

Contents

1	Motivation	3
2	Theoretical basics	3
2.1	Vibrational modes of molecules	3
2.2	Principle of a FTIR	3
2.3	Structure of proteins	5
3	Experimental setup	6
4	Results	7
5	Discussion	11
6	Conclusion	12

1 Motivation

In the past decades, the identification and characterization of molecules as well as their structure gained more and more relevance. The tools and methods which are developed recently to execute these investigations also improve regarding resolution and speed. However, the most common method, which is used to investigate molecules is spectroscopy [1]. There are many different spectroscopy types, which use different ways to excite the sample and detect the necessary information. Especially infrared spectroscopy is very interesting for the analysis of proteins because of the information that can be pulled out of the infrared light spectrum [2]. Furthermore, even bigger samples than proteins as microorganisms or bacteria can be examined by means of infrared spectroscopy. This makes infrared spectroscopy also relevant for clinical diagnostics regarding pathological examinations and other laboratory analysis. For example the blood serum can be analyzed by using near-infrared spectroscopy or the glucose level of diabetes patients can also be monitored. [3] Another big scope of application in infrared spectroscopy is quality improvement and management, including cosmetics, food, pharmaceuticals and much more. Because of its time-efficiency and wide spectrum, it has several advantages compared to other spectroscopy methods. [4]

2 Theoretical basics

2.1 Vibrational modes of molecules

The basic principle which is used in infrared spectroscopy is the absorption of specific wavelengths of radiation by molecules. Hereby characteristic wavelengths represent different structures in molecules. The use of infrared spectroscopy requires the investigated molecule to be infrared active, which means that there has to be a change of the dipole moment, when the molecule is excited which is visually a vibration of the dipole and by that the whole molecule. By the emitted radiation of the molecule, one can then get information about the structure and composition of the molecule [2]. To find out about the different possibilities of vibrations in a molecule, one has to think about the degrees of freedom. Since in space there exist three directions and each atom can move in every direction, the total number of possible movements is $3 \times N$, where N is the number of atoms. Since 3 of these degrees result in a rotation of the molecule and 3 more just move the molecule around in space, the total number of atom movements which include a change of the dipole moment is $3N-6$. These are called vibrational modes. However, there is an exception for linear molecules. Since a rotation around one axis has no effect at all the number of vibration modes is $3N-5$. These vibrational modes all represent a characteristic movement and can be divided into symmetrical and antisymmetrical vibrations. For bigger atoms, there are more possible vibration modes because of the higher number of atoms, which also results in more peaks in the spectrum. The way these spectrums are measured and analysed will be explained below [2].

2.2 Principle of a FTIR

Nowadays most of the spectrometers that are used are Fourier transform infrared spectrometers (FTIR) as shown in figure 1. The optical part of the spectrometer consists of a Michelson

interferometer. It consists of a monochromatic light source, from which the beam is emitted onto a beam splitter, which is a semipermeable mirror. The two light rays then split up in a right angle, while one is reflected of a fixed mirror and the other is reflected of a movable mirror as shown in the picture. The light ray which moves straight forward, moves through the semipermeable mirror 2 times less, which is the reason for the use of a compensation plate, which assures that there is no phase shift because of the semipermeable mirrors. The distance of the movable mirror is measured by a separate laser to assure that the distance that the light travels is known. The two reflected beams then either constructively or destructively interfere, depending on the position of the movable mirror [2]. Next, the light hits

