

PREDIZIONE DELLA STRUTTURA: Dinamica Molecolare

Prof. Ranieri Urbani

Calcola la TRAIETTORIA di un sistema molecolare =
la configurazione molecolare in funzione del tempo, ovvero come
variano nel tempo le posizioni, le velocità e le accelerazioni degli
atomi della molecola.

**Corso di
CHIMICA DELLE MACROMOLECOLE I
a.a. 2021-2022**

Dinamica molecolare

Permette lo studio di processi dinamici complessi che avvengono nei sistemi biologici. Studia sia transizioni conformazionali che vibrazioni locali, ad esempio:

stabilità dei polimeri

variazioni conformazionali

Ripiegamento dei polimeri

Interazioni polimero-ligando

Drug Design

Determinazione di strutture 3D di macromolecole biologiche

.....

Tecnica di simulazione che permette lo studio del moto e delle proprietà di un sistema di particelle.

- **Moti localizzati** (da 0.01 a 5 Å, da 10^{-15} a 10^{-1} s)
 - fluttuazioni atomiche
 - movimenti di catene laterali
 - movimenti di loop
- **Moti a corpo rigido** (da 1 a 10 Å, da 10^{-9} a 1 s)
 - movimenti di eliche
 - movimenti di domini
 - movimenti di subunità
- **Moti su larga scala** (> 5 Å, da 10^{-7} a 10^{-3} s)
 - transizioni eliche-coil
 - associazione/dissociazione
 - folding/unfolding

Dinamica molecolare

A T ambiente si dimostra che vale l'approssimazione della **meccanica classica**, purchè i moti abbiano **tempi caratteristici del ps (10^{-12} s)** o siano più lunghi.

Una struttura polimerica viene approssimata ad un network di sfere (particelle di Lennard-Jones) con carica puntiforme al loro centro, collegate da molle (legami covalenti).

Le particelle si trovano in un **campo di energia potenziale** che si modifica in funzione delle posizioni raggiunte dagli atomi del sistema.

Dinamica Molecolare

Calcola la **TRAIETTORIA** ossia le variazioni nel tempo delle posizioni, delle velocità e delle accelerazioni degli atomi della molecola.

Il sistema viene trattato secondo la legge del moto newtoniano

Le simulazioni di dinamica molecolare si basano sulla seconda legge di Newton:

$$\vec{F}_i = m_i \vec{a}_i \quad i = 1, \dots, N \quad (\text{numero atomi del sistema})$$

La traiettoria è generata da integrazioni simultanee dell' equazione del moto di Newton

$$\mathbf{F}_i = \mathbf{m}_i \cdot \mathbf{a}_i$$

per tutti gli atomi del sistema molecolare

Tenendo presente che la forza si può esprimere come **gradiente dell'energia potenziale**:

$$\mathbf{F}_i = - dV/d\mathbf{r}_i$$

si combinano le due equazioni e si ottiene:

$$- dV/d\mathbf{r}_i = \mathbf{m}_i d^2\mathbf{r}_i/dt^2$$

che collega la derivata dell'energia potenziale alle variazioni di posizione in funzione del tempo. E' questa che viene **integrata**.

Per calcolare una **traiettoria** c'è bisogno:

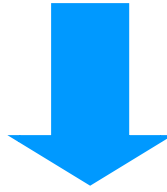
1. delle posizioni iniziali degli atomi (coordinate atomiche)
2. delle velocità iniziali
3. delle accelerazioni

1. le posizioni iniziali r_i si ricavano da strutture sperimentali (cristallografia raggi X, NMR ecc.) o ottenute con modeling;

2. le velocità iniziali v_i si ottengono dalla distribuzione di Boltzmann delle velocità ad una data temperatura;

3. le accelerazioni sono determinate dal gradiente dell'energia potenziale.

$$F_i = m_i a_i = - dV/dr_i$$



Le posizioni e le velocità iniziali ($t = 0$) determinano le posizioni e le velocità a tutti gli altri tempi t .

In pratica si considerano intervalli di integrazione finiti Δt .

Δt tipicamente va da 0.1 a 10 fs per i sistemi molecolari

→ una simulazione di 100 ps coinvolge 10^5 - 10^6 intervalli di integrazione.

Dinamica molecolare

L'integrazione delle equazioni del moto consente di ottenere una **traiettoria** che descrive la variazione nel tempo delle posizioni, velocità ed accelerazioni delle particelle. Le simulazioni di dinamica molecolare sono in genere calcolate sulla scala dei **ns** (10^{-9} s).

$$\vec{F}_i = m_i \vec{a}_i \quad i = 1, \dots, N \quad (\text{numero atomi del sistema})$$

$$\vec{F}_i = -\vec{\nabla}_i V(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = -\left(\frac{\partial V}{\partial x_i}, \frac{\partial V}{\partial y_i}, \frac{\partial V}{\partial z_i}\right) \quad \text{dove } V \text{ è energia potenziale}$$

Per moti uniformemente accelerati ($a = \text{costante}$):

$$\vec{a}_i = d\vec{v}_i/dt \quad \Rightarrow \quad \vec{v}_i(t) = \vec{a}_i t + \vec{v}_i(t_0)$$

ma

$$\vec{v}_i = d\vec{r}_i/dt \quad \Rightarrow \quad \vec{r}_i(t) = \vec{v}_i t + \vec{r}_i(t_0)$$

quindi

$$\vec{r}_i(t) = \frac{1}{2} \vec{a}_i t^2 + \vec{v}_i(t_0) t + \vec{r}_i(t_0) \quad \text{dove} \quad \vec{a}_i = - (1/m_i) \vec{\nabla}_i V$$

Funzione Energia Potenziale

Per calcoli di dinamica molecolare su sistemi biologici vengono utilizzate funzioni energia potenziale empiriche capaci di riprodurre le proprietà fisiche del sistema fornendo un buon compromesso fra accuratezza ed efficienza computazionale.

Limitazioni: tali funzioni energia potenziale non possono riprodurre drastici cambi nella struttura elettronica degli atomi del sistema. Per esempio non possono essere modellati eventi come la formazione o la rottura di legami covalenti (in questo caso occorre un approccio quanto-meccanico).

Funzione Energia Potenziale

Per velocizzare il calcolo del potenziale, si possono introdurre dei cutoff per le interazioni di non legame (per esempio, cutoff di 12 Å per le interazioni di Coulomb).

Manca un termine esplicito per le interazioni di legame a idrogeno.

Per una dinamica realistica bisognerebbe esplicitare il solvente (una o due shell di idratazione della macromolecola), che scherma parzialmente le cariche q_i e q_j .

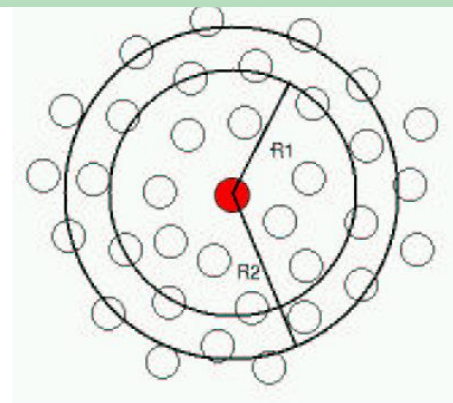


Figura 4.1: Nella figura viene mostrato l'uso del metodo del doppio cutoff (*twin range cutoff*). La particella centrale interagisce sempre con le particelle che si trovano dentro il raggio ($R1$) mentre con quelle entro il raggio $R2$ ogni n passi.

Dinamica molecolare

L'integrazione delle equazioni del moto non è banale, perché il moto degli atomi modifica i valori delle energie potenziali di interazione e quindi delle forze agenti sugli atomi stessi.

L'integrazione delle equazioni del moto viene fatta per intervalli di tempo molto piccoli ($1 \text{ fs} = 10^{-15} \text{ s}$) in modo tale da mantenere costanti le quantità fisiche rilevanti.

La lunghezza totale di una simulazione è dell'ordine dei ns (10^{-9} s).

Gli algoritmi matematici utilizzati per l'integrazione sono:

- Leap frog
- Verlet

Algoritmo di Verlet

Il potenziale U è funzione delle posizioni di tutti gli atomi del sistema

E' a causa della complessità di tale funzione che le equazioni del moto devono essere risolte **numericamente**

E' necessario discretizzare le equazioni, ovvero passare ad una descrizione in cui il tempo diventa una variabile discreta (δt).

Scegliamo un intervallo di tempo (time step) δt piccolo ed attraverso lo sviluppo in serie di Taylor scriviamo (per una particella):

$$\begin{aligned} \mathbf{r}(t + \delta t) &= \mathbf{r}(t) + \mathbf{v}(t)\delta t + \frac{1}{2}\mathbf{a}(t)\delta t^2 + \frac{1}{6}\dot{\mathbf{a}}(t)\delta t^3 \cdot \\ \mathbf{r}(t - \delta t) &= \mathbf{r}(t) - \mathbf{v}(t)\delta t + \frac{1}{2}\mathbf{a}(t)\delta t^2 - \frac{1}{6}\dot{\mathbf{a}}(t)\delta t^3 \cdot \end{aligned}$$

Sommando e sottraendo le due relazioni e ricordando $F = ma$ otteniamo *l'algoritmo di Verlet*:

$$\begin{aligned} \mathbf{r}(t + \delta t) &= \mathbf{r}(t - \delta t) + 2\mathbf{r}(t) + \frac{\mathbf{F}(t)}{m}\delta t^2 \\ \mathbf{v}(t) &= \frac{\mathbf{r}(t + \delta t) - \mathbf{r}(t - \delta t)}{2\delta t} + \mathcal{O}(\delta) \end{aligned}$$



in pratica:

- Si specificano le posizioni a $t-dt$ e t .
- Dalle posizioni si calcolano le forze al tempo t
- Si calcolano le posizioni al tempo $t+dt$.
- Si calcolano le velocità

Un altro algoritmo molto usato è il **leap-frog** (a salto di rana):

$$\begin{aligned}\mathbf{r}(t + \delta t) &= \mathbf{r}(t) + \mathbf{v}(t + \frac{\delta t}{2})\delta t \\ \mathbf{v}(t + \frac{\delta t}{2}) &= \mathbf{v}(t - \frac{\delta t}{2}) + \frac{\mathbf{F}(t)}{m}\delta t,\end{aligned}$$

così chiamato perchè velocità e posizione non sono calcolati allo stesso tempo ma **"saltellano"** fra tempi sfalsati di **$t/2$** .

Il controllo della T

- la simulazione dinamica coinvolge tecniche di valutazione e di controllo della T che si realizzano in genere utilizzando dei **termostati (virtuali)** opportuni.
- La temperatura è una grandezza che traduce a livello macroscopico la situazione a livello atomico mediante il termine dell'energia cinetica
- Tale relazione passa attraverso il valore della velocità degli atomi (**distribuzione di Maxwell-Boltzmann**)

$$p(\mathbf{v})d\mathbf{v} = \left(\frac{m}{2\pi kT} \right)^{3/2} e^{-\frac{m\mathbf{v}^2}{2kT}} 4\pi \mathbf{v}^2 d\mathbf{v}$$

La regolazione della T

- Le velocità assegnate come valori iniziali del processo di integrazione vengono determinate in modo tale che **sia soddisfatta la distribuzione di Maxwell-Boltzmann una volta scelto il valore T**
- Durante l'esperimento virtuale tale **distribuzione muta nel tempo** mano a mano che il calcolo procede.
- Dovendo T essere **costante**, il sistema preposto alla termostatazione deve configurarsi come una tecnica che permetta di generare la corretta configurazione statistica.

La regolazione della T

- **Metodi di Termostatazione:**
- Si basano sul concetto di “bagno di termostatazione virtuale”(Berendsen , Langevin)



- Valutano la **differenza tra T impostata e T calcolata** dall'energia cinetica. Questa differenza riscalata in funzione di un parametro di rilassamento permette di calcolare un fattore correttivo delle velocità (intesa come distribuzione di probabilità)

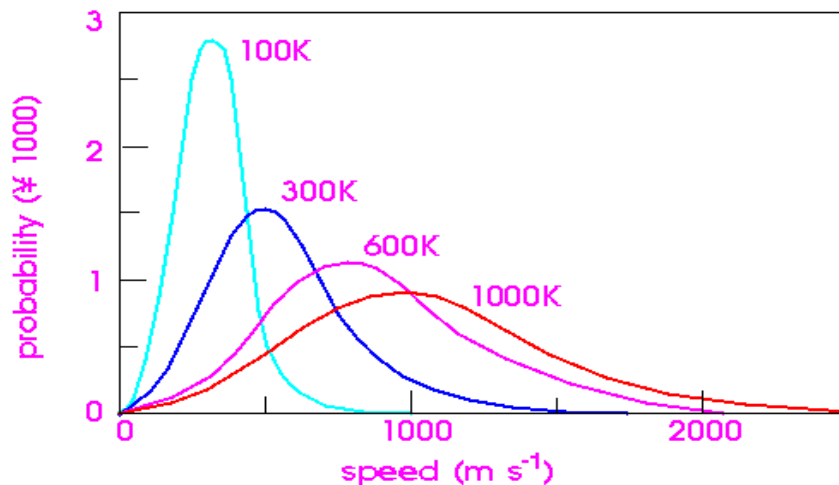
Quindi la distribuzione delle velocità iniziali sono derivate da una distribuzione Maxwelliana random corrispondente alla T desiderata:

$$p(\vec{v}_i) = \sqrt{\frac{m_i}{2\pi K_B T}} \exp [-m_i \vec{v}_i^2 / 2K_B T]$$

$p(\vec{v}_i)$ = probabilità che l'atomo i -esimo abbia velocità $\vec{v}_i$ alla temperatura T .

L'iniziale distribuzione random delle velocità viene scelta in modo tale che non ci sia momento totale della quantità di moto, cioè:

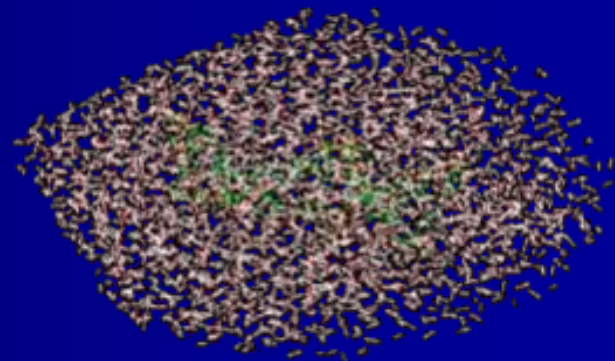
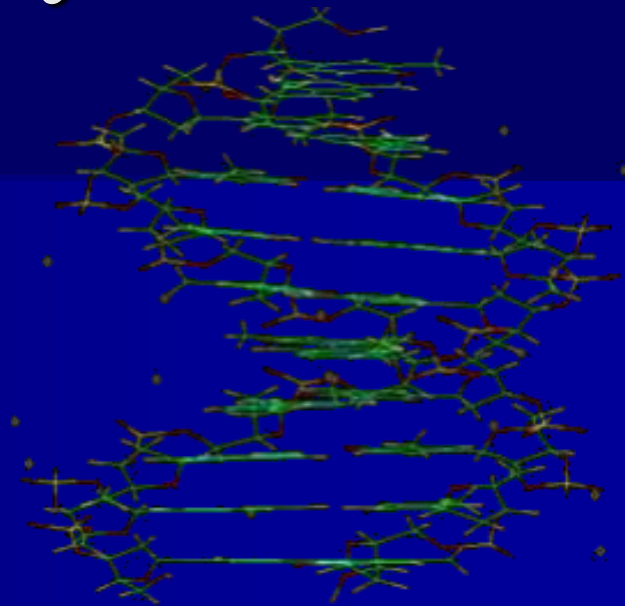
$$\vec{P} = \sum_{i=1}^N m_i \vec{v}_i = 0$$



$$\frac{3}{2} N k_b T(t) = \sum_i \frac{1}{2} m_i v_i^2(t)$$

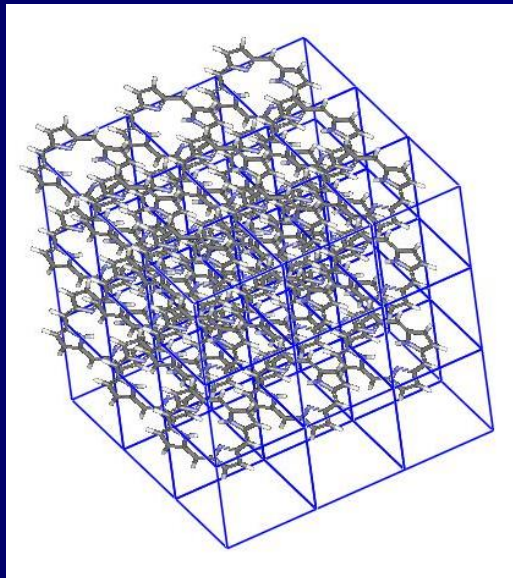
Le «periodic boundary conditions» (PBC)

- Per riprodurre le proprietà reali di un bulk un box di molecole non basta !
- Ciò significherebbe che oltre loro c'è il vuoto?
- Per impedire che le molecole vadano nello spazio e per permettere che un numero piccolo di molecole reproduca il bulk si utilizzano le **PBC**.



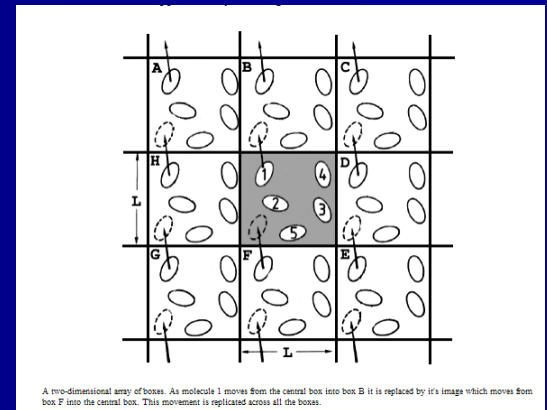
Le periodic boundary conditions PBC

- L'idea risolutiva per questo problema risale al **1912** ed è opera di Born e von Karman.



Questa tecnica genera delle condizioni al contorno periodiche, PBC: ossia **il box viene moltiplicato nello spazio generando una serie di celle tutte uguali una accanto all'altra il cui insieme forma un bulk reale!**

Viene trattata un'unica cella detta originaria (dipende dal sistema fisico)



PBC: “un box di molecole riproduce le proprietà reali di un bulk”

Misura delle proprietà ottenute dagli esperimenti di MD

- La dinamica molecolare permette di determinare *esattamente* l'evoluzione spazio temporale della struttura e le proprietà delle molecole;
- Il processo di integrazione porta all'ottenimento di un set di valori che è costituito dalle **posizioni** $\mathbf{r}(t)$, dalle **velocità** e dalle **accelerazioni** di ogni atomo nel periodo di tempo analizzato
- Bisogna però collegare queste informazioni con delle **grandezze macroscopiche** “caratterizzanti” dello stato termodinamico del sistema

Misura delle proprietà ottenute dagli esperimenti di MD: Insiemi statistici

- Per definire lo stato termodinamico di un gruppo di molecole è necessario disporre di una serie di grandezze come
 - il numero N di atomi
 - il tipo di atomi e come sono legati
 - la temperatura T
 - la pressione P
 - il volume V
 - L'energia globale del sistema.

- Sono dunque **$N, P, T, e V$** a dover essere ricercate

Si fissa un insieme di grandezze termodinamiche

- Le si applica al modello
- Si fa la simulazione e si calcola la sua energia

Misura delle proprietà ottenute dagli esperimenti di MD: Insiemi statistici

- Si introducono così gli INSIEMI STATISTICI:

$$\frac{3}{2}Nk_bT(t) = \sum_i \frac{1}{2}m_i v_i^2(t).$$

Sistema microcanonico	N, E, V costanti	per mantenere E costante la T fluttua per continua conversione di en.cinetica in potenziale.
Sistema canonico	N, T, V costanti	per mantenere la T costante, le velocità fluttuano secondo una statistica gaussiana
Sistema isotermo-isobarico	N, T, P costanti	per mantenere P costante, vengono lasciati fluttuare il volume e pertanto le distanze

Criticità della MD

- Il grado di attendibilità è legato all'accuratezza con cui è stata sviluppata l'equazione dell'energia (legata al MODELLO del FF)
- I risultati di una simulazione molecolare sono quindi determinati unicamente dalla natura del **modello teorico** usato:
 - il **confronto** dei risultati di una simulazione con i dati sperimentali è l'unico test non ambiguo per l'accuratezza del modello scelto
 - differenze tra misure sperimentali accurate e dati virtuali possono essere attribuiti senza dubbio al **fallimento** del modello scelto nel rappresentare il comportamento molecolare.

