

Protein Vibrational Spectrum Calculations Using Dielectric Coupling in Carbonyls

Jacki Werner
(Dated: October 31, 2010)

The infrared spectrum of a protein depends upon its secondary structure. A simulation was created to calculate absorption spectra of proteins in the Amide I band due to dielectric in carbonyls using candidate protein configurations. Spectra from the simulation can be compared to actual spectroscopic data to help verify proposed protein configurations.

Background

Proteins chains of amino acids. The backbone structure of a protein consists of a repeating pattern of two carbon atoms followed by one nitrogen atom. This three atom group makes up a single amino acid, as seen in Figure 1. The nitrogen is single-bonded to a hydrogen, and one carbon is double-bonded to an oxygen, forming what is known as a carbonyl, seen in Figure 2.

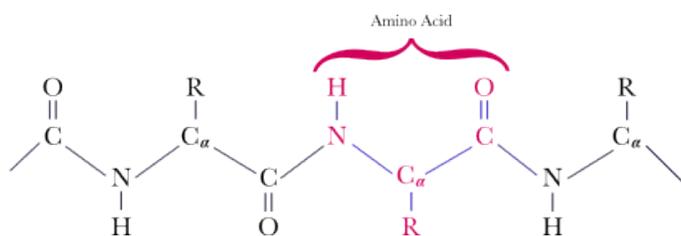


FIG. 1. Protein Backbone with Amino Acid

The carbon designated by C_α (“C-alpha” or “alpha carbon”) is bonded to a side group (indicated by the R in Figure 1) which determines the amino acid’s identity [3].

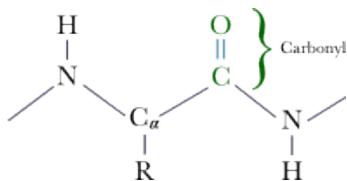


FIG. 2. A Single Carbonyl

A carbonyl vibrates with a natural frequency which can be modeled by a simple harmonic oscillator. Carbonyls are electric dipoles due to uneven sharing of electrons in the covalent bond. The oxygen is more electronegative than the carbon and thus attracts electrons more strongly, and so it gains a net negative charge while the carbon gains a net positive charge (Figure 3).

The carbonyl dipoles are strong enough to cause the carbonyl vibrations to couple across the protein. Thus total vibrations in a protein due to carbonyls can be modeled as a system of coupled harmonic oscillators. Carbonyls have two infrared spectrum absorption bands, known as the Amide I and Amide II bands [4]. In the

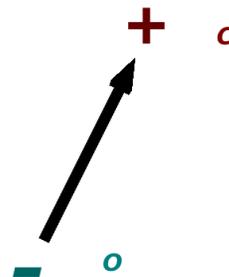


FIG. 3. Carbonyl Dipole Moment

range of wavenumbers which make up the Amide I band (1600 - 1700 cm⁻¹) [4], up to 80% of vibrations in a protein can come from only carbonyls [4]. Thus an approximate infrared spectrum of a molecule in the range of the Amide I band can be theoretically calculated from only the carbonyl positions and orientations.

Method

The coupling strength of two carbonyls depends upon the carbonyl’s dipole orientation and also the distance between the two carbonyls. The distance between carbonyls was taken to be the distance between their centers of mass. In this calculation, the carbonyls modeled as classical simple harmonic oscillators rather than quantum mechanical harmonic oscillators. As coupling quantum mechanical oscillators is significantly more complex, we decided to first create a simpler version of the program to test for viability of this method. Two carbonyls m and n are coupled by the equation

$$d_{mn} = \left(\frac{\kappa^2 e^2}{4\pi\epsilon_0} \right) \frac{\hat{\mu}_m \cdot \hat{\mu}_n - 3(\hat{r}_{mn} \cdot \hat{\mu}_m)(\hat{r}_{mn} \cdot \hat{\mu}_n)}{r_{mn}^3}$$

where μ_m and μ_n are the unit vectors indicating direction of the dipole moments (oriented towards the positively-charged carbon.) r_{cm} is the distance between the centers of mass pointing from carbonyl m to carbonyl n although either direction would give the same results as the two dot products $\hat{r}_{mn} \cdot \hat{\mu}_m$ and $\hat{r}_{mn} \cdot \hat{\mu}_n$ when multiplied by each other would cancel out any sign given to \hat{r}_{cm} . The electric field of an electric dipole is proportional to $\frac{1}{r^3}$,

unlike electric monopoles, which go by $\frac{1}{r^2}$. The coupling strength is determined also by the charges of the dipole and the electric permittivity of the medium the proteins are in. e is the charge of a single electron, and κ is the partial charge constant for the carbonyls. For this simulation, we used $\kappa \approx .4$, using the carbonyl dipole moment calculated by Hiroyuki Aoki and Tsutomu Suzuki [1], by the equation

$$\text{dipole moment} * \frac{\text{Bohr radius}}{\text{average carbonyl length}}$$

