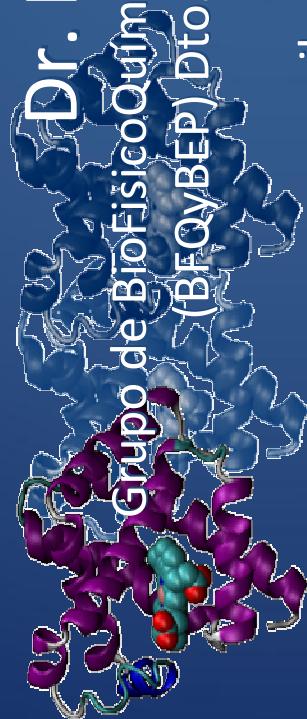


# HPC en Biomoleculas

## Como aprovechar mas computadoras para obtener la respuesta correcta



Dr. Marcelo A. Martí

Gruppo de BioFísicoQuímica y Bioinformática Estructural de Proteínas  
(BFQyBEP) Dto. de Química Biológica e INQUIMAE  
FCEN-UBA-CONICET

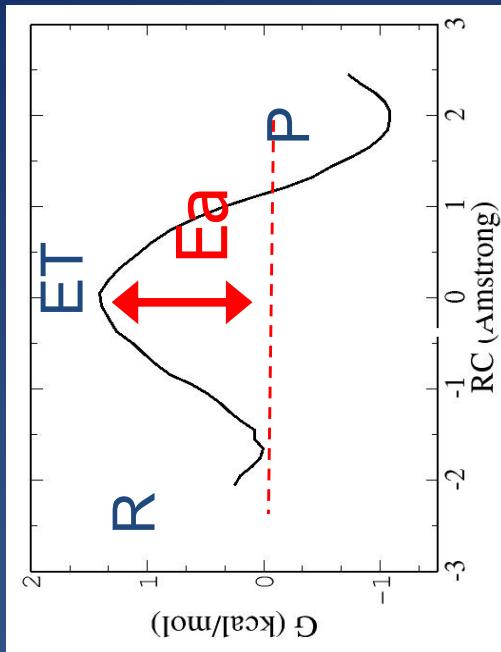
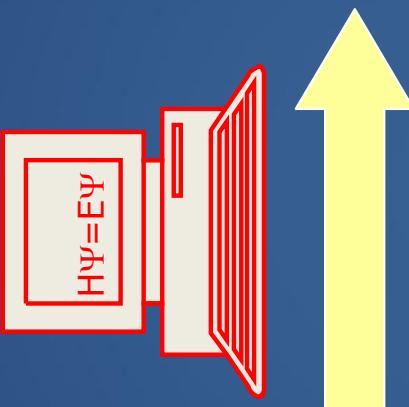
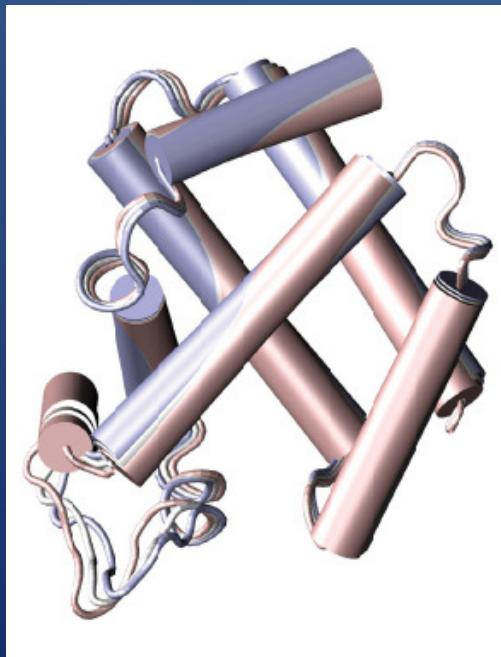
mail: marti.marcelo@gmail.com



# ¿Qué respuestas me da la simulación?

- Lo importante es hacer la pregunta correcta
- La respuesta “debería” ser algo potencialmente medible experimentalmente
- A escala molecular la termodinámica gobierna los procesos
- Hay que traducir la pregunta “biológica” en una pregunta “termodinámica”

¿Como me da la respuesta la simulación?  
Calcular la energía libre asociada al proceso!



$$\Delta G = RT \ln(K_{eq})$$

- Mutaciones *in-silico* para ver efecto de aminoácidos particulares

$$V_r = A e^{-(G_{act}/RT)}$$

Pensemos Ejemplos!

# ¿Por que es importante Calcular la Energía Libre?

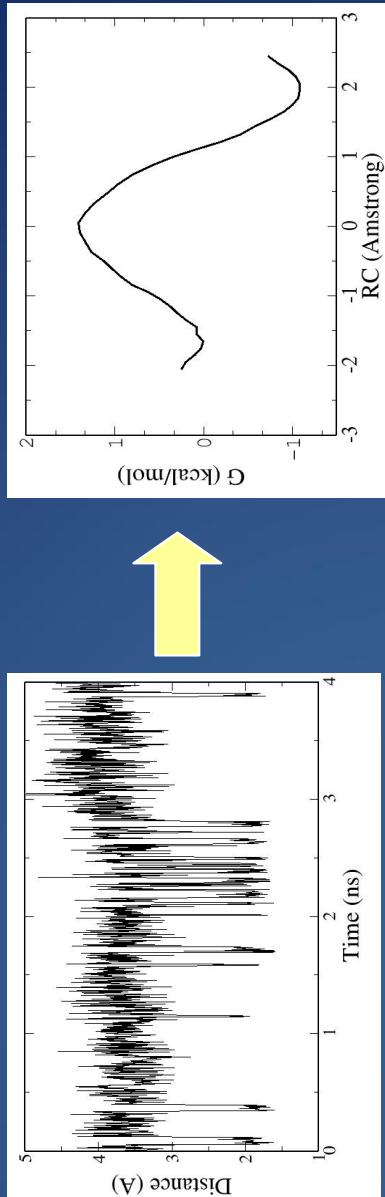
- Nos dice si un proceso (Químico, Físico y/o biológico) es, o no espontáneo
- Ejemplo  $H_2O$  sólida  $\Rightarrow H_2O$  líquida ( $T > 0C$ )
- Termodinámicamente:
  - $\Delta A = \Delta E - T\Delta S$  (a V, T ctes..)
  - $\Delta G = \Delta H - T\Delta S$  (a P, T ctes...)
- Ahora  $H = E + P\Delta V$  (en fases condensadas  $H \approx E$ )
- En Simulación E es función de las coordenadas ( $r^N$ ) Sale directamente.....el problema es “S”?

# $S = Entropía$

- ¿Cuándo es importante  $S$ ?
  - ¿que es la entropía?
- Mezcla de Gases
- En el famoso “efecto hidrofobico”
- Unión Proteína-Ligando (droga-receptor)
  - Plegamiento de Proteínas
  - Mecanismos de reacción
- $S$ , es un fenómeno “ESTADÍSTICO”
  - $S$ , es desorden...
  - $S$ , es el numero de de microestados posibles...
    - ( $S = k \ln W$ )
    - pero tambien....
  - $\delta S = q_{rev}/T$
  - (definición termodinámica)

# ¿Cómo obtengo $\langle A \rangle$ o $\langle G \rangle$ de una simulación?

- Muestrear mucho para llegar al equilibrio, tal que:  
$$G(r_C) = -kT \ln(r_C)$$



- O, usar muchas computadoras y combinar las simulaciones para obtener el  $G(r_C)$ , como?

# Un poco de Termodinámica estadística

- Promedio colectivo => “ $\langle X \rangle$ ”
- Hipótesis ergodica:  $\langle X \rangle_t = \langle X \rangle$  de N partículas,
  - con N => infinito.
- $\langle X \rangle = \sum X_i P_i$ 
  - (donde  $P_i$  es la probabilidad de obtener  $X_i$ )
- $P_i = \exp(-E_i/kT) / Q$ 
  - $Q$  es la función de partición  $Q = \sum \exp(-E_i/kT)$
- $\langle A \rangle = \langle E \rangle - T \langle S \rangle$ 
  - $\langle E \rangle$  sale directo de la simulación pero ¿ $\langle S \rangle$ ?
- $\langle S \rangle = -k \sum P_i \ln P_i$ 
  - (necesito conocer TODAS las  $P_i$ , MUY difícil...)
- $\langle A \rangle = -kT \ln Q$ 
  - es muy difícil de calcular => veamos por que....

# El problema del Muestreo..

$$\langle E \rangle = kT^2/Q \delta Q/\delta T$$

$$\delta Q/\delta T = \sum E_i \exp(-E_i/kT) / T^2$$

Entonces:

$$\langle E \rangle = k \sum E_i \exp(-E_i/kT)/Q$$

Pero...

$$\exp(-E_i/kT)/Q = P$$

Entonces:

$$\langle E \rangle = k \sum E_i P_i$$

Los estados de alta energía  $E_i$  alto, contribuyen poco.

Ej:  $E_i=10kT \Rightarrow P_i=0.00005$ ,  
 $E_i=1kT \Rightarrow P_i=2.7$

En cambio....

$$\langle A \rangle = -kT \ln Q \dots$$

$$\text{Con } Q = \sum \exp(-E_i/kT)$$

Ahora notemos que:

$$1 = \sum \exp(E_i/kT) \exp(-E_i/kT)$$

$$\text{Reemplazo..} \langle A \rangle = kT \ln(\sum \exp(E_i/kT) \exp(-E_i/kT)/Q)$$

$$\text{Pero: } \exp(-E_i/kT)/Q = P_i$$

Entonces

$$\langle A \rangle = kT \ln(\sum P_i \exp(E_i/kT)) \text{ Ahora los valores poco probables, } E_i \text{ alto pesan MUCHO!}$$

# Umbrella Sampling

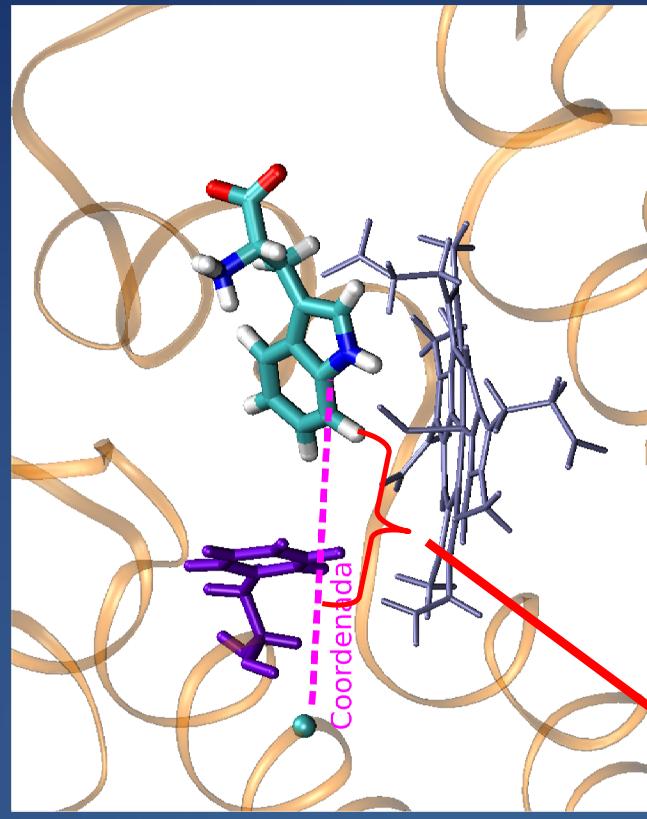
# Umbrella Sampling

- El método se basa en utilizar un campo de fuerzas modificado para muestrear MEJOR las zonas poco probables naturalmente.. Ej barreras..
- La modificación  $W$  es por lo general un potencial armónico ( $W(r) = k(r - r_o)^2$ ) que fuerza al sistema a estar en la zona cercana a  $r_o$ , para recuperar los promedios NO modificados se usa la siguientes ecuaciones
- Sea el potencial modificado  $E_w(r) = E(r) + W(r)$ , y la distribución de probabilidades resultante  $P_w$ .. que da lugar a un Valor medio modificado  $\langle X \rangle_w$ , el valor medio inicial se obtiene de acuerdo con:  
$$\langle X \rangle = \langle X \rangle_w / \langle \exp(-W(r)/kT) \rangle_w$$
- si lo que buscamos es el valor medio de  $P(r)$ ...  
$$P(r) = P_w(r) / \exp(W(r)/kT)$$
- por lo que la energía libre en función de  $r$ ,  $A(r)$  se obtiene de acuerdo a:
- $A(r) = -kT \ln P(r) - W(r)$

