{[x(i+2) - x(i-2)]} \quad (2)$$

The corresponding temperature could be calculated by equation (3) [16]:

$$\text{temperature(i)} = \frac{[\text{x(i} + 2) + \text{x(i} - 2)]}{2} \quad (3)$$

2.5 Melting curve fitting

Since the evaluation method described in *"2.4 Data treatment and analysis of the data"* is still a bit imprecise, the measured data was also fitted via Matlab. In order to perform the fitting properly an own fitting function was devised. It is based on a non-weighted leastsquare procedure, which is already implemented in Matlab.

In the following chapter the fitting with Matlab in general and in the case of the Differential Scanning Fluorimetry data and the UV-Vis data will be described.

The reason why the fitting is performed is to get a fitting function, to get more precise melting temperatures T_m and to see if there is any baseline caused by the fluorescent dyes, the RNA itself or something else in the sample mixture.

2.5.1 Fitting Method and Options

The first step is to look at the progress of the curve to be fitted and to evaluate which kind of function it could be. With the data of this project it seemed to be a mixture of a polynomial function and an exponential function at first, but after some runs of fitting it was clear that this is not the right choice. After a literature research it became clearer that it is a sigmoid function. The following formula (4) is the basic form of a sigmoid function [17]:

$$S(x) = \frac{1}{1 + e^{-x}} \quad (4)$$

After the formula had been found and verified, the actual fitting process could be started. In figure (9) each step of the curve fitting is pointed out graphically. The curve was fitted to the measured data points by evaluating the unknown parameters. The first step was to import the measured data into Matlab to be able to work with it, after that the needed constants and the fitting options for the fitting procedure were set. If the fitting function and the values and constants were chosen correctly, a proper fit can be performed. If the fitted curve overlaps with the measurement curve, the fitting was successful. Otherwise, in case there are too many differences, the fitting function and the other values have to be reevaluated in order to get a better fit in the next run.

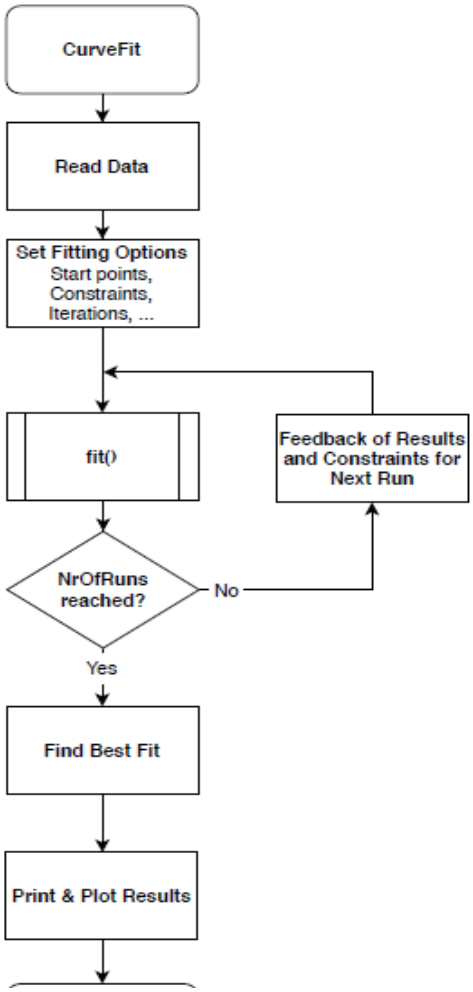

Figure 9: Flowchart of the fitting procedure

3. Fitting results

3.1 UV-Vis Spectroscopy

The following table (1) shows the selected options and the corresponding values for fitting a UV-Vis measurement of RNA (10mer A and B; 5 µM) and the fluorescent dye RiboGreen (400 ng / µl).

Table 1: Fitting options for the fitting procedure of the measurement “UV-Vis A+B 5 uM / RiboGreen400x”

Option	Value
Method	NonLinearLeastSquare
Algorithm	Trust - Region
Maximum iterations	800
Number of steps	100
Tolerance	10^{-18}
Start values	A = 0.964874429; B = 0.143905282; C = 1.2; D = 53; E = 4.7; G = 0.001
Initial constraint coefficients	up = 1.3; low = 0.7

All measurements via UV-Vis Spectrometry were fitted in this way. The measurement described in this chapter is the only one that will be described in detail, otherwise the scope of this report would be exceeded.

The nonlinear least square method with the Trust-Region algorithm was used. This algorithm allows to set constraints for the fitted parameter results. The number of maximum iterations was limited on 800, the number of steps on 100. In addition to that the fitting run aborts earlier if the change in parameters per iteration is smaller than the set tolerance which is 10^{-18} in this specific case.

After some rounds of fitting and literature research the following fitting function (5) was selected [17]:

$$f(T) = (T \cdot G) + \left(A - \left(\frac{B}{C + e^{\left(\frac{T-D}{E}\right)}} \right) \right) \quad (5)$$

From this it can be concluded that it is a slightly modified sigmoid function mixed with an exponential function, which indicates an exponential baseline.

