

# 3CL SARS-CoV-2 3CL Protease Analyzing

Yan Haoming

September 27

September 28

October 11

October 18

2024

# Chapter 1

## Introduction

### 1.1 Protein Purification

#### 1.1.1 *E.coli* Lysis: Homogenizer

To extract the protein expressed by *E.coli* we should homogenize the bacteria to get its lysate. Homogenizer is used to break down the cell wall of bacteria considering its relatively thicker membrane compared with animal cells.

#### 1.1.2 Column Chromatography

Before analyzing the protein, we should purify it first with the help of column chromatography including affinity chromatography and filtration chromatography. The affinity chromatography is used first for prominent separation while the filtration chromatography is used for further separation from which we can also deduce the molecular weight of the protein roughly.

#### 1.1.3 Analyzing: SDS-PAGE

SDS-PAGE is the apparatus and methods for the electrophoresis of protein. SDS (detergent) is used to denature the protein and grant negative charge to them. Instead of using agarose sugar, polyacrylamide gel is used to form the lane for electrophoresis.

### 1.2 Enzyme Kinetics

#### 1.2.1 Verifying Michaelis Menton Equation



$$V_0 = \frac{k_{cat}E_0[S]}{K_m + [S]} = \frac{V_{max}[S]}{K_m + [S]}$$

The fluorescent substrate relates to **FRET**, which means the substrate will be fluorescent only when catalyzed by protease. Specifically speaking, DABCYL plays the role of dark quencher<sup>[5]</sup> while the EDANS<sup>[3]</sup> as the donor of fluorescence. The substrate (polypeptide) is stuck in the middle as shown in Figure 1.1.

The overall consequence is that, when the substrate is consumed, then, it will show fluorescence. Further process of reaction will lead to more intensive fluorescence which can be detected by **ELISA** in the unit of RFU.

#### 1.2.2 Detecting Half-maximal Inhibitory Concentration

The structure of inhibitor is shown in Figure 1.2. It will repress the reaction catalyzed by 3CL protease probably by competitive binding to the enzyme. And we will determine the **Half-maximal Inhibitory Concentration**,  $IC_{50}$  through speed of reaction.

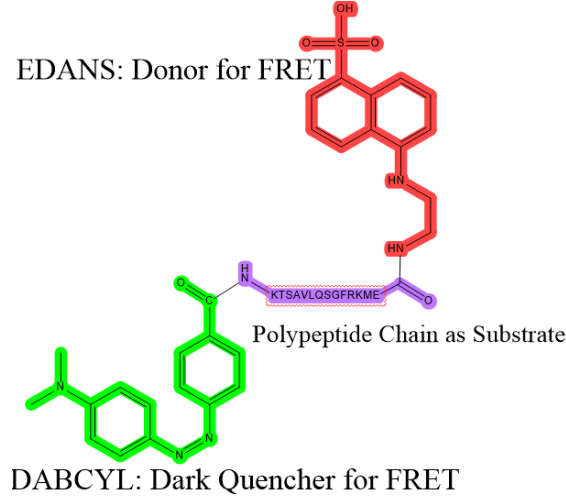


Figure 1.1: Structure of Fluorescent Substrate

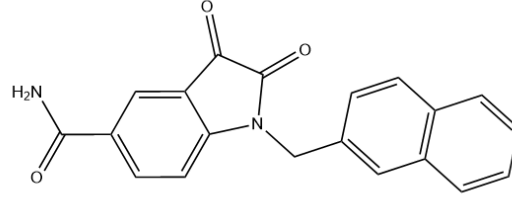


Figure 1.2: Structure of Inhibitor

### 1.3 Structure Detection: Structure Detection: CD Spectroscopy

The secondary structure of protein can be assessed by **Circular Dichroism Spectroscopy (CD Spectroscopy)** <sup>[6]</sup>. The spectroscopy harnesses the differences in absorption of left-handed versus right-handed circularly polarized light. As for the range of wave length, the far UV region (190 to 250 nm) is used to analysis. The  $\alpha$  helix and  $\beta$  sheets have characteristic CD spectra due to their peptide bond folding states.

To completely define the secondary structure of polypeptide segment, we should also figure out the  $\phi$  and  $\psi$  angles of each amino acid residues. The  $\phi$  and  $\psi$  stand for the important dihedral angles in a peptide.

$$\phi = \text{dihedral}(C - N - C_{\alpha} - C) \quad (1.1)$$

$$\psi = \text{dihedral}(N - C_{\alpha} - C - N) \quad (1.2)$$

Two quantities are often used to describe the result:  $\theta$ ,  $\Delta\epsilon$  and  $[\theta]$ . The transformation and definition of them are shown below:

$$\Delta\epsilon = \epsilon_L - \epsilon_R \quad (1.3)$$

$$\Delta A = A_L - A_R \quad (1.4)$$

$$\Delta A = \Delta\epsilon l c \quad (1.5)$$

$$[\theta] = \frac{100\theta}{lcN_{\text{amino acid residues}}} \quad (1.6)$$

$$\tan \theta = \frac{E_R - E_L}{E_R + E_L} \approx \theta \quad (1.7)$$

$$\theta \approx \frac{\sqrt{I_R} - \sqrt{I_L}}{\sqrt{I_R} + \sqrt{I_L}} \quad (1.8)$$

$$\theta \approx \Delta\epsilon l c \times \frac{\ln 10}{4} \times \frac{180}{\pi} \quad (1.9)$$

In the above equations,  $\theta$  is called ellipticity,  $[\theta]$  is **mean residue molar ellipticity** and  $\Delta\epsilon$  is **molar circular dichroism**.

## 1.4 Thermodynamic Parameters Measuring

**Isothermal Titration Calorimetry (ITC)** is used to determine thermal dynamic parameters including enthalpy, entropy, equilibrium constant. During the measurement, both reference cell and sample cell are heated to maintain the constant temperature while the input of heat impulse will be recorded precisely for calculation.

The procedure and logic behind it are as following:

- Figure out the stoichiometric ratio  $N$  according to the corresponding position of maximal slope.
- Determine  $\Delta H$  from the difference between final and initial heat.
- Fit the curve to get the equilibrium constant  $K$  (corresponds to the value of maximal slope).
- Calculate  $\Delta G = -RT \ln K$
- Calculate  $\Delta S$  by the definition of  $\Delta G = \Delta H - T\Delta S$ .

## 1.5 Model of Protein Binding

### 1.5.1 Virtual Screening: By Molecular Docking

We use the method of **Molecular Docking** to predict the binding orientation and affinity of twenty different Ligands, which is able to achieve enough accuracy with a relatively short period. The docking engine chosen by us is AutoDock Vina.

### 1.5.2 Interaction Analysis

After determining the most suitable ligand, I analyzed the interaction between the ligand and receptor in PyMOL with the purpose of providing an interpretation of its strong affinity.

# Chapter 2

## Experimental Process

### 2.1 Protein Purification

#### 2.1.1 *E.coli* Lysis: Homogenizer

##### Materials

- Buffer Lysis
  - PBS (Phosphate Buffered Saline:  $\text{HPO}_4^{2-} / \text{H}_2\text{PO}_4^-$ ): 20 mM (pH=7.3)
  - NaCl: 500 mM
  - Imidazole: 10 mM
  - DTT (Dithiothreitol: Reducing Agent): 1 mM
  - PMSF (Phenylmethanesulfonyl Fluoride: Serine Protease Inhibitor): 1 mM

##### Procedure

- Stir the buffer lysis containing *E.coli* until without obvious lumps and precipitates in an ice bath.
- Filter the mixture (50 mL) by gauze and then pour into the **homogenizer**<sup>[4]</sup> with 10 mL  $\text{H}_2\text{O}$  following the below settings:
  - Raise the pressure to 1,000 bar for 2 minutes.
  - Pause 1 minute.
  - Repeat the cycle 3 times.
- Centrifugate the liquid at 14,000 rpm for 30 minutes.
- Strain the product through a filter (0.45  $\mu\text{m}$ ) into a graduated cylinder.