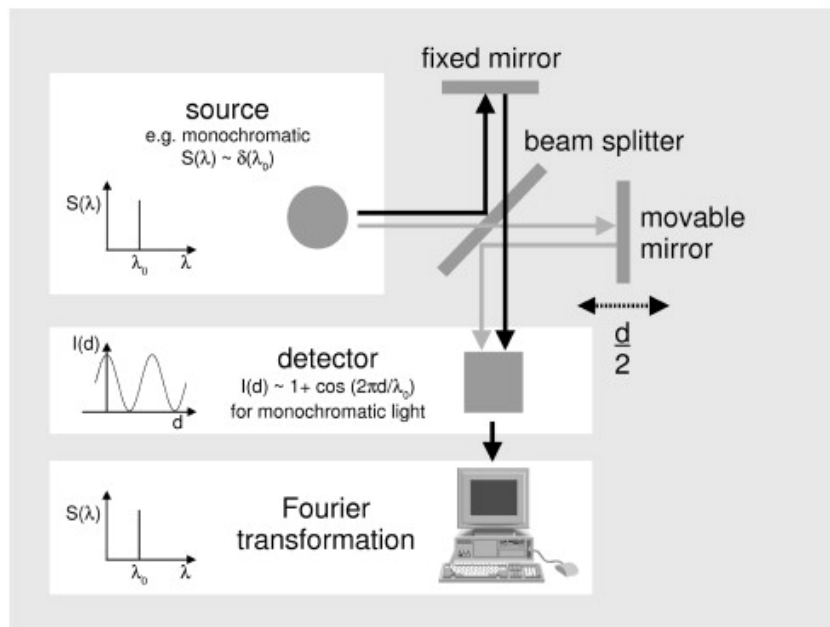


Figure 1: The basic structure of an FTIR. [2]

the sample which is placed in front of the detector. After that the detector measures the interferogram of the sample which is then transformed by Fourier transformation into an intensity spectrum for the one used wavelength. Figure 1 also shows the goal of the fourier transofrmation which is to get a the dependance of the intensity of one specific wavelength. Typically for FTIR this spectrum ranges from wavenumbers around 700cm^{-1} to 4000cm^{-1} . Hereby typical resolutions are $2\text{-}4\text{cm}^{-1}$. Advantages of the FTIR method over conventional dispersive spectroscopy methods, where a slit is used for interference, is that FTIRs use more light which results in higher light intensity and by that also higher resolution. Another advantage resulting of that feature is a faster measurement compared to other methods, which makes FTIR one of the most common methods to investigate structures as proteins [5]. In general when using FTIR there is always a background measurement, to increase the quality of the measurement. By using the same background for many different samples, one can filter the signal which is created by the sample by subtracting the background signal of the experimental setup [3].

2.3 Structure of proteins

Proteins are omnipresent in living species. They consist of amino acid chains, which are folded and connected by peptide bonds. Proteins can be determined by their secondary structure. In the alpha-helix structure the amino acid chains are arranged in a helix while hydrogen bonds form between CO and NH groups and build the so called peptide bonds. Beta sheets correspond to connected hydrogen bonds and are typically around 5 amino acids long[6]. Compared to that beta-sheets show hydrogen bonds between different strands of amino acids. One can differentiate different types of beta-sheet structure, which are flat, bend or rod-shaped. In infrared spectroscopy, especially the CO and NH bonds vibrate and show specific peaks in the spectrum, which then can give conclusion about the protein structure and even the amino acids, the protein consists of. The peptide bonds show a total number of 9 vibrational modes. They are named -A, -B and Amide-I-VII while Amide-A is the absorption band in the spectrum with the highest respective wavenumber and Amide-VII is the one with the lowest wavenumber. Since there is an overlap of peaks in the spectrum of water and -C=O , the samples to be measured are often solved in deuterium prior to the experiment which shifts the peak of the spectrum and by that makes a measurement possible. [1]

3 Experimental setup

The first step before conducting the experiment was the preparation of the samples. In the experiment, three different proteins were used. γ -globulin which consists of antibodies of the class IgG, Papain, which is a protease from the papaya plant and myoglobin which naturally exists in cardiac muscles were investigated via FTIR. Herefore, the three different proteins were given into a D_2O solution. For each protein about 6mg of the protein were dissolved in $20\mu l$. The transmission cell had a size of 20mm and a sample thickness of $25\mu m$ was required which resulted in a total amount of $30\mu l$ sample solution per protein. After preparation the sample cell was put together, and the different measurements were conducted. The sample cell conducted of a metal frame, which had teflon spacers inside, that secured and held the CaF_2 -windows (calcium fluoride) in place. Between these windows, the sample was placed. For the measurements, there was first a background measurements, to have a reference for the sample measurements. The first measurement used air as a background and the empty cell as the reference. The second measurement stage had the empty cell as background and H_2O and D_2O as samples. The last measurement stage used D_2O as background and D_2O and the proteins for sample measurement. The FTIR spectrometer used the OPUS software for evaluation and calculation of the spectra. The measurement range was $1000cm^{-1}$ to $4000cm^{-1}$ and the resolution was $2cm^{-1}$.

