

Máster Universitario en Bioinformática

Proteómica y Bioinformática Estructural

Curso académico 2024-2025



Universidad
Internacional
de Valencia

Dra. Magdalena Nikolaeva Koleva

magdalena.nikolaeva@professor.universidadviu.com

10/10/2024

De:



Planeta Formación y Universidades

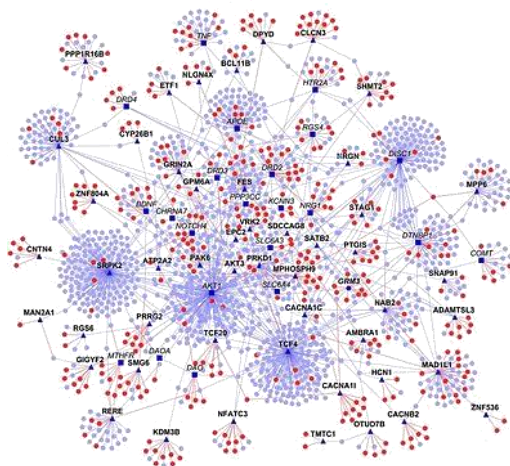
Sesión 8

Docking entre proteínas (I): *Docking* de cuerpo rígido



Universidad
Internacional
de Valencia

¿Qué vimos en la sesión anterior?



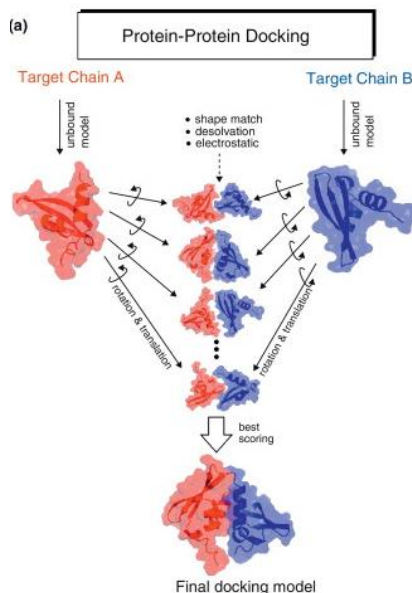
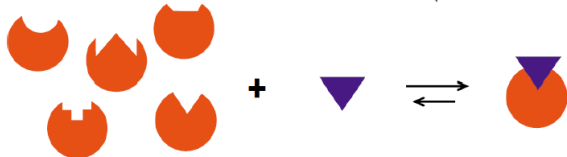
1894: LLAVE Y CERRADURA (E. FISCHER)



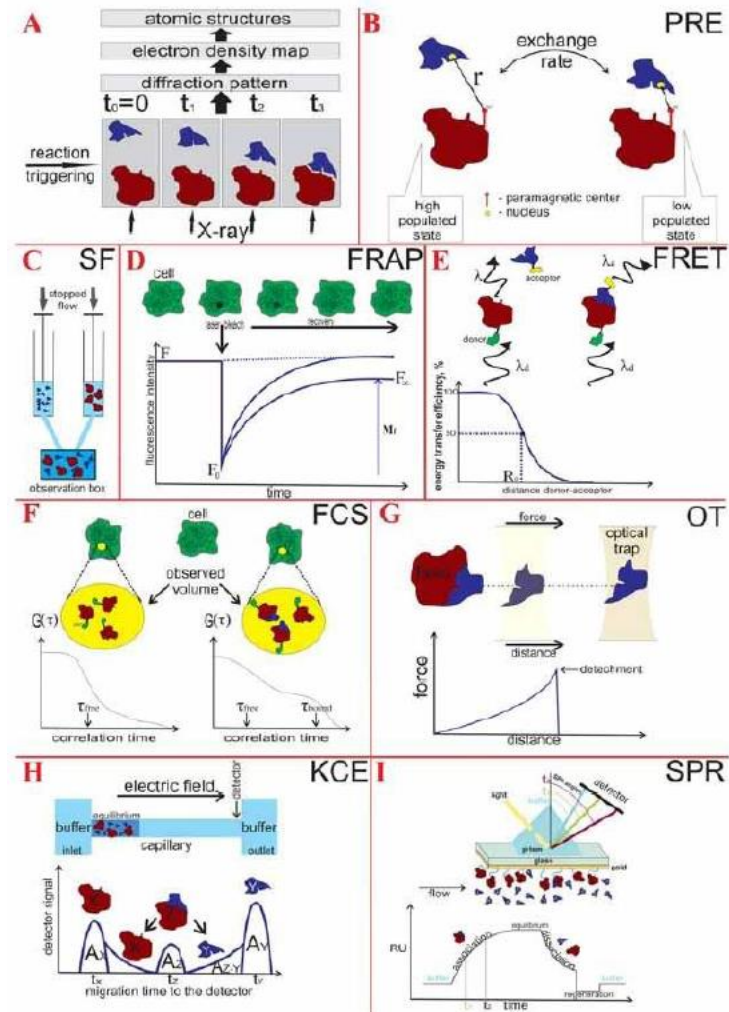
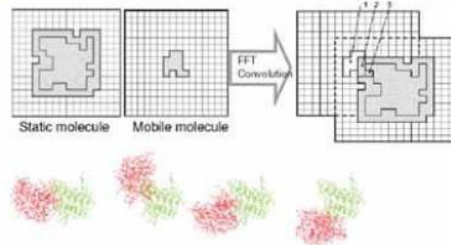
1958: AJUSTE INDUCIDO (D. E. KOSHLAND)



1999: SELECCION CONFORMACIONAL (R. NUSSINOV)



Búsqueda exhaustiva (FFT, superficies)



Temario - Contenidos

Tema 6. Modelización de interacciones por *docking*

6.1. Búsqueda de orientaciones entre proteínas

6.2. Puntuación de modelos de *docking*

6.3. Flexibilidad conformacional en *docking*

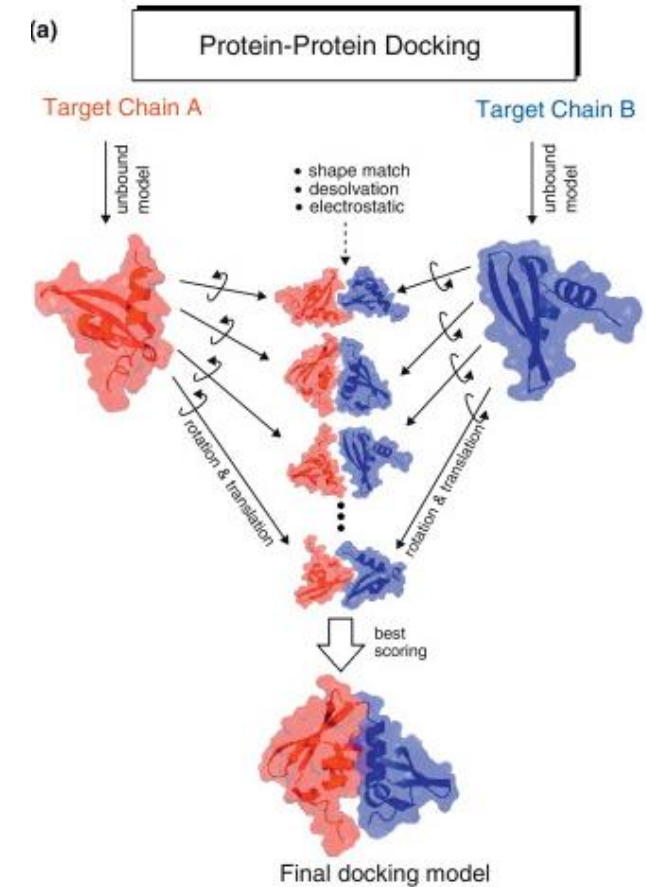
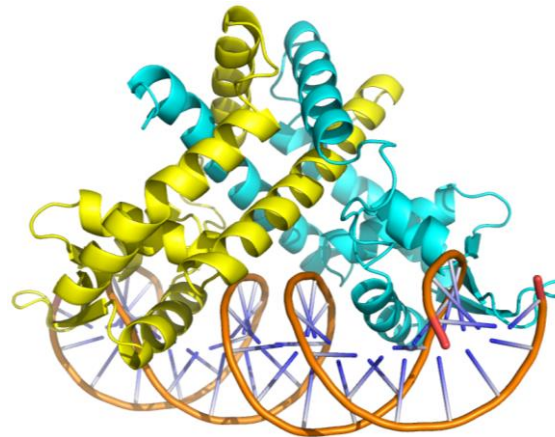
6.4. Evaluación de las predicciones de *docking*

Docking

Se trata de un método que predice la posición y orientación que adoptan dos moléculas al unirse entre sí para formar un complejo estable.

