

Algoritmo para Simplificar Trayectorias de Plegamiento de Proteínas

Luis Garreta, Mauricio Martinez, Pedro A. Moreno

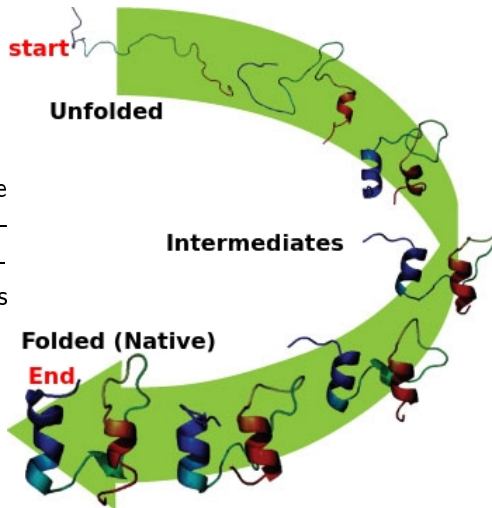
Grupo de Investigación en Bioinformática
Universidad del Valle
Cali, Colombia

8 de noviembre de 2019

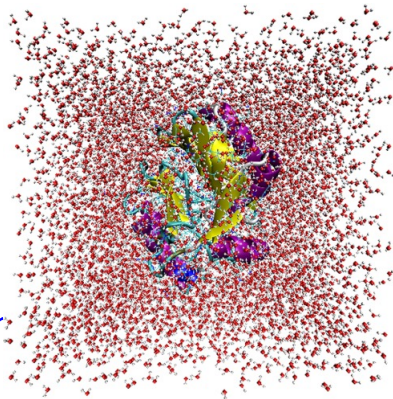


- 1 Plegamiento de Proteínas
- 2 Trayectorias de Plegamiento
- 3 Reducción de Trayectorias
- 4 Algoritmo Propuesto
- 5 Resultados

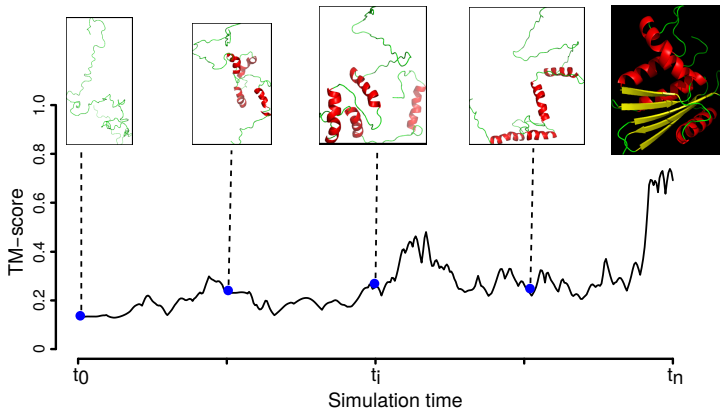
Las proteínas para ser funcionales se pliegan desde un estado inicial o *desplegado* hasta alcanzar su estado plegado o *nativo*, pasando por distintos *estados intermedios*.



- Las simulaciones del plegamiento de proteínas son una de las *principales herramientas* para entender este proceso.
- La técnica más usada para estas simulaciones es la *Dinámica Molecular*.



Estas simulaciones producen *trayectorias de plegamiento* que describen los distintos eventos o estructuras 3D (*conformaciones*) que la proteína adquiere en su camino hacia su estado nativo *en función del tiempo*.



Avances en software y hardware han permitido simulaciones más largas y para más proteínas



Computación Distribuida
(e.g. Folding@home)

Simulaciones en el
orden de los
microsegundos



Supercomputadores
(e.g. Anton, Anton2)

Simulaciones en el
orden de los
milisegundos



Arreglos de GPUs
(e.g. Nvidia)

Simulaciones en el
orden de los
microsegundos

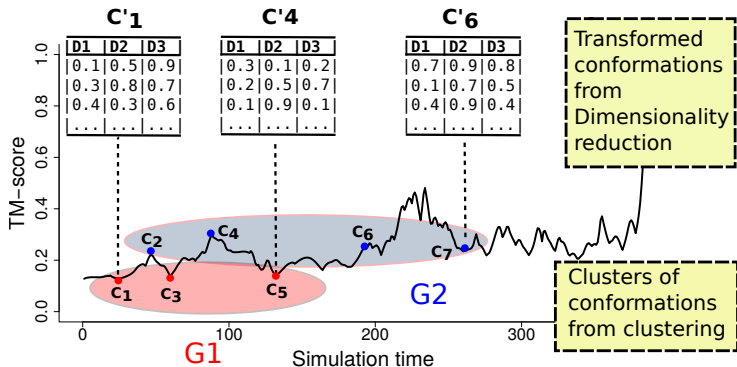
- Diferentes tecnologías (supercomputadores, GPUs, Clusters)
- Tiempos de simulación ya llegan al orden de los milisegundos.
- *Trayectorias cada vez más frecuentes y mucho más extensas*



