Foundation of DNA Nanoengineering

Farros Alferro [1] **Ibuki Kawamata** [2]

1. Introduction

In biology, DNA often conveys genetic data. DNA is utilized as a well defined building component, similar to a Lego block, in the scientific subject of DNA nanoengineering to create objects at the nanoscale known as DNA nanostructures. The DNA nanostructure is a synthetic construct that resembles natural molecules and was created to artificially implement cell-like functionalities. One illustration is a useful chemical that can pierce a cell membrane like a needle and create a hole. This hole is then used to destroy a bad cell, such as a cancer cell [1].

Due to the time constraints and difficulty of the experiment, it is not possible to create DNA structures in this experiment 2. This experiment focuses on the foundation for creating these structures: the construction of a DNA double helix by measuring the melting temperature T_M.

2. Principles

2.1. DNA Structure

DNA (*deoxyribonucleic acid*) is the hereditary material in humans and almost all other species. Although the molecule DNA was identified in 1869, its role in genetic heredity was not proven until 1943. Then, in 1953, James Watson and Francis Crick discovered that the structure of DNA is a double-helix polymer, a spiral made up of two DNA strands twisted around one another, with the help of the research of biophysicists Rosalind Franklin and Maurice Wilkins. The discovery advanced scientists' understanding of DNA replication and hereditary control of cellular functions significantly [2].

Although scientists have made minor changes to the Watson and Crick model or elaborated upon it since its inception in 1953, the model's four major features remain the same today. These features are as follows [3]:

[1]Computer Vision Lab., Tohoku University, Sendai, Japan [2]Molecular Robotics Lab., Tohoku University, Sendai, Japan. Correspondence to: Farros Alferro <farros.alferro.t3@dc.tohoku.ac.jp>, Ibuki Kawamata <ibuki.kawamata@tohoku.ac.jp>.

Theme 2 Report on *Laboratory Experiment II* which was done on November 11th, 2022. Copyright 2022 by the author(s).

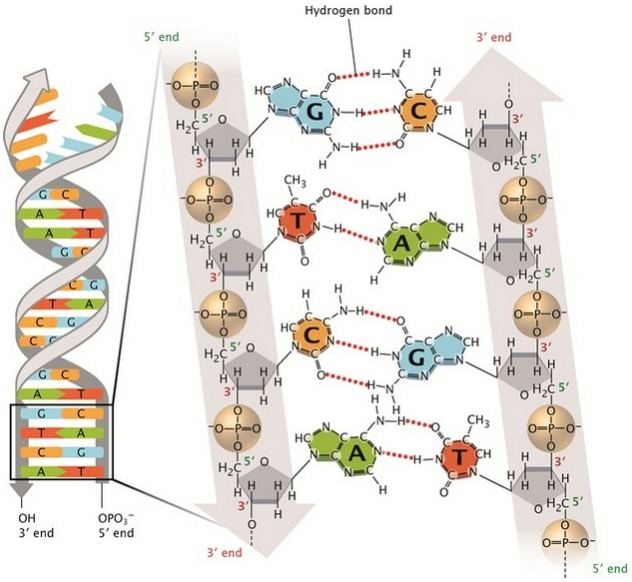


Figure 1. Two hydrogen bonds connect T to A; three hydrogen bonds connect G to C. The sugar-phosphate backbones (grey) run anti-parallel to each other, so that the 3' and 5' ends of the two strands are aligned [3].

- DNA is a double-stranded helix with hydrogen bonds connecting the two strands. A bases are always paired with Ts, and C bases are always paired with Gs, which is compatible with Chargaff's rule and explains for it.
- Most DNA double helices are right-handed; that is, if you hold your right hand out with your thumb pointed up and your fingers curled around your thumb, your thumb represents the helix's axis and your fingers represent the sugar-phosphate backbone. Z-DNA is the only kind of DNA that is left-handed.
- The DNA double helix is anti-parallel, which implies that the 5' end of one strand is coupled with the 3' end of its corresponding strand (and vice versa). Figure 1 shows how nucleotides are linked together by their phosphate groups, which connect the 3' end of one sugar to the 5' end of the next sugar.
- Not only are the DNA base pairs linked by hydrogen bonding, but the outside edges of the nitrogen-containing bases are also exposed and potentially hy-

drogen bonded. These hydrogen bonds allow other molecules, such as proteins that play critical roles in DNA replication and expression, easy access to the DNA (Figure 1).