Los componentes del complejo pueden ser:

- **2 proteínas** (presente tema) – complejos proteicos
- Proteína y un ligando pequeño (08MBIF)
- Proteína y ácido nucleico (ADN/ARN)



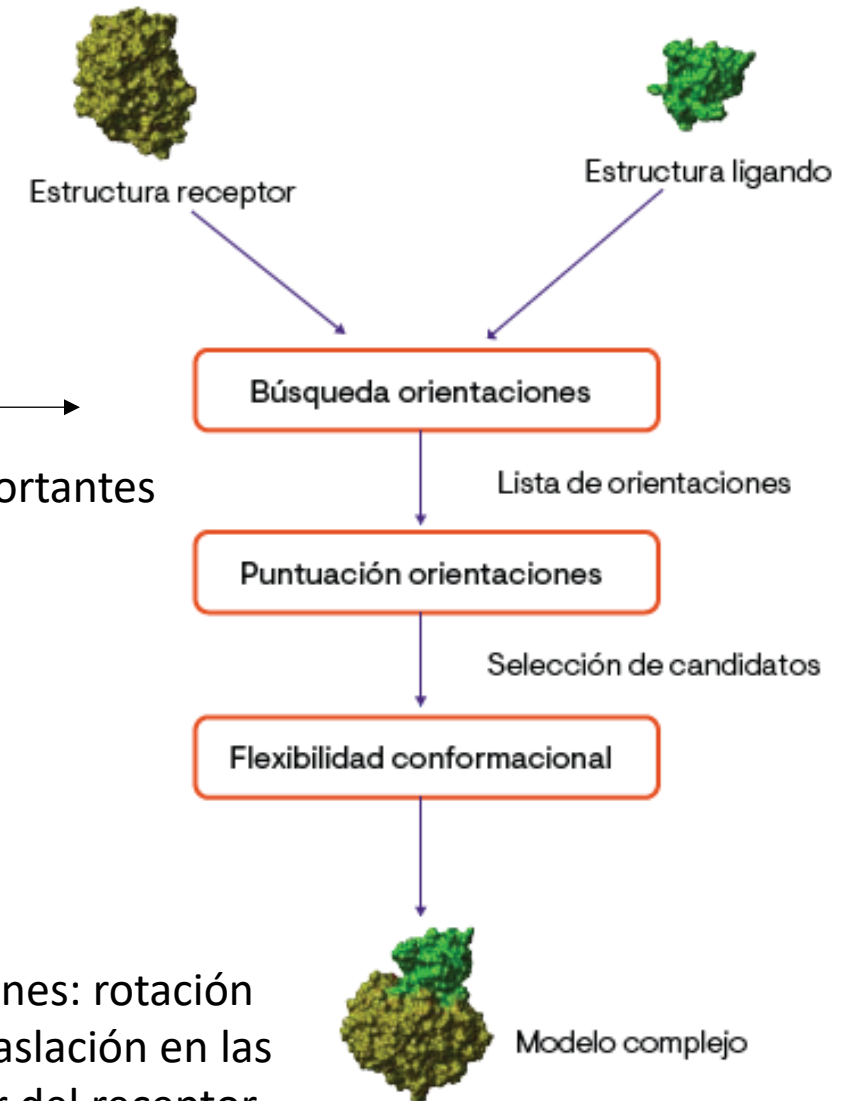
Docking

Se trata de un método que predice la posición y orientación que adoptan dos moléculas al unirse entre sí para formar un complejo estable.

Los componentes del complejo pueden ser:

- **2 proteínas** (presente tema) – complejos proteicos
- Proteína y un ligando pequeño (08MBIF)
- Proteína y ácido nucleico (ADN/ARN)

Aspectos importantes



Problema de 6 dimensiones: rotación en las 3 dimensiones y traslación en las 3 dimensiones alrededor del receptor

Temario - Contenidos

Tema 6. Modelización de interacciones por *docking*

6.1. Búsqueda de orientaciones entre proteínas

6.2. Puntuación de modelos de *docking*

6.3. Flexibilidad conformacional en *docking*

6.4. Evaluación de las predicciones de *docking*

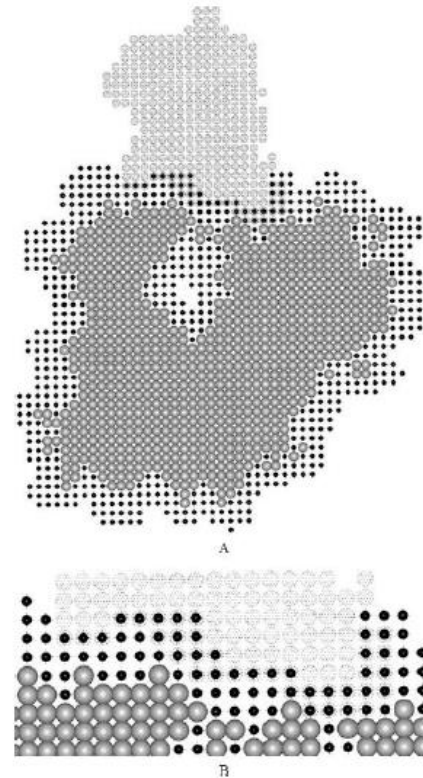
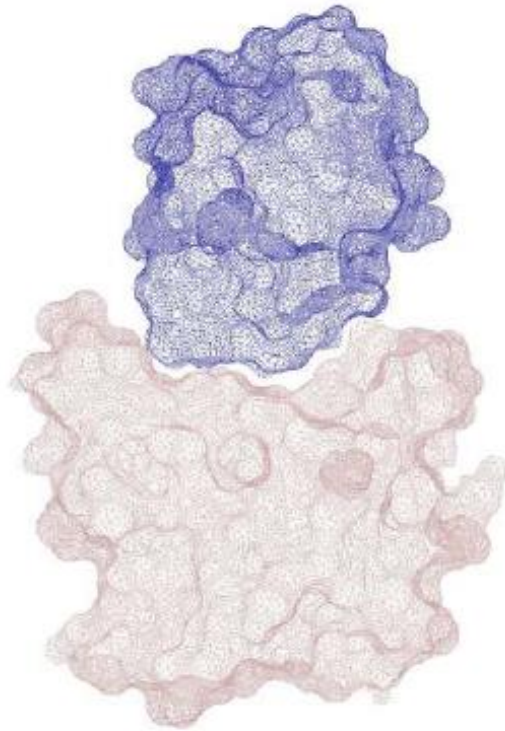
Docking de cuerpo rígido

- Muestreo exhaustivo: métodos FFT
- Muestreo exhaustivo: fragmentación geométrica
- Muestreo estocástico: basado en energías