Stage	Sample	Background or Sample mode	Scan time
1	air	B	64
	empty cell	S	64
2	empty cell	B	128
	H_2O	S	128
	D_2O	S	128
3	D_2O	B	256
	proteins	S	256

Table 1: Experimental stages and measurements. B stands for background and S for sample mode

4 Results

To determine the thickness of the spacer in between the two cells and by that also the thickness of the protein samples which were used, the measurement of the empty transmission cell was used. With the formula

$$d = \frac{n}{2(v_1 - v_2)} \quad (1)$$

where n is the number of fringes within a spectral region, v_1 and v_2 are the wave number boundaries of this spectral region and d is the thickness of the spacer or the sample thickness later. With $v_1=3907.8 \text{ cm}^{-1}$ $v_2=1582.4 \text{ cm}^{-1}$ and $n=12$ a total thickness of $25.8 \mu\text{m}$ was determined. To understand the usage of deuterium as a solvent instead of water figure 2 shows the peaks of the respective measurements for D_2O and H_2O . As it can be seen in the figure, the relevant wavenumbers in the range of $1600\text{-}1700 \text{ cm}^{-1}$ show a peak in absorption for H_2O and a relative low value for deuterium. This means that if the measurement was conducted with water as a solvent, this would result in a high peak in this area of interest, which is much higher than the once of the proteins, which would make the experiment not effective for this special wavenumber range. Compared to the peak in absorption of 2.8, the highest peaks in the protein measurements still showed values below 1, which underlines the choice of deuterium as a solvent. Because of that, further measurements with the respective proteins of interest were conducted with deuterium as a solvent.

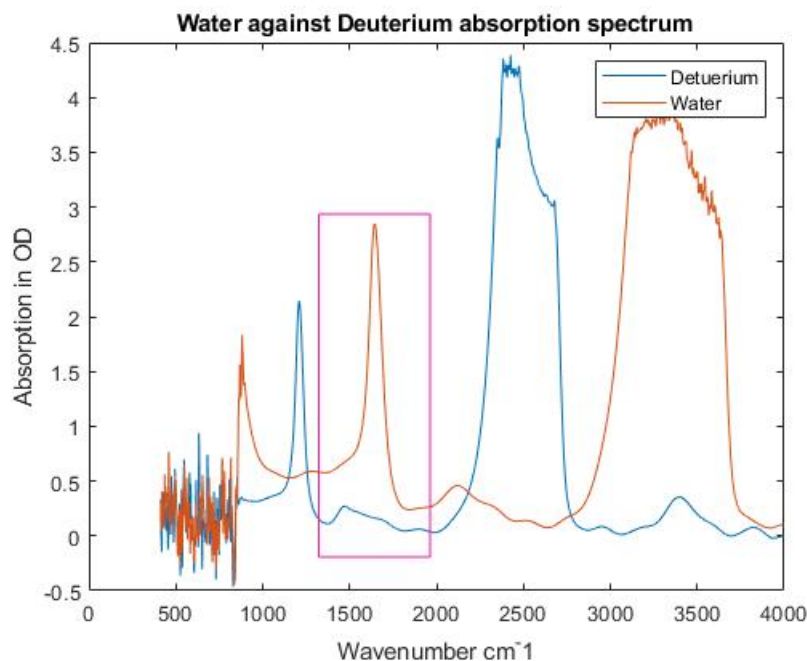


Figure 2: The absorption spectrum of water and deuterium in comparison with the relevant peak marked in pink

To analyze the structure of the investigated proteins, the spectrum of the wavenumbers $1600\text{-}1700 \text{ cm}^{-1}$ were chosen, because the Amid-1 band is located in that range. First, a baseline correction was done to level the data borders to zero and get comparable results.

To get information about the percentage of β -sheets, not ordered, α -helices and loops, a 5 gaussian fit was used, which divides the main peak of the function in the respective range into 5 separate gauss functions, which then correspond to the different mentioned structure. The first fit was done with fixed positions and fixed widths according to the table in the experiment instructions. The figures 3-5 show the resulting gauss fits. It was assured that the values for the gauss fit of the structure which basically looked like formula 2, had values of a which did not exceed the maximum of the measurement values, b was chosen with boundaries according to the given experiment introduction and c should not exceed 50.

$$f(x) = a1 * \exp(-((x - b1)/c1)^2) \quad (2)$$

However, since the gauss fits are more accurate with free widths, because they are then chosen according to the measurement data, these functions show better results and will be used for the following analysis. The figures are shown in figures 6-8. To get information about the influence on the total absorption of each structure and by that also the percentage of each structure in the proteins, the integral of each gauss function was calculated and summed up. Then each integral was divided by the total area under curve, which then resulted in the percentage amount of the structure in the protein. The second β -sheet was neglected. Table 2 shows the percentage of the respective bands, while the 5th peak which also corresponds to beta-sheets is ignored.