#### 2.1.2 Affinity Chromatography: Immobilized $\text{Ni}^{2+}$ Metal Ion + His-tag

##### Materials

- Buffer A
  - PBS (Phosphate Buffered Saline:  $\text{HPO}_4^{2-} / \text{H}_2\text{PO}_4^-$ ): 20 mM (pH=7.3)
  - NaCl: 500 mM
  - Imidazole: **10 mM**
- Buffer B
  - PBS (Phosphate Buffered Saline:  $\text{HPO}_4^{2-} / \text{H}_2\text{PO}_4^-$ ): 20 mM (pH=7.3)
  - NaCl: 500 mM
  - Imidazole: **500 mM**
- Buffer C

- PBS (Phosphate Buffered Saline:  $\text{HPO}_4^{2-} / \text{H}_2\text{PO}_4^-$ ): 40 mM (pH=7.3)
- NaCl: 100 mM
- EDTA: 1 mM

## Procedure

We purified 3CL protease fist by affinity chromatography. Immobilized  $\text{Ni}^{2+}$  ion had been attached to the column while His-tag (six histidine in the terminal of protein) had been fused with 3CL protease.

The  $\text{Ni}^{2+}$  column has the capacity of 5 mL. We washed the column before loading using buffer A and buffer B in turn (A(25 mL)-B(25 mL)-A(15 mL)-B(15 mL)-A(15 mL)) as shown in the blue curve in Figure 3.1 and Figure 3.2. We set pressure difference ( $\Delta P = 0.3$  MPa) to be the control threshold with the pre-column pressure ( $P_{pre} = 0.5$  MPa). The flow rate was set to be 1 mL/min for above operation.

Then we loaded 55 mL sample after lysis into the column with the flow velocity to be 2 mL/min. After that, buffer A was constantly pumped into to column to elute impurities until the signal of UV absorbance went back near 0 mAU. Once the eluting was complete, we started to raise the proportion of buffer B linearly to form a linear gradient. The threshold of scope to collect product was set to be 20% buffer elution (within the 0% to 100% buffer elution gradient). The fraction collecting was performed automatically with the amount of 1.8 mL/tube. Finally, we collected tubes from number 3 to number 8, totally  $6 \times 1.8 = 10.8$  mL.

Condensation with centrifuging filtering and solvent changing was used condense 10.8 mL solution to nearly 1.5 mL. The centrifuging was in  $4^\circ\text{C}$  with the rotating speed of 3,500 rpm for 10 minutes. In each turn, we decanted 5 mL solvent for the Gel-filtration Column to replace the buffer for  $\text{Ni}^{2+}$  column containing lots of imidazole.

### 2.1.3 Gel-filtration Chromatography

#### Procedure

The Gel-filtration Column has the capacity of 24 mL. We set pressure difference ( $\Delta P = 1.8$  MPa) to be the control threshold with the pre-column pressure ( $P_{pre} = 5.0$  MPa). Careful washing of the column and sample ring was performed. As for the sample ring, 4 mL ultra-pure water with 4 mL buffer assay were used to wash the pipe. The two direction of washing(up to down and vice versa) with careful injector operation was included to ensure any bubble would be excluded from the column.

The loading amount should be 500  $\mu\text{L}$  with the concentration of 1 mg/mL. To ensure the sample concentration to be around 1 mg/mL, we detected the concentration of the sample by nano drop. The result was 6.737 Abs equivalent to  $6.737 \times 1.07 = 7.21$  mg/mL. That means around 70  $\mu\text{L}$  concentrated protease was needed with 430  $\mu\text{L}$  buffer assay. However we enlarged the system by two fold which means diluting 140  $\mu\text{L}$  concentrated protease with 860  $\mu\text{L}$  buffer assay in case of failure.

We collected according to the UV absorbance as shown in Figure 3.4 and Figure 3.3. Then condensed the sample as previous process in affinity chromatography.

$$A = \epsilon cl \quad (2.1)$$

$$A = 6.737 \text{ Abs} \quad (2.2)$$

$$\epsilon = 0.937 \text{ L/g} \cdot \text{cm} \quad (2.3)$$

$$l = 1 \text{ cm} \quad (2.4)$$

$$c \approx 7.21 \text{ mg/mL} = \frac{7.21 \text{ g/L}}{34 \times 10^3 \text{ g/mol}} = 212.1 \text{ } \mu\text{M} \quad (2.5)$$

### 2.1.4 Analyzing: SDS-PAGE

#### Procedure

In this experiment, we used precast polyacrylamide gel. However, we still need to prepare the buffer(500 mL) for electrophoresis by hands.

Four type of samples were prepared for SDS-PAGE to check the effect of protein purification.

Table 2.1: Substrate Solution Configuration in Microplate

Number of Sample	1	2	3	4	5	6	7	8
Substrate Concentration ( $\mu\text{M}$ )	500	750	1000	1250	1500	2000	2500	3000
DMSO Solvent ( $\mu\text{L}$ )	9	8.5	8	7.5	7	6	5	4
5000 $\mu\text{M}$ Concentrated Substrate ( $\mu\text{L}$ )	1	1.5	2	2.5	3	4	5	6

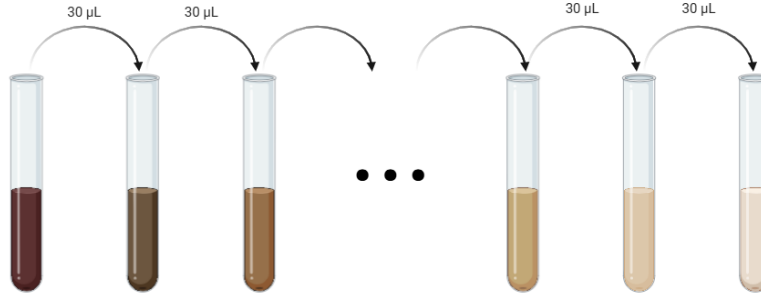


Figure 2.1: Serial Dilution of Inhibitor

- No.1 *E.coli* lysate
- No.2 Supernate of lysate after centrifuging
- No.3 Sample after affinity chromatography
- No.4 Sample after gel-filtration chromatography

In each well, we put 20  $\mu\text{L}$  sample with 5  $\mu\text{L}$  loading buffer except for No.3 which only left 8  $\mu\text{L}$ . To control the concentration in consistency, only 2  $\mu\text{L}$  loading buffer was added to No.3.

We mixed the solution by **Vortex Genie 2** and centrifuge for instant. Then heated at 94  $^{\circ}\text{C}$  for 3 minutes. Finally, loaded marker (2  $\mu\text{L}$ ) as well as the four types of sample (10  $\mu\text{L}$ ) into the polyacrylamide gel in sequence.

The electrophoresis apparatus was set at 150 V working for 38 minutes.

After electrophoresis, the gel was stained using Coomassie Brilliant Blue with heating in microwave oven for three times. In each turn, the paint was heated until boiling. 3% Acetic acid- $\text{H}_2\text{O}$  was used to destain with the similar procedure in microwave oven.

## 2.2 Enzyme Kinetics

The following experiments involve tedious configuration of solutions with different concentrations, including substrate, 3CL Protease (Enzyme), and inhibitor. To summarize and visualize the com-

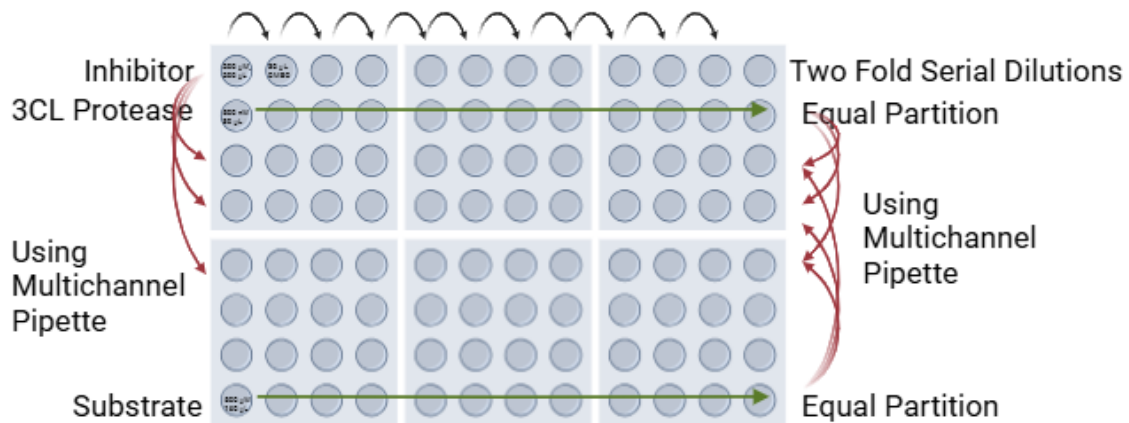


Figure 2.2: Solution Configuration in Microplate

plicated process of dilution, I have drawn the Figure 2.2. Generally speaking, we used **multi-channel pipette** to transfer a whole row of solution to other three different rows because of the three independent parallel groups. Before that, however, we must use **single-channel pipette** to adjust the concentration in one row. The **two fold serial diluting**, as shown in Figure 2.1, was used for inhibitor while substrate was diluted one by one differently as shown in Table 2.1.

