

# Trajectory of Information Entropy During Peptide Folding

Mehmet Sarikaya<sup>1,5,6,7,8</sup>, Siddharth Rath<sup>1,5,6</sup>, Tatum Hennig<sup>1,3,4</sup>, David Corbo<sup>1,2</sup>

## Abstract

Some peptides are known to form stable secondary structures due to their occupation of lower energy states. These folded peptides theoretically have a greater information entropy upon folding, but this has not been experimentally proven. One such peptide, LK14, which reliably folds into an  $\alpha$ -helix at an interface, is used as a case study here to prove that uncertainty in electron energy values increases upon formation of stable secondary structures. Using Python and the SciPy ecosystem we create atom adjacency matrices of each frame of the trajectory and weight these matrices by the atoms' respective counts of valence electrons. Performing an elementwise multiplication of these matrices with normal random Hermitian matrices yields the Hamiltonian. Analyzing the eigenvalues of this matrix shows that the  $\alpha$ -helical form of the peptide grows closer to the Wigner semicircle distribution. This signals that the folded peptide has reached a state of greater information entropy.

## 1. Introduction and Background

It is well established in the literature that LK14 folds into an  $\alpha$ -helix when at an interface. The statistical behavior of the eigenvalues of random matrices is also well established. The electronic consequences of peptides folding, however, is not as well established. With the ultimate goal of

---

<sup>1</sup> Genetically Engineered Materials Science and Engineering Center, University of Washington

<sup>2</sup> College of Engineering, University of Washington

<sup>3</sup> Atmospheric Chemistry, University of Washington

<sup>4</sup> Applied Mathematics, University of Washington

<sup>5</sup> Materials Science and Engineering, University of Washington

<sup>6</sup> Molecular Engineering and Science Institute, University of Washington

<sup>7</sup> Chemical Engineering, University of Washington

<sup>8</sup> Oral Health Sciences, University of Washington

creating a hyperparameter for a physics-informed AI that will have the ability to both predict peptide function from sequence as well as peptide function from structure, a better understanding of these consequences is required. Two direct consequences of the changing electron adjacencies of a peptide over its folding is the Shannon's entropy and the energy of the system. Here we perform a preliminary analysis of how those two properties change before and after a peptide has completely folded.

## 2. Experimental Approach

The primary tools used in this work are the molecular dynamics simulation software, Maestro, by Schrodinger, Python, along with the Python ecosystem SciPy. Our representative models, the extended and fully-alpha frames of LK14 were obtained in Schrodinger, then exported as pdb files. We parsed these pdb files using a Python program which then put the data into an atom distance matrix.

```
for line in pdb_file:
    lineno += 1
    if line.startswith('ATOM'):
        at_obj = PDBAtom(line)
        atoms.append([at_obj.x, at_obj.y, at_obj.z])
        val_atoms.append(at_obj.valence_count)
    elif line.startswith('END'):
        frames.append(atoms)
        atoms = []
        val_frames.append(val_atoms)
        val_atoms = []
pdb_file.close()

base = np.zeros((len(frames), len(frames[0]), 3))
for i in range(len(frames)):
    for j in range(len(frames[i])):
        for k in range(len(frames[i][j])):
            base[i][j][k] = frames[i][j][k]
dists = np.reshape(base, (len(frames), 1, len(frames[0]), 3)) - \
np.reshape(base, (len(frames), len(frames[0]), 1, 3))
dists = dists**2
```

```
dists = dists.sum(3)
dists = np.sqrt(dists)
```

```

r = np.random.normal(size=(elec_count, elec_count))
rt = np.transpose(r)
h = (r + rt) / np.sqrt(2 * elec_count)

```

The pairs of adjacency matrices and random matrices are multiplied elementwise. This yields the Hamiltonian. We take the eigenvalues from this matrix. The distributions of these eigenvalues reveal the Shannon's entropy of the system. In the following figures a trend toward filling the Wigner semicircle distribution is seen as the adjacency values increase. Within each adjacency value the  $\alpha$ -helix eigenvalue distribution is closer to a semicircle than the eigenvalue distribution of the extended structure.

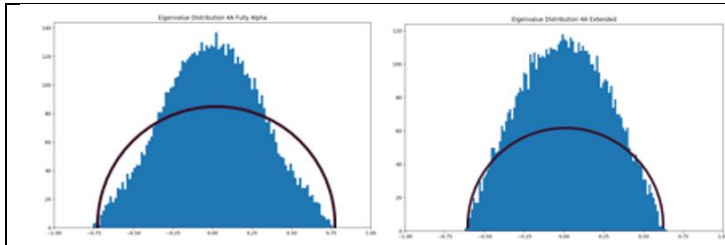


Fig 3 – Eigenvalue distributions of the extended and  $\alpha$ -helical forms of LK14 with adjacency cutoff of 4Å

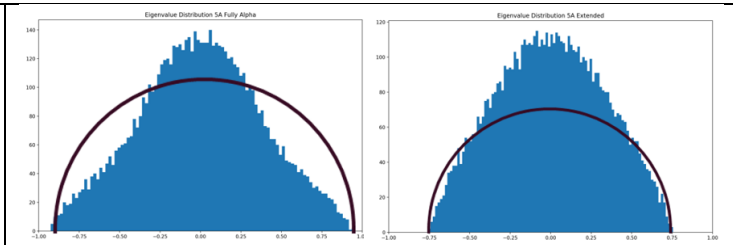


Fig 4 – Eigenvalue distributions of the extended and  $\alpha$ -helical forms of LK14 with adjacency cutoff of 5Å

