

Capítol 3. Mètodes

3.1 Introducció

L'aplicació de la Química Quàntica ha permès la modelització molecular dels sistemes biològics, augmentant els coneixements que tenim d'aquests sistemes moleculars així com la capacitat de predicció de les seves propietats. Però malgrat aquests avenços, l'interès creixent per estudiar sistemes macromoleculars com els àcids nucleics, els enzims i els receptors cel·lulars ha fet que a la pràctica el volum de càlcul requerit faci impossible aplicar aquests coneixements a sistemes amb més d'unes desenes d'àtoms. Per aquest motiu s'han hagut de desenvolupar "aproximacions" que permetin tractar computacionalment aquests sistemes macromoleculars. En general, quan més econòmica respecte al volum de càlcul sigui l'aproximació, menor serà la seva eficàcia i amb menys exactitud predirà les propietats del sistema estudiat i major serà la incertesa dels seus resultats.

D'altra banda, quan major sigui la varietat de mètodes i aproximacions de què es disposi, més grans seran les possibilitats de descriure correcta i acuradament el sistema i per tant de predir les seves propietats estructurals, dinàmiques i de reactivitat. De forma genèrica cada mètode es fonamenta en uns principis fisicoquímics, un desenvolupament matemàtic més o menys complex i, finalment, un programa informàtic que permet efectuar els càlculs. Això implica que per tal d'aplicar un mètode ha d'existir un programa informàtic que reculli les característiques del model físic i les formuli matemàticament. El coneixement d'aquestes bases físiques i matemàtiques és essencial per poder aplicar la metodologia a sistemes d'interès fisicoquímic i extreure informació vàlida de les dades obtingudes.

A continuació es descriuen, si bé necessàriament amb una extensió limitada, els principals fonaments de les metodologies que s'han utilitzat en aquesta tesi doctoral.

3.2 Química Quàntica

S'inclou amb aquesta denominació a un conjunt de mètodes basats tant en l'equació d'Schrödinger com en la teoria funcional de la densitat [1], la química quàntica és una eina fonamental per aprofundir en les característiques fisicoquímiques de les molècules. Segons les possibilitats del model electrònic utilitzat per resoldre l'equació d'Schrödinger es podrà calcular amb una gran exactitud per una banda l'energia molecular i per l'altra les propietats moleculars del sistema. Des de fa una dècada el grup QCPE (*Quantum Chemical Program Exchange*) de la Universitat d'Indiana ha contribuït enormement a la distribució d'excel.lents programes de química quàntica per a tots els grups d'investigació mundial, com els programes MOPAC [2] i AMPAC [3], que inclouen mètodes quàntics semiempírics i els programes GAUSSIAN [4], GAMES [5] i HONDO [6] per a càlculs *ab initio* entre altres.

En aquest capítol s'introdueixen els mètodes teòrics [6-8] basats en la resolució de l'equació de Schrödinger independent del temps (equació 3.1), en la que \hat{H} representa l'operador hamiltonià per al qual cal obtenir les funcions pròpies, Ψ és la funció d'ona que descriu completament al sistema que té una energia E .

$$\hat{H}\Psi = E\Psi \quad (3.1)$$

Excepte per als sistemes més simples, resulta impossible resoldre exactament l'equació 1, de manera que si es desitja abordar l'estudi de sistemes d'interès químic cal emprar alguna certa aproximació. Les diferents aproximacions proposades [6-8] han donat lloc a diversos mètodes de la mecànica quàntica com els *ab initio* (punt 2.1) que permeten estudiar qualsevol sistema sense disposar d'informació empírica addicional. No obstant, en molts casos d'interès, aquest plantejament té uns requeriments computacionals difícilment assumibles. Per aquesta raó es van desenvolupar els *mètodes semiempírics* (punt 3.2.2), la filosofia dels quals consisteix en prescindir de part de la rigorositat dels plantejaments *ab initio* a canvi d'introduir paràmetres que permetin reproduir els resultats experimentals.

3.2.1 Mètodes *ab initio*

Els mètodes *ab initio* parteixen de l'aproximació Born-Oppenheimer, consistent en tractar per separat el moviment dels nuclis i dels electrons. Això és possible ja que la massa dels electrons és molt més petita que la dels nuclis i per tant la seva velocitat és molt superior. Tal com s'explica a continuació, aquesta aproximació simplifica el hamiltonià i per tant facilita la resolució de l'equació de Schrödinger. Per a un sistema de N nuclis i M electrons, el hamiltonià (\hat{H} ; en unitats atòmiques) de l'equació 3.1 pren la forma:

$$\hat{H} = -\sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{A=1}^M \frac{1}{2m_A} \nabla_A^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>1}^N \frac{1}{r_{ij}} + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}} \quad (3.2)$$

on m_A és la relació de la massa del nucli A respecte a la massa de l'electró, Z_A és el nombre atòmic del nucli A , ∇_i^2 i ∇_A^2 són operadors laplacians que impliquen la diferenciació respecte a les coordenades de l'electró i i de l'àtom A , respectivament.

En l'equació 3.2, cada terme és un operador que defineix els diferents components de l'energia del sistema: 1) energia cinètica dels electrons, 2) energia cinètica dels nuclis, 3) atracció electrostàtica entre nuclis i electrons, 4) repulsió electrostàtica entre electrons i 5) repulsió electrostàtica entre nuclis. Gràcies a l'aproximació de Born-Oppenheimer podem descartar el segon i el cinquè terme, per definir el hamiltonià electrònic (\hat{H}_{el}):

$$\hat{H}_{el} = -\sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>1}^N \frac{1}{r_{ij}} \quad (3.3)$$

La solució de l'equació Schrödinger corresponent a \hat{H}_{el} és la funció d'ona electrònica (\hat{H}_{el}) i aquesta és la que es pretén resoldre amb els diferents mètodes de la química quàntica, depenent de la qualitat i quantitat del conjunt d'orbitals atòmics que es tractin, és el que s'anomena base i pot tenir diferents dimensions per representar cadascun dels àtoms de la molècula. La base mínima que més s'utilitza és la STO-3G, en la que cada àtom es representa amb els seus orbitals ocupats total o parcialment (1s, 2s, 2p_x, 2p_y, 2p_z en el cas de l'àtom de carboni) i alhora cadascun d'aquests orbitals es representa mitjançant 3 funcions de Gauss o gaussianes. Existeixen altres bases atòmiques més complexes, des de la 3-21G a les 6-31G**.

Cal tenir en compte que quan més sofisticada sigui la base, més temps de càlcul requerirà i aquest pot ser la quarta potència del nombre de funcions base que s'utilitzin en l'aproximació.