### 2.2.1 Verifying Michaelis Menton Equation

#### Procedure

The independent variable in Michaelis Menton Equation is the concentration of substrate  $[S]$ . Thus to verify the equation we must configure the substrate solution of different concentrations. We started by dissolving 2 mg solid substrate into 192  $\mu\text{L}$  DMSO solvent, which has the concentration of 5000  $\mu\text{M}$ . Then, we adjusted the ratio of DMSO and concentrated substrate to achieve 8 different concentrations of substrate solution, as shown in Table 2.1.

The Composition of the System (100  $\mu\text{L}$ ):

- Buffer Assay (80  $\mu\text{L}$ , 80%)
  - $\text{Na}_2\text{HPO}_4 / \text{NaH}_2\text{PO}_4$  (40 mM)
  - NaCl (100mM)
  - EDTA (1mM)
  - Triton X-100 (0.1%)
- SARS-CoV2-3CL Protease (10  $\mu\text{L}$ , 10%) (before mixing: 500 nM, after mixing: 50 nM)
- Fluorescent Substrate (10  $\mu\text{L}$ , 10%) (before mixing the different concentrations can be seen in Table 2.1).
- Temperature: 25°C

3CL protease was preserved in fridge at  $-80^\circ\text{C}$ , which was then melted in ice box. The initial concentration of 3CL protease was 206  $\mu\text{M}$ , and then we took 2  $\mu\text{L}$  diluting it with 2 mL Buffer Assay to get 500 nM protease solution. The buffer assay was prepared by adding 50  $\mu\text{L}$  Triton X-100 into 50 mL buffer solution at the very beginning.

Both buffer assay and 3CL protease were partitioned equally into eight different wells in the same microplate, and were mixed using multichannel pipette in other three rows. Fluorescent substrate was stored in another microplate while was prepared according to Table 2.1. Also, multichannel pipette was used to transfer substrate into the system.

We detected the speed of reaction by the concentration variation of fluorescent substrate which was identified by **enzyme-linked immunosorbent assay (ELISA)**. The original data is shown in Figure 3.6.

### 2.2.2 Detecting Half-maximal Inhibitory Concentration

#### Procedure

The inhibitor is consist of competitive inhibitor, uncompetitive inhibitor and mixed inhibitor. Instead of changing substrate concentration simultaneously, we varied inhibitor concentration  $[I]$  only and keep  $[S]$  under control. Also, we focused on **Half-maximal Inhibitory Concentration ( $\text{IC}_{50}$ )** only, without concerns about modified factors to  $V_{max}$  or  $K_m$ .

The following procedure was used to prepare to inhibitor solution with different concentration:

- sample 20  $\mu\text{L}$  concentrated inhibitor(20 mM) solution with 180  $\mu\text{L}$  DMSO as additional solvent. Then mix in 1<sup>st</sup> well of microplate, which has the concentration of 200  $\mu\text{M}$ .
- prepare 30  $\mu\text{L}$  DMSO in each well of 1<sup>st</sup> raw.
- take 30  $\mu\text{L}$  solution from 1<sup>st</sup> to 2<sup>nd</sup> and mix in 2<sup>nd</sup> well.
- repeat the step for the following wells to perform two fold serial dilution as shown in Figure 2.1.



Table 2.2: Configuration of Solution for CD Spectroscopy

No.	Guanidine Final Concentration (M)	Guanidine Volume (mL)	CD Buffer (mL)	Protein (mL)
1	1.2	0.2	0.8	1
2	2.4	0.4	0.6	1
3	3.6	0.6	0.4	1
4	4.8	0.8	0.2	1
5	6.0	1.0	0.0	1

To get the solution of substrate, we decanted 1.92 mL DMSO into 2 mg fluorescent substrate then separated it into twelve different wells in the same row. As shown in Figure 2.2, the substrate and inhibitor were in different rows. To be noted that, in reality, they were also stored in separated microplates to avoid unwanted illumination of fluorescent substrate.

The Composition of the System (100  $\mu$ L):

- Buffer Assay (80  $\mu$ L, 80%)
  - $\text{Na}_2\text{HPO}_4$  /  $\text{NaH}_2\text{PO}_4$  (40 mM)
  - NaCl (100mM)
  - EDTA (1mM)
  - Triton X-100 (0.1%)
- SARS-CoV2-3CL Protease (10  $\mu$ L, 10%) (before mixing: 500 nM, after mixing: 50 nM)
- Fluorescent Substrate (5  $\mu$ L, 5%) (before mixing: 500  $\mu$ M, after mixing: 25  $\mu$ M)
- Inhibitor (5  $\mu$ L, 5%) (after mixing the serial concentration can be seen in Figure 3.9).
- Temperature: 25°C

## 2.3 Structure Detection: Structure Detection: CD Spectroscopy

### Materials

- CD Buffer (pH=7.4):  $\text{KH}_2\text{PO}_4$  20 mM 19.8 mL +  $\text{K}_2\text{HPO}_4$  20 mM 80.2 mL
- Guanidinium Chloride 12M
- Protein Solution 2 mg/mL

### Procedure

We took the measurement from the wavelength of 190 nm to 340 nm in the cuvette with light path equals 10 mm. The corresponding sample (protein+guanidine) dosage is 2 mL where 1 mL coming from 2 mg/mL protein the other 1 mL coming from diluted guanidinium chloride. Therefore, we diluted and mixed the protein with guanidine according to the Table 2.2.

Before measuring we installed the apparatus and filling the light path with  $\text{N}_2$ . The mixture of guanidinium chloride with protein was put in ice box for 1 minute to denature the protein. However we adapted the more practical control of time by limiting the whole process, including adding protein to cuvette then guanidine, to be 3 minutes. Finally we took measurement for five different samples.

## 2.4 Thermodynamic Parameters Measuring

For practical purposes, we chose the reaction of  $\text{MgCl}_2$  and EDTA to be measured instead of the binding of inhibitor with protein. The measurements were performed twice under 25°C and 30°C respectively with the parameters shown in Table 2.3.

Since most of the steps were performed automatically, the only process that remained difficult was loading or injecting the solution into the cells. Strict 280  $\mu$ L was injected into the cells confirmed by the steady curve in the initial delay period.

Table 2.3: Experimental Parameters of ITC

Total Injections	16
Cell Temperature ( $^{\circ}\text{C}$ )	25 and 30
Preference Power (pCal/sec)	11
Initial Delay (sec)	60
Syringe Concentration (mM)	5
Cell Concentration (mM)	0.40
Stiring Speed (rpm)	1000

## 2.5 Model of Protein Binding

### 2.5.1 Virtual Screening: By Molecular Docking

We used AutoDock Vina to be the docking engine for the experiment, which had been installed on our laptops' linux environment (windows subsystem). There were twenty ligand candidates to be the inhibitor and we would choose the one having maximal binding affinity. To perform molecular docking in Vina, a key step was determining the center and size of vina box for searching. I uploaded the 7EN8 protein data to the website called CAVITYPLUS<sup>[2]</sup> to get box size and box center which corresponds to the binding pocket of the protein.

I screened the inhibitor according to the maximal binding affinity of each ligand.

### 2.5.2 Interaction Analysis

After selecting the most suitable ligand out, I observed the ligand and receptor in the software called PyMOL. The nearby residues surrounding the ligand were kept while irrelevant residues were hidden for clarify. Given the type of amino acid residues and the distance and conformation, I can deduce the interaction between the ligand and receptor.

## Chapter 3

# Results and Discussion

### 3.1 Protein Purification

#### 3.1.1 *E.coli* Lysis: Homogenizer

As the initial step, homogenization must succeed, which was proved by the subsequent results. However, for the homogenization itself, the clarifying solution was the indication of success.