# En la práctica

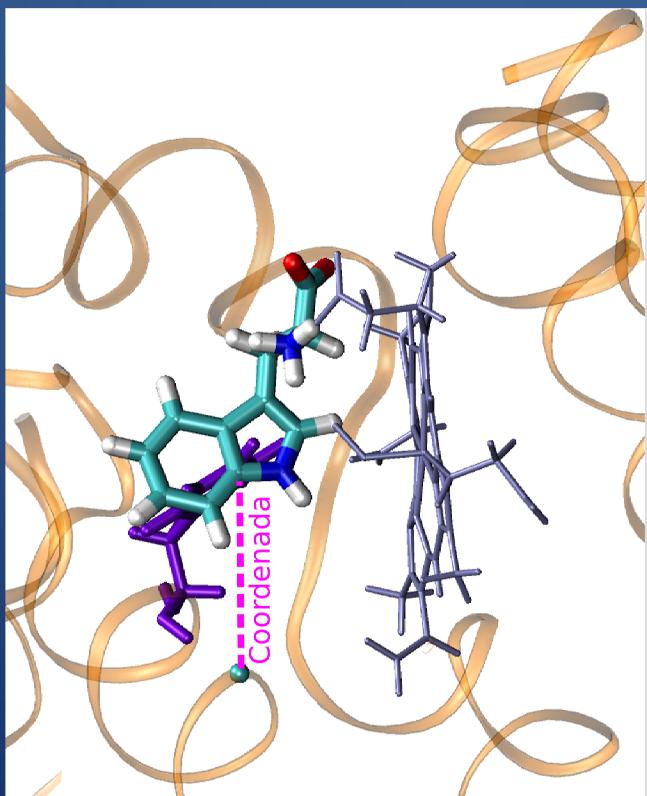
- Se elije una coordenada de reacción “ $r$ ” representativa del proceso y un rango,  $r_{\min} - r_{\max}$
- Se realizan distintas simulaciones (cada una es llamada Ventana) en las cuales se varía  $W(r)$  modificando en cada caso ambos  $k$  y  $r_0$ .
- De cada simulación se monitorea  $r(t)$ , y se obtiene el histograma correspondiente que corresponde a  $P(r)$ .
- Se obtiene:  $A(r_i) = -KT \ln P(r) - W(r)$ 
  - para el rango MUESTREADO de  $r$  (llámémoslo  $r_i$ )
- $A(r_{\min} - r_{\max})$  se obtiene COMBINANDO las distintas  $A(r_i)$
- Nótese que:
  - Dado que el CERO de  $A(r_i)$  es arbitrario puedo mover cada  $A(r_i)$  arriba o abajo para combinarla con la ventana siguiente
  - Lo importante es que entre dos ventanas adyacentes haya superposición de  $A(r_i)$  en un rango razonable de  $r_i$  (aprox 0.2-0.5 Å)

# Casos Reales: Equilibrio entre los dos sitios de unión de la hIDO



Conformación 2

Sitio de unión inicial

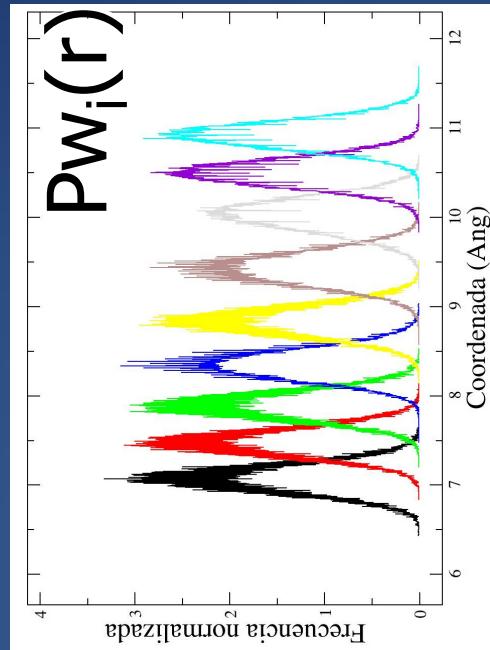
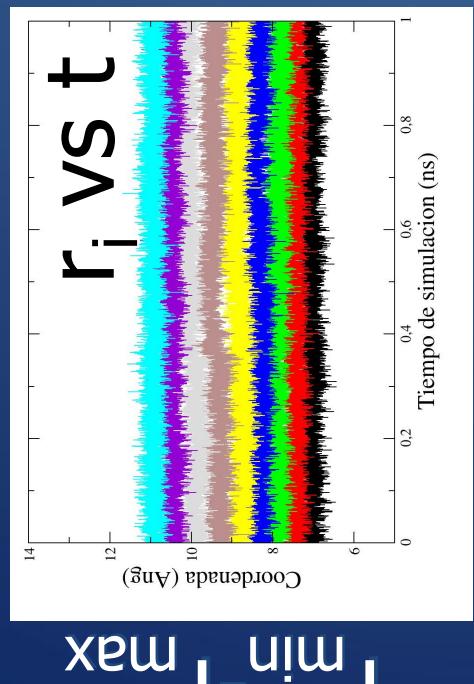


Conformación 1

Conformación Reactiva

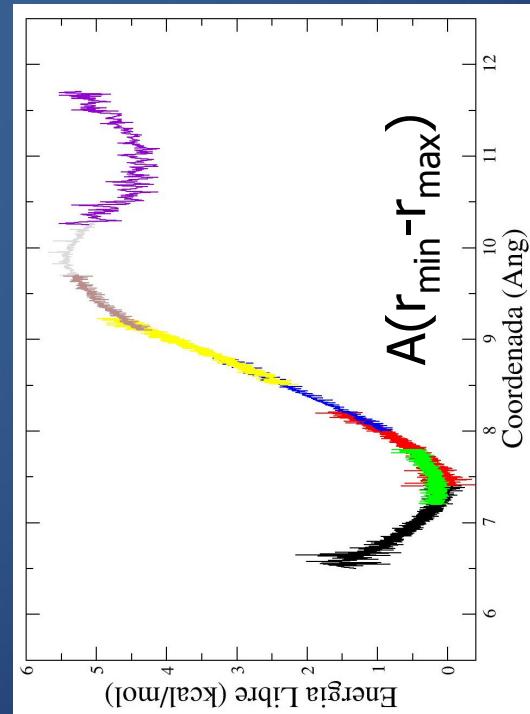
Divido la amplitud de la RC en segmentos,  
cada uno en una maquina distinta!

# Casos Reales: Equilibrio entre los dos sitios de unión de la hIDO



**MD: con distintos  $W(r)$**

$$A(r_i) = -KT \ln P(r) - W(r)$$



Cada VENTANA en una maquina separada!

Obtención de energía libre de dinámicas de no equilibrio:  
la igualdad de Jarzynski

## Obtención de energía libre de dinámicas de no equilibrio: la inigualdad de Jarzynski

- Demostración: 1996 C. Jarzynski PRL: Nonequilibrium Equality for Free energy Differences”
- Supongamos un sistema que evoluciona de A=>B INFINITAMENTE LENTO por medio de una FUERZA EXTERNA, como el sistema siempre esta en equilibrio:
  - $Wr = \Delta A_{AB}$
- Ahora si la evolución de A=>B, ocurre a una **velocidad finita** ( $v$ ) el  $Wi$  (trabajo irreversible) depende de la condición microscópica inicial y:
  - $\langle Wi \rangle \geq \Delta A_{AB}$
  - La diferencia entre  $\langle W \rangle$  y  $\Delta A_{AB}$ , es el  $W$  disipado, asociado al aumento de “S” en un proceso IRREVERSIBLE
- Lo que voy a demostrar es que:
  - $\langle \exp(-\beta Wi) \rangle = \exp(-\beta \Delta A_{AB})$  con  $\beta = 1/kT$
  - y el resultado es INDEPENDIENTE de  $v$

# Demonstración

- Hay que probar que:
  - $\Delta A_{AB} = -1/\beta \ln \langle \exp(-\beta W_i) \rangle$
- Analicemos un sistema aislado:
  - Definamos el  $W_i$  (es el trabajo que se hace sobre el sistema debido a la fuerza que mueve el sistema de  $A \Rightarrow B$ ).
    - Sea  $\lambda$  una coordenada espacial generica del sistema tal que en  $A$   $\lambda=0$  y en  $B$   $\lambda=1$ , y sea  $H$ , el hamiltoniano del sistema:
  - $W_i = \int \lambda \delta H / \delta \lambda \, d\lambda$ 
    - (es la integral de Fuerza x distancia a lo largo de toda la coordenada)
  - Ahora recordemos que el cambio de  $\lambda_{0 \Rightarrow 1}$  se hace en realidad a una velocidad finita o sea en un tiempo “ts”.
    - Donde  $\lambda=0$  a  $t_0$  y  $\lambda=1$  a  $t_s$ . Por lo que la integral la puedo hacer en tiempo.
  - Sea además  $Q$  la función de partición del sistema, en el estado inicial tengo al sistema descrito por:  $t_0 Q_0$

# Demonstración

- Cuando el sistema evoluciona irreversiblemente lo hace mediante distintas trayectorias del ensemble, determinadas por la siguiente distribución de probabilidades:  $f(z,t_0)$ 
  - $f(z,t_0) = \exp(-\beta H(z,t_0)) / Q_0$ 
    - “z” indica la conformación inicial que da lugar a una trayectoria particular
- En este sistema puedo calcular el promedio colectivo exponencial de Wz de acuerdo a la fórmula genérica:  $\langle x \rangle = \int x_i p(x_i)$  para todos los “i”
  - $\langle \exp(-\beta Wz) \rangle = \int f(z,t_S) \exp(-\beta W(z,t_S)) dz$
  - Donde:  $f(z,t_S)$  probabilidad del estado z a  $t_S$ ,  $W(z,t_S)$  es el W de dejar evolucionar al sistema con configuración inicial z hasta  $t_S$ , y  $dz$  indica que la integral es sobre todas las conformaciones de “z”
- Ahora para un sistema AISLADO: “W = ΔE” pq no hay intercambio de calor, ergo no hay  $ΔS$
- Entonces:  $W(z,t_S) = ΔE_{t0=>t_S} = H(z,t_S) - H(z,t_0)$ 
  - diferencia de energía del sistema entre estado inicial y final

# Demostración

- Ahora se invoca al teorema de Liouville que dice que la densidad de estados en el espacio de fases se conserva a lo largo de cualquier trayectoria o en otras palabras que:  $f(z, t_0) = f(z, t_S)$
- Si combino entonces las ecuaciones anteriores tengo que:
  - $\langle \exp(-\beta W_z) \rangle = \int f(z, t_S) \exp(-\beta W(z, t_S)) dz$
  - $\langle \exp(-\beta W_z) \rangle = \int f(z, t_0) \exp(-\beta W(z, t_S)) dz$
  - $\langle \exp(-\beta W_z) \rangle = 1/Q_0 \int \exp(-\beta H(z, t_0)) \exp(-\beta W(z, t_S)) dz$
  - $\langle \exp(-\beta W_z) \rangle = 1/Q_0 \int \exp(-\beta (H(z, t_0) + H(z, t_S) - H(z, t_0))) dz$
  - $\langle \exp(-\beta W_z) \rangle = 1/Q_0 \int \exp(-\beta H(z, t_S)) dz$
  - pero  $\int \exp(-\beta H(z, t_S)) dz = Q_{tS}$
  - $\langle \exp(-\beta W_z) \rangle = Q_{tS}/Q_0$
  - Por definición  $\Delta A_{AB} = -1/\beta \ln (Q_B/Q_A)$
  - Entonces:  $-1/\beta \ln \langle \exp(-\beta W_z) \rangle = -1/\beta \ln (Q_{tS}/Q_0)$
  - Pero recordemos que  $Q_{tS}$  es  $Q_B$  y  $Q_0$  es  $Q_A$  por lo que:
    - $-1/\beta \ln \langle \exp(-\beta W_z) \rangle = \Delta A_{AB}$
    - **QUE ES LO QUE QUERIAMOS PROBAR.**

# Demonstración

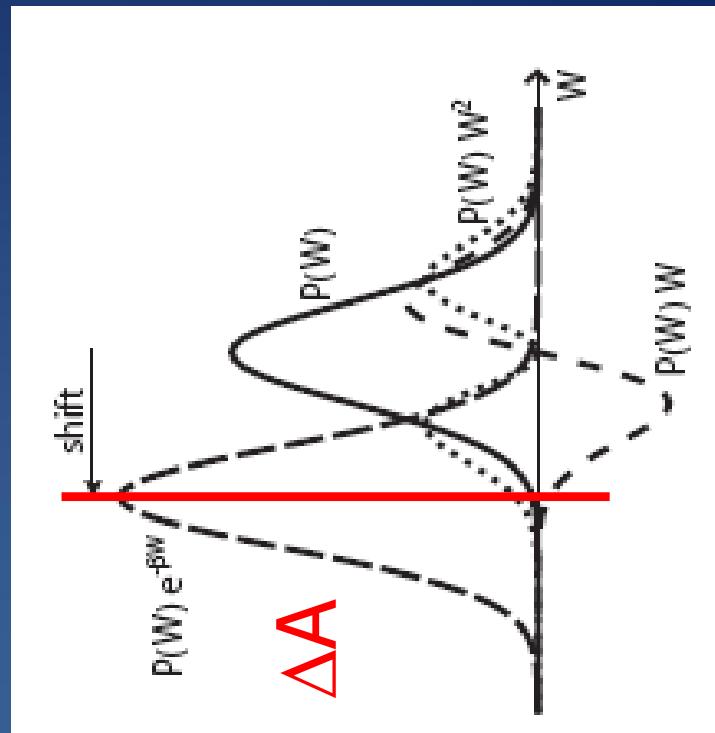
- Para obtener información termodinámica relevante los experimentos no se llevan a cabo en sistemas aislados, sino por lo general, en sistemas acoplados a fuentes de energía “Baño termostático”
- Para probar que la igualdad de Jarzynski vale también en estos casos suponemos que:
  - El sistema TOTAL gobernado por el hamiltoniano  $H_{Tot}$  esta AISLADO con:
    - $H_{Tot} = H_S + H_R + h_{Int}$
  - Donde  $H_s$  es el hamiltoniano del sistema de interés,  $H_R$  el del reservorio y la  $h_{Int}$  interacción entre ambos. Se asume que  $h_{Int}$  Es chica comparada con  $H_R$ , tal que en el sistema de interés el espacio de fases esta determinado por  $H_S$  indepnde de lo que ocurra en  $H_R$  (weak coupling regime)

# Demonstración

- Sea  $Q_T$  la función de partición del sistema tota aislado vale la demostración anterior:
  - $\langle \exp(-\beta W_Z) \rangle = Q_T(t_S)/Q_T(t_0)$
- Analizando la ecuación vemos que:
  - el lado derecho " $Q_T(t_S)/Q_T(t_0)$ " solo depende del estado inicial  $t_0$  y del final  $t_S$ , por lo tanto vale para cualquier valor de  $t_S$ .
- Pero entonces si,  $t_S \Rightarrow \infty$  vale que:  $W_r = \Delta A$ ,  
y entonces  $\langle \exp(-\beta W_r) \rangle = \exp(-\beta \Delta A)$
- Entonces como vale para  $t_S \Rightarrow \infty$  vale para cualquier  $t_S$  que es lo que queríamos probar!
- NOTA: la clave de esta segunda parte de la demostración es que  $W_r = \Delta A$  cosa que SOLO vale si el acoplamiento: ( $h_{int}$ ) es chico.