La **simulazione MD** è in molti aspetti simile ad un vero e proprio **esperimento di laboratorio**, perciò gli *step* da seguire per ottenere una simulazione realistica sono correlati alla sua controparte reale:

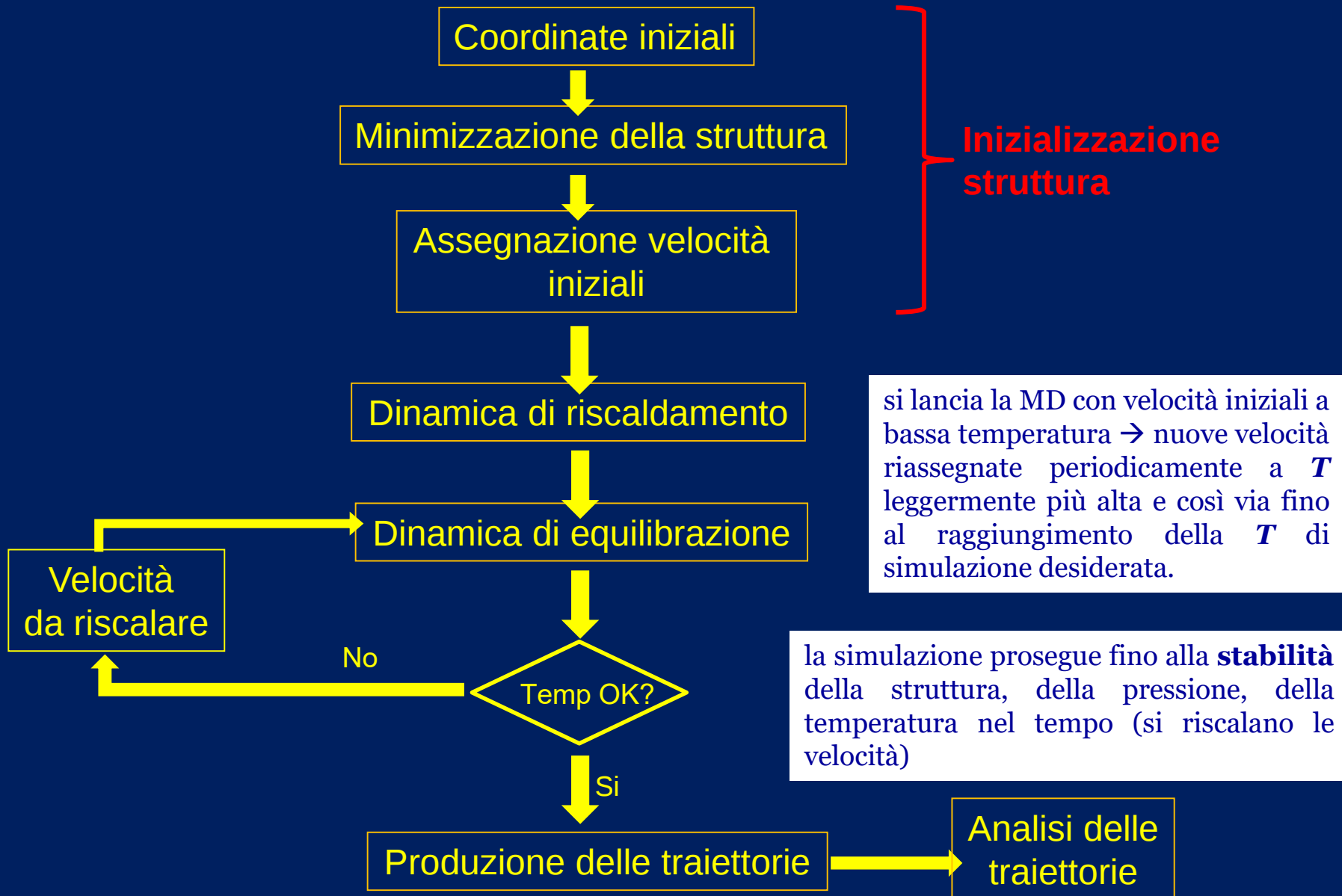
1. Per prima cosa occorre preparare il campione da simulare, costituito da N particelle, risolvendo le equazioni di Newton affinché le proprietà temporali siano quelle di equilibrio (**processo di equilibratura**);

2. **Simulazione dell'evoluzione del sistema** monitorando le grandezze termodinamiche importanti: **Energia, Temperatura, Volume, Pressione e Numero di particelle devono essere costanti**; questa è la vera parte di simulazione di dinamica molecolare in cui avviene la produzione dei dati da analizzare;

3. A simulazione conclusa, dopo aver definito delle opportune osservabili, **si misurano le grandezze fisiche di interesse al variare del tempo**.

Questa quindi è la parte di analisi dati che ci permetterà di trarre delle conclusioni alla fine del nostro esperimento computazionale.

Set-up e running di una Simulazione MD



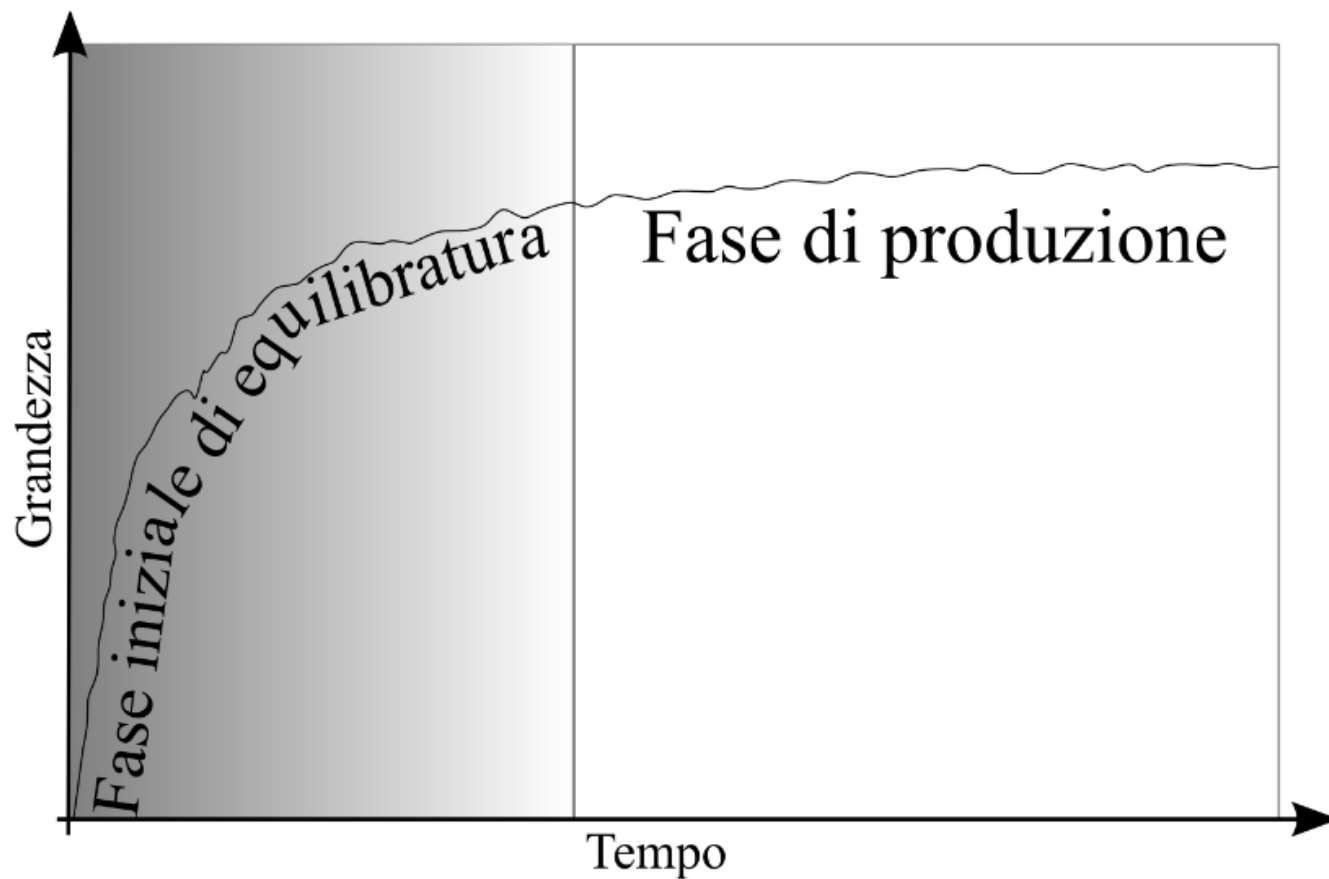


Figura 2 Fase di equilibratura e fase di produzione per una generica grandezza durante una simulazione di dinamica molecolare.

Analisi dei risultati

- calcolo della “root mean squared deviation” (**rmsd**) di atomi della molecola rispetto ad una struttura di riferimento (non necessariamente quella a $t = 0$):

$$\text{rmsd}(r_1, r_2) = \sqrt{\frac{1}{N} \sum_i (r_{1,i} - r_{2,i})^2}$$

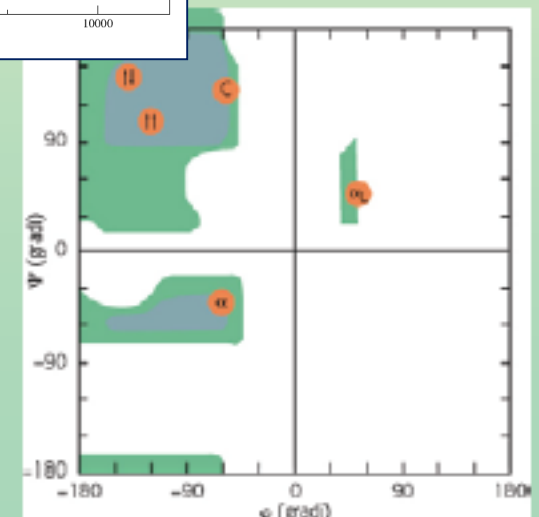
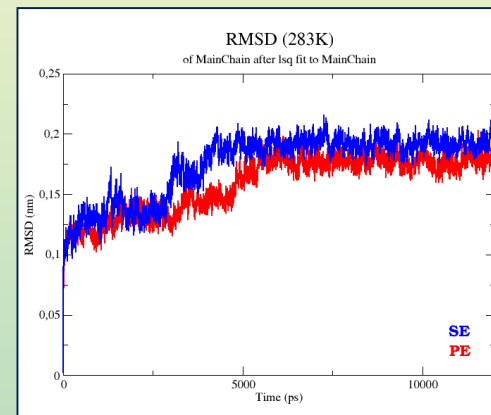
$$r_1 = (x_{1,1}, \dots, x_{1,N})$$

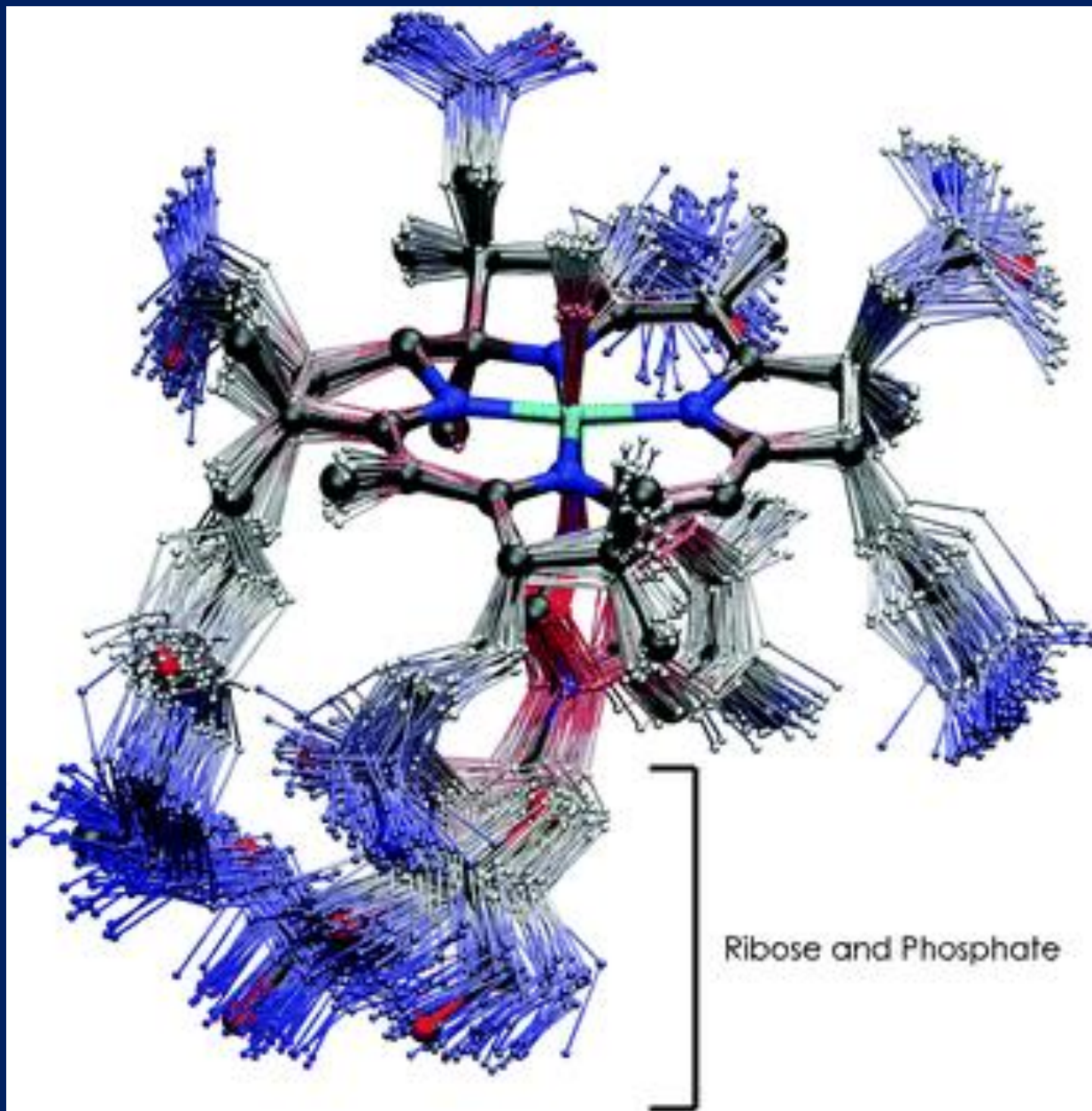
$$r_2 = (x_{2,1}, \dots, x_{2,N})$$

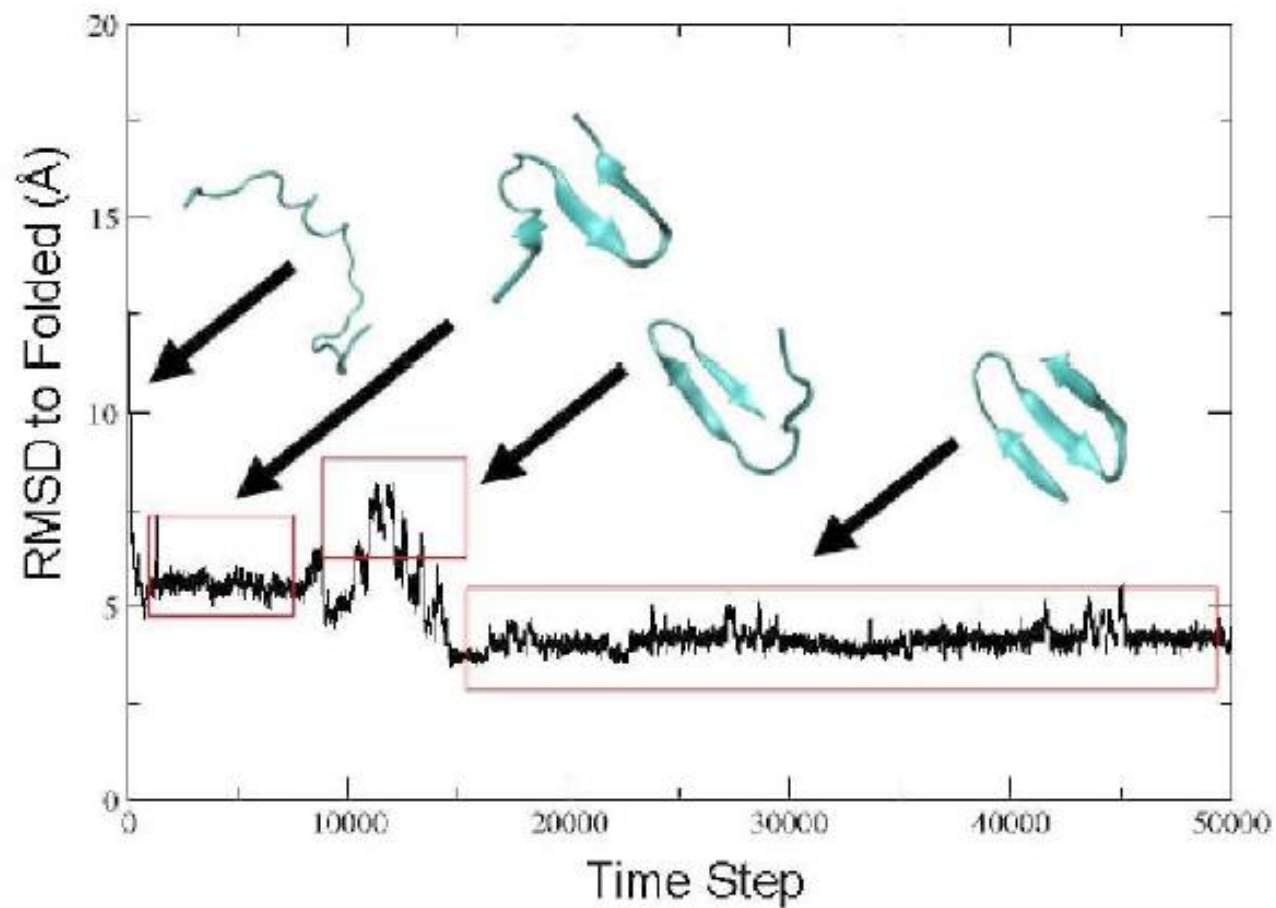
- calcolo delle **fluttuazioni di rmsd**

$$\text{rmsd}_i^{\text{fluct}} = \sqrt{\frac{1}{N_f} \sum_f (r_i^f - r_i^{\text{ave}})^2}$$

- andamento del **plot di Ramachandran** nel tempo (variazione della struttura secondaria).