	Percentage of the secondary structure			
	Beta-sheets	not ordered	Alpha-helix	loops
Gammaglobulin	29.25	23.29	21.74	25.71
Myoglobin	25.58	28.76	26.98	18.68
Papain	19.23	24.03	29.81	26.92

Table 2: Experimental percentage of the secondary structure for gammaglobulin, myoglobin and papain

Another factor to get knowledge about the quality of the experiment is the position of the peaks of the respective gauss functions, which are shown in table 3 and later compared to theoretical values.

	Peaks of the gauss fit cm^{-1}				
	Beta-sheets	not ordered	Alpha-helix	loops	Beta-sheets
Gammaglobulin	1625	1637	1650	1666	1684
Myoglobin	1629	1643	1652	1666	1682
Papain	1630	1644	1656	1668	1682

Table 3: Peaks of the gaussfit for gammaglobulin, myoglobin and papain with free widths and fixed positions according to figures 6-8

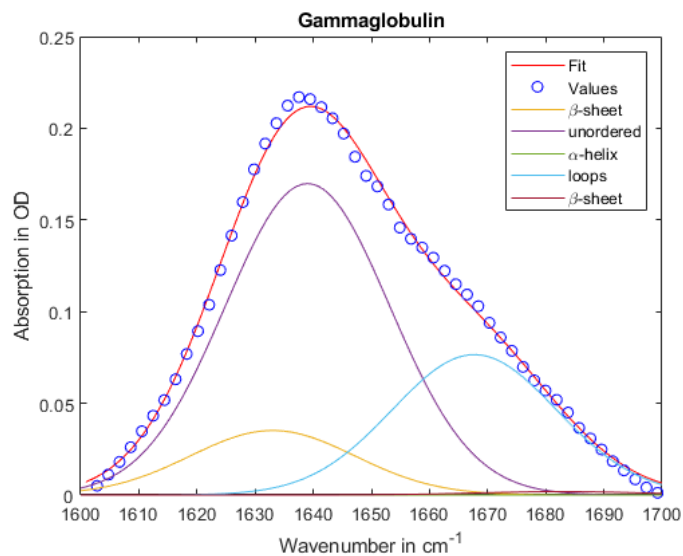


Figure 3: Gammaglobulin gaussian fit with fixed widths and positions

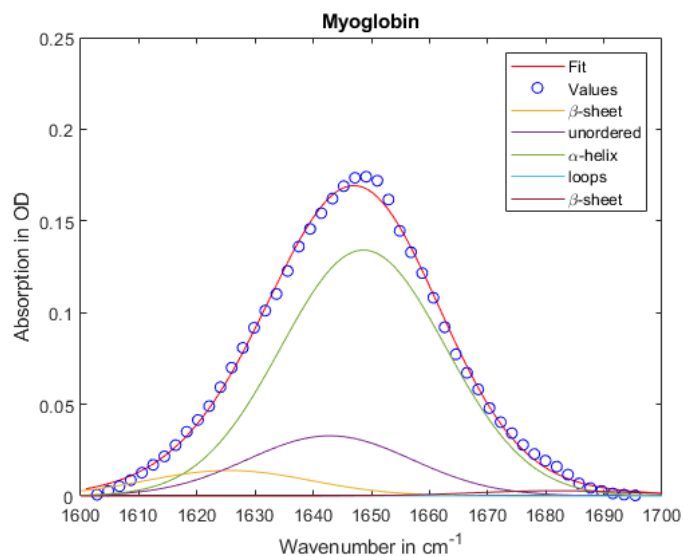


Figure 4: Myoglobin gaussian fit with fixed widths and positions

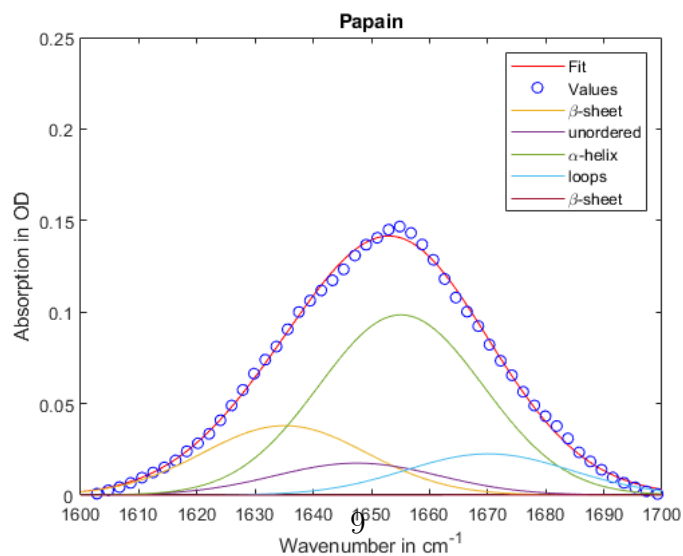


Figure 5: Papain gaussian fit with fixed widths and positions

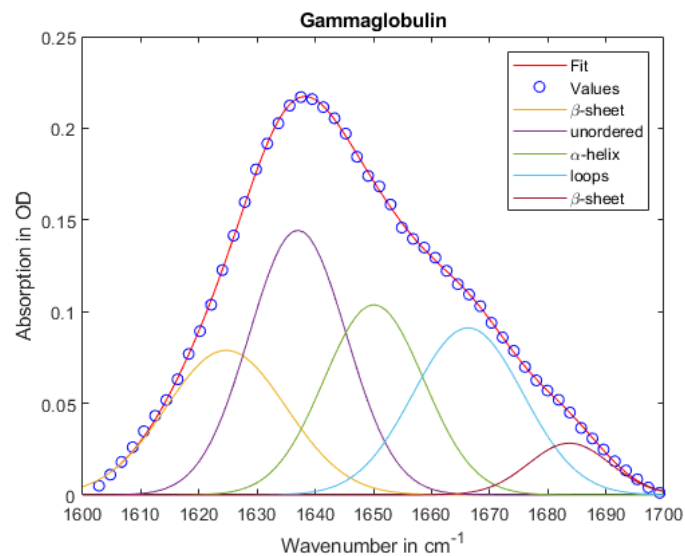


Figure 6: Gammaglobulin gaussian fit with free widths and fixed positions

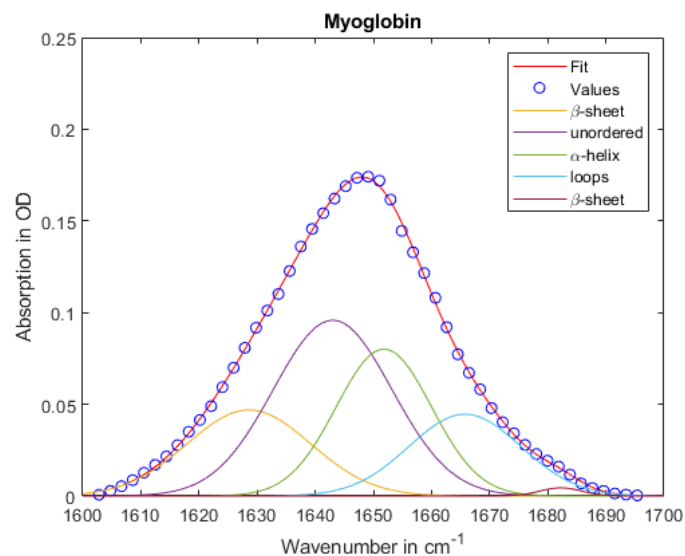


Figure 7: Myoglobin gaussian fit with free widths and fixed positions

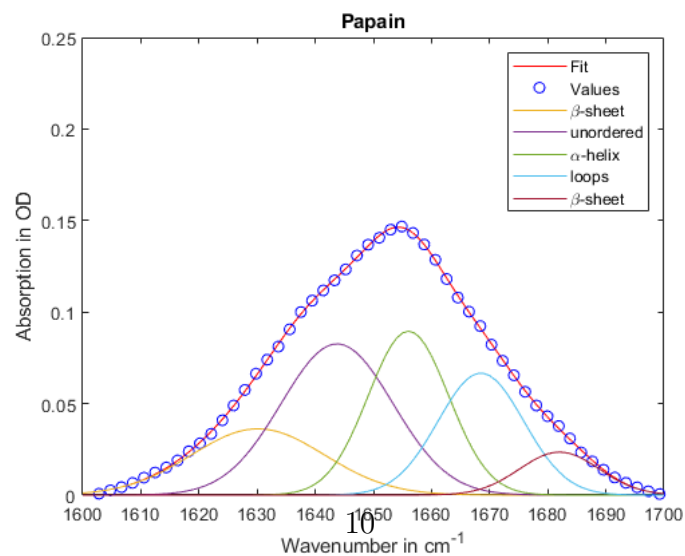


Figure 8: Papain gaussian fit with free widths and fixed positions

5 Discussion

To evaluate the quality of the experimental procedure, this section provides comparison of the measured values to theoretical values. The theoretical values are given in the experimental introduction. Comparing the values of the percentages of the different sheets (tables 2 and 4), it can be seen that only the measurements for papain were in