

# MD Simulation of a $\beta$ -Peptide

Guido Putignano<sup>a</sup>, Lorenzo Tarricone<sup>a</sup>

<sup>a</sup>*Department of Biosystems Science and Engineering, ETH Zürich, Basel, Switzerland*

---

## Abstract

The exploration of peptides' dynamic behaviour in solution plays a pivotal role in deciphering the intricate phenomena of protein folding and conformational transitions. This study employs molecular simulations to investigate the equilibrium between folding and unfolding in a peptide that, based on experimental evidence, can adopt a hairpin (folded) or a more disordered (unfolded) conformation in a specific solvent. The primary focus centres on unravelling the intricate interplay between ionic strength and temperature in shaping the peptide's conformational landscape. Utilizing molecular dynamics simulations, we visualize the spatiotemporal dynamics of the peptide and employ three key analyses. Firstly, we calculate the atomic positional root-mean-square deviation (RMSD) from a canonical hairpin structure. Secondly, we determine the end-to-end distance of the peptide, providing insights into its structural transitions. Finally, we compute the electrostatic interaction energies between solute-solute and solute-solvent, shedding light on the impact of solvent environment and ionic presence. In our investigation, we observe oscillatory behaviour at elevated temperatures, indicating the dynamic nature of the peptide. Furthermore, in the presence of a single chloride ion  $Cl^-$  or under high-temperature conditions, the peptide exhibits heightened instability. A noteworthy finding is the observation that Solute-Solvent and Solvent-Solvent electrostatic energies display an intriguing correlation pattern. In the absence of ions, these energies exhibit an anti-correlation, while in the presence of ions, no substantial correlation is discernible. Our study aims to provide comprehensive insights into the behaviour and properties of protein folding and conformational transitions through computational analysis. This research contributes to the advancement of our understanding of molecular dynamics and has implications for the broader field of protein science.

*Keywords:*  $\beta$ -Peptide, Classical Simulation, Biomolecular Systems, Electrostatic Interactions.

---

## 1. Conformational Behavior

The first activity of our study was to describe the conformation of the peptide. The conformation in those cases remains stable, and there are not many noticeable differences between the three cases.

### 1.1. Configuration A (300K, No Ions)

In this case, the peptide exists at a relatively low temperature of 300 Kelvin, and there are no counter-ions present. At this temperature, the peptide is likely to adopt a stable, folded conformation. The movements of the peptide molecule in space are somewhat restricted due to the lower thermal energy. However, it still exhibits some thermal fluctuations, causing it to vibrate and rotate to some extent. The absence of counter-ions means there are no additional electrostatic interactions to influence the peptide's conformation. Overall, the peptide's conformation can be described as relatively stable, with limited spatial movement.

### 1.2. Configuration C (380K, No Ions)

In this case, the temperature is significantly higher at 380 Kelvin, and there are no counter-ions present. The increased temperature imparts more thermal energy to the peptide, leading to greater vibrational and rotational movements. The peptide is more dynamic and can explore a wider range of conformations compared to Configuration A. It may sample multiple conformations and experience greater flexibility due to the increased thermal energy. However, the absence of counter-ions still means that the conformational behaviour is primarily influenced by temperature.

### 1.3. Configuration E (340K, With an Ion)

In Configuration E, the temperature is intermediate at 340 Kelvin, and there is an ion present. The peptide's behaviour is influenced by both temperature and the presence of the counter-ion. The ion introduces electrostatic interactions, which can affect the peptide's conformation. The ion's charge can lead to attractive or repulsive forces between the peptide and the ion, influencing the peptide's spatial orientation. This can result in more pronounced rotational movements or changes in the peptide's overall conformation compared to Configuration A. The temperature, although not as high as in Configuration C, still contributes to thermal fluctuations, allowing the peptide to explore different conformations.