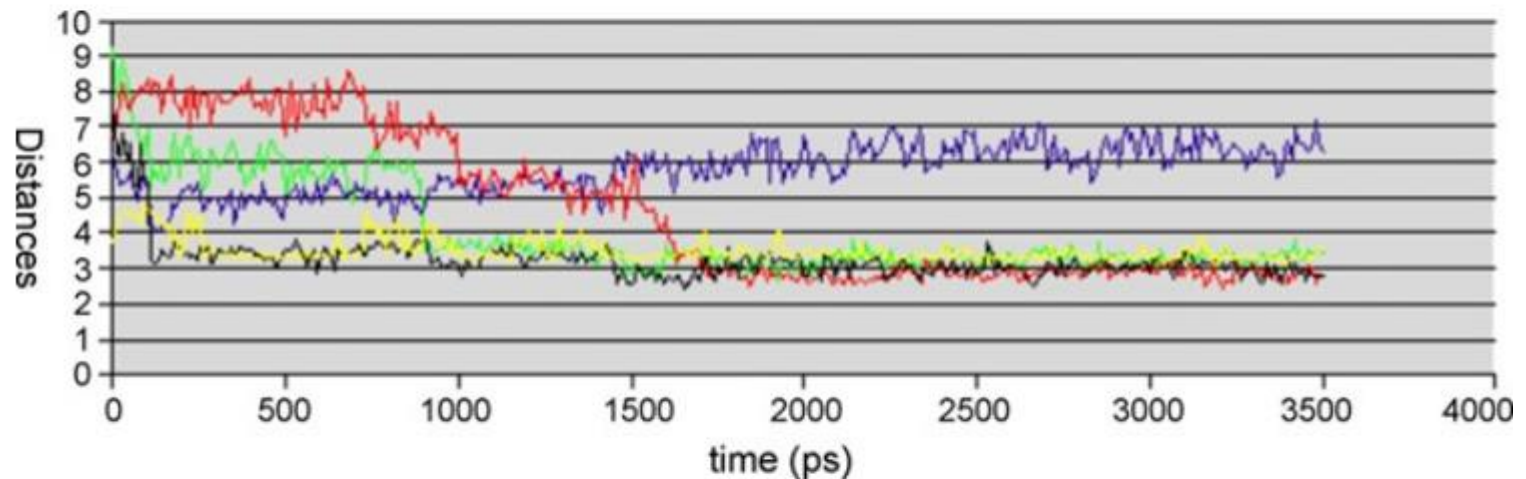




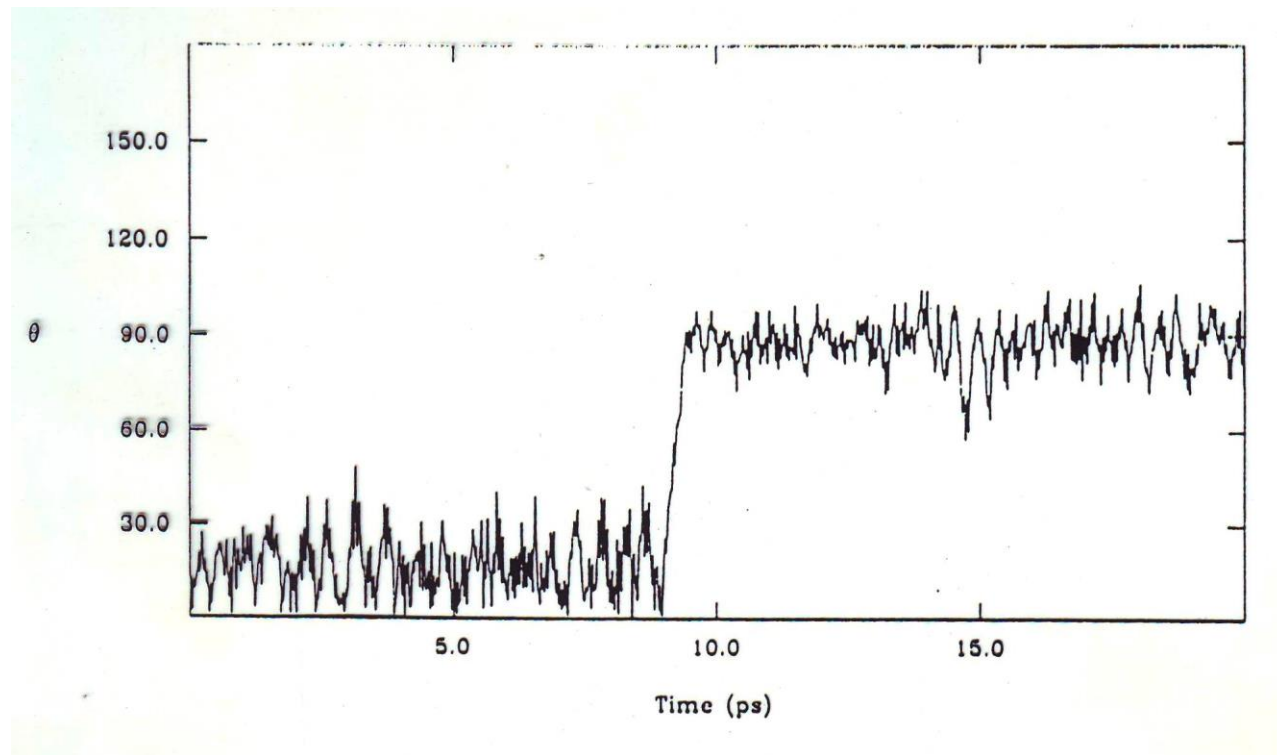


Esempio di Traiettorie

- simulazione MD già' equilibrata dopo 2 ns (dato non mostrato)
- Monitorano **stabilita' ponti salini** tra aa durante l'intera simulazione (distanza nel tempo tra aa coinvolti nel ponte salino)

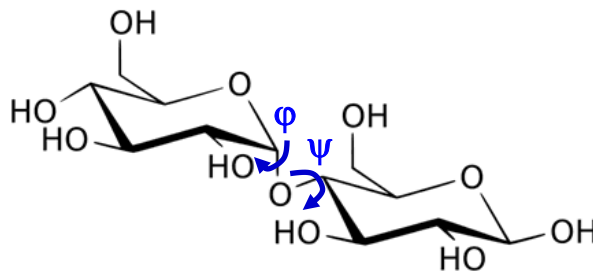


- **Distanza** tra Asp700 (su HECT) e Lys9 (Ubch7) aumenta dopo 2 ns oltre la soglia di 6 Å indicando un'interazione più debole e non fondamentale al mantenimento complesso
- **Distanza** tra Asp681 (HECT) e Lys64 (Ubch7) stabilizzato durante simulazione e loro distanza diminuisce dopo 2 ns.



1. Proprietà termodinamiche.
2. Proprietà strutturali.
3. Proprietà dinamiche.

Esempio: maltosio

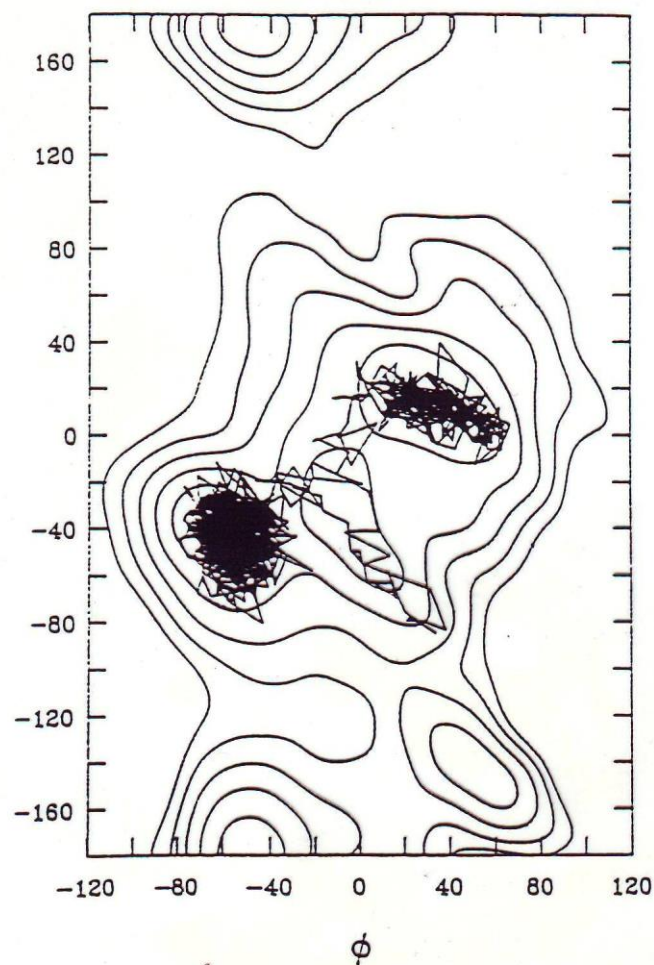
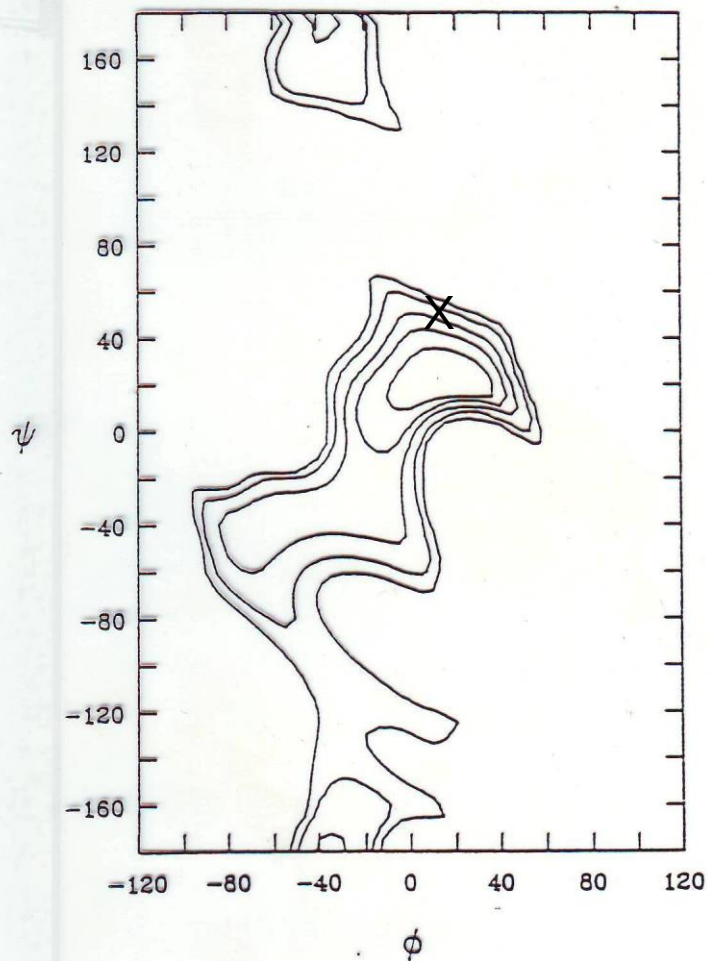


Residuo rigido

MM

MD (+MM)

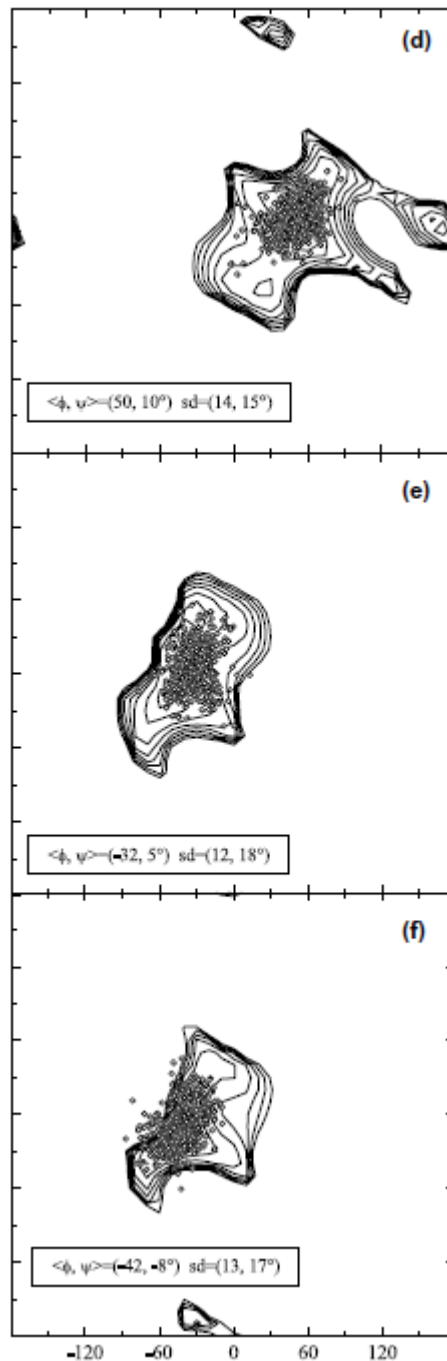
Residuo flessibile



Disaccaride A-B

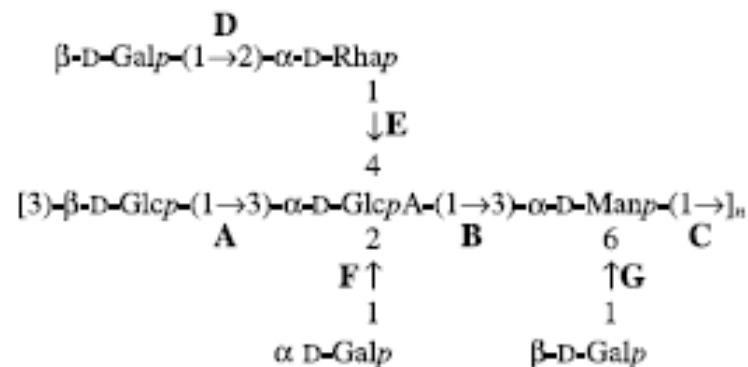
Disaccaride B-C

Disaccaride C-A



Conformational features of cepacian: the exopolysaccharide produced by clinical strains of *Burkholderia cepacia*

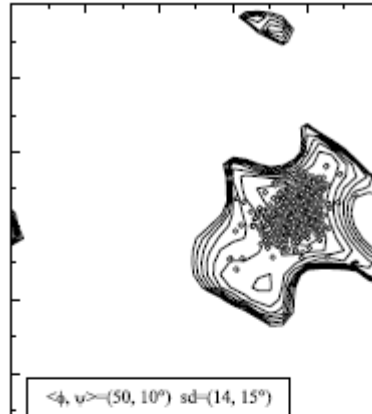
Carlos E. Sampaio Nogueira,^a Jose R. Ruggiero,^a Paola Sist,^b Paola Cescutti,^b Ranieri Urbani^{b,*} and Roberto Rizzo^b



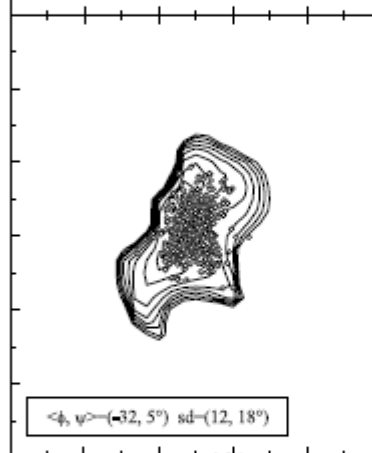
SOLO CATENA PRINCIPALE



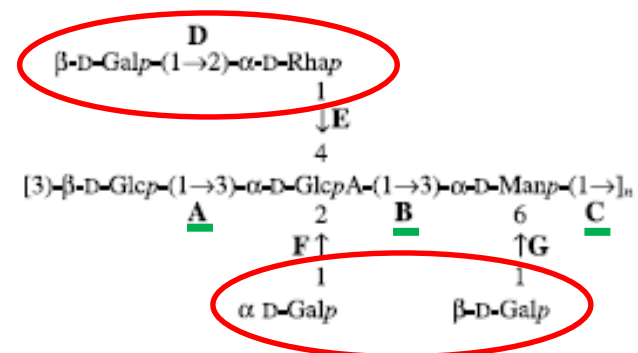
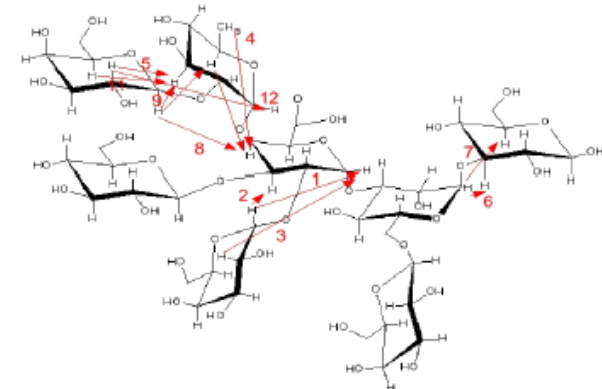
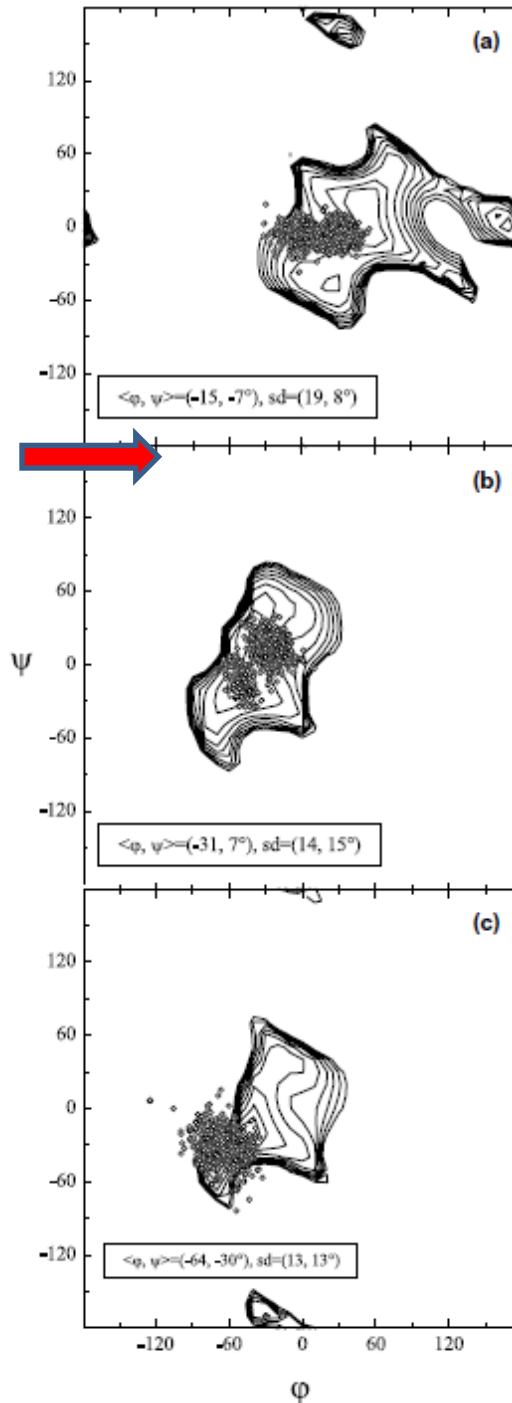
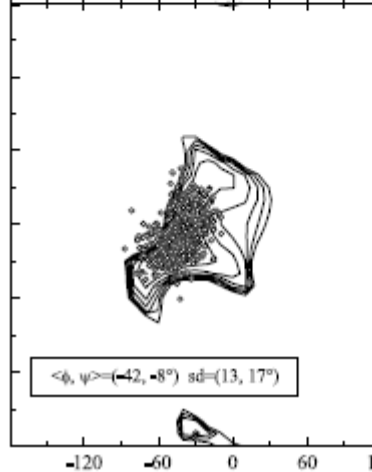
Disaccharide A-B