3.2.1.1 Mètode d'Hartree-Fock

Segons el principi variacional, l'energia obtinguda a partir d'una funció d'ona aproximada, serà sempre superior a l'energia real. En conseqüència, com millor sigui la funció d'ona, menor serà l'energia associada, ja que a la funció d'ona òptima li correspon un valor mínim d'energia. Imposant la condició que la primera derivada de l'energia (δE) sigui zero, juntament amb el requeriment que els orbitals moleculars han de ser ortogonals, el mètode Hartree-Fock (HF) aconsegueix resoldre l'equació de Schrödinger d'una manera iterativa. En aquest procés, el primer pas és generar una funció de prova, a continuació es construeix l'operador de Fock, es resol l'equació de valors i funcions pròpies i es repeteix el procediment fins arribar a la convergència. Per aquest motiu el mètode Hartree-Fock s'anomena també del camp autoconsistent (*Self Consistent Field*; SCF).

Les funcions que utilitza el mètode Hartree-Fock s'expressen com una combinació lineal antisimètrica de productes d'*spin-orbitals* (equació 3.4) que adopta la forma d'un determinant de Slater. Els *spin-orbitals* són funcions d'ona mono-electròniques que depenen de les coordenades espacials i de l'*spin*.

$$|\Psi_0\rangle = |\chi_1\chi_2\cdots\chi_i\cdots\chi_N\rangle \quad (3.4)$$

Aquests *spin-orbitals*, a la seva vegada, es representen com a combinació lineal d'orbitals mono-electrònics (també anomenats funcions de base) que sovint es corresponen amb orbitals atòmics. Aquesta estratègia es coneix amb el nom de CLOA (Combinació Lineal d'Orbitals Atòmics), esquematitzada en l'expressió 3.5, on K és el nombre de funcions de base f que descriuen l'*spin-orbital* χ_i , i c_{vi} són els coeficients que multipliquen a cada orbital atòmic.

$$\chi_i = \sum_{v=1}^K c_{vi}\phi_v \quad (3.5)$$

Les funcions de base poden ser del tipus Slater, però generalment s'utilitzen funcions gaussianes degut a que computacionalment resulten més eficients. L'optimització de la funció de prova consisteix, doncs, en modificar els coeficients c_{vi} de les combinacions lineals, és a dir, trobar el conjunt de coeficients que satisfaci:

$$\frac{\partial E}{\partial c_{vi}} = 0 \quad (3.6)$$

Els mètodes HF proporcionen resultats molt dependents del nombre de funcions de base utilitzades. Anomenem base mínima al nombre més petit de funcions de base que poden acomodar tots els electrons del sistema. Les anomenades *split-valence* són segurament les bases més emprades en l'estudi de sistemes d'interès bioquímic i s'anomenen així perquè descriuen els electrons interns amb una sola funció de base mentre que n'utilitzen dues o més per als electrons de valència. També s'acostumen a afegir funcions addicionals (polarització, difuses) per millorar la descripció de la densitat electrònica en les regions més allunyades dels nuclis. A mesura que anem augmentant el nombre de funcions de base, la qualitat dels resultats també augmenta i l'energia del sistema disminueix, fins assolir el límit Hartree-Fock a partir del qual, encara que s'augmenti el nombre de funcions de base, ja no es pot seguir reduint el valor de l'energia.

La majoria de sistemes d'interès bioquímic presenten tots els electrons aparellats. Aquest tipus de sistemes, sovint anomenats de capes tancades, són tractats mitjançant l'aproximació RHF (*Restricted* Hartree-Fock). En aquesta aproximació els *spin-orbitals* χ_i i χ_{i+1} es descriuen amb la mateixa funció espacial però diferent funció de *spin* (equació 3.7).

$$\left. \begin{aligned} \chi_i(x) &= \varphi_j(r)\alpha(\omega) = \varphi_j \\ \chi_{i+1}(x) &= \varphi_j(r)\beta(\omega) = \bar{\varphi}_j \end{aligned} \right\} \quad (3.7)$$

D'aquesta manera, la funció d'ona expressada com determinant de Slater s'ajustarà a l'expressió:

$$|\Psi_0\rangle = |\varphi_1\bar{\varphi}_1 \dots \varphi_i\bar{\varphi}_i \dots \varphi_N\bar{\varphi}_N\rangle \quad (3.8)$$

Per aquest tipus de sistema, Roothaan i (independentment) Hall van proposar una manera de reformular les equacions de Hartree-Fock en forma de matrius, podent així resoldre-les mitjançant tècniques estàndards d'àlgebra que poden ser implementades amb relativa facilitat en programes informàtics.

Una aproximació intrínseca del mètode HF és l'assumpció que cada electró interacciona amb la resta d'electrons de manera promitjada, és a dir, que els electrons es mouen en un camp generat per la distribució electrònica promig generada per la resta dels electrons. Donat que aquest tractament no considera la correlació electrònica (no existeix interacció explícita entre electrons), el mètode del camp autoconsistent o HF pot proporcionar bons resultats per a algunes propietats però, si es pretén estudiar propietats en les que la correlació electrònica pugui ser rellevant, caldrà emprar mètodes més sofisticats.

3.2.2 Mètodes Semiempírics

Els mètodes semiempírics van ser desenvolupats amb la intenció de poder estudiar sistemes grans en un temps en que les limitacions d'ordre computacionals eren molt importants. Avui en dia aquestes limitacions han minvat notablement, però segueixen sent mètodes útils per estudiar sistemes de grans dimensions. Tots els mètodes semiempírics introdueixen simplificacions als mètodes *ab initio* que facin menys costós el càlcul. Tanmateix, per compensar les simplificacions, s'introdueixen paràmetres ajustats a partir de dades experimentals o càlculs *ab initio*.

Donat que la part més lenta del procés SCF és la manipulació de les integrals, una primera aproximació consisteix en considerar explícitament només els electrons de valència, és el que s'anomena aproximació del *core* que considera exclusivament als electrons de valència, els quals es mouen dins d'un camp fix generat per el core de nuclis i electrons interns.

Una segona aproximació va ser l'anomenada *aproximació de base mínima* que utilitzen els mètodes AM1 i MNDO, aquests mètodes fan servir una base mínima d'orbitals d'Slater. La darrera i més recent de les aproximacions que utilitzen els mètodes semiempírics consisteix en l'aplicació del solapament diferencial nul (*Zero differential Overlap*, ZDO).

Aquesta aproximació també ha estat adaptada en els hamiltonians MNDO i AM1. D'acord amb aquesta aproximació la superposició entre parells d'orbitals diferents és zero per a tots els elements de volum $d\nu$:

$$\phi_\mu \phi_\nu d\nu = 0 \quad (\mu \neq \nu) \quad (3.9)$$

Aquesta aproximació comporta l'ortogonalització de la base d'orbitals d'Slater, i com a conseqüència un gran estalvi en el nombre d'integrals que cal calcular. Els mètodes semiempírics poden ser classificats segons el grau d'aplicació de l'aproximació ZDO:

- CNDO (*Complete Neglect of Differential Overlap*)
- INDO (*Intermediate Neglect of Differential Overlap*)
- MINDO (*Modified INDO*)
- NDDO (*Neglect of Diatomic Differential Overlap*)
- MNDO (*Modified Neglect of Diatomic Overlap*)

Així, mentre el primer mètode negligeix totes les integrals bielectròniques, els dos darrers només negligeixen el solapament diferencial entre orbitals atòmics situats sobre àtoms diferents. De l'aproximació MNDO [10] se'n deriven els mètodes semiempírics AM1 [11] i PM3 [12] que són els més coneguts.