6.1. Búsqueda de orientaciones entre proteínas

Docking de cuerpo rígido

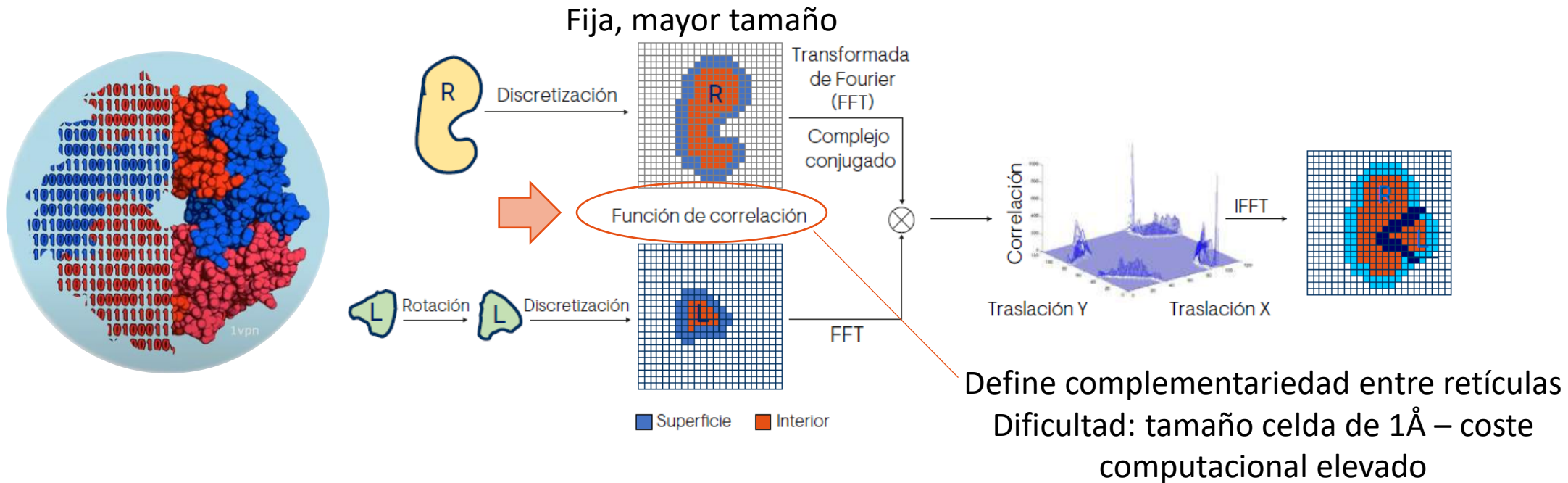
- **Muestreo exhaustivo o búsqueda sistemática: métodos FFT**
 - Encaje geométrico: las superficies de interacción entre dos proteínas encajan a la perfección
 - Explorar todas las posibles orientaciones y evaluar complementariedad



6.1. Búsqueda de orientaciones entre proteínas

Docking de cuerpo rígido

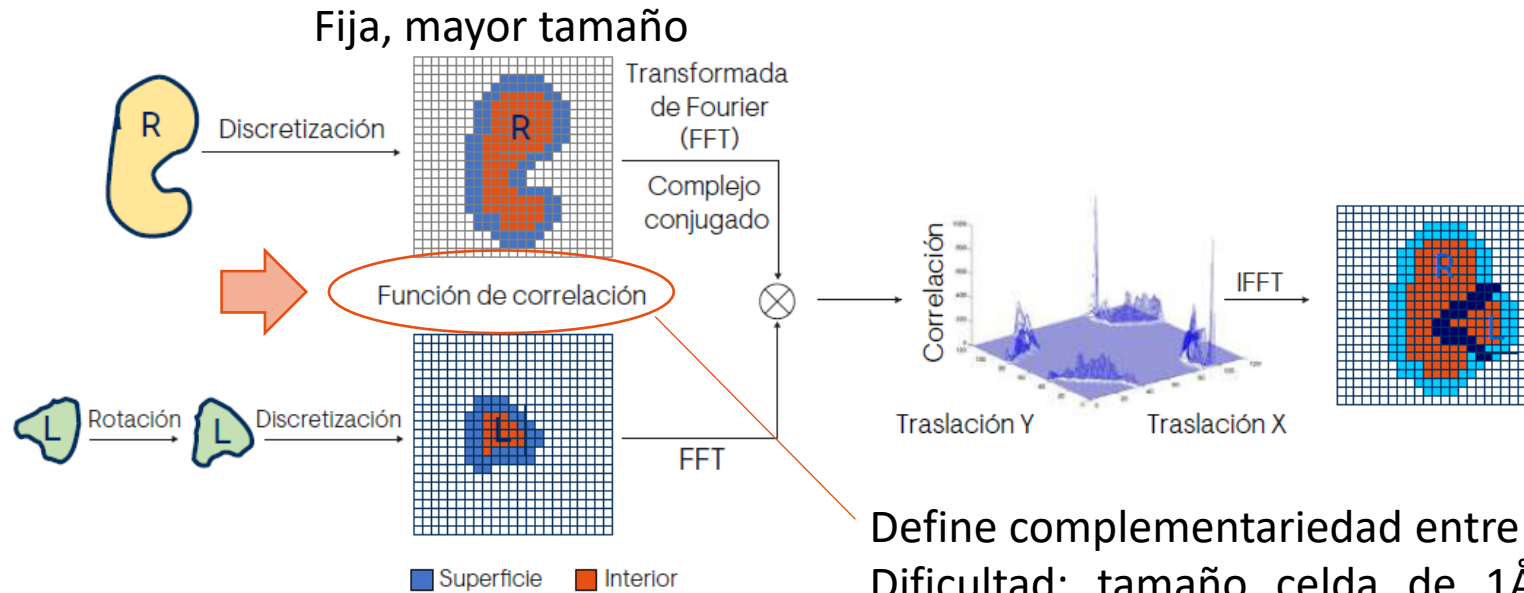
- **Muestreo exhaustivo o búsqueda sistemática: métodos FFT**
 - Encaje geométrico: las superficies de interacción entre dos proteínas encajan a la perfección
 - Explorar todas las posibles orientaciones y evaluar complementariedad
 - Descripción explícita (ALTO COSTE COMPUTACIONAL)
 - Simplificación mediante retícula tridimensional (0/1) – explorar millones de orientaciones



6.1. Búsqueda de orientaciones entre proteínas

Docking de cuerpo rígido

- Muestreo exhaustivo o búsqueda sistemática: **métodos FFT**



Define complementariedad entre retículas
Dificultad: tamaño celda de 1Å – coste computacional elevado.

- FFT – algoritmo que disminuye el tiempo de cálculo de n^2 a $n \cdot \log_2(n)$
- Hasta 10000 veces más rápido

6.1. Búsqueda de orientaciones entre proteínas

Docking de cuerpo rígido

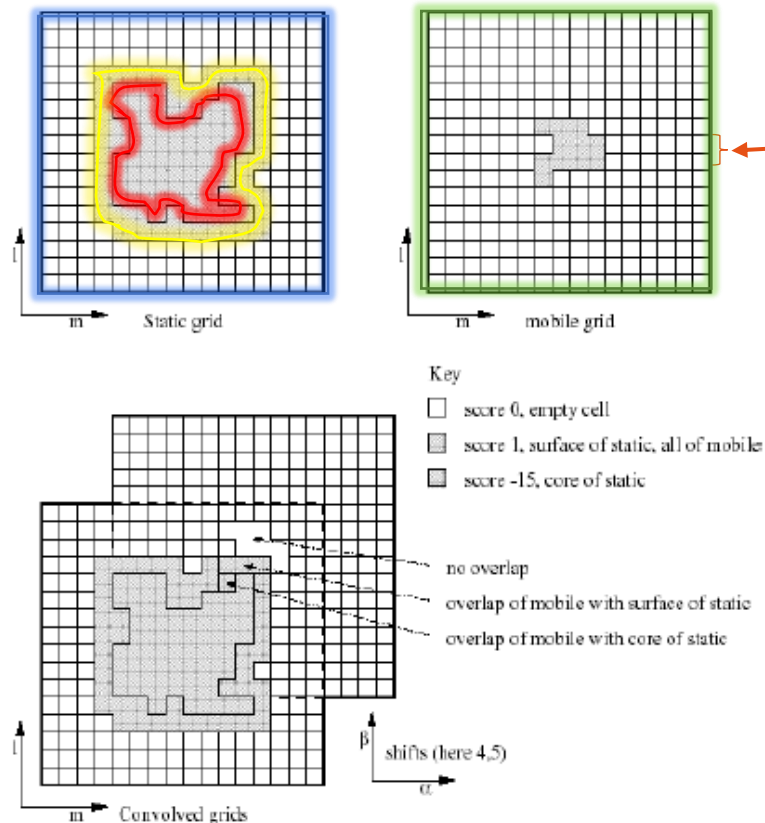
- **Muestreo exhaustivo o búsqueda sistemática: métodos FFT**
 - Programas:
 - **FTDOCK**: hasta 10000 modelos. Función de evaluación geométrica (posible incorporar electrostática).
 - **ZDOCK**: consideraciones geométricas + funciones energéticas (físicas/empíricas)
 - **GRAMM-X**: funciones de puntuación empíricas
 - Desventajas:
 - Difícil introducir flexibilidad conformacional (retícula 3D).
 - No apropiado para proteínas que sufren grandes cambios conformacionales durante interacción.