## Corolarios importantes de la demostración

- Siendo que la distribución de posibles trabajos es Gaussiana, se puede expandir el promedio exponencial de  $W$  según una serie dando lugar a la siguiente relación importante:
  - $\Delta A = \langle W \rangle - \beta/2 \sigma^2$
  - donde  $\sigma^2 = \langle W^2 \rangle - \langle W \rangle^2$
  - (es el desvió estandard de los  $W$ )
- Esta formula sirve para **testear la convergencia**, ya que  $\sigma^2$  tarda mas en converger que:  $\langle \exp(-\beta Wr) \rangle$



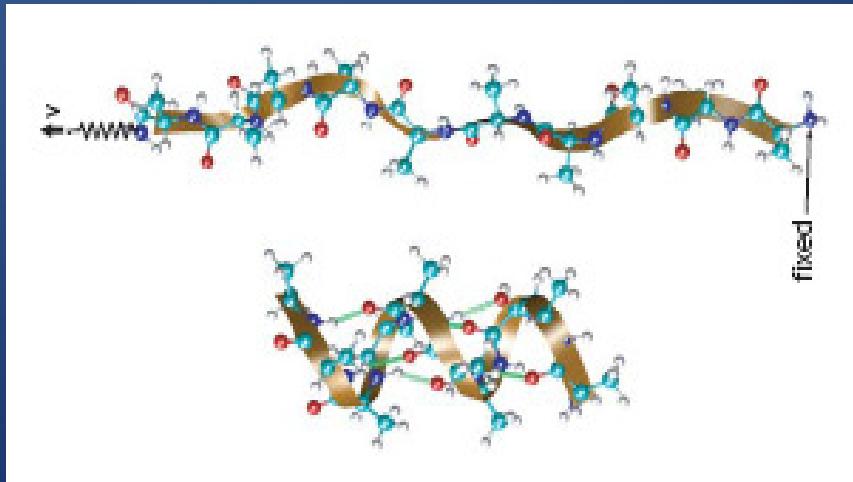
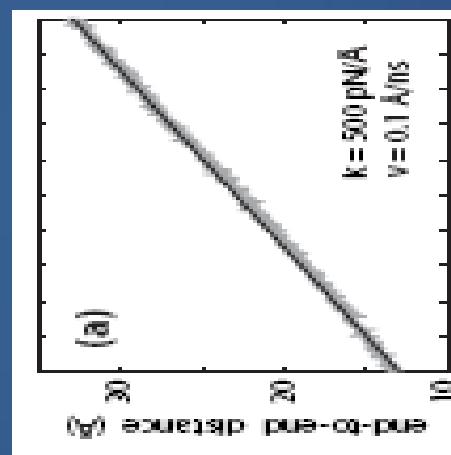
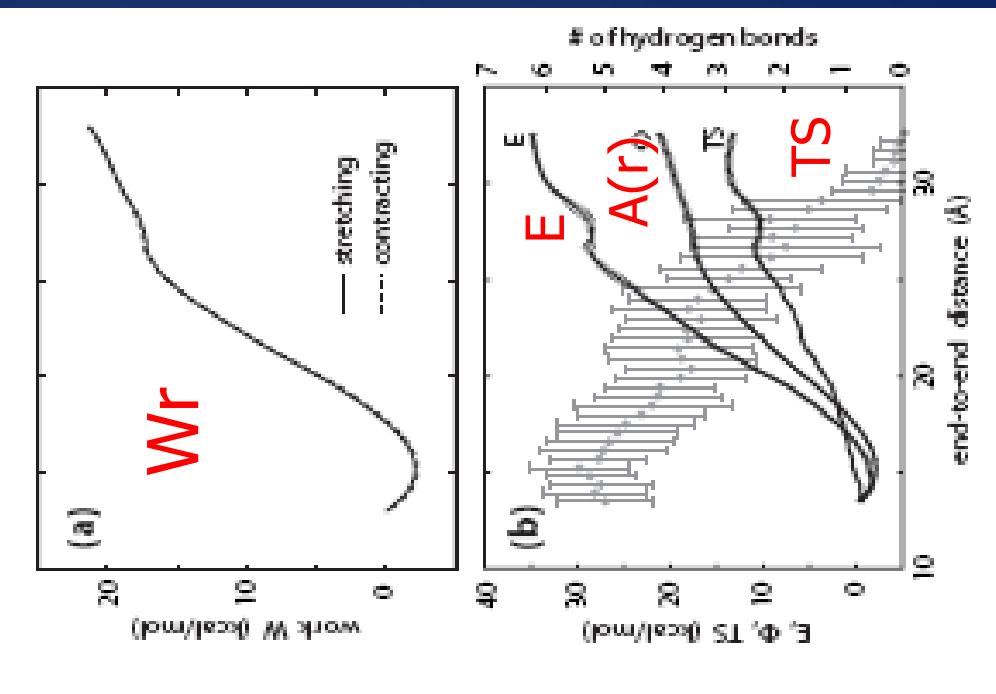
- En la práctica además:
  - Si:  $\sigma^2 \gg kT$
  - $\langle \exp(-\beta Wr) \rangle$  esta dominado por valores muy lejos de  $\langle W \rangle$ , lo cual son eventos muy raros, y requiere mayor muestreo.

# Como lo aplico a SC

- Se agrega al H del sistema un potencial armónico dependiente del tiempo:
  - $E_{ext} = k(r - r_o)^2$ , con  $r_o = r_i + vdt$
  - “r” es la RC que “guia” al sistema de A=>B
- “v” es el **parametro clave**
- El W en función de “r” se calcula como la integral de la Fuerza externa:
  - $Wi(r) = \int r \delta E_{ext} / \delta r dr$
- A partir de una MD de equilibrio en A, selecciono configuraciones de partida ( $C_i$ ) y las llevo al estado B, y para cada una calculo el trabajo  $Wi(r)$
- Utilizando la “JE” calculo  $A(r)_{A=>B}$
- En la practica calculo  $A(r)_{A=>B}$  y  $A(-r)_{B=>A}$

# Ejemplo: Desnaturalización de hélice de deca-alaninas

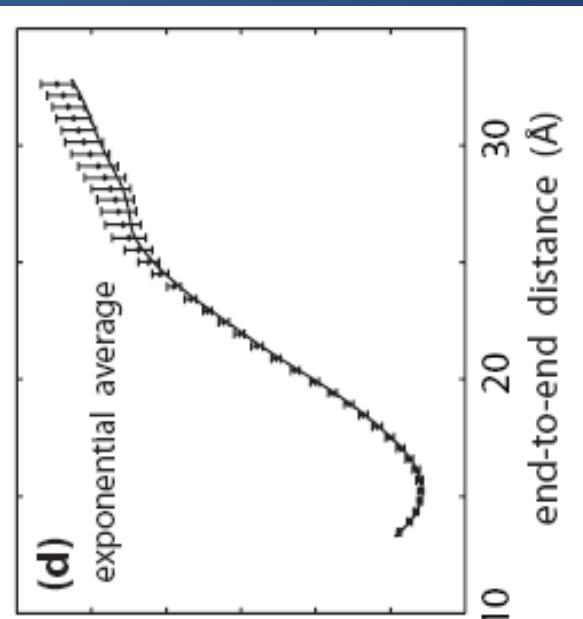
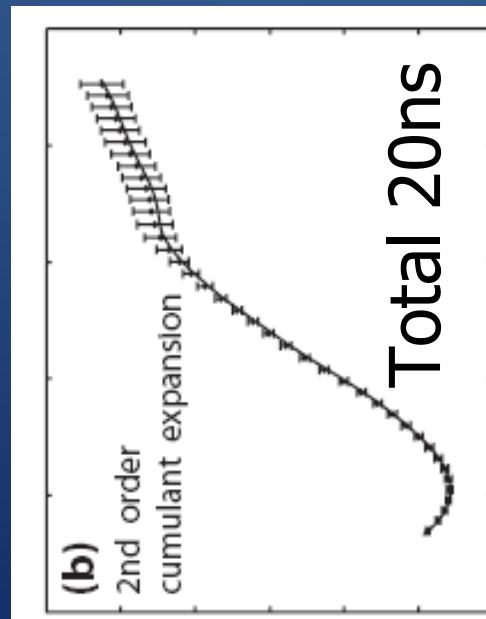
1ero Proceso Reversible: 200ns  
RC: distancia punta-punta  
 $v=0.1\text{\AA/ns}$



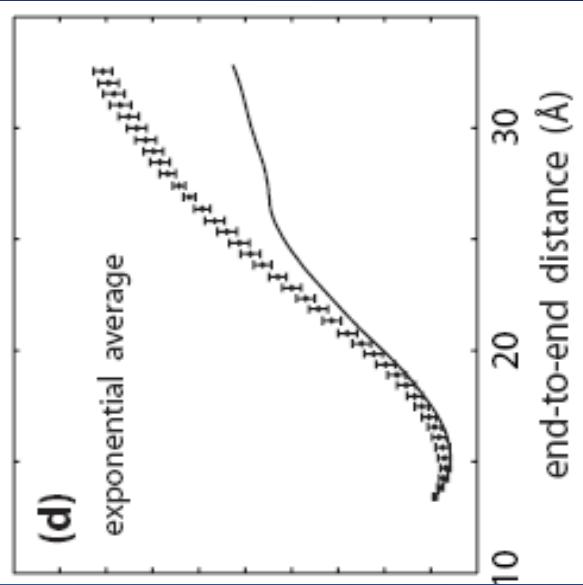
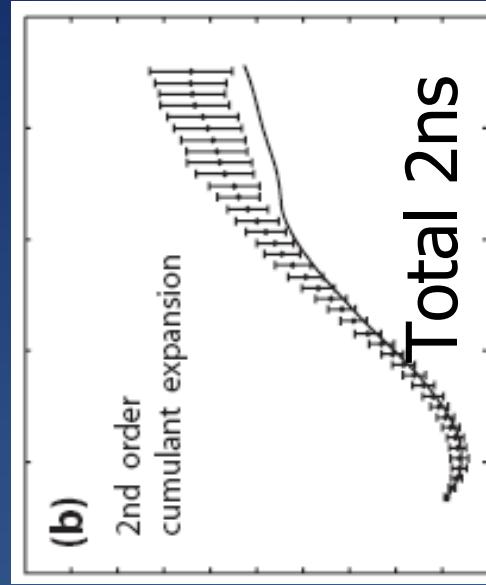
## Ejemplo: Desnaturalización de hélice de decá-alaninas. Proceso Irreversible:

10 Dinámicas de  $v=10\text{ \AA}/\text{ns}$ ,  $\nu=100\text{ \AA}/\text{ns}$

$$\langle W \rangle - \beta/2 \sigma^2$$

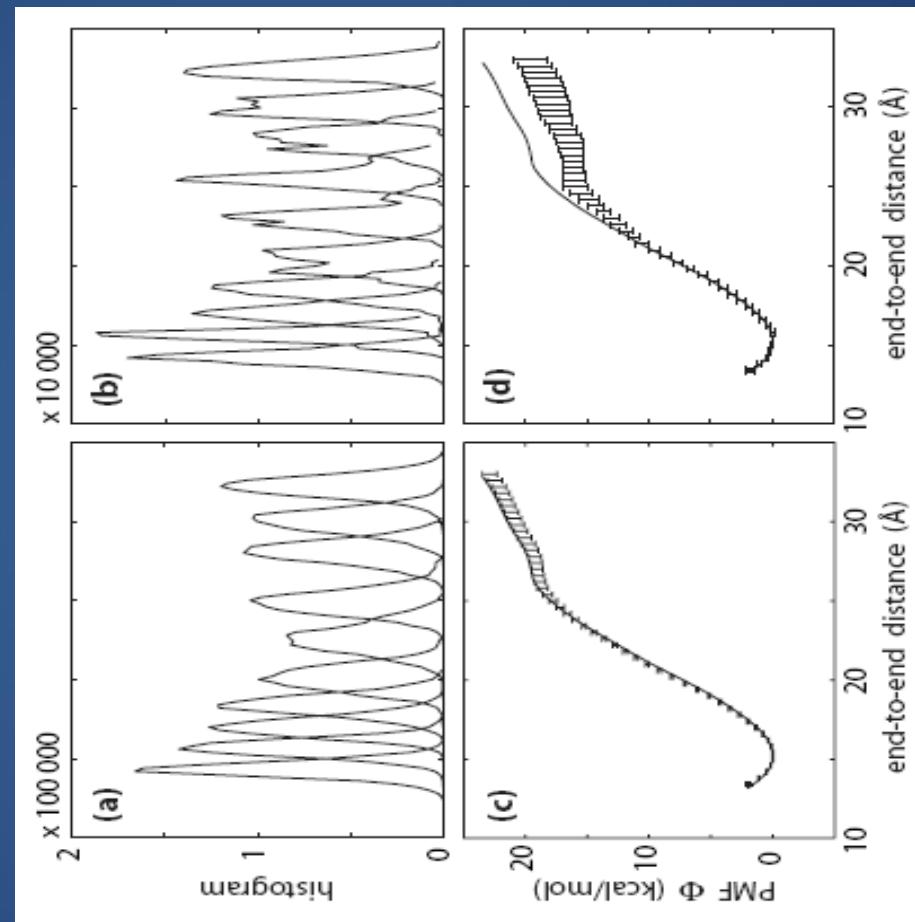


$$\langle \exp(-\beta W_Z) \rangle$$



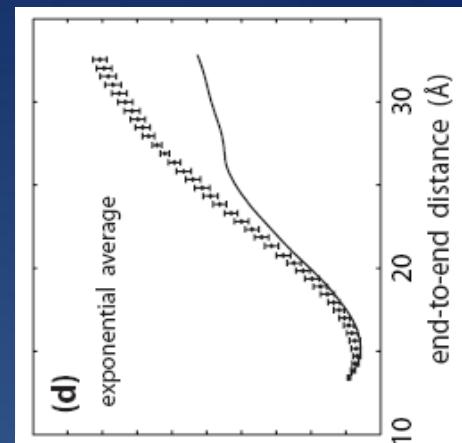
Ejemplo: Desnaturalización de hélice de decáalaninas. Comparación con Umbrella Sampling

## Umbrella Sampling Total 2ns

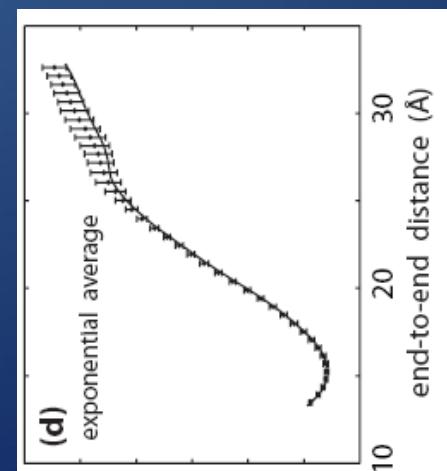


## Total 20ns

$$\langle \exp(-\beta W_z) \rangle$$



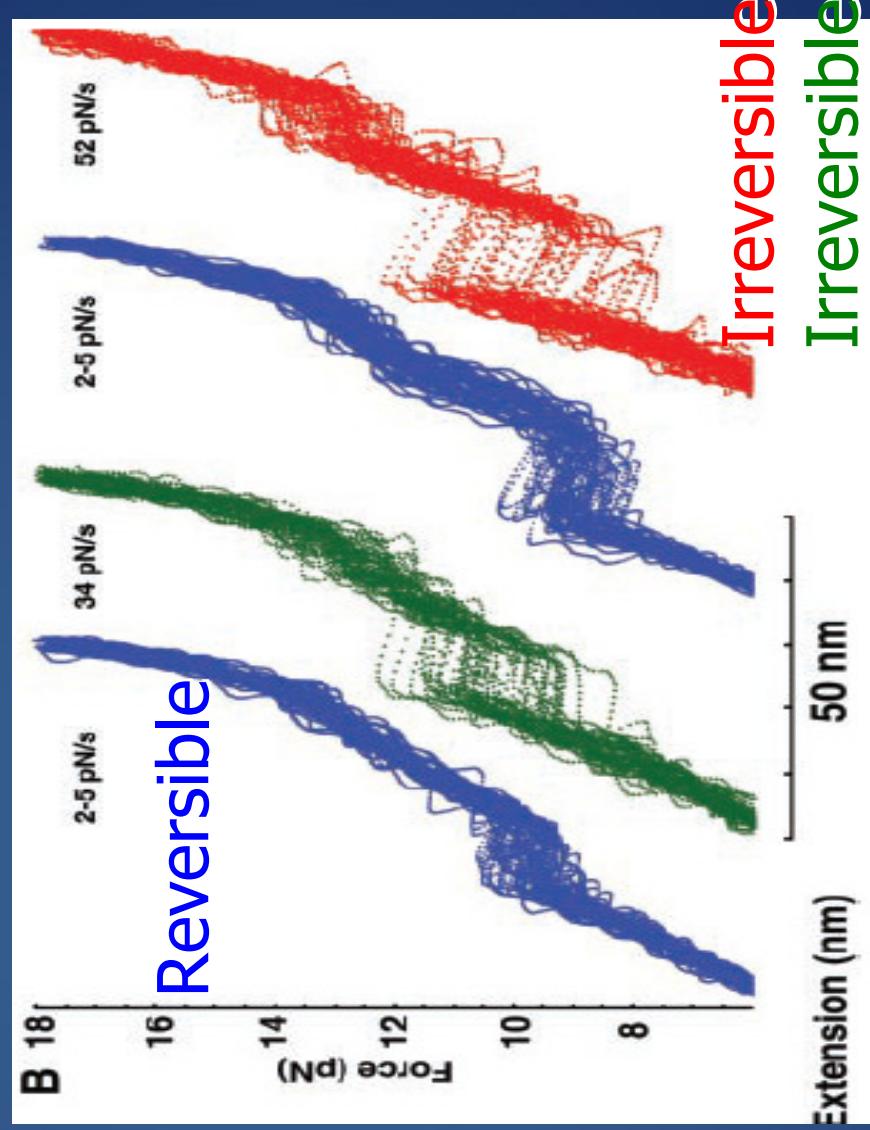
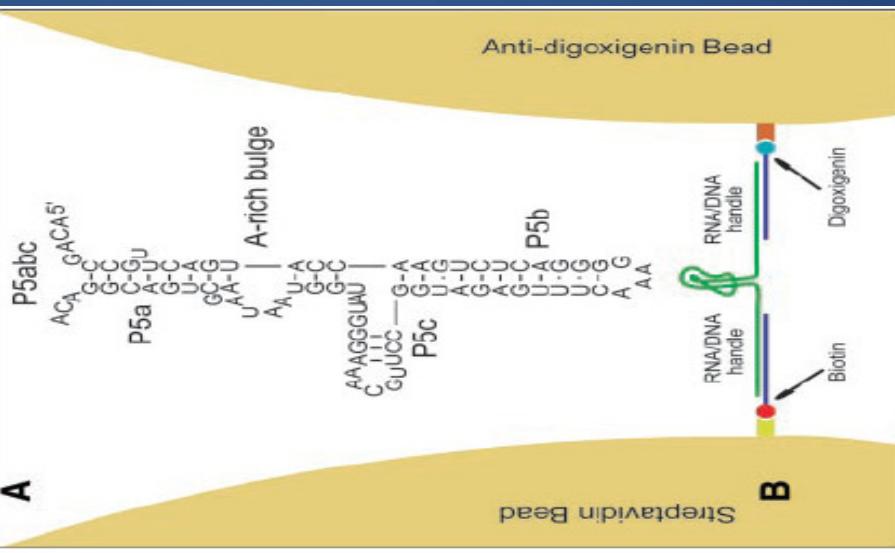
$$\langle \exp(-\beta W_z) \rangle$$



# Demostración Experimental. Desnaturalización de un RNA

Science. Bustamante 2002

## Resultados



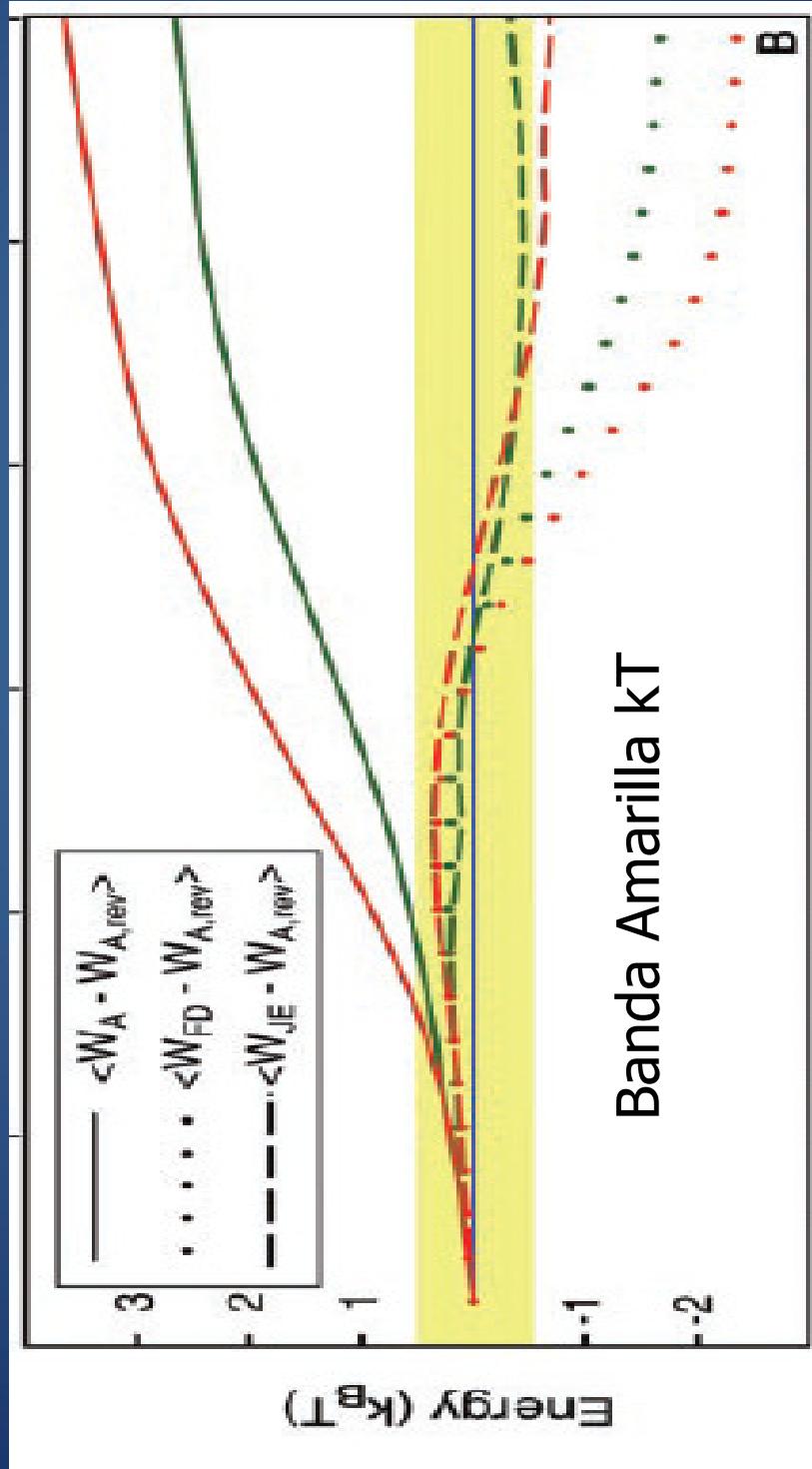
El Sistema

# Demostración Experimental. Desnaturalización de un RNA

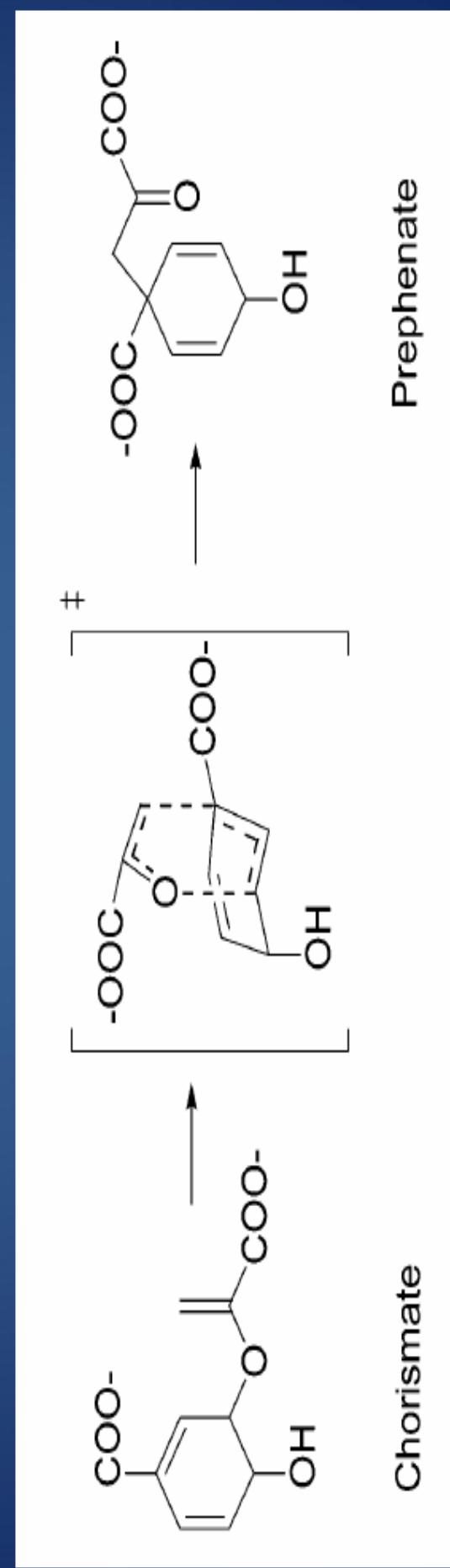
Science. Bustamante 2002

## Resultados

Solid  $\langle W \rangle$   
Dotted  $\langle W \rangle - \beta/2 \sigma^2$   
Dashed  $\langle \exp(-\beta Wz) \rangle$