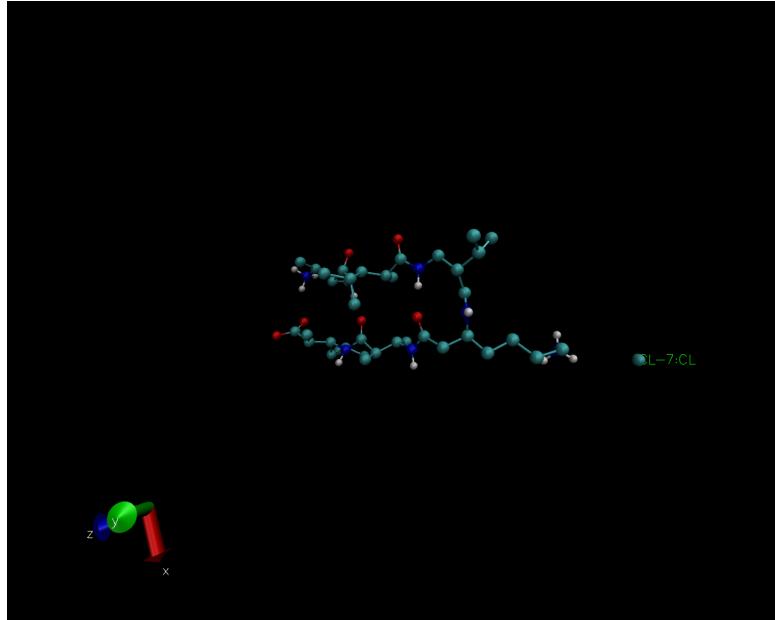


Figure 1: Picture of the pepetide with the ion.

## 2. Boiling point and conformational changes

Determining the state of methanol is a critical aspect of our investigation. It is accomplished by closely monitoring the variations in box volume. In the event of complete methanol vaporization, a substantial increase in the box volume is anticipated. Conversely, when methanol remains in its liquid state, the box volume should demonstrate a relatively consistent behavior.

A valuable approach to discerning methanol's state is to compare the mean box volume at different temperatures, as demonstrated in the accompanying illustration. It's important to note that the box dimensions are similar (calculated volume in the following sections), as computed in subsequent sections. Nevertheless, it is plausible that, at the conclusion of our simulation, the structure subjected to the higher temperature may display a greater degree of expansion in comparison to the one subjected to the lower temperature. This variation in expansion can furnish valuable insights into the state of methanol within the confines of our simulation conditions.

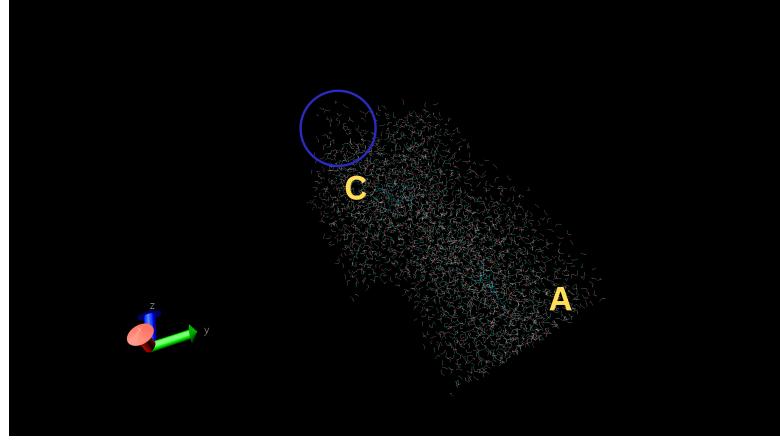


Figure 2: Representation of condition C and A, describing their orientation in space at the final time point.

### 3. Root Mean Square Deviation (RMSD)

The root-mean-square deviation (RMSD) serves as a quantitative metric for assessing structural disparities between two distinct configurations. It provides a valuable means to systematically investigate how the conformation of a peptide may undergo variations under different conditions.

#### 3.1. Display Results

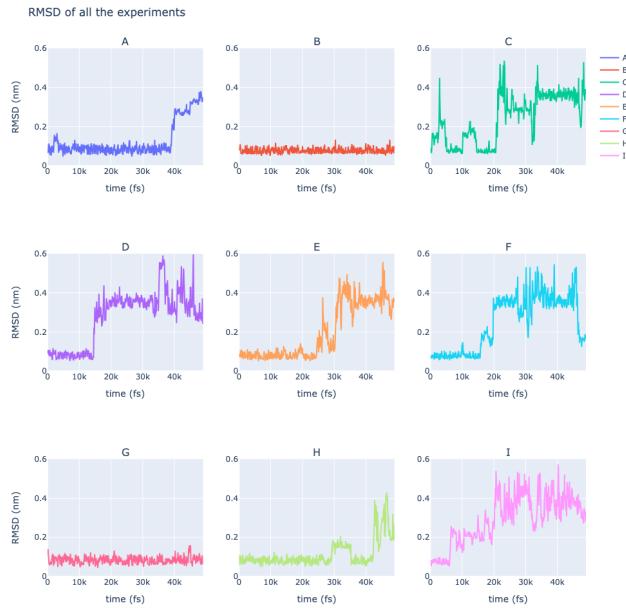


Figure 3: Root Mean Square Deviation of all the experiment

#### 3.2. Summary table

Here is possible to see a summary table depending on the previous plot

ions	300K	340K	380K
none	A(*)	B(o)	C(+)
1 Cl <sup>-</sup>	D(*)	E(*)	F(+)
11Cl <sup>-</sup> + 10Na <sup>+</sup>	G(o)	H(*)	I(*)

