

Modelado Molecular

PRÁCTICA MODELO TRIPEPTIDO AQA

Fernando Lucas Ruiz

14 de marzo de 2024



ÍNDICE

1. Introducción	3
2. Construir el modelo	3
2.1. Generamos la unidad celular cúbica	4
2.2. Fase de equilibrado	5
3. Análisis	7
3.1. Temperatura	8
3.2. Energía total y cinética	9
3.3. Radio de giro	11
3.4. Distancias de enlace	12
3.5. Ángulos de enlace	14
3.6. Ángulos diedros	15
3.7. Velocidades	18
4. Análisis extendido	20
4.1. Temperatura	21
4.2. Velocidad de un átomo en el espacio	22
4.3. Plot de Ramachandran	23

INTRODUCCIÓN

En el campo de la biofísica y la bioquímica, la comprensión detallada de las proteínas y su comportamiento bajo diferentes condiciones ambientales es fundamental para descifrar los misterios de los procesos biológicos y el diseño de terapéuticos específicos. Esta práctica estudia del tripéptido AQA mediante simulaciones de dinámica molecular utilizando GROMACS, una potente herramienta de simulación para la modelización de sistemas bioquímicos. Centrándose en la construcción del modelo de tripéptido AQA, desde la generación de su topología hasta el estudio termodinámico, este informe ilustra la preparación esencial para el análisis termodinámico.

CONSTRUIR EL MODELO

El tripéptido a simular es AQA (aqa.pdb) como se muestra en la **Figura 1**.

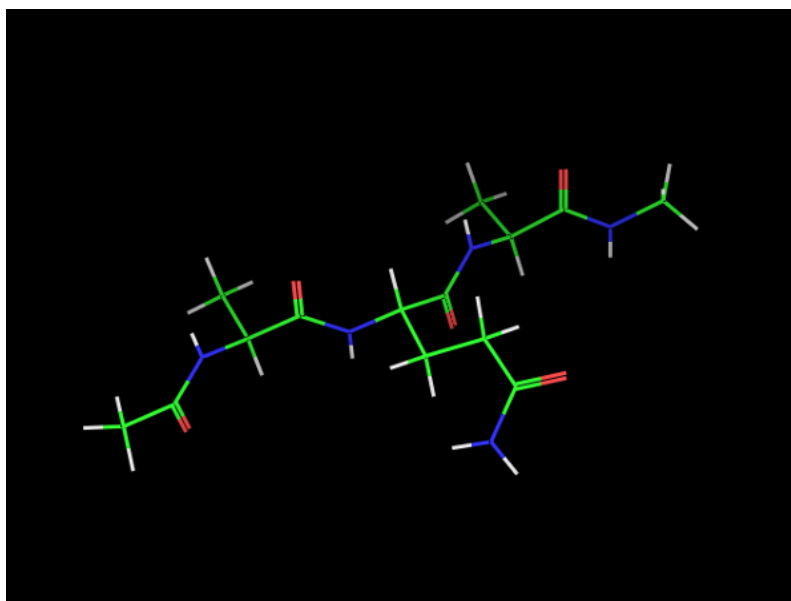


Figura 1: Proteína AQA.

La molécula está compuesta por tres residuos: dos de alanina y uno de glutamina en la posición central. Mediante la inspección del archivo aqa.pdb (usando el comando 'less aqa.pdb'), podemos confirmar que los extremos N-terminal y C-terminal incluyen seis átomos cada uno, como se anticipaba. Cada residuo de alanina posee un carbono α y un carbono β , estando este último unido a tres átomos de hidrógeno. La glutamina es un residuo más complejo: al carbono β le sigue un carbono γ , que a su vez está enlazado a un carbono δ . Este último está doblemente enlazado a un oxígeno y a un nitrógeno, añadiendo diversidad a la estructura

molecular.

Generamos la topología a raíz de los datos que tenemos del archivo `aqa.pdb` para que GROMACS pueda utilizarlo. Para ello utilizamos la siguiente línea de código:

- `gmx pdb2gmx -f aqa.pdb -o aqa.gro -p aqa.top -ter`

donde `-f` es el archivo input; `-o` es el archivo de salida de gromacs; `-p` el archivo de salida con la topología; `-ter` selección interactiva de terminals.

Con ello generamos dos archivos **aqa.gro** y **aqa.top**, con la información necesaria para siguientes pasos.

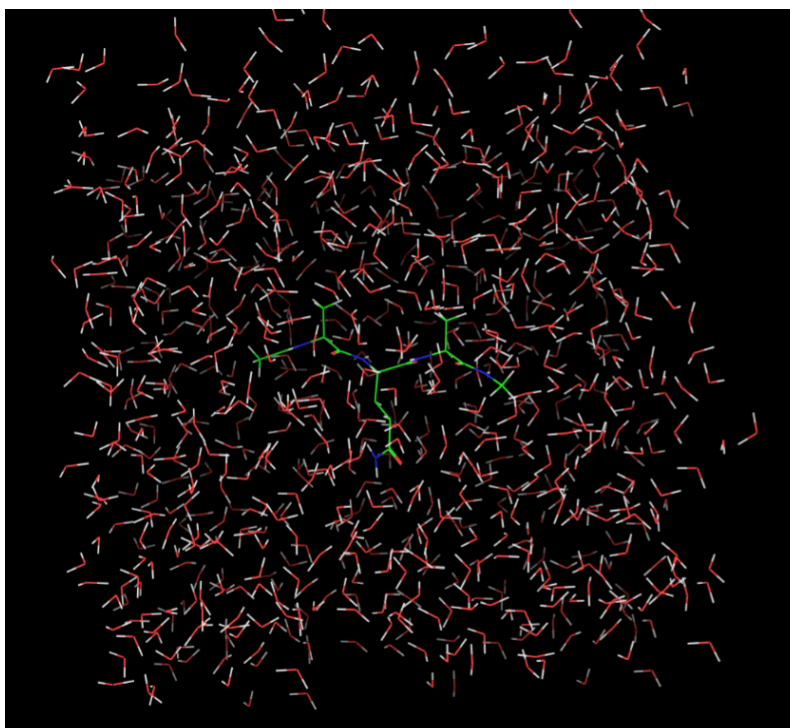


Figura 2: Proteína AQA con moléculas de agua.

2.1 GENERAMOS LA UNIDAD CELULAR CÚBICA

Ahora vamos a crear el cubo donde vamos a estudiar nuestro modelo, en el cual luego introduciremos las moléculas de agua. Para ello introducimos la siguiente línea de código:

- `gmx editconf -f aqa.gro -o aqa-box.gro -bt cubic -box 3.0 3.0 3.0`

donde `-f` es el archivo input; `-o` es el archivo de salida de gromacs; `bt` es el tipo de caja (triclinico, cubic, dodecahedro u octaedro); `-box` es la longitud en milímetros de los lados de la caja

Se genera un archivo llamado **aqa-box.gro**

El siguiente paso es introducir las moléculas de agua en el cubo (**Figura 2**). Para ello ejecutamos la siguiente línea de código:

- `gmx solvate -cp aqa-box.gro -cs -o aqa-box-solv.gro -p aqa.top`

Donde `-cp` es el fichero de entrada con la estructura de la caja junto a la proteína; `-cs -o` el archive de salida; `-p` es el archivo con la topología de la proteína. Se genera un archivo llamado **aqa-box-solv.gro** con la proteína dentro de la caja junto a las moléculas de agua alrededor (**Figura 2**).

2.2 FASE DE EQUILIBRADO

Esta etapa es crucial para simular nuestra molécula bajo determinadas temperaturas, lo que permitirá analizar las energías y velocidades de los átomos en variadas condiciones. Emplearemos dos niveles de temperatura específicos: 298 K y 400 K. Dentro del archivo **equiNVT.mdp** especificamos los parámetros de equilibrio para nuestros modelos moleculares. Generamos dos archivos de configuración, **equiNVT_298K.mdp** y **equiNVT_400K.mdp**, cada uno ajustado a las condiciones térmicas respectivas. En la entrada 'ref_t' del archivo de configuración, asignamos el valor de 298 o 400 para reflejar la temperatura correspondiente.

```
integrator          = md          ; leap-frog integrator
dt                  = 0.0005      ; 0.5 fs
nsteps              = 800000      ; 400 ps
...
nstxout             = 4000        ; Save coordinates every 0.1 ps
nstvout             = 4000        ; Save velocities every 0.1 ps
nstfout             = 0           ; Save forces
nstlog              = 4000        ; Update log every 0.1 ps
nstenergy           = 4000        ; Save energies every 0.1 ps
nstcalcenergy       = 4000
...
ref_t               = 298 or 400 ; Above fluid phase transition
```

El código para ejecutarlo es:

- `gmx grompp -f equiNVT_298k.mdp -c aqa-box-solv.gro -p aqa.top -o aqa-a_298K.tpr`
- `gmx grompp -f equiNVT_400k.mdp -c aqa-box-solv.gro -p aqa.top -o aqa-a_400K.tpr`

Donde `-f` es el archivo donde están los parámetros; `-c` el archive con la proteína dentro de la caja con sus moléculas de agua; `-p` la topología; `-o` el archive de salida.