#### 3.1.2 Column Chromatography

##### Affinity Chromatography: Immobilized $\text{Ni}^{2+}$ Metal Ion + His-tag

The tracking of UV absorbance changing and conductance changing with varied buffer B proportion can be seen in Figure 3.1 and Figure 3.2. The major peak in the middle of the Figure 3.1 corresponds to the impurities (protein that has UV absorbance accompanied by nucleic acid). Meanwhile the conductance drops down (Figure 3.2) due the poor conductivity of protein compared with inorganic salts.

From the Fourth Column in Figure 3.5 we can clearly see the effective purification produced by affinity chromatography compared with the lysis mixture in second and third column.

##### Gel-filtration Chromatography

The effect can be seen in Figure 3.5. Compared with the fourth column, the last column after gel-filtration chromatography does not show additional purification. Probably, the solution had been purified to an extreme extend after affinity chromatography, which leads to the unobvious improvement.

Besides, the molecular weight can be estimated according to the volume where we observed the peak of UV absorbance in Figure 3.4. The UV peak locates near 10.1 mL, and therefore the molecular weight is around 43 kDa (between 13.7 kDa to 43 kDa), which is consistent with the

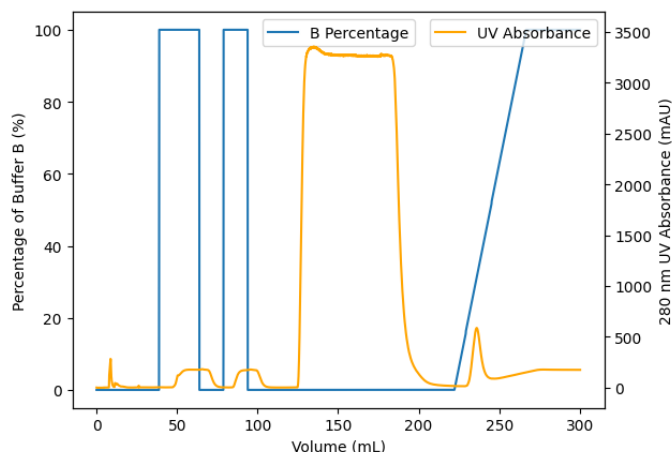


Figure 3.1: Affinity Chromatography: UV Absorbance with Percentage of Buffer B

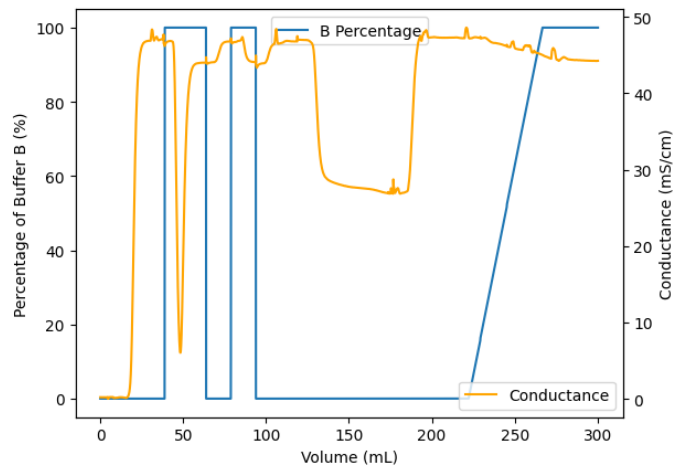


Figure 3.2: Affinity Chromatography: Conductance with Percentage of Buffer B

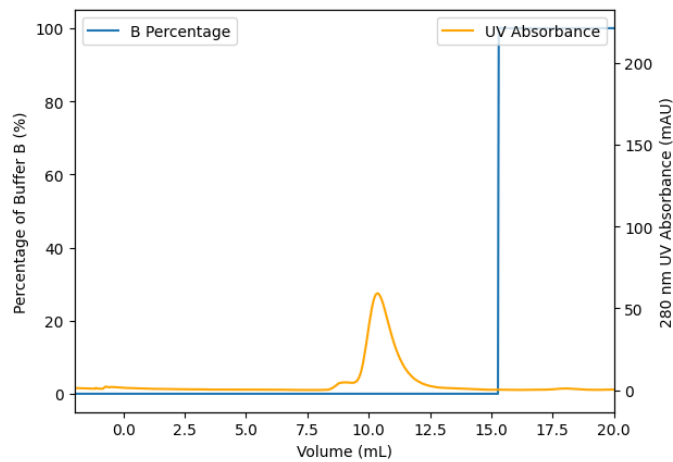


Figure 3.3: Gel-filtration Chromatography: UV absorbance with Percentage of Buffer B

more exact result gotten from SDS-PAGE as following.

molecular weight  $\approx 43$  kDa (according to Gel-filtration Chromatography)

### 3.1.3 Analyzing: SDS-PAGE

As shown in Figure 3.5, the effect of purification is obvious. The most effective method should be affinity chromatography while other steps are also playing an important role in the whole process. Besides, we can also figure out the molecular weight of 3CL protease which is near **34 kDa**.

molecular weight  $\approx 34$  kDa (according to SDS-PAGE)

The result of 34 kDa should be more accurate than the one obtained from Gel-Filtration Chromatography above.

## 3.2 Enzyme Kinetics

### 3.2.1 Verifying Michaelis Menton Equation

The original data gotten from **ELISA** is visualized in Figure 3.6. In each subfigure, I have drawn the three parallellised groups of same substrate concentration. It should be noted that, the several initial points were omitted for keeping linear parts only in accordance with the assumptions of Michaelis Menton Equation (the  $[ES]$  has reached steady state, not changing with time).