Table 1: Table reassuming the trend of the RMSE value in different experimental conditions

### 3.3. Description of the plots

When analysing the plots, it was possible to visualise three different behaviours. As a hypothesis and in order to make our analysis more rigorous, we considered that "the peptide stays close to the hairpin structure throughout the simulation" corresponds to oscillations of the RMSE value being under 0.1nm, "the peptide drifts away from the hairpin structure and never comes back" if the value of the RMSE went steadily over 0.4nm and "the structure oscillates between hairpin and non hairpin" between The value of the RMSE was oscillating inside this interval. Note that these values are arbitrary and have been chosen to facilitate the description of the plots. The prediction of minor oscillations at a stable equilibrium state was anticipated, as it would have been impractical to expect their absence, given the influence of thermal fluctuations.

According to the data plots generated, a consistent observation is the emergence of oscillatory behavior, particularly evident at elevated temperatures (C and F). We attribute this phenomenon to the heightened energy of the molecule, enabling it to periodically explore both conformations.

Moreover, an analysis of the data plots indicates that in the presence of a single chloride ion  $Cl^-$  or under high-temperature conditions, the peptide exhibits increased instability. This observation could be attributed to the absence of ions causing the two charged ends of the molecule to attract each other, with this interaction becoming even more pronounced in the presence of two ions, effectively "locking" the molecule in its hairpin structure. Conversely, in the presence of a negative ion (as clearly demonstrated above), the positive end of the molecule forms a bond with the ion, causing a shift in the electron cloud away from the positive end.

## 4. End to end distance

### 4.1. Displaying Results

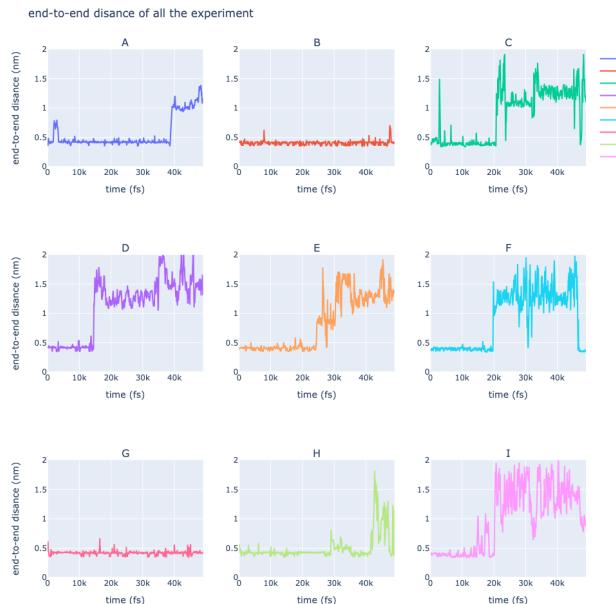


Figure 4: End to end distance of the peptide depending of the experiments

#### 4.2. Discussion

Upon scrutinizing the end-to-end distance data, it becomes evident that the observed trends closely align with those highlighted by the Root Mean Square Error (RMSE) metric. Notably, the end-to-end distance measurement exhibits a degree of noise, rendering the trends somewhat less distinct, especially when it comes to identifying oscillatory behavior. It is plausible that the unique characteristics of the molecule and the experimental conditions contribute to this variance in data quality.

This concordance between the two measurement approaches is likely attributed to the fact that the hairpin conformation represents the configuration that minimizes the molecule's end-to-end distance. Consequently, any deviation from this hairpin structure, as captured by the RMSE, correlates with a proportional increase in the end-to-end distance.