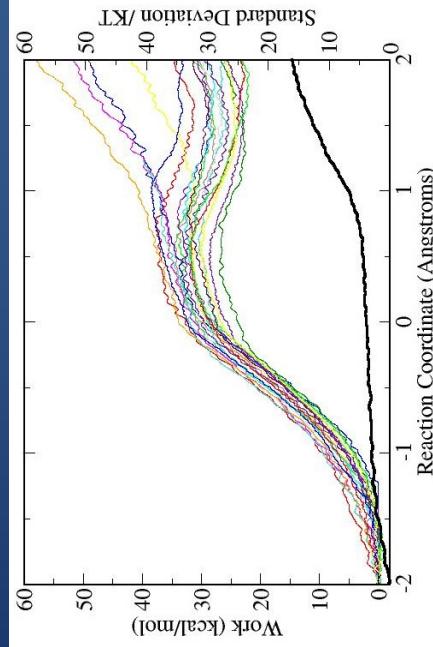


# Aplicaciones de JE: Calculo de Energía libre de reaccion, el caso de la Corismato Mutasa

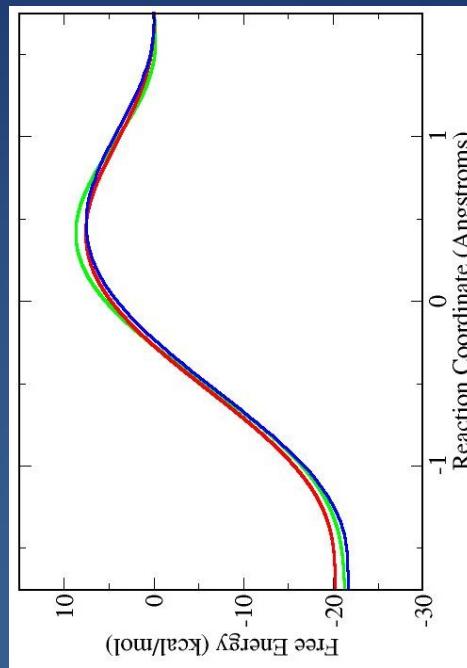


# Aplicaciones de JE: Calculo de Energía libre de reaccion, el caso de la Corismato Mutasa

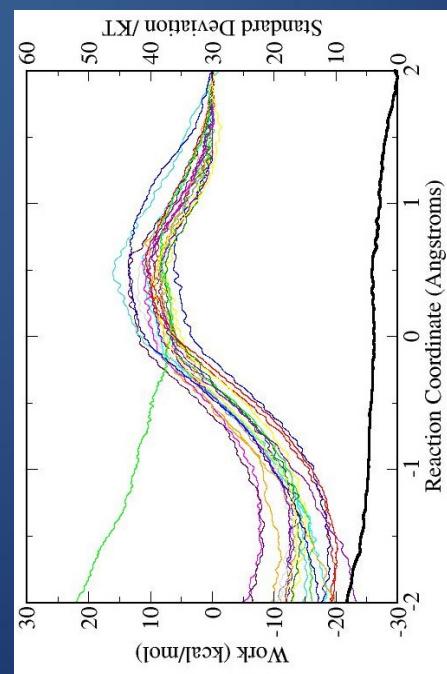
**Wz, forward reaction**



$-1/\beta \ln <\exp(-\beta W_z)>$



**Wz, back reaction**

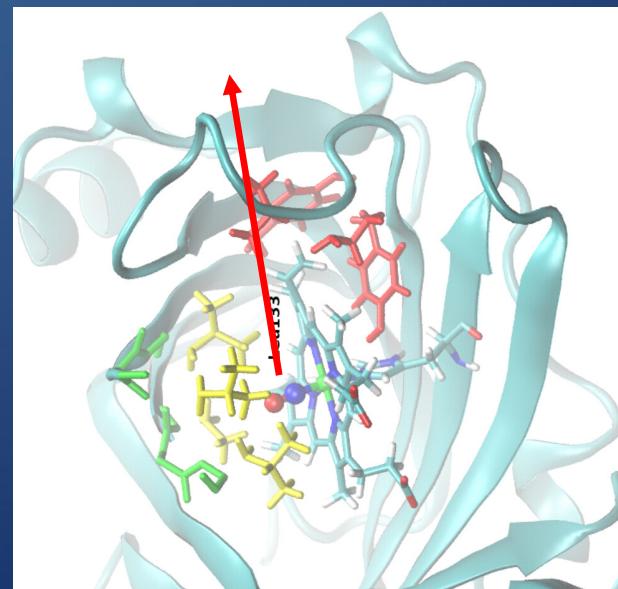


# Aplicaciones de JE: Migración de ligandos

## pequeños en Hemoproteínas

- Recordemos las nitroforina (proteínas que transportan NO)
- La liberación del NO se produce por un cambio conformacional pH dependiente.

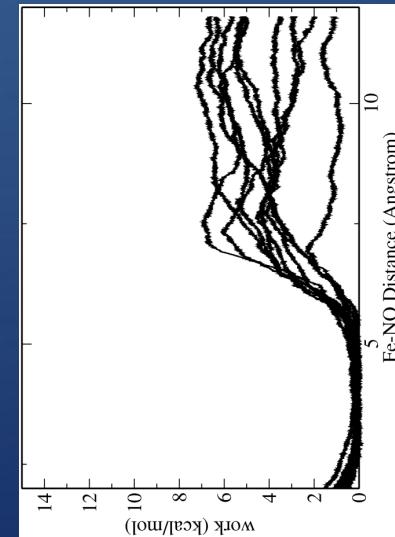
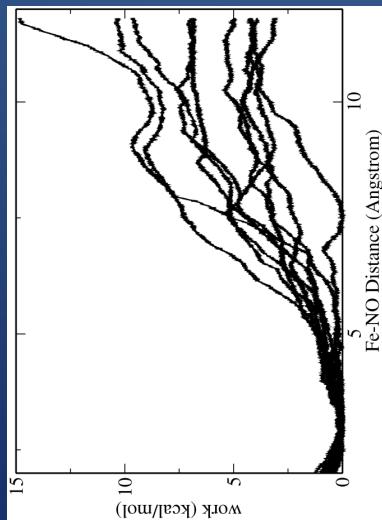
- En la saliva del insecto pH bajo (5.5) NO es retenido
- En la sangre pH alto (7.4) el NO es liberado
- La NP es una hemoproteína y el NO se une al grupo Hemo
- ¿Podemos medir la barrera de energía libre para la salida del NO en la NP4 en las conformaciones abierta y cerrada?



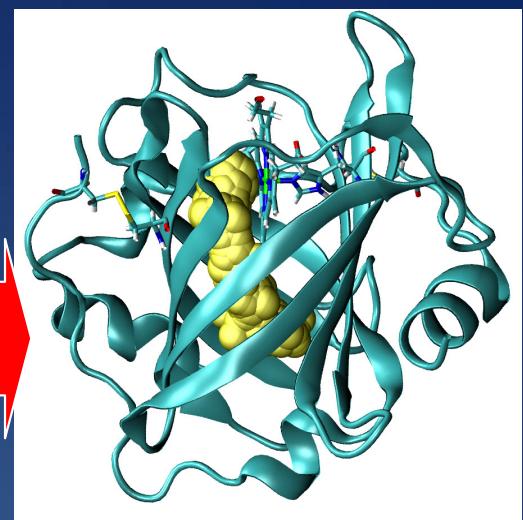
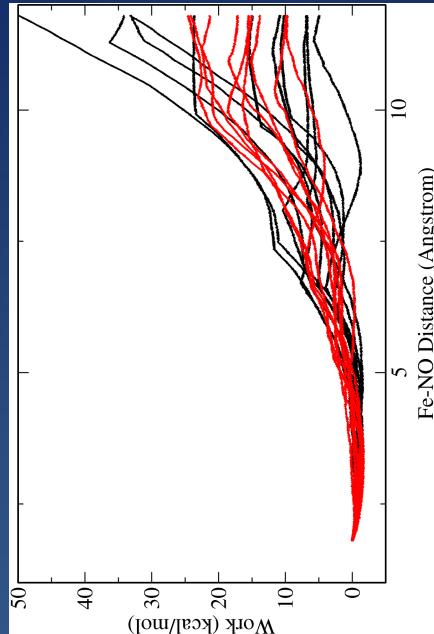
Utilizo como RC: distancia Fe-N  
RCini 2 Å => RCfin 12 Å  
Selecciono 20 conformaciones para cada  
estado de una MD inicial con NO unido (25ns)  
Velocidad: 10 Å/ns

# Aplicaciones de JE: Migración de ligandos pequeños en Hemoproteínas

pH 7.5, 2 sets de 10 MD y sus  $W_i$

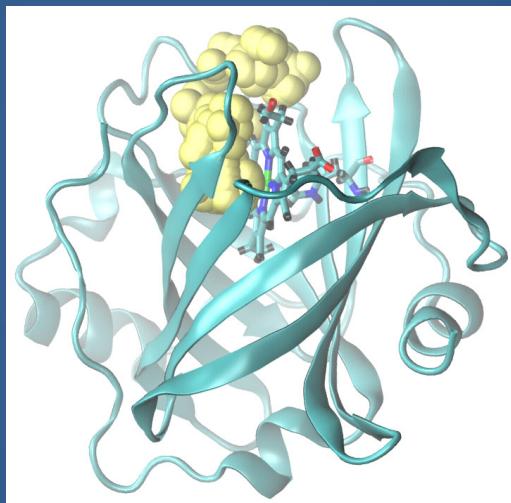


pH 5, 20 MDs y sus  $W_i$



/ El NO o sale  $W > 10$  kcal/mol  
O va hacia adentro  $5 < W < 20$  kcal/mol

/



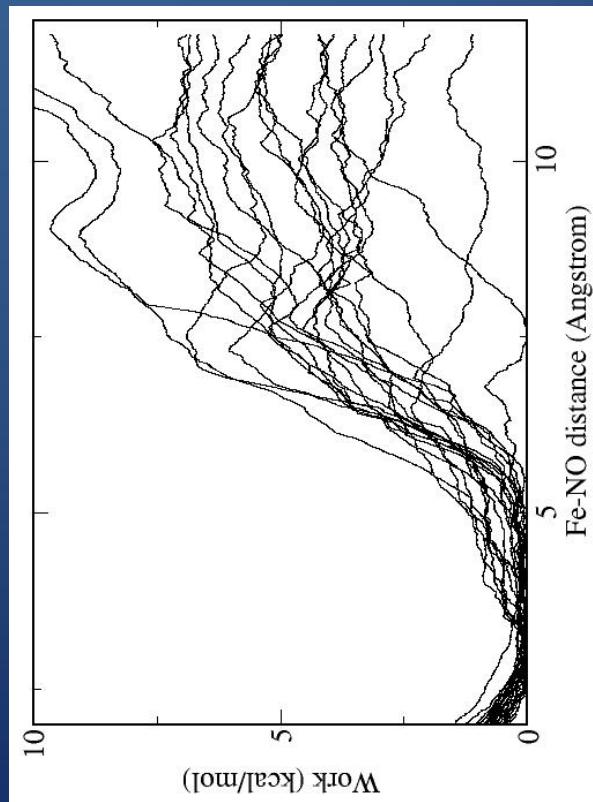
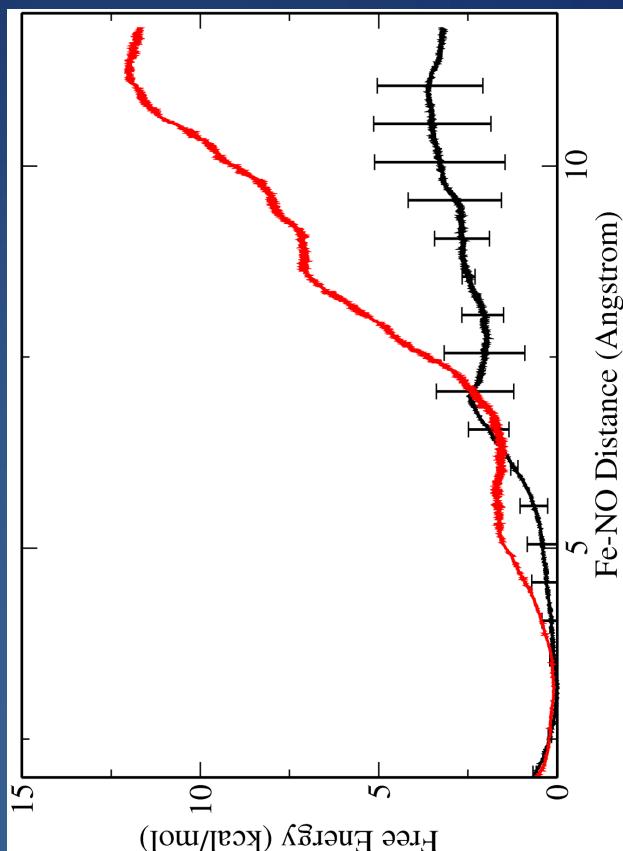
/