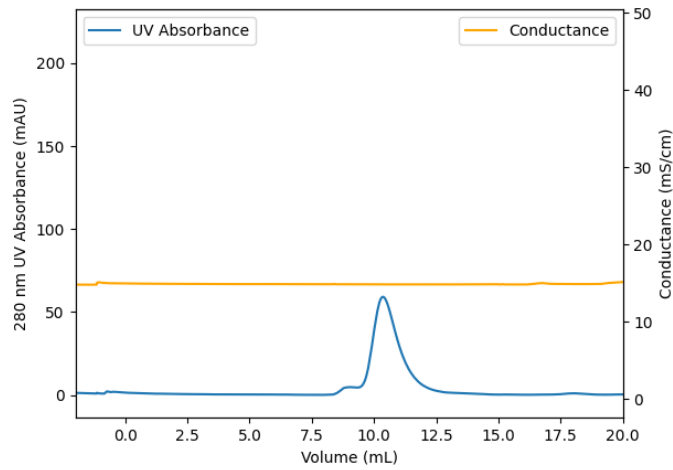


Figure 3.4: Gel-filtration Chromatography: UV absorbance with Conductance

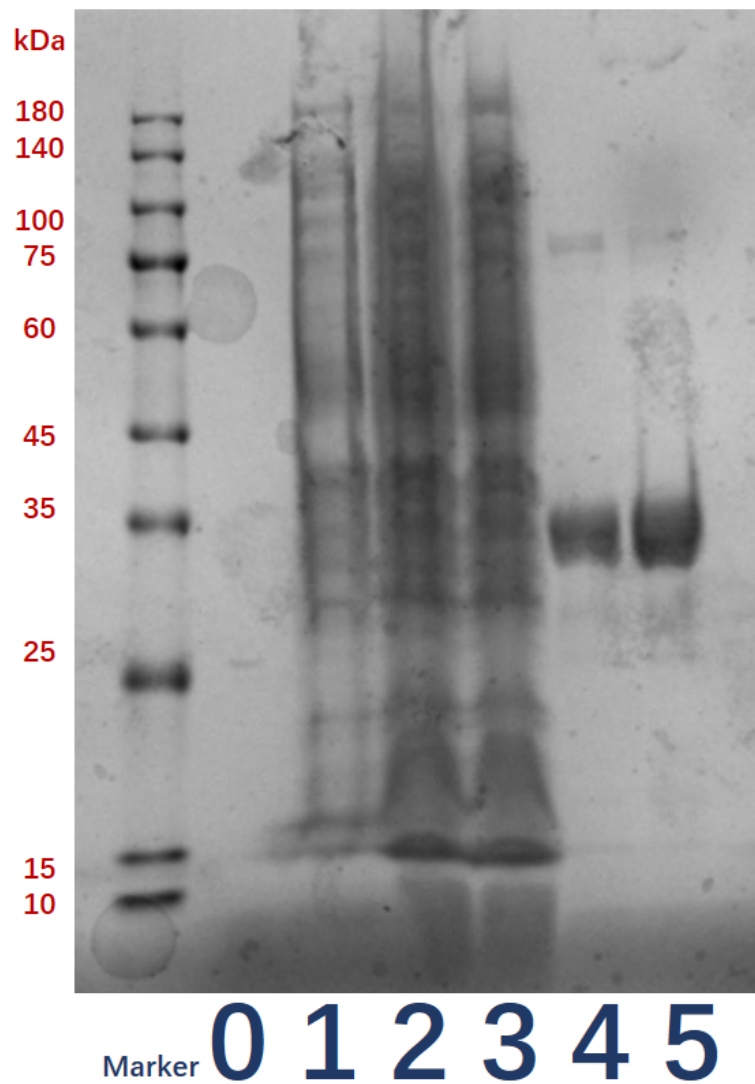


Figure 3.5: Results of SDS-PAGE: The Second Column is the solution after lysis; The Third Column is from supernate; The Fourth Column is after affinity chromatography; The Last One is after gel-filtration chromatography. The First Column has the same composition as the second one just with additional loading buffer. The leftmost one is marker.

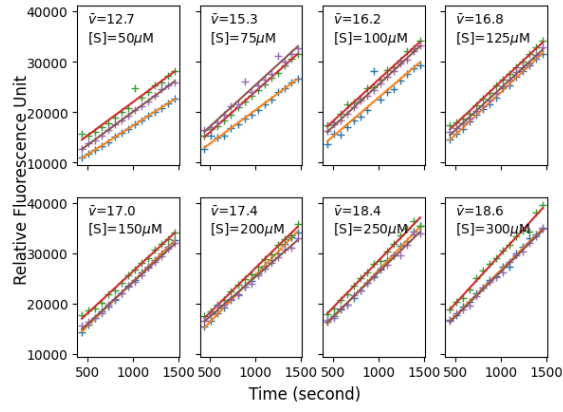


Figure 3.6: Reaction Rate versus Substrate Concentration: 8 groups of experiments, each has 3 parallel groups

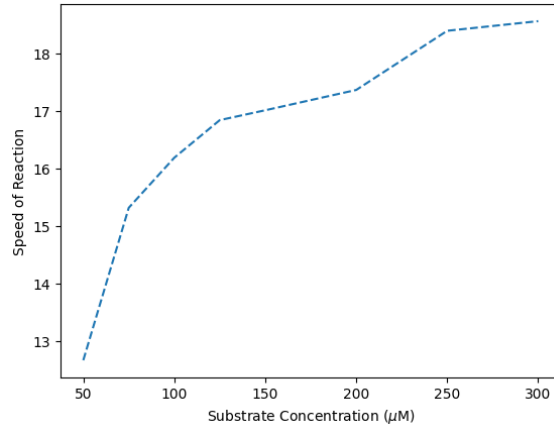


Figure 3.7: Reaction Rate versus Substrate Concentration: in normal axes

I performed linear regression to each of them, then averaged the slope which represents the speed of reaction.

The eight slopes (speed of reaction) were plotted with respect to substrate concentration  $[S]$  then. As shown in Figure 3.7, the consequence is a curve which was similar to the prediction of Michaelis Menton Equation.

To further analyze the results, however, I converted the axes to the reciprocal of speed  $V_0^{-1}$  and substrate concentration  $[S]^{-1}$ . The benefit of reciprocal plotting is that we can operate linear regression again.

$$\frac{1}{V_0} = \frac{K_m}{V_{max}} \cdot \frac{1}{[S]} + \frac{1}{V_{max}}$$

$$y = kx + b$$

The vertical intercept  $b$  is  $\frac{1}{V_{max}}$  and the horizontal intercept corresponds to  $-\frac{1}{K_m}$ , which are also annotated in Figure 3.8.

$$V_{max} = 20 \text{ RFU/second}$$

$$K_m = 28.99 \text{ } \mu\text{M}$$

### 3.2.2 Detecting Half-maximal Inhibitory Concentration

The reaction speed was also detected by **ELISA**. Twelve subfigures are shown in Figure 3.9 and each of them corresponds to a different inhibitor concentration  $[I]$ . There are two intriguing phenomena in the figure. Firstly, one group in  $[I]=10 \text{ } \mu\text{M}$  and one group in  $[I]=1.25 \text{ } \mu\text{L}$  exist extreme fluctuations which may caused by bubbles in solution. Secondly, the reaction speed has negative

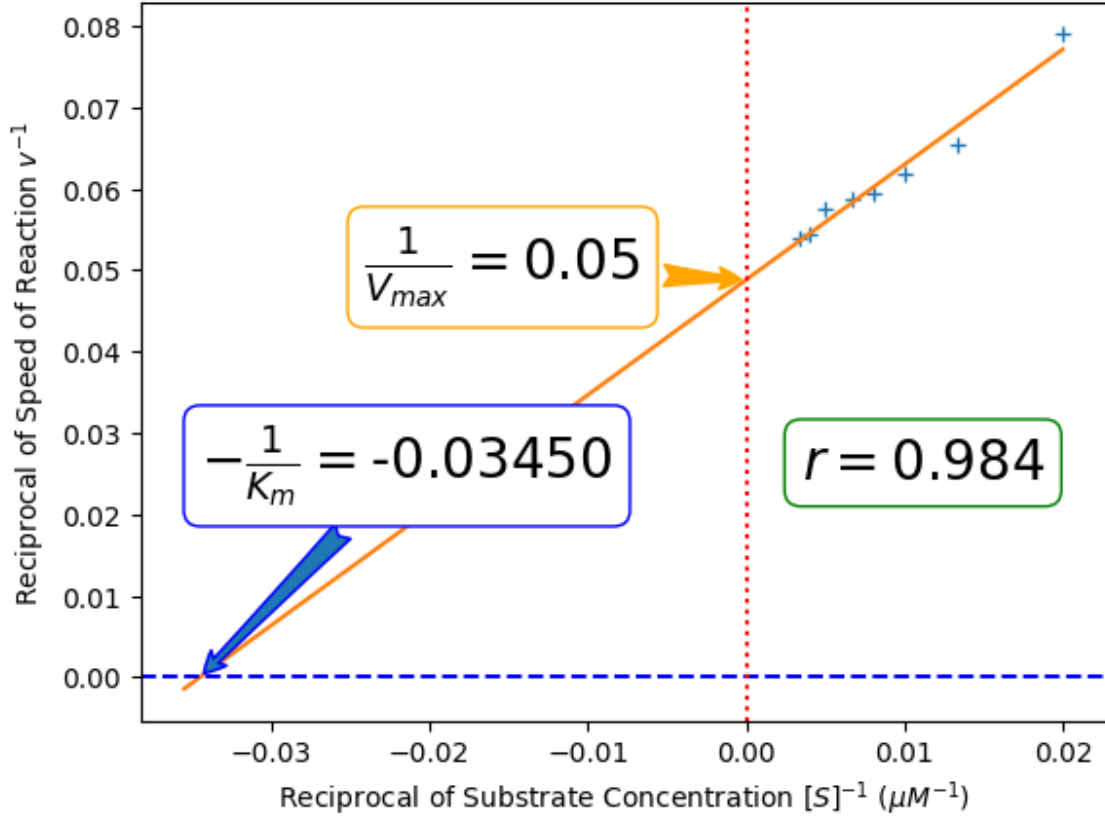


Figure 3.8: Reaction Rate versus Substrate Concentration: in reciprocal axes

value in high concentration of inhibitor. The major cause of that absurd phenomenon is probably the unsteady state with fluctuation at early initial period. To eliminate the negative value and disturbance of it, I also decided to omit the early four points, which should be reasonable and appropriate.

