Corso di Laurea Specialistica in Chimica delle Molecole di Interesse Biologico

Struttura e Dinamica di Biomolecole - 2008/2009

Laboratorio 7 - (12 maggio 2009)

---

1 Risultati delle simulazioni di 50 ns

Anche la simulazione più lunga dopo circa $20ns$ si porta in una configurazione non-nativa anche se non totalmente elongata

Va considerato che a $300K$ sperimentalmente la frazione di folded è 50%.

Correzioni da fare:

1. portare temperatura a 277 K
2. cambiare potenziale
3. abbassare frequenza del termostato (in certi casi con $300cm^{-1}$ l'integrazione numerica dà errore)

---

2 Calcolo del ΔG di *unfolding/refolding* da simulazioni di Dinamica Guidata

Usiamo l'uguaglianza di Jarzynski applicata ad una distribuzione gaussiana del lavoro $A \to B$:

$$\Delta G = \bar{w} - \frac{\beta\sigma^2}{2}$$

dove

$$\bar{w} = \frac{1}{N}\sum w_i$$

$$\sigma = \sqrt{\frac{1}{N-1}\sum(w_i - \bar{w})^2}$$

Nella versione `5` di `ORAC`, è possibile eseguire in parallelo una serie di processi $A \to B$. Si fa in due fasi:

1. campionamento dei microstati $A_1, A_2, \ldots, A_N$ dallo stato termodinamico $\mathcal{A}$
2. esecuzione di N processi di trasformazione $A_i \to B_i$ (N.B. la distribuzione dei B_i non verrà di equilibrio)
3. analisi dei risultati e applicazione uguaglianza di Jarzynski

2.1 campionamento dello stato di equilibrio di partenza

- si fa una simulazione normale a partire da una configurazione dello stato di partenza
- si salvano configurazioni (come dati di *restart*) ad intervalli regolari

```
&INOUT
    ...
    RESTART
       write 1000.0 SAVE_ALL_FILES ../RESTART_A/native
    END
&END
```

L'intervallo deve essere sufficiente a ridurre al minimo la correlazione tra uno stato e il successivo; ovviamente il run deve essere lungo abbastanza da accumulare almeno N configurazioni

- Per comodità si salvano anche topologia e parametri in un file "binario"

```
&PARAMETERS
    READ_TPG_ASCII ../../lib/amber03-release.tpg
    READ_PRM_ASCII ../../lib/amber03-release.prm
    WRITE_TPGPRM_BIN ../chigno_A.prmtpg
    ...
&END
```

- Il programma si lancia in "parallelo" su un solo processore:

```
mpiexec -n 1 orac-p < jarz-p-pre.in
```

  Il programma gira sulla sottodirectory `PAR0000` (è a partire da questa directory che si intendono i percorsi relativi dei file specificati nell'input!).

  Le configurazioni sono salvate sulla directory `../RESTART_A` come file `native0000.rst`, `native0001.rst`, etc.

2.2 esecuzione degli N processi in parallelo

- distruggere `PAR0000/`
- creare un unico input che verrà copiato nelle directory `PAR0000`, `PAR0001`, ...:
  - le simulazioni partono dai file di restart e dal file binario topologia-parametri quindi non si devono mettere i blocchi
    - `&SETUP`
    - `&SOLVENT`
    - `&SOLUTE`

    ma solo

    ```
    &PARAMETERS
        READ_TPGPRM_BIN ../chigno_A.prmtpg
    &END
    ```
  - oltre al consueto input per Jarzynski si deve aggiungere la lettura dei file di restart (eventualmente a partire da un certo indice)

    ```
    &INOUT
        ASCII 9000.0 OPEN jarz-p.pdb
        RESTART
          rmr ../RESTART_A/native 10
        END
        PLOT STEER_ANALYTIC 900.0 OPEN WRKa.1
    &END
    ```
- Il programma si lancia in parallelo su N processori:

```
mpiexec -n 64 orac-p < jarz-p.in
```

  creerà N sottodirectory chiamate `PAR0000`, `PAR0001`, etc.

2.3 analisi e applicazione formula di Jarzynski

1. Il lavoro totale di una traiettoria è riportato nell'ultima riga del file relativo, ultima colonna.

   Si può creare un file con tutti i lavori finali con il comando:

```
tail -n 2 PAR*/p.WRK| awk '/bond/ {print $7}' > dw
```

2. Calcolare media e σ:

$$\bar{w} = \frac{1}{N}\sum w_i$$

$$\sigma = \sqrt{\frac{1}{N-1}\sum (w_i - \bar{w})^2}$$

Media e σ si possono calcolare anche direttamente con

```
avg dw
```

3. Calcolare la differenza di energia libera:

$$\Delta G = \bar{w} - \frac{\beta\sigma^2}{2}$$

Tenendo conto che a $T = 300$, $\beta^{-1} = 2.5\ kJ/mol$, questa formula si può programmare con

```
avg dw |awk 'NR==2 {print $2-.2*$4}'
```