2.2. Nearest Neighbor Method

By examining nearest-neighbor interactions, the thermodynamics of enthalpy change (ΔH°), entropy change (ΔS°), and free energy change (ΔG°) for nucleic acid helix creation and melting temperature (T_m) for helix melting may be described [4]. As stated in the equations, the value of ΔH° and ΔS° can be calculated by adding the initial values $\Delta H^\circ_{Initial}$ or $\Delta S^\circ_{Initial}$ and the sum of stacking parameters ΔH°_{Stack} or ΔS°_{Stack}. The summation approach is the same whether the column to lookup in the table is ΔH° or ΔS°. The stacking parameter's value varies based on the pair of surrounding nucleotides, and its value is found empirically, as detailed in the table 1.

Sequence Pair	ΔH° ($kcal\ mol^{-1}$)	ΔS° ($cal\ mol^{-1}\ K^{-1}$)
Initial	0.6	-9.0
5'-AA-3'/3'-TT-5'	-8.0	-21.9
5'-AC-3'/3'-TG-5'	-9.4	-25.5
5'-AG-3'/3'-TC-5'	-6.6	-16.4
5'-AT-3'/3'-TA-5'	-5.6	-15.2
5'-CA-3'/3'-GT-5'	-8.2	-21.0
5'-CC-3'/3'-GG-5'	-10.9	-28.4
5'-CG-3'/3'-GC-5'	-11.8	-29.0
5'-CT-3'/3'-GA-5'	-6.6	-16.4
5'-GA-3'/3'-CT-5'	-8.8	-23.5
5'-GC-3'/3'-CG-5'	-10.5	-26.4
5'-GG-3'/3'-CC-5'	-10.9	-28.4
5'-GT-3'/3'-CA-5'	-9.4	-25.5
5'-TA-3'/3'-AT-5'	-6.6	-18.4
5'-TC-3'/3'-AG-5'	-8.8	-23.5
5'-TG-3'/3'-AC-5'	-8.2	-21.0
5'-TT-3'/3'-AA-5'	-8.0	-21.9

Table 1. Improved thermodynamic parameters for DNA/DNA double helix initiation and propagation in $1M$ NaCl buffer. Estimated errors in ΔH°, ΔS° and ΔG°_{37} are $\pm 0.3\ kcal\ mol^{-1}$, $\pm 1.3\ cal\ mol^{-1}\ K^{-1}$ and $\pm 0.1\ kcal\ mol^{-1}$, respectively.

2.3. Hybridization and Melting Temperature

In genomics, hybridization is the process by which two complementary single-stranded DNA or RNA molecules join together to generate a double-stranded molecule (Figure 2). The bonding is determined by the correct base pairing between the two single-stranded molecules [5]. DNA denaturation (melting), on the other hand, is the process of separating ds-DNA (double-stranded DNA) into two ss-DNA (single-stranded DNA). This cooperative unwinding is sometimes referred to as helix-coil unwinding or melting transition. DNA melting happens across a narrow temperature range, causing changes in its physical properties. Since the 1950s, it has been known that heating a DNA solution above room temperature causes strand separation [6].

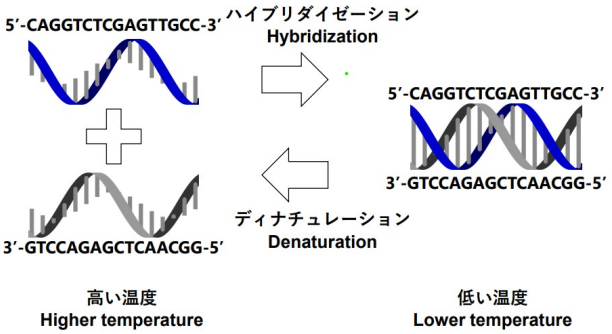


Figure 2. Illustration of denaturation and hybridization processes [1].