Se generan dos archivos **aqa-a_298K.tpr** y **aqa-a_400K.tpr**

Tras la ejecución, la ausencia de advertencias sobre desequilibrios de carga en nuestro modelo indica que el tripeptido es eléctricamente neutro. Esto nos permite obviar cualquier paso adicional de neutralización de

cargas para este tripeptido específico.

El procedimiento siguiente implica el equilibrio térmico de dos modelos preparados para operar bajo dos diferentes condiciones de temperatura. Para efectuar este proceso, es necesario crear un script de Shell que posteriormente enviaremos a la cola de trabajos 'ecq-k'. El script de Shell necesario para ejecutar la simulación con el modelo a 298 K es:

```
#!/bin/bash
#
#SBATCH -p eck-q
#SBATCH --chdir=/home/alumno14/modelado_molecular/tarea/2-equilibration
#SBATCH -J equilibrado
#SBATCH --cpus-per-task=1

date
gmx mdrun -s aqa-a_298K.tpr -mp aqa.top -c aqa_298K.g96 -nt 1
date
```

Donde -s es el archivo de entrada; -mp el archivo con la topología; -c el archivo de salida.

El shell script utilizado para ejecutar la simulación con el modelo a 400K es similar al empleado para 298K, con la diferencia de que se refiere a los archivos correspondientes a 400K. Procedemos a lanzar ambos scripts mediante el comando sbatch, lo que resulta en la creación de seis archivos: dos archivos slurm con nombres **slurm_[número]**, que contienen registros del proceso en la cola de ejecución, dos archivos **aqa_298K.g96** y **aqa_400K.g96**, que almacenan las posiciones y velocidades de cada átomo, y dos archivos de registro **aqa_298k.log** y **aqa_400k.log**, que ofrecen información detallada sobre los modelos de simulación.

ANÁLISIS

Para calcular las trayectorias de ambos modelos, pasaremos a la fase de análisis. Inicialmente, es necesario editar el archivo `runNVT.mdp` para reflejar las condiciones de nuestros experimentos. A partir de este archivo, generaremos dos nuevos archivos, uno para cada condición térmica: **`runNVT_298K.mdp`** y **`runNVT_400K.mdp`**, correspondientes a las temperaturas de los modelos que vamos a analizar.

```
ref_t                = 298 or 400                ; Avobe fluid phase trasition
```

Cuando tengamos los archivos, ejecutamos el siguiente código para generar los ficheros .tpr:

- `gmx grompp -f runNVT_298k.mdp -c aqa_298.g96 -p aqa.top -o aqa_298k.tpr`
- `gmx grompp -f runNVT_400k.mdp -c aqa_400.g96 -p aqa.top -o aqa_400k.tpr`

Donde -f el archivo con los parámetros; -c archivo .g96; -p el archico con la topología; -o fichero de salida.

Se generan dos archivos **`aqa_298k.tpr`** y **`aqa_400k.tpr`**.

Procederemos a reenviar la tarea de equilibrado a la cola de trabajos 'ecq-k'. Para ello, crearemos dos scripts Shell distintos, uno para cada uno de los ficheros correspondientes a los modelos.

```
#!/bin/bash
#
#SBATCH -p eck-q
#SBATCH --chdir=/home/alumno14/modelado_molecular/tarea/3-run
#SBATCH -J equilibrado_run
#SBATCH --cpus-per-task=1

date

gmx mdrun -s aqa_298k.tpr -mp aqa.top -c aqa_298K.g96 -o aqa_298k.trr ...
...-e aqa_298k.edr -nt 1

date
```

Donde -s es el fichero .tpr; -mp el fichero con la topología; -c el archivo .g96; -o el fichero de salida con las trayectorias; -e el fichero de salida con las energías.

Finalmente después de lanzar a la cola ecq-k tendremos 4 ficheros nuevos: **`aqa_298k.trr`**, **`aqa_298k.edr`**, **`aqa_400k.trr`** y **`aqa_400k.edr`**

Si ejecutamos ahora el código podremos ver las cinéticas y trayectorias de nuestros modelos.

- `gmx traj -f aqa_298k.trr -s aqa.tpr -oxt cartoon_298k.pdb`
- `gmx traj -f aqa_400k.trr -s aqa.tpr -oxt cartoon_400k.pdb`

3.1 TEMPERATURA

Lo primero de todo vamos a ver si la temperatura de nuestros modelos están correctamente simuladas. Para calcular la temperatura de nuestros modelos, debemos ejecutar el siguiente código para cada uno de los modelos con temperaturas distintas.

- `gmx traj -f aqa_298k.trr -s aqa_298k.tpr -xvg none -ot temp_aqa_298k.xvg`
- `gmx traj -f aqa_400k.trr -s aqa_400k.tpr -xvg none -ot temp_aqa_400k.xvg`

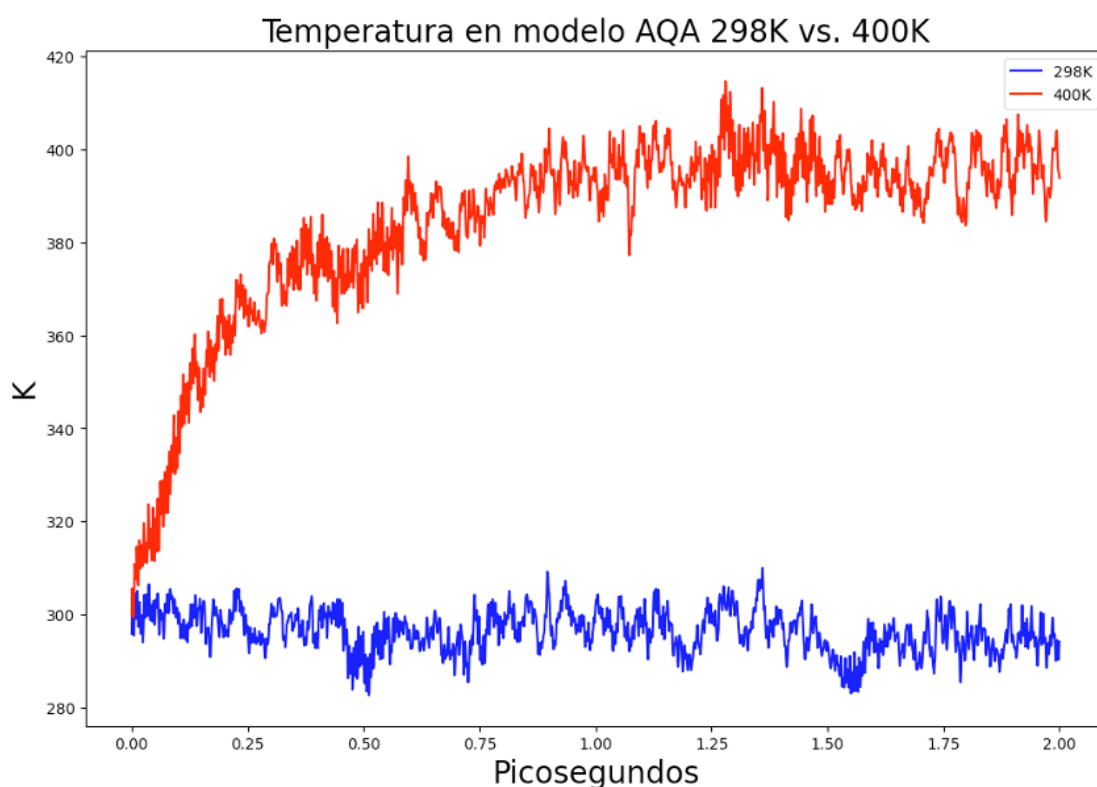


Figura 3: Temperatura de los dos modelos de la proteína AQA.

La simulación realizada a 298K revela una notable estabilidad térmica a lo largo del tiempo, lo que señala que el sistema ha logrado un equilibrio térmico satisfactorio. Por otro lado, la simulación efectuada a 400K nos muestra una dinámica inicial más volátil, con un aumento progresivo en la temperatura hasta estabilizarse, lo que sugiere un proceso de calentamiento o equilibrio donde la temperatura se incrementa de manera controlada hasta alcanzar el punto de 400K deseado (**Figura 3**).

Es interesante observar que en ambas simulaciones existen fluctuaciones alrededor de sus respectivas temperaturas objetivo. Estas variaciones son características esperadas y representan la interacción energética dinámica entre los átomos y el termostato implementado en la simulación, reafirmando que el sistema no está completamente inerte sino que está participando en un intercambio energético constante que mantiene el

equilibrio térmico.

