



Universitat de Barcelona

Facultat de Química

Departament de Química Física

i

Parc Científic de Barcelona

Centre especial de Recerca en Química Teòrica

**MECANISMES DE PREACTIVACIÓ DE
SUBSTRAT EN 1,3-1,4- β -GLUCANASA**
MODELITZACIÓ MITJANÇANT DINÀMICA MOLECULAR
DE PRIMERS PRINCIPIS

Memòria presentada per XEVI BIARNÉS FONTAL
per tal d'optar al títol de Doctor per la Universitat de Barcelona.
Programa de doctorat de Biotecnologia (bienni 2003-2005).

Aquesta tesi ha estat realitzada al Parc Científic de Barcelona, sota la direcció
de la Dra. Carme Rovira i Virgili, Professora d'Investigació ICREA,
comptant amb la tutoria del Dr. Fausto Sanz Carrasco, Catedràtic del Departament de
Química Física de la Universitat de Barcelona.

Barcelona, setembre de 2007

capítol II

Fonaments Metodològics

II.1. Mètode *Car-Parrinello* de Dinàmica Molecular (CPMD)

El mètode de *Car-Parrinello* és un mètode de dinàmica molecular de primers principis¹¹⁴ basat en la Teoria del Funcional de la Densitat (DFT). La DFT proveeix d'un esquema per a calcular l'energia total d'un sistema poliatòmic a partir de les seves coordenades atòmiques (R_i), essent una eina força útil per a l'estudi de processos en que hi ha importants reorganitzacions electròniques, com ara la catàlisi enzimàtica i els processos d'interacció proteïna-l·ligand¹¹⁹. Segons aquesta teoria, l'energia de l'estat fonamental d'un sistema d'electrons que interactuen entre ells i que estan subjectes a un potencial extern $V(r)$ és una única funció de la densitat electrònica ($\rho(r)$) i que pot ser obtinguda mitjançant la minimització del funcional de l'energia respecte la densitat¹⁰⁴:

$$E^{DFT} = \min_{\rho(r)} E[R_i, \rho(r)] \quad \text{eq.8}$$

Respecte d'altres tècniques de la química quàntica, la DFT té l'avantatge que presenta una bona relació entre la precisió el cost computacional. És per això que es la base de la majoria de mètodes de dinàmica molecular de primers principis. L'esquema general del funcionament d'un mètode de dinàmica molecular basat en la DFT és el següent:

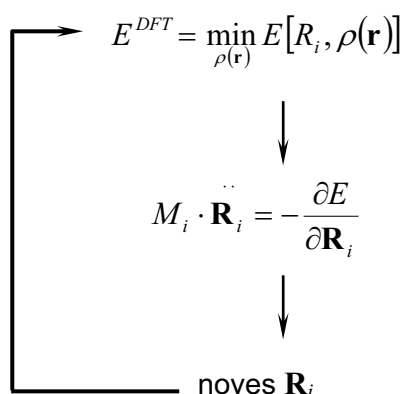


Figura II.1. Esquema general d'una dinàmica molecular de primers principis basada en DFT

En primer lloc, es minimitza el funcional de la densitat, on la densitat electrònica s'expressa en termes d'orbitals monoelectrònics $\{\psi_i(\mathbf{r})\}$ (anomenats orbitals de *Kohn-Sham*¹⁰⁵) i es calcula l'energia total del sistema. Posteriorment s'avaluen els gradients respecte les coordenades nuclears. En aquest punt, a través de la resolució de les equacions del moviment, s'obtenen les noves coordenades nuclears. A partir de la nova posició dels nuclis, cal repetir de nou el procés. Aquest mètode s'anomena també mètode de dinàmica molecular de *Born Oppenheimer* (BO-MD), en tant que durant tota la simulació, el sistema electrònic es manté en l'estat fonamental.

Computacionalment parlant, el procés de minimitzar el funcional i obtenir els nous orbitals monoelectrònics és el pas limitant en tot aquest procés. Com que això s'ha de realitzar moltes

vegades durant la simulació (per exemple, una simulació de 5 ps emprant un pas de simulació de 0.5 fs, requereix 10000 minimitzacions del funcional de la densitat), el cost computacional de la dinàmica molecular de primers principis (BO-MD) és molt més elevat que el d'una dinàmica clàssica (basada en l'ús del camp de forces).

Va ser arrel d'un treball del 1985 de *Car i Parrinello*^{iv} quan les tècniques de dinàmica molecular de primers principis van experimentar una empenta més important. La principal característica del mètode CP es que els orbitals de *Kohn-Sham* s'inclouen en les equacions del moviment com a nous graus de llibertat. Les equacions del moviment passen a dependre doncs dels nuclis i dels electrons, els quals evolucionen simultàniament. L'avantatge és que no cal minimitzar el funcional de la densitat a cada pas de simulació. Com a contrapartida, el pas de simulació ha de ser molt més petit (0.1 – 0.2 fs mentre que en una dinàmica BO-MD és de ~1 fs) a fi de descriure apropiadament el moviment del sistema electrònic. Els graus de llibertat electrònics seran els coeficients de l'expansió dels orbitals de *Kohn-Sham* en una base d'ones planes^v eq.9, que pren la següent forma:

$$\psi_i = \sum_j c_{ij} e^{i\mathbf{G}_j \cdot \mathbf{r}} \quad \text{eq.9}$$

on $\{c_{ij}\}$ són els corresponents coeficients de l'expansió, i $\{\mathbf{G}_j\}$ són vectors de la xarxa recíproca (s'assumeix condicions de periodicitat de contorn). Aquests coeficients aniran variant a mesura que la simulació avanci, és a dir, a mesura que s'integrin les equacions del moviment.

El mètode CP es basa en l'ús d'un *Lagrangià* estès (eq.10) on els graus de llibertat electrònics es tracten com a partícules clàssiques.

$$\mathcal{L} = E_N^{kin} + E_{el}^{kin} - E^{KS} + \sum_{ij} \Lambda_{ij} \left(\int d\mathbf{r} \psi_i^*(\mathbf{r}) \psi_j(\mathbf{r}) - \delta_{ij} \right) \quad \text{eq.10}$$

on $E_N^{kin} = \sum_N \frac{1}{2} M_N \dot{\mathbf{R}}_N^2$ és l'energia cinètica del nuclis, M_N i $\mathbf{R}_N$ són les masses i les posicions, respectivament; $E_{el}^{kin} = \sum \mu \int d\mathbf{r} |\dot{\psi}_i(\mathbf{r})|^2$ és el terme d'energia cinètica “fictícia” associat al subsistema electrònic $\{\psi_i(\mathbf{r})\}$, la massa fictícia μ és un paràmetre que controla la escala de temps del moviment electrònic (veure més endavant); E^{KS} és l'energia potencial electrònica calculada mitjançant DFT; Λ_{ij} són els multiplicadors de *Lagrange* associats a la

^{iv} “Unified approach for molecular dynamics and density-functional theory”, Phys. Rev. Lett. 1985.¹¹⁴
Actualment un dels 5 treballs més citats de la revista Phys. Rev. Lett.

^v l'ús d'una base d'ones planes simplifica el càlcul de les forces sobre els àtoms, que en una dinàmica CPMD es calculen més sovint que en una dinàmica BO.

necessitat de mantenir l'ortonormalitat dels orbitals al llarg de la simulació. L'energia total CP és una constant del moviment i ve donada per $E_{tot}^{CP} = E_{el}^{kin} + E_N^{kin} + E^{KS}$.