El NO siempre sale por el mismo camino,  
 $2 < W_i < 10$  kcal/mol

# Aplicaciones de JE: Migración de ligandos pequeños en Hemoproteínas

Wi pH 7.5

Perfil de energía libre resultante  
Rojo pH 5  
Negro pH 7.5

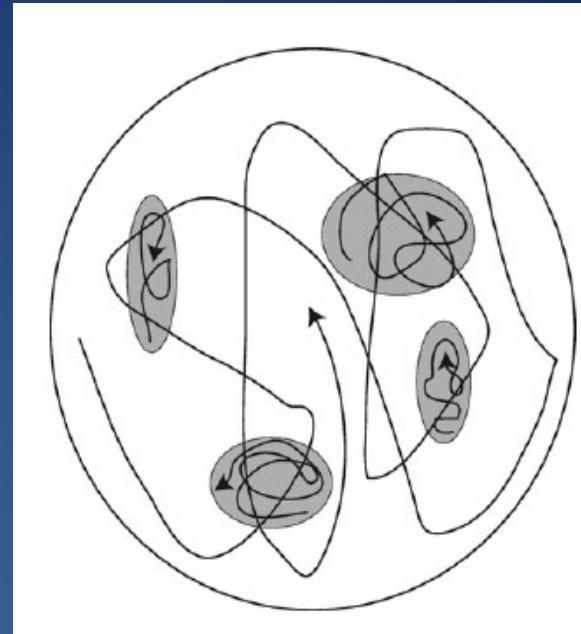


¿Cómo puedo mejorar el muestreo  
del espacio conformational?

Replica Exchange Molecular  
Dynamics

# El problema del muestreo

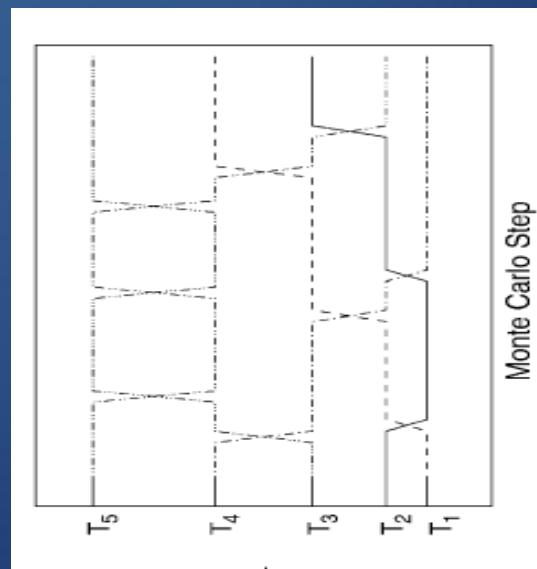
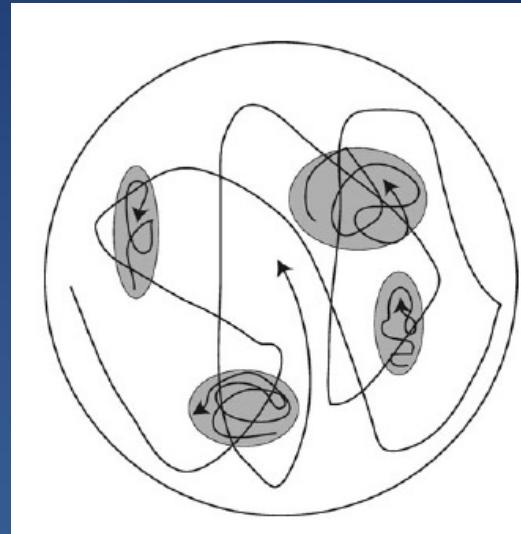
- En la MD convencional (all-atom) los sistemas usualmente quedan atrapados en los mínimos locales, debido a la rugosidad de la superficie (valles dentro de valles).
- ¿Cómo salir del mínimo local?
- Los tiempos simulables x MD son 10-100ns (quizá 1us), pero muchos procesos interesantes (plegamiento, cambios conformacionales u alóstéricos) ocurren en la escala de us a ms.
- ¿Cómo lograr estudiarlos por MD?



Respuestas:  
Utilizar representaciones simplificadas,  
Coarse-Grain  
Utilizar REMD.....

# Replica Exchange MD (REMD)

- REMD: el cada sistema es simulado en M-replicas (usualmente a distintas T,  $REMD(T)$ )
- Cada una se simula x MD convencional y periódicamente se intercambian las replicas de T vecinas de acuerdo a una probabilidad de transición  $P(X=>X')$
- En total cada replica explora un amplio rango de T, lo que se supone le permite cruzar barreras mas altas mas rápidamente. El sistema realiza un “random walk” en T, y por ende en el espacio configuracional



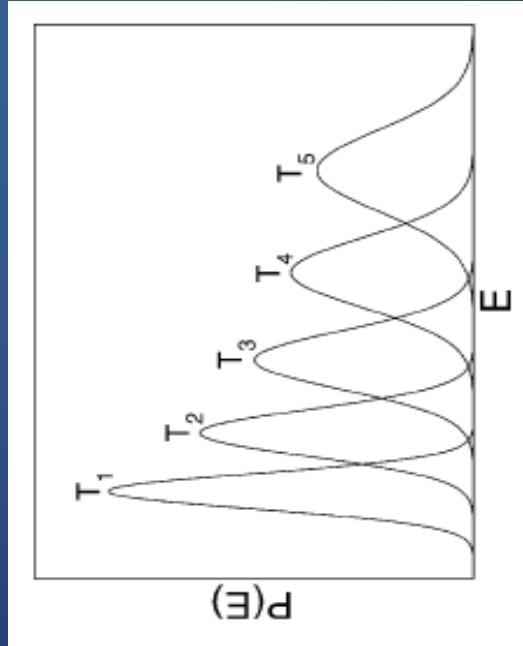
Monte Carlo Step

# Replica Exchange MD (REMD)

- $P(X_n \rightarrow X_m)$  Criterio de metrópolis (tipo monte carlo)
  - vale 1, si  $\Delta < 0$
  - vale  $\exp(-\Delta)$  si  $\Delta > 0$
  - Donde  $\Delta = (\beta_m - \beta_n)(E(X_m) - E(X_n))$
- Las velocidades se reescalan si hay intercambio según:
  - $v_{new} = v_{old} V(T_m/T_n)$
- En una simulación de REMD hay que verificar que:
  - el % de intercambios (Acceptance Ratio o AR) sea alto y uniforme
  - Tíene que haber una buena superposición en la distribución de energías entre las replicas.

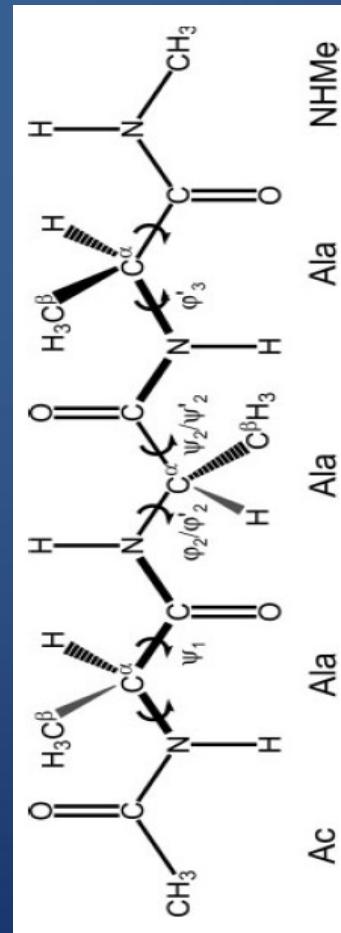
Table 2  
Acceptance ratios of replica exchange in REMI

Pair of temperatures (K)	Acceptance ratio
200 ↔ 239	0.15
239 ↔ 286	0.16
286 ↔ 342	0.15
342 ↔ 409	0.14
409 ↔ 489	0.13
489 ↔ 585	0.13
585 ↔ 700	0.13
700 ↔ 836	0.14
836 ↔ 1000	0.15



# Ejemplo REMD(T): Validación Parámetros dihedros Backbone de AMBER ff99SB

- Campos de fuerza clásicos la clave son los parámetros.
- Como obtener “testear” parámetros para los dihedros del Backbone.
- Tradicionalmente los CF son muy Hélicos
- Objetivo: Mejorar parámetros para el backbone AMBER
- Método de obtención: Comparación con ab-initio
- Métodos de “testeo”: Cálculo de REMD(T) del espacio de fases conformacional en tetrapetidos Ace-GLY<sub>3</sub>-NMe y Ace-ALA<sub>3</sub>-NMe



AMBER Backbone dihedrals:

Glicina SOLO:

$\Phi = \text{C-N-CA-C}$

$\Psi = \text{N-CA-C-N}$

Resto de los residuos  $\Phi/\Psi +$

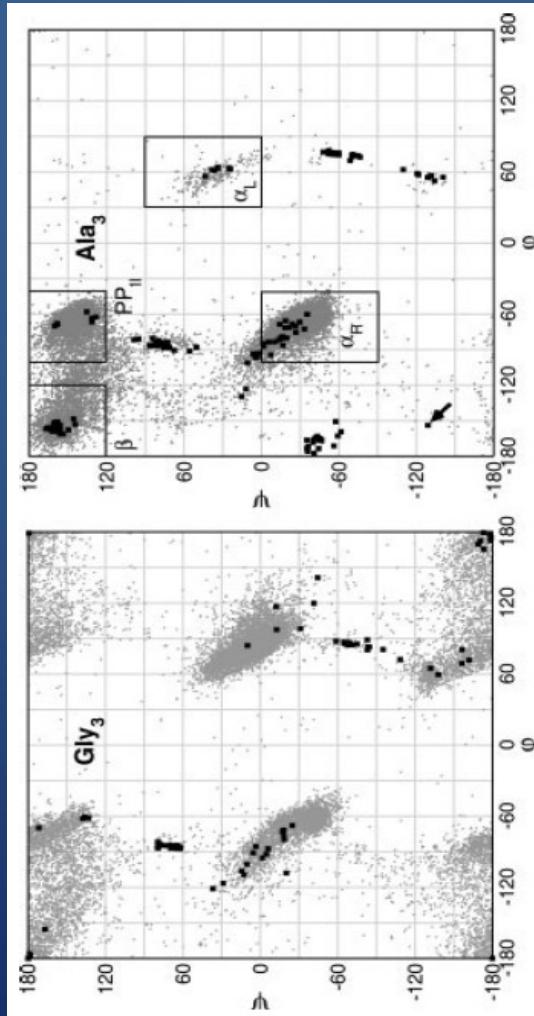
$\Phi' = \text{C-N-CA-CB}$

$\Psi' = \text{CB-CA-C-N}$

$$E_{\text{dihedral}}(\theta) = \sum_{n=1}^3 V_n (1 + \cos(n\theta - \gamma_n))$$

# Ejemplo REMD(T): Validación Parámetros dihedros Backbone de AMBER ff99SB

Obtención de los parámetros: búsqueda conformacional QM (HF/631-G\*+LMP2)  
28 confórmeros para Gly<sub>3</sub> y 51 para Ala<sub>3</sub> (puntos negros)



