Atomic Interactions as Linear Systems: Hydrogen Bonds in the Alpha Helix of Crambin

Stanley Nicholson[1], Dr. David Minh[2], Dr. Bob Eisenberg[1,3]
IIT Applied Mathematics[1] and Chemistry[2], Rush University Dept. of Physiology[3]

Abstract

We analyze the properties of the stochastic trajectories of atoms in simulations of molecular dynamics (MD) using the methods of stochastic signal analysis, well established and used in the engineering literature. Hydrogen bonds in proteins form critical structural elements such as alpha helices. Using spectral analysis and linear time invariant systems represented as frequency functions (also known as transfer functions) we provide quantitative descriptions of hydrogen bonds along with other interactions such as salt bridges. We present the versatility and applicability of the method by analyzing all pairwise interactions along with possible mechanical models associated with our analysis.

Introduction to Linear Systems

We work in the frequency domain by taking the Fourier transform of our signals. Each frequency point is associated with a sine curve. Take a time series $x(t)$ and $y(t)$ and look at its Fourier transform $X(f)$ and $Y(f)$, respectively. We define $H(f)$ to be a frequency function as [1]:

$$H(f) = Y(f)/X(f)$$

We will demonstrate an example later. The energy a signal has is defined as follows and is equivalent both in time and frequency.

$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 dt = \int_{-\infty}^{\infty} |X(f)|^2 df$$

This motivates the definition of the auto and cross power functions:

$$G_{xx}(f) = X(f)X^*(f) = |X(f)|^2$$
$$G_{xy}(f) = X(f)Y^*(f)$$

Estimation of the powers is much more robust through methods like Welch's and the periodogram than Fourier transform estimations. [2]

$$H(f) = \hat{G}_{xy}(f)/\hat{G}_{xx}(f)$$

We now arrive at the pinnacle of the technique with the coherence function $C_{xy}(f)$ which states there exists a frequency function ***when coherence is equal to 1***. [1]

$$C_{xy}(f) = \frac{|G_{xy}(f)|^2}{G_{xx}(f)G_{yy}(f)}$$

Spring Model

Take a damped spring with a driving input white noise force.

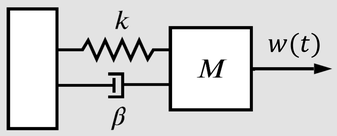


Figure 1: Damped spring system

$$M\ddot{x} + \beta\dot{x} + kx = w(t)$$

Taking the Laplace transform of both sides gives us the formula of the frequency function.

$$|H(j\omega)| = \frac{X(j\omega)}{W(j\omega)} = \frac{1}{\sqrt{(-M\omega^2 + k)^2 + (\beta\omega)^2}}$$

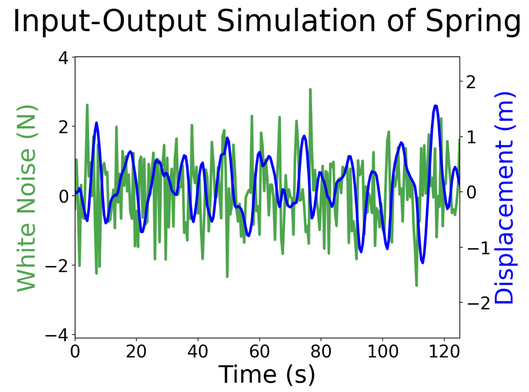


Figure 2: Simulating the damped spring system with a white noise force input (denoted in green) and the system's response (blue).

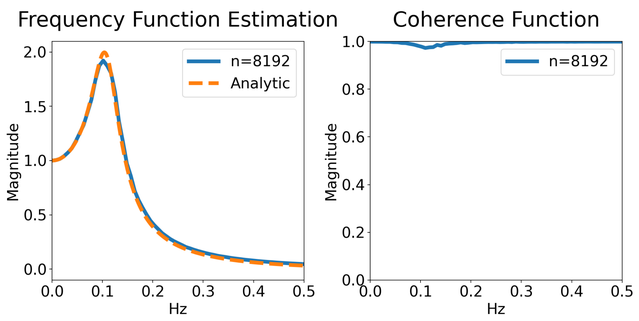


Figure 3: Estimation of the frequency function's magnitude via the power estimation technique outlined in the introduction. We are as confident in our estimation, at a particular frequency, as when the coherence function is close to 1.

The Biology and Results

Molecular dynamics (MD) simulates proteins in solvent solving Newton's equations at femtosecond time scales. Proteins consistent of well defined structures such as alpha helices and beta sheets held together by hydrogen bonding. This motivates us to analyze the hydrogen bond **N**-H…**O** as an input-output relation. We take the 3D trajectory of nitrogen and oxygen, calculate the displacement and define each as an input and output respectively.