3.3 Metodologia d'estudi d'Estats de Transició

3.3.1 Estudi del mecanisme de reacció

La metodologia per estudiar els mecanismes radicalaris es fonamenta en tècniques teòriques cada cop més similars a les utilitzades en química experimental. En la química teòrica el programa informàtic és equivalent a l'equipament del laboratori experimental, ja que constitueix l'eina habitual de treball.

En termes teòrics una reacció o procés biològic es pot descriure com:



sent TS un estat de transició, no han d'existir necessàriament ni el reactiu B ni el producte D, si bé poden combinar-se tres o més reactius i poden formar-se més productes. En la nostra proposta teòrica *A*, *B*, *C* i *D* corresponen a espècies estables i *TS* representa al *saddle point* (punt estacionari) situat en la coordenada de reacció que és el camí que connecta als reactius amb els productes [13-14] i mitjançant la qual es produeix la reacció.

Abans de caracteritzar qualsevol reacció i proposar possibles camins de reacció cal optimitzar les geometries de totes les espècies estables. Cada optimització necessita d'un càlcul per separat, igualment es requereix obtenir les constants de força per les geometries obtingudes. Un sistema en estat fonamental (*ground state*) ha de tenir totes les constants de força positives. Si s'obté alguna constant de força negativa, aleshores, cal un refinament de la geometria, per la qual caben diverses estratègies: la més simple seria pertorbar mínimament la posició dels àtoms causants de la constant de força negativa i llavors reoptimitzar la geometria [15].

Un cop hem definit la geometria dels reactius i dels productes, podem caracteritzar l'estat de transició, fet que en ocasions pot representar un càlcul complicat.

Una de les tècniques més utilitzades per la localització d'estats de transició és a partir de l'estudi del camí de reacció de mínim gradient (*reaction path*) [16].

La coordenada de reacció sol ser un paràmetre geomètric d'enllaç que és sensible a la transformació dels reactius en productes, en aquest sentit poden utilitzar-se variables com la distància d'enllaç ó l'angle d'enllaç com coordenades de reacció.

Representant les superfícies d'energia potencial tal i com es mostra en la figura 9 obtindríem gràficament el camí que connectaria al/s reactiu/s amb el/s productes passant per l'estat de transició, aquest camí és la coordenada de reacció també anomenada intrínseca.

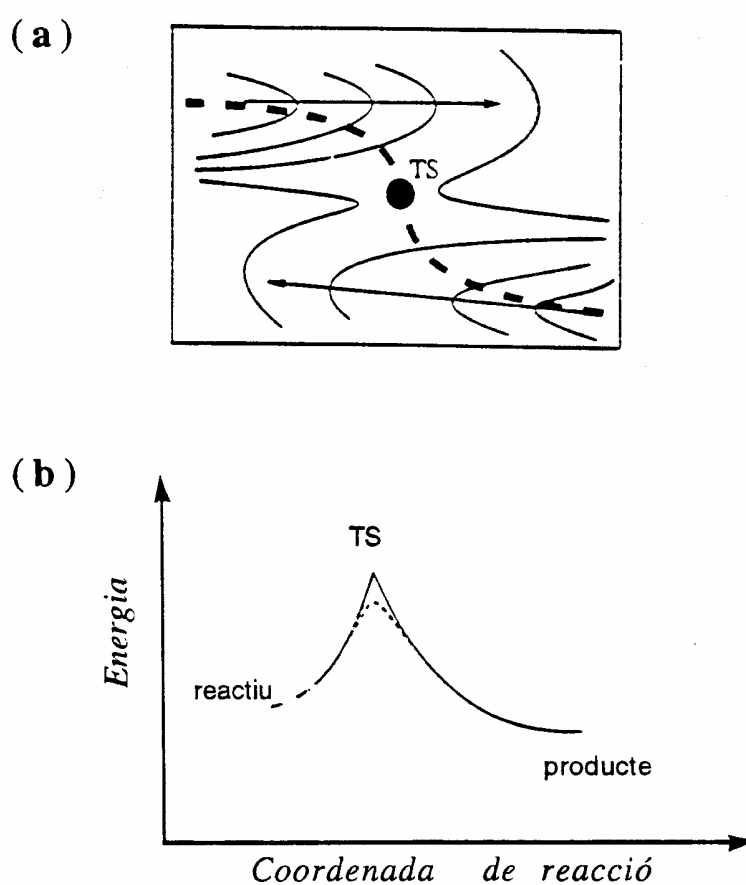


Figura 9. a) Superfícies d'energia potencial per la transformació d'un reactiu en un producte. La línia discontinua correspon a la coordenada de reacció intrínseca que connecta el reactiu amb el producte passant per l'estat de transició (TS). Les fletxes assenyalen els *paths* optimitzats per el reactiu i producte. b) Representació de la coordenada de reacció envers l'energia de formació, en línia discontinua s'observa el perfil de la coordenada de reacció intrínseca i en línia contínua s'observa el perfil dels *paths* pel reactiu i producte.

El perfil energètic de la coordenada de reacció pot obtenir-se representant la coordenada de reacció (en abscisses) envers l'entalpia de formació (en coordenades), tal i com mostra la figura 9b, el punt més elevat d'aquest perfil energètic és l'estat de transició l'optimització del qual conduiria a l'obtenció de la geometria o punt per el que realment passa la coordenada de reacció (figura 9a).

En certs casos, on l'estat de transició és simètric (per exemple, l'abstracció de l'hidrogen d'un metà es poden aplicar simplificacions de simetria, però en altres, el *path* de reacció és tan complicat que fa que l'obtenció de l'estat de transició sigui aproximat a partir de dues direccions oposades, aquest mètode és coneix com la tècnica de *saddle* i habitualment s'utilitza per localitzar estats de transicions en moltes reaccions [17].

Una vegada s'ha realitzat l'aproximació a l'estat de transició, cal utilitzar tècniques de minimització de gradient per tal d'optimitzar-lo. A mode d'exemple els mètodes més clàssics són el de Bartels [16] i el de McIver-Komornicki [14]. Per tenir la certesa que hem obtingut un estat de transició ens cal avaluar-lo calculant les seves constants de força. Llevat de casos molt excepcionals, tot estat de transició ha de tenir una constant de força negativa. Sempre cal tenir en compte, que pot donar-se la possibilitat que encara que s'obtingui un estat de transició, aquest pot no ser un veritable estat de transició ja que no connecta el/s reactiu/s amb el/s producte/s [18].