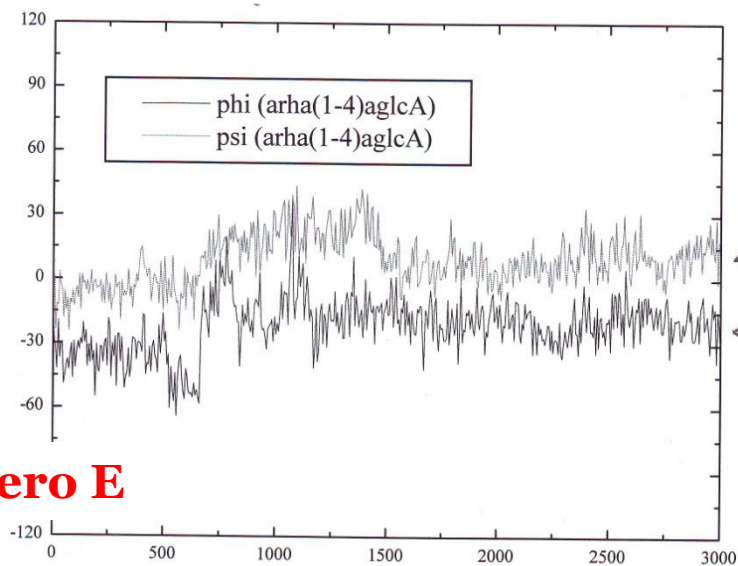
Disaccharide B-C



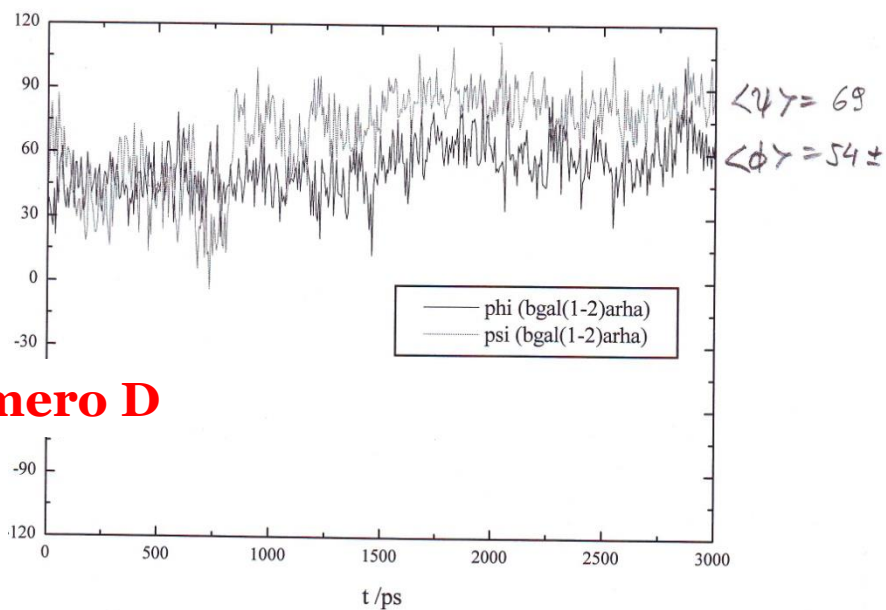
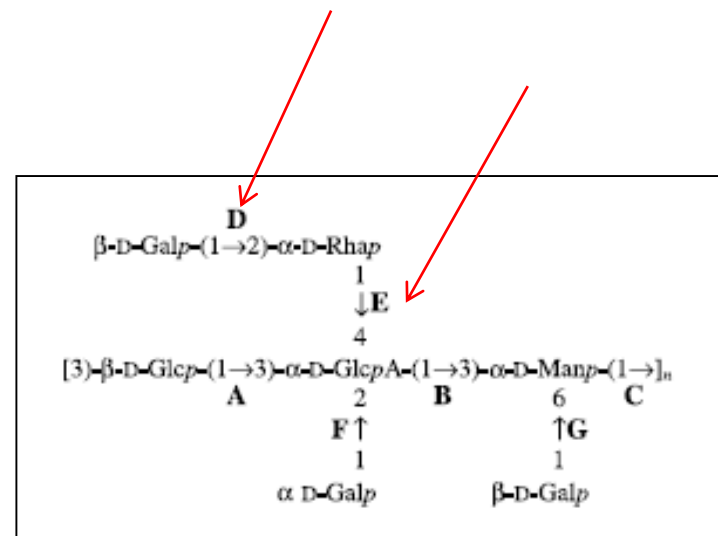
Disaccharide C-A



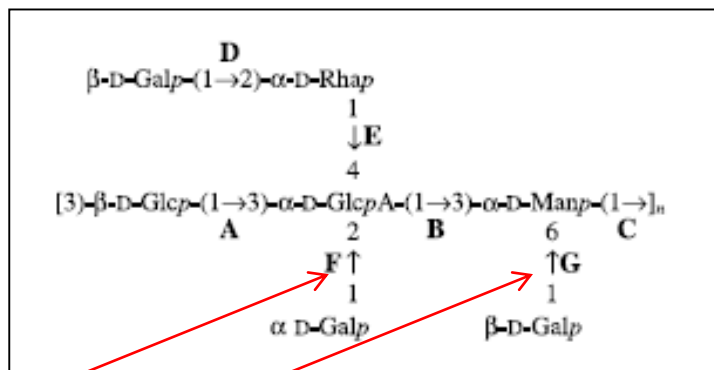
100% DMSO



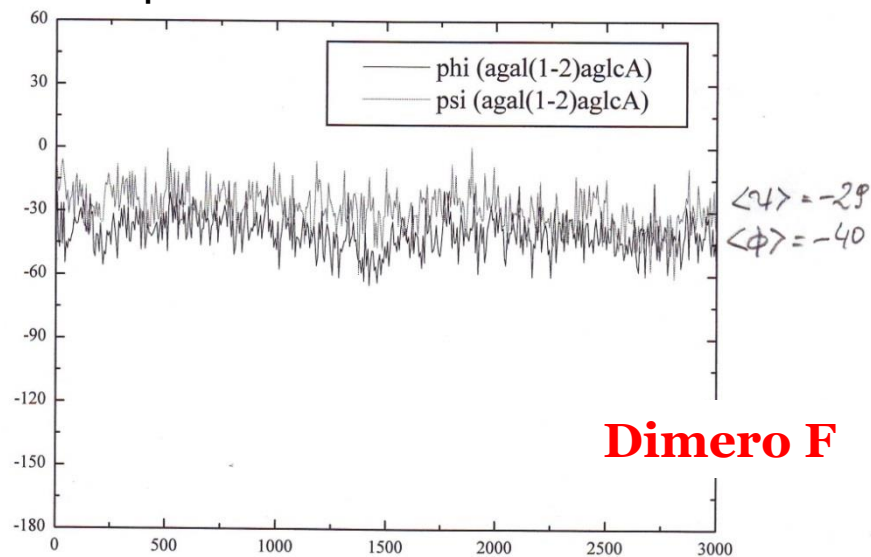
Dimero E



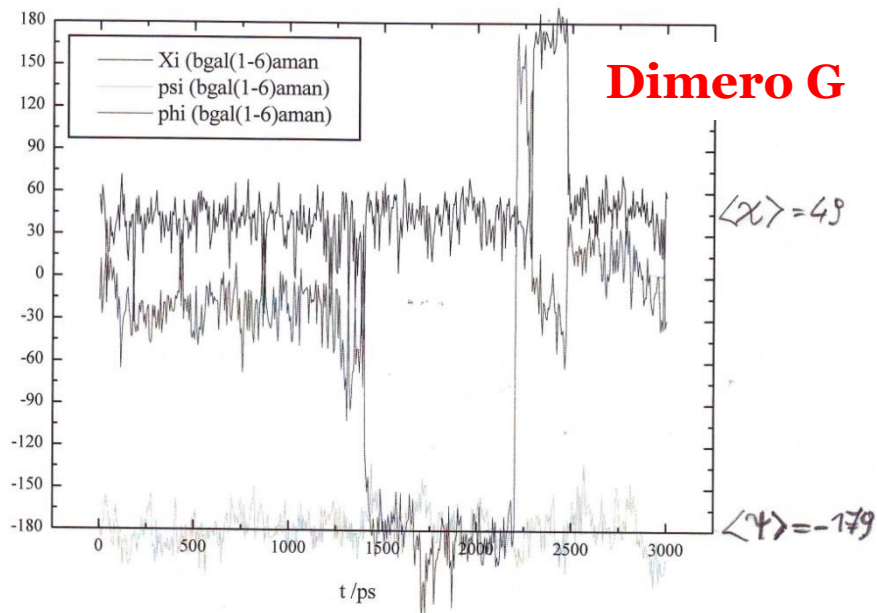
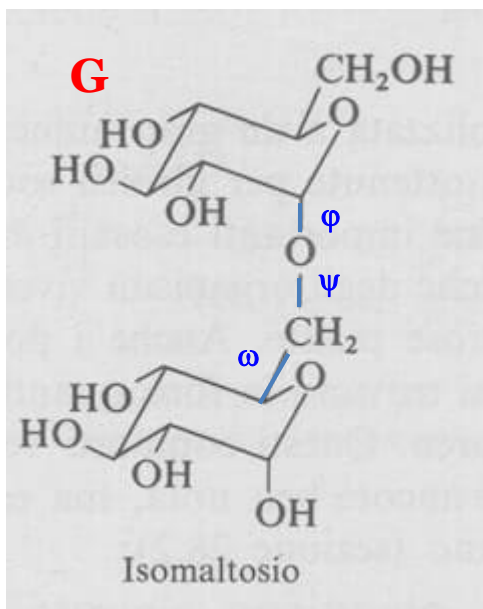
Dimero D



acqua



Dimero F



Dimero G

Macromolecular properties of cepacian in water and in dimethylsulfoxide

Yury Herasimenka,^{a,†} Paola Cescutti,^a Carlos E. Sampaio Noguera,^b
 Josè R. Ruggiero,^b Ranieri Urbani,^a Giuseppe Impallomeni,^c Flavio Zanetti,^d
 Stéphane Campidelli,^{e,‡} Maurizio Prato^e and Roberto Rizzo^{a,*}

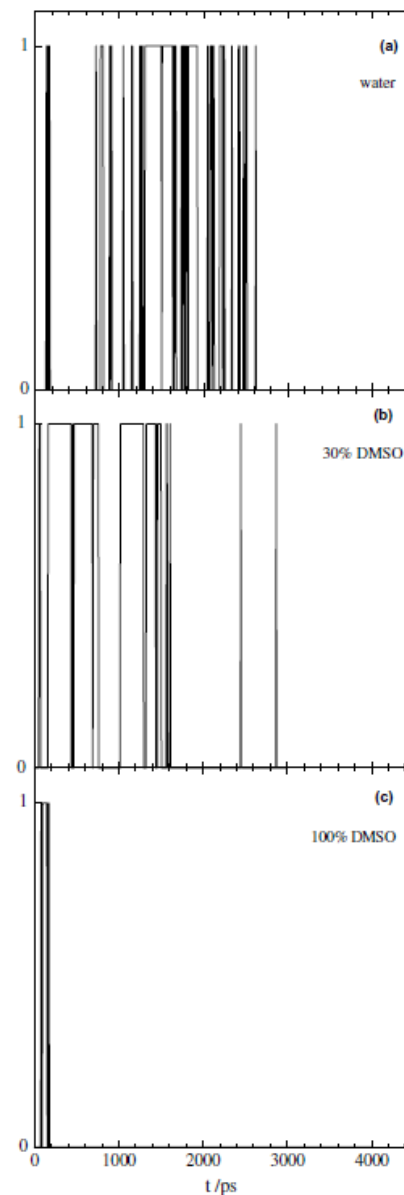


Figure 10. Occurrence of hydrogen bonding during the 3 ns simulation time for the 3-OH (α -D-Rhap)–O-6 (β -D-Glcp) interaction in (a) water, (b) 30% Me₂SO–water, and (c) 100% Me₂SO solutions.

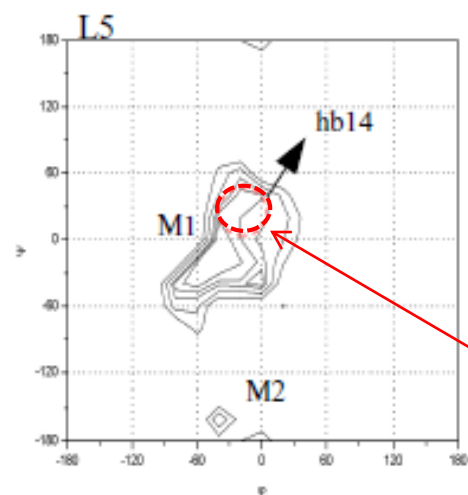


Figura 13: Mapa de Energia da ligação L5,
 α -man(1 $\rightarrow$ 3) β -glc.

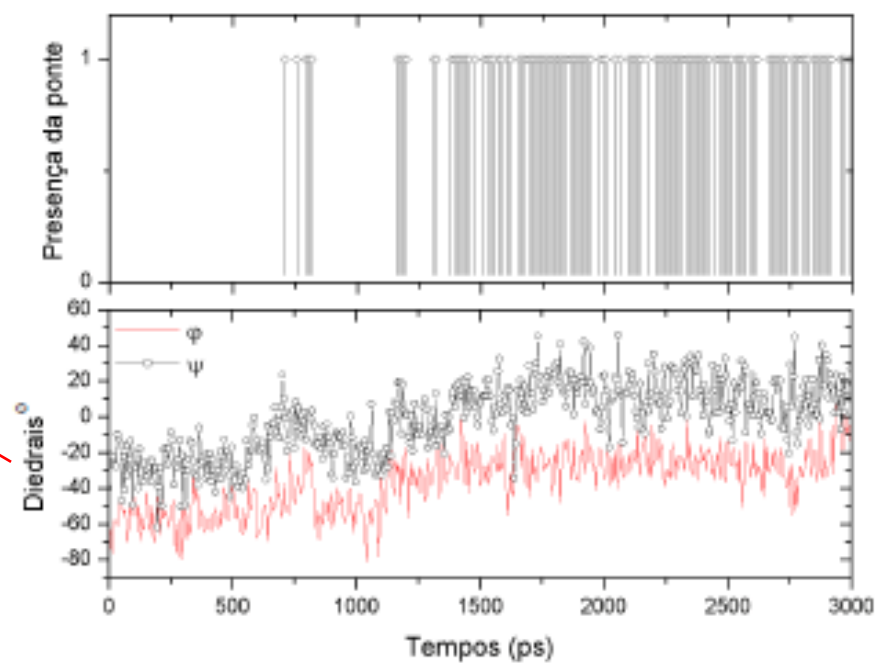


Figura 14: A dependência da ponte hb14 dos valores dos ângulos conformacionais da ligação L5.

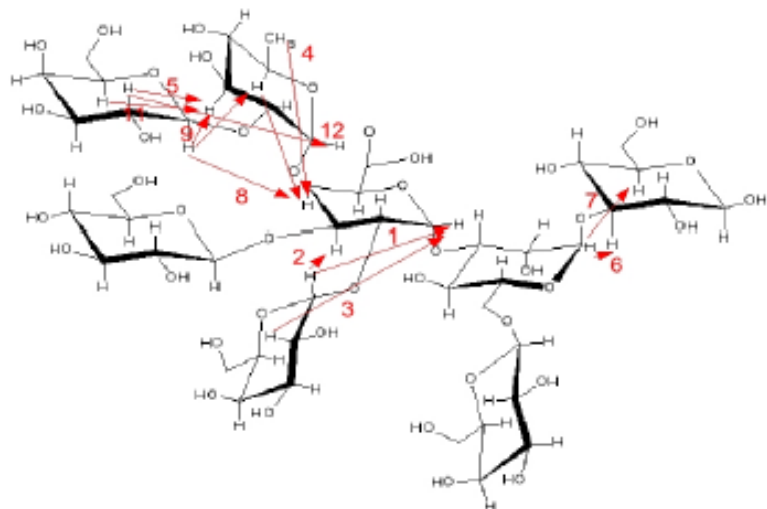


Figura 27: Esquema com os contatos de NMR. A numeração coincide com aquela da Tabela 6, na página 54, que exibe as distâncias para os respectivos contatos.

Eq. di Karplus

$$^3J_{H,H} = C \cos 2\phi + B \cos \phi + A$$

dove

- 3J è la costante di accoppiamento
- ϕ è l'angolo diedro
- A , B e C sono dei parametri ottenuti empiricamente i cui valori dipendono dagli atomi e dai sostituenti presenti nella molecola.



N. do Contato	Prótons envolvidos	Dist. NMR	Dist. MD	MD std
1	$\alpha\text{-gal}(1) \rightarrow \alpha\text{-glcA}(1)$	1,8	2,1	0,24
2	$\alpha\text{-gal}(1) \rightarrow \alpha\text{-glcA}(2)$	2,6	2,5	0,25
3	$\alpha\text{-glcA}(1) \rightarrow \beta\text{-gal}(3)$	4,6	4,9	0,25
4	$\alpha\text{-glcA}(4) \rightarrow \alpha\text{-rha}(6)$	4,6	4,4	0,39
5	$\alpha\text{-rha}(1) \rightarrow \beta\text{-gal}(2)$	4,5	4,5	0,25
6	$\alpha\text{-man}(1) \rightarrow \beta\text{-glc}(3)$	2,2	2,2	0,28
7	$\alpha\text{-man}(1) \rightarrow \beta\text{-glc}(5)$	4,1	4,3	0,33
8	$\beta\text{-gal}(1) \rightarrow \alpha\text{-glcA}(4)$	4,3	4,3	0,25
9	$\beta\text{-gal}(1) \rightarrow \alpha\text{-rha}(3)$	4,3	4,5	0,40
10	$\beta\text{-gal}(1) \rightarrow \alpha\text{-rha}(5)$	4,4	5,4	0,33
11	$\alpha\text{-rha}(2) \rightarrow \beta\text{-gal}(5)$	4,2	3,6	0,27
12	$\alpha\text{-rha}(2) \rightarrow \beta\text{-gal}(2)$	4,5	4,2	0,11
13	$\alpha\text{-rha}(5) \rightarrow \alpha\text{-glcA}(4)$	3,8	3,6	0,33

Tabela 6: Comparação das distâncias (em Angstroms) interprótons obtidas por ressonância magnética com as obtidas através de dinâmica molecular. O número entre parênteses indica a que carbono está ligado o próton.

Proprietà Termodinamiche

- **Energia Interna:**

$$u = \langle E \rangle = \frac{1}{M} \sum_{i=1}^M E_i$$

- **Capacità termica**

$$C_v = \{ \langle E^2 \rangle - \langle E \rangle^2 \} / k_B T^2$$

- **Pressione:**

$$P = \frac{1}{V} \left[N k_B T - \frac{1}{3} \sum_{i=1}^N \sum_{j=i+1}^N r_{ij} f_{ij} \right]$$

- **Temperatura:**

$$K = \sum_{i=1}^N \frac{|P_i|^2}{2m_i} = \frac{k_B T}{2} (3N - N_c)$$

ESEMPI di grandezze termodinamiche calcolabili con MD:

Calcolo della costante di diffusione

Equazioni di Fick

$$J_z = -D \frac{\partial N}{\partial z}$$

$$\frac{\partial N}{\partial z} = D \frac{\partial^2 N(z, t)}{\partial z^2}$$

Equazioni di Einstein

$$6Dt = \frac{1}{N} \sum_{i=1}^N [r_i(t) - r_i(0)]^2$$

Calcolo della temperatura del sistema

$$\frac{1}{2} \sum_i^N m_i v_i^2 = \frac{1}{2} N_f k_b T$$

$$T = \frac{1}{N_f k_b} \sum_i^N m_i v_i^2$$

Alcuni esempi di possibili confronti tra dati sperimentali e proprietà simulate dalla MD

Proprietà atomiche	Metodi sperimentali	Gas	Soluzione	Membrana	Cristallo
Struttura					
-Posizioni	Diffrazione dei raggi X, neutroni o elettroni	✓			✓
-Distanza	Risonanza magnetica		✓		
-Orientazione	NMR				✓
Mobilità					
-fattori di Debye-Waller	Diffrazione dei raggi X, neutroni				✓
-fattori di occupanza	Diffrazione dei raggi X, neutroni				✓
Dinamiche					
-Frequenze di vibrazione	Spettroscopia nell'infrarosso	✓			✓
-Velocità di rilassamento	NMR e spett. ottiche		✓	✓	
-Diffusione	NO spin label		✓	✓	
Termodinamiche					
-densità	densitometria		✓	✓	✓
-Entalpia di evaporazione			✓		
-Viscosità e conduttanza	viscosimetria		✓	✓	

Solvatazione

E' importante poiché molte molecole sono in soluzione

Le molecole di solvente polare si dispongono attorno allo ione o alla molecola con il lato di polarità opposta rispetto alla carica (parziale) ed esercitano un campo elettrico

Questo campo elettrico dà luogo ad un **potenziale di interazione** che deve essere aggiunto all'Hamiltoniana dell'energia

Nei metodi di Meccanica e Dinamica Molecolare si aggiunge semplicemente un contributo elettrostatico nell'espressione empirica dell'energia e si considera poi una costante dielettrica nell'interazione elettrostatica tra i vari atomi



- Effetto solvente** {
1. Trattamento **IMPLICITO** (modello CONTINUO)
 2. Trattamento **ESPLICITO** (modello ALL-ATOMS)

Effetto solvente { 1. Trattamento IMPLICITO (modello CONTINUO)
2. Trattamento ESPLICITO (modello ALL-ATOMS)

Solvation Models

- Explicit solvent models
 - Fixed charge models: SPC, SPC/E, TIP3P, TIP4P, TIP5P, ST2,...
 - Polarizable water models: TIP4P/FQ, POL5, MCDHO,...
- Implicit Solvent models
 - Poisson-Boltzman solver (Delphi, Honig)
 - Generalized Born Model (Still)
 - Karplus' EEF1 model
 - Benoit Roux's Spherical Solvent Boundary Potential (SSBP)

Termodinamica di solvatazione

Energia libera di solvatazione: *variazione di energia libera nel trasferimento di una molecola dal vuoto al solvente.*

$$\Delta G_{solv} = \Delta G_{elett} + \Delta G_{VdW} + \Delta G_{cav} + \Delta G_{HB}$$

The diagram illustrates the components of the free energy of solvation equation. Four boxes are connected to the terms in the equation by arrows:

- A box labeled "elettrostatico" has an arrow pointing to ΔG_{elett} .
- A box labeled "Van der Waals" has an arrow pointing to ΔG_{VdW} .
- A box labeled "cavitazione" has an arrow pointing to ΔG_{cav} .
- A box labeled "Legame idrogeno" has an arrow pointing to ΔG_{HB} .