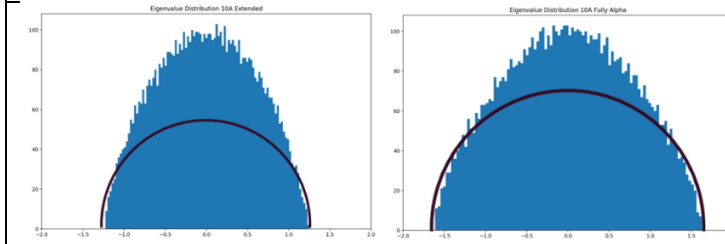


Fig 5 – Eigenvalue distributions of the extended and  $\alpha$ -helical forms of LK14 with adjacency cutoff of 6Å

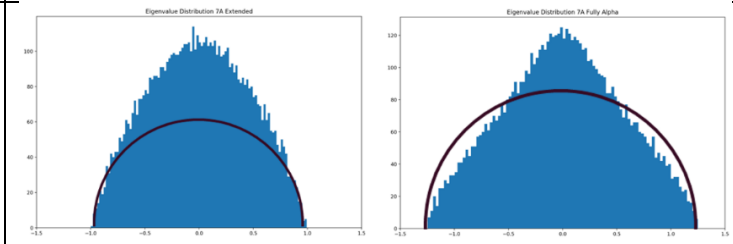


Fig 6 – Eigenvalue distributions of the extended and  $\alpha$ -helical forms of LK14 with adjacency cutoff of 7Å

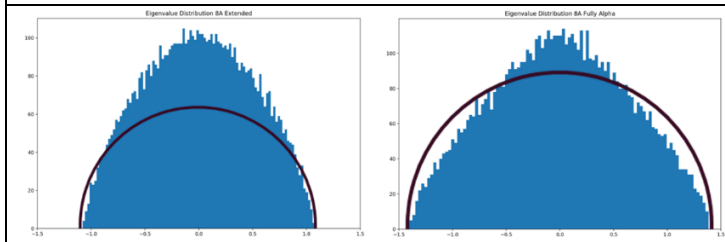


Fig 7 – Eigenvalue distributions of the extended and  $\alpha$ -helical forms of LK14 with adjacency cutoff of 8Å

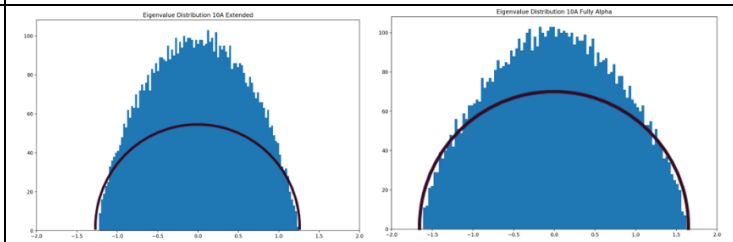


Fig 8 – Eigenvalue distributions of the extended and  $\alpha$ -helical forms of LK14 with adjacency cutoff of 10Å

### 3. Discussion

In addition to all the work presented above, significant work was done towards creating the actual trajectory of entropy and energy over time. We produced quite a few figures depicting these trajectories, but those results were hampered by a couple things. The first was that the molecular dynamics simulation of LK14 was not at an interface, rather it was in solution. Because of this, we did not quite get the fully- $\alpha$ -helical structure we were looking for. Some  $\alpha$ -helical tendencies were observed, and we did see the entropy and energy graphs display changes in the frames with those tendencies. However, we were also using equations for single-state entropy and energy rather than continuous entropy and energy. We have just solved for the continuous entropy equation and will be implementing that soon. We hope to solve for the energy equation soon as well. In addition to the current molecular dynamics data we have which only includes the atoms in the peptide, we plan to include water molecules surrounding the peptide into our analysis in the future. We also plan on exploring the possibility of adding new electronic associations of the peptide into our data, such as electronic resonances within each amino acid in the peptide.

### 4. Conclusions

From the eigenvalue distributions created in this work we can see that the  $\alpha$ -helical forms of LK14 have greater information entropy than the extended forms of LK14. In this work we have set a foundation of code with which we can analyze molecular dynamics data in the future. We have also acquainted ourselves with the mathematics necessary to continue making progress on this project.

### 5. Acknowledgements

This research is supported by the DMREF Program of NSF through the MGI platform under DMR# 1629071, 1848911, and 1922020 at GEMSEC, Genetically Engineered Materials Science and Engineering Center, University of Washington. We thank T. Trogon (UW) for helpful discussions.

## 6. References

Joanna R. Long, Nathan Oyler, Gary P. Drobny, and Patrick S. Stayton, "Assembly of  $\alpha$ -helical Peptide Coatings on Hydrophobic Surfaces," *Journal of the American Chemical Society* **2002** 124(22), 6297-6303 DOI: 10.1021/ja011624n

Bertrand Eynard, Taro Kimura, and Sylvain Ribault, "Random Matrices," **2015** arXiv:1510.04430

Yi-Kai Liu, "Statistical Behavior of the Eigenvalues of Random Matrices," Mathematics Junior Seminar **Spring 2001**, Princeton University