3.2 ENERGÍA TOTAL Y CINÉTICA

Analizar la energía cinética y la energía total en simulaciones de dinámica molecular, como las realizadas en GROMACS para un tripéptido como AQA, aporta insights cruciales a una multitud de aplicaciones científicas. En el ámbito microscópico, la energía cinética es una función directa de la temperatura, vinculando la agitación térmica de las partículas con la cantidad de calor presente en el sistema. Comprender esta relación es clave para interpretar cómo la temperatura influye en la actividad y el comportamiento molecular.

Las proteínas, con su complejidad y su rol central en la biología, son especialmente sensibles a los cambios de temperatura. Al monitorear las variaciones de energía con la temperatura en simulaciones, se pueden identificar aquellos momentos críticos donde la proteína se despliega o desnaturaliza, fenómenos que son fundamentales para comprender su estabilidad y funcionalidad. Las energías en juego también reflejan la flexibilidad de la proteína y sus movimientos intramoleculares, detalles que son esenciales para desentrañar los mecanismos subyacentes a su función biológica. A temperaturas elevadas, se pueden producir cambios en la energía cinética que sugieran una alteración de las interacciones moleculares, tales como el debilitamiento de los enlaces de hidrógeno, lo que a su vez podría tener implicaciones en la integridad de las estructuras secundarias y terciarias de la proteína. Estos cambios son esenciales para la predicción de la interacción de las proteínas con posibles fármacos y para ajustar las condiciones experimentales, como ensayos de actividad enzimática.

Para conseguir nuestras energías totales y cinéticas de nuestros modelos, ejecutamos el siguiente código en GROMACS y mostramos los resultados en el **Cuadro 1**:

- `gmx energy -f aqa_298k.edr -s aqa_298k.tpr -o total-energy_298k.xvg -xvg none`
- `gmx energy -f aqa_400k.edr -s aqa_400k.tpr -o total-energy_400k.xvg -xvg none`
- `gmx energy -f aqa_298k.edr -s aqa_298k.tpr -o kinetic-energy_298k.xvg -xvg none`
- `gmx energy -f aqa_400k.edr -s aqa_400k.tpr -o kinetic-energy_400k.xvg -xvg none`

	Total (kJ/mol)		Cinética (kJ/mol)	
	298 k	400 k	298 k	400 k
Average:	-24868.1	-17902	9707.92	12585.1
Err.Est:	48	830	27	270
RMSD:	152.024	1819.66	143.725	654.01
Tot-Drift:	38.0643	5107.56	-143.567	1687.6

Cuadro 1: Energía total y cinética.

Como vemos en la **Figura 4**, la energía total a 298K del sistema se estabiliza rápidamente y permanece bastante constante durante el transcurso de la simulación. Esto sugiere que el sistema alcanza un estado de

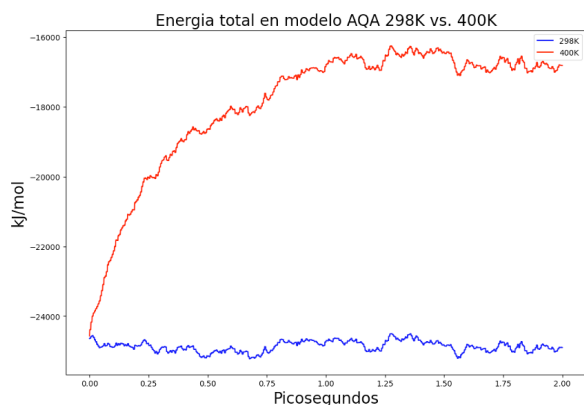


Figura 4: Energía Total de la proteína AQA a 298K y 400K.

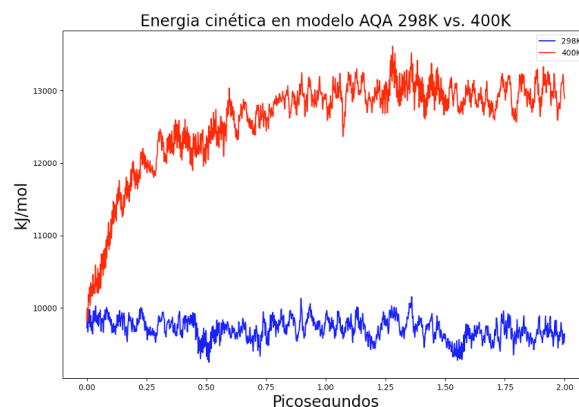


Figura 5: Energía cinética de la proteína AQA a 298K y 400K.

equilibrio térmico y se mantiene estable en esta temperatura más baja. Por otro lado, la línea roja que representa la simulación a 400K asciende gradualmente hasta estabilizarse. Aunque también alcanza un estado de equilibrio, lo hace en un nivel de energía total mucho más alto en comparación con la simulación a 298K. Esto es coherente con la expectativa de que a temperaturas más altas, las moléculas del sistema tienen una energía cinética mayor debido a la agitación térmica incrementada. Las fluctuaciones en ambas temperaturas son mínimas, lo que indica que ambos sistemas están en equilibrio a lo largo del período de simulación observado. No obstante, es notable que el sistema a 400K requiere más energía para mantenerse en equilibrio, lo cual es típico debido a la mayor energía térmica de las moléculas a esta temperatura. Esto puede influir en la estabilidad de la estructura del tripéptido, su dinámica y sus posibles interacciones con otros biomoléculas o fármacos en el entorno.

En la **Figura 5** de la energía cinética observamos que la línea azul (298K) se mantiene relativamente constante a lo largo del tiempo, con fluctuaciones dentro de un rango limitado. Esto indica que, a 298K, las partículas del sistema tienen menor energía cinética, lo cual es consistente con la expectativa de que a menor temperatura, hay menor agitación térmica. Por otro lado, la línea roja (400K) comienza y se mantiene en un nivel significativamente más alto que la línea azul. Esta diferencia se alinea con la comprensión de que a mayor temperatura, las partículas adquieren mayor energía cinética debido a la agitación térmica aumentada. La línea roja muestra una mayor variabilidad y no parece disminuir o estabilizarse en un valor más bajo, lo que sugiere que el sistema mantiene una energía cinética elevada a lo largo de toda la simulación.

La comparación directa entre las dos temperaturas indica que el sistema a 400K tiene más energía para facilitar procesos endotérmicos, como reacciones químicas o cambios conformacionales, que podrían no ser posibles a 298K. Estos datos son importantes para comprender la estabilidad de la proteína y su comportamiento en condiciones de temperatura variadas, y pueden tener implicaciones en la funcionalidad de la proteína en diferentes entornos biológicos o industriales.

3.3 RADIO DE GIRO

El radio de giro es una medida de la extensión tridimensional de una molécula y es particularmente útil en el estudio de macromoléculas complejas como las proteínas. En el contexto de simulaciones de dinámica molecular, el estudio del radio de giro puede servir para varios propósitos importantes como proporcionar una medida cuantitativa de qué tan compacta o extendida está una proteína o un polímero en el espacio. Un radio de giro pequeño indica una conformación más compacta, mientras que un valor más grande sugiere una conformación más extendida o desplegada. Al observar cómo cambia el radio de giro a lo largo del tiempo, se pueden detectar cambios en la estructura de la proteína, como el plegamiento, desplegamiento o agregación. También puede indicar la tendencia de una proteína a formar agregados; las proteínas con un radio de giro que aumenta significativamente con el tiempo pueden estar propensas a la agregación. La medición del radio de giro a diferentes temperaturas puede ofrecer información sobre la estabilidad térmica de la proteína y cómo la temperatura afecta su estructura global.

Para calcular los radios de giro (**Figura 6**) utilizamos la siguiente línea de código:

- `gmx gyrate -f aqa_298k.trr -s aqa_298k.tpr -o aqa_298k.xvg -xvg none`
- `gmx gyrate -f aqa_400k.trr -s aqa_400k.tpr -o aqa_400k.xvg -xvg none`