Modello del solvente continuo

Il solvente agisce come una perturbazione del nostro sistema in fase gassosa e questo concetto è alla base dei modelli del solvente “continuo”.

L'energia libera di solvatazione (ΔG_{sol}) include tre componenti:

$$\Delta G_{sol} = \Delta G_{elec} + \Delta G_{vdw} + \Delta G_{cav} + \Delta G_{HB}$$

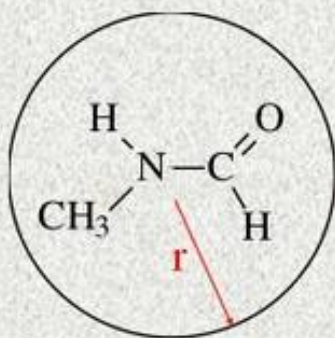
dove ΔG_{elec} è la componente elettrostatica (la polarizzazione del solvente con costante dielettrica ϵ).

ΔG_{vdw} rappresenta le interazioni di van der Waals tra il soluto ed il solvente; questa può essere ulteriormente suddivisa in un termine repulsivo, ΔG_{rep} , ed in un termine di dispersione attrattiva, ΔG_{disp} .

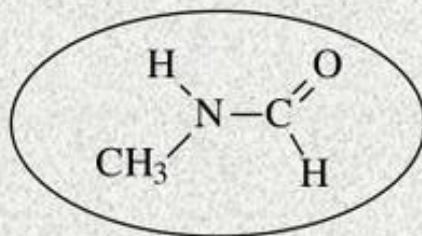
ΔG_{cav} è l'energia libera richiesta per formare la **cavità** del soluto all'interno del solvente. Questa componente è positiva e comprende la penalizzazione entropica associata alla riorganizzazione delle molecole di solvente attorno al soluto ed al lavoro fatto contro la pressione del solvente nel creare la cavità.

1a. Solute Cavity Size and Shape

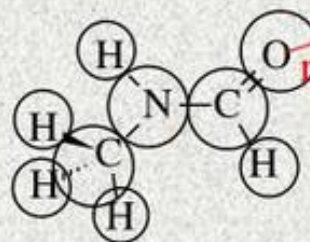
Spherical
(Born)



Ellipsoidal
(Onsager)

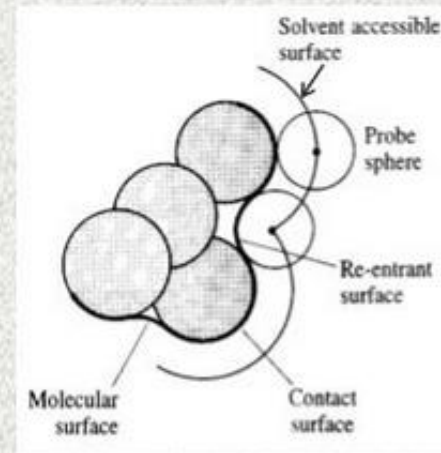
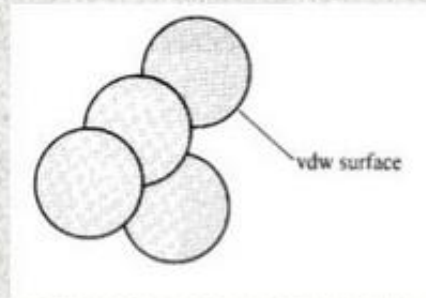


van der Waals
(Kirkwood)



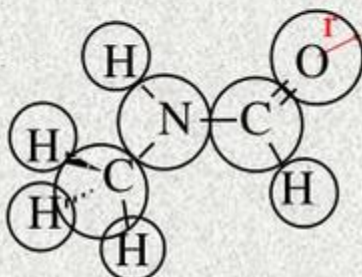
1b. Solvent-Accessible Surface

- A solvent-accessible surface is made by connecting the center of a “rolling” sphere (e.g., 1.5 Å radius) around the vdW (electron density iso-) surface of the molecule.
- This method excludes pockets that are inaccessible even to small solvent molecules.



3b. Charge Distribution of Solute

(assuming vdW sized spheres for each atom)



$$\Delta G_{el} = - \left(1 - \frac{1}{\epsilon}\right) \frac{q^2}{2r}$$

(charge = q)

$$\Delta G_{el} = - \frac{\epsilon - 1}{2\epsilon + 1} \cdot \frac{\mu^2}{r^3} \left[1 - \frac{\epsilon - 1}{2\epsilon + 1} \cdot \frac{2\alpha}{r^3} \right]^{-1}$$

(dipole + polarizability)

$$\Delta G_{el} = - \frac{\epsilon - 1}{2\epsilon + 1} \cdot \frac{\mu^2}{r^3}$$

(dipole = μ)

(this calculation is summed over all atoms)

SOLVATAZIONE

Trattamento **IMPLICITO**

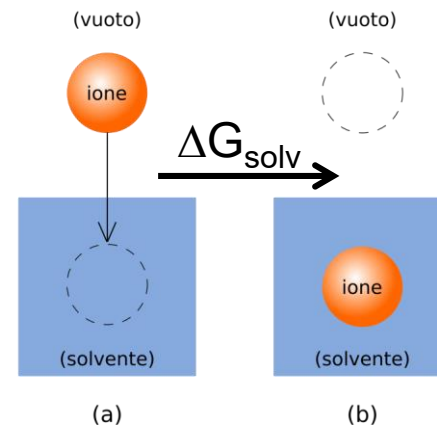
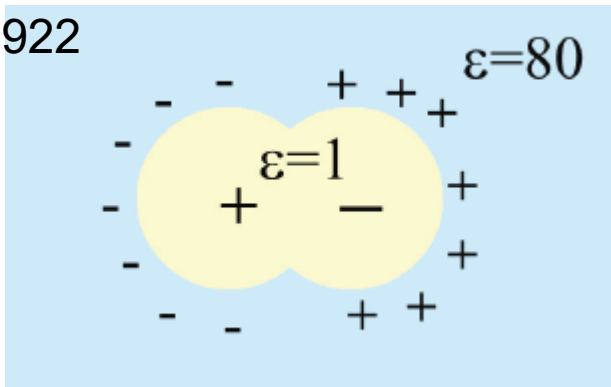
Il modo implicito più semplice è quello di includere una costante elettrostatica nel termine elettrostatico dell'energia potenziale, usando una costante dielettrica effettiva dipendente dalla distanza:

$$\epsilon_{\text{eff}} = \epsilon(r_{ij}).$$

Le molecole di solvente sono implicitamente considerate modificando opportunamente le interazioni tra i nuclei come se essi subissero l'effetto "medio" cioè statistico della presenza di un gran numero di molecole di solvente. In questo caso il costo computazionale è notevolmente ridotto rispetto alla trattazione con solvente esplicito.

Modello di Born

1922



$$\Delta G_{\text{sol}} = \Delta G_{\text{elec}} + \Delta G_{\text{vdw}} + \Delta G_{\text{cav}}$$

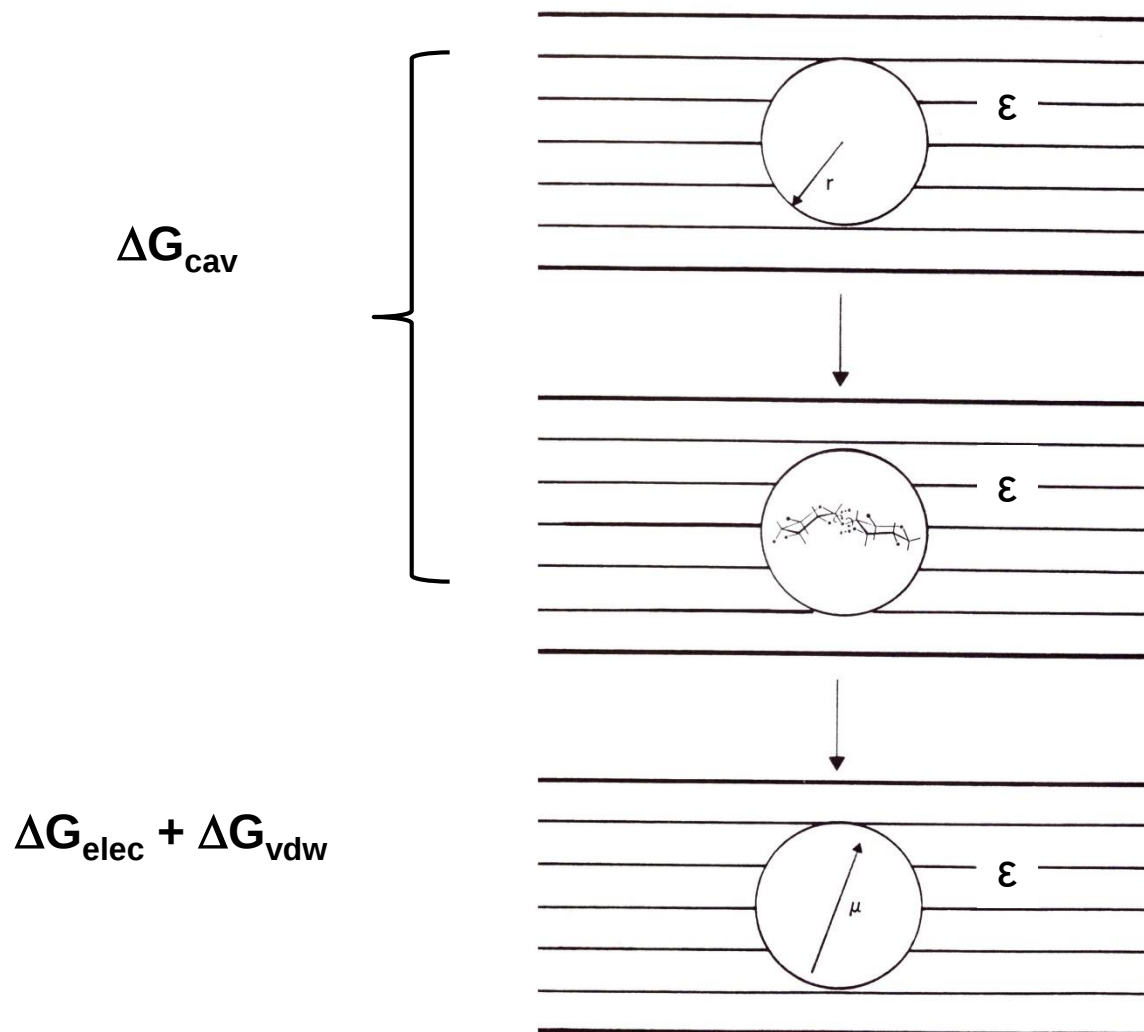
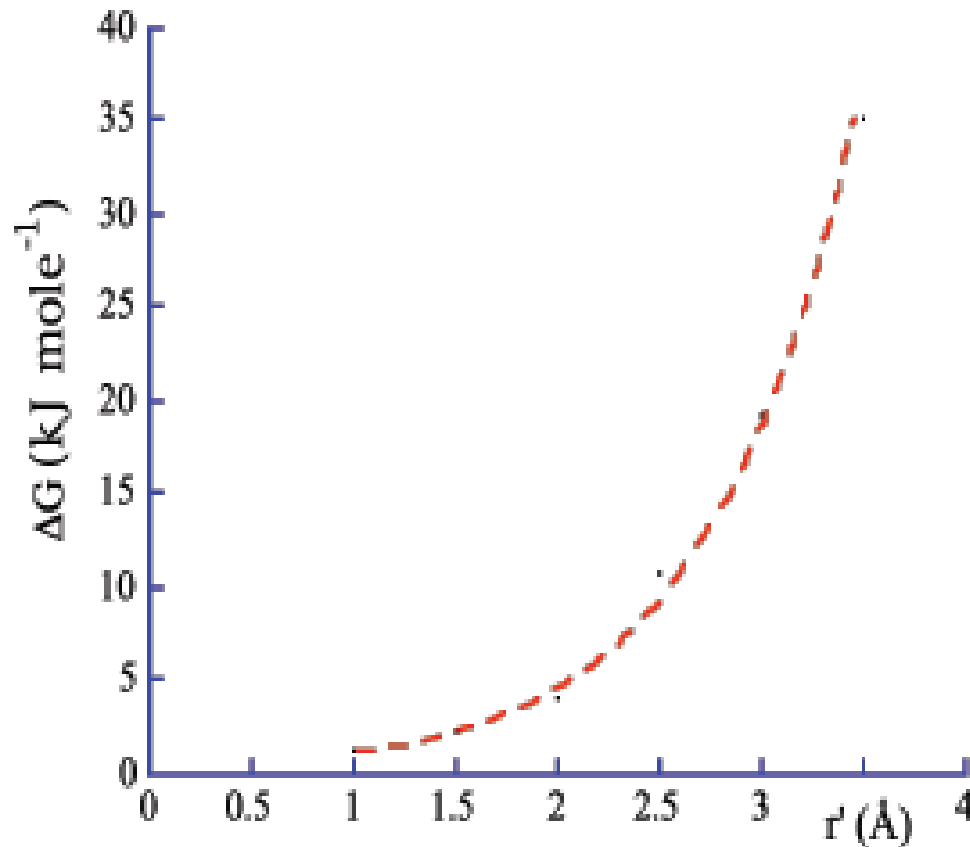


Figure 1 Scheme of the theoretical approach for the calculation of the solvation energies. First, a cavity of proper size is created in the solvent. Secondly, the real molecule, with a given conformational energy, is inserted in the cavity. Thirdly, the surrounding molecular layer of the solvent interacts with the solute molecules with given charge placement

L'energia libera per la creazione di una cavità aumenta fortemente con le dimensioni della cavità stessa.



Variazione dell'energia libera in funzione del raggio della cavità

Trattamento ESPLICITO (All-atoms)

Consiste nell'includere molecole di solvente nel sistema. Le molecole forniscono direttamente l'effetto di schermo (o almeno la parte legata al riorientamento delle stesse).

Condizioni al contorno

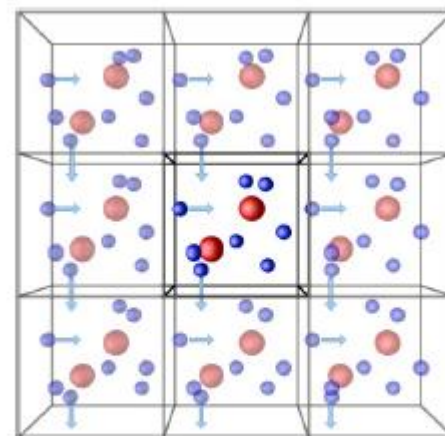
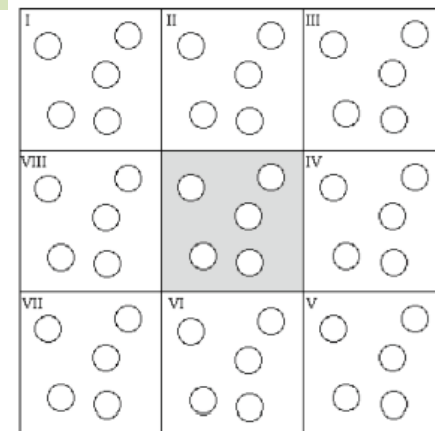
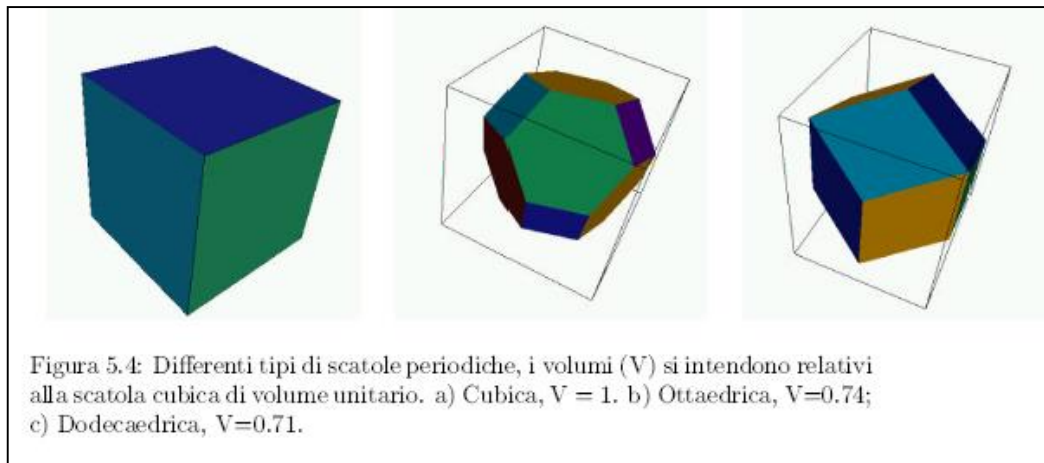


Figura 2.1. Rappresentazione delle condizioni periodiche al contorno: l'unico sistema realmente simulato è quello centrale ma grazie a questa condizione è come se fosse circondato da sistemi identici. Fonte: isaacs.sourceforge.net

Per realizzare **la scatola** si deve conoscere la densità del liquido **ad una certa temperatura**.

Questa ci permette di calcolare il volume medio occupato da una singola molecola.

Se ***d*** è la densità in kg m⁻³, allora il **volume in nm** occupato da una molecola è uguale a:

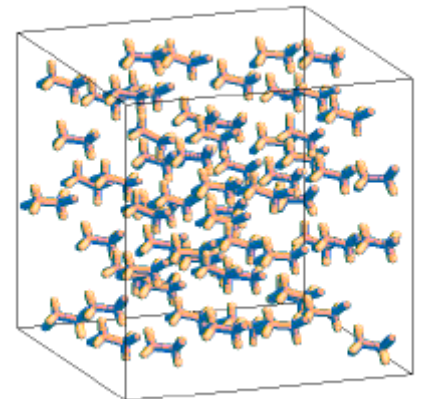
$$V_m = \frac{1000M}{dN_a} 10^{21}$$

Dal volume (***V_m***) si ottiene il lato (***l***) del volumetto che contiene la molecola.