The start values could be determined based on the shape of the measured melting curve and the individual measurement data. Thereby the A is a specific instinctive point of the curve, B is the maximum absorbance minus the minimum absorbance and D is the x-axis value of the inflection point of the curve. The parameters C, E, G influence the shape or linearity / curvature of the curve. At these points the fitting curve becomes more linear the larger the start value is and the more angular the smaller the value becomes. Hence the points can be partially identified by reading the data and partially by multiple tests of different values. For each further step of the fitting, the result of the respective previous step is used as the new start value. To get the constraints, the respective start value is multiplied by the constraint coefficients. The constraints are narrowed every 5th step by reducing these coefficients according to the following equation (6):

$$\text{coefficients}_{\text{new}} = \text{coefficients}_{\text{old}}^{\frac{2}{\text{step}}} \quad (6)$$

The initial constraint coefficients are up, for upper constraint, and low, for lower constraint. To determine these coefficients and the equations (6) several trials were necessary, whereby the errors were taken into consideration. The reason for the constraints is to prevent a runaway of the parameters and to get a valid convergence. After the complete fitting progress is finished, the results are logged and plotted.

In the following figure (10) the fitted curve of the measurement “UV-Vis A+B 5 uM / RiboGreen400x” is shown.

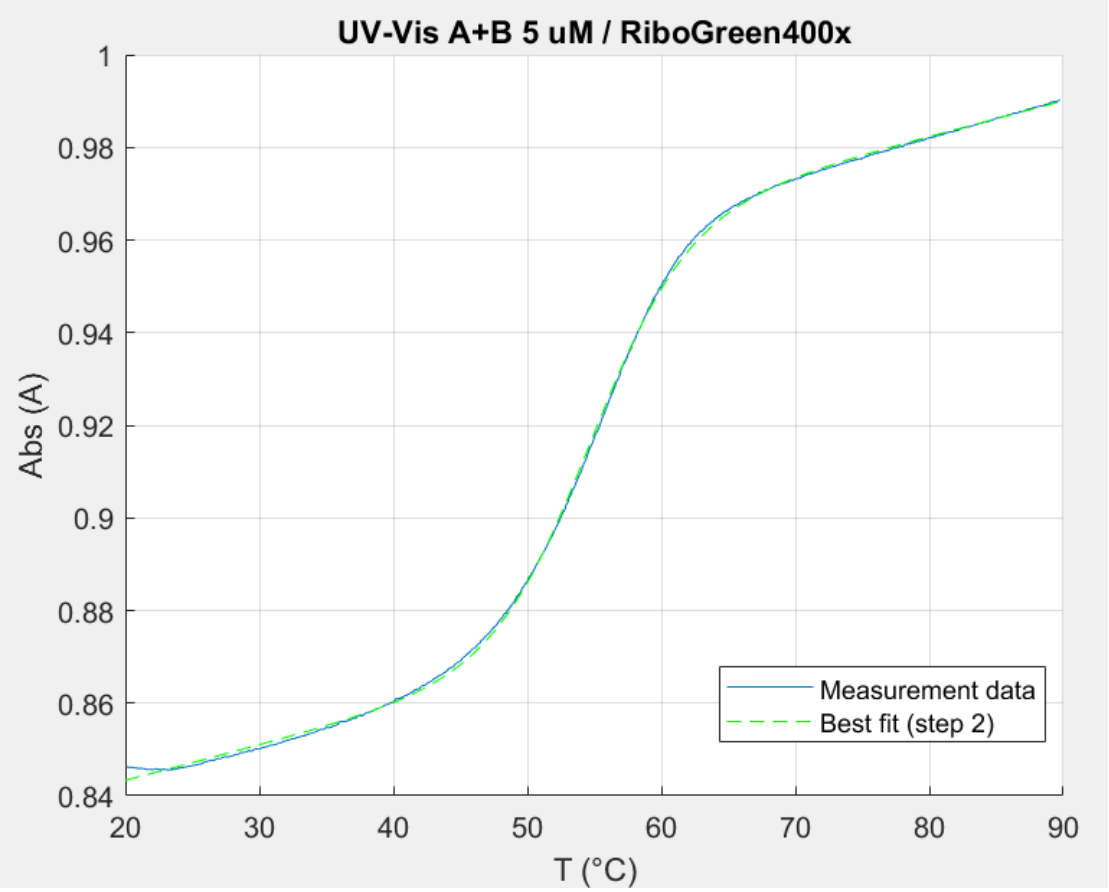

Figure 10: Measured melting curve and the best fit of “UV-Vis A+B 5 uM / RiboGreen400x”.

It can be seen that the fit matches the measured data quite well, there are small deviations at the beginning of the curve, which is not to severe. The following table (2) shows the melting curves parameters of the best fit.

Table 2: Parameters of the best fit of the melting curve for the data “UV-Vis A+B 5 uM / RiboGreen400x”

A	B	C	D	E	G
0.9215	0.1252	1.3400	53.6824	3.6211	7.6097e-04

3.2 Differential Scanning Fluorimetry

The following table (3) shows the selected options and the corresponding values for fitting a Differential Scanning Fluorimetry measurement of RNA (10mer A and B; 5 µM) and the fluorescent dye RiboGreen (400 ng / µl).