6.1. Búsqueda de orientaciones entre proteínas

Docking de cuerpo rígido

- Muestreo exhaustivo o búsqueda sistemática: **métodos FFT**
 - Programas:
 - **FTDOCK**: hasta 10000 modelos. Función de evaluación geométrica (posible incorporar electrostática).

Busca la superficie de la proteína, que es lo que señala en amarillo, y el core, que es lo que señala en rojo. Si el core está muy encerrado le pone una penalización

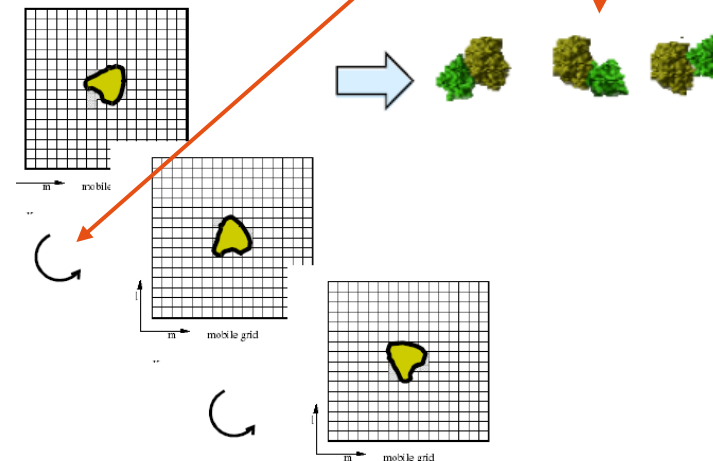


```
ftdock -static receptor.pdb -mobile ligand.pdb
```

```
-noelec -calculate_grid 1.2 -angle_step 12
```

```
-internal -15 -Surface 1.3 -keep 3
```

```
-out complex.ftdock > complex.ftdock.log
```



Temario - Contenidos

Tema 6. Modelización de interacciones por *docking*

6.1. Búsqueda de orientaciones entre proteínas

6.2. Puntuación de modelos de *docking*

6.3. Flexibilidad conformacional en *docking*

6.4. Evaluación de las predicciones de *docking*

Docking de cuerpo rígido

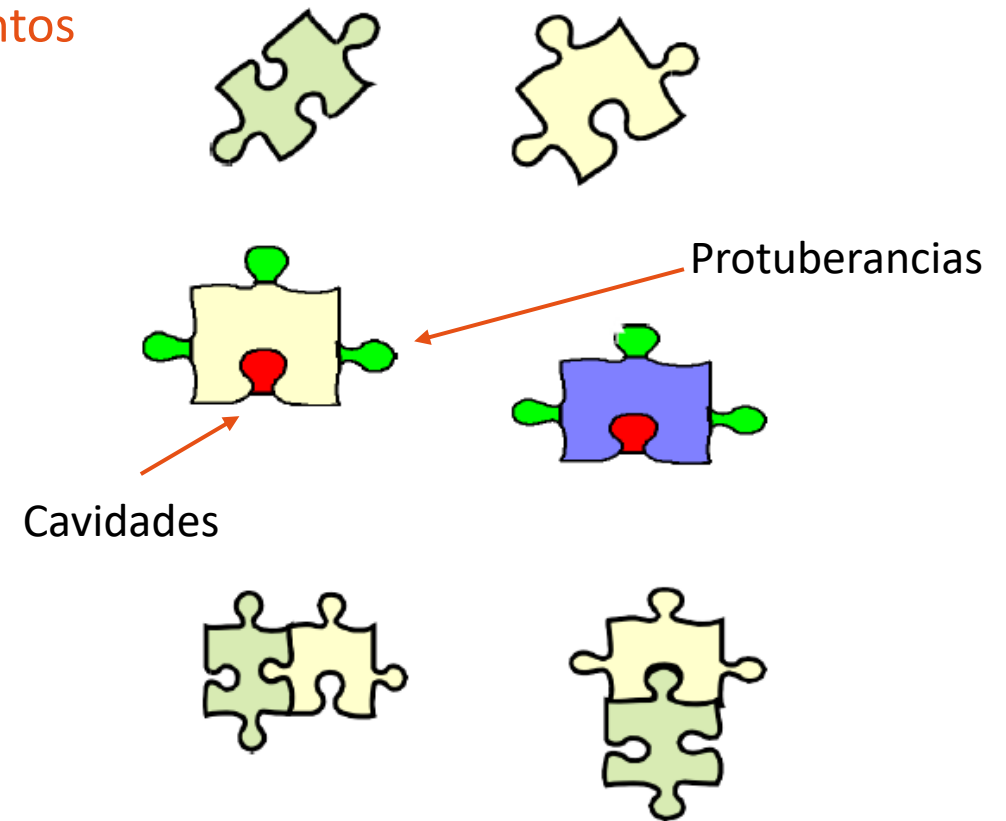
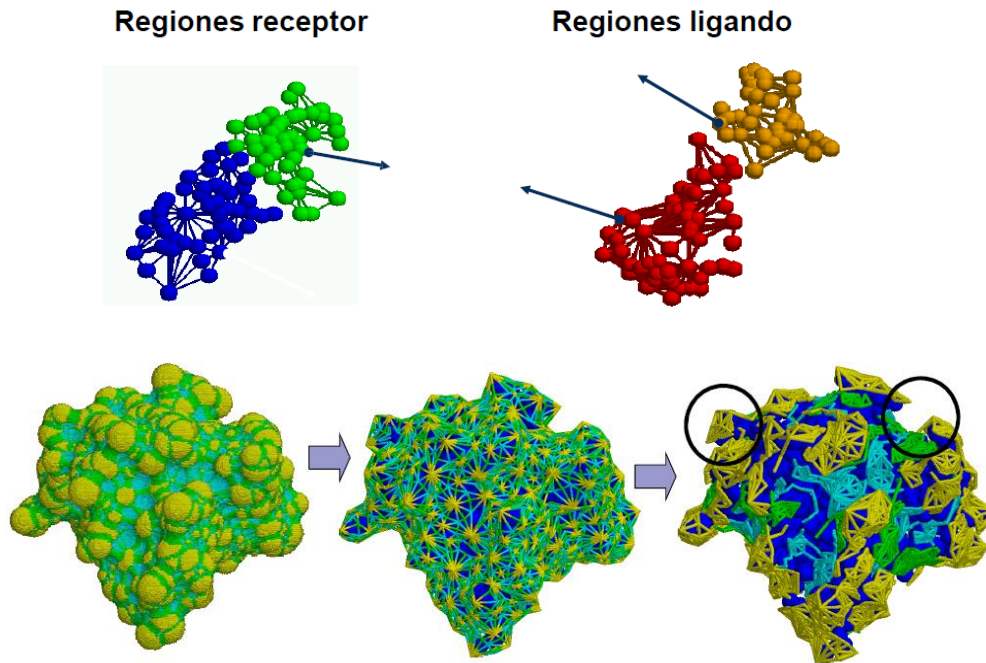
- Muestreo exhaustivo: métodos FFT
- Muestreo exhaustivo: fragmentación geométrica
- Muestreo estocástico: basado en energías

6.1. Búsqueda de orientaciones entre proteínas

Docking de cuerpo rígido

- Muestreo exhaustivo o búsqueda sistemática: **métodos de fragmentación geométrica (geometric hashing)**

- Representación simplificada de la superficie proteica.
- Características relevantes representadas con **pocos puntos** (cavidades/protuberancias).