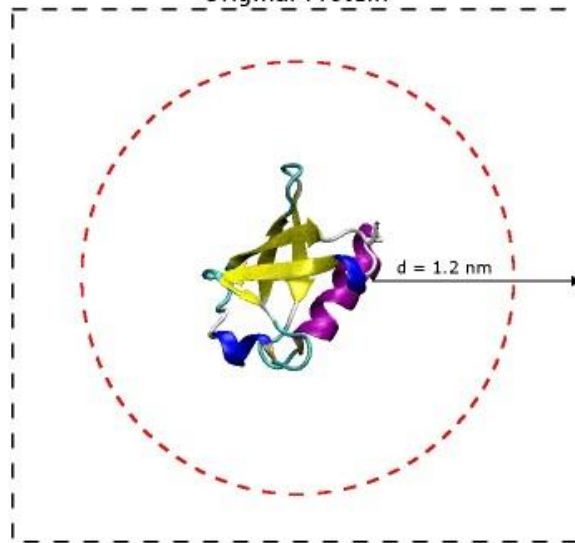
Per esempio, nel caso del 1,2-dicloroetano abbiamo che $M = 98.96\text{g}$ e $d = 1246.3\text{ kg m}^{-3}$, da cui si ottiene $l = \sqrt[3]{0.132} = 0.509\text{ nm}$.

Il metodo replica il volumetto di dimensione (0.509 x 0.509 x 0.509) avente la densità del liquido (ad una data ***T***) e la molecola in essa contenuto in un reticolo di dimensione voluta.

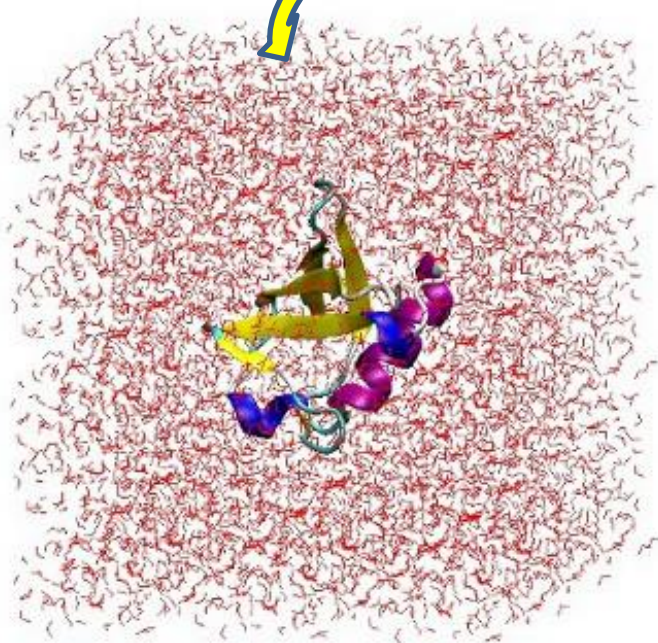
Le molecole vengono distribuite ed orientate in modo *random*.



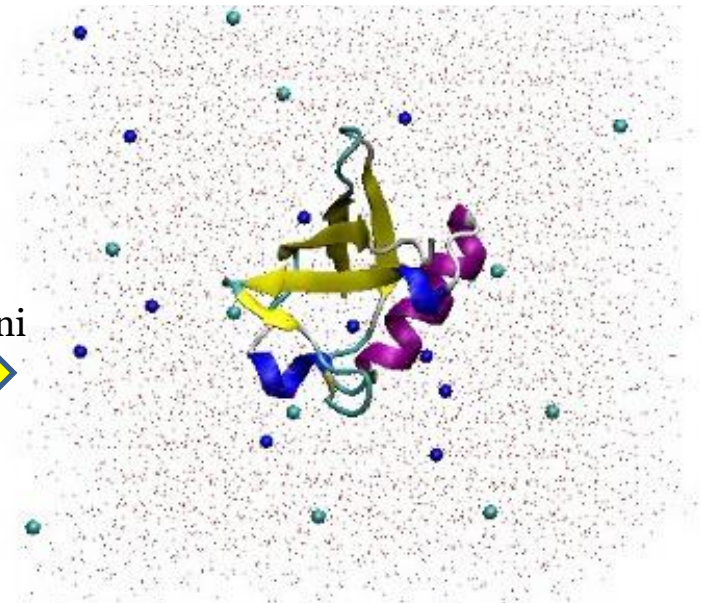
Original Protein



Addizione di solvente



Addizione di ioni



Ottimizzazione geometria sistema (soluto + solvente + controioni)

Modello TIP3P dell'acqua (*transferable intermolecular potential 3P*)

Descrive accuratamente le interazioni della molecola con le altre molecole del solvente ed anche con il soluto stesso (**potenziali L-J e coulombiani**)

Questo modello è caratterizzato dalla presenza di tre siti di interazione e risulta molto accurato

$$V_{ab}^{\text{TIP3P}} = \sum_i^{\text{on a}} \sum_j^{\text{on b}} \frac{k_c q_i q_j}{r_{ij}} + \frac{A}{r_{OO}^{12}} - \frac{B}{r_{OO}^6}$$

dove $k_C = \frac{1}{4\pi\epsilon_0} = 332.1 \text{ Å}^2 \text{ kcal/mol}$ è la costante elettrostatica, q_i sono le cariche parziali relative alla carica degli elettroni, che nel modello TIP3P vengono settate pari a:

$$q_O = -0.834e \quad q_H = +0.417e \quad (4.20)$$

mentre r_{ij} è la distanza tra due siti atomici carichi, ovvero:

$$r_{OH} = 0.9572 \text{ Å} \quad r_{OO} = 1.5139 \text{ Å}. \quad (4.21)$$

I parametri A e B sono i parametri di L-J, che in questo modello vengono settati a:

$$A = 582.0 \cdot 10^{-3} \text{ kcal Å}^{12} / \text{mol} \quad B = 582.0 \text{ kcal Å}^6 / \text{mol}$$

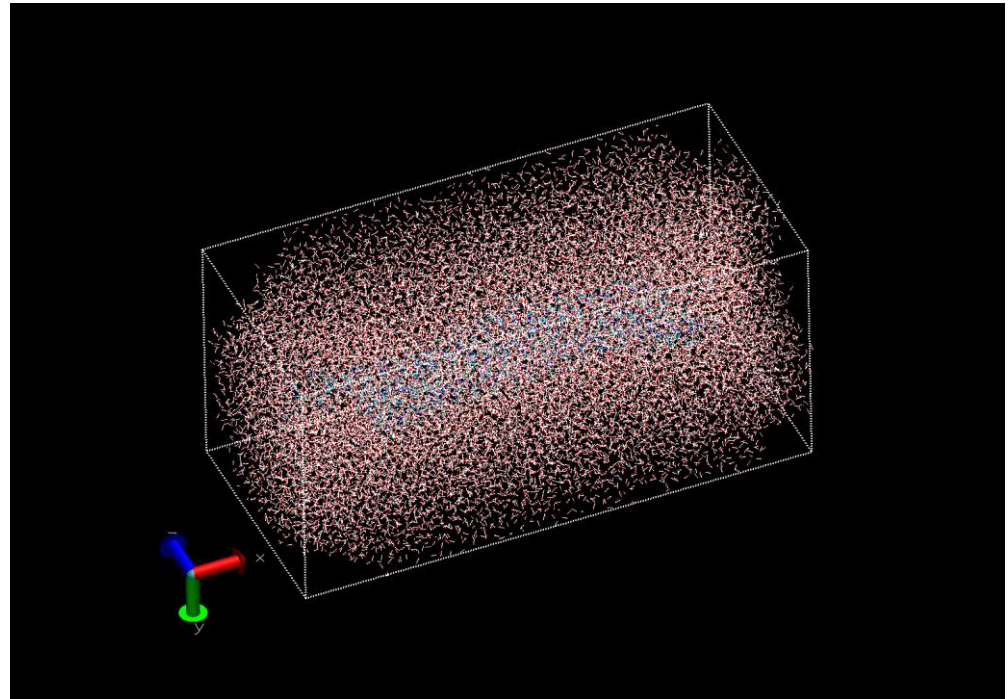
Esempio:
scatola periodica di molecole d'acqua

$P = 1 \text{ atm}$

$T = 300 \text{ K}$

Dimensione della scatola periodica: $60 \text{ \AA}$

N di molecole = **22487**



l'amiloide $\beta 42$ ($A\beta 42$). Questo peptide, lungo 42 aminoacidi, originato dall'*Amyloid Precursor Protein* (APP), è il maggior costituente delle placche amiloidee ritrovate dall'analisi *post mortem* del cervello di pazienti affetti da morbo di Alzheimer.

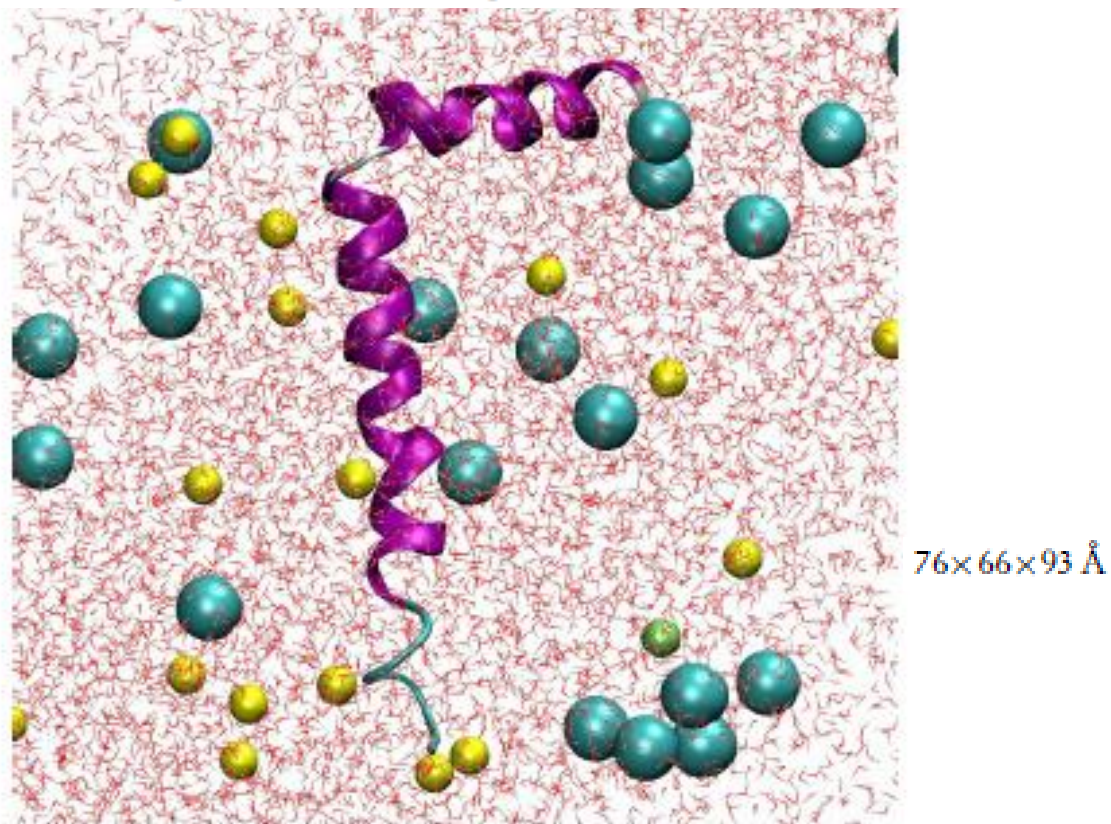
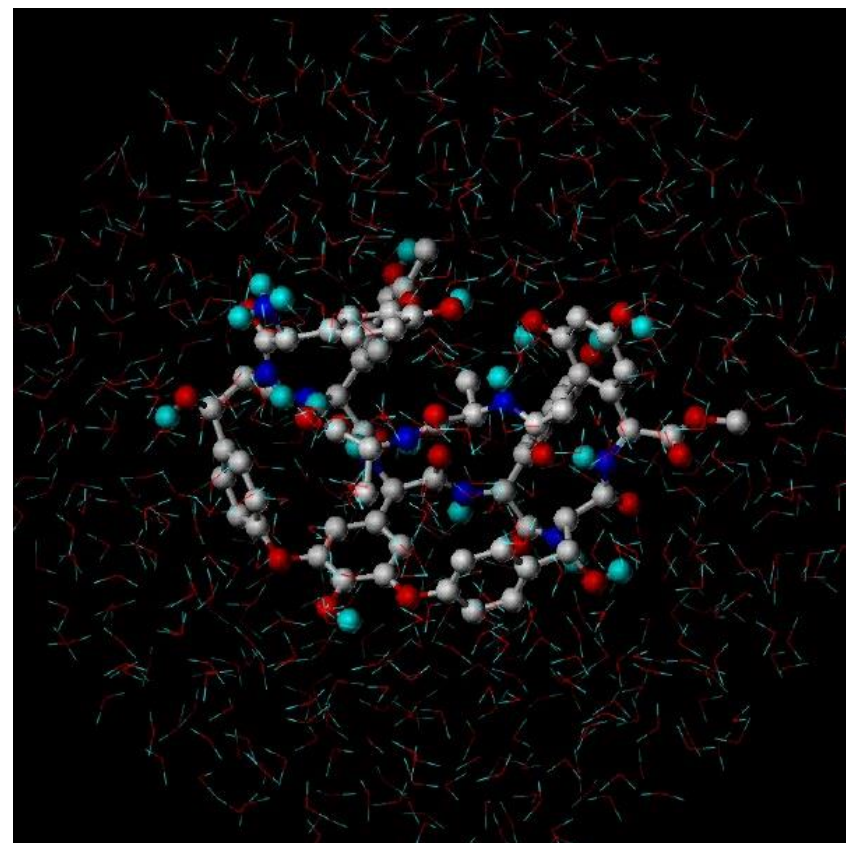
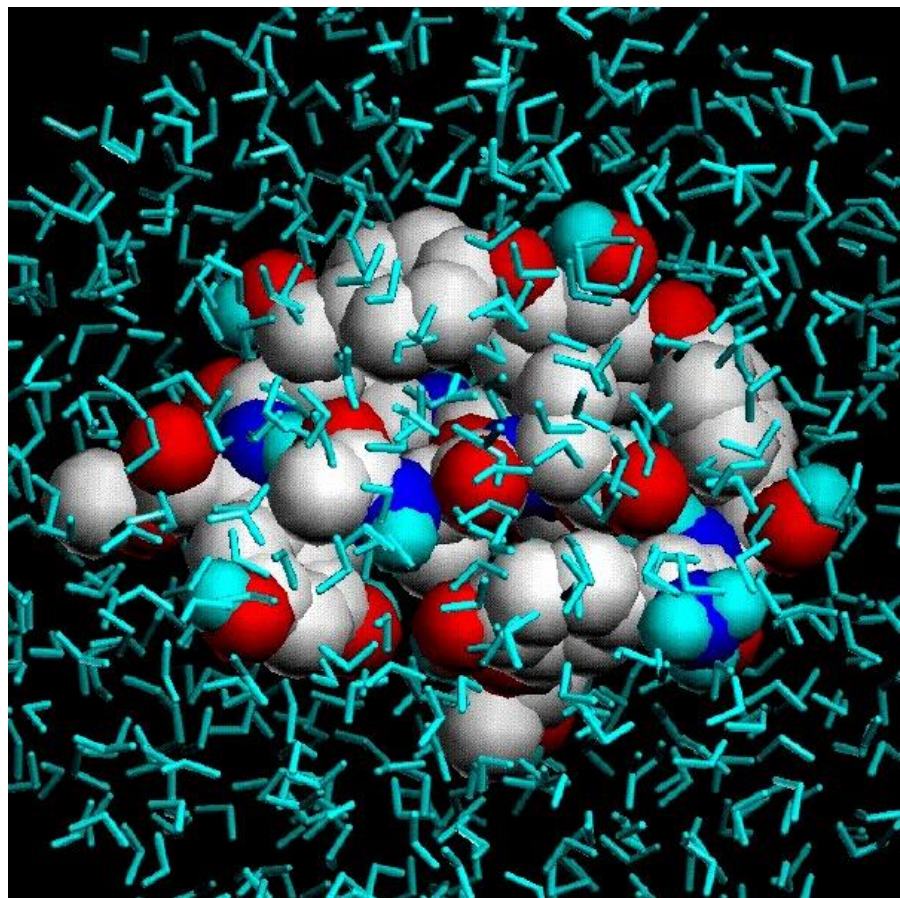
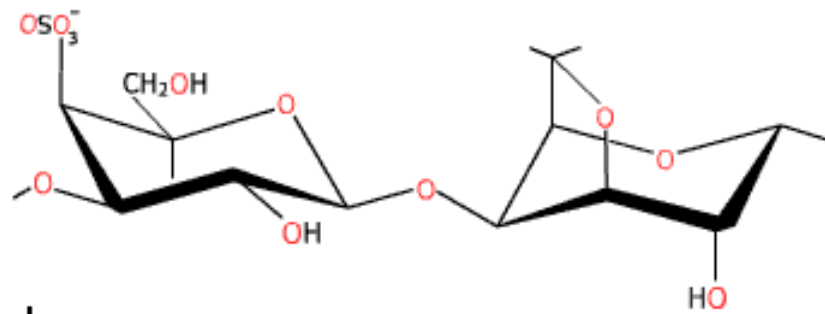


Figura 2.3: Set-up molecolare di partenza dell' $A\beta 42$ in soluzione acquosa che mima l'ambiente extra-cellulare. Impostando un valore di Forza Ionica di 140mM in NaCl, il box d'acqua era sufficientemente grande da contemplare anche la presenza di uno ione Ca^{2+} (sfera di Van der Waals color verdino). In giallo sono rappresentate come sfere di Van der Waals i cationi Na^{+} ; in celeste gli anioni Cl^{-} . In viola sono rappresentate le due α -eliche del motivo *Helix-Kink-Helix* della struttura 1IYT così come ottenuta da NMR. Quest'immagine è stata ottenuta con VMD [31]. La struttura secondaria viene assegnata con il plugin STRIDE [32].



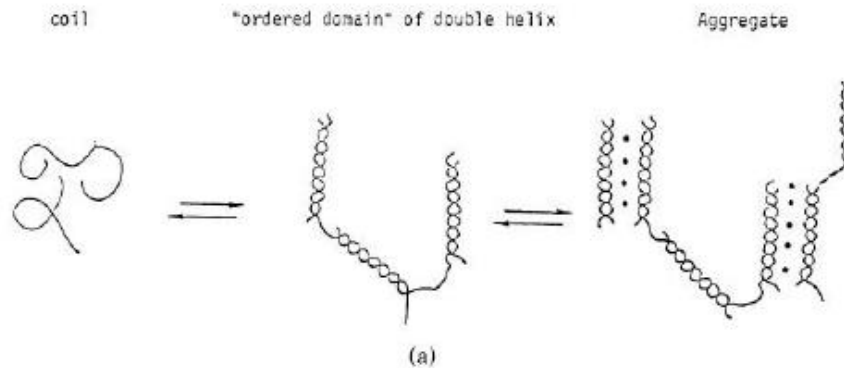
K-Carrageenan



kappa

Il «caso» del carragenano

Modello di Morris



Modello di Smidsrød

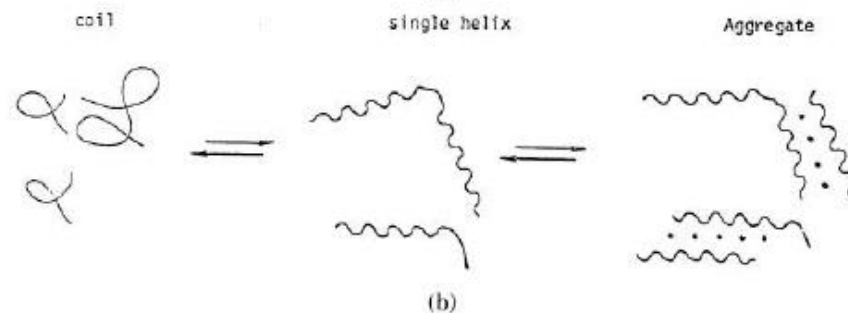


Fig. 1. Schematic models for gel formation in κ -carrageenan (a) by Morris et al. (Ref. 3) and (b) by Smidsrød (Ref. 4). K^+ counterions indicated by solid circles (●).

