

# Practical Course on Molecular Dynamics and Trajectory Analysis

## Episode 3: Preparing and editing the system

Jordi Villà i Freixa

Universitat de Vic - Universitat Central de Catalunya  
Facultat de Ciències, Tecnologia i Enginyeries (FCTE)

*jordi.villa@uvic.cat*

MD Course and Trajectory Analysis  
Concepcion, January 2026

# Contents

## 1 Episodio 3: Preparación y edición del sistema

- Objetivos y flujo
- Modelado guiado
- Calidad de la estructura
- Reparación y completado
- Campos de fuerza
- Ligandos y cargas
- Solvatación
- Iones y electrostática
- Membranas
- Minimizado
- Validación
- Exportación

# Objetivos del episodio

- Preparar sistemas listos para simular con OpenMM.
- Estandarizar topologías y coordenadas.
- Generar entradas reproducibles y verificables.

# Flujo de preparación

- ① Lectura de estructura (PDB/MOL2/SDF).
- ② Reparación y completado de residuos.
- ③ Protonación y asignación de cargas.
- ④ Solvatación, iones y caja.
- ⑤ Minimizado y validación.

# Archivos de entrada

- PDB: coordenadas y metadatos experimentales.
- MOL2/SDF: ligandos y pequeñas moléculas.
- Topología y parámetros se asignan después de la lectura.

# Identificadores esenciales

Residuo = (cadena, número, nombre),  
Átomo = (tipo, elemento, carga).

- Los identificadores consistentes evitan errores en el campo de fuerza.

# Reproducibilidad

- Guardar scripts de preparación.
- Documentar versiones de librerías y parámetros.
- Evitar pasos manuales no trazables.

# Modelado conforme a la guía oficial

- OpenMM User Guide (§4.1-4.6) sugiere reconectar hidrógenos, solvatación y adición de membranas antes de parametrizar.
- El repositorio de ejemplos incluye `simulateAmber.py` + `simulateCharmm.py`, y `argon-chemical-potential.py` para verificar energías libres en líquidos simples.
- Podemos reutilizar la infraestructura para el dipéptido de alanina y el complejo proteína-ligando, pero extendiendo los scripts con variables de membrana o polímeros coarse-grained.

# Sistemas adicionales

- `coarse_grained_polymer.py` (OpenMM Cookbook) crea topologías bead-spring que muestran cómo definir masas y enlaces manualmente.
- Ese mismo enfoque se conecta con el script `argon-chemical-potential.py`, donde se parametrizan fuerzas de Lennard-Jones y se mide energía libre de inserción.
- Los modelos biomoleculares se benefician al combinar estos ejemplos ligeros con la solvatación y el tratamiento de ligandos pesados en Amber.

# Membranas y solventes

- Añadir solvente con ‘Modeller.addSolvent’ y escoger OPC/TIP3P (User Guide §4.2).
- Si hay membranas, usar ‘Modeller.addMembrane’ y luego `simulateAmber.py` con barostato anisotrópico.
- Guardar la topología final para reproducir exactamente el sistema (OpenMM App §4.6).

# Calidad del PDB

- Resolución y factores B informan sobre la incertidumbre.
- Regiones flexibles suelen presentar huecos o baja ocupación.

# Residuos faltantes

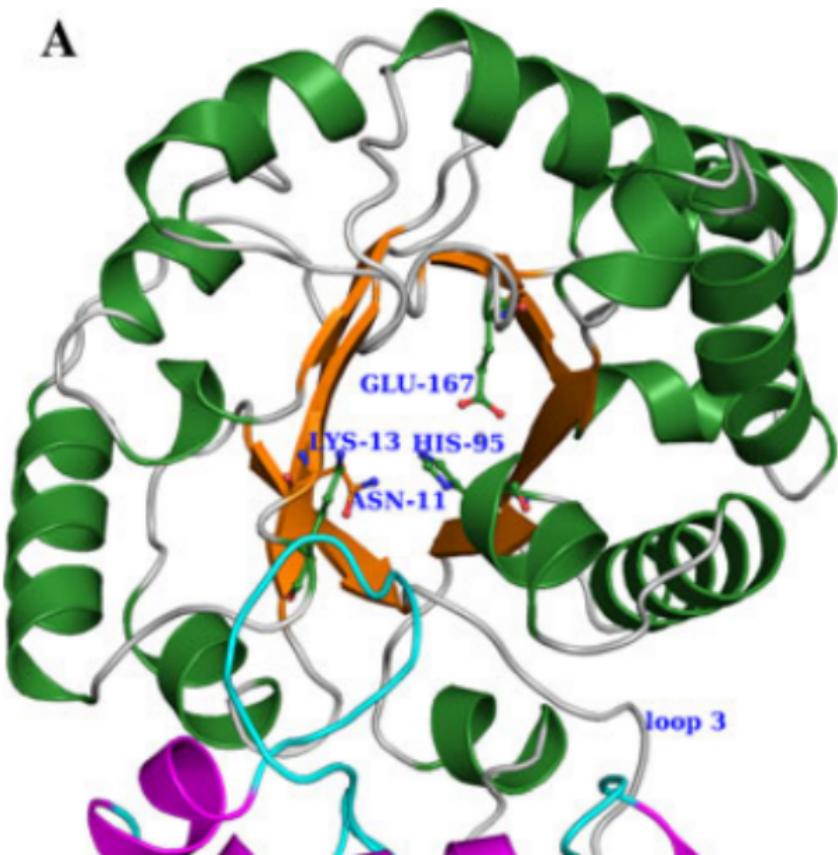
- Los PDB pueden carecer de segmentos completos.
- La reconstrucción requiere inferir geometría y estereoquímica.