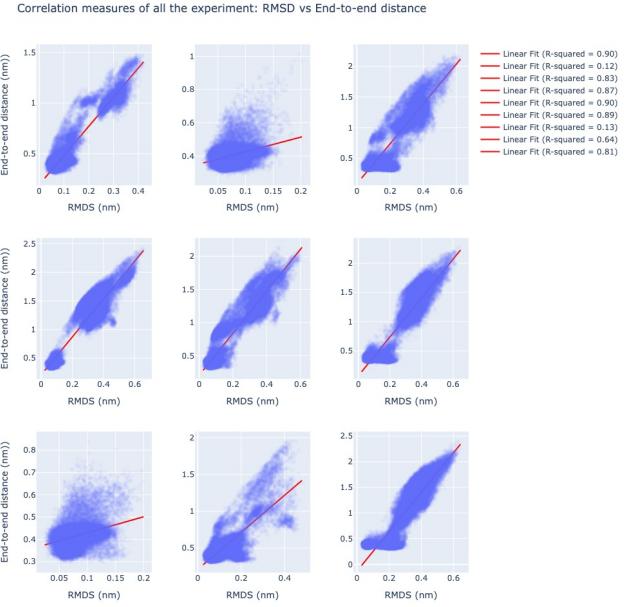


Figure 5: Correlation between RMSD and End-to-End distance

## 5. Energy

### 5.1. Display Results

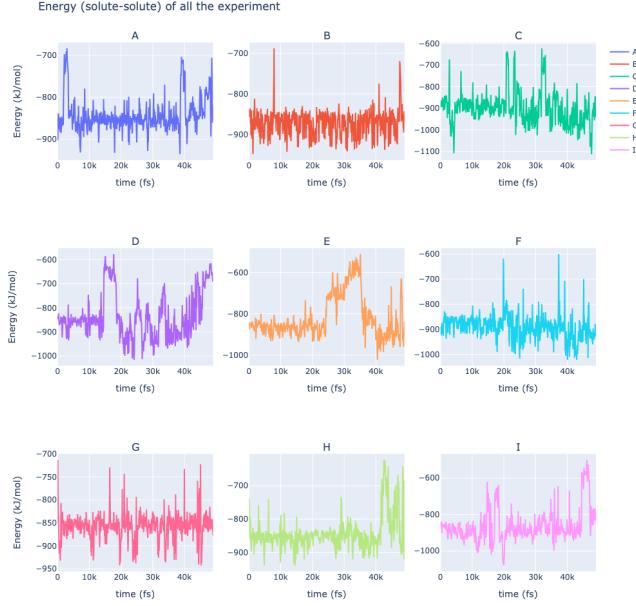


Figure 6: Solute Solute electrostatic energy for all the experiments

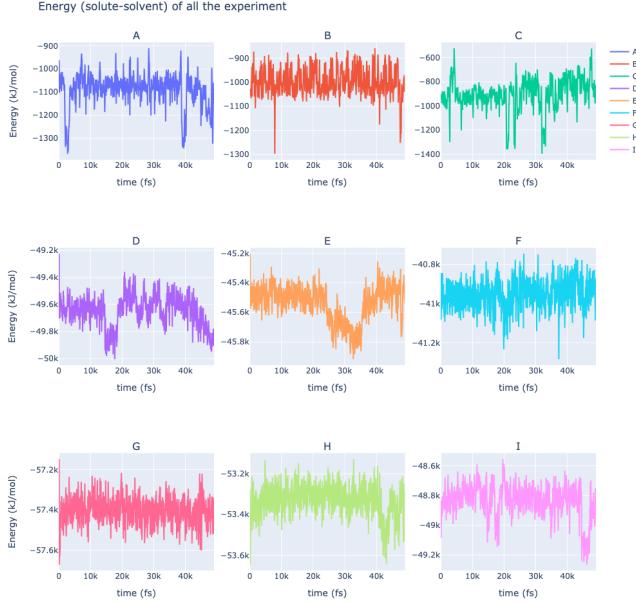


Figure 7: Solute Solvent electrostatic energy for all the experiments

### 5.2. Discussion

When examining the Solute-Solvent and Solvent-Solvent electrostatic energies, it is apparent that defining a specific trend can be challenging. To gain a more comprehensive understanding of these values, it is beneficial to

investigate their correlation. The correlation analysis reveals a noteworthy observation: in the absence of ions, these values exhibit an anti-correlation, whereas in the presence of ions, no significant correlation is discernible. This is further substantiated by the R-squared values of the corresponding plots.

Upon comparing the trends in energy levels with the ones observed for RMSE and end-to-end distance, as previously elucidated, it becomes evident that the latter two parameters tend to correlate with the Solute-Solvent electrostatic energy. This correlation sheds light on how the presence of bonds and the consequent changes in electrostatic energy may impact the conformational behavior of the molecule. Specifically, an increase in the number of bonds between the solvent and the molecule tends to result in a more open conformation, as it exposes a greater surface area to the surrounding solvent molecules.

Furthermore, at higher temperatures, a consistent trend emerges, with both types of electrostatic energies increasing. This is apparent when examining the y-axis values, where for experimental conditions C, F, and I, these values consistently surpass those for other cases. This phenomenon is likely due to the greater conversion of electrostatic energy into kinetic energy at higher temperatures, assuming the conservation of energy within our closed system.