L'esquema general d'una dinàmica CPMD és el següent:

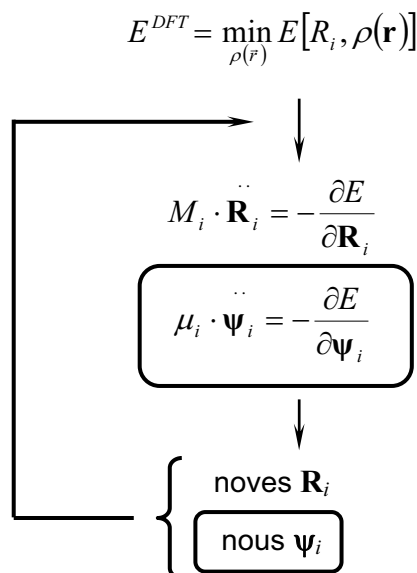


Figura II.2. Esquema d'una dinàmica molecular de Car-Parrinello (CPMD).

A partir de la integració numèrica de les equacions del moviment, s'obté l'evolució temporal de la posició dels nuclis (R_i) i de la dels orbitals electrònics (Ψ_i), que es mouen sobre la superfície d'energia potencial corresponent l'estat electrònic fonamental del sistema. És a dir es mouen alhora els nuclis i els electrons. Això permet que per al següent pas de simulació ja no sigui necessari tornar a minimitzar el funcional de la densitat a fi de recalculer els orbitals monoelectrònics, estalviant així temps computacional.

II.1.1. Temps de Simulació i Massa Electrònica Fictícia

L'estabilitat d'una simulació CPMD depèn fortament del valor de la massa electrònica fictícia (μ) i del pas de simulació (Δt) escollits per al càlcul. Com en tot càlcul de dinàmica molecular, el màxim pas de simulació que es pot emprar està relacionat amb la màxima freqüència vibracional del sistema. En un càlcul CPMD, les freqüències més elevades es corresponen ara amb els graus de llibertat electrònics. Aquestes freqüències han d'estar separades de les freqüències nuclears per tal de garantir l'adiabaticitat entre el moviment electrònic i el nuclear, a fi de què no hi hagi transferència energètica entre ells¹²⁷. Quan això s'aconsegueix, el sistema electrònic no es separa de la superfície de BO, oscil·lant al voltant de l'estat fonamental. Diversos estudis han demostrat que les forces sobre els àtoms són pràcticament les mateixes que s'obtenen en una dinàmica BOMD i, a més, el petit error en les forces atòmiques presenta unes oscil·lacions que es cancel·len durant la simulació¹¹⁶.

Com que les freqüències electròniques són necessàriament més elevades que les nuclears, el pas de simulació (Δt) ha de ser menor que en dinàmica molecular clàssica o BOMD. Típicament, el pas de simulació oscil·la entre els 0.1 i 0.3 fs. La màxima freqüència electrònica, depèn del valor escollit per a la massa electrònica fictícia (μ). Com més elevat és el valor de μ , menor és la màxima freqüència. Per la seva banda, la mínima freqüència electrònica, $\omega_e^{\min}$, és de l'ordre de $\sqrt{E_{gap}/\mu}$, on E_{gap} és la separació HOMO–LUMO del sistema¹¹⁷. Típicament, s'usen valors de μ entre 500 i 1500 unitats atòmiques en combinació amb valors per al pas de simulació (Δt) entre 0.12 i 0.24 fs. De totes maneres, és convenient fer diferents tests de μ i Δt per a cada sistema.

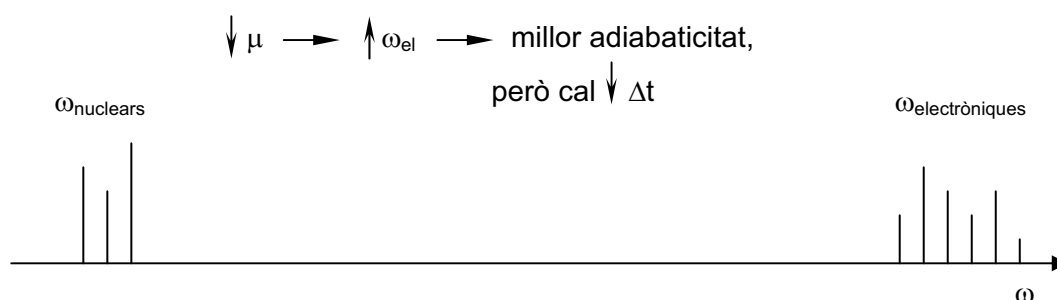


Figura II.3. Separació de les freqüències vibracionals associades al moviment nuclear i al moviment electrònic en una simulació CPMD

Una manera de controlar que el sistema electrònic es manté molt proper a la superfície de BO, és seguir l'evolució de l'energia cinètica associada al moviment electrònic (E_{el}^{kin}). Aquesta ha de tenir un valor petit (3 o més ordres de magnitud inferior a l'energia cinètica nuclear) i el que és més important, s'ha de mantenir constant durant tota la dinàmica.

La Figura II.4 extreta d'un estudi de dinàmica molecular d'aigua líquida il·lustra el comportament del sistema emprant dues masses electròniques diferents¹²⁷. Quan s'empra un valor de $\mu=800$ u.a. (en vermell) s'observa que la temperatura disminueix. Això és conseqüència d'una transferència d'energia cinètica dels nuclis als electrons. Per això, E_{el}^{kin} va incrementant a mesura que evoluciona la simulació i això és un indicatiu que el sistema electrònic s'està separant de la superfície de BO. La conseqüència principal d'aquest fet és que les forces no són les corresponents a l'estat fonamental i els resultats de la simulació no serien vàlids. Per tant, en aquest cas el valor de 800 u.a. és massa gran. En canvi, 400 u.a. representa un valor adient ja que manté l'adiabaticitat (E_{el}^{kin} oscil·la al voltant d'un valor constant) i conserva l'energia total (E_{tot}).

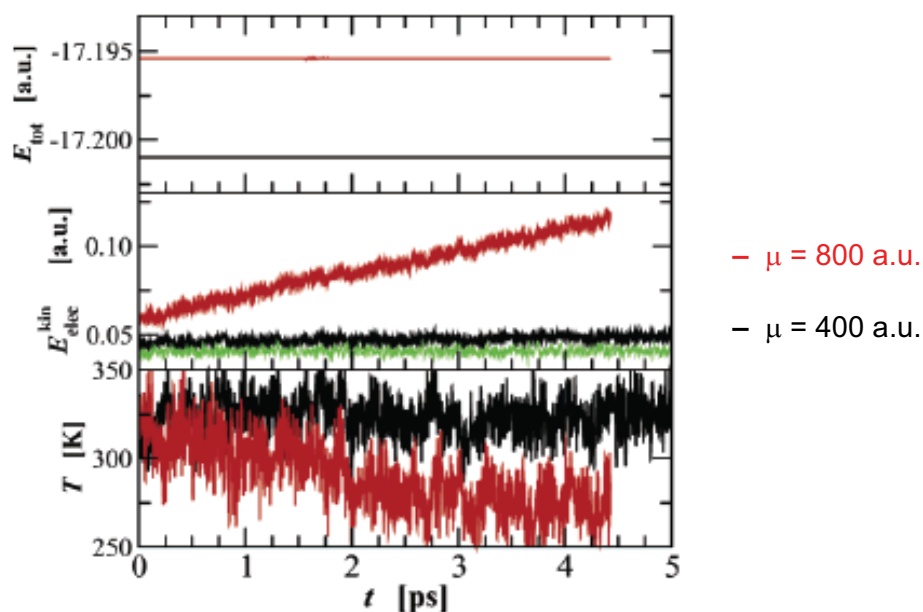


Figura II.4. Evolució de (a) l'Energia Total (E_{tot}), (b) Energia Cinètica Electrònica (E_{el}^{kin}) i (c) Temperatura (T) en una simulació CPMD d'aigua líquida emprant dos valors de μ diferents.¹²⁷