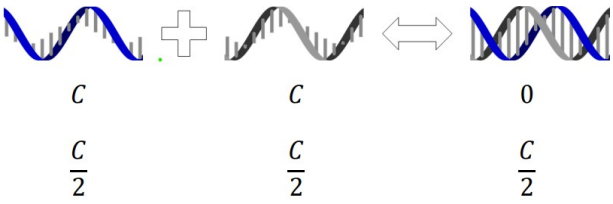


Figure 3. Concentration of each molecule in hybridization process [1].

As previously stated, DNA melts when heated to a certain temperature. The melting temperature is the temperature at which the concentration of single- and double-stranded DNAs is the same when a solution containing complementary DNAs is heated from low to high temperature [1]. The melting temperature is determined by a number of factors, including the length of the DNA (shorter sections melt more quickly), the nucleotide sequence composition, and salt concentration (ionic strength of the added salt), and is typically between 50°C and 100°C [6].

It is possible to theoretically estimate the melting temperature by combining the equilibrium equation and the free energy equation. Initially, the concentration of both ss-DNA is equal to C as depicted in Figure 3. After heating and hybridization, each strand's concentration becomes half its initial value, and the formed ds-DNA also has a $\frac{C}{2}$ concentration. The equilibrium constant can then be determined by dividing the right-hand side concentration by the product of

the left-hand side concentrations:

$$K = \frac{\frac{C}{2}}{(\frac{C}{2})^2} = \frac{2}{C} \tag{1}$$

On the other hand, The relation between the free energy ΔG° of double-stranded DNA and the equilibrium constant K is calculated as follows:

$$\Delta G^\circ = -RT \ \ln K \tag{2}$$

where R is the gas constant and T is the absolute temperature. Simultaneously, free energy G° can be expressed using ΔH° (change in enthalpy) and ΔS° (change in entropy), as indicated in below equation:

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ \tag{3}$$

Because ΔS° is a negative value, ΔG° rises as temperature rises, making double-stranded DNA unstable. When the temperature rises, it becomes more difficult to stay in ds-DNA, resulting in denaturation, and the ratio of ss-DNA rises. When the temperature drops, ss-DNAs attach to each other, and the ratio of ds-DNA increases.

By substituting $T = T_M$, one can derive the theoretical value of T_M from equation (1) $\sim$ (3):

$$\begin{aligned} \Delta H^\circ - T_M \Delta S^\circ &= -RT \ \ln \frac{2}{C} \\ \Delta H^\circ &= T_M (S^\circ - R \ln \frac{2}{C}) \\ T_M &= \frac{H^\circ}{(S^\circ + R \ln \frac{C}{2})} \end{aligned} \tag{4}$$

2.4. Förster Resonance Energy Transfer

It is impossible to examine DNA directly due to its size of 2 nm, which is far smaller than the wavelength of visible light. T_M is then measured experimentally using an indirect method involving a fluorophore known as FRET (Förster resonance energy transfer). A fluorophore and quencher-modified DNA are introduced for measuring purposes, as shown in Figure 4. When exposed to excitation light in the single-stranded condition, a fluorophore glows green. When two DNA molecules hybridize and create a double helix shape, a quencher trails the fluorophore. Even when the excitation light is irradiated, the fluorescence becomes very faint because energy transfers from the fluorophore to the quencher. As a result, ss-DNA has strong fluorescence properties, whereas ds-DNA has weak fluorescent properties [1].

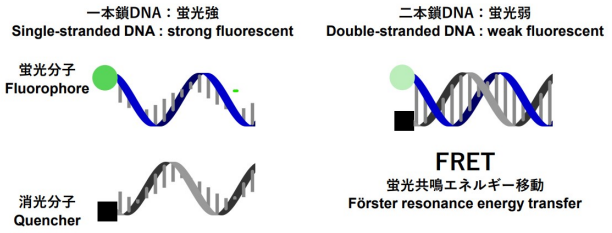


Figure 4. Illustration of FRET (Förster resonance energy transfer) method [1].

3. Experiment

For this experiment, two complementary DNA and a buffer of NaCl solution are mixed with different volumes and concentrations. The DNA sequence used in this experiment is 5'-GTCAGTTCCAG-3', and its complementary 5'-CTGGAACTGAC-3'. There are four configurations in total, where each configuration results in the DNA concentration of $0.5\mu M$, $1\mu M$, $2\mu M$, and $4\mu M$ by varying the volume of both the DNA solution and the buffer, as shown in Table 2. Moreover, three copies are created for each configuration in order to address error issue.