3.3.2 Localització i identificació d'estats de transició

En l'apartat anterior hem vist que una estructura de transició en una reacció química correspon a un punt estacionari que es troba situat entre dos valls de mínima energia potencial, sent ell un màxim local al llarg d'una sola direcció i un mínim per a tota la resta de direccions ortogonals restants. Aquesta única direcció s'anomena vector de transició i correspon al vector propi (*eigenvector*) i està associat amb el valor propi (*eigenvalue*) de la matriu de constants de força. Aquest valor propi per una estructura d'estat de transició sempre és negatiu.

En general, l'orientació del vector de transició no es coneix *a priori* i per tant, ha de ser determinada en el decurs de l'optimització de l'estat de transició [19].

En termes teòrics la funció energia potencial que depèn de les $3N$ coordenades cartesianes de N àtoms del sistema pot definir-se per $E(q_1, q_2, \dots, q_n)$. El gradient de la funció energia potencial es designa com g i té com components a les derivades parcials de E respecte a cadascuna de les coordenades tal i com es representa en l'equació 3.10:

$$g_i = \frac{\partial E(q)}{\partial q_i} \quad i = 1, 2, 3, \dots, 3N \quad (3.10)$$

La matriu de la segona derivada parcial de l'energia potencial també anomenada matriu de la constant de força es designa per A i matemàticament pot expressar-se per l'equació 3.11:

$$A_{ij} = \frac{\partial^2 E(q)}{\partial q_i \partial q_j} \quad (3.11)$$

Per tal que un punt q^T sigui estat de transició d'una reacció química, cal que s'acompleixin les següents condicions [14,20]:

1 - q^T ha de ser un punt estacionari

Això significa, que el gradient d'energia avaluat en el punt q^T ha de ser zero:

$$g(q^T) = 0$$

Aquesta condició assegura la invariabilitat topològica dels estats de transició en el sentit que la condició s'acompleix independentment del sistema de coordenades utilitzat [20].

2 - La matriu de la constant de força A^T en l'estat de transició ha de tenir un i tan sols un valor propi negatiu (eigenvalue)

Els valors propis de A^T quantifiquen la curvatura de la superfície energia potencial al llarg de cada direcció principal definida per cada vector propi (eigenvector). La presència d'un valor propi negatiu ens indica que el punt q^T correspon a un *saddle point*.

Un mínim local en la regió de l'estat de transició podria presentar una constant de força sense cap valor propi negatiu, amb el qual hauria de catalogar-se com un

intermedi estable o de vida llarga. En cas contrari, és a dir que tingués més d'un valor propi negatiu, aleshores, indicaria que existeix un altre *saddle point* de menor energia i per tant caldria refinar la metodologia per poder aïllar-lo.

3 - q^T ha de ser el punt energètic més elevat en una línia contínua que connecti els reactius amb els productes

Aquesta condició consisteix en un altre recurs per a identificar als reactius i als productes que estan en el mateix camí conjuntament amb l'estat de transició q^T . En una reacció d'una sola etapa, la línia correspon a aquell traçat del perfil en què l'energia disminueix progressivament des del valor de l'estat de transició al valor dels reactius o productes.

Per una reacció amb més d'una etapa, la condició passa per identificar q^T com el *saddle point* de major energia. Un mètode senzill de construcció d'una línia podria ser variant les coordenades entre els valors dels reactius o dels productes i calculant l'energia en el major número de punts al llarg de la línia que connecta els reactius amb els productes, però aquesta tècnica presenta inconvenients ja que cal prestar atenció als possibles canvis estereoquímics o als conformacionals però que no es corresponen realment a un canvi d'espècie química.

En el mateix *saddle point* la línia ha de tenir un component diferent de zero al llarg de tota la direcció de la curvatura més negativa. Aquesta direcció, ve definida per el vector propi que correspon al valor propi negatiu de A^T i pot ser utilitzada per iniciar la minimització d'energia seguint el sentit al llarg d'aquesta direcció que conduiria als reactius i en sentit oposat conduiria als productes tal i com es mostrava en la figura 9.

Si la minimització no convergeix en els reactius o productes d'interès, aleshores apareix la possibilitat que la reacció consti de varies etapes i per tant caldrien noves investigacions del mecanisme de reacció per pensar en l'existència de varies etapes o intermedis de reacció.

4 - q^T ha de tenir la norma de gradient més baixa

Aquesta condició i l'anterior són sempre complementàries. El gradient de l'energia potencial s'expressa matemàticament per l'equació 3.12 on cada g_i correspon a un component del gradient:

$$\sigma(q) = \sum_{i=1}^{3N} g_i^2(q) \quad (3.12)$$

La funció norma de gradient σ és la funció suma dels quadrats per aquest fet sempre serà positiva exceptuant en els punts estacionaris d'energia E , per els quals la funció pren els valors de zero i correspon a mínims locals que poden localitzar-se pel mètode dels mínims quadrats.

El problema per la minimització de σ sol ser el seleccionar una geometria de partida que sigui la més acurada possible, ja que per definició l'optimització de E sempre donarà una geometria d'equilibri, en tant que la minimització de σ pot portar a un mínim, màxim o *saddle point* de E . A més pot donar-se el cas que la superfície de σ pugui tenir dos o més mínims locals addicionals on σ no sigui zero, donat que el nombre de mínims de σ és major que el nombre de mínims de E .

Cal tenir present que el grau de convergència del mètode de minimització dependrà en major o menor mesura del grau d'aproximació al vertader mínim. En càlculs de geometries d'equilibri, es pot obtenir una geometria inicial a partir de la geometria experimental, principi en que es fonamenten tots els mètodes semiempírics, però les geometries dels estats de transició no han estat mesurades experimentalment degut a la natura inestable dels enllaços en aquestes espècies químiques, és per això que en principi resulta molt difícil preveure la seva estructura.

Malgrat tot moltes reaccions consisteixen en la ruptura o extensió d'un enllaç. Aleshores pot ser possible "forçar" aquesta situació mitjançant la variació de petits increments d'aquest enllaç des del valor dels reactius als valors dels productes o a l'inrevés, però sempre minimitzant l'energia respecte als graus de llibertat que queden per a cada valor de l'increment de la longitud d'enllaç.

Si aquest procediment realment està obtenint un punt molt proper a l'estat de transició, inicialment σ augmentarà, aleshores passarà per un màxim i finalment començarà a decreixer quan el *saddle point* ja sigui proper. Els graus de llibertat que han estat optimitzats haurien de canviar suaument des dels valors dels reactius fins als valors dels productes. Una vegada s'ha assolit el valor màxim de σ , la minimització hauria d'obtenir en principi el *saddle point*.

El procés anteriorment descrit de "forçar" la reacció fins a trobar la regió de l'estat de transició pot ser no vàlid en certs casos. Per altra banda, la quantificació de σ a més de les constants de força, dels valors i vector propis pot representar en reaccions complexes una despesa computacional realment extraordinària.

3.3.3 Minimització de la norma de gradient

La minimització de la funció norma de gradient σ es porta a terme per diversos mètodes, sent els més clàssics el mètode de Bartels o bé amb el de McIver-Komornicki que habitualment també es coneixen per les *keywords* NLLSQ i SIGMA respectivament amb què s'utilitzen en els programes de càlcul. Ambdós mètodes es fonamenten en la teoria dels mínims quadrats [21].