I performed linear regression again to each group and averaged the parallel groups to get the mean speed of reaction. Then the plot of mean speed  $\bar{V}_0$  versus inhibitor concentration was drawn. From Figure 3.10, we can identify the **Half-maximal Inhibitory Concentration  $IC_{50}$** , which is  $IC_{50} = 0.095 \mu M$ . Somehow, the value is higher than the reported value  $IC_{50} = 0.095 \mu M$ .

$$IC_{50} = 0.095 \mu M$$

The  $IC_{50} = 0.095 \mu M$  measured by us was twice as the reported value  $IC_{50} = 0.045 \mu M$ , which may be caused by slightly different experimental conditions or by partial inefficiency of inhibitor solution.

Besides, we can also figure out the inhibitory rate with different inhibitory concentration  $[I]$ . The intuitive recognition can be acquired in Figure 3.10 while the precise values are shown in Table 3.1.

### 3.3 Structure Detection: Structure Detection: CD Spectroscopy

The raw data gotten from CD spectrometer is shown in Figure 3.11. Besides, we have already known the following parameters:

$$\text{Number of Residues} = 306 \text{ residues} + 5 \text{ histidines} = 311 \quad (3.1)$$

$$\text{Concentration} = 29.4 \mu M \quad (3.2)$$

$$\text{Pathlength} = 1 \text{ cm} \quad (3.3)$$

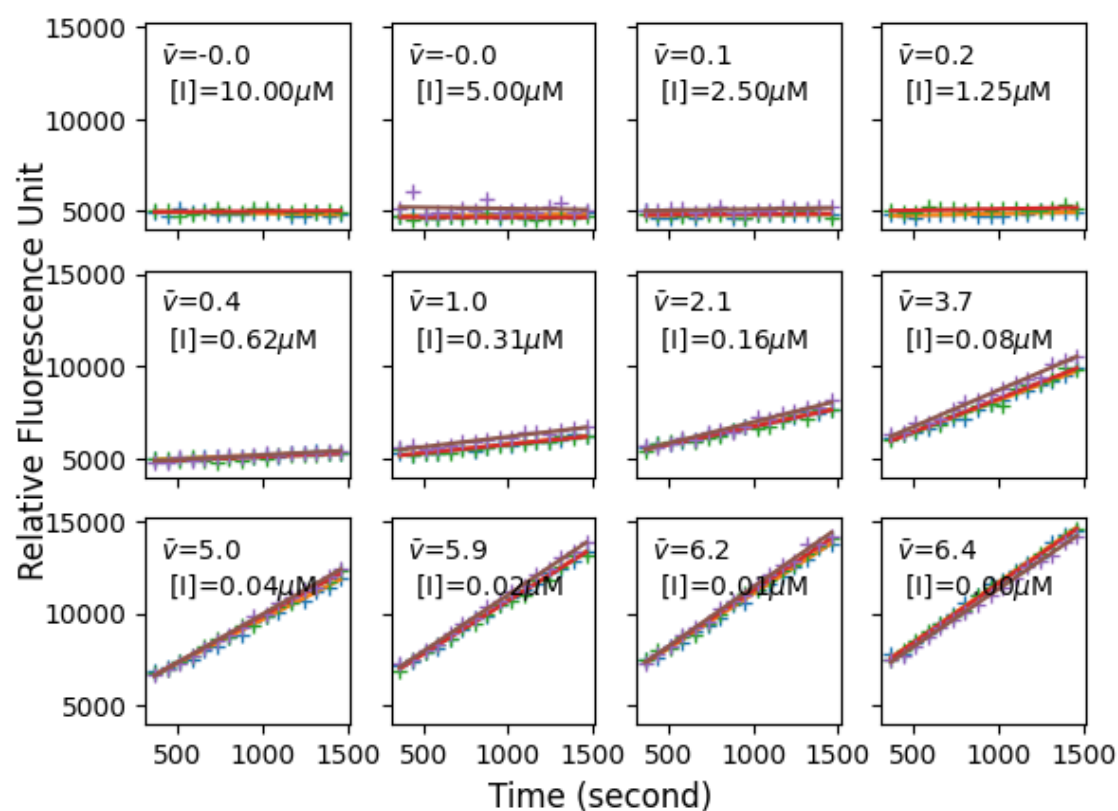


Figure 3.9: Reaction Rate versus Inhibitor Concentration: 12 groups of experiments, each has 3 parallel groups

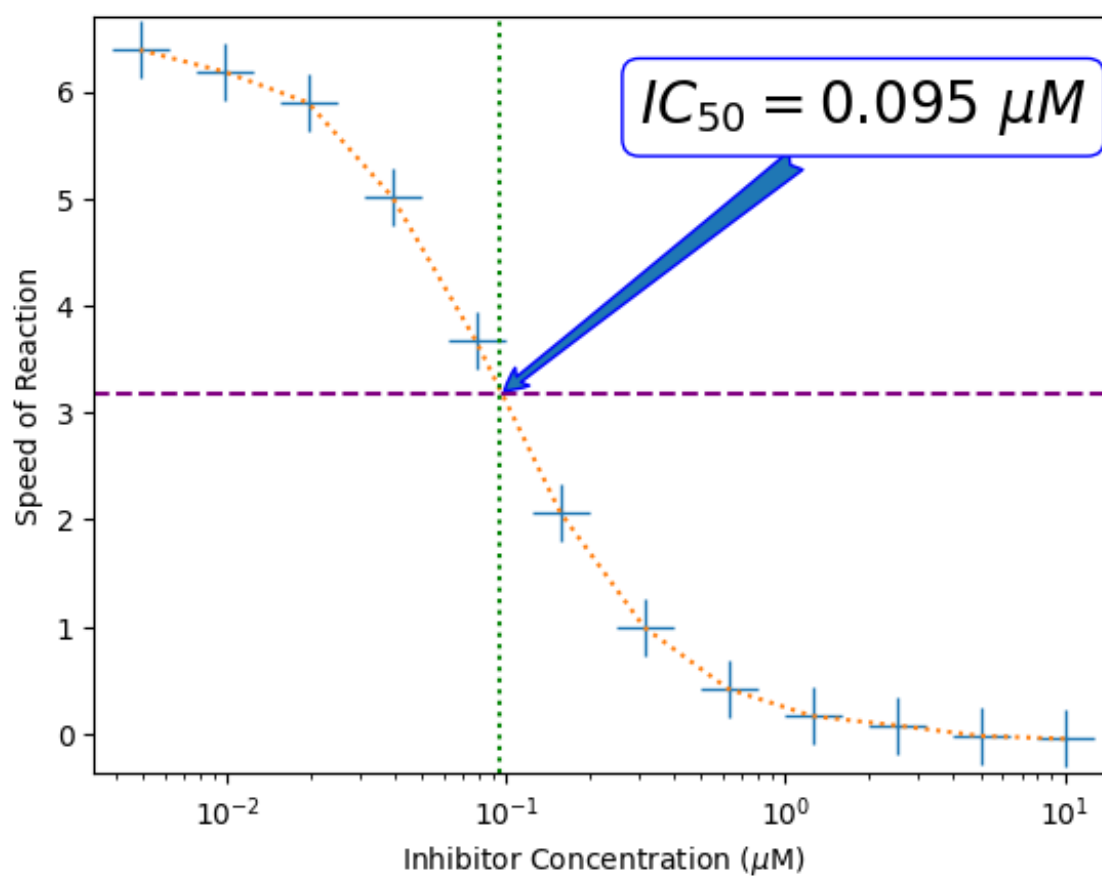


Figure 3.10: Reaction Rate versus Inhibitor Concentration: determining  $IC_{50}$



Table 3.1: Inhibitory Rate with Different Inhibitor Concentration

	Inhibitor Concentration ( $\mu\text{M}$ )	Inhibitory Rate
0	10.00	100.00 %
1	5.00	99.59 %
2	2.50	98.07 %
3	1.25	96.72 %
4	0.62	92.77 %
5	0.31	83.84 %
6	0.16	67.12 %
7	0.08	42.29 %
8	0.04	21.31 %
9	0.02	7.65 %
10	0.01	3.15 %
11	0.00	0.00 %