II.2. Mètode QM/MM de Dinàmica Molecular

Si bé els mètodes clàssics de dinàmica molecular permeten analitzar sistemes grans mitjançant la utilització de camps de força, i per la seva banda els mètodes de primers principis permeten analitzar sistemes petits sense necessitat de parametrització prèvia, existeix la possibilitat d'emprar un mètode híbrid que s'aprofita dels beneficis d'ambdós mètodes.

Els mètodes híbrids QM/MM de dinàmica molecular tracten una part del sistema mitjançant dinàmica clàssica, i l'altra (la regió on es produeixen importants reorganitzacions electròniques) mitjançant dinàmica molecular de primers principis¹²⁸⁻¹³².

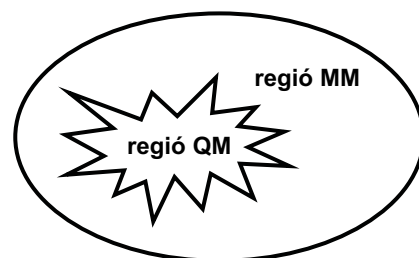


Figura II.5. Definició de les regions en els models híbrids.

La regió QM està sotmesa a efectes de la regió MM i viceversa. Per exemple, el potencial creat pels àtoms MM sobre la regió QM intervé en forma d'un potencial extern que s'inclou en el funcional (seguint el formulisme DFT). Alhora, l'energia de la interacció QM – MM inclou la interacció electrostàtica de les càrregues dels àtoms MM amb la densitat electrònica de la regió QM.

En aquesta tesi s'ha emprat el mètode QM/MM desenvolupat per A. Laio, J. Vandevondele i U. Röthlisberger¹³² que combina la dinàmica CPMD amb la dinàmica molecular clàssica, emprant el camp de forces GROMOS o AMBER.

II.2.1. Interaccions Electroestàtiques QM–MM

En el mètode CPMD QM/MM, les interaccions electroestàtiques entre la regió MM i la regió QM es tracta en tres graus diferents de simplificació. Així, la regió MM es divideix en tres subregions diferents. En cadascuna s'empra un nivell de càlcul diferent¹³².

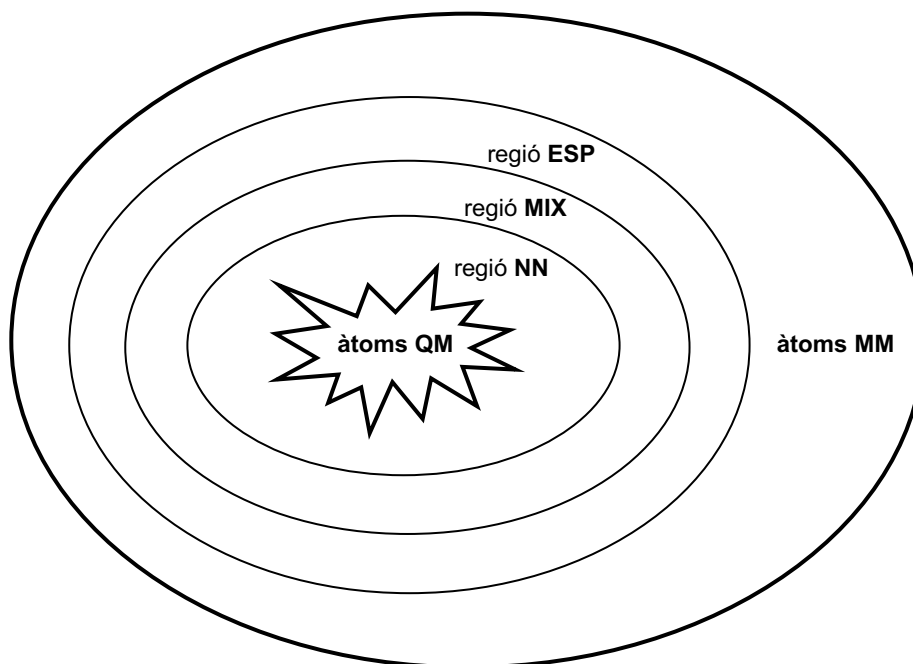


Figura II.6. Regions d'Interaccions Electroestàtiques

Per als àtoms clàssics presents en la primera subregió (regió NN) es defineix un model d'interacció electroestàtica entre la càrrega puntual de cada àtom clàssic i la densitat electrònica del sistema quàntic. Les càrregues puntuals del sistema clàssic són suavitzades en l'espai mitjançant una distribució de càrrega tipus potencial (v_j). D'aquesta manera s'evita l'anomenat efecte 'spill-out' que provocaria una forta migració de densitat electrònica cap al punt localitzat d'una càrrega positiva. En la segona subregió (regió MIX) es calculen de la mateixa manera únicament els àtoms amb una càrrega superior a 0.1.

$$E_{electrostatic}^{NN} = \sum_{j \in NN} q_j \int \rho(r) v_j(r - r_j) \quad \text{prenent} \quad v_j(r) = \frac{r_{cj}^4 - r^4}{r_{cj}^5 - r^5} \quad \text{eq.11}$$

Els àtoms de la tercera subregió (regió ESP), juntament amb els àtoms amb càrrega inferior a 0.1 de la regió MIX interaccionen amb la densitat electrònica a través d'una interacció clàssica *Coulombiana* (càrrega – càrrega). Les càrregues dels àtoms en la regió MM són les càrregues del camp de forces, mentre que les càrregues per als àtoms de la regió QM són obtingudes dinàmicament a partir de la densitat electrònica mitjançant el mètode R-ESP¹³³.

Per a la resta d'àtoms, més enllà de la regió ESP, la interacció amb el sistema quàntic es calcula a través de multipols de quart ordre, la qual cosa disminueix el cost computacional.

II.2.1.1. Càrregues ESP i R-ESP

En la present tesi, s'ha calculat sovint les càrregues ESP de diferents sistemes en fase gas. En el procés del càlcul de càrregues ESP, el primer pas és calcular, mitjançant mètodes mecanoquàntics, el Potencial Electrostàtic Molecular (PEM) generat a diferents punts d'una superfície que engloba el sistema. A continuació, es procedeix a ajustar diferents càrregues puntuals (ubicades en els centres nuclears del sistema) de manera que reproduïxin el potencial electrostàtic en els punts de l'esfera¹³⁴.