Molecular Dynamics Simulation of a Double-Helical β -Carrageenan Hexamer Fragment in Water

K. Ueda*

Department of Physical Chemistry, Faculty of Engineering, Yokohama National University, 79-5 Tokiwadai, Hodogaya-ku, Yokohama 240-8501, Japan

A. Imamura

Department of Chemistry, Faculty of Science, Hiroshima University, Kagamiyama, Higashi-Hiroshima 739, Japan

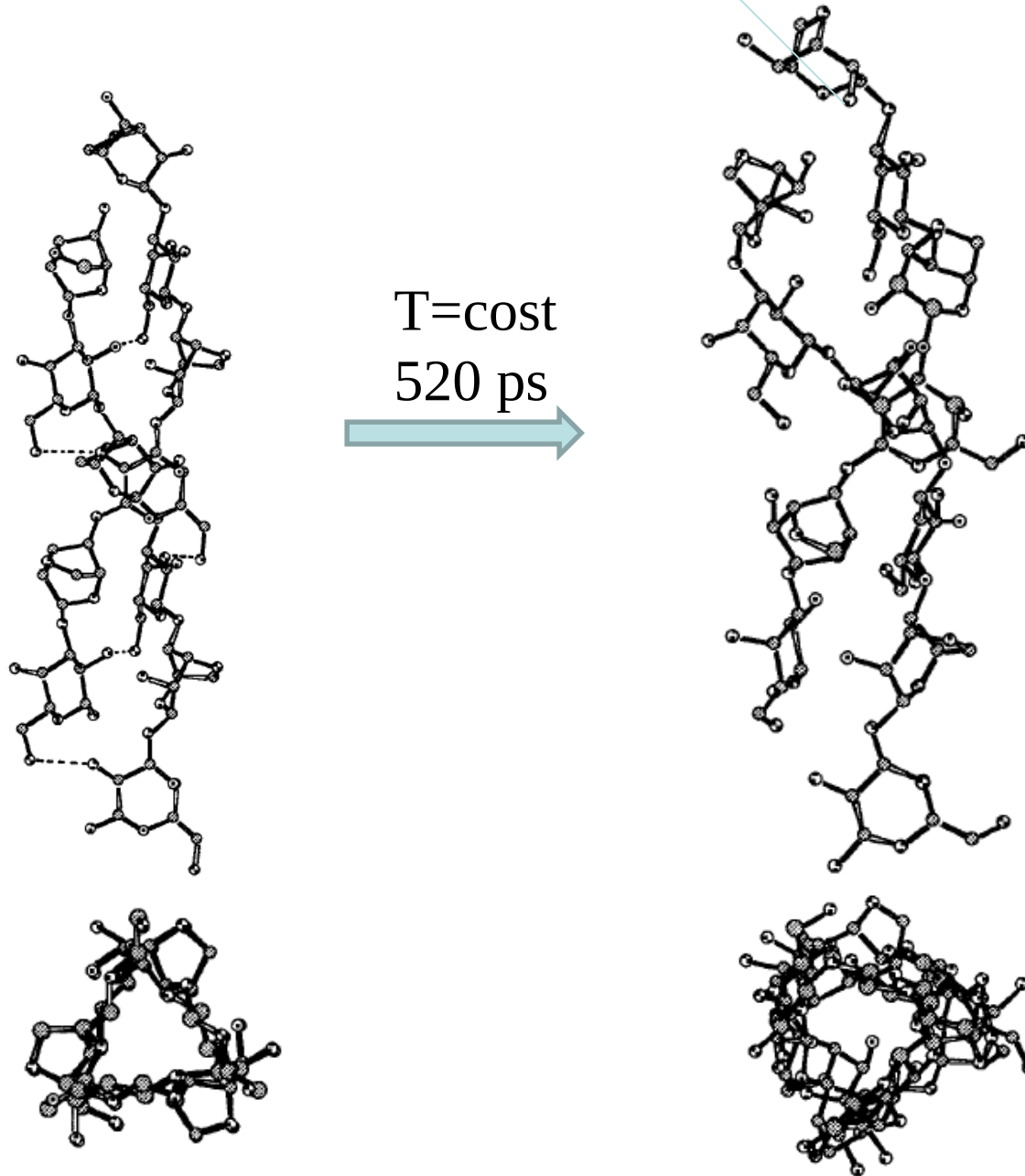
J. W. Brady*

Department of Food Science, Stocking Hall, Cornell University, Ithaca, New York 14853

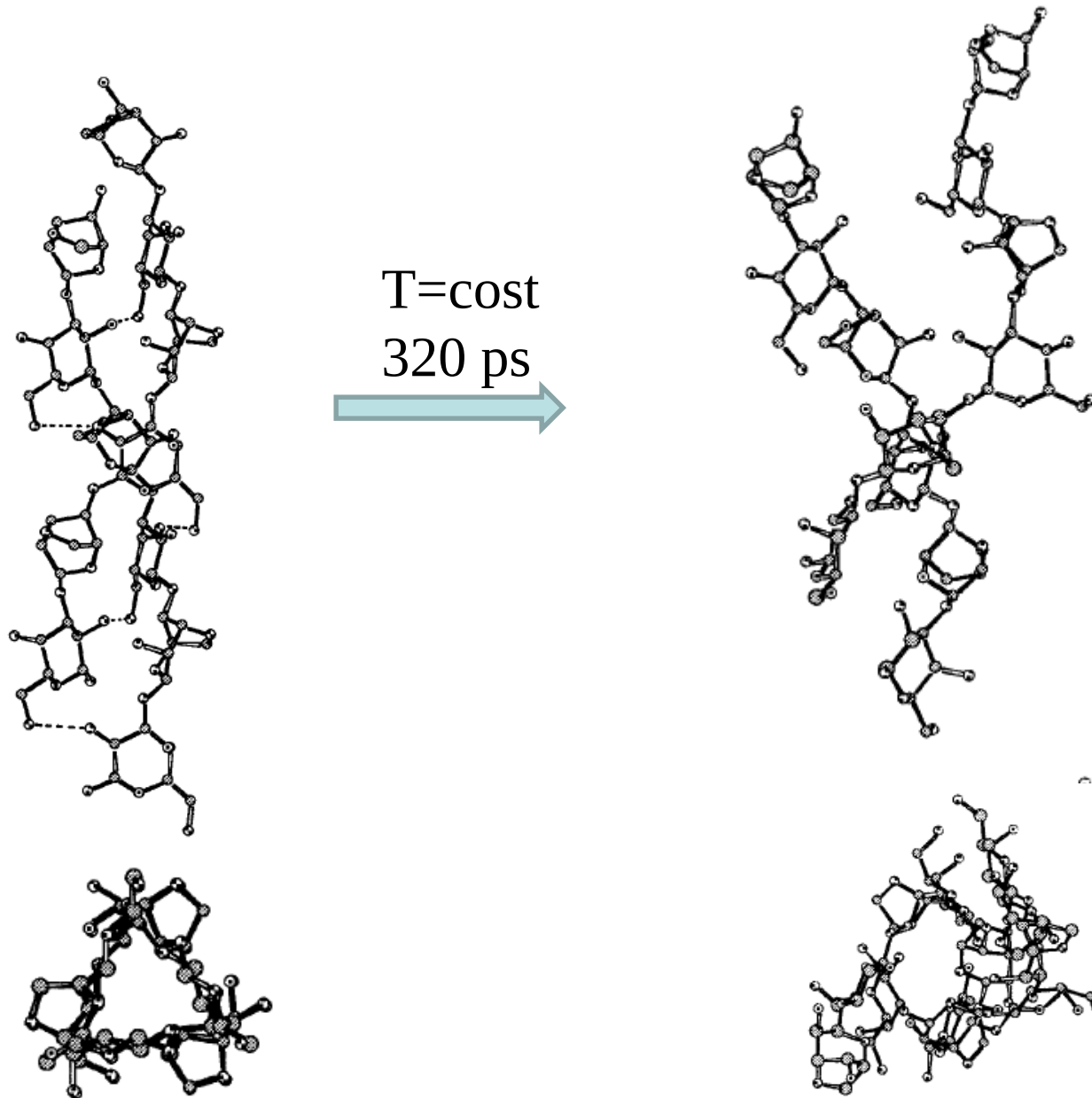
Received: July 16, 1997; In Final Form: February 18, 1998



Simulazione MD di k-carragenano **nel vuoto**



Simulazione MD di k-carragenano **in acqua**



Dinamica Molecolare del DNA

Le difficoltà del calcolo sono dovuti principalmente al fatto che gran parte delle trasformazioni molecolari avviene su scale di tempo da **tre a sei ordini di grandezza** più lunga delle più lunghe simulazioni MD di un sistema dalle dimensioni tipiche di un gene (ns).

Ad esempio:

- l'RNA polimerasi trascrive il DNA ad un tasso di circa 14 ms/nucleotide negli eucarioti;
- processi «veloci» come l'espulsione del DNA dei batteriofagi hanno tempi di traslocazione superiore a 10 s/bp;
- scale temporali per la strutturazione del DNA nel capsido virale sono di 10 ms/ bp;
- La formazione dei nucleosomi ha scale temporali dell'ordine dei 10 ms.

Per affrontare e semplificare uno qualsiasi di questi fenomeni biologicamente rilevanti il calcolo richiede l'introduzione di modelli «**coarse-grained**».

Simulazione del DNA

Molteplici modellizzazioni di differenti tipologie di acidi nucleici sono disponibili in letteratura ed il loro utilizzo dipende in modo cruciale da quale caratteristica si vuole simulare nel sistema

«full-atoms»

completa rappresentazione atomica, nella quale tutti gli atomi, solventi e ioni inclusi, sono considerati esplicitamente in un database che descrive con precisione la struttura terziaria della molecola

«coarse-grained»

gruppi di atomi sono modellizzati mediante la costruzione di unità fondamentali (**beads**) connesse tra loro a formare una catena ed aventi delle determinate proprietà di interazione tra i suoi costituenti (utilizzo di force-field particolari)



Ad una prima analisi superficiale, l'approccio di MD **full-atom** sembrerebbe essere la migliore tecnica possibile per simulare il comportamento di un sistema atomico (molti più dettagli chimici)

Ma nella pratica, la presenza di un gran numero di gradi di libertà che debbono essere presi in considerazione la **MD** risulta computazionalmente molto esigente anche per la simulazione di piccoli sistemi (N_{sys} attorno ai 20,000 atoms) formati da doppi filamenti di DNA corti immersi in soluzione acquosa, limitando di fatto le scale di lunghezza e tempo che possono essere investigate.

In un modello **coarse-grained** invece, la descrizione del DNA può essere semplificata riducendo in modo significativo il numero di gradi di libertà, considerando l'effettiva interazione tra unità coarse-grained al posto della più dettagliata interazione full-atom: l'obiettivo è di includere il **minimo** livello di dettaglio richiesto per descrivere tutte le proprietà fisiche delle strutture da analizzare.

I parametri del force-field, che in questo caso non modellizzano le interazioni tra atomi bensì tra le unità che costituiscono il modello, utilizzati nell'approccio coarse-grained sono usualmente scelti in accordo con i dati sperimentali a disposizione.

Modelli COARSE-GRAINED

modelli più raffinati approssimano il DNA a singolo filamento come una catena semiflessibile di sfere stericamente repulsive, che possono portare carica e quindi interagire con un campo esterno, mentre le interazioni monomero-monomero sono state trascurate in questi modelli.

Knotts et al. (2007) hanno introdotto un raffinato modello «bead-spring» di sfere legate in modo elastico e che distingue i gruppi di zucchero e fosfato sulla catena principale e le singole molecole di base.

Un modello di DNA in cui il fosfato e lo zucchero i gruppi sono rappresentati da un residuo sferico e le basi da un ellissoide rigido.

Le interazioni che descrivono legami covalenti, angoli e diedri del DNA sono ottenuti da simulazioni *all-atoms* all'equilibrio di filamenti corti di DNA

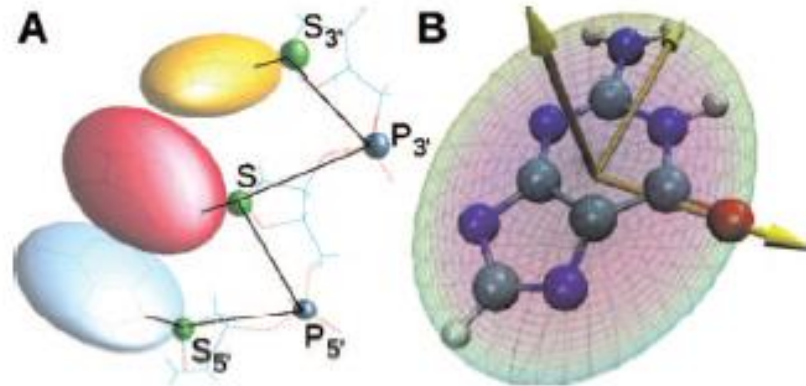


FIG. 1. Our coarse-graining scheme overlaid onto an all-atom visualization. Panel (a) shows a section of three nucleotides: spherical sugar (s) and phosphate (p) residues together with ellipsoidal bases. Panel (b) shows a close up of a base with ellipsoid overlay. The axes indicate the principal axes of the base and are aligned with the principal axes of the ellipsoid, which determine the base's initial orientation.

$$|V_{\text{total}} = V_{\text{bond}} + V_{\text{angle}} + V_{\text{dihedral}} + V_{\text{stack}} + V_{\text{bp}} + V_{\text{ex}} + V_{\text{qq}}, \quad (1)$$

where

$$V_{\text{bond}} = \sum_i^{N_{\text{bond}}} [k_1(d_i - d_{0_i})^2 + k_2(d_i - d_{0_i})^4], \quad (2a)$$

$$V_{\text{angle}} = \sum_i^{N_{\text{angle}}} \frac{k_\theta}{2} (\theta_i - \theta_{0_i})^2, \quad (2b)$$

$$V_{\text{dihedral}} = \sum_i^{N_{\text{dihedral}}} k_\phi [1 - \cos(\phi_i - \phi_{0_i})], \quad (2c)$$

$$V_{\text{stack}} = \sum_{i < j}^{N_{\text{st}}} 4\epsilon \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right], \quad (2d)$$

$$V_{\text{bp}} = \sum_{\text{base pairs}}^{N_{\text{bp}}} 4\epsilon_{\text{bp}_i} \left[5 \left(\frac{\sigma_{\text{bp}_i}}{r_{ij}} \right)^{12} - 6 \left(\frac{\sigma_{\text{bp}_i}}{r_{ij}} \right)^{10} \right], \quad (2e)$$

$$V_{\text{qq}} = \sum_{i < j}^N \frac{q_i q_j}{4\pi\epsilon_0\epsilon_k r_{ij}} e^{-r_{ij}/\kappa_D}.$$

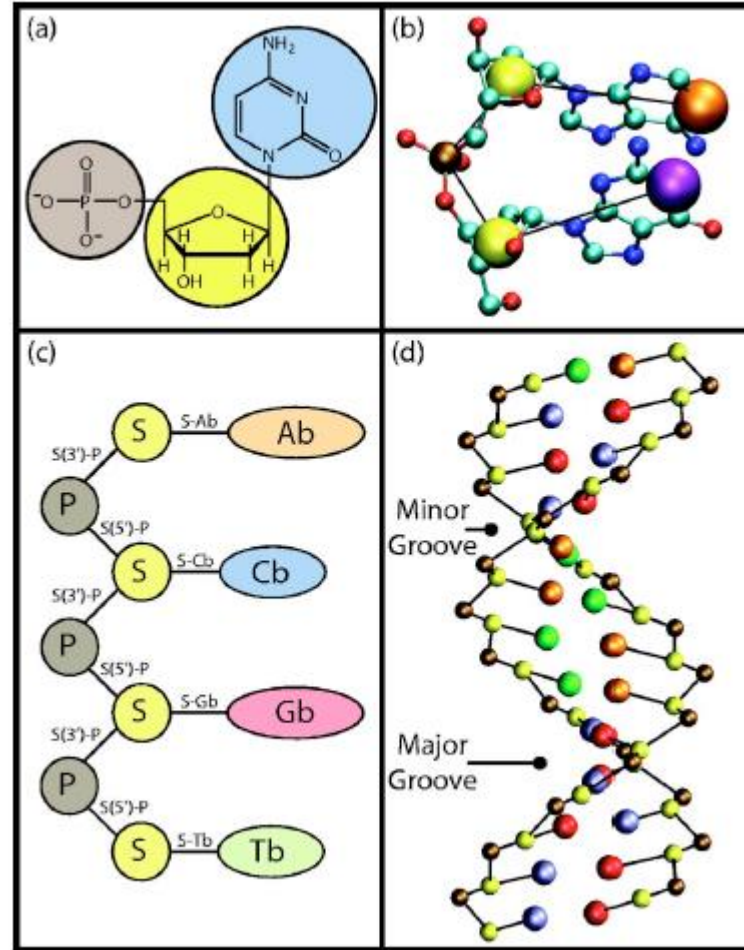


FIG. 1. Schematic representation of the mesoscale model of DNA. Panel (a) Grouping of the atoms for each coarse grain site for a cytosine nucleotide. Panel (b) Atomistic to coarse grain site mapping for the monophosphate dinucleotide 5'-GA-3'. Panel (c) Single-strand topology illustrating the polarity of the strands. Panel (d) Model of a 13 bp oligonucleotide.

TABLE III. Values for geometric parameters found in the potential energy function. A phosphate can bind to a sugar in either a 5' or a 3' sense. [See panel (c) of Fig. 1.] Thus, S(5')-P represents a bond between a phosphate and a sugar belonging to the same nucleotide while S(3')-P joins together neighboring residues. The bond angle P-(5')S(3')-P consists of both types of bonds. [Note: S(5')-P=P-(5')S.]

Bond	d_0 (Å)	Bond angle	θ_0 (degree)
S(5')-P	3.899	S(5')-P-(3')S	94.49
S(3')-P	3.559	P-(5')S(3')-P	120.15
S-Ab	6.430	P-(5')S-Ab	113.13
S-Tb	4.880	P-(3')S-Ab	108.38
S-Cb	4.921	P-(5')S-Tb	102.79
S-Gb	6.392	P-(3')S-Tb	112.72
		P-(5')S-Cb	103.49
		P-(3')S-Cb	112.39
		P-(5')S-Gb	113.52
		P-(3')S-Gb	108.12
Dihedral angle	ϕ_0 (degree)	Nonbonded	Length (Å)
P-(5')S(3')-P-(5')S	-154.80	σ_{ij}	Interaction specific
S(3')-P-(5')S(3')-P	-179.17	σ_{bpAT}	2.9002
Ab-S(3')-P-(5')S	-22.60	σ_{bpGC}	2.8694
S(3')-P-(5')S-Ab	50.69	σ_0 (mismatched bases)	$2^{-1/6}(1.0)$
Tb-S(3')-P-(5')S	-33.42	σ_0 (otherwise)	$2^{-1/6}d_{\text{cut}}$
S(3')-P-(5')S-Tb	54.69		
Cb-S(3')-P-(5')S	-32.72	d_{cut}	$\langle\sigma_{ij}\rangle \approx 6.86$
S(3')-P-(5')S-Cb	54.50		
Gb-S(3')-P-(5')S	-22.30		

TABLE II. Values for energy parameters found in the potential energy function.

Parameter	Value
k_1	ϵ
k_2	100 ϵ
k_θ	400 $\epsilon/(\text{radian})^2$
k_φ	4 ϵ
ϵ	0.26 kcal/mol
ϵ_{bpGC}	4 ϵ
ϵ_{bpAT}	$\frac{2}{3}\epsilon_{\text{bpGC}}$

Approccio Bottom-Up

In ambito informatico la strategia di programmazione bottom-up è una tecnica nella quale le singole parti che costituiscono un sistema sono esaminate nel dettaglio.