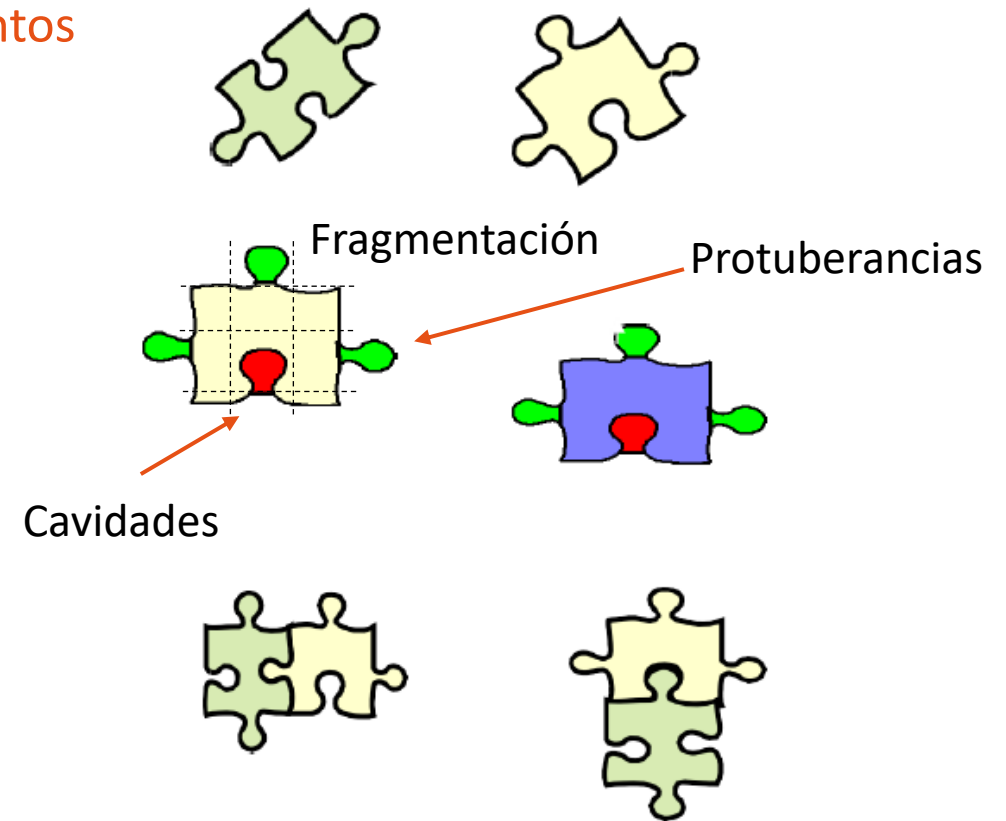


6.1. Búsqueda de orientaciones entre proteínas

Docking de cuerpo rígido

○ Muestreo exhaustivo o búsqueda sistemática: métodos de fragmentación geométrica (geometric hashing)

- Representación simplificada de la superficie proteica.
- Características relevantes representadas con pocos puntos (cavidades/protuberancias).
- **Fragmentación** en regiones de pequeño tamaño para definir orientaciones.
- Posterior evaluación con proteínas enteras.
- Desventajas:
 - Difícil introducir flexibilidad conformacional (retícula 3D).
 - No apropiado para proteínas que sufren grandes cambios conformacionales durante interacción.
- Programa PatchDock



Temario - Contenidos

Tema 6. Modelización de interacciones por *docking*

6.1. Búsqueda de orientaciones entre proteínas

6.2. Puntuación de modelos de *docking*

6.3. Flexibilidad conformacional en *docking*

6.4. Evaluación de las predicciones de *docking*

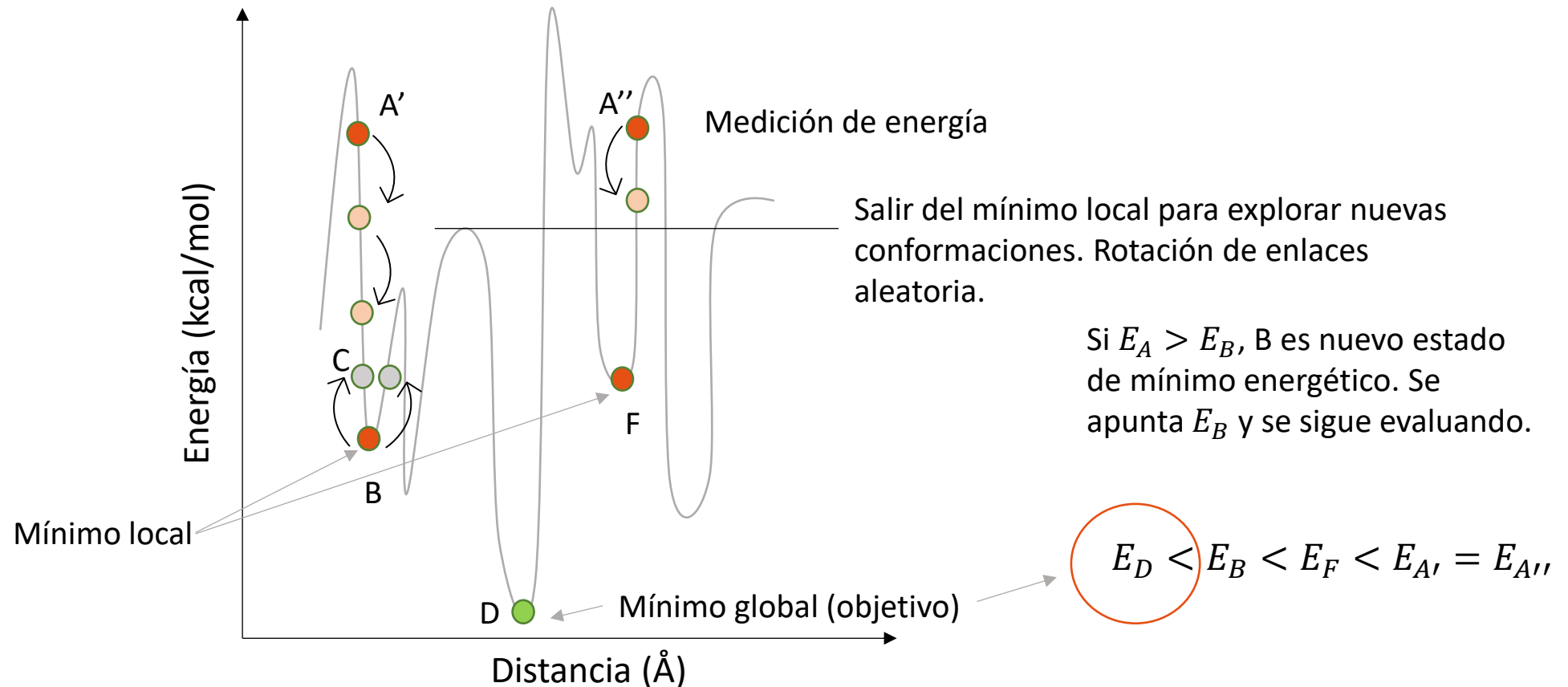
Docking de cuerpo rígido

- Muestreo exhaustivo: métodos FFT
- Muestreo exhaustivo: fragmentación geométrica
- Muestreo estocástico: basado en energías

6.1. Búsqueda de orientaciones entre proteínas

Docking de cuerpo rígido

- **Muestreo estocástico: métodos basados en energía (Método Monte-Carlo)**
- En vez de explorar todas las posibles orientaciones y evaluarlas energéticamente (costoso), se exploran unas pocas orientaciones al azar (junto con sus términos energéticos). Se siguen estrategias de optimización para identificar la mejor orientación en términos energéticos.