Si assumim que el punt q^0 en la proximitat a l'estat de transició es conegut. Aleshores expandint $\sigma(q)$ en una sèrie de Taylor sobre q^0 fins a segon ordre obtindrem l'expressió de l'equació 3.13:

$$\sigma(q) = \sigma(q^0) + (q-q^0) V^0 + \frac{1}{2} (q-q^0) B^0 (q-q^0) \quad (3.13)$$

On V^0 és un vector gradient amb uns components que venen definits per:

$$V_i^0 = \left[\frac{\partial \sigma(q)}{\partial q_i} \right] \quad (3.14)$$

ó bé per la notació vectorial:

$$V_i^0 = (\nabla \sigma)_{q=q_0} \quad (3.15)$$

On ∇ és l'operador gradient (un vector columna). Els elements de la matriu B venen definits per:

$$B_{ij}^0 = \left[\frac{\partial^2 \sigma(q)}{\partial q_i \partial q_j} \right]_{q=q_0}$$

(3.16)

ó:

$$B^0 = (\nabla \nabla' \sigma)_{q=q_0} \quad (3.17)$$

El punt mínim q^{min} ha de ser un punt estacionari, per tant, a partir de l'equació 3.18:

$$\nabla \sigma = 0 = \nabla^0 + B^0 (q^{min} - q^0) \quad (3.18)$$

Si σ és una funció quadràtica pura de q , això és, si els termes d'ordre més elevat de les equacions 3.13 i 3.18 són zero, aleshores el punt mínim q^{min} s'obtindrà a partir de l'equació 3.19 sempre i quan la matriu B^0 sigui definida com valor positiu:

$$q^{min} = q^0 - (B^0)^{-1} \nabla^0 \quad (3.19)$$

D'altra banda, q^{min} serà un *saddle point* o un màxim. Si σ no correspon a un valor quadràtic pur, llavors q^{min} serà sols una aproximació al mínim. L'equació 3.19 es resol iterativament substituint q^0 per q^{min} i avaluant les derivades en les equacions 3.14 i 3.16 sobre el nou punt i substituint en 4.10 tornant a obtenir q^{min} , aquest procés es repeteix fins a complir:

$$(B^0)^{-1} \nabla^0 = 0 \quad (3.20)$$

Aquest mètode es coneix com el mètode de *Newton-Raphson* generalitzat, presenta l'avantatge de convergir ràpidament en la proximitat d'un mínim, però té l'inconvenient que necessita del càlcul de la primera i segona derivades de σ .

Per la localització dels estats de transició en la superfície d'energia potencial ve definida per l'equació 3.21 com el producte dels components de gradient d'energia potencial:

$$\sigma = g' \quad (3.21)$$

per tant:

$$\nabla^0 = \nabla \sigma = 2g \quad (3.22)$$

i:

$$B^0 = \nabla \nabla' \sigma = 2 (A \quad A \quad + \quad C) \quad (3.23)$$

On A és la matriu constant de força ja definida en l'equació 3.11. En quant a la matriu consta dels elements fins a la tercera derivada de l'energia potencial:

$$C_{ij} = \sum_K^{3N} g_k \left[\frac{\partial^2 g_k}{\partial q_i \partial q_j} \right] \quad (3.24)$$

En arribar a l'estat de transició $C=0$ aleshores podríem escriure l'equació 3.23 com :

$$B^0 = 2 \quad A \quad A \quad , \quad (3.25)$$

Aquesta equació constitueix una aproximació al mètode generalitzat dels mínims quadrats. El seu interès radica en una simplificació del càlcul depenent només del càlcul de les constants de força de A convergint sempre a un punt mínim de σ . Aquest procediment utilitza les direccions linealment independents definides per les coordenades q_{ij} , qualsevol nou punt de partida requerirà la recerca d'un mínim al llarg d'aquesta nova direcció, la nova matriu B^0 es construeix a través de l'equació 3.17.

El procés es repeteix fins que $g=0$ dins d'una tolerància específica.

Des del moment que es necessita l'invers de B^0 en l'equació 3.19, el mètode no pot tolerar dependències lineals de q_r .

Si s'utilitzen les $3N$ coordenades cartesianes existiran sis direccions linealment dependents: tres que corresponen a translació pura i tres a rotació pura. Per tant, utilitzant sis de les q_{ij} per definir les coordenades cartesianes d'una molècula, quedaran $3N-6$ graus de llibertat linealment independents per la minimització de σ .

Per exemple, el sistema de coordenades es pot definir de forma que el primer àtom) es situï a l'origen, el segon àtom en l'eix de les z i el tercer (que no és colineal amb els dos primers) es localitzi en el pla xz . De forma que en la minimització de σ sabem que $x_1=y_1=z_1=x_2=y_2=z_2=0$ i aquestes sis coordenades poden eliminar-se de la llista de les $3N$ q_{ij} . Tanmateix els components de gradient d'aquestes sis coordenades poden ser eliminats ja que si no hi han forces externes al sistema deuen ser zero quan els components restants del gradient $3N-6$ són zero.

Malgrat això, l'experiència demostra que en el mètode de Powell la convergència millora quan s'utilitzen tots els components de gradient. També s'ha trobat que el mètode genera una matriu B^0 complexa que no pot ser formalment invertida. Però aquest inconvenient s'ha solucionat seguint tres etapes per invertir B^0 : primer diagonalitzant B^0 , segon invertint els valors propis i tercer transformant-los amb els vectors propis de B^0 per obtenir finalment B^{0-1} .

Si un dels valors propis és zero (o numèricament insignificant), aleshores la recerca unidimensional del mètode de Powell té lloc al llarg de la direcció del vector propi de B^0 corresponent a aquest valor propi zero. Des del moment en què les geometries d'equilibri poden ser determinades mitjançant la minimització de σ el mètode de Powell és comparable als mètodes d'optimització de geometria.

3.3.4 Càlcul de les constants de força

Des del moment en què les constants de força corresponen a les segones derivades parcials de l'energia, poden ser avaluades mitjançant la teoria de la pertorbació [21]. És convenient i computacionalment eficient avaluar-les numèricament utilitzant el mètode de les pertorbació que passarem a descriure molt resumidament.

A partir de l'equació 3.11 un element de la matriu constant de força A^T en l'estat de transició pot escriure's com:

$$A^T_{ij} = \left[\frac{\partial^2 E(q)}{\partial q_i \partial q_j} \right]_{q=q^T} = \left[\frac{\partial g_i(q)}{\partial q_j} \right]_{q=q^T} = \left[\frac{\partial g_i(q)}{\partial q_i} \right]_{q=q^T} \quad (3.26)$$

Segons els mètodes que utilitzem, el vector g s'obté en el càlcul de l'energia potencial, és convenient considerar una columna A^T :

$$A^T_j = \left(\frac{\partial g_j}{\partial q_i} \right)_{q=q^T} \quad (3.27)$$

on A^T_j denota la columna j de A^T .