# AltLocs y ocupación

$$\sum_k \text{occ}_k \leq 1.$$

- Elegir la conformación dominante o promediar según objetivo.

# Sitio activo en PDB



# Superposición y RMSD

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N \|\mathbf{r}_i - \mathbf{r}_i^{\text{ref}}\|^2}$$

- Verifica coherencia con estructuras de referencia.

# PDBFixer/Modeller

- Rellena residuos, corrige nombres y quita ligandos no deseados.
- Prepara el sistema para asignar campo de fuerza.

# Inserción de residuos

- Se usa información geométrica y estereoquímica.
- Minimizado local para aliviar choques.

# Protonación y pH

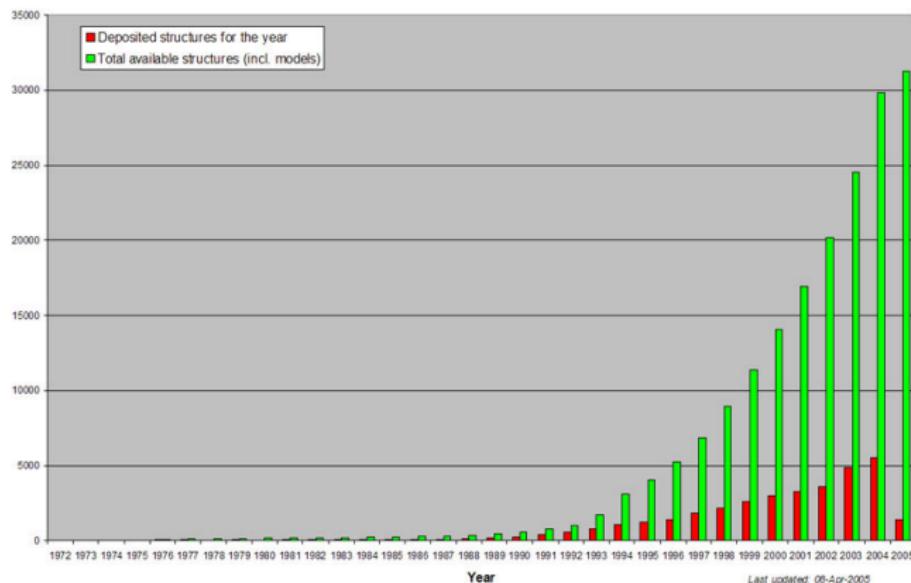
$$\frac{[A^-]}{[HA]} = 10^{pH - pK_a}.$$

- Determina estados de carga de residuos titulables.

# Estados de carga

- Ajustar histidinas (HID/HIE/HIP), ASP/GLU, LYS/ARG.
- Coherencia con el entorno del sitio activo.

# Crecimiento estructural



# Descomposición del potencial

$$U = U_{\text{enlace}} + U_{\text{ángulo}} + U_{\text{dihedro}} + U_{\text{no enlazado}}.$$

- Base del cálculo de fuerzas y energías.

# Enlaces y àngulos

$$U_{\text{enlace}} = \sum_b k_b (r_b - r_b^0)^2,$$

$$U_{\text{àngulo}} = \sum_a k_a (\theta_a - \theta_a^0)^2.$$

- Aproximación armónica alrededor del equilibrio.

# Dihedros

$$U_{\text{dihedro}} = \sum_d \frac{V_d}{2} [1 + \cos(n_d \phi_d - \gamma_d)] .$$

- Controlan barreras rotacionales y conformaciones.

## No enlazados

$$U_{\text{LJ}} = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right],$$
$$U_C = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}}.$$

- LJ y Coulomb dominan interacciones a distancia.

# Selección del campo de fuerza

- Proteínas: AMBER/CHARMM/OPLS.
- Ligandos: OpenFF u otros parametrizadores.
- Validar compatibilidad con el sistema.

# Consistencia de parámetros

- Evitar mezclar campos de fuerza sin reglas claras.
- Revisar unidades y escalas de energía.