6.1. Búsqueda de orientaciones entre proteínas

Docking de cuerpo rígido

- **Muestreo estocástico: métodos basados en energía (Método Monte-Carlo)**
 - Al explorar menos conformaciones permiten mayor explicitad (nivel atómico) en interacción y funciones de puntuación.
 - Cierta flexibilidad conformacional.
- **ICM-DISCO** – (*Docking and Interface Side-Chain Optimization*) usa potenciales energéticos precalculados en retículas 3D.

$$E_{unión} = E_{Van\ der\ Waals} + E_{electrostática} + E_{puentes\ hidrógeno} + E_{hidroicidad}$$

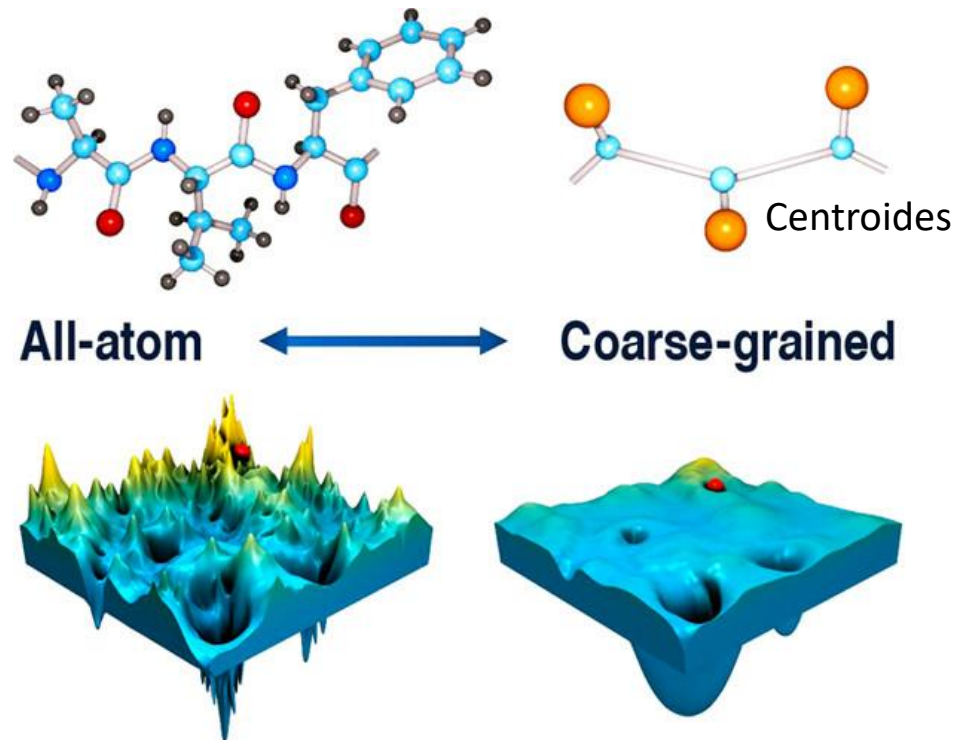
Algunos componentes limitados para
no introducir grandes desviaciones

- Exploración Monte-Carlo o **función de probabilidad (criterio Metrópolis** – Método Monte-Carlo en cadena de Markov, usado en muestreo de distribuciones multidimensionales, permite obtener una secuencia de muestras aleatorias según función de probabilidad).
- Los modelos se optimizan introduciendo cierta flexibilidad conformacional.

6.1. Búsqueda de orientaciones entre proteínas

Docking de cuerpo rígido

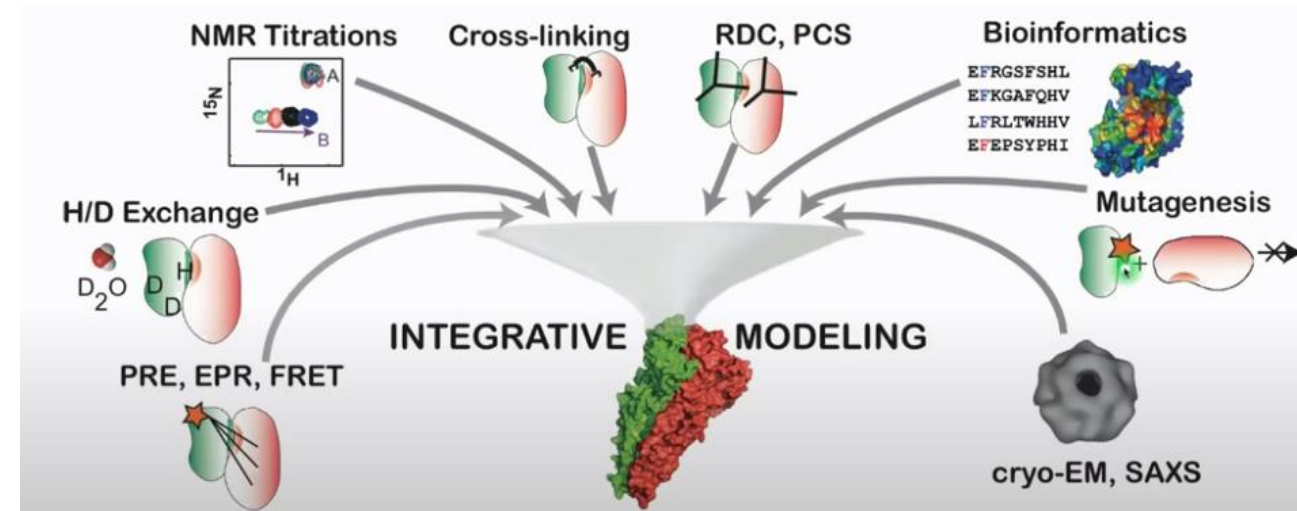
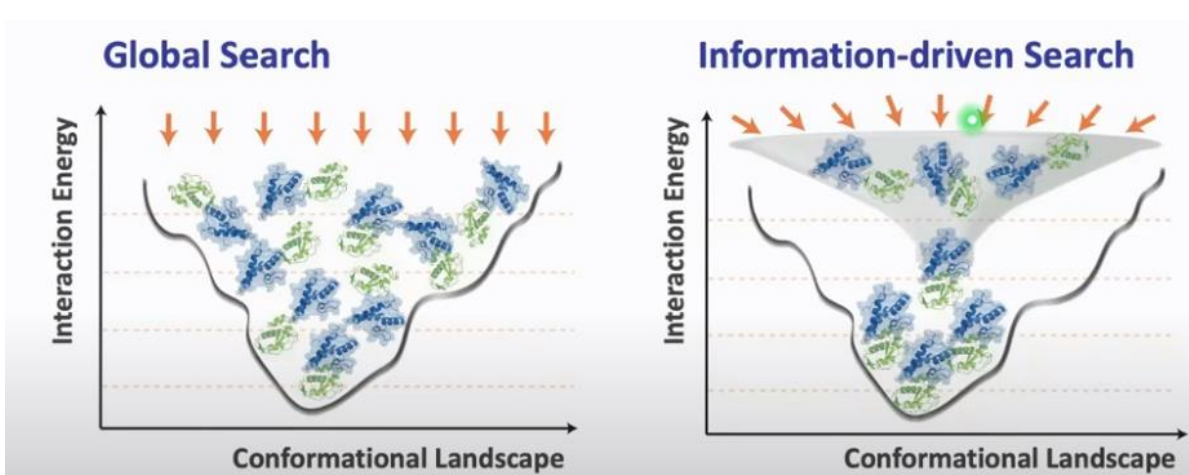
- **Muestreo estocástico: métodos basados en energía (Método Monte-Carlo)**
- Programa **RosettaDock** (*coarse-grained*)
 - Usa representación simplificada de las cadenas laterales (centroides).
 - Descripción energética – pares de residuos (simplificación).
 - Flexibilidad conformacional



6.1. Búsqueda de orientaciones entre proteínas

Docking de cuerpo rígido

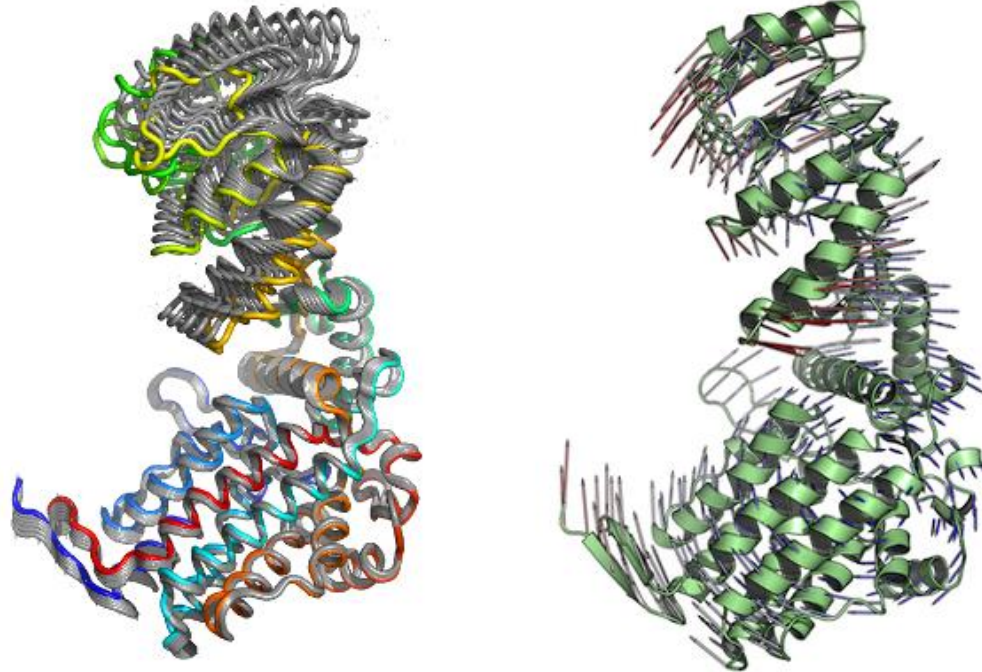
- **Muestreo estocástico: métodos basados en energía (Método Monte-Carlo)**
- Programa **HADDOCK** (*information-driven docking*)
 - Las orientaciones de docking provienen de dinámica molecular con campos de fuerza simplificados
 - Se nutre de información basada en conocimiento sobre residuos importantes para la interacción (estudios de mutagénesis, NMR, predicciones previas, etc.) (*integrative modeling*).