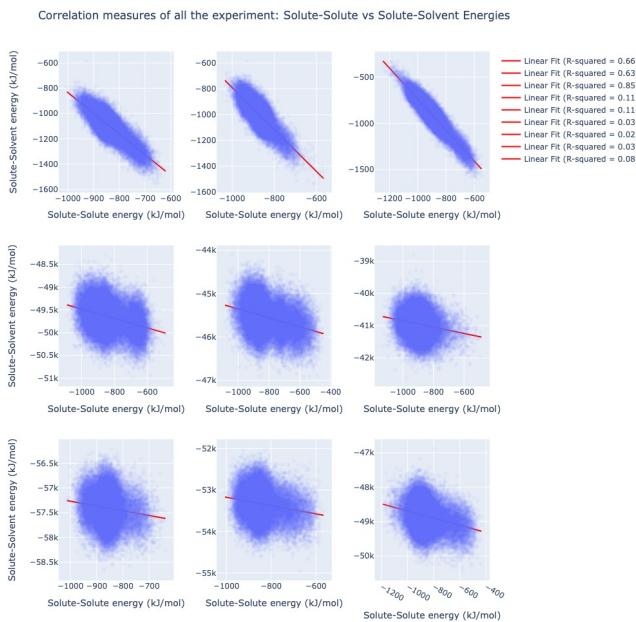


Figure 8: Correlation between Solute-Solute & Solute-Solvent Energies

## 6. Main remarks

### 6.1. Folded & unfolded conformations

In order to characterise the folding-unfolding equilibrium of the peptide there is either the possibility to monitor the relative populations of folded and unfolded conformers, or to start with an initial configuration. Overall, there are several advantages and disadvantages for both methods.

Advantages and Shortcomings of the Perturbation Approach	
Advantages & Disadvantages	Explanation
Faster Timescale (+)	One significant advantage of the perturbation approach is its ability to provide insights into how a folded conformation responds to changes in environmental conditions on a relatively short timescale, such as 50 ns. This allows researchers to study dynamic processes and responses to external factors more efficiently.

Energetic Landscape (+)	By subjecting the folded conformation to different conditions, you can gather information about the energetic stability of the conformation and how it responds to changes in temperature, pressure, solvent, or other factors. This can help identify the conditions under which the folded state is maintained and when it may unfold.
Mechanistic Insights (+)	This approach can provide mechanistic insights into the stability of the folded conformation. For example, you can observe whether specific interactions, like hydrogen bonds or hydrophobic interactions, are disrupted by changes in the environment, providing details about the forces that stabilize the folded state.
Lack of Equilibrium Information (-)	The perturbation approach does not provide direct information about the equilibrium constant, which is crucial for quantifying the thermodynamic stability of different conformations. It cannot give you the exact population ratio of folded and unfolded states under specific conditions.
Equilibrium at Initial Configuration (-)	When studying our molecule, our initial condition is with a specific conformation or structure. However, we don't know how the system would evolve starting from a different conformation or structure.
Timescale Limitation (-)	The perturbation approach may not capture slow kinetic processes or rare events that occur on timescales longer than the simulation duration (e.g., folding or unfolding transitions that take place on a much longer timescale than 50 ns). Equilibrium sampling over a more extended period is necessary to fully understand the folding-unfolding equilibrium.
Limited Sampling of Phase Space (-)	The perturbation approach starts from a specific conformation and explores a limited region of phase space around that conformation. It may miss the exploration of alternative conformations or pathways that are relevant to the folding-unfolding process.

Overall, as seen during the lecture, the relaxation time for a peptide goes from 10 ns to 100 ns. that value increases even more when considering a protein, reaching a value of up 1 s. In our case, our timescale was 50 ns. So starting from an initial consideration was ideal.

System type	Relaxation time (indicative)
gas	~ 1 ps
pure liquid	~ 10 – 100 ps
small organic molecule in solution	~ 10 ps – 1 ns
short peptide in solution	~ 10 – 100 ns
lipid aggregation in solution	~ 10 – 100 ns
protein in solvent	~ 1 ms – 1 s

Figure 9: Slide describing the total relaxation time needed depending on the considered system

When contemplating the use of reverse simulations, it is essential to recognize their lack of realism. In an actual system, atom velocities exhibit a random distribution, and there is no assurance of spontaneous system refolding. In reality, the probability of refolding from an unfolded state depends on numerous factors, such as peptide sequence, temperature, and pH. For certain peptides, the likelihood of refolding is exceedingly low, resulting in an extended period of an unfolded state (which may not be observable within a mere 50 ns simulation). Furthermore, the system cannot spontaneously refold without an external energy source, and reversing velocities fails to account for temperature variations and energy dissipation that accompany the folding and unfolding processes. Therefore, it is an erroneous assumption to equate the timescale of protein folding with that of its unfolding.