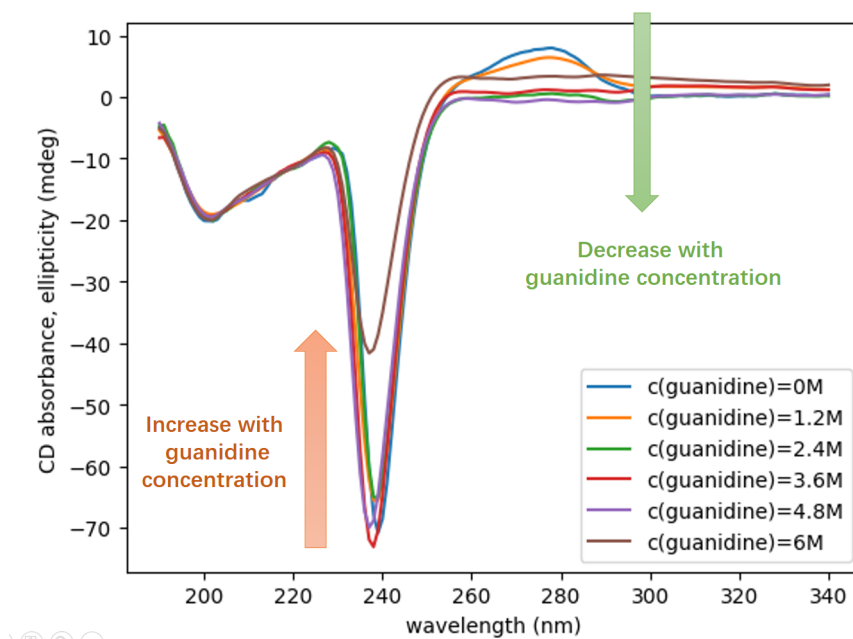


Figure 3.11: CD Absorbance

After submitting the raw data to the website<sup>[1]</sup> to get the proportion of different secondary structures (alpha helix and beta sheets) as shown in Figure 3.12, we still could not tell the differences between the five groups. However, we could still figure out the influence of guanidinium chloride.

As shown in Figure 3.11, in the far-UV region (190 nm to 260 nm), especially the peak near 240 nm, the CD absorbance (ellipticity) increased with guanidinium chloride concentration. While, in the near-UV region (250 nm-320 nm), especially the CD absorbance peak extending from 260 nm to 280 nm, decreased and eventually diminished with the increasing guanidinium chloride concentration. This phenomenon reflects the influence of guanidine to the secondary structure of protein, however, lacking of more concrete implication about the percentage of alpha-helix and beta-sheet.

To analyze the reasons of the phenomenon, we can observe the pie chart shown in Figure 3.12. The proportion of alpha helix is rather small while the unstructured portion accounts for nearly half of the protein. Thus, I deduced the denaturing of protein even in the absence of guanidine. Besides, the CD spectroscopy could not provide extremely accurate measurement which could only verify assumptions rather than figure out the protein structure from bottom.

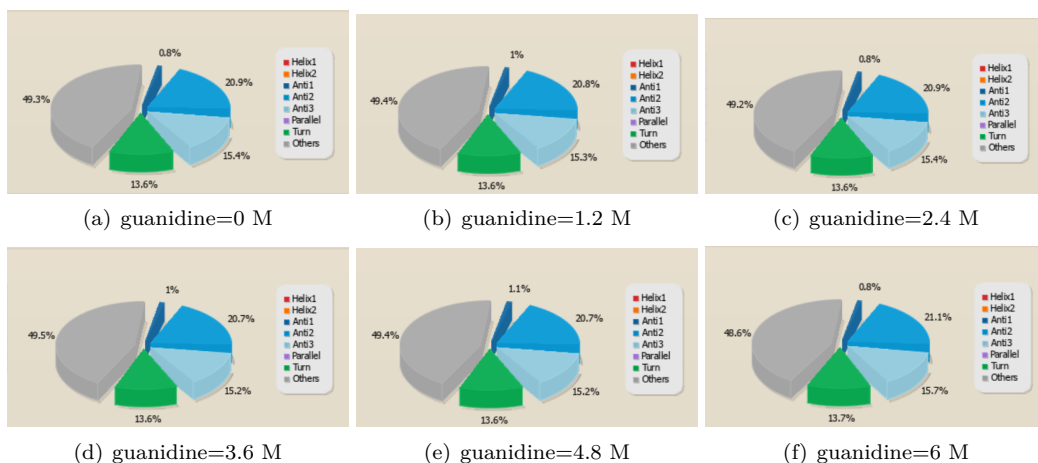


Figure 3.12: Secondary Structure of 3CL Protease with Different Guanidine Concentration

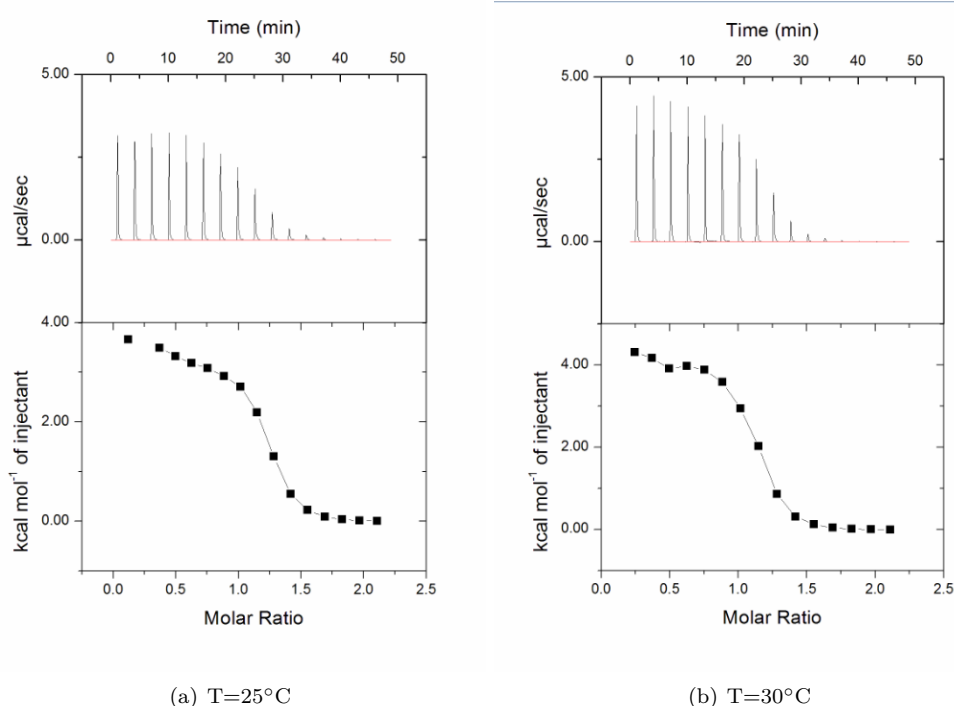


Figure 3.13: Titration Curve Measured by ITC

### 3.4 Thermodynamic Parameters Measuring

The ITC titration curve can be seen in Figure 3.13. In the process of analysis, we neglect the second points because the large deviation must be caused by error rather than internal mechanism. Based on the titration curve, the curve fitting was performed automatically, providing us with specific thermodynamic parameters as shown in Table 3.2. As for the reaction under consideration,  $\Delta H$  has positive value indicating that the reaction absorbs heat from environment. While the increasing equilibrium constant  $K$  with Temperature further confirmed the theory.

Besides, there is a small difference in the accuracy between the two group. The group with high temperature (30°C) has less error and higher accuracy, which is also intuitive. In Figure 3.13, the 25°C group as negative slope at initial period which is not conform to thermal physical theory and must lead to less accuracy of measurement.