Solution	Volume (μL)	Concentration (μM)	
		Original	Final
DNA (GTCAGTTCCAG)	Y	2 or 8	X
DNA (CTGGAACTGAC)	Y	2 or 8	X
Buffer	Z		
Total	20		

Table 2. Configurations of DNA solution.

When the combined solutions are finished, the test tubes containing them are vibrated by VORTEX [7] and whirled by a centrifuge machine. The test tubes are then placed inside the PCR (Polymerase Chain Reaction) amplification equipment [8]. By irradiating the excitation light from the light source at the top to the DNA solution at the bottom, this equipment can read the fluorescence of the solution. The fluorescence is then measured by the left detector, while changing the temperature from $25°C$ to $75°C$ by $0.5°C$ per 5 seconds. The illustration of the equipment is depicted in Figure 5.

3.1. Experimental Melting Temperature Measurement

We collect data from the experiment on the relationship between temperature and fluorescence intensity. When this

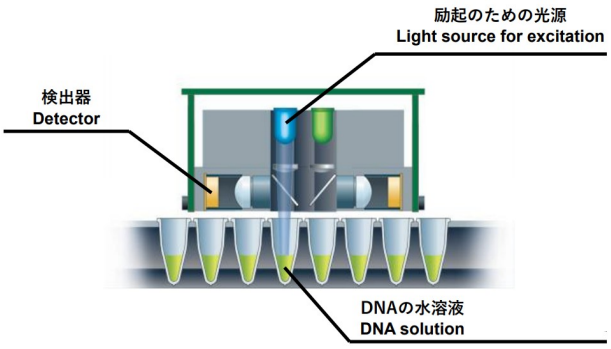
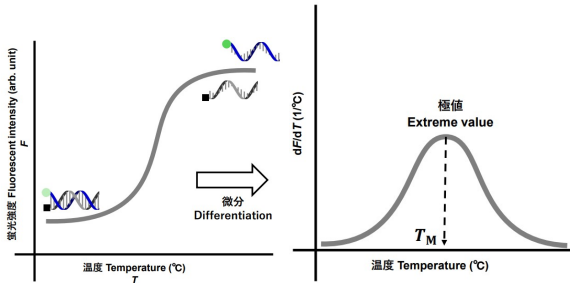


Figure 5. Equipment used to measure fluorescent intensity [1].

data is plotted, it forms a curve, as illustrated in Figure 6. As already stated in Section 2.3, the fluorescence intensity is low when the temperature is low, indicating a high ds-DNA ratio and vice versa. One approach for finding T_M is to examine the graph's inflection point, which implies differentiating Figure 6(a) and obtaining the greatest value, as illustrated in Figure 6(b). Because the data is discrete, the difference in temperatures is used to discriminate.



(a) Before differentiation (b) After differentiation

Figure 6. Curves of relation between temperature and fluorescent intensity before and after differentiation [1].

3.2. Experimental Enthalpy and Entropy Measurement

Enthalpy change (ΔH°) and entropy change (ΔS°) can be experimentally obtained by transforming equation (4) as follows:

$$\frac{1}{T_M} = \frac{R}{\Delta H^\circ} \ln \frac{C}{2} + \frac{\Delta S^\circ}{\Delta H^\circ} \tag{5}$$

Using this equation, a fit line can be drawn, and the value of the slope and the y-interception from the linear relationship between $\ln \frac{C}{2}$ and $\frac{1}{T_M}$ are acquired. This graph is called the Van't Hoff plot (Figure 7). The slope is used to compute ΔH°, which later will be used to determine ΔS° altogether with the y-intercept value.

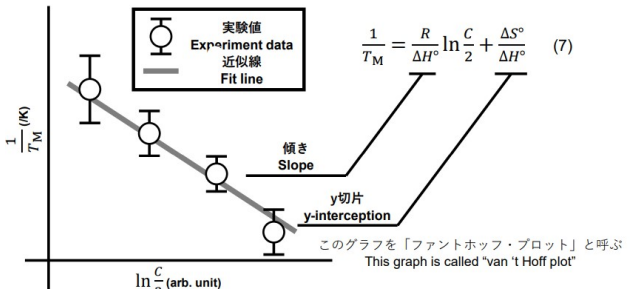


Figure 7. Linear relationship between $\ln \frac{C}{2}$ and $\frac{1}{T_M}$ [1].