En el mètode de pertorbació finita les derivades de l'anterior equació s'avaluen numèricament utilitzant diferències finites. En primer lloc g s'avalua en diferents punts espaiats sobre g^T i separats mitjançant un increment fixat h en q_j .

$$\delta g_m(j) = g(q_1^T, q_2^T, \dots, q_j^T + mh \dots q_n^T) \quad m=-n-2, -1, 0, 1, 2, \dots n \quad (3.28)$$

L'increment h s'escull de forma que sigui el mateix per qualsevol j . Per a un valor de n establert, s'ha demostrat que la magnitud de h pot escollir-se de forma que l'error quedi minimitzat. En el cas dels càlculs de la constant de força, hi han dos significats addicionals en l'estimació de l'error.

El primer és el grau d'asimetria de la matriu de constants de força. Des del moment que la matriu es calculada columna per columna A_j^T s'ha d'aproximar a A_{ij}^T , d'aquesta forma el major element de la matriu $|A_j^T - A_{ij}^T|$ és una estimació aproximada de l'error.

En segon lloc, si són emprades les $3N$ coordenades cartesianes A^T tindrà sis valors propis de valor zero. Els corresponents vectors propis poden també examinar-se per conèixer l'exactitud amb què corresponen a la translació i rotació pures i a més si el sistema guarda o no alguna simetria, la resta de vectors propis seran útils per saber la precisió amb què es transformen segons les representacions apropiades.

Un factor important a tenir en compte és la gran despesa computacional que comporten aquests càlculs. Es desitjable reduir al màxim el consum de CPU mantenint el càlcul de les constants de força a una exactitud raonable. En principi des que compleix $g(q)^T=0$ correspondria a una regió de la superfície d'energia potencial propera a l'estat de transició i l'equació 3.28 es simplificaria en:

$$\delta g_m^0(j) = \delta g_{-m}^0(j) \quad (3.29)$$

per tant, només dos punts per m positius (o negatius) necessiten ser calculats. Això suposa per el càlcul de les constants de força d'un estalvi de la meitat de la despesa computacional. A la pràctica, per localitzar l'estat de transició amb suficient precisió fent servir l'equació 3.29 ja s'inverteix un gran temps de càlcul en la minimització de σ .

3.3.5 Mètodes d'optimització d'estats de transició

Sigui $f(x)$ una funció que expandim en sèrie de Taylor fins a segon ordre:

$$f(x) = f(x_0) + (x-x_0)^T \nabla f(x_0) + \frac{1}{2} (x-x_0)^T \nabla^2 f(x_0) (x-x_0) \quad (3.30)$$

on $x = (x_1, x_2, \dots, x_n)$ i $x_0 = (x_1^0, x_2^0, \dots, x_n^0)$

derivant respecte a x obtindrem :

$$\nabla f(x) - \nabla f(x_0) = [\nabla^2 f(x_0) (x-x_0)] \quad (3.31)$$

Per tant si a la darrera equació l'imposem la condició de punt estacionari de la pàgina 38, és a dir: $\nabla f(x) = 0$ la primera derivada de la funció també serà zero i tindrem el mètode *Newton-Raphson generalitzat*. Després de reordenar l'equació 3.31 tindrem:

$$x = x_0 - H^{-1}(x_0) g(x_0) = x_0 + p(x_0) \quad (3.32)$$

on $H^{-1}(x_0) = [\nabla^2 f(x_0)]^{-1}$ i $g(x_0) = x - x_0$ (3.33)

La condició suficient per localitzar un punt mínim estacionari és: $H(x_0) > 0$ és a dir que l'hessiana sigui en tot moment positiva.

Uns altres mètodes utilitzats per localitzar estats de transició que constitueixen punts estacionaris són els mètodes *quasi Newton-Raphson* també coneguts com de mètrica variable, ja que tenen utilitat per funcions no quadràtiques, en aquests mètodes l'estratègia consisteix en assegurar sempre la positivitats de l'hessiana a través d'un algoritme matemàtic α que alhora calcula la segona derivada:

$$x_{i+1} = x_i - \alpha_i A(x_i) g(x_i) \quad (3.34)$$

on $A(x_i) = H^{-1}(x_i)$ i d'aquí $x_{i+1} = x_i - \alpha_i p(x_i)$ (3.35)

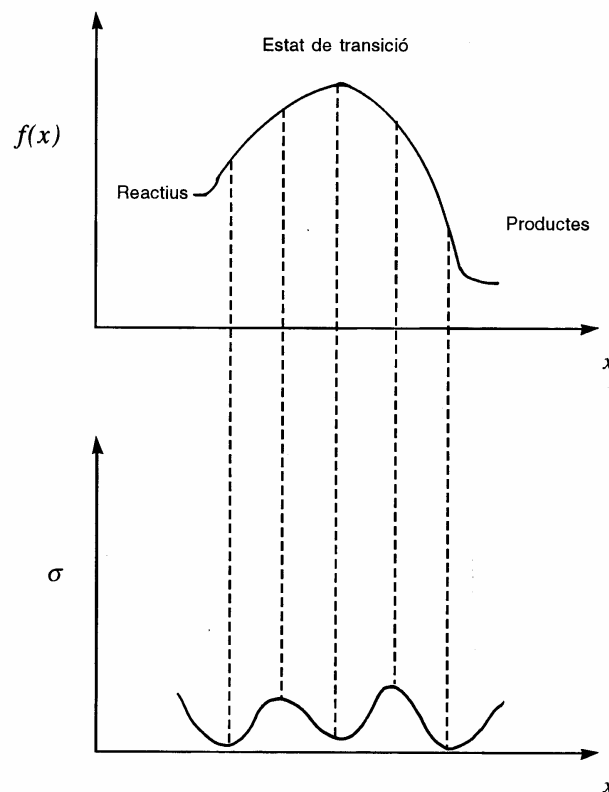


Figura 10. Estratègia seguida per alguns mètodes per localitzar estats de transició (McIver - Komornicki, 1972). Consisteix en transformar matemàticament l'estat de transició (inicialment màxim) en un mínim estacionari a partir de la minimització de la norma de gradient $f(x)$.

De forma molt resumida el procediment general dels mètodes *quasi Newton-Raphson* seria el següent:

- 1r) Obtenció de la matriu inversa de l'hessiana aproximada a $A(x_i)$ en el punt x_i
- 2n) Recerca del mínim de la funció $f(x_{i+1})$ en la direcció del vector $p(x_i)$, la variació del paràmetre α ens assegurarà que el vector s'hi troba en la direcció de recerca, sent ortonormal a ella, matemàticament:

$$\begin{bmatrix} \frac{\partial f(x_{i+1})}{\partial \alpha_1} \end{bmatrix} = \begin{bmatrix} \frac{\partial f(x_{i+1})}{\partial x_{i+1}} \end{bmatrix} \times \begin{bmatrix} \frac{df(x_{i+1})}{d\alpha_1} \end{bmatrix} = \frac{\mathbf{g}^T(x_{i+1}) \times \mathbf{p}(x_i)}{0} \quad (3.36)$$

- 3r) Comprovació amb el test de convergència (ex. *Peter test* per recerca lineal o un altre *Everstest*) si no s'assoleix la convergència pot ser degut a la gran diferència de gradients ($\mathbf{g}_i - \mathbf{g}$), llavors cal retornar al punt 1r.