Se busca minimizar

$$\text{AAE} = 1/N(N-1) \sum_{i=1}^N \sum_{j>i}^N |E_{\text{QM}}^{o,i}(j) - E_{\text{MM}}^{o,i}(j)|$$

$$\text{MAE} = \max_{i,j>i} |E_{\text{QM}}^{o,i}(j) - E_{\text{MM}}^{o,i}(j)|$$

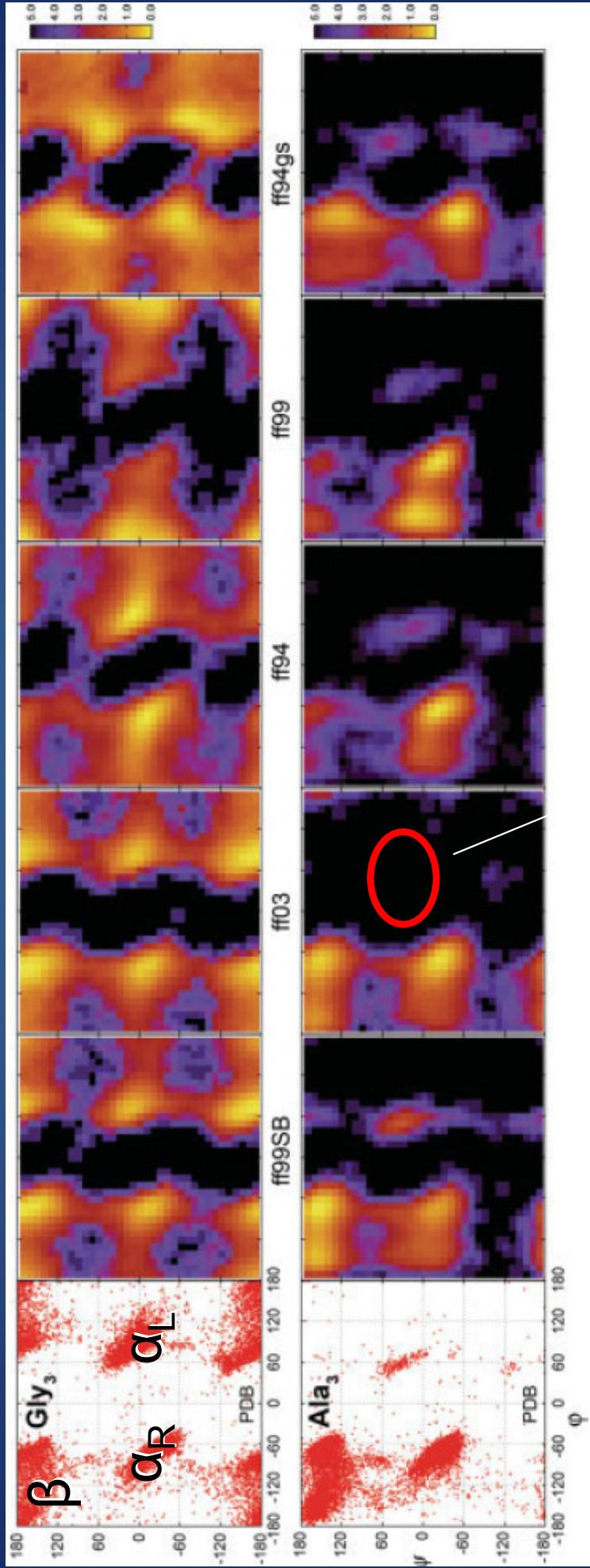
	V <sub>1</sub>	γ <sub>1</sub>	V <sub>2</sub>	γ <sub>2</sub>	V <sub>3</sub>	γ <sub>3</sub>
ϕ	C-N-C <sup>α</sup> -C	0.00	0	0.27	0	0.42
ψ	N-C <sup>α</sup> -C-N	0.45	π	1.58	π	0.55
ϕ	C-N-C <sup>α</sup> -C <sup>β</sup>	2.00	0	2.00	0	0.40
ψ	N-C-C <sup>α</sup> -C <sup>β</sup>	0.20	0	0.20	0	0.40

## Resultados:

# Ejemplo REMD( $\tau$ ): Validación Parámetros

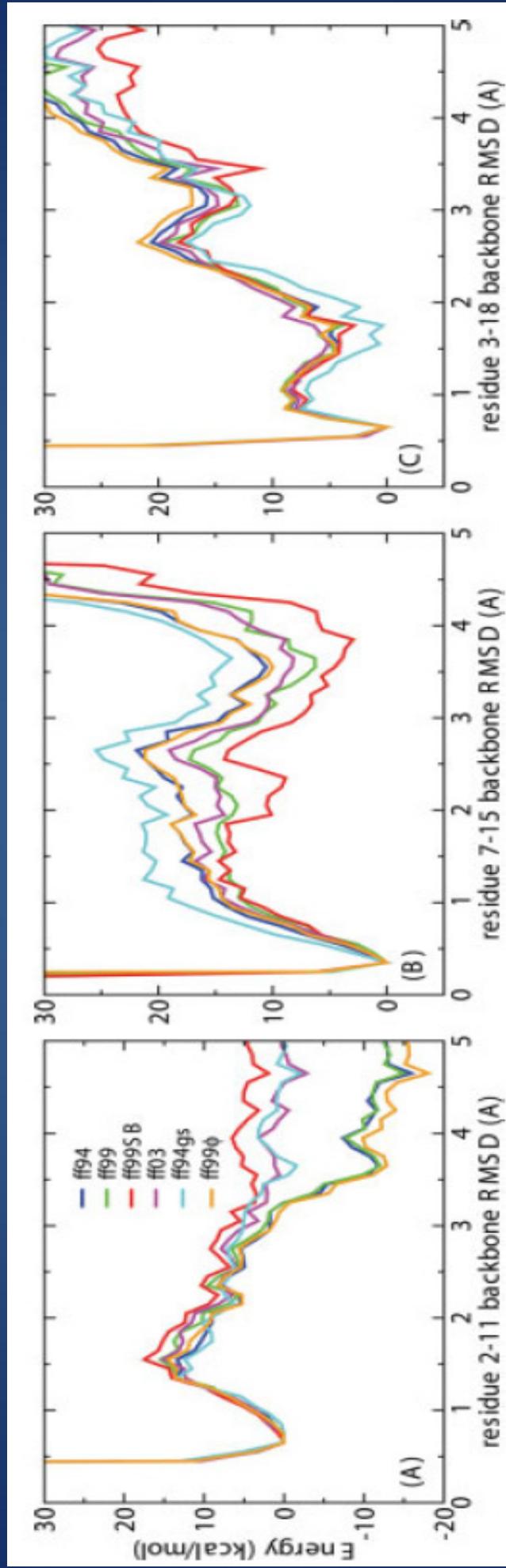
dihedros Backbone de AMBER ff99SB

REMD( $\tau$ ): 26 replicas, 267-571K, AR 30%



# Ejemplo REMD( $T$ ): Validación Parámetros dihedros Backbone de AMBER ff99SB

REMD( $T$ ): 26 replicas, 267-571K, AR 30%



B-Hairpin

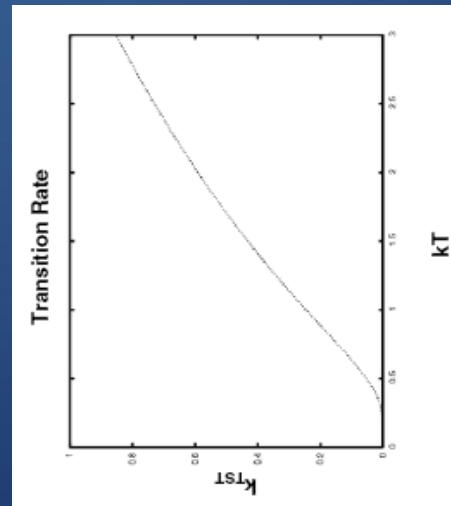
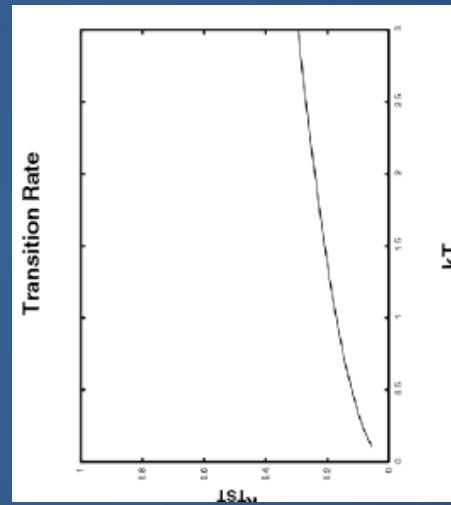
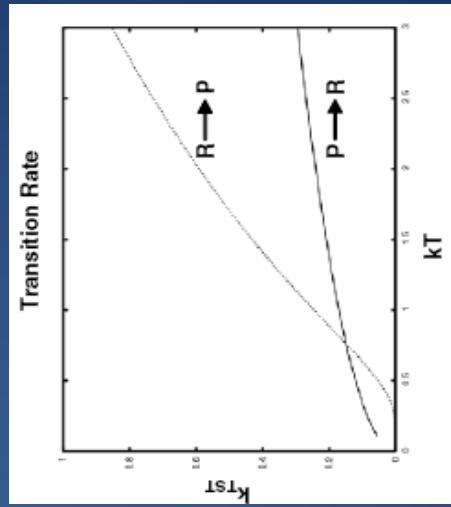
$\alpha$ -Helix

Trp-Cage

Rojo ff99SB, Verde ff99

# Contras REMD(T)

- El problema pbal de REMD son que M escala rápidamente con los grados de libertad del sistema, según:
  - $M = \nu(f) \ln(T_{\max}/T_{\min})$ 
    - si queremos llegar a altas T y sistemas grandes M aumenta rápidamente.
  - Al aumentar T la velocidad de cruce de barreras NO es homogéneo. Las barreras con contenido enáltico se cruzan mas rápido que las que poseen contenido entrópico.
    - En “folding” se acelera velocidad de desplegamiento MAS que la de Plegamiento.



$$\Delta E_{act} \approx \Delta S_{act}$$
$$\Delta E > 0 \text{ y } \Delta S > 0$$

$$\Delta E_{act} \ll \Delta S_{act}$$
$$\Delta E = \Delta S = 0$$

$$\Delta E_{act} \gg \Delta S_{act}$$
$$\Delta E = \Delta S = 0$$

# REMD( $H$ )

- El formalismo es completamente análogo a REMD( $T$ ) pero ahora  $P(X_n \Rightarrow X_m)$ , depende del hamiltoniano  $H_i$ , y la conformación  $X_i$  con:

$$\Delta = \beta \{ [H_m(X_m) + H_n(X_n)] - [H_m(X_n) + H_n(X_m)] \}$$

- Esto permite:
  - Reducir los GL relevantes del sistema (y por ende el # de replicas)
  - Cambiar el Hamiltoniano Ej: Combinar All-atom y CG
- Nótese que REMD( $T$ ) es un caso particular de REMD( $H$ )  
Donde:  $E_m = S_m E_n$  (factor de corrección de  $T$ )  
 $\exp(-\beta_m E) \Leftrightarrow \exp(-\beta S_m E)$

# REMD( $H$ ) para reducir “f”

- Hay que dividir el sistema en GL relevantes (Ej: la proteína ( $r$ )) y GL irrelevantes (Ej: el solvente ( $x$ ))
- Si escribo  $H$  según:  
$$E(r,x) = E(r) + E(r,x) + E(x)$$
Asumiendo que el acoplamiento  $E(r,x)$  es pequeño.

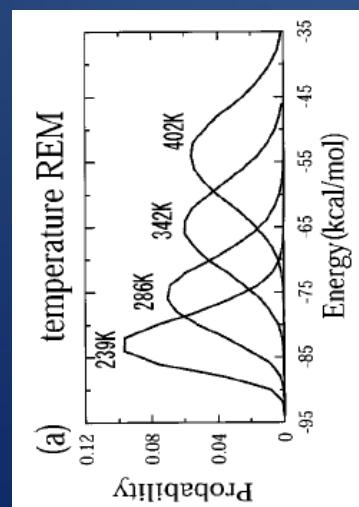
- Puedo definir las replicas en  $\neq T$  de acuerdo a:  
$$E_m(r,x) = S_m E(r) + S_m E(r,x) + E(x)$$

Nótese que al escalar solo los GL “ $r$ ” el # de GL relevantes es menor y se requieren menos replicas para obtener una simulación adecuada.