the range of the theoretical values, while the experiment values for gammaglobulin and myoglobin are not comparable with the theoretical values. In theory the percentage of myoglobins β -sheets is 0 while the experiment showed a values of 25.58%. These strongly deviating results could have different sources. First the experiment was not conducted in a clean room, which means that temperature control as well as possible impurities could have affected the result. Furthermore there was no delay before measurement which is usually done to eliminate the water amount in the measurement chamber with nitrogen. Also the FTIR-device which was used, showed some issues and one of the mirrors had some scratches on the surface. Another source of error could be the data analysis. It is unknown how the theoretical data was collected and analysed. This could also lead to deviations from the experimental values due to different analysis techniques. Furthermore the baseline fit which was done could also influence the outcome of the experiment. Last, there was also no quality control of the used protein solutions, which also can lead to different experiment outcomes if the used samples are too old or were damaged in the past.

	Percentage of the secondary structure			
	Beta-sheets	not ordered	Alpha-helix	loops
Gammaglobulin	54	-	7.4	-
Myoglobin	0	7-11	77-88	7-12
Papain	22-29	25-30	26-29	18-23

Table 4: Theoretical percentage ratios of the secondary structure

When comparing the theoretical peaks of the secondary structures as shown in table 5 and the experimental values in table 3, one can see that the peak positions correspond quite well to the theoretical values. There are small deviations of around $pm10cm^{-1}$, which could also be a result of the measuring or analysis inaccuracies.

	Secondary structure band position				