Table 3: Fitting options for the fitting procedure of the measurement “Differential Scanning Fluorimetry A+B 5 uM / RiboGreen400x”

Option	Value
Method	NonLinearLeastSquare
Algorithm	Trust - Region
Maximum iterations	800
Number of steps	100
Tolerance	10^{-18}
Start values	A = 0.0055666; B = 0.805190988; C = 0.7; D = 57; E = 2.5; F = 1.733; G = 0.04819; H = 0.04832
Initial constraint coefficients	up = 1.3; low = 0.7

In conclusion, the same procedure was used as for the fitting for the UV-Vis measurement. The only differences are in the fitting function and the starting points. For clarity, all DSF measurements were fitted in this way. The measurement described here is the only one that will be described in detail, otherwise the scope of this report would be exceeded.

After some rounds of fitting and literature research the following fitting function (7) was selected [17]:

$$f(T) = \left(\mathrm{A} + \left(\frac{\mathrm{B}}{\mathrm{C} + \mathrm{e}^{\left(\frac{\mathrm{T}-\mathrm{D}}{\mathrm{E}}\right)}} \right) \right) + \left(-(F - G) \cdot e^{\left((-H \cdot T) + G\right)} \right) \quad (7)$$

It can be noticed that there are some similarities to the fitting function of the UV-Vis data (5), which can be explained with the argument that the base of all the obtained melting curves is a sigmoid function. As it was seen in the other function, this one also has an exponential baseline. Additionally, a gaussian baseline was identified. This is the reason for the mixture of a sigmoid, an exponential and a gaussian function.

In this case the start values are also a bit different. Instead of having 6 unknown parameters there are 8 parameters to determine. As already mentioned, some parts of the equation have similar function as the one of the UV-Vis data. Therefore, A is still a specific instinctive point of the curve, B also is the maximum absorbance minus the minimum absorbance and D still is the x-axis value of the inflection point of the

curve. The variables, E, G also influence the shape or linearity / curvature of the curve. At these points the fitting curve becomes more linear the larger the start value is and expresses a higher curvature the smaller the value becomes. These values were determined by trials and the involved errors again. The additional parameters F and H represent the fluctuations at the beginning of the melting curve. Thereby H has an influence on whether there is a peak or not, so it makes a difference if it is a simple sigmoid curve or whether a baseline influences the shape of the curve. F on the other hand influences the steepness of the peak, the higher the value, the steeper the peak. The rest of the fitting process is performed as explained in *"3.1 Matlab: UV-Vis Spectrometry"*.

In the following figure (11) the fitted curve of the measurement "Differential Scanning Fluorimetry A+B 5 uM / RiboGreen400x" is shown graphically.