4. Results and Discussions

4.1. Assignment 1

Discussion: *How many DNA molecules are there in a solution of* 20 µL *with concentration of* 4 µM*?*

$$\begin{aligned} M &= \frac{n}{V} \rightarrow n = M \cdot V \\ n &= 4 \cdot 10^{-6} M \ \times 20 \cdot 10^{-6} L \\ n &= 8.0 \cdot 10^{-11} mol \end{aligned} \tag{6}$$

Where M, n, and V are molar concentration, moles of solute, and liters of solution, respectively. Calculating the number of molecules:

$$\begin{aligned} molecules &= n \times N_A \\ &= 8.0 \cdot 10^{-11} mol \ \times 6.02 \cdot 10^{23} mol^{-1} \\ &= 4.8 \cdot 10^{13} molecules \end{aligned} \tag{7}$$

Where N_A is the Avogadro Constant valued $6.02 \cdot 10^{23}\ mol^{-1}$.

4.2. Assignment 2

Discussion: *Measure experimental value of melting temperature T_M for each concentration.*

The plot of the relationship between fluorescent intensity and temperature for each concentration and its copies is shown in Figure 8(a). It can be seen that for one measurement of concentration $4\mu M$, the intensity value is already high even when the temperature is low. An incorrect procedure might cause this during the experiment: only putting one kind of the ss-DNA (either the + or -) into the test tube. This results in the ss-DNA not combining in low temperatures, or in other words, fluorophore intensity is not reduced by the quencher. During the lecture, it is told not to use the $4\mu M$ concentration as it might skew the data. However, in this report we will use the $4\mu M$ concentration as a comparison for assignment 6.

Furthermore, the derivative of Figure 8(a) is also depicted in Figure 8(b). From this plot, we obtained the melting

temperature T_M for each concentration from the maximum point of each graph and averaging it across the three samples, which is shown in Table 3. To calculate the derivative, we divides the fluorescent intensity difference with the temperature difference between two consecutive points. On the other hand, since the number of samples is reduced by one, the temperature value is also processed by taking the mean between two consecutive points [9]:

$$(x_i',\ y_i') = \left(\frac{(x_{i+1} + x_i)}{2},\ \frac{(y_{i+1} - y_i)}{(x_{i+1} - x_i)}\right) \tag{8}$$

where $(x_i',\ y_i')$ denotes the i$^{\text{th}}$ derivative coordinate.

From Table 3, we can observe that there is a tendency for melting temperature to be higher when the concentration is also high. One plausible reason that might explain this is that as the concentration is high, the ds-DNA ratio is higher, resulting in a stronger hydrogen bond between one pair of ss-DNA. Thus, extra energy is needed to break this bond, which means a higher temperature. On the other hand, concentration $4\mu M$ has lower melting temperature value than concentration $2\mu M$ that might be caused by the experimental error stated above.

4.3. Assignment 3

Discussion: *Derive equation (4) from equations (1) $\sim$ (3).*

This assignment has been discussed in Section 2.3

4.4. Assignment 4

Discussion: *By nearest neighbor method, compute theoretical values of H° and S° of the DNA sequence used in the experiment.*

For enthalpy (ΔH°) :

$$\begin{aligned}\Delta H^\circ &= H^\circ_{initial} + H^\circ_{GT} + H^\circ_{TC} + H^\circ_{CA} \\ &\quad + H^\circ_{AG} + H^\circ_{GT} + H^\circ_{TT} + H^\circ_{TC} \\ &\quad + H^\circ_{CC} + H^\circ_{CA} + H^\circ_{AG} \\ \Delta H^\circ &= 0.6 - (9.4 + 8.8 + 8.2 + 6.6 \\ &\quad + 9.4 + 8.0 + 8.8 + 10.9 + 8.2 + 6.6) \\ \Delta H^\circ &= -84.3\ cal\ mol^{-1}\end{aligned} \tag{9}$$