En el cas del mètode CPMD QM/MM, l'ajust es fa de manera que les càrregues reproduïxin el potencial electrostàtic generat la regió QM en diferents punts de la regió MM¹³³. Aquests punts, enlloc d'agafar-se en una superfície esfèrica externa es corresponen ara a la ubicació dels àtoms de la regió NN. A més, s'afegeix una restricció harmònica (k) de manera que les càrregues (q^{fit}) no es desviïn d'un valor de referència (q^{ref}). Les càrregues q^{ref} es calculen mitjançant el mètode *Hirshfeld*¹³⁵ a partir de la densitat electrònica del sistema. El procés d'ajust consisteix en minimitzar la funció σ^2 (eq.12). Les càrregues resultants d'aquest ajust s'anomenen càrregues R-ESP.

$$\sigma^2 = \sum_{j \in NN} \left(\sum_{i \in QM} \frac{q_i^{fit}}{r_{ij}} - V_j \right)^2 + k \sum_{i \in QM} (q_i^{fit} - q_i^{ref})^2 \quad \text{eq.12}$$

II.2.2. Regió Frontera QM–MM

No tots els sistemes proteïna – lligand permeten una separació entre la regió analitzada quànticament i la regió clàssica, en la que no es vegin involucrats enllaços covalents. En aquests casos, hi ha determinats enllaços que travessen la interfície entre les dues regions, que caldrà tractar de manera especial.

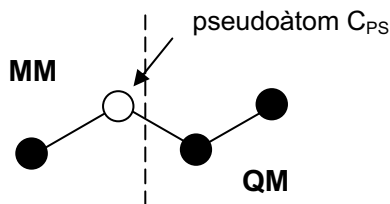


Figura II.7. Interfície entre la regió QM i la MM a través d'un enllaç

Un enllaç insaturat en la regió QM és una forta pertorbació de l'estructura electrònica del sistema. Per aquest motiu, s'ha de saturar l'àtom fronterer a fi de minimitzar aquesta pertorbació. Una de les aproximacions més emprades, es basa en assignar un pseudopotencial monovalent a l'àtom fronterer, substituint així un dels àtoms del sistema real. Aquest pseudopotencial (anomenat C_{ps}) és un pseudopotencial empíric que conté un únic electró de valència (pseudopotencial hidrogenoide) i està parametritzat per tal de reproduir les característiques d'un enllaç simple C—C.

Les zones de tall entre la regió MM i la regió QM han de ser escollides de tal manera, que dins del possible s'allunyin de la zona realment d'interès, per exemple, dels àtoms involucrats en el procés químic que es pretén seguir per dinàmica.

II.3. Mètode d'Acceleració d'Esdeveniments (metadinàmica)

La metadinàmica proveeix d'un marc unificat per al càlcul d'energies lliures i per accelerar la successió d'esdeveniments en simulacions de dinàmica molecular¹³⁹. L'algoritme està basat en una reducció de dimensionalitat del sistema. Això requereix la identificació prèvia d'un conjunt de variables col·lectives (s), les quals són funció de les coordenades del sistema, i són capaces de descriure el procés d'interès. L'algoritme de la metadinàmica s'explica a continuació per a un cas ideal en què es pretén accelerar l'exploració de l'espai conformacional d'un sistema de coordenades atòmiques x en base a una sola variable col·lectiva s . L'algoritme permetrà també extreure el perfil d'energia lliure en funció de s . La variable col·lectiva s és una funció explícita d'un subconjunt de coordenades x' del sistema ($s=f(x')$), com ara un angle, una distància o un índex de coordinació entre altres. L'evolució del sistema es simula mitjançant dinàmica molecular clàssica a una temperatura T determinada (també podria seguir altres esquemes de dinàmica, com per exemple dinàmica *Monte Carlo*⁸⁷, dinàmica *Langevin*⁸⁸, ...). La dinàmica del sistema ve definida pel potencial d'interacció interatòmic ($V(x)$) definit com a funció de les coordenades atòmiques x (camp de forces). A la Figura II.8 es representa el perfil del potencial d'interacció V en funció de la coordenada s . L'evolució del sistema al llarg de la dinàmica pot ser projectada sobre aquest perfil (bola de color gris).

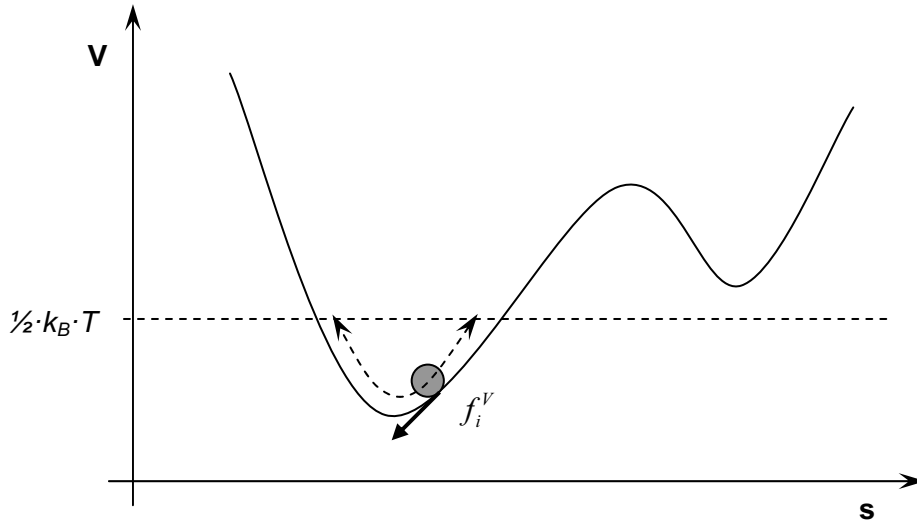


Figura II.8. Evolució del sistema sobre el perfil d'energia potencial.

La simulació comença en un dels mínims d'aquest potencial. A la temperatura T de simulació, el sistema oscil·la al voltant d'aquest mínim (oscil·lacions representades amb la corba discontinua). La força que actua en cada pas de simulació sobre els àtoms del sistema ve definit pel gradient d'aquest potencial:

$$f_i^V = -\frac{\partial V}{\partial r_i} \quad \text{eq.13}$$

Aquestes forces són les que, a través de les equacions del moviment de Newton, propaguen el moviment dels àtoms del sistema. Cada cert interval de temps t' , s'afegeix un petit terme gaussià al potencial V centrat en el valor de s a temps t' . Així, en qualsevol moment de la simulació, el potencial total és:

$$V_T = V + V_G \quad \text{on} \quad V_G = \sum_{t'} w \cdot e^{-\frac{(s(t)-s(t'))^2}{2(\delta s)^2}} \quad \text{eq.14}$$

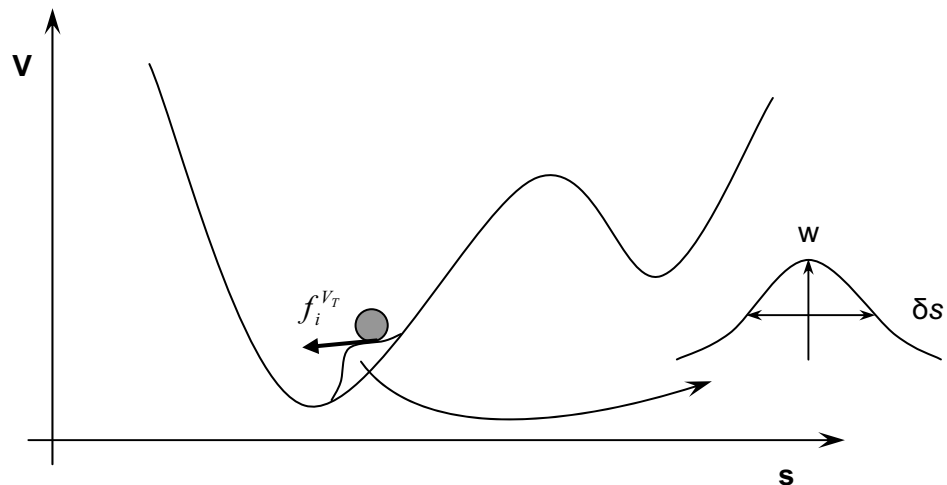


Figura II.9. Evolució del sistema sobre el perfil d'energia potencial modificat.