# Cargas parciales

$$\sum_i q_i = Q_{\text{total}}.$$

- El total debe coincidir con el estado de protonación.

# Ajuste electrostático

- RESP y AM1-BCC derivan cargas de potenciales electrostáticos.
- La distribución de cargas afecta energías de unión.

## Escaneo de dihedros

$$E(\phi) = E_0 + \sum_k V_k \cos(k\phi - \gamma_k).$$

- Ajusta perfiles rotacionales en ligandos.

# Asignación de tipos

- Tipos atómicos determinan parámetros LJ y enlaces.
- Inconsistencias producen energías no físicas.

# Validación del ligando

- Revisar geometría, quiralidad y cargas netas.
- Comparar con referencias experimentales o QM.

## Elección de caja

- Cúbica, ortorrómbica o dodecaédrica.
- Compromiso entre coste y distancia a la imagen.

## Número de moléculas de agua

$$N_{\text{H}_2\text{O}} \approx \rho V \frac{N_A}{M_{\text{H}_2\text{O}}}.$$

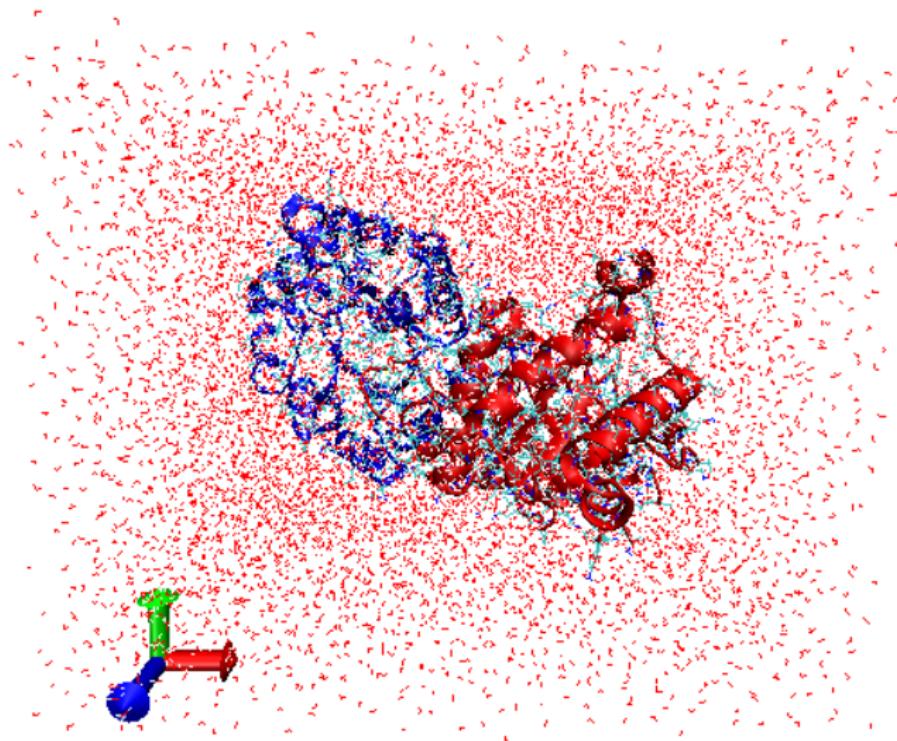
- Estima el tamaño de caja según densidad objetivo.

# Concentración iónica

$$N_{\text{ion}} = C V N_A.$$

- Ajustar molaridad para condiciones fisiológicas.

# Caja solvatada



# Modelos de agua

- TIP3P, SPC/E, TIP4P: distintas densidades y dinámicas.
- Elegir el modelo compatible con el campo de fuerza.

## Recorte de solvente

- Mantener un colchón mínimo alrededor del soluto.
- Evita interacciones artificiales con la imagen periódica.

# Chequeo de densidad

- Verificar densidad tras equilibrado NPT.
- Ajustar tamaño si la densidad se desvía.

# Neutralización

- Neutralizar la carga neta mejora estabilidad numérica.

# Fuerza iónica

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

- Controla apantallamiento electrostático.

# Longitud de Debye

$$\lambda_D = \sqrt{\frac{\varepsilon k_B T}{2N_A e^2 I}}.$$

- Define el rango de interacciones electrostáticas en solución.

## PME y corte

- PME trata Coulomb de largo alcance de forma eficiente.
- Elegir corte consistente con la caja.