For entropy (ΔS°) :

$$\begin{aligned}\Delta S^\circ &= S^\circ_{initial} + S^\circ_{GT} + S^\circ_{TC} + S^\circ_{CA} \\ &\quad + S^\circ_{AG} + S^\circ_{GT} + S^\circ_{TT} + S^\circ_{TC} \\ &\quad + S^\circ_{CC} + S^\circ_{CA} + S^\circ_{AG} \\ \Delta S^\circ &= -9.0 - (25.5 + 23.5 + 21.0 + 16.4 \\ &\quad + 25.5 + 21.9 + 23.5 + 28.4 + 21.0 + 16.4) \\ \Delta S^\circ &= -232.1\ cal\ mol^{-1}\ K^{-1}\end{aligned} \tag{10}$$

4.5. Assignment 5

Discussion: *Compute theoretical value of melting temperature T_M for each concentration used in the experiment.*

The theoretical value of melting temperature is calculated for each concentration using equation 4 by substituting the value of ΔH° and ΔS° obtained from assignment 4.4 as well as the ideal gas constant $R = 1.987\ cal\ K^{-1}\ mol^{-1}$ [10]:

$$T_M = \frac{-84.3 \times 10^3}{(-232.1 + 1.987 \cdot \ln \frac{C}{2})} \tag{11}$$

where C denotes the DNA solution concentration.

The value of melting temperature is shown in Table 3. Similar to experimental value, it can be seen that the melting temperature is higher for higher concentration solution. Thus, it can be concluded that the experimental result agrees with the theoretical result.

Concentration	Melting Temprature (K)		Difference
(μM)	Theoretical	Experimental	(%)
4.0	326.52	333.57	2.16
2.0	324.79	335.07	3.17
1.0	323.08	331.57	2.63
0.5	321.38	331.40	3.12

Table 3. Comparison between theoretical and experimental value of melting temperature. % difference is calculated with respect to theoretical value.

However, the experimental value is typically a little percentage greater than the theoretical value (measured in Kelvin). Several factors could account for this disparity: (1) The differentiation method used, which is the central difference. There might be more accurate methods, such as the forward difference, backward difference, or Newton's series. (2) The experimental result is not deemed credible because the temperature value rise is relatively considerable at $0.5°C$. There is a potential that the intensity varies rapidly between these consecutive values at some point, which might be the melting temperature of that curve. A more accurate result can be obtained by decreasing the increment.

4.6. Assignment 6

Discussion: *(Optional) Measure experimental values of ΔH° and ΔS° using Van't Hoff plot.*

The Van't Hoff plot is depicted in Figure 9. Here, as already mentioned, we also created Van't Hoff plot without concentration $4\mu M$ to show how this particular point affects the value of ΔH° and ΔS°. The value of the slope

and y-intercept are already included in the plots. As already mentioned in Section 3.2, we calculate the value of ΔH° and ΔS° using the aforementioned variables, which are displayed in Table 4.

The fit line was calculated using formulas [11]:

$$a = \frac{N\sum_{i=1}^{N} x_i\, y_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i}{\Delta} \tag{12}$$

$$b = \frac{\sum_{i=1}^{N} x_i^2 \sum_{i=1}^{N} y_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} x_i\, y_i}{\Delta} \tag{13}$$

$$\Delta = N\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2 \tag{14}$$

$$y = ax + b \tag{15}$$

where the equation of the fit line is represented in equation 15. From the line equation of Figure 9(a), we can see that the concentration $4\mu M$ does affect the final value by shifting the slope higher, almost half of the slope of Figure 9(b). Even though the coordinate of concentration $4\mu M$ is still in the same range as the other concentrations, its presence significantly deviates the slope due to the data's high sensitivity (very small values).

Moreover, it can be observed that the ΔH° and ΔS° values obtained from the plot without $C = 4\mu M$ are much closer to the theoretical value than the one with concentration $4\mu M$. From this, we can deduce that $4\mu M$ is indeed a measurement error, and it is better to omit this data.

	Theoretical	Experimental	
		Without $4\mu M$	With $4\mu M$
ΔH°	-84.3	-83.3	-152.4
ΔS°	-232.1	-221.7	-429.6

Table 4. Comparison between theoretical and experimental values of ΔH° $[kcal\ mol^{-1}]$ and ΔS° $[cal\ mol^{-1}\ K^{-1}]$.