Table 3.2: Thermodynamic Parameters Measured by ITC

Temperature	25°C	30°C
N	$1.16 \pm 0.0134$ Sites	$1.07 \pm 0.00904$ Sites
K	$2.14 \times 10^5 \pm 4.81 \times 10^4 \text{ M}^{-1}$	$2.70 \times 10^5 \pm 4.49 \times 10^4 \text{ M}^{-1}$
$\Delta H$	$3431 \pm 67.69 \text{ cal} \cdot \text{mol}^{-1}$	$4160 \pm 57.64 \text{ cal} \cdot \text{mol}^{-1}$
$\Delta S$	$35.9 \text{ cal} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$	$38.6 \text{ cal} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$

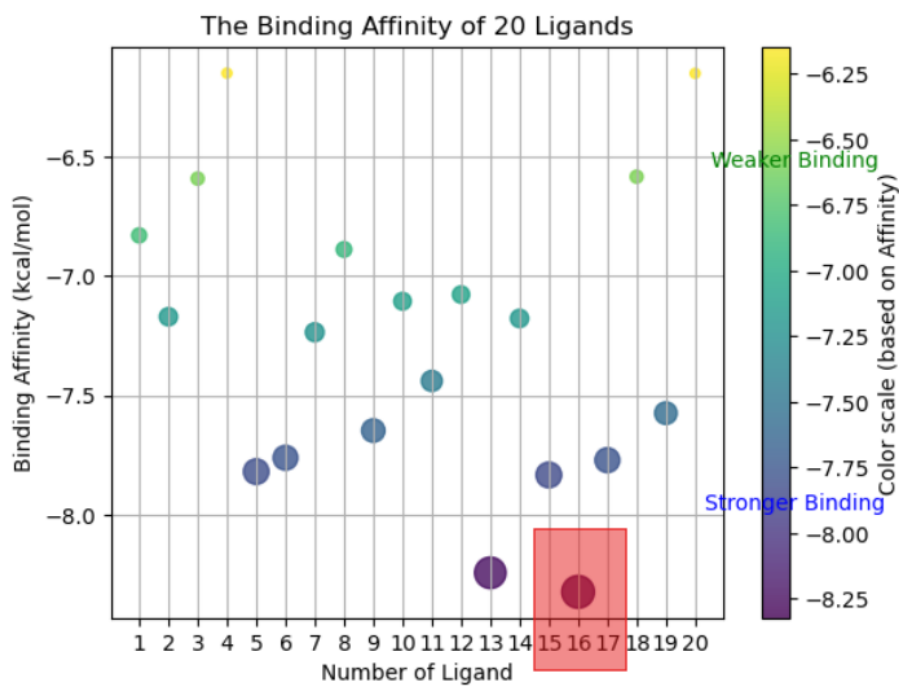


Figure 3.14: Virtual Screening of Ligand

Table 3.3: Binding Affinity of Ligands

Number of Ligand	Affinity (kcal/mol)
1	-6.829
2	-7.170
3	-6.591
4	-6.149
5	-7.819
6	-7.761
7	-7.235
8	-6.888
9	-7.647
10	-7.105
11	-7.439
12	-7.077
13	-8.243
14	-7.177
15	-7.833
16	-8.323
17	-7.771
18	-6.583
19	-7.574
20	-6.150

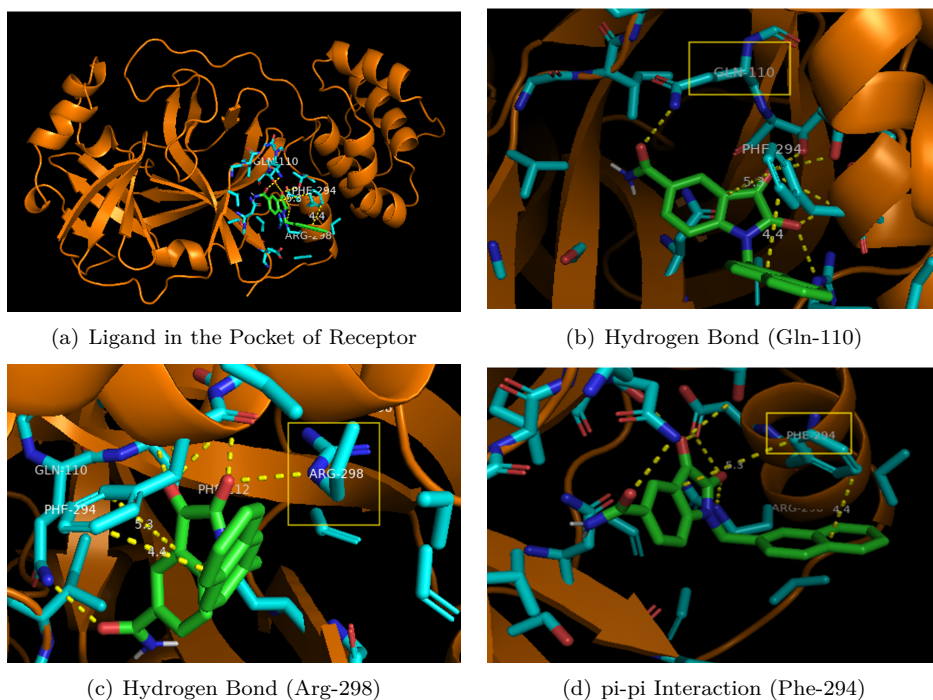


Figure 3.15: Interaction Analysis

## 3.5 Model of Protein Binding

### 3.5.1 Virtual Screening: By Molecular Docking

According to the computing at pkumdl website<sup>[2]</sup>, the center of vina box should be (-5.5, -7.75, 26.25) with the size of (12.0, 15.5, 18.5).

$$\text{center} = (-5.5, -7.75, 26.25)$$

$$\text{size} = (12.0, 15.5, 18.5)$$

The Table 3.3 contains the binding affinity of each ligand while Figure 3.14 provides a visualization for it. From both of them, we can easily pick out the most suitable ligands that is number 16 ligand.

### 3.5.2 Interaction Analysis

The overview of the ligand with receptor can be seen in Figure 3.15. Zooming in, we can observe in detail. There are two major types of interaction between the ligand and receptor. One is hydrogen bond while the other one is pi-pi interaction.

As for the hydrogen bond, the amide group in the residue of glutamine in the 110 position provide hydrogen atom. The H atom in the guanidino group of arginine at position of 298 is also the hydrogen donor for hydrogen bond. The acceptor of hydrogen atoms are the oxygen atoms in ligand.

As for the pi-pi interaction, the approximation of phenylalanine at 294 position to the aromatic rings of the ligand indicates the existence of pi-pi interaction.

## Chapter 4

# Conclusion

I was amazed by the carefully designed experimental classes which walked us through the wonderful process of analyzing protein. From my point of view, the set of experiments follows the logic: from unknown to known; from structure and function to application.

We purified the protein first, because that was the cornerstone for any further analysis involving protein. It would be difficult to make conclusion in the following experiments if there had been impurities in the protein solution.

- Purification: Column Chromatography

As for the analysis of protein, there are several aspects:

- Molecular weight: SDS-PAGE
- Catalytic Ability (Kinetics): ELISA (enzyme-linked immunosorbent assay)
- Thermodynamic and Equilibrium Property: ITC
- Secondary Structure: CD Spectroscopy

Given the study of protein is related to pharmaceutical development, especially for 3CL protease which has been the target of drugs to treat COVID-19. The design and verification of drugs (called inhibitors or ligands alternatively) should be put in the center of the stage:

- Design(screening): Molecular Docking
- Design(rational): Interaction analysis
- Verification:  $IC_{50}$  inhibitory rate detection also by ELISA

From the aspect of learning, we have gotten hands on experiences of several classical theory, models and experiments in the realm of biochemistry and computational chemistry:

- Thermodynamic Parameters Detection
- Enzyme Kinetics: Michaelis Menton Equation
- ...

Hope this experience could lay a solid foundation for future exploration in the world of science.

## Appendix A

# Equipments and Apparatus

- FPLC: Produced by GE HealthCare Life Sciences
- Fast Blue Protein Staining Solution: Produced by Biosharp Life Sciences
- Imaging System: Fusion Fx Vilber Lourmat
- ELISA: Produced by BioTeK
- Nanodrop: Nanodrop2000 produced by Thermo Scientific
- CD Spectrometer: MOS 450 AF/CD (Biologic)
- ITC: MicroCal iTC200 (Malvern)
- Centrifuge: Allegra X-30R Centrifuge, Beckman Coulter

# Bibliography

- [1] Bestsel. <http://bestsel.elte.hu/>.
- [2] Cavityplus. <http://pkumdl.cn:8000/cavityplus/index.php#/computation>.
- [3] Edans. <https://en.wikipedia.org/wiki/EDANS>.
- [4] Homogenization. [https://en.wikipedia.org/wiki/Homogenization\\_\(chemistry\)#cite\\_note-1](https://en.wikipedia.org/wiki/Homogenization_(chemistry)#cite_note-1).
- [5] Quencher. [https://en.wikipedia.org/wiki/Dark\\_quencher](https://en.wikipedia.org/wiki/Dark_quencher).
- [6] Michael M. Cox David L. Nelson. *Principle of Biochemistry*. W. H. FREEMAN AND COMPANY • New York, 6th edition, 2013.