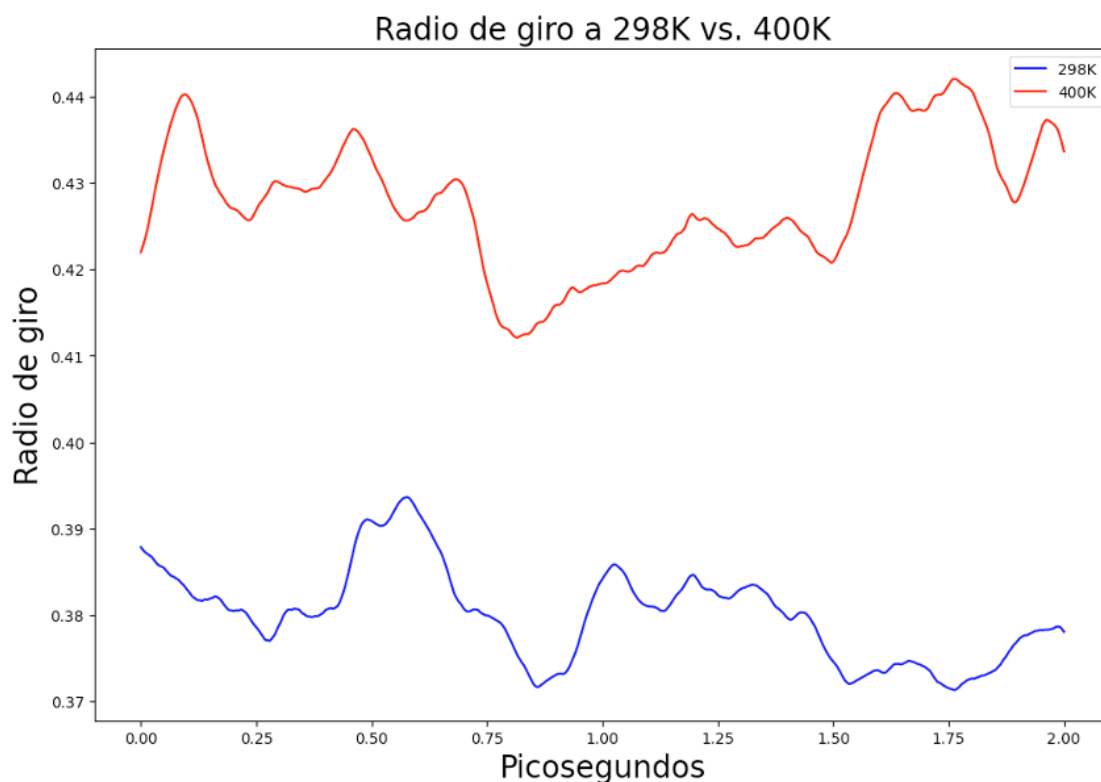


Figura 6: Radio de giro de la proteína AQA.

El gráfico (**Figura 6**) ilustra una comparativa detallada del radio de giro en función del tiempo, expresado

en picosegundos, para dos distintas temperaturas: 298 Kelvin, representada por la línea azul, y 400 Kelvin, denotada por la línea roja. El análisis de las variaciones en las trayectorias del radio de giro puede desvelar insights cruciales acerca de la influencia de la temperatura en la conformación y el comportamiento dinámico de la molécula o del sistema en estudio.

Con el incremento de la temperatura, se observan típicamente variaciones en el radio de giro, atribuibles a la incrementada energía térmica a disposición de las moléculas. Este aumento energético puede inducir una mayor movilidad molecular y, potencialmente, alteraciones estructurales significativas. Dichos cambios estructurales se manifiestan a través de una ampliación o reducción del radio de giro, lo cual depende de si la molécula en cuestión se dilata o se condensa en respuesta al ascenso térmico. En el escenario descrito, se registra un incremento en el radio de giro correspondiente al aumento de la temperatura, sugiriendo una tendencia de la proteína a adoptar una forma más expandida.

3.4 DISTANCIAS DE ENLACE

El análisis de la variabilidad en la distancia entre pares de átomos a lo largo del tiempo puede arrojar luz sobre la manera en que la temperatura influye en sus interacciones mutuas, facilitando así una comprensión más profunda de las transiciones conformacionales, la dinámica de las interacciones moleculares y la integridad estructural bajo un rango de condiciones termales. Estos cambios observados pueden ser un reflejo de variaciones en la flexibilidad de la molécula, su estabilidad intrínseca o la naturaleza de las interacciones entre regiones específicas de la molécula, aspectos todos ellos que revisten gran relevancia en los estudios de dinámica molecular. Estos estudios son cruciales para descifrar el comportamiento de sistemas biológicos, la funcionalidad de materiales y la química de complejos moleculares.

Para profundizar en este análisis, seleccionaré dos átomos vinculados por una relación fisicoquímica directa. He optado por investigar la distancia entre el Carbono δ y el Nitrógeno del grupo amino en el residuo de Glutamina, y entre el Carbono α y el Carbono β del segundo residuo de Alanina. La observación meticulosa de estas uniones específicas nos proporcionará un entendimiento detallado de cómo las condiciones térmicas afectan estas interacciones moleculares esenciales.

En el archivo **distances.ndx** ponemos lo siguiente:

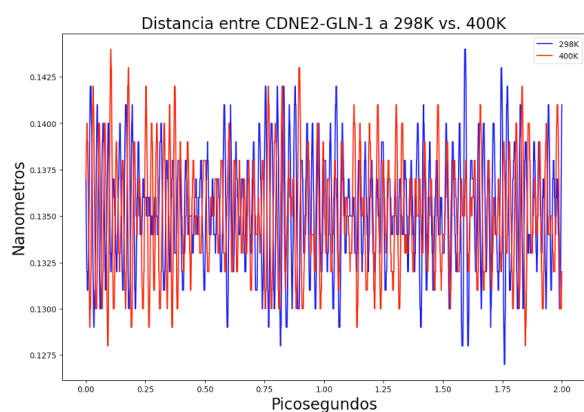
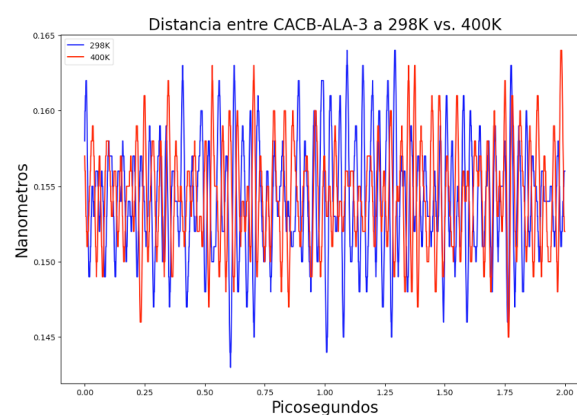
```
[CDNE2-GLN-1]
27 29
[CACB-ALA-3]
36 38
```

Y ejecutamos seguidamente el código y mostramos los resultados en el **cuadro 2**:

- *gmx distance -f aqa_298k.trr -s aqa_298k.tpr -n distances.ndx -oall dist_1_aqa_298k.svg -xvg none*
- *gmx distance -f aqa_298k.trr -s aqa_298k.tpr -n distances.ndx -oall dist_2_aqa_298k.svg -xvg none*
- *gmx distance -f aqa_400k.trr -s aqa_400k.tpr -n distances.ndx -oall dist_1_aqa_400k.svg -xvg none*
- *gmx distance -f aqa_400k.trr -s aqa_400k.tpr -n distances.ndx -oall dist_2_aqa_400k.svg -xvg none*

	CDNE2-GLN-1		CACB-ALA-3	
	298 k	400 k	298 k	400 k
Number of samples:	2001	2001	2001	2001
Average distance:	0.13527 nm	0.13511 nm	0.15417 nm	0.15436 nm
Standard deviation:	0.00304 nm	0.00298 nm	0.00378 nm	0.00336 nm

Cuadro 2: Distancia media entre átomos.

Figura 7: Distancia de enlace entre el Carbono δ y el Nitrógeno del grupo amino de la Glutamina.Figura 8: Distancia de enlace entre el Carbono α y el Carbono β del segundo residuo de Alanina.

En la primera gráfica (**Figura 7**) se muestran las variaciones térmicas en la distancia entre átomos entre el Carbono δ y el Nitrógeno del grupo amino en el residuo de Glutamina, bajo dos diferentes condiciones de temperatura. Los resultados a 298 K están representados en azul, mientras que los de 400 K se muestran en rojo. Se aprecia que las fluctuaciones son similares para ambas temperaturas, corroborando nuestras hipótesis iniciales basadas en el promedio de distancias observadas en el **Cuadro 2**. Esta consistencia sugiere que la distancia interatómica en el enlace estudiado mantiene una estabilidad estructural a pesar de las variaciones térmicas significativas.

Similar a la primera gráfica, en esta se muestra la distancia entre el Carbono α y el Carbono β del segundo residuo de Alanina a 298 K y 400 K (**Figura 8**). De nuevo, la tendencia es que a mayor temperatura, no hay un cambio en las fluctuaciones de la distancia. Lo cual puede sugerir que las interacciones y la estructura alrededor de estos átomos son resistentes al cambio térmico o que la región es intrínsecamente más rígida.

3.5 ÁNGULOS DE ENLACE

Estudiar el ángulo entre tres átomos en un tripeptido como AQA (Ala-Gln-Ala) a diferentes temperaturas puede proporcionar información detallada sobre la conformación y la dinámica del péptido, lo que es crucial para entender su función y comportamiento en diferentes entornos. Los ángulos de enlace determinan la forma tridimensional del péptido, la cual está intrínsecamente ligada a su funcionalidad. Estudiar estos ángulos puede ayudar a identificar conformaciones específicas que pueden ser críticas para la interacción con otros biomoléculas o para la actividad biológica del péptido. Al observar cómo varían los ángulos con la temperatura, podemos obtener una comprensión de la flexibilidad del péptido. A temperaturas más altas, las moléculas tienen más energía y por lo tanto pueden explorar un rango más amplio de conformaciones debido al aumento en la vibración y rotación de los enlaces. Pueden ocurrir cambios en los patrones de movimiento o en la estructura general a ciertas temperaturas (transiciones de fase). Estos pueden indicar eventos importantes como el plegamiento de proteínas, la formación de estructuras secundarias (α -hélices, β -láminas) o cambios en el estado de agregación. En resumen, al estudiar los ángulos entre átomos en péptidos a diferentes temperaturas, podemos obtener una imagen detallada de la relación entre estructura, energía y función, lo que tiene aplicaciones en bioquímica, diseño de fármacos, y biotecnología.