. For simplicity, we used $\frac{1}{4\pi\epsilon_0}$, the permittivity of free space, but to get more accurate results the permittivity of the medium holding the proteins could easily be substituted if it is known. We created a matrix to couple each carbonyl with all other carbonyls:

$$\hat{d}_{mn} = \left(\frac{\kappa^2 e^2}{4\pi\epsilon_0} \right) \frac{1}{2} \sum_{m \neq n} \frac{\hat{\mu}_m \cdot \hat{\mu}_n - 3(\hat{r}_{mn} \cdot \hat{\mu}_m)(\hat{r}_{mn} \cdot \hat{\mu}_n)}{r_{mn}^3}$$

The factor of $\frac{1}{2}$ accounts for the fact that each unique pair appears twice in the matrix ($d_{mn} = d_{nm}$).

Thus the coupling matrix appears:

$$\hat{d}_{mn} = \begin{bmatrix} 0 & d_{12} & d_{13} & \cdots & d_{1n} \\ d_{21} & 0 & d_{23} & \cdots & d_{2n} \\ d_{31} & d_{32} & 0 & \cdots & d_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & d_{n3} & \cdots & 0 \end{bmatrix}$$

We also took into account the vibrations of the individual carbonyls.

$$\frac{k}{M} \hat{I} = \begin{bmatrix} \omega_0^2 & 0 & 0 & \cdots & 0 \\ 0 & \omega_0^2 & 0 & \cdots & 0 \\ 0 & 0 & \omega_0^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \omega_0^2 \end{bmatrix}$$

The simulation was set up to take only one value for the frequency for all carbonyls. We would like to possibly improve the simulation by using different natural frequency values for each carbonyl depending on the carbonyl's orientation and location, and also the medium the protein is in, possibly taking these values from the program AMBER. The frequency is calculated from a given vibrational wavenumber by $f = \frac{v}{\lambda}$ where v is the velocity of light in a medium. We used $v = c$ here for simplicity, converting the wavenumber from cm^{-1} into m^{-1} .

Adding the vibrational matrix to the coupling matrix gave us a dynamical matrix for the system, which described all of the protein's carbonyl vibrations.

$$\hat{D} = \left[\frac{k}{M} \hat{I} + \frac{\hat{d}_{mn}}{M} \right]$$

$$\hat{D} = \frac{1}{M} \begin{bmatrix} \omega_0^2 M & d_{12} & d_{13} & \cdots & d_{1n} \\ d_{21} & \omega_0^2 M & d_{23} & \cdots & d_{2n} \\ d_{31} & d_{32} & \omega_0^2 M & \cdots & d_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & d_{n3} & \cdots & \omega_0^2 M \end{bmatrix}$$

Eigenvalues of the dynamical matrix are used to calculate the allowed vibrational frequencies for the system, as each allowed ω^2 is an eigenvalue.

$$\hat{D}\vec{v} = \lambda\vec{v}$$

$$\omega_n = \sqrt{\lambda_n} = k_n c \Rightarrow k_n = \frac{\sqrt{\lambda_n}}{c}$$

Once we had calculated the allowed vibrations, we needed to account for coupling to a light source, the infrared light used for spectroscopy. The absorption intensity of each frequency was calculated as a proportion.

$$I(\omega) \propto \sum_{\alpha=1}^N \left(\frac{N_B(\omega_\alpha) + 1}{\omega_\alpha} \right) \sum_{j=1}^N \sum_{k=1}^N (\hat{\mu}_j \cdot \hat{\mu}_k) (v_{\alpha j}^{-1} v_{\alpha k}^{-1})$$

$N_B(\alpha)$ is the Bose-Einstein distribution value for that frequency.

$$N_B(\omega_\alpha) = \frac{1}{\exp(\hbar\omega_\alpha/k_B T) - 1}$$

v_i^{-1} and v_j^{-1} are vectors from the inverted eigenvector matrix.