5. Conclusion

From this experiment, we learn about DNA structure and DNA nanoengineering. We also study to determine the melting temperature of DNA for different concentration by examining the inflection point of intensity vs. temperature plot. Finally, we discuss enthalpy and entropy change using two methods: Nearest neighbor method and Van't Hoff plot.

References

[1] Ibuki Kawamata. "Foundation of DNA Nanoengineering". In: 2022. URL: https://drive.google.com/file/d/1DjgBkl4xwVQFw2gCoArn-OGmfpUHxwVI/view (see pp. 1–4).

[2] T Britannica. "DNA". In: *Encyclopedia Britannica* (Aug. 2022). URL: https://www.britannica.com/science/DNA (see p. 1).

[3] L Pray. "Discovery of DNA structure and function: Watson and Crick". In: *Nature Education* 1.1 (2008), p. 100. URL: https://www.nature.com/scitable/topicpage/discovery-of-dna-structure-and-function-watson-397/ (see p. 1).

[4] N. Sugimoto. "Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes". In: *Nucleic acids research* 24.22 (1996), pp. 4501–4505. URL: https://doi.org/10.1093/nar/24.22.4501 (see p. 2).

[5] Ellen Sidransky. *Hybridization*. May 2022. URL: https://www.genome.gov/genetics-glossary/hybridization (see p. 2).

[6] G. Khandelwal and J. Bhyravabhotla. "A phenomenological model for predicting melting temperatures of DNA sequences". In: *PloS one* 5.8 (2010). URL: https://doi.org/10.1371/journal.pone.0012433 (see p. 2).

[7] *Vortex-Genie 2*. URL: https://www.scientificindustries.com/vortex-genie-2.html (see p. 3).

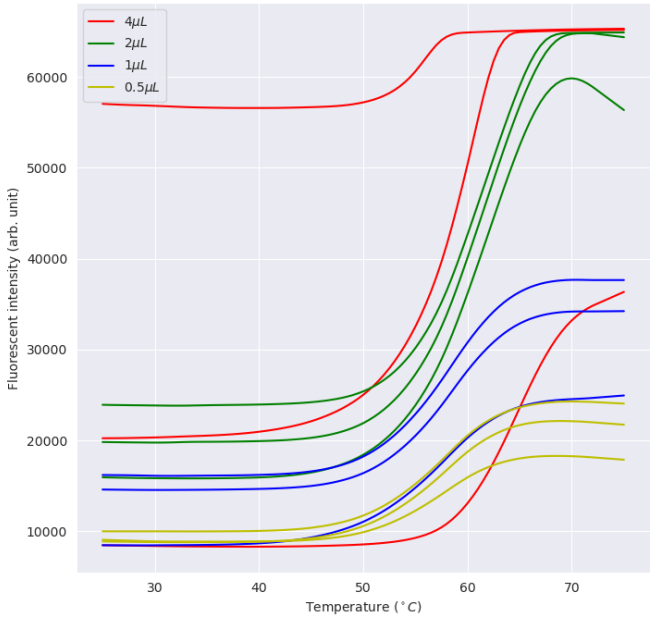
[8] *Real-Time PCR Systems*. URL: https://www.bio-rad.com/en-jp/category/real-time-pcr-systems?ID=059db09c-88a4-44ad-99f8-78635d8d54db (see p. 3).

[9] HagesLab. *How to: Numerical Derivative in Python*. 2020. URL: https://www.youtube.com/watch?v=utRKIlOZbtw (see p. 5).

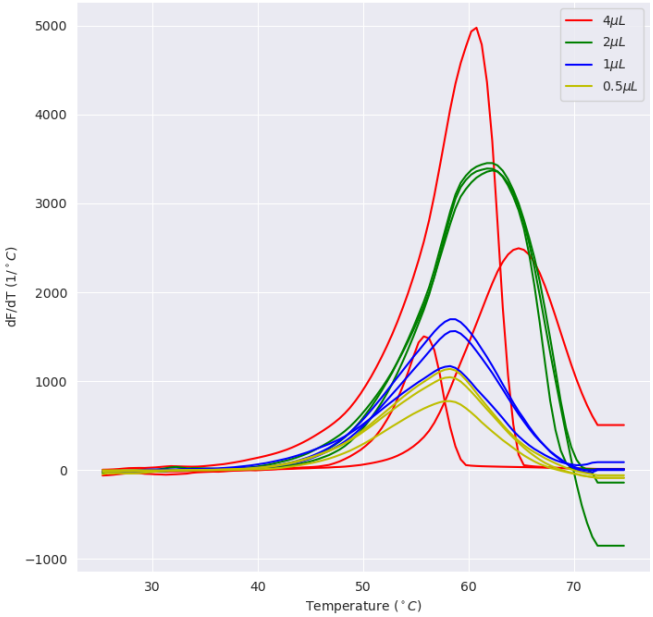
[10] *Gas Constant*. URL: https://www.chemeurope.com/en/encyclopedia/Gas_constant.html (see p. 5).