FONAMENTS METODOLÒGICS

Ara, a cada pas de dinàmica, la força que actua sobre els àtoms del sistema és la suma de dues components: una provinent del potencial d'interacció interatòmic V , i l'altre del potencial històric V_G , segons els corresponents gradients:

$$f_i^{V_T} = f_i^V + f_i^{V_G} = -\frac{\partial V}{\partial r_i} - \frac{\partial V_G}{\partial r_i} = -\frac{\partial V}{\partial r_i} - w \cdot \frac{\partial \left(\sum_{t'} e^{-\frac{(s(t)-s(t'))^2}{2(\delta s)^2}} \right)}{\partial r_i} \quad \text{eq.15}$$

Cal tenir present que la trajectòria de metadinàmica no és una trajectòria d'equilibri, ja que cada cert temps t' es modifiquen les forces que actuen sobre el sistema. Cada vegada que s'afegeix un terme gaussià al potencial històric V_G , cal deixar equilibrar el sistema ja que s'ha modificat el potencial total. Això requereix que en aquest mètode s'hagi d'acoblar un termòstat adient per assegurar que la distribució de velocitats es mantingui constant a la temperatura de simulació.

Aquesta modificació del potencial al llarg de la dinàmica penalitza les regions de l'espai s ja visitades, i accelera l'exploració de l'espai conformacional, guiat per la variable col·lectiva s . Com que el potencial històric compensa iterativament el potencial subjacent, **un sistema que evolucioni amb metadinàmica tindrà tendència a escapar d'un mínim d'energia a través de l'estat de transició més proper**. Aquesta propietat és probablement la qualitat més important de la metadinàmica i la que la fa una eina flexible i eficient per explorar nous camins de reacció (accelerar esdeveniments). Si les variables col·lectives s'escullen de manera adient, el sistema evolucionarà a través del camí que hauria seguit en un temps major de simulació si s'hagués simulat amb dinàmica molecular ordinària.

La metadinàmica permet també obtenir informació quantitativa sobre propietats d'equilibri del sistema. Aquesta capacitat es basa en l'observació que el **potencial històric V_G** , si s'ha construït de manera adient, **es correspon amb una estimació de l'energia lliure** en funció de s en les regions explorades pel sistema fins a temps t .

$$\lim_{t \rightarrow \infty} V_G(s, t) = \sum_{t'} w \cdot e^{-\frac{(s(t) - s(t'))^2}{2(\delta s)^2}} \approx G(s) \quad \text{eq.16}$$

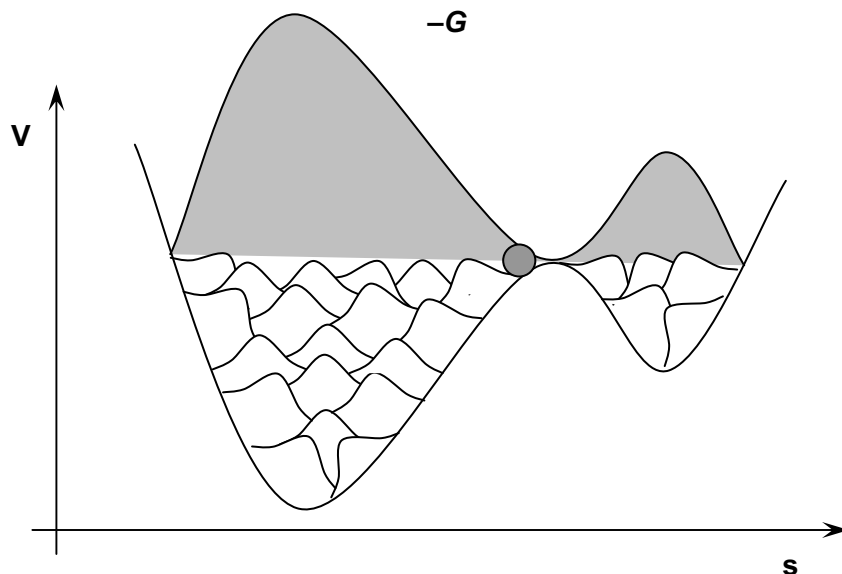


Figura II.10. Metadinàmica completada, i reconstrucció del perfil d'energia lliure (G).

Aquesta propietat no es deriva de cap identitat termodinàmica. Inicialment es va postular de manera heurística¹³⁹ i posteriorment s'ha verificat empíricament en diferents sistemes de creixent complexitat. De fet recentment s'ha demostrat que aquesta propietat deriva de principis generals i pot ser demostrada rigorosament per a un sistema que evoluciona en una dinàmica de *Langevin*¹⁴¹.

Els paràmetres w i δs que defineixen el potencial històric V_G influeixen en la qualitat del perfil d'energia lliure reconstruït¹⁴⁰. Per a cada problema específic, aquests paràmetres s'han d'escollir amb cura a fi d'assegurar un correcte compromís entre precisió i eficiència computacional. Valors grans de w i δs permetran una ràpida exploració de l'espai de les variables col·lectives, però al preu d'una baixa precisió.

II.3.1. Variables Col·lectives

Les variables col·lectives s , en les que actua el potencial històric V_G , s'han d'escollir de manera que donin suficient llibertat al sistema per fer el que faria en una dinàmica sense guiar. Això pot semblar una tasca difícil ja que no es disposa d'una recepta *a priori* per trobar un conjunt adient de variables col·lectives o per comprovar si la descripció obtinguda amb el conjunt seleccionat és prou acurada. De totes maneres, l'experiència en sistemes específics ha demostrat que es poden definir variables col·lectives generals de manera que s'obtinguin camins de reacció raonables i nous estats estables que no eren predits inicialment. Per exemple, el mètode pot ser emprat per simular transicions de fase en cristalls emprant les

arestes de la caixa de simulació com a variables col·lectives¹⁴⁹. En fer això, el sistema escapa de l'estructura local en que s'ha iniciat la simulació cap a una nova estructura, en la que no s'ha fet cap assumpció, i que ocasionalment podria correspondre's al mínim global d'energia lliure en les condicions termodinàmiques d'interès. En el camp de la química, el mètode ha estat emprat extensivament en conjunció amb el mètode CPMD¹⁴², introduint índex de coordinació generalitzats que poden discriminar entre diferents camins de reacció^{142,150}, o directament permeten la reconstrucció de l'energia lliure en diferents dimensions^{151,152}. En el camp de la biofísica i de la determinació de complexes enzim-ligand^{147,148}, així com de la ciència de materials, s'han definit variables col·lectives específiques de cada sistema les quals permeten observar un esdeveniment específic (per exemple, la translocació d'un antibiòtic a través d'un porus de membrana¹⁵³, o la nucleació de la fase líquida condensada en aigua¹⁵⁴). Una variable col·lectiva (CV) per a ser emprada en metadinàmica ha de complir els següents requisits:

- La coordenada de reacció del procés ha de ser una combinació de les CVs escollides.
- Les CVs han de ser capaces de diferenciar entre els estats inicials i final del procés.
- Tots els modes de reacció amb temps característics similars als de les CVs han de ser inclosos.
- La funció matemàtica que defineix les CVs ha de ser contínua i derivable. Poder avaluar $\delta s / \delta r_i$ és necessari per a l'algoritme (eq.15).