## 6.2. Molar concentration

In order to calculate the molar concentration, it was important to identify the right dimensions of our system.

Depending on our system, the volume varies. Temperature is the most influential factor in this regard. Therefore, we can consider systems at the same temperature to have the same volume

Given our simulation file (for T = 380K), we found these values:

```
CRYST1 50.268 50.268 50.268 90.00 90.00 90.00 P1
```

In this case:

- **CRYST1:** This is a label indicating that the line contains information about the crystal unit cell.
- **50.268 50.268 50.268:** These three numbers represent the lengths of the edges of the unit cell in Angstroms (Å). In this case, all three cell edges have a length of 50.268 Å. This information defines the size of the crystal unit cell.
- **90.00 90.00 90.00:** These three numbers represent the angles between the edges of the unit cell. Each angle is measured in degrees. In this case, all three angles are 90.00 degrees. This information defines the angles between the edges of the unit cell and indicates whether the crystal structure is cubic or orthorhombic.
- **P 1:** This part refers to the space group symmetry of the crystal. In this case, "P 1" indicates that the crystal belongs to the simplest possible space group, which is the P1 space group. This means there is no additional symmetry beyond the translational symmetry defined by the unit cell dimensions.

Given:

Length = 50.268 Å

Width = 50.268 Å

Height = 50.268 Å

Number of peptide molecules= 1

Avogadro's number  $\approx 6.022 \times 10^{23}$

We can make our calculation:

1. Calculate the volume of the simulation box in Å<sup>3</sup>:

$$\text{Volume } (\text{\AA}^3) = \text{Length } (\text{\AA}) \times \text{Width } (\text{\AA}) \times \text{Height } (\text{\AA})$$

2. Convert the volume from Å<sup>3</sup> to liters:

$$\text{Volume (liters)} = \text{Volume } (\text{\AA}^3) \times (1.0 \times 10^{-27} \text{ L}/\text{\AA}^3)$$

3. Determine the number of moles of the peptide in the simulation box:

$$\text{Number of moles} = \frac{1}{\text{Avogadro's number}}$$

4. Calculate the effective molar concentration of the peptide in the solution:

$$\text{Effective Molar Concentration} = \frac{\text{Number of moles}}{\text{Volume (liters)}} = 0.01307 \text{ mol/L} \approx 13.07 \text{ mmol/L}$$

For T = 300 K

```
CRYST1 47.858 47.858 47.858 90.00 90.00 90.00 P 1 1
```

For T = 340 K

```
CRYST1 48.927 48.927 48.927 90.00 90.00 90.00 P 1 1
```

We can obtain a table for all the examples

ions	300K	340K	380K
none	A(15)	B(14)	C(13)
1 Cl <sup>-</sup>	D(15)	E(14)	F(13)
11Cl <sup>-</sup> + 10Na <sup>+</sup>	G(15)	H(14)	I(13)

Table 3: Table describing the approximative value of Molar Concentration [mM]

**Effective Molar Concentration of the Peptide:** The effective molar concentration obtained from the simulation is based on the number of peptides in the simulation box and the box's volume. It represents a local concentration within the simulation, but it doesn't necessarily reflect the concentration in a bulk solution. In an experimental setup, you would typically have a much larger volume of solvent, and the peptide concentration would be uniform throughout the entire solution. The simulation does not typically include all of the components of a real experimental sample. For example, the simulation may not include other proteins, salts, or buffer molecules. These other components can affect the solubility and stability of the peptide, and can also affect the ionic strength of the solution.

### 6.3. Ionic strength

Ionic strength ( $I$ ) is calculated using the formula:

$$I = \frac{1}{2} \sum_i c_i z_i^2$$

Where:

- $I$  represents the ionic strength.
- $c_i$  represents the concentration of each ion species.
- $z_i$  represents the charge of each ion species.

Taken T = 370 K as an example

To make it possible we have to calculate  $c_{\text{Na}}$  and  $c_{\text{Cl}}$  in both cases.