La funció $f(x)$ a optimitzar a partir de la qual hem d'obtenir un mínim estacionari sol ser generalment l'energia molecular, l'entalpia de formació o bé la norma del gradient. Aquesta última és la que fan servir els mètodes SIGMA o NLLSQ (McIver-Komornicki, 1972) de forma que transformen l'estat de transició en un mínim estacionari tal i com es representa matemàticament per l'equació 3.37 i gràficament en la figura 10.

$$\sigma(x) = \sum_{i=1}^I g_i^2(x) \quad \text{on} \quad g_i(x) = \frac{\partial f(x)}{\partial x_i} \quad (3.37)$$

Finalment assenyalarem que aquests mètodes asseguren amb una certa fiabilitat la localització dels mínims locals o punts estacionaris, per aquest motiu es troben implementats en la majoria de programes de càlcul, en la taula1 descrivim alguns d'ells:

Funció d'optimització $f(x)$	Nom mètode	Implementat en el programa
Energia molecular	DAVIDSON-FLETCHER-POWELL [22-23]	MOPAC 6.0 [2]
	BROYDEN-FLETCHER-GOLDFARB-SHANNO (BFGS) [24]	
	CHAIN [26]	AMPAC 6.0 [3]
	SCHLEGEL [19] MURTAGH-SARGENT [25] BERNY [27]	GAUSSIAN-94 [4]
Norma del gradient	SIGMA (POWSQ) [14]	MOPAC 6.0
	NLLSQ (Bartels) [16]	
	SADDLE [28]	
	TS [29]	
$f(x)$ com màxim	SCHLEGEL[19]	GAUSSIAN-94

Taula 1. Mètodes més utilitzats per a la localització de punts estacionaris, s'indica la funció que optimitzen així com els programes de càlcul en els quals es troben implementats

3.4 Mecànica Clàssica

Una filosofia completament diferent als mètodes quàntics és la de fer servir les lleis de la física clàssica per estudiar sistemes químics, bioquímics o farmacològics. En aquesta aproximació no descriurem els electrons, sinó que els àtoms seran representats

per esferes interpenetrables, amb càrregues i masses característiques i els enllaços seran tractats com molles deformables.

La dificultat inicial per fer aquest tipus de tractament és que necessitem tot un conjunt de paràmetres empírics (masses, radis i càrregues atòmiques, duresa de les esferes, distàncies d'enllaç....) de vegades gens trivials d'obtenir. Al conjunt de paràmetres i formulacions matemàtiques que descriuen el sistema se'ls anomena camp de forces (*force fields*). S'han desenvolupat diversos camps de forces (AMBER [30], GROMOS [31], CHARMM [32]...) que es diferencien en els valors dels paràmetres, el procés de parametrització, el tractament de diversos tipus d'àtoms (*all-atom / united atom force fields*), afegir o no un terme explícit per als ponts d'hidrogen, etc.

3.4.1 Camps de Forces AMBER i GROMOS

En un camp de forces l'energia potencial del sistema ve definida per la funció $V_{\text{sist}}(r, p)$ que depèn de les coordenades r del sistema i dels paràmetres p del camp de forces i es divideix entre termes enllaçants o *bonded* (V_{bond}) i no enllaçants o *non-bonded* (V_{nonb}) tal i com mostra l'equació 3.38:

$$V_{\text{sist}} = V_{\text{bond}} + V_{\text{nonb}} \quad (3.38)$$

Els termes enllaçants inclouen contribucions degudes als enllaços covalents (V_{enl}), als angles d'enllaç (V_{ang}) i a la torsió dels angles dihedres propis (V_{dih}) i impropis (V_{imp}).

$$V_{\text{bond}} = V_{\text{enl}} + V_{\text{ang}} + V_{\text{dih}} + V_{\text{imp}} \quad (3.39)$$

$$V_{\text{nonb}} = V_{\text{vdw}} + V_{\text{ele}} \quad (3.40)$$

D'altra banda els termes no enllaçants (equació 3.40) venen definits per un terme d'atracció-repulsió tipus Lennard-Jones que reproduïx les forces de van der Waals (V_{vdw}) i un terme electrostàtic (V_{ele}).

Els termes enllaçants només cal calcular-los per aquells grups de dos (enllaç), tres (angle), o quatre àtoms (dihedres propis i impropis) que s'hagin definit com a tals en la topologia. Per contra, els termes no enllaçants cal calcular-los en principi per a totes les parelles d'àtoms possibles, amb les dues limitacions següents:

- Àtoms veïns exclosos. No es calculen els termes no enllaçants entre parells d'àtoms units directament (interaccions 1-2) o a través de dos enllaços (interaccions 1-3), perquè a distàncies curtes aquests termes tindrien un pes exageradament gran i són innecessaris donat que els termes enllaçants ja reproduïxen la interacció d'aquests parells. En el camp de forces GROMOS, les interaccions 1-4 (aquelles que es donen entre parells d'àtoms units a través de 3 enllaços) també s'exclouen quan es tracta d'àtoms que formen part o estan units a nuclis aromàtics.
- Parells d'àtoms a una distància major d'un límit establert. Donat que les interaccions no enllaçants decauen amb la distància, en sistemes grans s'imposa un límit (generalment entre 8 i 14 Å) més enllà del qual aquestes no es calculen.

3.4.1.1 Termes d'Enllaç Covalent

Es pot reproduir el perfil propi de l'energia d'enllaç per a un ampli rang de distàncies mitjançant expressions exponencials, com un potencial de Morse, o polinòmiques, però donat que normalment simularem sistemes amb distàncies d'enllaç properes a l'equilibri, els enllaços covalents es poden descriure mitjançant la llei de Hooke (oscil·lador harmònic) o equacions similars. Definint una distància d'enllaç d'equilibri (d_0) i una constant de força d'enllaç (K_e), per a un sistema amb N_e enllaços, el programa AMBER calcula l'energia d'enllaç seguint la llei de Hooke:

$$V_{\text{enl}} = \sum_{n=1}^{N_e} K_{e_n} [d_n - d_{0_n}]^2 \quad (3.41)$$

En el cas de GROMOS s'utilitza una versió lleugerament modificada de l'equació anterior que reproduïx el mateix tipus de perfil:

$$V_{\text{enl}} = \sum_{n=1}^{N_e} \frac{1}{4} K_{e_n} [d_n^2 - d_{0_n}^2]^2 \quad (3.42)$$