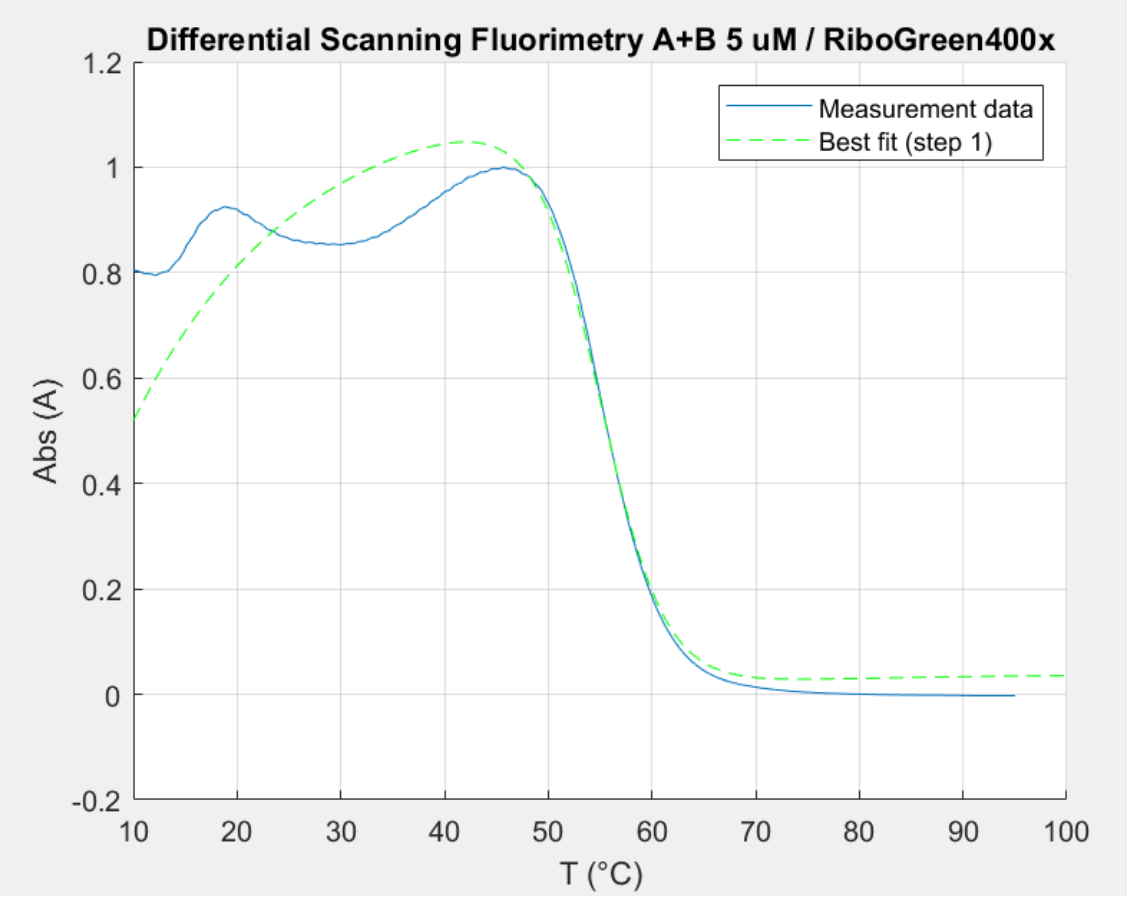

Figure 11: Measured melting curve and the best fit of "Differential Scanning Fluorimetry A+B 5 uM / RiboGreen400x".

It can be noticed that the fit is not as good as the one concerning the UV-Vis data, which is explained by the fact that the DSF data is much more complicated. As already mentioned, it is a sigmoid function with an exponential and a gaussian baseline. Especially the part until ~49 °C is hard to fit because of the fluctuating absorbance. The explanation for these fluctuations is most likely that the baselines, which occur because of A-mer and B-mer RNA, influence the progression quite a lot. The fluorescence dye also influences the shape of the curve. During the measurements it was found out that the smaller the dilution of the fluorescence dyes, the more severe the fluctuations respectively peaks up to a temperature about 50 °C becomes. Thereby it does not matter which dye is used, since this effect occurs with all of the used fluorescent dyes.

The following table (4) shows the melting curves parameters of the best fit.

Table 4: Parameters of the best fit of the melting curve for the data “Differential Scanning Fluorimetry A+B 5 uM / RiboGreen400x”

A	B	C	D	E	F	G	H
0.0039	0.7911	0.7123	56.0917	3.0942	1.2131	0.0337	0.0628

3.3 Baselines

For a better understanding of the baseline issue, here is a short explanation. If no interference occured, the melting curve should be a simple sigmoid function. However, this was not the case in the measurements done during this project. Since the curves always had deviations a so-called baseline correction was performed during fitting. In other words, the fitting function was modified such that the deviation was removed. These baselines can be introduced by the reagents in the samples or, for example, by the RNA itself. The problem is that the baselines can include peaks or fluctuations in the melting curve, which can significantly affect the results [18].

As an example, 10-mer A and 10-mer B RNA was measured and fitted. The following figure (12) shows the measured data and the fit for the 10-mer A RNA baseline with Ribogreen measured via Differential Scanning Fluorimetry.

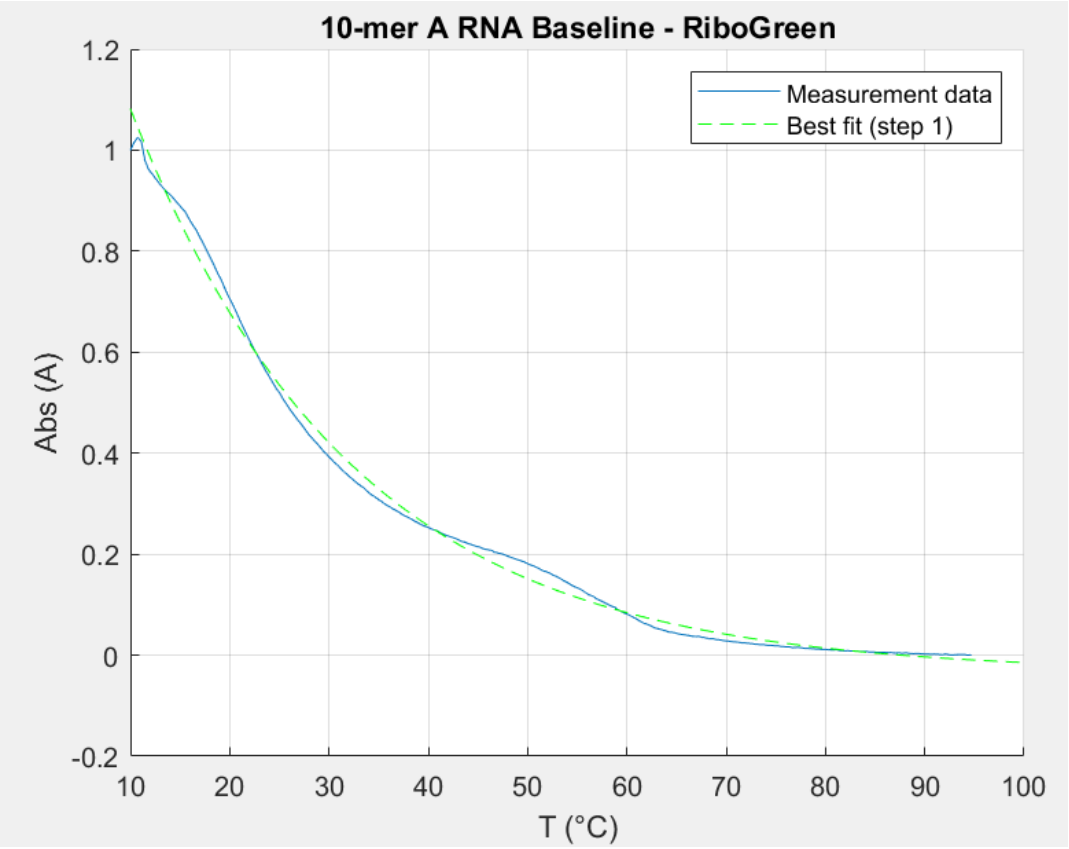

Figure 12: Baseline of 10-mer A RNA measured with RiboGreen (400x) and via Differential Scanning Fluorimetry.