Para esta práctica voy ahora a calcular los ángulos de dos combinaciones de átomos. El primero es la combinación entre el Carbono γ , el Carbono δ y el nitrógeno del grupo amino del residuo de glutamina. La segunda combinación es el Carbono α , el Carbono y el Oxígeno del residuo de la primera alanina Creamos un archivo llamado `distances_angles.ndx` con lo siguiente:

```
[CG-CD-NE2-GLN-2]
24 27 29
[CA-C-O-ALA-1]
9 15 16
```

Para ello utilizamos el siguiente código y mostramos los resultados en el **Cuadro 3**:

- `gmx angle -f aqa_298k.trr -n distances_angles.ndx -ov angle_1_aqa_298k.xvg -xvg none`
- `gmx angle -f aqa_298k.trr -n distances_angles.ndx -ov angle_2_aqa_298k.xvg -xvg none`
- `gmx angle -f aqa_400k.trr -n distances_angles.ndx -ov angle_1_aqa_400k.xvg -xvg none`
- `gmx angle -f aqa_400k.trr -n distances_angles.ndx -ov angle_2_aqa_400k.xvg -xvg none`

La visualización del comportamiento dinámico del ángulo la cadena lateral del residuo de glutamina en sus átomos Carbono γ , el Carbono δ y el nitrógeno del grupo amino, a través de la primera gráfica de átomos (**Figura 9**), revela una notable flexibilidad en esta región molecular. A una temperatura de 298K, representada por la línea azul, se observa una serie de fluctuaciones que ilustran la diversidad de movimientos conforma-

	CG-CD-NE2-GLN-2		CA-C-O-ALA-1	
	298 k	400 k	298 k	400 k
<angle>:	117.024	116.58	121.64	121.602
<angle ² >:	13705.3	13603.2	14802.8	14795
Standard deviation:	3.28516	3.50888	2.56386	2.81238

Cuadro 3: Ángulo entre átomos.

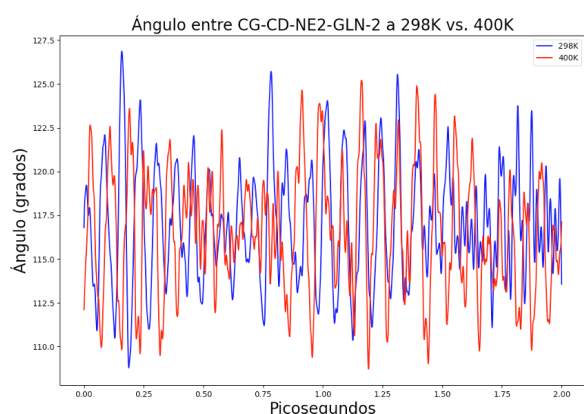


Figura 9: Ángulo de enlace entre el Carbono γ , el Carbono δ y el nitrógeno del grupo amino del residuo de glutamina.

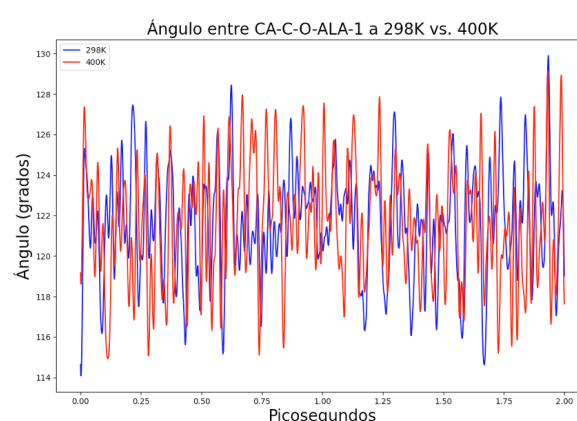


Figura 10: Ángulo de enlace entre el Carbono α , el Carbono y el Oxígeno del residuo de la primera alanina.

cionales accesibles al residuo en estas condiciones térmicas. Al escalar la temperatura a 400K, demarcada por la línea roja, se percibe un patrón de fluctuación similar, sugiriendo que la variación en la temperatura hasta este punto no induce cambios significativos en la amplitud de movimiento conformacional de este grupo específico de átomos, tal y como se confirmó previamente en los datos del **Cuadro 3**.

La segunda gráfica del análisis de ángulos (**Figura 10**) expone la relación entre el grupo carboxilo al final del primer residuo de alanina y cómo responde ante dos temperaturas distintas. A pesar de la rigidez típica del grupo carboxilo, un elemento estructural clave de la proteína, la comparativa muestra que, tanto a 298K (azul) como a 400K (rojo), la variabilidad de sus átomos es similar. Esto sugiere que la estabilidad de esta zona de la alanina no se ve alterada significativamente entre las temperaturas estudiadas, un dato interesante que apunta a la robustez de esta región molecular en condiciones térmicas variables.

3.6 ÁNGULOS DIEDROS

La exploración de los ángulos diedros, o de torsión, es esencial en la elucidación de la estructura y la dinámica de un tripeptido a nivel molecular, especialmente bajo la influencia de variaciones térmicas. Estos ángulos críticos ofrecen una representación dinámica de las posibles conformaciones espaciales que las cadenas polipeptídicas pueden adoptar. Los cambios observados en estos ángulos reflejan la diversidad conformacional

y las preferencias estructurales, que están directamente influenciadas por las barreras energéticas. A medida que la temperatura se incrementa, los átomos ganan energía cinética, ampliando el rango de ángulos diedros posibles y revelando la flexibilidad intrínseca de la molécula bajo distintas condiciones termodinámicas.

El conocimiento de cómo fluctúan estos ángulos con la temperatura no solo arroja luz sobre la estabilidad de ciertas conformaciones estructurales sino que también indica cómo el calor puede promover o inhibir los procesos de plegamiento y desplegamiento, fundamentales para la función proteica. Estos cambios en la torsión pueden afectar tanto las interacciones internas como las intermoleculares, influyendo en la interacción del péptido con sustratos o fármacos, lo que es crucial para la creación de terapéuticos específicos.

El estudio detallado de la respuesta de los ángulos diedros frente a variaciones de temperatura tiene aplicaciones prácticas en el diseño de biomateriales y sistemas de administración de fármacos, donde la estabilidad térmica es un parámetro clave. La integración de datos experimentales con simulaciones computacionales puede mejorar la precisión de nuestras predicciones sobre el comportamiento de las proteínas en distintos estados funcionales o condiciones ambientales.

Para cuantificar estos ángulos diedros en el tripéptido AQA, hemos seleccionado los ángulos de la primera alanina y de la glutamina. Utilizando los siguientes códigos, podemos obtener los valores que nos permiten analizar estos ángulos críticos y sus variaciones durante la simulación.

- `gmx rama -f aqa_298k.trr -s aqa_298k.tpr -o rama_aqa_298k.svg -xvg none`
- `gmx rama -f aqa_400k.trr -s aqa_400k.tpr -o rama_aqa_400k.svg -xvg none`

Luego escogemos los ángulos que queremos estudiar. En este caso he escogido los átomos de la primera alanina y de la glutamina central.

- `grep ALA-2 rama_aqa_298k.svg | awk 'print $1' | cat -n >phi-ala-2_298k.dat`
- `grep ALA-2 rama_aqa_298k.svg | awk 'print $2' | cat -n >psi-ala-2_298k.dat`
- `grep ALA-2 rama_aqa_400k.svg | awk 'print $1' | cat -n >phi-ala-2_400k.dat`
- `grep ALA-2 rama_aqa_400k.svg | awk 'print $2' | cat -n >psi-ala-2_400k.dat`
- `grep GLN-3 rama_aqa_298k.svg | awk 'print $1' | cat -n >phi-gln-3_298k.dat`
- `grep GLN-3 rama_aqa_298k.svg | awk 'print $2' | cat -n >psi-gln-3_298k.dat`
- `grep GLN-3 rama_aqa_400k.svg | awk 'print $1' | cat -n >phi-gln-3_400k.dat`
- `grep GLN-3 rama_aqa_400k.svg | awk 'print $2' | cat -n >psi-gln-3_400k.dat`

Y unimos los ficheros para tenerlos en uno a cada temperatura y ángulo diedro.