Problema

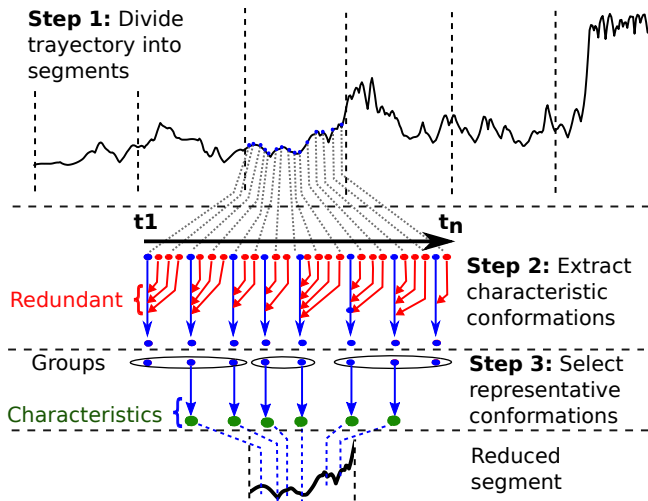
El procesamiento y análisis se vuelve complejo por el gran número de conformaciones y se necesita **reducirlas**!

- **Reducción de dimensionalidad:** PCA, nMDS, Isomap, diffusion maps
- **Agrupamientos:** aglomerativos como kmeans, jerárquicos como single-linkage



Características Algoritmo Propuesto para Reducción de Trayectorias de Plegamiento

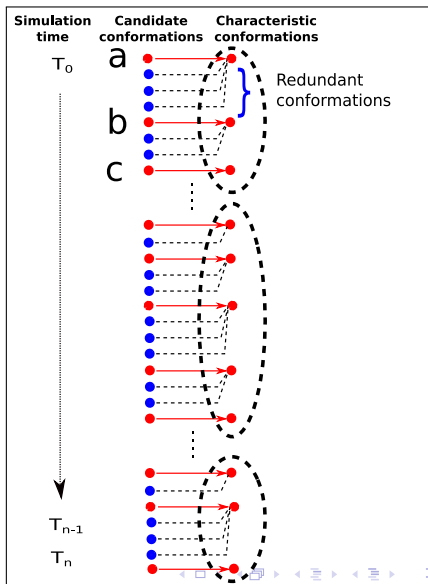
- Sea rápido:
 - ▶ Reduzca eficientemente trayectorias largas
- Preserve los eventos principales:
 - ▶ Busque conformaciones disimilares
- No pierda la información estructural:
 - ▶ Genere conformaciones de proteínas
- No pierda la información temporal:
 - ▶ Genere trayectorias reducidas
- Sea paralelizable:
 - ▶ Reduzca localmente, utilice los procesadores multinucleó

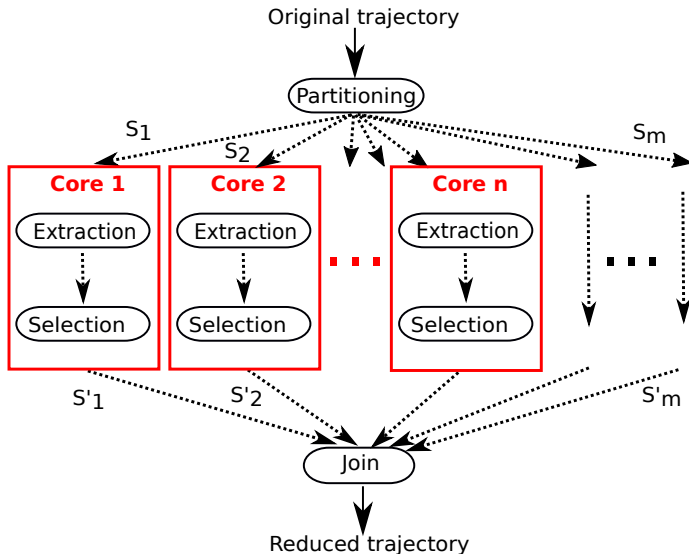


Reducción Rápida Local (Step 2)