While the intensity depends upon the orientations of the dipoles for each carbonyl in relation to all other carbonyls in the protein, we also needed to average over all possible orientations for the molecules. The intensity was calculated for each allowed frequency ω_α .

Protein Forms

We explored the infrared spectra of three protein secondary structures, alpha helices, beta helices, and beta sheets. These were used as a way of checking the program's output, as they each vibrate in different wavenumber ranges and proteins are often have sections in these configurations. An alpha helix (Figure 4) is a tight spiral that is right-handed, meaning that when pointing the thumb of your right hand along the helix axis, it spirals up in the direction your fingers curl [3]. In an alpha helix all carbonyls have the same orientation (Figure 5).

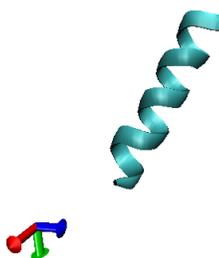


FIG. 4. Protein Backbone in an Alpha Helix

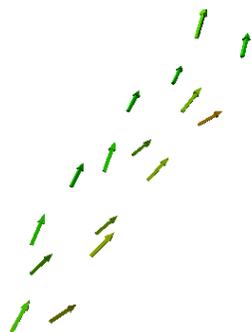


FIG. 5. Carbonyl Orientations in an Alpha Helix

A beta helix (Figure 6) is a triangular spiral, larger than an alpha helix, where half of the carbonyls are anti-parallel to the other half along the same axis (Figure 7).

The carbonyls along the protein's backbone alternate up/down directions. A beta sheet has the same carbonyl orientations as a beta helix, but not the spiral form (Figures 8 and 9). It is formed from long parallel rows of protein backbone, not necessarily from the same protein.

Results

We obtained preliminary spectra for alpha helix, beta helix, beta sheet, and beta hairpin structures using $k = 1674$ as the wavenumber to calculate ω_0 . The intensities were divided by the number of carbonyls in each protein structure for better comparison. The results shown in Figures 10 - 13 include points for the intensity of each allowed wavenumber, as well as a total spectrum obtained by adding a gaussian spread to each point.

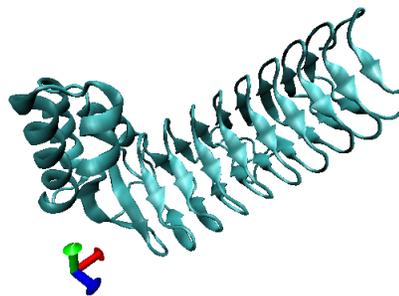


FIG. 6. Protein Backbone in a Beta Helix



FIG. 7. Carbonyl Orientations in a Beta Helix

Future Work

Since so many approximations were made to make the basic version of this program, an important next step is to calibrate the results using the spectra of proteins with well-known structures. Things to change include the permittivity of the medium containing the protein as well as the speed of light in that medium, among other things. Also, using AMBER to obtain initial vibrational frequencies for carbonyls might greatly improve the program's results.

Acknowledgments

I would like to thank my advisor, Dr. Daniel Cox, for his help this summer as well as Dr. Rena Zieve for all of her work making this program possible. Also, I would like to thank the National Science Foundation for funding this research program.

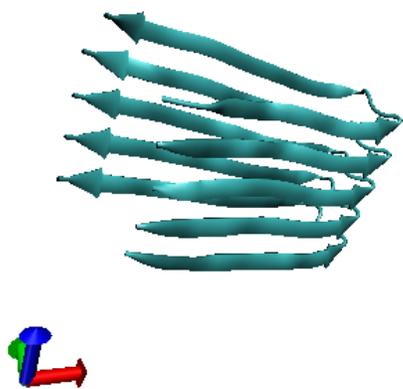


FIG. 8. Protein Backbone in a Beta Sheet

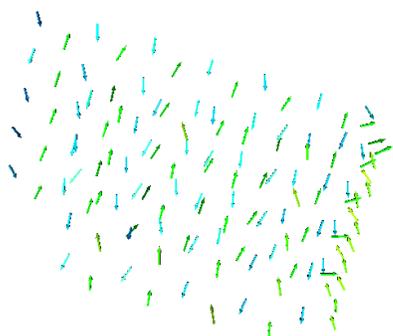


FIG. 9. Carbonyl Orientations in Beta Sheet

Bibliography

- [1] Aoki, Hiroyuki and Suzek, Tsutomu. 1987. *Dielectric Properties of Thermooxidized Poly(Vinyl Formal)* Journal of Polymer Science Part B: Polymer Physics; 25 (2): 369-378.
- [2] Ashcroft, N. W. and Mermin, N. D. 1976. *Solid State Physics*. Saunders College, Philadelphia.
- [3] Nelson, Philip. 2004. *Biological Physics: Energy, Information, Life*. W. H. Freeman and Company, New York: 48, 53.
- [4] Kemp, William. 1975. *Organic Spectroscopy*. John Wiley & Sons, New York: 18-19.