6.1. Búsqueda de orientaciones entre proteínas

Docking de cuerpo rígido

- **Muestreo estocástico: métodos basados en energía (NMA – *normal mode analysis*)**
- Calcula los estados vibracionales de la proteína para estudiar ligera flexibilidad conformacional (movimientos armónicos colectivos de gran amplitud que no están al alcance de simulaciones de dinámica molecular).
 - Modelo de redes elásticas (ENM – *elastic network model*) - tiene en cuenta solo Ca.



Temario - Contenidos

Tema 6. Modelización de interacciones por *docking*

6.1. Búsqueda de orientaciones entre proteínas

6.2. Puntuación de modelos de *docking*

6.3. Flexibilidad conformacional en *docking*

6.4. Evaluación de las predicciones de *docking*

Docking de cuerpo rígido

- Funciones de puntuación en *docking*
- Integración de información externa

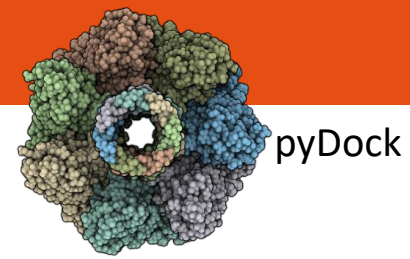
6.2. Puntuación de modelos de *docking*

Docking de cuerpo rígido

○ Funciones de puntuación

- Funciones (**pre**) de puntuación durante la generación de orientaciones a evaluar (simplificadas, deben ser rápidas).
- Funciones (**post**) de puntuación de modelos finales (complejas, específicas de cada método). Diseñadas para mejorar el filtrado, calidad y refinamiento de los mejores modelos.
 - **Complementariedad de forma** (la más básica). En métodos basados en FFT también sirve para **discretizar** otros términos energéticos basándose en la retícula (*grid*) – GSC (*grid-based shape complementarity*).
 - **Complementariedad geométrica.** Busca la complementariedad de forma dependiente de la curvatura entre superficies moleculares.
 - Los programas de *docking* basados en FFT intentan encontrar la mejor interacción entre las capas de superficie restando una penalización de choque entre los interiores proteicos.
 - Potenciales precalculados describiendo términos de **Van der Waals, electrostática, puentes de hidrógeno e hidrofobicidad**. Diferente ponderación de cada término dependiendo de la fase de evaluación.
 - Anterior + **desolvatación**
 - **Restricciones de distancia de residuos** (corta, larga - LSC (*long range shape complementarity*)).

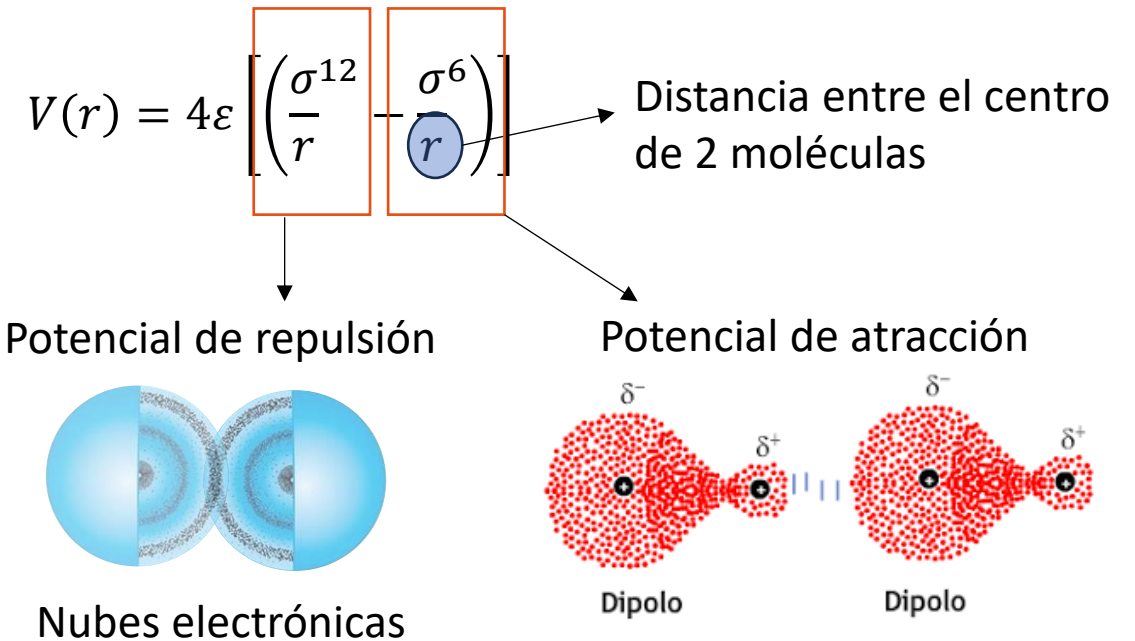
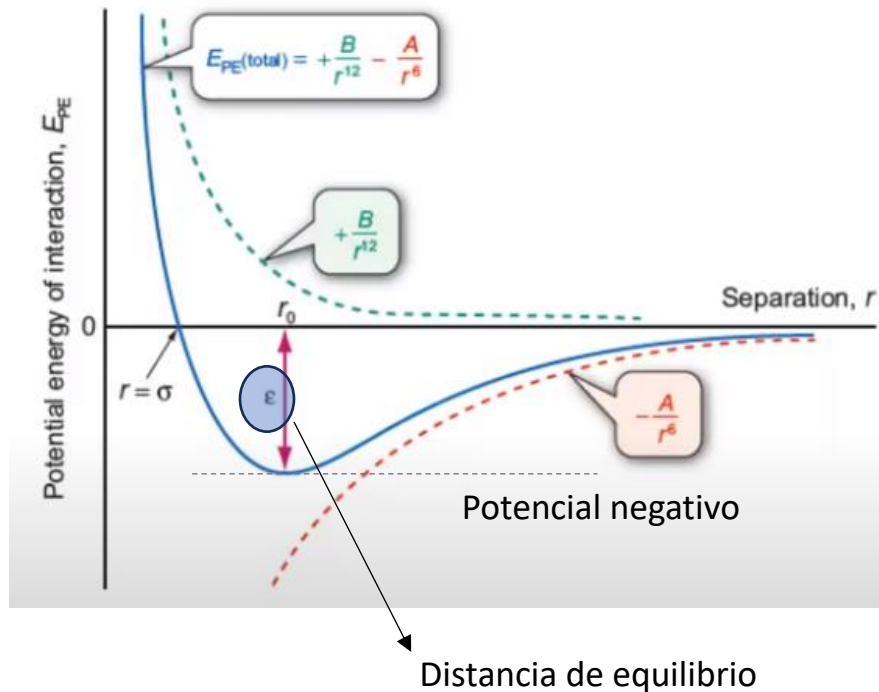
6.2. Puntuación de modelos de *docking*