The figure shows that the baseline has the progression of an exponential function.

The following equation (8) was used as a fitting function [22]:

$$f(T) = (A - B) \cdot e^{\left((-C \cdot T) - B\right)} \quad (8)$$

The following table (5) shows the selected options and the corresponding values which were used for fitting the Differential Scanning Fluorimetry measurement of the Baseline of 10-mer A RNA and RiboGreen.

Table 5: Fitting options for the fitting procedure of the measurement “10-mer A RNA Baseline – RiboGreen (400x)”

Option	Value
Method	NonLinearLeastSquare
Algorithm	Trust - Region
Maximum iterations	800
Number of steps	20
Tolerance	10^{-18}
Start values	A = 1.733; B = 0.03627; C = 0.04832
Initial constraint coefficients	up = 1.1; low = 0.9

Concerning the startpoint values, there are three unknown parameters. A influences the beginning point on the y-axis and the slope of the function. B and C both influence the shape and the shift on the y-axis. The following table (6) shows the melting curves parameter of the best fit.

Table 6: Parameters of the best fit of the melting curve for the data “10-mer A RNA Baseline – RiboGreen”

A	B	C
1.7818	0.0343	0.0448

Although there are some small fluctuations between the measured data and the fit, they are still within the acceptable range.

The following figure (13) shows the measured data and the fit for the 10-mer B RNA Baseline with Ribogreen measured via Differential Scanning Fluorimetry.

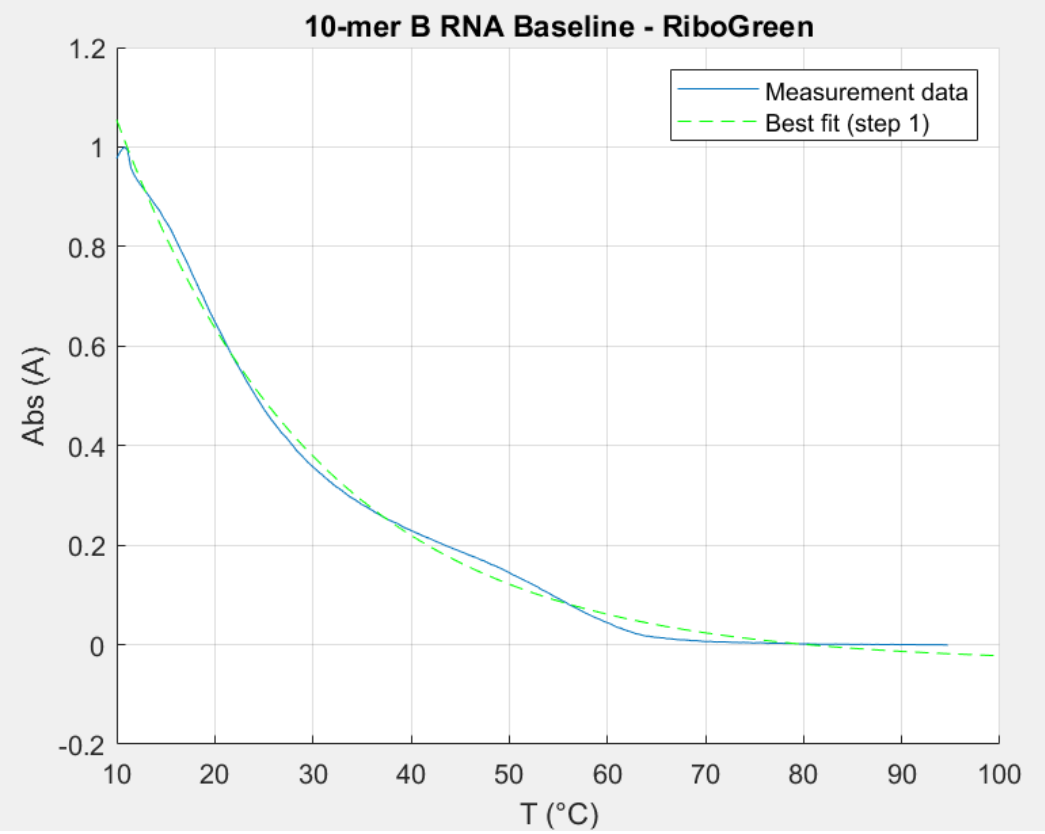

Figure 13: Baseline of 10-mer B RNA measured with RiboGreen and via Differential Scanning Fluorimetry (400x).

As well as in the other baseline measurement and fit, it is obvious that the baseline has the progression of an exponential function. To be exact, the same fitting equation (8) as in the other baseline measurement was used. The following table (7) contains the selected options and the corresponding values which were used for fitting the DSF measurement of the Baseline of 10-mer B RNA and RiboGreen.