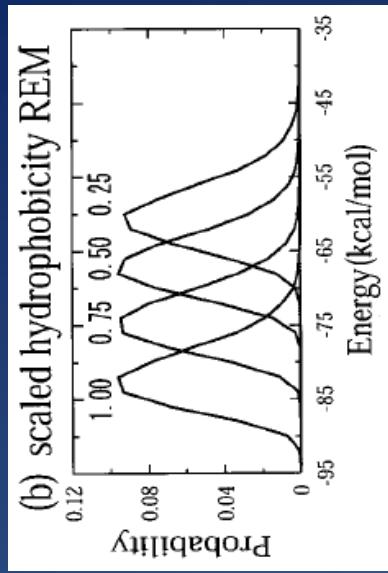
# REMD( $H$ ): cambiando la hidrofobicidad del medio

- Puedo hacer replicas con diferente hydrofobicidad del medio.
  - En Solvente ( $S_v$ ) implícito puedo:
    - mover la constante dieléctrica) o..
    - si utilizo una función efectiva para la energía de solvatación, esta usualmente se escribe según:
- $$E_{Sv} = -\sum \gamma_{SASA} q_i$$
- La  $E_{Sv}$  depende de la superficie expuesta( $\gamma_{SASA}$ ) y de la carga del átomos ( $q_i$ )
- Las replicas se pueden escalar según:  
 $E_{Sv_m} = S_m E_{Sv}$   
Donde  $S=1$ , tengo medio fisiológico, y si  $S_m < 1$  medio mas hidrofóbico.
  - Este método puede pensarse como una analogía de lo que hacen las chaperonas in-vivo que presentan un medio con zonas hidrof/hidrofil para facilitar el plegamiento

# REMD( $H$ ): cambiando la hidrofobicidad del medio



REMD( $T$ ) vs REMD( $H$ )



# Integración Termodinámica

# Integración Termodinámica

- Que pasa si quiero medir el  $\Delta A$  de una transformacion quimica o de algun otro proceso que NO conozco el camino ej  $\text{CH}_4 \Rightarrow \text{CH}_3\text{Cl}$  en agua?
- Cómo obtengo el  $\Delta A$  del proceso?
- Supongamos que  $A$  es función continua de algún parametro externo “ $\lambda$ ”
- Si quiero medir el  $\Delta A$  del estado A al estado B entonces:  
$$\Delta A_{A \rightarrow B} = \int \delta A(\lambda) / \delta \lambda \, d\lambda$$

# ¿Cómo mido $\Delta A_{A \Rightarrow B} = \int \delta A(\lambda) / \delta \lambda \, d\lambda$ , en una simulación?

- Recordemos:  $A(\lambda) = -kT \ln Q(\lambda)$  ... entonces..
- $\Delta A = \int 1/Q \, \delta Q(\lambda) / \delta \lambda \, d\lambda$ 
  - Con:  $Q = \langle e^{-E/kT} \rangle$ , reemplazo:
- $\delta Q / \delta \lambda = 1/kT \langle \delta E / \delta \lambda \exp{-E/kT} \rangle$
- Reemplazando para  $\Delta A$ :
- $\Delta A = - \int \langle \delta E / \delta \lambda \rangle d\lambda,$ 
  - Y esto es lo que hay que evaluar...

$$\Delta A = - \int <\delta E / \delta \lambda> d\lambda$$

- En palabras: hay que determinar la integral entre ambos estados, del valor medio de la diferencia de energía (entre ambos estados) respecto de  $\lambda$ .
- Para determinar  $\delta E / \delta \lambda$  entre los sistemas A y B se define:
  - $E(\lambda) = E_A \lambda + (1-\lambda)E_B$
  - $\delta E / \delta \lambda = E_A - E_B$
- Entonces:
  - por lo que el valor medio sale directo de la simulacion.
  - La integral se determina realizando distintas simulaciones (Ventanas) para distintos valores de  $\lambda$  (entre 0 y 1) y se calcula el valor medio de  $\delta E / \delta \lambda$  en función de  $\lambda$ .
  - Luego se calcula la integral numérica o graficamente.

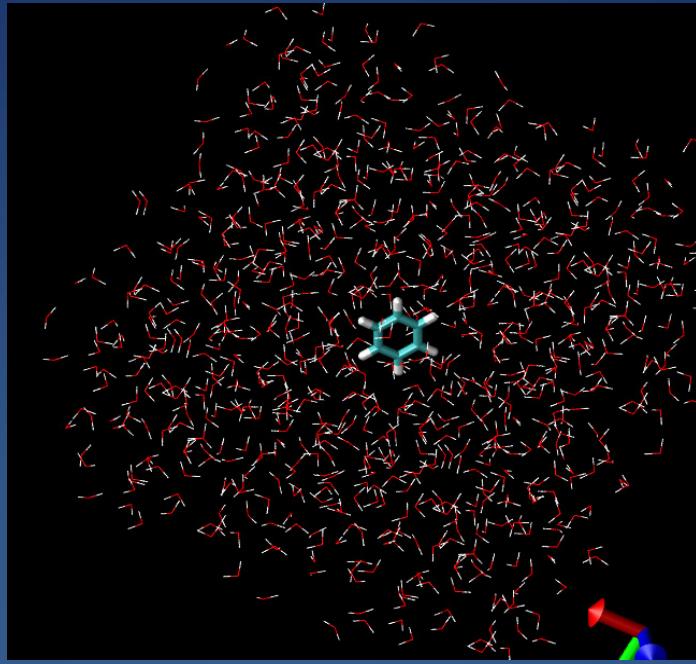
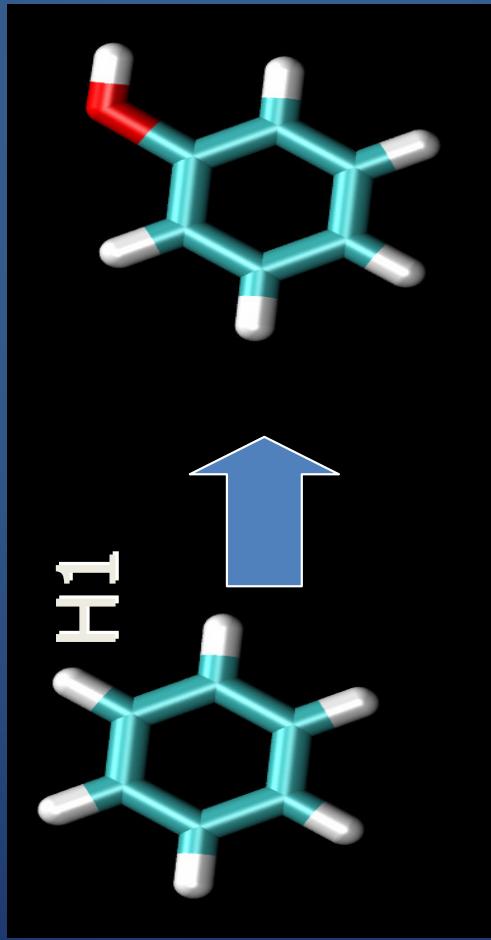
# En la práctica

- Simulo el sistema con  $H = E(\lambda)$
- Si simulo el sistema A la energía es  $E_A$ , si simulo B la energía es  $E_B$ ..
- Pero si simulo ambos sistemas ACOPLADOS por  $\lambda$   
$$E(\lambda) = E_A \lambda + (1-\lambda) E_B$$
 y  
$$F = -\delta E(\lambda) / \delta r$$
- La dinámica es guiada por la combinación de hamiltonianos
- Ambos sistemas A y B comparten las coordenadas atómicas y la energía de cada uno se calcula de manera usual.
- En la práctica construyo DOS hamiltonianos  $H_A$  y  $H_B$  pero hago una SOLA dinámica o seas las fuerzas que guían A y B son las mismas y se calculan como  $F = -\delta E(\lambda) / \delta r$

Ejemplo: Conversion de:

Benzeno (A) => Fenol (B) en Agua

Los dos sistemas se simulan acoplados, tienen las mismas coordenadas para todos los atomos del anillo y el solvente. Pero en A tengo un H, y en B un OH

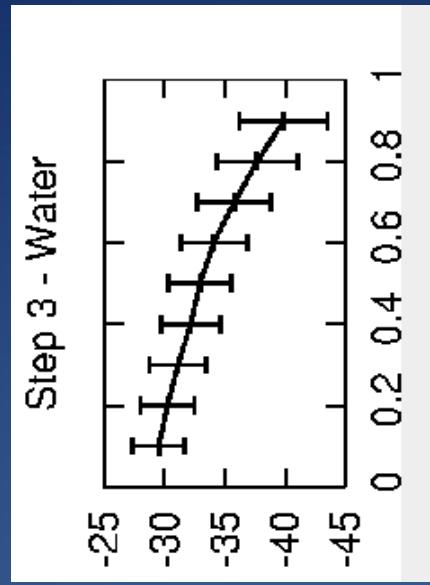
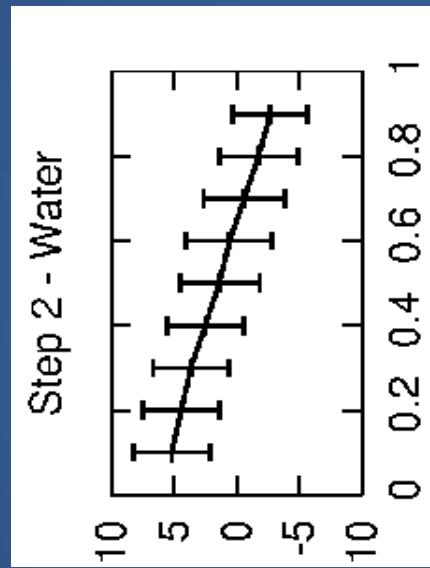
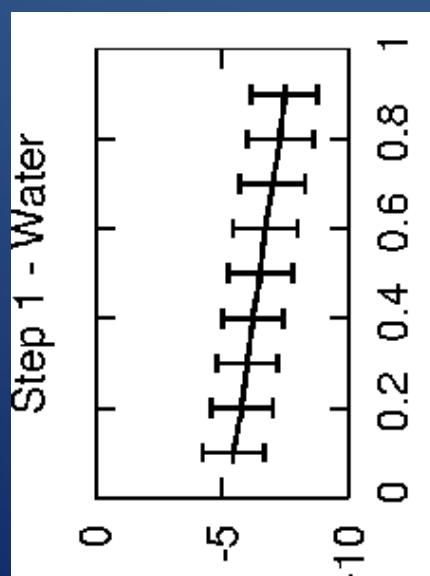


El proceso se hace en tres pasos  
1ro: Convierbo la carga de H1 a cero  
2do: Cambio el LJ de H1 => LJ de OH  
3ro: Agrego carga al OH

Ejemplo: Conversion de:

Benzeno (A) => Fenol (B) en Agua

Para cada uno: se simulan 9 ventanas con:  $\lambda=0,1$  a  $\lambda=0,9$



$$\Delta A_1 = -6.5$$

$$\Delta A_2 = 1.4$$

$$\Delta A_3 = -33.8$$

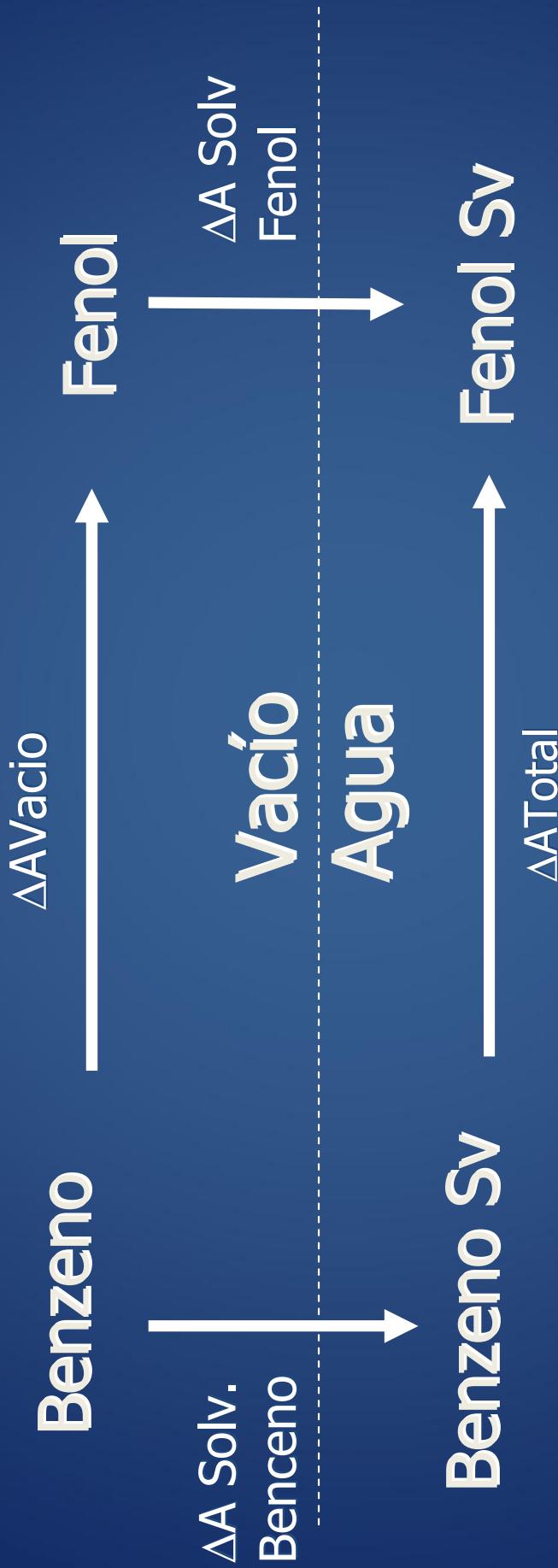
$$\Delta A_{\text{TOTAL}} = \Delta A_1 + \Delta A_2 + \Delta A_3 = -38.9 \text{ kcal/mol.}$$

Es este el cambio en energía libre "real"? (se cerro el ciclo termodinamico?)

Que valor experimental representa?

Ejemplo: Conversion de:

Benzeno (A) => Fenol (B) en Agua



$$\Delta A_{\text{exp}} = \Delta S_v \text{Benceno} - \Delta S_v \text{Fenol} = \Delta A_{\text{Total}} - \Delta A_{\text{Vacío}}$$

SIEMPRE HAY QUE RESTAR EL " $\Delta A$ " del MISMO PROCESO EN VACÍO  
PARA CORREGIR POR AUTOINTERACCION DEL CAMPO DE FUERZAS

Para que se usa....

- Cambios químicos en solvente explícito
- Cambios en la energía de solvatación
- ΔΔA de Unión entre dos Drogas diferentes
- Cambio de energía debido a mutaciones en proteínas..
- ...cualquier proceso para el que pueda acopiar dos Hamiltonianos: A y B

Muchas Gracias por la Atención

Dudas?

Preguntas?

Inquietudes?