II.3.2. Metadinàmica Acoblada al *Lagrangià* de *Car-Parrinello*

Si es fa servir la metadinàmica per simular reaccions químiques en conjunció amb el mètode CPMD, el potencial històric (V_G) ha de forçar les variables col·lectives a creuar barreres energètiques elevades (de l'ordre de la desena de kcal·mol⁻¹) en temps molt curts (al voltant de pocs ps). Això implica que s'injecta molta energia en els graus de llibertat associats amb les variables col·lectives. Això pot provocar importants deshomogeneïtats en la distribució de temperatura en el sistema, i introduir inestabilitats en la dinàmica. Una solució a aquests problemes, en l'esperit dels mètodes *Parrinello-Rahman* i *Car-Parrinello*, és introduir un *Lagrangià* estès¹⁴² en el que les variables col·lectives s són acoblades a unes variables col·lectives auxiliars s' . Aquestes estan acoblades a les variables col·lectives reals a través d'un potencial harmònic (constant de força k) de forma:

$$V_H = \frac{1}{2} k (s' - s(x))^2 \quad \text{eq.17}$$

Les noves variables col·lectives auxiliars s' són tractades com graus de llibertat addicionals, assignant-los-hi una massa fictícia (M) i per tant tenen associada una energia cinètica fictícia:

$$K_{s'} = \frac{1}{2} M v_{s'}^2 \quad \text{eq.18}$$

D'aquesta manera, la dinàmica d'aquests graus de llibertat extrems pot ser controlada explícitament a través de termòstats adequats. La trajectòria de s' pot fer-se suficientment suau de manera que l'estabilitat de l'algorisme es manté sota control.

La forma modificada del potencial total que actua sobre el sistema real, i que substitueix l'eq.14 és el següent:

$$V_T = V + V_H = V + \frac{1}{2} k (s' - s(x))^2 \quad \text{eq.19}$$

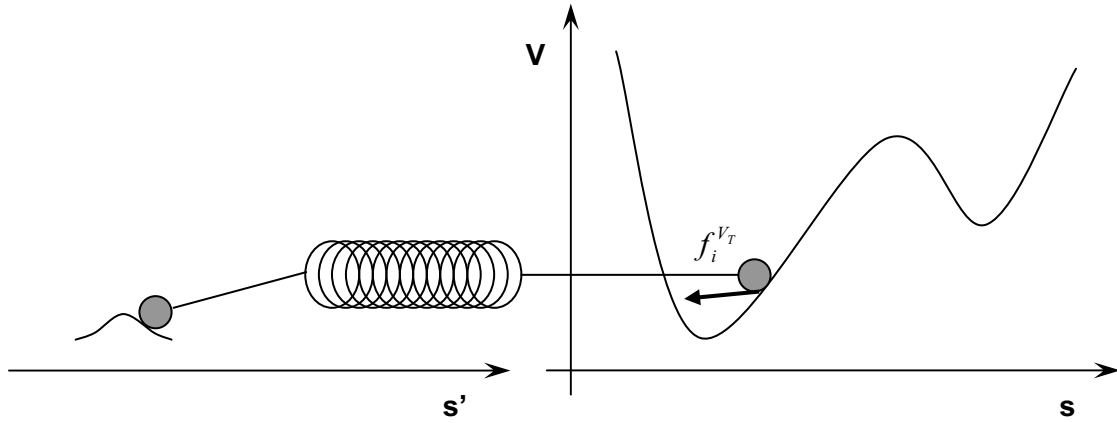


Figura II.11. Metadinàmica acoblada al Lagrangiana de Car-Parrinello.

Així mateix, la força total que experimenta el sistema real, i que substitueix l'eq.15 és:

$$f_i^{V_T} = f_i^V + f_i^{V_H} = -\frac{\partial V}{\partial r_i} - \frac{\partial V_H}{\partial r_i} = -\frac{\partial V}{\partial r_i} - \frac{1}{2} k \frac{\partial (s' - s(x))^2}{\partial r_i} \quad \text{eq.20}$$

Ara el potencial històric (V_G) enlloc d'actuar sobre el sistema real, actua en l'espai de la variable col·lectiva auxiliar s' , i va construint-se a partir de termes gaussians centrats ara en s' . La dinàmica en l'espai de la variable col·lectiva auxiliar s' està guiada, doncs, per les forces que provenen del potencial harmònic V_H i del potencial històric V_G :

$$f_{s'}^{V_T} = f_{s'}^{V_H} + f_{s'}^{V_G} = -\frac{\partial V_H}{\partial s'_i} - \frac{\partial V_G}{\partial s'_i} = -\frac{1}{2} k \frac{\partial (s' - s(x))^2}{\partial s'_i} - w \cdot \frac{\partial \left(\sum_{t'} e^{-\frac{(s'(t) - s'(t'))^2}{2(\delta s')^2}} \right)}{\partial s'_i} \quad \text{eq.21}$$

La massa fictícia M i la constant d'acoblament k del potencial harmònic, determinen la velocitat a què la partícula en l'espai s' evoluciona en el temps respecte els graus de llibertat iònics. Si la massa M és gran, la variable col·lectiva auxiliar s' és lenta, de manera que les freqüències corresponents a les variables col·lectives estan separades respecte les corresponents als graus de llibertat iònics i electrònics¹⁴³.

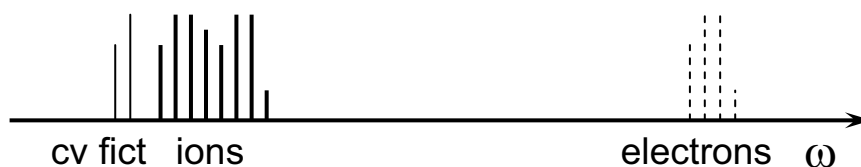


Figura 11.12. Espectre de freqüències idealitzat d'una metadinàmica acoblada al Lagrangiana CPMD.

Idealment, amb una combinació de M i k adient, les variables col·lectives en el sistema real s'oscil·len al voltant de la variable col·lectiva auxiliar s' . El treball de *Einsing et al.*¹⁴³ ofereix una recepta per a escollir adequadament aquests valors. En condicions de separació adiabàtica, les forces que actuen sobre el sistema fictici, són una estimació de la derivada de l'energia lliure respecte les variables col·lectives. De fet, en el límit de M elevades, la variable col·lectiva s' està fixada, de manera que les forces són un promig dinàmic sobre els graus de llibertat iònics i electrònics. Així doncs, s' evoluciona sota l'efecte d'una força promig que és la derivada de l'energia lliure respecte s' tal com en els mètodes de *umbrella sampling* i *constrained dynamics*^{90,91}.