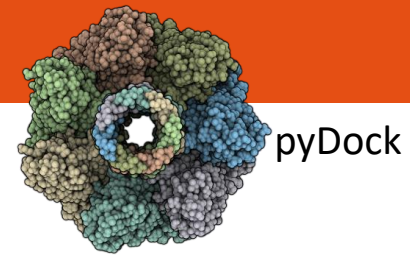
Docking de cuerpo rígido

○ Funciones de puntuación

- Aplicadas en multitud de métodos de *docking*. Función de puntuación basada en 3 términos energéticos
 - Energía de Van der Waals (usa el descriptor de Lenard-Jones)



6.2. Puntuación de modelos de *docking*



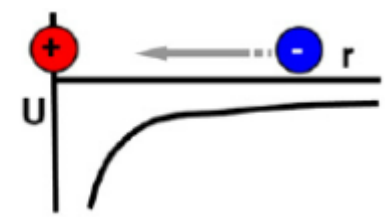
Docking de cuerpo rígido

○ Funciones de puntuación

- Aplicadas en multitud de métodos de *docking*. Función de puntuación basada en 3 términos energéticos
 - Energía electrostática

$$\sum_{i < j} \sum \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}$$

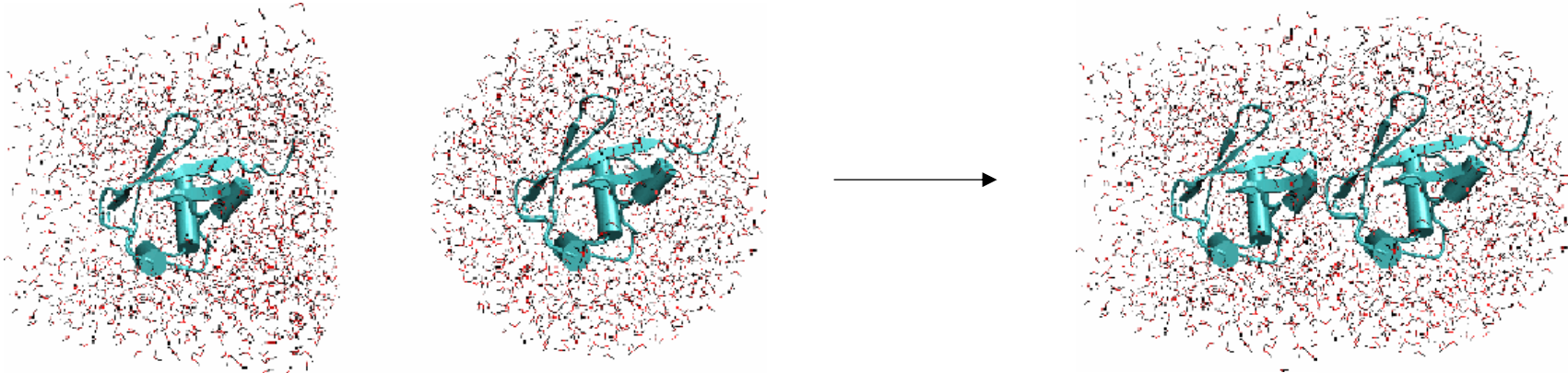
Carga 1 (i) Carga 2 (j)

A diagram illustrating the electrostatic interaction between two point charges. On the left, a red circle with a '+' sign represents 'Carga 1 (i)'. On the right, a blue circle with a '-' sign represents 'Carga 2 (j)'. A horizontal line connects them, with a double-headed arrow indicating the distance 'r'. Below the line, a curve shows the potential energy 'U' as a function of distance 'r', with the curve rising sharply as 'r' decreases.

Distancia que las separa (r)

Permisividad en el vacío (ϵ)

- Energía de desolvatación (ASA) complejo vs proteínas separadas



6.2. Puntuación de modelos de *docking*

Docking de cuerpo rígido

○ Funciones de puntuación

- Potenciales estadísticos usando bases de datos de estructuras de complejos.
- **Aprendizaje profundo y redes neuronales** (DOVE y GNN-DOVE).
- Funciones de puntuación clásicas siguen una relación lineal entre componentes (no siempre cierta) a diferencia de las funciones de puntuación basadas en aprendizaje automático.

Temario - Contenidos

Tema 6. Modelización de interacciones por *docking*

6.1. Búsqueda de orientaciones entre proteínas

6.2. Puntuación de modelos de *docking*

6.3. Flexibilidad conformacional en *docking*

6.4. Evaluación de las predicciones de *docking*

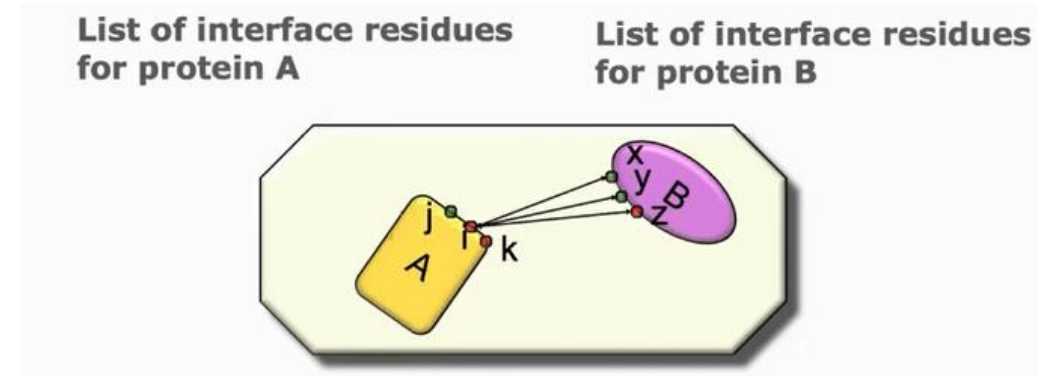
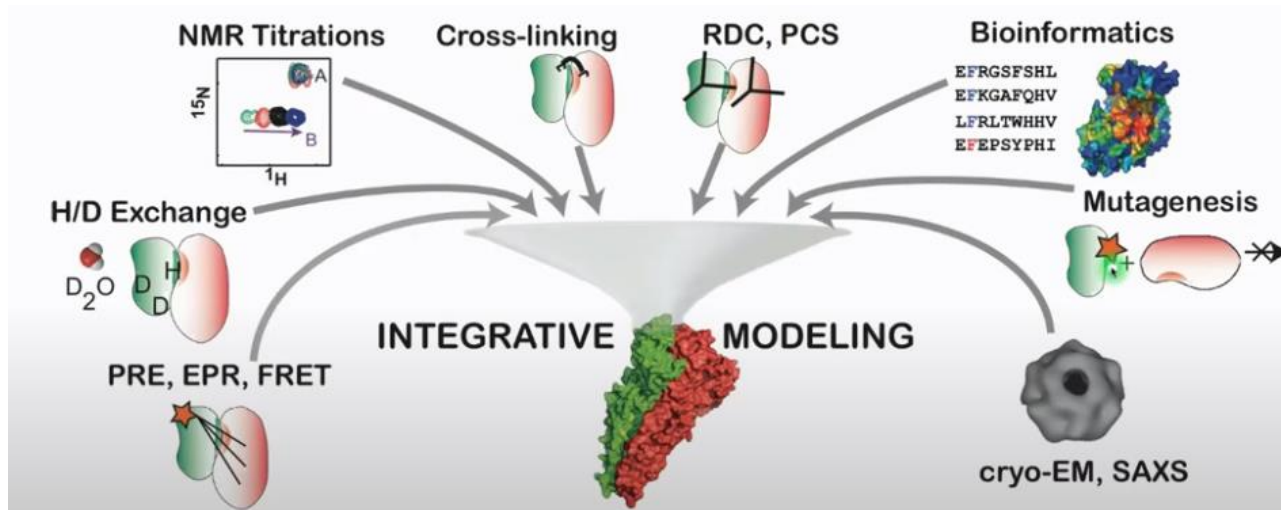
Docking de cuerpo rígido

- Funciones de puntuación en *docking*
- Integración de información externa

6.2. Puntuación de modelos de *docking*

Docking de cuerpo rígido

- Integración de información externa (modelado integrativo)
 - HADDOCK
 - PyDockRST – introducción de restricciones en cuanto a residuos importantes para la interacción según datos experimentales





viu

Universidad
Internacional
de Valencia

universidadviu.com

De:
 Planeta Formación y Universidades