Table 7: Fitting options for the fitting procedure of the measurement “10-mer B RNA Baseline – RiboGreen (400x)”

Option	Value
Method	NonLinearLeastSquare
Algorithm	Trust - Region
Maximum iterations	800
Number of steps	20
Tolerance	10^{-18}
Start values	A = 1.733; B = 0.03627; C = 0.04832
Initial constraint coefficients	up = 1.1; low = 0.9

For the start point values there are again three unknown parameters. A influences the beginning point on the y-axis and the slope of the function. B and C both influence the shape and the shift on the y-axis. The following table (8) shows the melting curves parameters best fit.

Table 8: Parameters of the best fit of the melting curve for the data “10-mer B RNA Baseline – RiboGreen”

A	B	C
1.8050	0.0363	0.0483

Again, it is noticeable that the fitted values are very close to the measured values. They are also similar to the parameters of the best fit of the 10-mer A RNA baseline measurement, which can be explained by the fact that two complimentary strands of RNA are used for this project.

3.4 Melting Temperature

In the following table (9) there is a small overview of the melting temperatures T_m obtained over the course of this project. As the project concentrates on the method of Differential Scanning Fluorimetry, only temperatures of this method are included. However, it should be noted that there are no significant differences between the final melting temperatures of UV-Vis and DSF. The intention behind this is to point out the temperature differences of the various fluorescent dyes. The melting temperature T_m can be determined from the melting curve at the inflection point.

Table 9: Overview of the obtained melting temperature T_m obtained via Differential Scanning Fluorimetry using different dyes and different concentrations.

Dye/Concentration	Melting Temperature (°C)
RiboGreen200x	56.34
RiboGreen / 400x	56.14
OLIGreen / 200x	60.81
OLI Green/400x	59.06
OLI Green/800x	58.48
OLI Green/4000x	59.26
TOTO-1 Iodide/100x	54.78
TOTO-1 Iodide/200x	52.25
SYBR Green/10,000x	56.14
SYBR Green/20,000x	56.34
Acridine Orange/50,000x	55.17
Acridine Orange/100,000x	56.53

4. Discussion

4.1 DMSO Concentration

DMSO or Dimethyl sulfoxide can have a significant effect on the stability of RNA. However, the contact of RNA and DMSO is almost unavoidable since the fluorescent dyes are dissolved in this agent. The effect of DMSO is mostly recognizable by the altered RNA structure and the reduced ligand binding [19].

In the following table (10) and table (11) the results of a Differential Scanning Fluorimetry measurement are shown. The usual conditions were used, which means 50 mM phosphate buffer, pH 7, 1mM EDTA, 5mM NaCL are put into the wells. The difference is the varying DMSO concentrations, to show if the agent has any influence on the course of this experiment. In addition to that the fluorescent dye RiboGreen in a dilution of 400 ng / µl was used, the result of this measurement are described in table (10).

Table 10: Results of a Differential Scanning Fluorimetry performed with different DMSO concentrations and the fluorescent dye RiboGreen (400x).

Fluorescent Dye	DMSO percent by Volume [%]	Melting Temperature T_m [°C]
400x RiboGreen	0.25	51.75
	0.50	51.12
	1.00	50.54
	2.00	51.07

To be certain that there are not only deviations in the run with the fluorescent dye RiboGreen, a measurement with OLIGreen was carried out as well. In table (11) the result of the measurement with the fluorescent dye OLIGreen in a dilution of 400 ng / µl is displayed.

Table 11: Results of a Differential Scanning Fluorimetry performed with different DMSO concentrations and the fluorescent dye OLIGreen (400x).

Fluorescent Dye	DMSO percent by Volume [%]	Melting Temperature T_m [°C]
400x OLIGreen	0.25	52.39
	0.50	54.38
	1.00	53.54
	2.00	53.54

In the case of the measurement with RiboGreen there is a temperature fluctuation of ± 0.68 °C, the OLIGreen-Mix measurement on the other hand shows fluctuation of ±1.99 °C. 0.68 °C is not too severe, but a difference of 2 °C is indeed unacceptable. The interesting part about this is, that the temperature decreases the higher the DMSO concentration gets, which can be argued with the fact that DMSO alters the RNA structure and reduces the ligand binding activity and therefore the melting temperature T_m drops.

4.2 RNA Concentration

As the project progressed, it became noticeable that the melting temperature T_m could be drastically influenced by the concentration of the RNA. In some cases, T_m differed over several degrees. This is a very important information because small differences in the volume of RNA e.g. because of false pipetting could lead to differences of the temperature. The following figure (14) shows the influence of the RNA (10-mer A + 10-mer B) concentration on the melting temperature T_m. The measurement was performed with Differential Scanning Fluorimetry and the fluorescent dye RiboGreen (400x).