In the first case,  $c_{\text{Cl}} = 0.01307 \text{ mol/L}$

$$I = \frac{1}{2} \cdot 0.01307 \text{ mol/L} \cdot (-1)^2 = 0.006535 \text{ mol/L}$$

In the second case there are 11 ions of  $\text{Cl}^-$  and 10 ions of  $\text{Na}^+$ . For this reason, we can employ an easy multiplication

$$I = \frac{1}{2} \cdot 11 \cdot 0.01307 \text{ mol/L} \cdot (-1)^2 + 10 \cdot 0.01307 \text{ mol/L} \cdot (1)^2 = \frac{1}{2} \cdot 21 \cdot 0.01307 \text{ mol/L} = 0.137235 \text{ mol/L}$$

ions	300K	340K	380K
none	A(0)	B(0)	C(0)
1 Cl <sup>-</sup>	D(8)	E(7)	F(7)
11Cl <sup>-</sup> + 10Na <sup>+</sup>	G(159)	H(149)	I(137)

Table 4: Table describing values of Ionic strength depending on different systems [mM]

Evidently, it is discernible that an elevated ionic strength corresponds to a heightened bonding strength, a relationship already observed in our system when additional ions were introduced into the solution. The primary distinction between the "effective" concentration in a simulation and the "actual" concentration in an experimental setting lies in the scale and volume. Simulations are typically conducted within finite-sized containers, wherein concentrations are constrained by the number of particles contained within those confines.

In contrast, experiments transpire in significantly larger volumes, affording a more uniform concentration profile that accurately represents the bulk solution. Furthermore, simulations often lack inclusion of all the components present in a genuine experimental sample. For instance, other proteins, salts, or buffer molecules may be absent from the simulation, although they significantly influence the solubility, stability of the peptide, and the overall ionic strength of the solution. It's worth noting that, in comparison to real-world experimental conditions, electrostatic interactions are often overemphasized in simulations.

#### 6.4. Force-field interaction

Discussing the values

**Table 5 Description of Molecular Interactions**

Interaction Type	Description
Bonds	Covalent bonds between atoms in the molecule, such as peptide backbone and side-chain bonds. Relatively stiff and do not significantly affect folding/unfolding equilibrium.
Angles	Bending of chemical bonds in the molecule. Contribute to maintaining geometry and secondary structure but are not dominant factors in folding/unfolding.
Dihedrals	Rotation about single bonds. Play a significant role in determining peptide's conformational flexibility and may influence folding/unfolding process.
Lennard-Jones (Nonbonded van der Waals) Interactions	Account for van der Waals forces between atoms. Crucial for maintaining compactness and structural stability of the folded state.
Electrostatic Interactions	Electrostatic forces between charged atoms or groups within the peptide. They have the most important influence in the stability of the molecule.

As said earlier, the treatment of electrostatic interactions in force fields, especially in classical molecular dynamics simulations, can lead to significant biases in results.

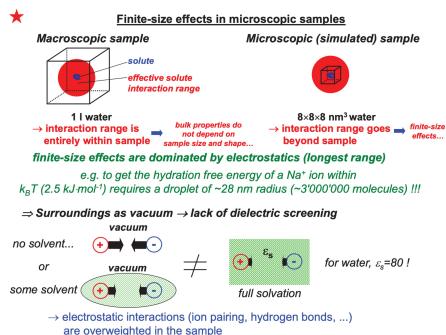


Figure 10: Slide describing the importance of electrostatic interactions

The two primary approximations that can introduce biases are:

**Table 6 Approximations in Electrostatic Interaction Treatment**

Approximation	Description
Cutoffs and Periodic Boundary Conditions	Many simulations use cutoff distances to truncate long-range electrostatic interactions. This is due to computational limitations. Using a cutoff introduces an artificial discontinuity in the electrostatic potential and may lead to errors, particularly when dealing with charged molecules like peptides. The use of cutoffs can result in poor representations of long-range electrostatic interactions.

**Table 6 Approximations in Electrostatic Interaction Treatment**

Approximation	Description
Force-Field Parameterization	Force fields rely on parameterization of terms like partial charges for atoms. These parameters are often optimized for a specific set of experimental data or to reproduce certain properties. If the force-field parameters for the peptide are not well-calibrated or transferable across various conditions or environments, they can introduce inaccuracies in simulations, affecting folding/unfolding behavior.

To mitigate these biases, researchers often employ techniques like Particle Mesh Ewald (PME) or Smooth Particle Mesh Ewald (SPME) for more accurate treatment of electrostatic interactions, removing the cutoff artifacts