Successivamente tutti questi costituenti elementari vengono connessi tra loro a formare delle parti più grandi e così via, fino ad arrivare alla descrizione completa dell'intero sistema

Approccio Top-Down

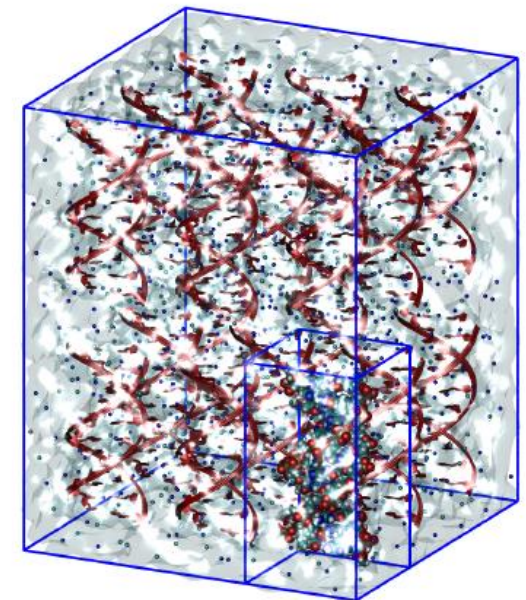
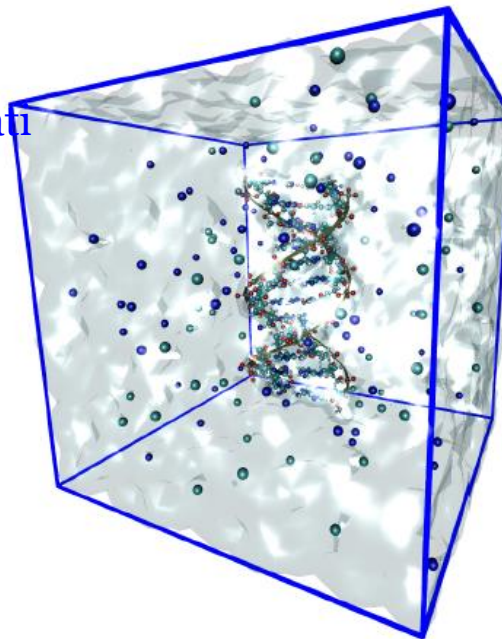
Viceversa nella teoria dell'informazione, il modello di programmazione dall'alto verso il basso chiamato approccio top-down, prevede un'analisi del problema che parte dapprima con l'esaminare il sistema nel suo insieme, formulando una visione generale e successivamente suddivide il sistema in parti più piccole, raffinando via via la modellizzazione con l'aggiunta di maggiori dettagli.

Contrariamente a quanto avviene nel modello bottom-up, nell'approccio top-down i force-field sono disegnati per riprodurre con ragionevole accuratezza una larga serie di proprietà termodinamiche, come ad esempio la temperatura di fusione dell'elica

Si eseguono diversi tipi di simulazione a seconda delle caratteristiche che si desiderano investigare distinguendo due categorie:

ESEMPIO:

- Simulazioni a singola molecola (sDNA), effettuate all'interno di box cubici a diverse dimensioni;
- Simulazioni a 18 DNA (multiDNA), ottenute replicando un box tetragonale contenente una singola molecola tre volte in ciascuna direzione x ed y e due volte lungo z, ottenendo così sistemi formati da $N_{\text{DNA}} = 3 \times 3 \times 2 = 18$.



Si aggiunge **sale**, inserendo separatamente ioni di cloro Cl^- e di sodio Na^+ , fino ad arrivare a concentrazioni 1M.

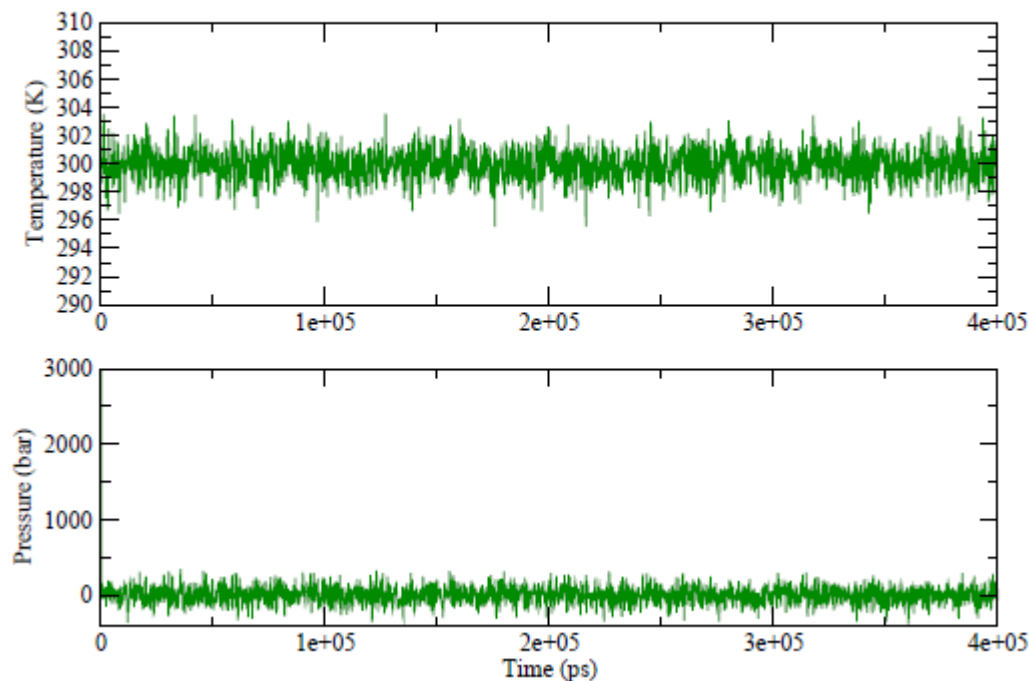
Questo accorgimento è particolarmente utile soprattutto nelle simulazioni a singola molecola per schermare elettrostaticamente il sistema limitando eventuali interazioni forti con se stesso dovute alle condizioni periodiche al bordo della cella

Prima di iniziare la simulazione vera e propria è necessario far rilassare il sistema per scongiurare eventuali posizionamenti degli atomi in posizioni sfavorevoli dal punto di vista energetico (minimizzazione dell'energia potenziale).

Prima di lanciare qualsiasi simulazione è necessario controllare che tali valori di energia raggiungano un **minimo**, altrimenti rischiamo di far partire la simulazione da una condizione di equilibrio instabile che porterebbe all'inutilizzabilità dei dati ottenuti.

Prima di passare all'analisi dati vera e propria è necessario controllare che tutte le variabili termodinamiche del sistema si mantengano all'equilibrio lungo tutta la durata della simulazione.

NVT Parameters - multiple ACC

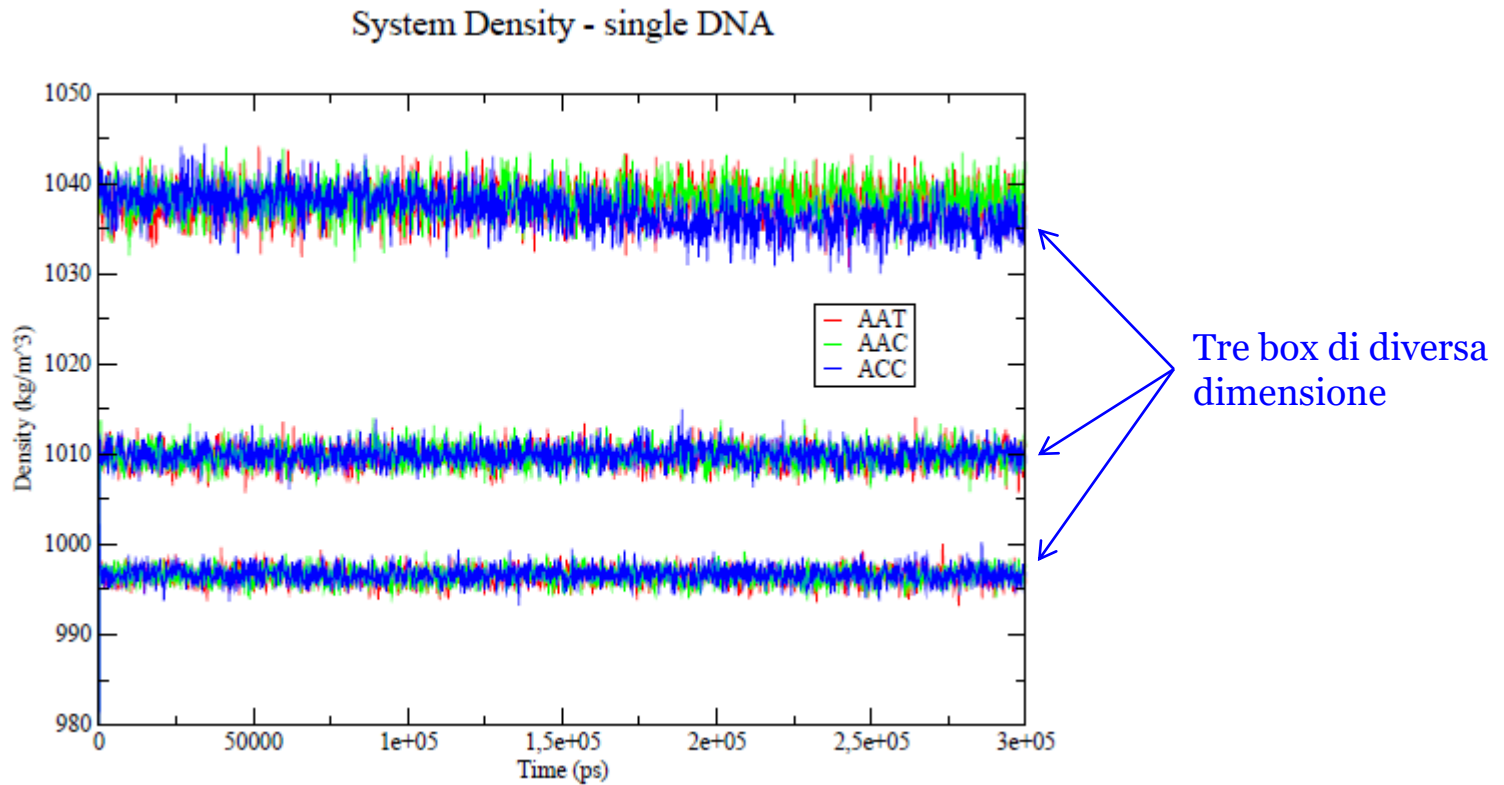


$$T_{system} = (299.929 \pm 0.002) K$$

$$P_{system} = (1.031 \pm 0.038) bar$$

$$E_{system}^{1DNA} = - (4.096 \pm 0.003) \cdot 10^5 \frac{kJ}{mol}$$

$$E_{system}^{18DNA} = - (1.505 \pm 0.005) \cdot 10^6 \frac{kJ}{mol}$$



Nei casi di una singola sequenza osserviamo che la densità è grossomodo stabile per tutte le dimensioni del box prese in esame ed inoltre non riscontriamo significative differenze al variare dell'acido nucleico.

differente è la situazione per quanto riguarda il più realistico sistema costituito da **18 DNA**: la densità del sistema risulta dipendente dalla sequenza presa in esame e dai grafici si evince chiaramente che il sistema costituito da acidi nucleici di tipo ACC, dovrebbe transire a concentrazioni maggiori, si stabilizza a densità maggiori.

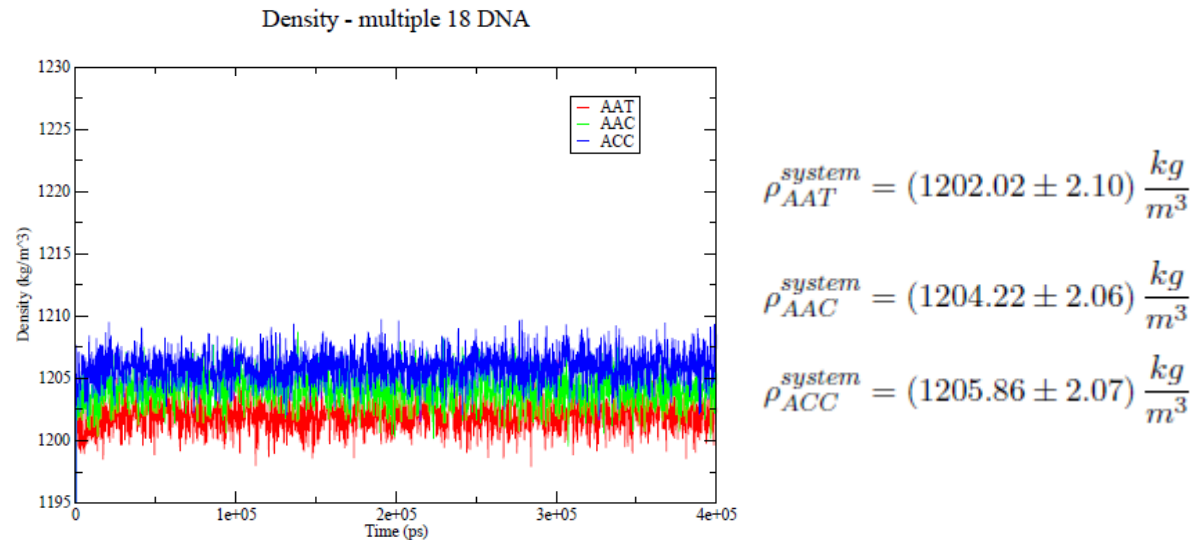


Figura 7.3: Fluttuazioni delle densità nei tre sistemi composti da 18 DNA ciascuno. Risulta particolarmente evidente come i valori di ρ si attestino rapidamente ai valori di equilibrio (da una stima grafica dopo 20ns circa il sistema ha raggiunto la densità media).

supponendo che le tre tipologie di sequenze abbiano lo stesso peso molecolare, questo comportamento ci fornisce un primo indicatore del fatto che la sequenza ACC possa essere la più piegata tra le tre sequenze prese in esame poiché una struttura piegata è geometricamente più compatta, quindi a densità maggiore, della sua controparte rettilinea.

Prima di passare all'analisi diretta delle osservabili bisogna lasciare al sistema tempo sufficiente per raggiungere l'equilibrio conformazionale (della sua struttura terziaria) per poi considerare nel calcolo delle medie solo i valori successivi a tale tempo di equilibrizzazione.

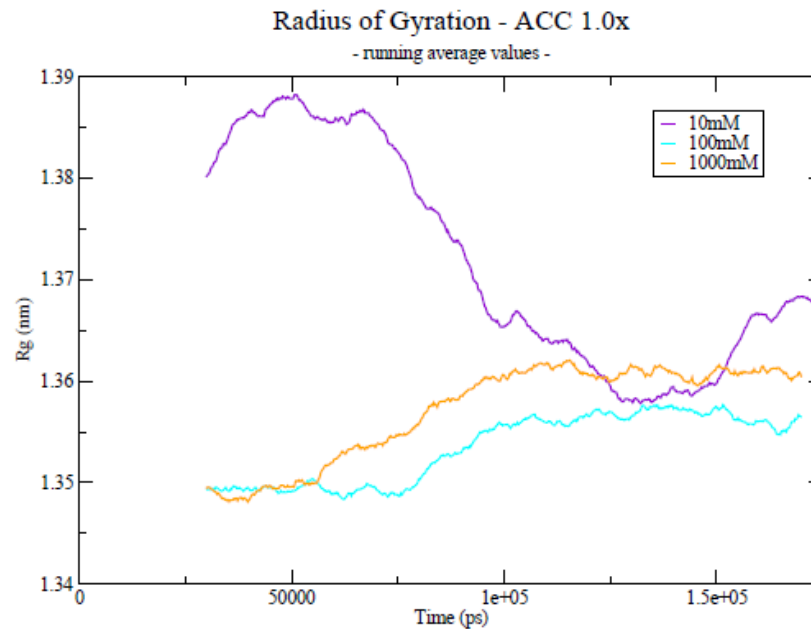
Una delle grandezze medie calcolate è il raggio di girazione:

$$R_g(t) = \sqrt{\frac{\sum_i \|\mathbf{r}_i(t)\|^2 m_i}{\sum_i m_i}}.$$

che ci permette di comprendere la «compattezza» di una molecola rispetto al suo centro di massa, un indicatore spesso utilizzato per caratterizzare macromolecole in soluzione.

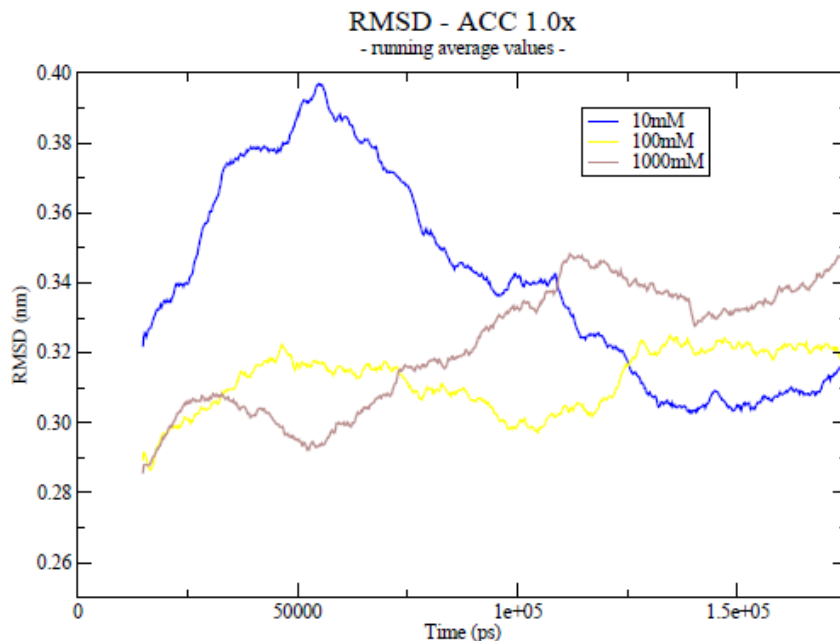
Dal grafico si osserva che il sistema immerso in soluzione acquosa con concentrazione salina pari a 10 mM nell'intervallo di tempo di 200 ns non raggiunge nessuna configurazione di equilibrio.

Al crescere della concentrazione salina, sia per 100 mM che per 1000 mM il raggio di girazione comincia a stabilizzarsi dopo circa 100 ns di simulazione.



la radice dello scarto quadratico medio (r.m.s.d.) della struttura rispetto alla struttura di riferimento iniziale, ottenuta calcolando la somma di tutti gli spostamenti quadratici medi di ogni singolo atomo dell'acido nucleico in funzione del tempo.

$$RMSD(t_1, t_2) = \left[\frac{1}{M} \sum_{i=1}^N m_i \| \mathbf{r}_i(t_1) - \mathbf{r}_i(t_2) \|^2 \right]^{\frac{1}{2}}$$



Con questa quantità possiamo stabilire effettivamente quando la forma della molecola comincia a stabilizzarsi in una struttura ben definita (*conformazione più stabile*).

In figura si osserva che la sequenza in soluzione 1 M raggiunge il plateau della media della RMSD a circa 100 ns mentre la soluzione 100 mM a circa 125 ns.

La sequenza a 10 mM presenta ampie fluttuazioni anche andando a tempi più lunghi (fino a 300 ns) dovuta ad uno stato di non-equilibrio derivato dalla scarsa schermatura delle cariche a tale forza ionica.

Si possono stimare i tempi entro i quali il sistema comincia a diffondere.
Dalla relazione di Einstein il coefficiente di diffusione è:

$$6Dt = \frac{1}{N} \sum_{i=1}^N [r_i(t) - r_i(0)]^2$$

Per la sequenze, eseguendo un fit lineare dei valori si ottengono i seguenti risultati per i coefficienti di diffusione per le tre forze ioniche

$$D_{10mM}^{ACC} = (0.2660 \pm 0.0082) \cdot 10^{-5} \frac{cm^2}{s}$$

$$D_{100mM}^{ACC} = (0.0994 \pm 0.1766) \cdot 10^{-5} \frac{cm^2}{s}$$

$$D_{1000mM}^{ACC} = (0.1214 \pm 0.0609) \cdot 10^{-5} \frac{cm^2}{s}.$$

Pertanto i tempi caratteristici con cui il sistema raggiunge una conformazione stabile, possono essere stimati come il reciproco del coefficiente di diffusione:

$$t \sim \frac{(1nm)^2}{D_A} \rightarrow t_{diffusione} \sim 10ns$$

I tre sistemi analizzati diffondono dopo tempi dell'ordine dei 10 ns.