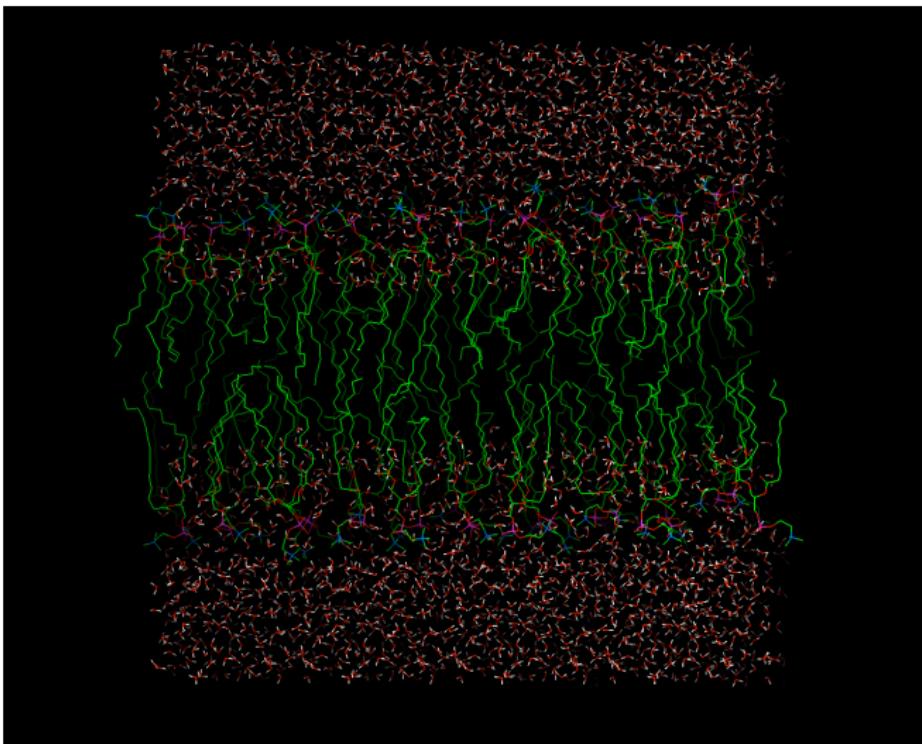
## Balance carga/solvente

- Comprobar que el sistema no esté sobrecargado de iones.

# Modelos de membrana

- Bicapas lipídicas requieren cajas anisotrópicas.
- La orientación debe ser coherente con la proteína.

# Bicapa lipídica



## Area por lípido

$$A_{\text{lip}} = \frac{A_{\text{caja}}}{N_{\text{lip}}/2}.$$

- Métrica clave para estabilidad de la membrana.

# Espesor de membrana

- Controla la exposición de dominios hidrofóbicos.
- Ajuste con minimizado y equilibrado.

# Restricciones iniciales

- Restringir la proteína evita colapsos durante equilibrado.

# Función objetivo

$$\min_{\mathbf{r}} U(\mathbf{r}).$$

- Reduce choques y mejora la estabilidad inicial.

# Descenso del gradiente

$$\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k \nabla U(\mathbf{r}_k).$$

- Método robusto para minimizar energías altas.

# Gradiente conjugado

- Más eficiente cerca del mínimo.
- Reduce el número de evaluaciones de fuerza.

# Restricciones armónicas

$$U_{\text{rest}} = \frac{k}{2} \sum_i \|\mathbf{r}_i - \mathbf{r}_i^0\|^2.$$

- Mantienen la estructura mientras se relaja el solvente.

# Criterios de convergencia

- Normas del gradiente y disminución de energía.
- Detener cuando las fuerzas son pequeñas.

# Chequeo geométrico

- Longitudes y àngulos razonables.
- Ausencia de choques estéricos.

# Energia por componente

- Revisar contribuciones LJ, Coulomb y enlaces.
- Identificar valores extremos.

## Comparación con referencia

$$\Delta U = U_{\text{nuevo}} - U_{\text{ref.}}$$

- Diferencias grandes indican problemas de parametrización.

# Chequeo de estabilidad

- Ejecutar un NVT corto para detectar explosiones numéricas.
- Inspeccionar energía potencial y temperatura.

# Solvente estable

- Verificar que no hay vacíos ni solapamientos.

# Lista de control

- PDB limpio, protonación correcta, sistema neutralizado.
- Archivo de topología guardado.

# Guardar el sistema

- Exportar PDB y topología con parámetros.
- Guardar estado de minimizado.

# Formatos de salida

- PDB, Amber prmtop/inpcrd, OpenMM XML.

# Trazabilidad

- Registrar scripts y configuraciones usadas.
- Versionar los parámetros.

# Semillas y reproducibilidad

- Guardar semillas de generadores aleatorios.
- Documentar condiciones iniciales.

# Estructura de carpetas

- Separar entradas, resultados y archivos temporales.
- Facilita auditoria y reutilización del sistema.

## Resumen del episodio

- Preparar el sistema es clave para resultados fiables.
- Las decisiones en protonación y solvatación impactan la dinámica.
- Una buena validación evita errores costosos.

# References I