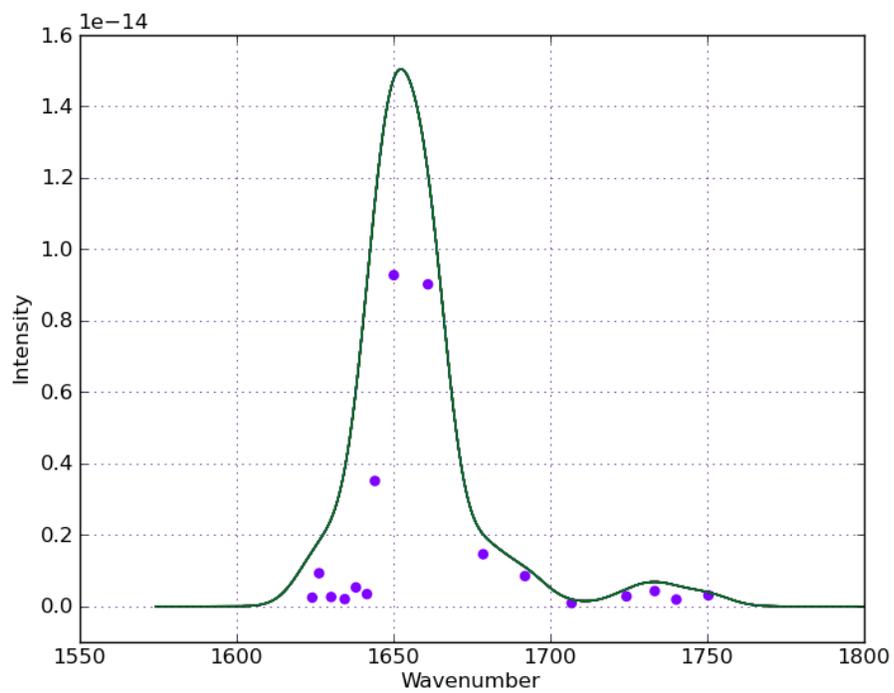


FIG. 10. Alpha Helix Spectrum

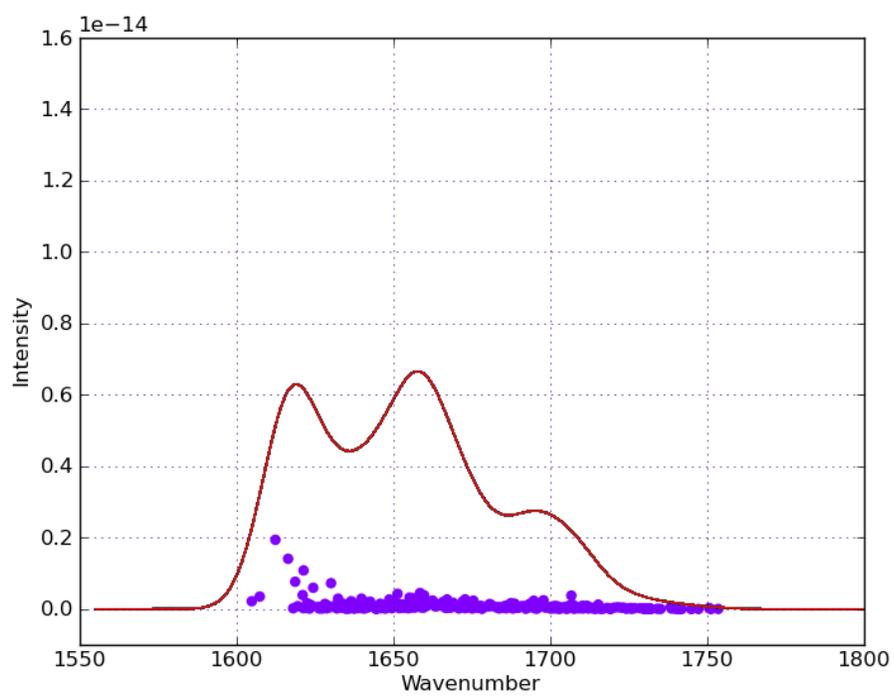


FIG. 11. Beta Helix Spectrum