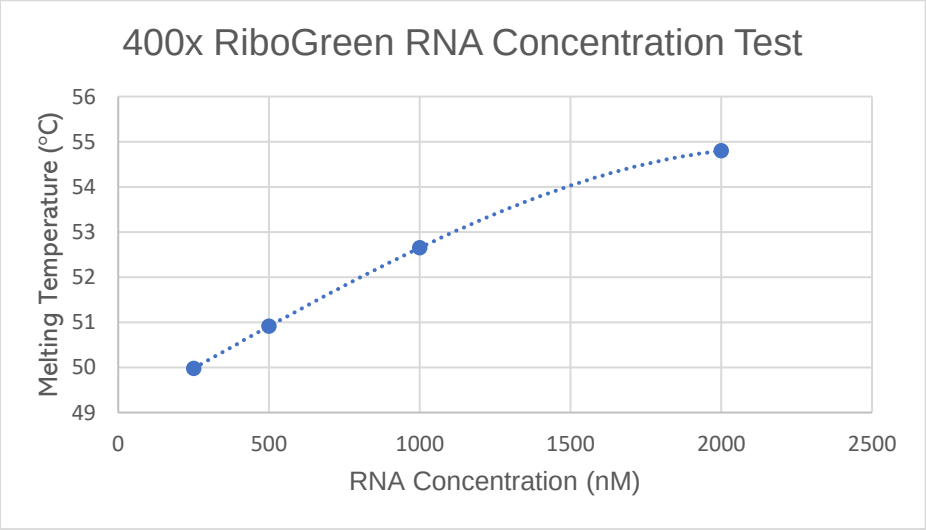

Figure 14: Differential Scanning Fluorimetry measurement to see if a fluctuating RNA concentration has an effect on the melting temperature; RNA = 10-mer A + 10-mer B; 400x RiboGreen; 50 mM phosphate buffer, pH 7, 1mM EDTA, 5mM NaCL, 2 % DMSO by volume.

In order to get more significant data concerning this issue, another measurement was performed. The only difference was that another fluorescent dye was used and that its dilution was much higher than the one of RiboGreen. The following figure (14)

shows how much the RNA (10-mer A + 10-mer B) concentration influences the melting temperature T_m. The measurement was performed with Differential Scanning Fluorimetry and the fluorescent dye SYBRGreen (8000x).

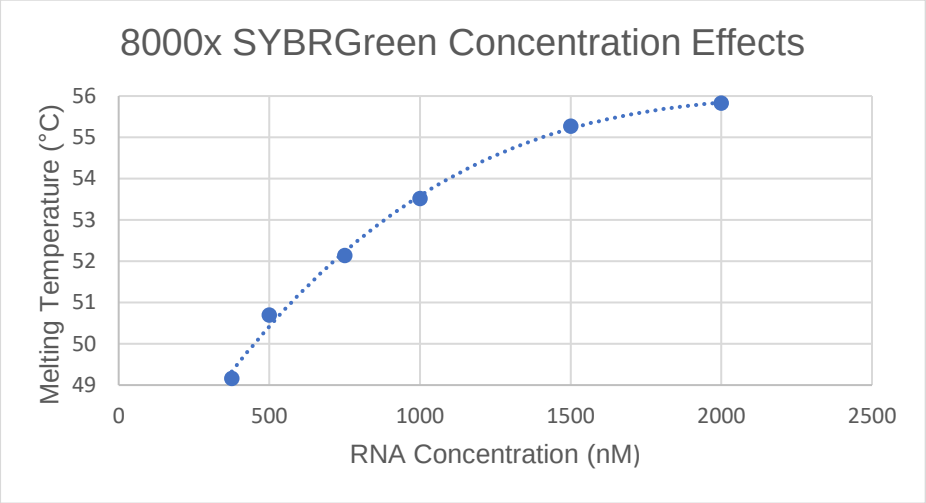

Figure 15: Differential Scanning Fluorimetry measurement to see if a fluctuating RNA concentration has an effect on the melting temperature; RNA = 10-mer A + 10-mer B; 8000x RiboGreen; 50 mM phosphate buffer, pH 7, 1mM EDTA, 5mM NaCL, 2 % DMSO by volume.

It can be noticed that there are significant differences between the two curves, thereby the measurement with SYBRGreen shows much more deviation than the one with RiboGreen. It is generally known that SYBRGreen is an intercalator which means the dye is known for insertion of molecules (in this case the fluorescent dye) into chemical compounds (here the RNA) without major structural changes during the insertion process. RiboGreen does not have that reputation, therefore the different reaction is justified. Another interesting information is that the dye to RNA ratio changes with each point, because the dye concentration stays the same whereas the RNA concentration is doubled [20].

In the following table (12) another example of this phenomenon is given in order to show how much the melting temperature is really influenced.

Table 12: Different Voluminal and RNA concentrations (10-mer A and B) in order to see the influence these factors have on the melting temperature of the RNA measured by Differential Scanning Fluorimetry; fluorescent dye: 400x RiboGreen.

Sample Volume	1 µM RNA	5 µM RNA
10 µL	50.7	56.1
20 µL	50.1	55.3
30 µL	49.8	56.5
40 µL	49.7	55.4

In an earlier stage of this project, it had already been tested whether RiboGreen could affect the melting temperature during a DSF measurement and no significant difference in temperature was detected. The RNA concentration, on the other hand, has a much more significant influence, as can be seen in the measurement data described above. However, it also depends on which fluorescent dye is used. Different measurements were performed, and significant differences were found. However, before a more precise statement can be made, further measurements must be carried out.