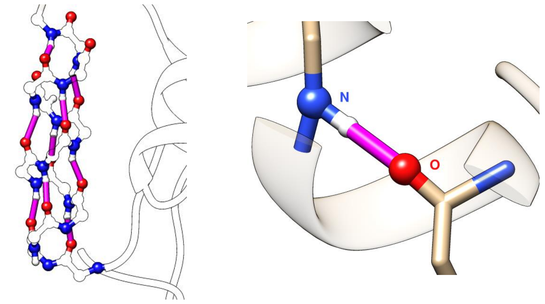


Figure 4a (left): The largest alpha helix of crambin (PDB: 1CRN).
Figure 4b (right): Particular hydrogen bond of crambin.

Estimation of the Alpha Helix

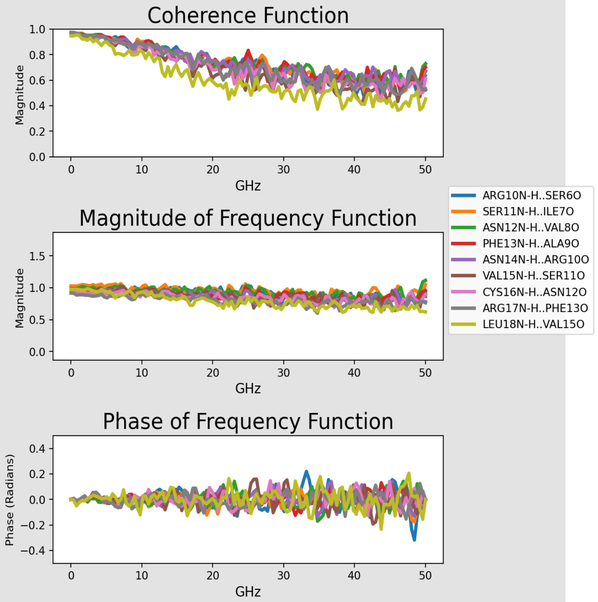
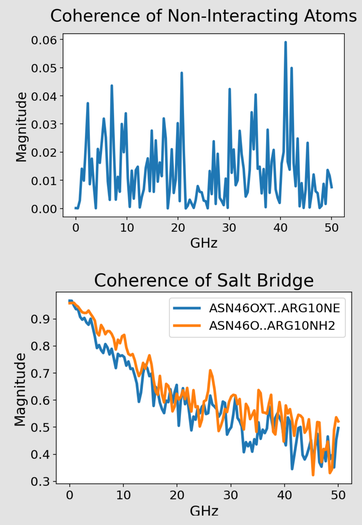


Figure 5: We carry out a nearly identical analysis to the spring model where instead we take our input and output to be the nitrogen and oxygen of the hydrogen bond. We see coherence at low frequencies suggesting a low frequency model of the H-bond.

Other Interactions



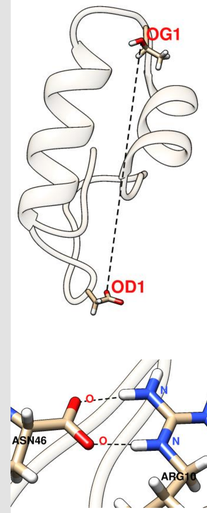


Figure 6a (top): Coherence between two atoms 24 Angstroms apart. Near zero coherence implies no frequency function exists.
Figure 6b (bottom): Coherence between well-studied salt bridge [3] interaction found in crambin. The result resembles **Figure 5** resulting in a similarly constant magnitude and phase response.

Conclusion

The displacement of nitrogen and oxygen is a linear system at low frequencies. This insight justifies the development of a mechanical model with effective spring constants. Similar results are obtained for averaging the turns of the alpha helix and performing the analysis between adjacent turns. Multiscale spring models can aid in the development of coarse grained models and further understanding of conformation changes as possibly linear systems. An extension of the technique as both for data mining and detailed model development are a promising future for the applicability of coherence to MD as a field.

References

[1] Bendat, J. S., & Piersol, A. G. (2011). *Random Data: Analysis and Measurement Procedures*. John Wiley & Sons.
[2] Welch, P. (1967). The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics*, *15*(2), 70–73.
[3] Bang, D., Tereshko, V., Kossiakoff, A. A., & Kent, S. B. H. (2009). Role of a salt bridge in the model protein crambin explored by chemical protein synthesis: X-ray structure of a unique protein analogue, [V15A]crambin-alpha-carboxamide. *Molecular BioSystems*, *5*(7), 750–756.

Acknowledgements

I would like to thank the amazing mentorship and guidance from Dr. Eisenberg and Dr. Minh along his financial support over the summer.