Basada en el algoritmo de agrupamiento rápido de Hobohm&Sander¹

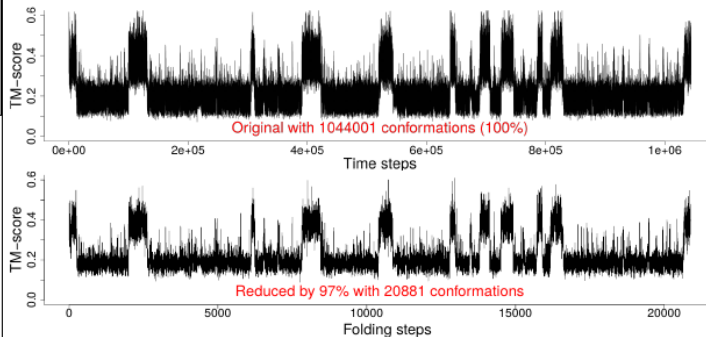
¹ U. Hobohm et al., Selection of representative protein data sets. Protein Sci. 1992 1: 409-417





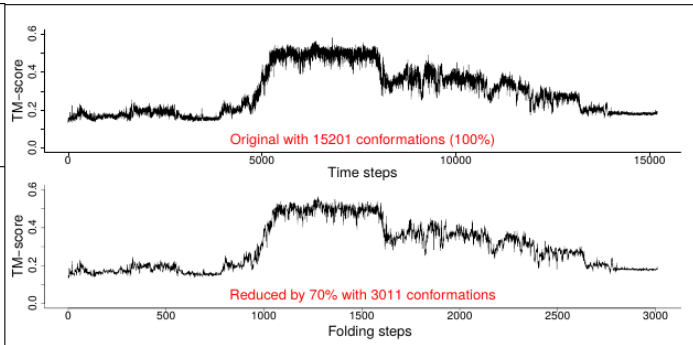
Reducción de la trayectoria de la Proteína Trp-cage

**A: Trp-cage
(PDB 2JOF)**



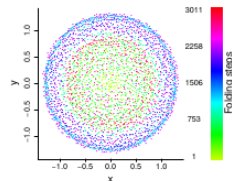
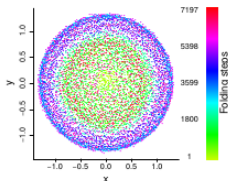
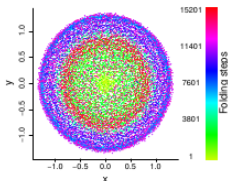
Reducción de la trayectoria de la Proteína Villin Head

**B: Villin headpiece
(PDB 2F4K)**

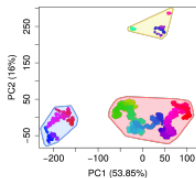


Comparación con otros métodos

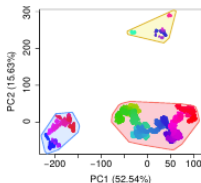
Patterns for
nMDS
reductions



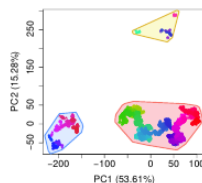
Structure for
clustering
reductions



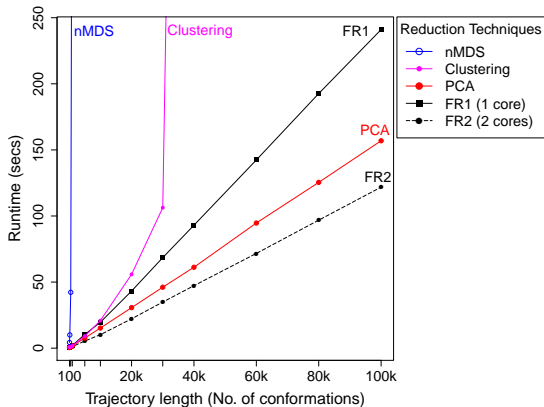
Original Trajectory

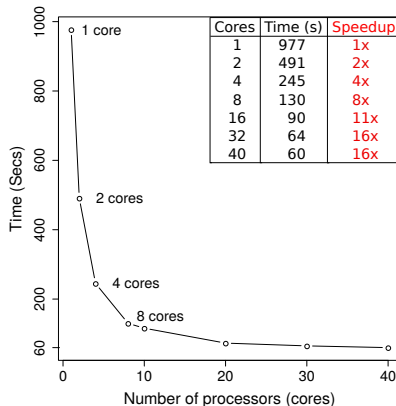


Trajectory reduced by 52%



Trajectory reduced by 80%





Gracias!

Preguntas?