5. Conclusion and Outlook

In conclusion, there is still some work to be done on the project. In the case of the UV-Vis measurements there is already a pretty clear overview of the overall situation, a suitable fitting function has been found and most of the problems which arose during the measurements have been solved. The method of Differential Scanning Fluorimetry, however, raises several questions. The first problem is that the fitting function is not ideal, although it is fairly certain that the changes in the melting curve is caused by an exponential and gaussian baseline. Consequently, a fitting function must be found in order to detect the exact melting temperature T_m and thus the values of the curve can be better understood. In general, measurements with DSF should be performed several more times with all fluorescent dyes to further confirm the significance of the melting temperatures obtained so far. Also, the new information that was discovered, like the possible influence of DMSO and the RNA concentration, the different reaction modes of the fluorescent dyes and not to forget the effect the sample volume has, is very important for the future research. These questions need to be answered in the future research at Silvers Lab to find out if Differential Scanning Fluorimetry really is as promising of a method to determine the stability of RNAs as it is suspected to be.

6. Bibliography

[1] Taschenbuch der Biochemie; Jan Koolman, Klaus – Heinrich Röhm, 2. Auflage; Stuttgart: 1997.

[2] Lexikon der Medizinischen Laboratoriumsdiagnostik: Proteinbiosynthese; J. Arnemann, A.M. Gressner, T. Arndt; Lexikon der Medizinischen Laboratoriumsdiagnostik. Springer Reference Medizin; Springer, Berlin, Heidelberg: 2019.

[3] Gentechnologie: Know-how mit Dual-Use-Aspekten; Mirko Himmler; Carl Friedrich von Weizsäcker-Zentrum für Naturwissenschaft und Friedensforschung; Universität Hamburg: 2015.

[4] The incredible power of proteins for our body; thecostaricanews.com; last call: 11.09.2020

[5] Biochemistry; Jeremy M. Berg, John L. Tymoczko, Lubert Stryer; 5th edition; W. H. Freeman; New York; 2002.

[6] DNA-based Applications in molecular electronics; Veikko Linko; faculty of Mathematics and Science; University of Jyväskylä; Finland: 2011.

[7] UV-Vis Spectrophotometer; shimadzu.com; last call: 25.09.2020.

[8] Instrumentelle Analytik / Experimente ausgewählter Analyseverfahren; Sergio Petrozzi; Wiley-VCH Verlag GmbH & Co; Weinheim: 2010.

[9] Lebensmittelanalytik; Reinhard Matissek, Gabriele Steiner, Markus Fischer; 5.Auflage; Springer Verlag; Berlin: 2014.

[10] UV Absorbance Spectroscopy of Biological Macromolecules; A. Rodger; Encyclopedia of Biophysics; Springer, Berlin, Heidelberg: 2013.

[11] High Throughput Screening for Food Safety Assessment: Fluorescence-based real-time quantitative polymerase chain reaction (qPCR) technologies for high throughput screening; C. Löfstöm, M.H. Josefsen, T. Hansen, M.S.R. Søndergaard, J. Hoorfar; Technical University of Denmark, Kongens Søburg, Denmark: 2015.

[12] Differential Scanning Fluorimetry for Monitoring RNA Stability; Robert Silvers, Heiko Keller, Harald Schwalbe, Martin Hengesbach; 2015.

[13] AGO-Driven Non-Coding RNAs: Small RNA proteome as disease biomarker: An incognito treasure of clinical utility; Jyoti Roy, Neha Jain , Garima Singh, Basudeb Das, Bibekanand Mallick; RNAi and Functional Genomics Lab, Department of Life Science, National Institute of Technology, Rourkela, India: 2019.

[14] Romeis Mikroskopische Technik; Maria Mulisch, Ulrich Welsch; 19. Auflage; Springer Spektrum; Heidelberg: 2015.

[15] Dye selection for live cell imaging of intact siRNA; Markus Hirsch, Dennis Strand, Mark Helm; Biol. Chem.; Vol. 393, pp. 23–35; 2012.

[16] Differential Scanning Fluorimetry for Monitoring RNA Stability: Supporting Information; Robert Silvers, Heiko Keller, Harald Schwalbe, Martin Hengesbach; 2015.

[17] Improving quantitative real-time RT-PCR reproducibility by boosting primer linked amplification efficiency; Ales Tichopad, Anamarija Dzidic, Michael W. Pfaffl; Institute of Physiology, FML-Weihenstephan, Center of Life and Food Science, Technical University of Munich, Munich, Germany: 2002.

[18] Baseline correction; mathworks.com; last call: 29.09.2020.

[19] Influence of Dimethylsulfoxide on RNA Structure and Ligand Binding; Janghyun Lee, Catherine E. Vogt, Mitchell McBrairty, Hashim M. Al-Hashimi; Department of Chemistry and Biophysics, University of Michigan: 2013.

[20] Use of fluorescent DNA-intercalating dyes in the analysis of DNA via ion-pair reversed-phase denaturing high-performance liquid chromatography; Bahrami AR, Dickman MJ, Matin MM, Ashby JR, Brown PE, Conroy MJ, Fowler GJ, Rose JP, Sheikh QI, Yeung AT, Hornby DP; Anal Biochem.: 2002