	Beta-sheets	not ordered	Alpha-helix	loops	Beta-sheets
Theoretical peak	1632	1640	1655	1675	1685

Table 5: Theoretical values for the positions of different the secondary structure elements

6 Conclusion

In the conducted experiment, it could be successfully shown, why deuterium is commonly used over water in examinations of protein structures. Furthermore the analysis of the secondary structure of papain showed similar values compared to literature values. The other two proteins myoglobin and gammaglobulin showed great deviations in there secondary structure analysis compared to theoretical values. The experiment also showed the applications, where multiple gauss fits are useful and baseline correction is necessary and gave the student a basic understanding of it. To improve the experiments results it could be conducted in a clean room with quality control of the used samples and devices.

References

- [1] Jose Luis R. Arrondo, Arturo Muga, Jose Castresana, and Felix M. Goñi. Quantitative studies of the structure of proteins in solution by fourier-transform infrared spectroscopy. *Progress in Biophysics and Molecular Biology*, 59(1):23–56, 1993.
- [2] Andreas Barth. Infrared spectroscopy of proteins. *Biochimica et biophysica acta*, 1767(9):1073–1101, 2007.
- [3] L. M. Ng and R. Simmons. Infrared spectroscopy. *Analytical chemistry*, 71(12):343R–350R, 1999.
- [4] Heinz W. Siesler, Yukihiro Ozaki, Satoshi Kawata, and Herbert Michael Heise, editors. *Near-Infrared Spectroscopy: Principles, instruments, applications*. Wiley-VCH and Ciando, Weinheim and München and München, 1. edition, [reprint] edition, 2008.
- [5] Dieter Baurecht. Schwingungs-spektroskopie. *Institut für Biophysikalische Chemie*, 2009.
- [6] Minli Xie and Richard L Schowen. Secondary structure and protein deamidation. *Journal of pharmaceutical sciences*, 88(1):8–13, 1999.