De manera anàloga al mètode original de metadinàmica, el potencial V_G penalitza les zones de l'espai s' ja visitades, accelerant així, a través del potencial harmònic l'exploració de l'espai conformacional s . A mesura que evoluciona la simulació, el potencial històric V_G en l'espai de s' compensa el potencial V en l'espai s . Si V_G es construeix adientment i es compleixen les condicions d'adiabaticitat, aquest potencial compensa el perfil d'energia lliure subjacent en l'espai s (Figura 11.13).

$$\lim_{t \rightarrow \infty} V_G(s', t) = \sum_{t'} w \cdot e^{-\frac{(s'(t) - s'(t'))^2}{2(\delta s)^2}} \approx G(s) \quad \text{eq.22}$$

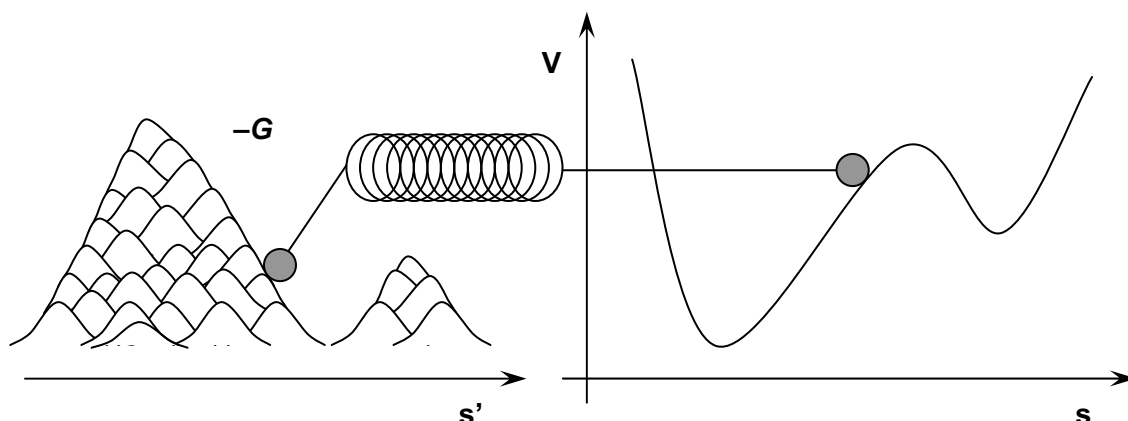


Figura II.13. Metadinàmica completada, i reconstrucció del perfil d'energia lliure (G).

II.3.3. Metadinàmica en Múltiples Coordenades

Un dels principals avantatges de la metadinàmica és la reducció de dimensionalitat del problema en estudi. Tots els graus de llibertat del sistema queden reduïts a la selecció de 2 ó 3 variables col·lectives que han d'incloure els modes característics del procés. Sovint, però, és necessari emprar un nombre major de variables col·lectives. En aquests casos, l'eficiència del mètode comença a disminuir a causa de l'increment en el cost computacional degut a la necessitat d'explorar un espai conformacional cada vegada major.

El temps requerit per escapar d'un mínim local en la superfície d'energia lliure (FES) ve determinat pel nombre de termes gaussians que s'han d'afegir al potencial històric V_G . Aquest número és proporcional a $(1/\delta s)^d$ on d és el número de variables col·lectives emprades en la simulació. Així doncs, l'eficiència del mètode escala exponencialment amb el nombre de dimensions involucrat. Si d és gran, la única manera de reconstruir eficientment la FES és emprar gaussianes de la mida del pou d'energia lliure. Lògicament, en aquests casos, la precisió es perd per complet.

Si bé la metadinàmica estàndard s'ha emprat eficientment per explorar diferents coordenades de reacció com ara en estudis de *docking* o translocació d'antibiòtics^{147,148,153}, aquesta no ha pogut ser aplicada per estudiar el plegament de proteïnes ja que aquest procés involucra un nombre elevat de variables col·lectives. Alternativament *Bussi et al.* han introduït una aproximació que combina metadinàmica amb un esquema d'intercanvi de rèpliques¹⁴⁴. En aquesta aproximació, es porten a terme intercanvis entre diferents rèpliques que evolucionen amb metadinàmica a diferents temperatures. D'aquesta manera no és necessari saber *a priori* les coordenades de reacció importants del sistema, ja que les barreres en les variables col·lectives desconegudes es creuen gràcies a la difusió per

temperatura (Figura II.14). Emprant aquest mètode ha estat possible obtenir l'energia lliure de plegament d'una β -hairpin en solvent explícit en un temps computacional menor que aplicant metadinàmica o intercanvi de rèpliques tradicional per separat¹⁴⁴. De totes maneres, segueix estan present la limitació en la mida de l'espai conformacional.

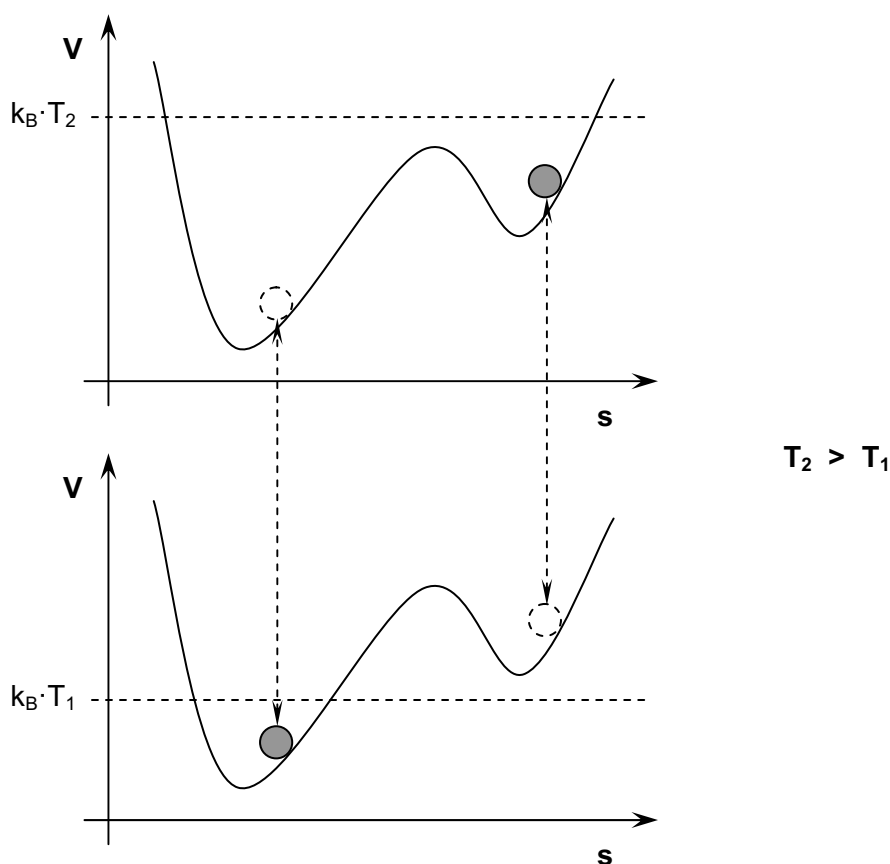


Figura II.14. Representació esquemàtica de dues metadinàmiques a diferents temperatures amb intercanvi de rèpliques. A temperatures més elevades es poden creuar barreres energètiques més altes.