**Insight: Particle Mesh Ewald (PME) and Smooth Particle Mesh Ewald (SPME)**

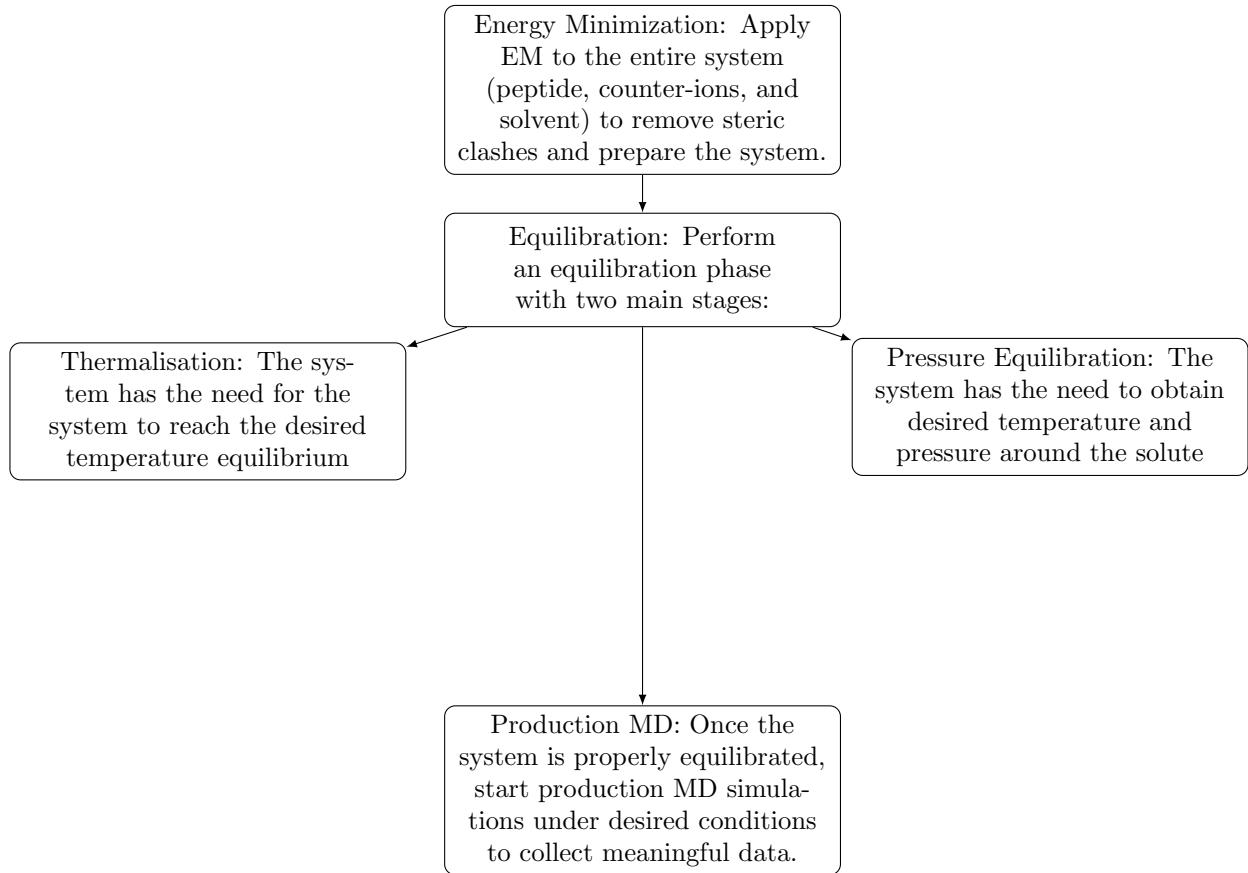
Particle Mesh Ewald (PME) and its extension, Smooth Particle Mesh Ewald (SPME), are essential techniques in molecular simulations for handling long-range electrostatic interactions in periodic systems. These methods use Fourier transforms and mesh grids to accurately compute electrostatic forces, offering significant improvements over simple cutoff-based approaches.

PME divides the electrostatic potential into short-range and long-range components, with the long-range part calculated using fast Fourier transforms. SPME further enhances this approach by introducing a damping function that results in a smoother and more precise treatment of long-range electrostatics.

In simulations, PME and SPME are indispensable for systems with periodic boundary conditions, especially those involving highly charged particles or strong electrostatic interactions. They play a pivotal role in providing a faithful representation of electrostatic forces, contributing to the accuracy of molecular dynamics simulations.

### 6.5. EM and Equilibria processes

In order to facilitate the relaxation of the system, the research efforts centered around the initial application of energy minimization (EM) before commencing the molecular dynamics (MD) simulations. In this particular scenario, the setup omitted the equilibration phase, both in terms of temperature and pressure. Notably, various temperature levels were taken into account in multiple experiments. Given the system's highly constrained dimensions, ensuring homogeneity of temperature throughout the environment assumed paramount importance. Conversely, from the perspective of pressure, it did not constitute the primary objective in our case. Nevertheless, it remains essential to consider pressure equilibration when feasible, as it can impact the system's behavior and stability.



## 7. Appendix

### Summary 1

You can find the list for the Google Colab Jupyter here.