- `join phi-ala-2_298k.dat psi-ala-2_298k.dat >ala-2_298k.dat`
- `join phi-ala-2_400k.dat psi-ala-2_400k.dat >ala-2_400k.dat`
- `join phi-gln-3_298k.dat psi-gln-3_298k.dat >gln-3_298k.dat`

• `join phi-gln-3_400k.dat psi-gln-3_400k.dat >gln-3_400k.dat`

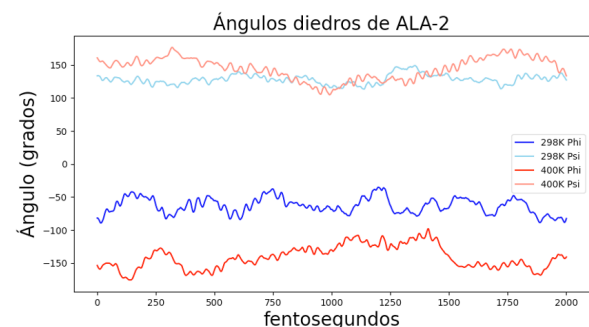


Figura 11: Ángulos diedros del primer residuo de Alanina.

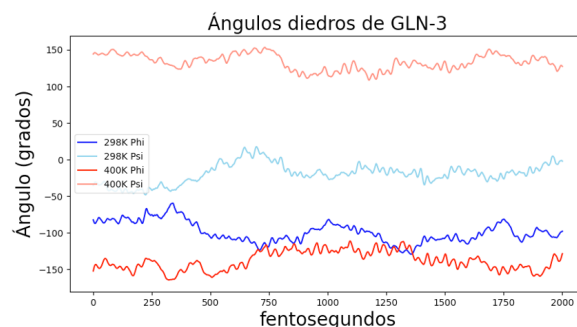


Figura 12: Ángulos diedros del residuo de Glutamina.

En la **Figura 11**, las líneas azul oscura (298K phi) y la línea azul clara (298K psi) parecen fluctuar dentro de un rango limitado, lo que sugiere una cierta estabilidad conformacional a la temperatura más baja. Esto indica que, a 298K, el residuo de Alanina tiende a mantener una conformación particular con menos cambio estructural a lo largo del tiempo. Al aumentar la temperatura a 400K, las líneas roja (400K phi) y roja clara (400K psi) también muestran fluctuaciones, pero con una amplitud aparentemente mayor, especialmente para el ángulo ψ . Esto implica que a una temperatura más alta, la molécula puede tener una mayor movilidad y explorar un conjunto más amplio de conformaciones, lo que refleja una mayor flexibilidad. A 298K, la diferencia entre las fluctuaciones de los ángulos ϕ y ψ es relativamente pequeña, lo que indica un comportamiento coherente en el espacio de configuración de la proteína. A 400K, parece haber una diferencia más significativa en el comportamiento de los ángulos ϕ y ψ , lo cual podría indicar que diferentes tipos de conformaciones estructurales son más accesibles o favorecidas a esta temperatura.

En la gráfica de la **Figura 12** se muestran los ángulos diedros ϕ y ψ del residuo de glutamina. A 298K, tanto el ángulo ϕ como el ψ presentan fluctuaciones a lo largo de la simulación. El rango de fluctuación es restringido, lo que sugiere una estabilidad estructural a esta temperatura más baja. Esto puede interpretarse como una preferencia por ciertas conformaciones de la cadena en esta temperatura. A 400K, los ángulos ϕ y ψ de la Glutamina exhiben una variabilidad más amplia. Especialmente notable es el ángulo ψ (400K psi), cuya línea roja clara muestra fluctuaciones significativas, indicando que la cadena lateral del residuo de glutamina tiene más libertad para explorar diferentes conformaciones a esta temperatura elevada. La diferencia en la estabilidad conformacional entre las temperaturas se puede deducir de la cantidad de cambio en los ángulos diedros. A 400K, la mayor amplitud de las fluctuaciones podría estar relacionada con una disminución en la estabilidad estructural o con un estado más dinámico de la proteína, permitiendo transiciones entre diferentes conformaciones.

3.7 VELOCIDADES

Analizar la velocidad de los átomos en un tripéptido a distintas temperaturas nos permite comprender su dinámica térmica y cómo esta interacciona con el medio circundante. Cada átomo, vibrando con energía cinética que varía directamente con la temperatura, actúa como un indicador de cómo los cambios térmicos afectan al sistema en su conjunto. Con el incremento de la temperatura, esperamos que la energía cinética, y por ende, la velocidad de los átomos, aumente. Este fenómeno ofrece un panorama cuantitativo sobre el impacto de la temperatura en la energía total del sistema.

En el terreno de las simulaciones de dinámica molecular, la velocidad atómica desempeña un papel esencial. Comparar estas velocidades en diferentes temperaturas nos permite refinar y validar nuestros modelos computacionales para que reflejen con mayor precisión el comportamiento físico real. En el mundo de las proteínas y los péptidos, donde la forma está intrínsecamente ligada a la función, la dinámica atómica es una pieza clave que puede influir significativamente en la estructura y la funcionalidad de estas macromoléculas.

Las variaciones en la velocidad atómica con la temperatura nos ofrecen pistas sobre cómo los tripéptidos responden a diferentes entornos fisiológicos. Un cambio en la temperatura puede precipitar transiciones entre diversos estados conformacionales. Las velocidades atómicas nos ofrecen un vistazo temprano a estas transiciones, lo cual es crucial para descifrar la estabilidad de las proteínas y sus mecanismos de plegamiento y desplegamiento.

Los estudios de velocidad pueden revelar cómo los átomos se ajustan tras las perturbaciones, cómo se modifican las vibraciones intramoleculares con la temperatura y cómo se puede alterar la estabilidad de las interacciones entre aminoácidos, afectando la formación y ruptura de enlaces, incluyendo enlaces de hidrógeno e interacciones hidrofóbicas. Tales dinámicas tienen implicaciones directas en las propiedades termodinámicas y de transporte, incluyendo la difusión, un aspecto de suma importancia en el diseño de fármacos, donde la movilidad de pequeñas moléculas es un factor crítico.

En contextos biotecnológicos y farmacéuticos, comprender la estabilidad térmica de los péptidos y las proteínas es esencial. Esta información es indispensable para el desarrollo de bioprocesos y la formulación de medicamentos, donde la resistencia a la variación de la temperatura es un criterio de diseño clave.

Para obtener datos precisos sobre las velocidades atómicas en nuestras simulaciones de GROMACS, ejecutamos los siguientes comandos específico.

- `gmx traj -f aqa_298k.trr -s aqa_298k.tpr -xvg none -ov vel_aqa_298k -len`
- `gmx traj -f aqa_400k.trr -s aqa_400k.tpr -xvg none -ov vel_aqa_400k -len`

He escogido las velocidades de los 5 primeros átomos.

En contra de lo que podíamos esperar, observamos que, con la excepción del primer átomo, la velocidad a 298K supera ligeramente a la registrada a 400K (**Figura 13**). Teóricamente, deberíamos anticipar un aumento

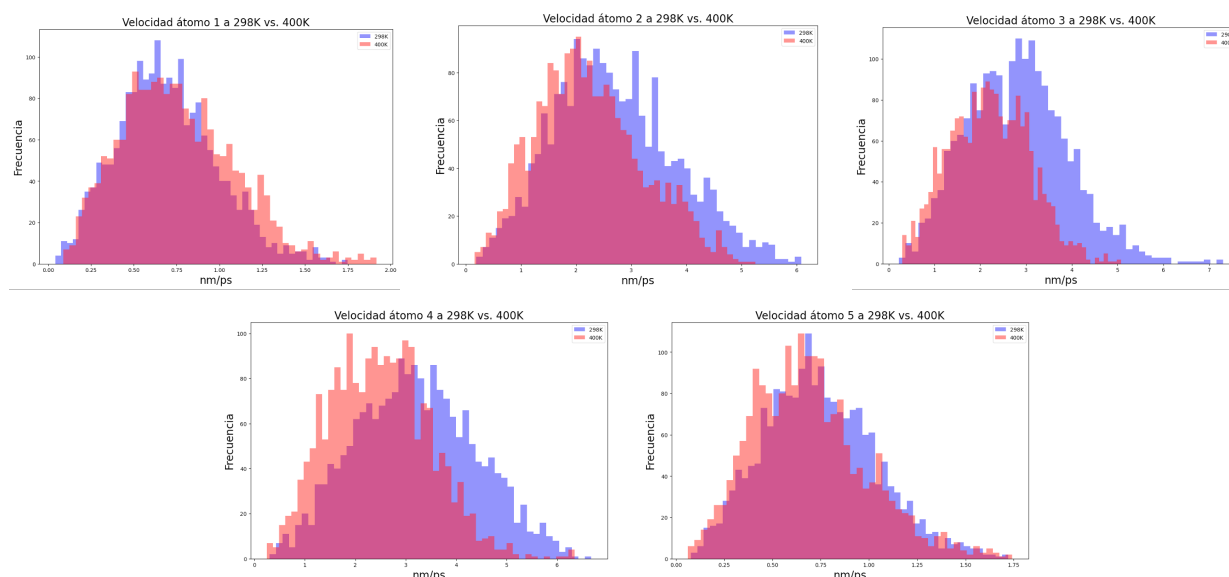


Figura 13: Velocidades de los 5 primeros átomos de AQA.

en la velocidad atómica con la temperatura, debido a que la energía cinética es proporcional a la temperatura. Sin embargo, este fenómeno aparentemente contra-intuitivo puede ser explicado por la naturaleza de las simulaciones de dinámica molecular, que están sujetas a fluctuaciones estadísticas, y que pueden ser más pronunciadas en sistemas de menor tamaño o durante simulaciones de corta duración. Por tanto, lo que observamos podría no representar una tendencia definitiva, sino más bien una variación momentánea específica de nuestro conjunto de datos.