Com a alternativa, recentment *Piana i Laio* han desenvolupat una ampliació del mètode de metadinàmica que permet l'eficient exploració de l'espai conformacional i reconstruir la FES en base a un nombre, en principi il·limitat, de variables col·lectives¹⁴⁵. Aquest mètode està basat en l'esquema d'intercanvi de rèpliques a diferents temperatures descrit al paràgraf anterior. En el mètode anterior el potencial històric V_G actua sobre un nombre limitat de variables col·lectives i la difusió en els graus de llibertat ortogonals s'aconsegueix a través de l'intercanvi entre rèpliques que corren a diferents temperatures. En aquest nou mètode, anomenat *bias exchange metadynamics* (BE-META), els intercanvis es porten a terme entre rèpliques que corren a la mateixa temperatura, però que estan guiades per diferents potencials històrics que actuen sobre diferents variables col·lectives: per exemple, en una simulació del plegament d'una proteïna emprant dues rèpliques, es podria fer servir la metadinàmica per explorar l'espai del nombre total de ponts d'hidrogen entre cadenes del *backbone* en una de les rèpliques (*ic*), i en l'altra activar l'exploració de l'espai del radi de gir (ω). Si s'accepta l'intercanvi, la trajectòria que prèviament estava guiada pel primer potencial

històric (V_1) continua la seva evolució guiada pel segon potencial històric (V_2) i viceversa (veure Figura II.15). D'aquesta manera, un gran nombre de variables col·lectives diferents poden ser emprades en diferents rèpliques. En principi, la dimensionalitat de l'espai explorat per metadinàmica pot ser tan gran que les barreres residuals ortogonals a les coordenades de reacció poden ser creuades en els temps de simulació disponibles. El resultat de la simulació no és una energia lliure en múltiples dimensions, sinó diferents projeccions de menor dimensió de la superfície d'energia lliure respecte totes les variables col·lectives.

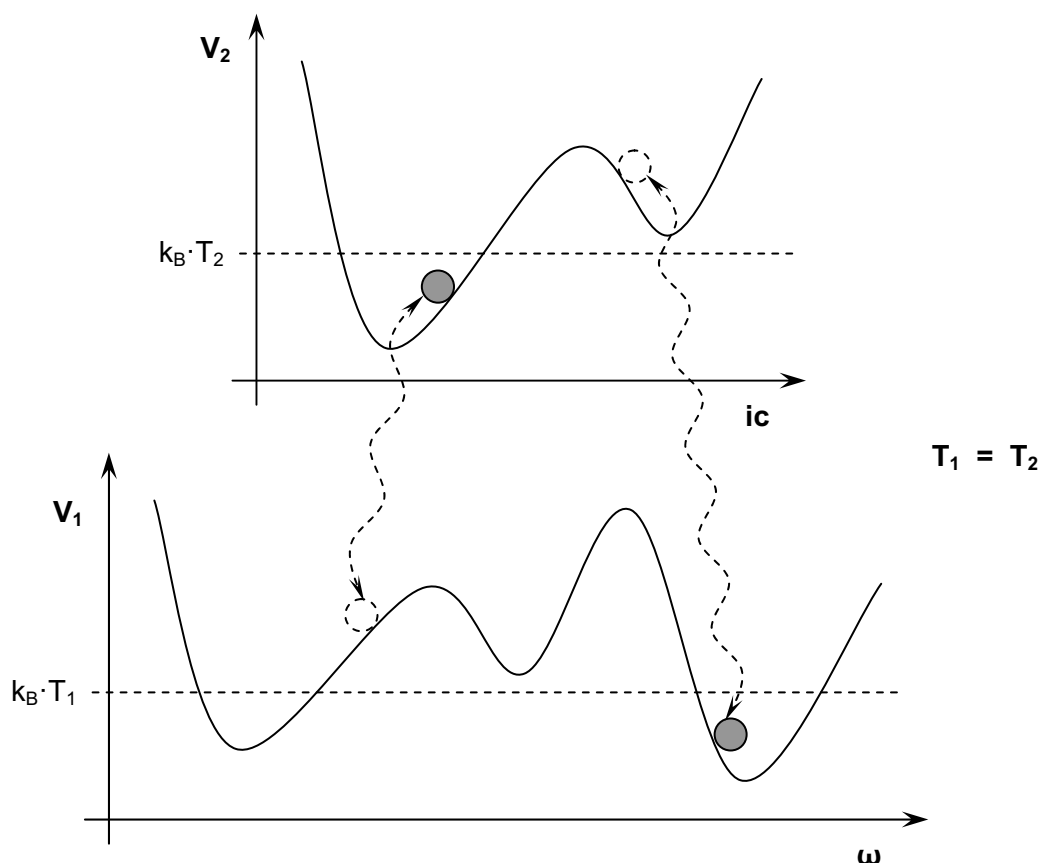


Figura II.15. Representació esquemàtica d'una metadinàmica amb intercanvi de rèpliques.

Els detalls del mètode BE-META es troben a la referència 145. Per un càlcul d'aquestes característiques cal definir un nombre NR de rèpliques del sistema simulades a la temperatura T . Cada una d'aquestes rèpliques està guiada per una variable col·lectiva diferent (o un conjunt petit de variables col·lectives) $s^\alpha(x)$, $\alpha=1,2, \dots$ NR. En cada rèplica s'acumula un potencial històric diferent $V_G(s^\alpha, t)$, el qual, després d'un temps de simulació prou llarg, amb metadinàmica estàndard, correspondria amb una estimació de l'energia lliure projectada sobre s^α . Si totes les variables col·lectives s^α fossin rellevants pel procés, la convergència en cada un dels perfils d'energia lliure seria extremadament lenta en la metadinàmica estàndard.

L'avantatge del mètode BE-META és que es permet un intercanvi entre les configuracions de diferents rèpliques, tal com en el mètode de *replica exchange*¹⁴⁶. S'escullen dues rèpliques (per exemple A i B) entre les NR rèpliques del sistema. L'intercanvi consisteix en permutar les coordenades x^A i x^B de les dues rèpliques. Com que les dues rèpliques evolucionen sota l'acció de dos potencials històrics diferents (V_G^A , V_G^B), l'intercanvi s'accepta amb una probabilitat p_{AB} definida com¹⁴⁶:

$$p_{AB} = \min \left(1, e^{\frac{V_G^A(x^A, t) + V_G^B(x^B, t) - V_G^A(x^B, t) - V_G^B(x^A, t)}{k_B T}} \right) \quad \text{eq.23}$$

Si el moviment s'accepta, les variables col·lectives de la rèplica A fan un salt de $sA(xA)$ a $sA(xB)$ i les de la rèplica B el fan de $sB(xB)$ a $sB(xA)$. Els salts en les variables col·lectives augmenten la capacitat de cada rèplica per difondre en l'espai de les variables col·lectives. Utilitzant aquest algoritme, cada configuració evoluciona sota l'acció d'un potencial històric que canvia cada vegada que s'accepta un intercanvi. Després d'uns quants intercanvis acceptats, cada configuració haurà explorat tot l'espai conformacional dividit en totes les variables col·lectives. Això incrementa considerablement la capacitat del sistema per explorar tot l'espai conformacional. Com en la metadinàmica tradicional, el potencial històric que actuava sobre cada rèplica (V_G) convergeix al valor negatiu de l'energia lliure.