[11] Wei Gao. In: *Measurement and Instrumentation*. Asakura Publishing Company, 2017, pp. 53–54. URL: https://www.asakura.co.jp/detail.php?book_code=20165 (see p. 6).

A. Result on Assignment 5



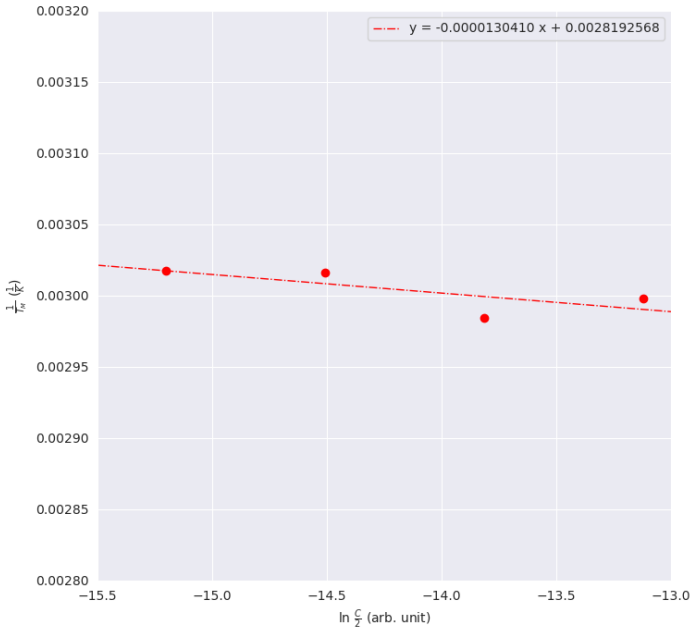
(a) Relationship between fluorescent intensity and temperature.



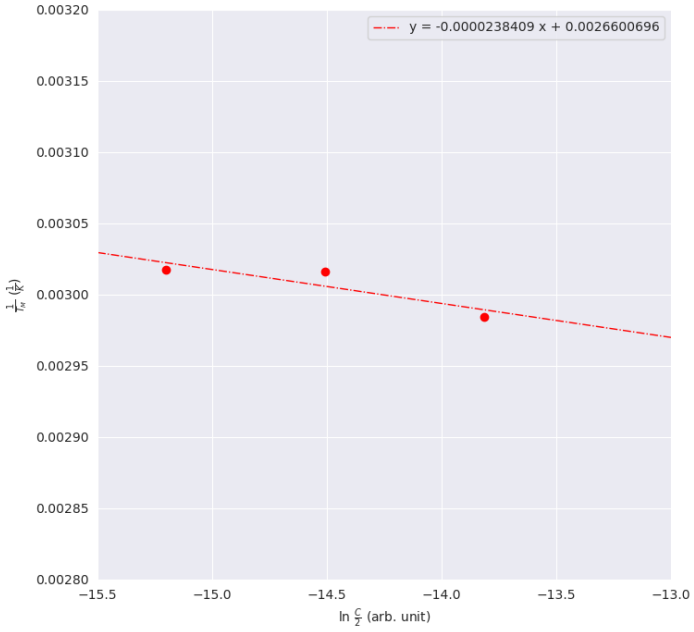
(b) Derivation of Figure 8(a).

Figure 8. Overall figures for Assignment 5

B. Result on Assignment 6



(a) Van't Hoff plot of the data.



(b) Van't Hoff plot of the data without $C = 4.0\mu M$.

Figure 9. Overall figures for Assignment 6