Otra explicación plausible reside en las peculiaridades del sistema molecular en estudio. Es posible que a temperaturas más altas, las interacciones intramoleculares se intensifiquen de tal manera que restrinjan el movimiento atómico, resultando en una disminución de la velocidad media. Esto podría reflejar un fenómeno real donde el calor adicional favorece la formación de interacciones no covalentes, como puentes de hidrógeno o interacciones de van der Waals, que limitan la libertad de los átomos para moverse a pesar de la energía cinética incrementada.

ANÁLISIS EXTENDIDO

Para continuar con la etapa opcional del experimento, se requerirá la creación de un modelo adicional que incorpore un periodo de equilibrado más extenso. Comenzaremos estableciendo una nueva carpeta denominada "5-run_optional". Este paso garantizará que los datos y archivos relacionados con esta nueva fase del experimento se mantengan organizados y separados de los anteriores.

El siguiente paso es realizar cambios en el archivo de configuración **runNVT.mdp**, que contiene los parámetros para la simulación de dinámica molecular en el ensamblado NVT (número de partículas, volumen y temperatura constantes). Es necesario ajustar los parámetros en este archivo para extender el tiempo de equilibrio, asegurándose así de que el sistema alcance un estado termodinámico aún más estable antes de proceder con cualquier análisis.

```
...
nsteps = 1000000      ; 500 ps
nstxout = 20          ; Save coordinates every 10 fs
nstvout = 20          ; Save velocities every 10 fs
nstlog = 20           ; Update log every 10 fs
nstenergy = 20        ; Save energies every 10 fs
...
```

Cuando tengamos el archivo, ejecutamos el siguiente código para generar el fichero .tpr:

- `gmx grompp -f runNVT.mdp -c aqa_298.g96 -p aqa.top -o aqa_298k.tpr`

donde -f el archivo con los parámetros; -c archivo .g96; -p el archico con la topología; -o fichero de salida.

Se generan dos archivos **aqa_298k.tpr** y **aqa_400k.tpr**

El procedimiento siguiente implica el equilibrio térmico del modelo. Para efectuar este proceso, es necesario crear un script de Shell que posteriormente enviaremos a la cola de trabajos 'ecq-k'. El script de Shell necesario para ejecutar la simulación con el modelo es:

```
#!/bin/bash
#
#SBATCH -p eck-q
#SBATCH --chdir=/home/alumno14/modelado_molecular/tarea/5-run_optional
#SBATCH -J equilibrado_run
#SBATCH --cpus-per-task=1
```

date

```
gmx mdrun -s aqa_298k.tpr -mp aqa.top -c aqa_298K.g96 -o aqa_298k.trr ...
```

```
...-e aqa_298k.edr -nt 1
```

date

Donde -s es el fichero .tpr; -mp el fichero con la topología; -c el archivo .g96; -o el fichero de salida con las trayectorias; -e el fichero de salida con las energías.

Finalmente después de lanzar a la cola ecq-k tendremos 2 ficheros nuevos: **aqa_298k.trr** y **aqa_298k.edr**.

4.1 TEMPERATURA

Para calcular la temperatura de nuestro modelo, debemos ejecutar el siguiente código.

- `gmx traj -f aqa_298k.trr -s aqa_298k.tpr -xvg none -ot temp_aqa_298k.xvg`

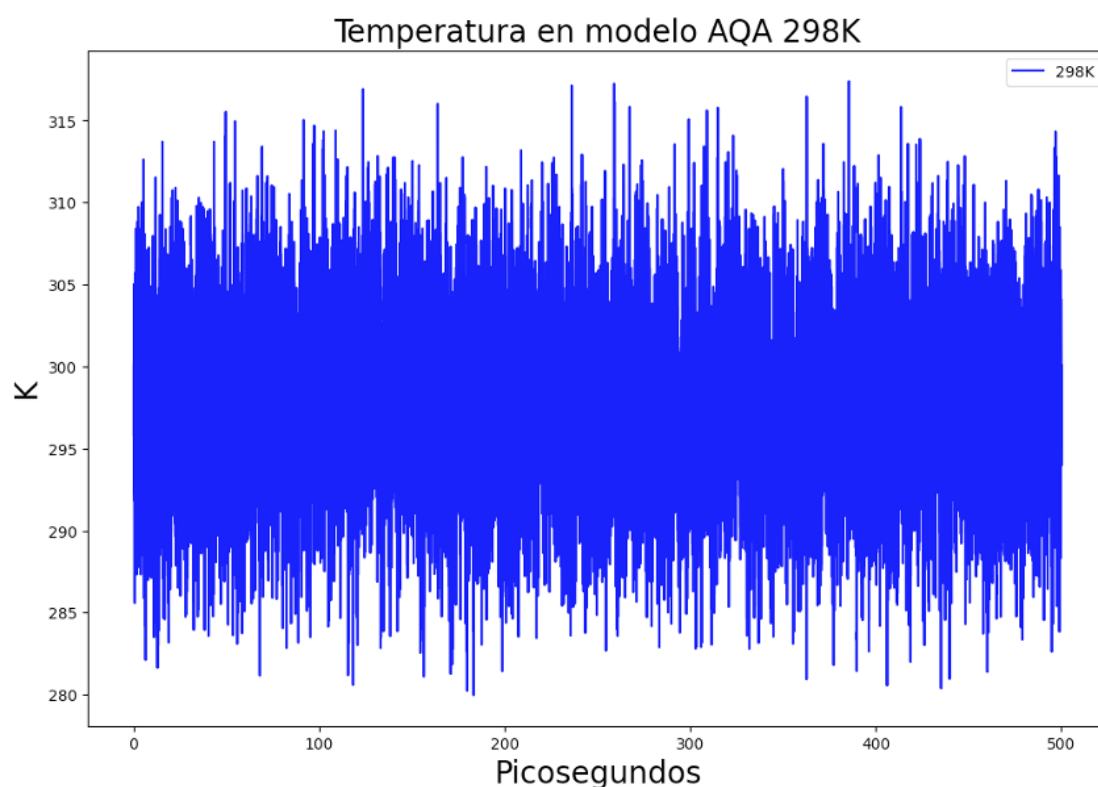


Figura 14: Temperatura del modelo AQA.

El análisis de la simulación (**Figura 14**) indica una estabilidad notable a lo largo de todo el proceso, con una media de temperatura que se mantiene alrededor de los 298K. La desviación observada, que oscila entre 280 y 316 Kelvin, entra dentro de los rangos normales esperados para este tipo de simulaciones y es indicativa

de la dinámica energética inherente al sistema. Estas fluctuaciones son una manifestación de la transferencia de energía constante entre los átomos y el termostato que regula la temperatura del sistema de simulación, lo cual es un componente crítico para mantener las condiciones específicas.

4.2 VELOCIDAD DE UN ÁTOMO EN EL ESPACIO

En una simulación de dinámica molecular, las velocidades de un átomo en distintos ejes (X, Y, Z) pueden proporcionar información sobre varios aspectos del comportamiento y la dinámica molecular del sistema. Las velocidades en distintos ejes pueden mostrar preferencias direccionales en el movimiento de los átomos. Por ejemplo, si un átomo se mueve principalmente a lo largo de un eje particular, esto puede indicar un alineamiento estructural o restricciones direccionales debido a enlaces químicos o interacciones con otros átomos o moléculas. Si las velocidades son significativamente diferentes entre los ejes, podría indicar anisotropía en las interacciones moleculares, como fuerzas intermoleculares que varían con la dirección, restricciones geométricas del sistema o la presencia de campos externos direccionales. Las diferencias en las velocidades a lo largo de los distintos ejes pueden estar relacionadas con la flexibilidad de la molécula. Las regiones flexibles podrían mostrar movimientos más amplios y velocidades más altas. En sistemas complejos, como membranas o interfaces, las velocidades en distintos ejes pueden verse influenciadas por el entorno. Por ejemplo, en una membrana, los átomos pueden moverse más libremente en el plano de la membrana que en la dirección perpendicular. La observación de las velocidades a lo largo del tiempo puede ayudar a identificar y caracterizar procesos dinámicos como vibraciones, rotaciones, y traslaciones de moléculas o dominios moleculares. Las velocidades de átomos en enzimas o sitios activos de proteínas pueden estar influenciadas por la presencia de sustratos o inhibidores y por la dinámica del proceso catalítico.