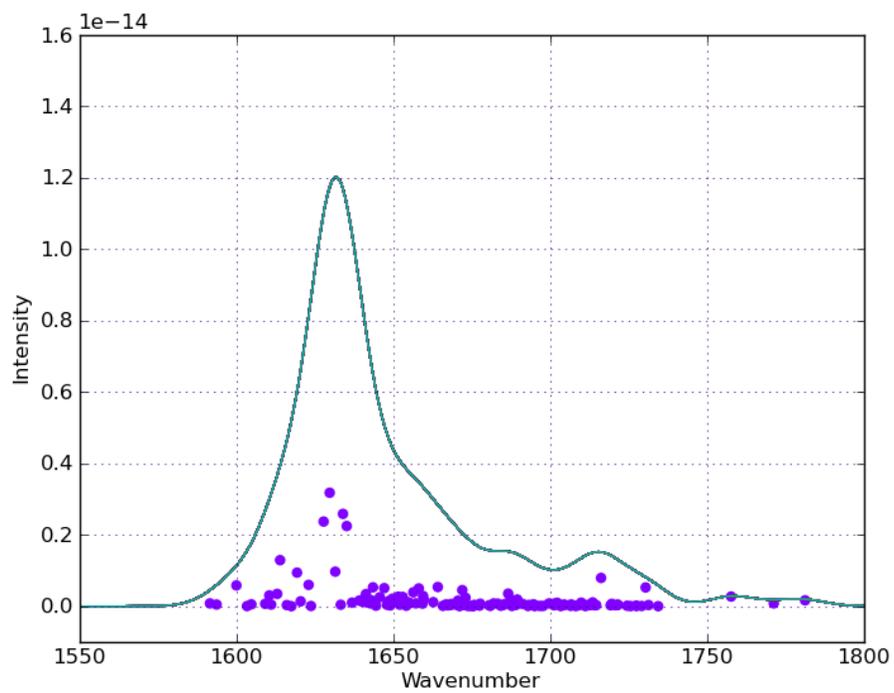


FIG. 12. Beta Sheet Spectrum

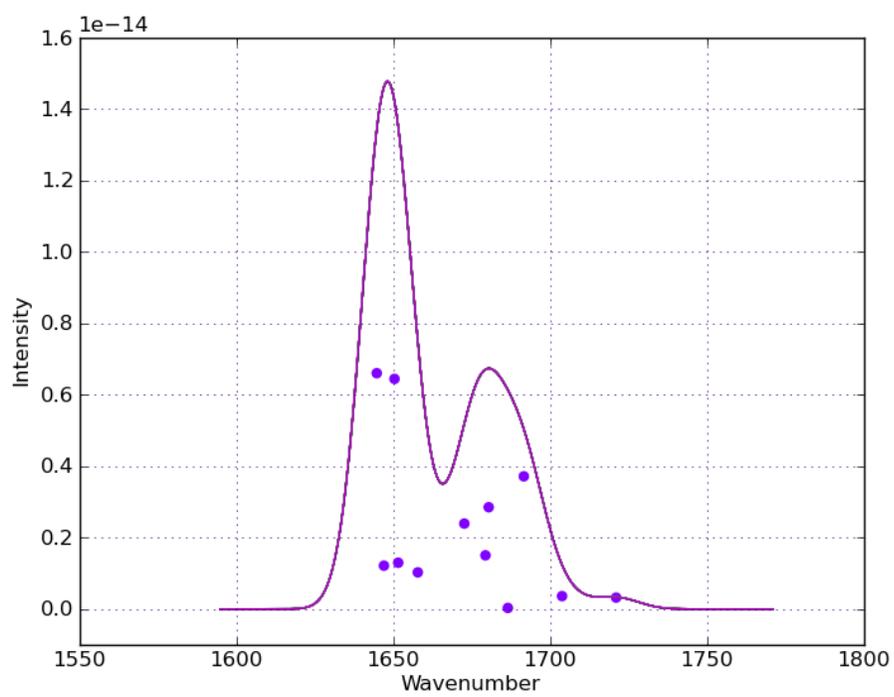


FIG. 13. Beta Hairpin Spectrum