Para extraer las velocidades los átomos debemos ejecutar el código siguiente:

- `gmx traj -f aqa_298k.trr -s aqa_298k.tpr -xvg none -ov vel_aqa_298k -len`

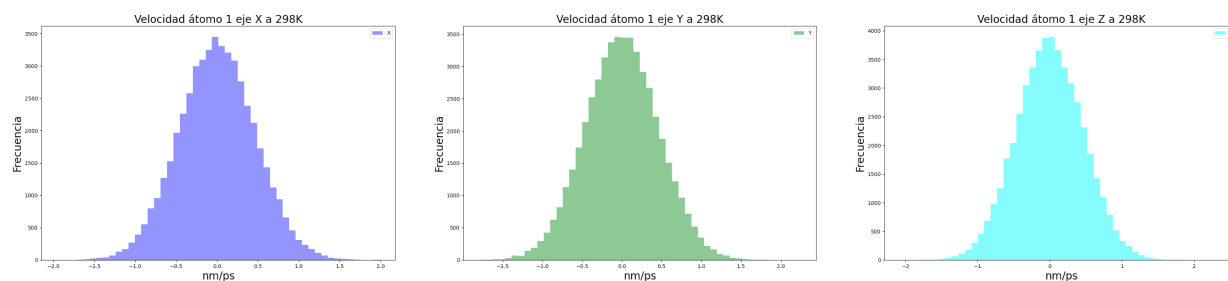


Figura 15: Velocidades en los 3 ejes del primer átomo de AQA.

Aquí mostramos la velocidad del primer átomo del tripeptido AQA en los tres ejes (**Figura 15**). Vemos que el átomo se mueve en los 3 ejes por igual en el rango entre -1,5 y 1,5.

4.3 PLOT DE RAMACHANDRAN

Los gráficos de Ramachandran son cruciales en la bioquímica y biología estructural, ya que proporcionan un medio visual para evaluar las conformaciones de las cadenas polipeptídicas en las proteínas y los péptidos. Permiten identificar las conformaciones más estables que los residuos de aminoácidos pueden adoptar en una proteína, basándose en los ángulos torsionales ϕ (phi) y ψ (psi). Estos ángulos representan las rotaciones alrededor de los enlaces N-C α y C α -C' (carbonyl carbon), respectivamente.

Estos gráficos son especialmente informativos al señalar las regiones del espacio de torsión ϕ - ψ que están asociadas con estructuras de hélice α , láminas β y giros, que son las estructuras secundarias fundamentales para la conformación en 3D de las proteínas. Además, son herramientas valiosas para investigar el efecto de mutaciones o modificaciones químicas en la estabilidad de las proteínas, permitiendo observar cómo los cambios en la secuencia de aminoácidos impactan la conformación estructural.

Además, los gráficos de Ramachandran son utilizados para validar modelos estructurales de proteínas obtenidos por técnicas como la cristalografía de rayos X, la resonancia magnética nuclear (RMN) o el modelado por homología, verificando que los ángulos torsionales caigan dentro de las regiones permitidas y favorecidas.

Para obtener los ángulos diedros de los tres residuos y analizar su conformación en el espacio de Ramachandran, aplicamos un conjunto de comandos específicos en nuestro software de dinámica molecular, que nos permiten realizar el análisis y la visualización pertinentes.

- `gmx rama -f aqa_298k.trr -s aqa_298k.tpr -o rama_aqa_298k.svg -svg none`

Luego escogemos el ϕ (phi) y ψ (psi) de cada residuo y lo guardamos en un fichero.

- `grep ALA-2 rama_aqa_298k.svg | awk 'print $1' | cat -n >phi-ala-2_298k.dat`
- `grep ALA-2 rama_aqa_298k.svg | awk 'print $2' | cat -n >psi-ala-2_298k.dat`
- `join phi-ala-2_298k.dat psi-ala-2_298k.dat >ala-2_298k.dat`
- `grep GLN-3 rama_aqa_298k.svg | awk 'print $1' | cat -n >phi-gln-3_298k.dat`
- `grep GLN-3 rama_aqa_298k.svg | awk 'print $2' | cat -n >psi-gln-3_298k.dat`
- `join phi-gln-3_298k.dat psi-gln-3_298k.dat >gln-3_298k.dat`
- `grep ALA-4 rama_aqa_298k.svg | awk 'print $1' | cat -n >phi-ala-4_298k.dat`
- `grep ALA-4 rama_aqa_298k.svg | awk 'print $2' | cat -n >psi-ala-4_298k.dat`
- `join phi-ala-4_298k.dat psi-ala-4_298k.dat >ala-4_298k.dat`
- `join phi-ala-2_298k.dat psi-ala-2_298k.dat >ala-2_298k.dat`
- `join phi-ala-2_400k.dat psi-ala-2_400k.dat >ala-2_400k.dat`
- `join phi-gln-3_298k.dat psi-gln-3_298k.dat >gln-3_298k.dat`
- `join phi-gln-3_400k.dat psi-gln-3_400k.dat >gln-3_400k.dat`

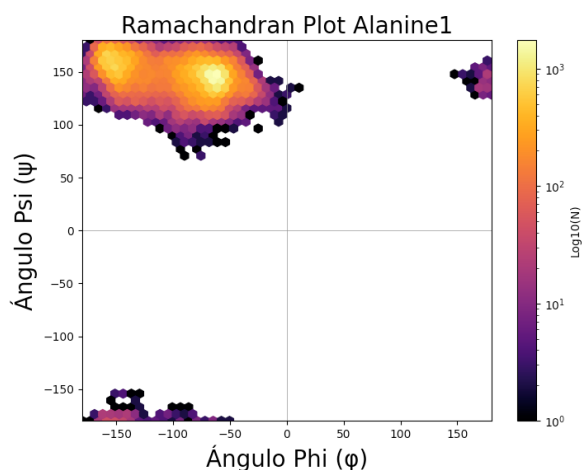


Figura 16: Plot ramachandran del primer residuo de Alanina.

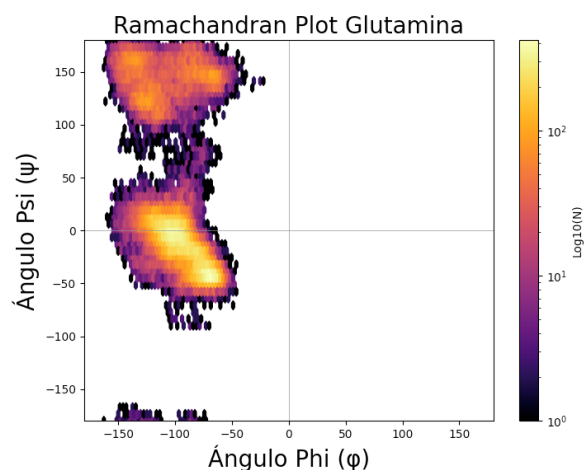


Figura 17: Plot ramachandran del residuo de Glutamina.

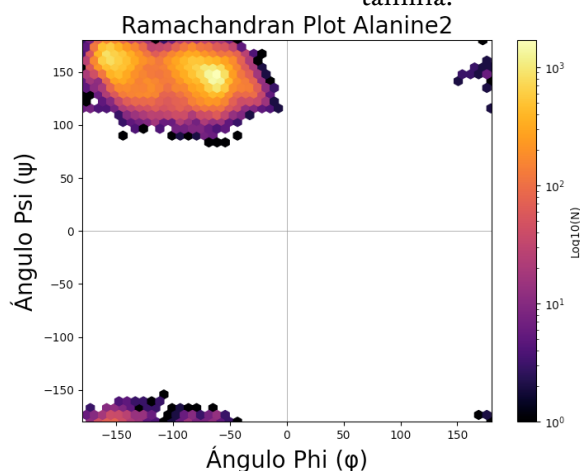


Figura 18: Plot ramachandran del segundo residuo de Alanina.

Los gráficos de Ramachandran para las dos Alaninas (**Figura 16**, **Figura 18**) muestra una acumulación de puntos, principalmente en la zona superior izquierda, lo que indica que estas conformaciones son energéticamente favorables para los residuos de alanina. Se destaca una predisposición por las conformaciones de láminas β extendidas, caracterizadas por un ángulo ϕ negativo y un ángulo ψ positivo, sugiriendo que la alanina, en su primera posición, tiende a esta configuración.

Para la Glutamina (**Figura 17**), el gráfico muestra una alta densidad de puntos en un área que se extiende a lo largo de valores negativos de ϕ , lo cual es típico para conformaciones de α hélice.

La barra de colores que acompaña los gráficos representa la cantidad de conformaciones en cada área, con las regiones más cálidas indicando una mayor frecuencia y las más frías una menor frecuencia de conformaciones. Estos gráficos indican que el tripéptido en cuestión muestra una preferencia significativa por adoptar conformaciones de láminas β por parte de las Alaninas y α hélice por parte de la Glutamina.