3.4.1.2 Termes d'Angles d'Enllaç.

Els angles d'enllaç també es poden descriure mitjançant un potencial harmònic, amb un angle d'equilibri θ_0 i una constant de força d'enllaç K_a . En el cas de GROMOS, s'ha escollit una funció trigonomètrica, que per a un sistema de N_a angles d'enllaç s'expressaria com:

$$V_{\text{ang}} = \sum_{n=1}^{N_a} \frac{1}{2} K_{a_n} [\cos\theta_n - \cos\theta_{0_n}]^2 \quad (3.43)$$

mentre que per al cas d'AMBER s'ha optat per una expressió equivalent a la d'enllaç:

$$V_{\text{ang}} = \sum_{n=1}^{N_a} K_{a_n} [\theta_n - \theta_{0_n}]^2 \quad (3.44)$$

3.4.1.3 Termes d'Angles Dihedres Propis

A diferència dels dos termes anteriors, els angles dihedres no tenen un sol valor d'equilibri. Per exemple, en la rotació de l'età cada hidrogen té tres mínims possibles (60° , 180° i 300°) i tres màxims (0° , 120° i 240°). Així que en el cas dels dihedres no podem restringir el moviment al voltant d'un punt de referència, sinó que ens caldrà una funció més complexa que reproduïxi el perfil energètic desitjat. Els camps de força AMBER i GROMOS han resolt aquest aspecte de diferent manera.

En el camp de forces d'AMBER s'ha adoptat l'expressió següent:

$$V_{\text{dh}} = \sum_{n=1}^{N_d} K_{d_n} [1 + \cos(m_n\phi_n - \delta_n)] \quad (3.45)$$

on N_d és el nombre de dihedres, K_d és la meitat de la magnitud de la barrera energètica, m és el nombre de barreres de potencial en 360° (periodicitat), ϕ és el valor que pren el dihedre i δ és el que anomenem fase (ambdós expressats en graus). La fase generalment pren valors de 0° o 180° , indicant la presència d'un màxim o un mínim respectivament a 0° . A cada dihedre se li pot aplicar una o més funcions com la descrita, podent així obtenir perfils amb diferents valors mínims i màxims, tal com es mostra en la figura 11.

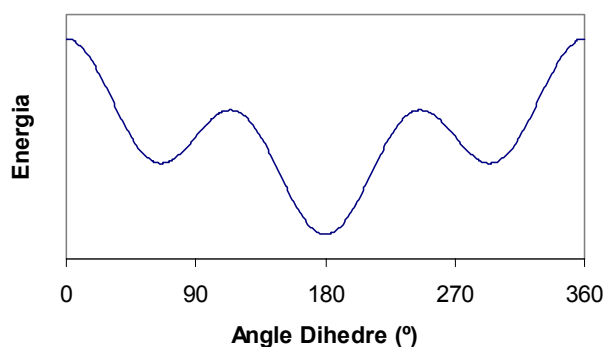


Figura 11. Perfil de l'energia potencial respecte al valor d'un angle dihedre descrit com la suma de dues funcions trigonomètriques.

En el camp de força GROMOS l'equació emprada és:

$$V_{\text{dih}} = \sum_{n=1}^{N_d} K_{d_n} [1 + \cos(\delta)\cos(m_n\phi_n)] \quad (3.46)$$

on K_d és la constant que donarà el valor d'energia en el màxim; δ és un paràmetre limitat als valors de 0 o π radians, de manera que $\cos(\delta)$ valdrà 1 o -1 ; m és un nombre enter entre 1 i 6 que indica el nombre de cicles idèntics que hi haurà per volta (periodicitat) i, com abans, ϕ és el valor que pren el dihedre.

Com s'aprecia és una formulació lleugerament diferent de l'emprada en AMBER, però completament equivalent per els valors de fase emprats comunament (0° i 180°). El camp de forces GROMOS no contempla la possibilitat d'aplicar més d'una funció a un sol dihedre, de manera que són els termes no enllaçants els que hauran de fer asimètric el perfil de rotació quan sigui necessari.

3.4.1.4 Termes d'Angles Dihedres Impropis

Els angles dihedres impropis (ξ) són termes útils per forçar determinats àtoms a mantenir-se en el pla format per uns altres tres àtoms. S'apliquen sobretot a sistemes aromàtics quan volem que un substituent tingui una capacitat molt limitada per sortir del pla del sistema. Com en el cas dels enllaços i dels angles d'enllaç, aquí un potencial harmònic amb una constant de força K_i centrat en un valor de referència (ξ_0) pot ser

suficient i és el que s'usa en GROMOS. Per a un sistema amb N_i angles dihedres impropis, l'expressió de l'energia es descriu:

$$V_{\text{imp}} = \sum_{n=1}^{N_i} \frac{1}{2} K_{i_n} [\xi_n - \xi_{0_n}]^2 \quad (3.47)$$

AMBER usa la mateixa expressió que per a un angle dihedre convencional però amb la fase (δ) fixada a 180° , i la periodicitat (m) al valor de 2. El potencial resultant té per tant mínims energètics a 0° i 180° :

$$V_{\text{imp}} = \sum_{n=1}^{N_i} K_{i_n} [1 + \cos(2\xi_n - 180)] \quad (3.48)$$

3.4.1.5 Termes de van der Waals

Les interaccions d'atracció i repulsió que es donen quan es posen en contacte dos àtoms, conegudes genèricament amb el nom d'interaccions tipus van der Waals, es simulen utilitzant un terme atractiu que depèn de la sisena potència de la distància entre els àtoms i i j (r_{ij}) i un altre repulsiu que és funció de la dotzena potència de la mateixa distància:

$$V_{\text{vdw}} = \frac{A}{(r_{ij})^{12}} - \frac{B}{(r_{ij})^6} \quad (3.49)$$

A i B són paràmetres que venen donats pel camp de força i que depenen del tipus d'àtoms que siguin i i j .

3.4.1.6 Termes Electroestàtics

Com sabem, el fet que uns elements siguin més electronegatius que altres, genera en cada molècula una distribució electrònica particular, amb unes regions amb càrrega positiva (dèficit d'electrons) i unes altres amb càrrega negativa (excés d'electrons). Com que les càrregues de diferent signe s'atrauen i les del mateix signe es repeleixen les molècules s'atrauen i es formen el que coneixem com interaccions electroestàtiques. Per poder reproduir-les mitjançant les lleis de la física clàssica, cal parametritzar unes càrregues q centrades típicament en el nucli dels àtoms. Una

vegada fet això, podem usar la llei de Coulomb per conèixer l'energia electrostàtica entre els àtoms i i j separats per un dielèctric ϵ_0 a una distància r_{ij} tal i com mostra l'equació 3.50:

$$V_{\text{ele}} = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (3.50)$$

La majoria de programes informàtics que usen camps de força permeten utilitzar diferents modificacions de la llei de Coulomb per tal d'introduir d'alguna manera l'efecte dielèctric del solvent. Més comuna és la utilització d'una constant dielèctrica dependent de la distància, com la que es mostra en l'equació 3.51 que fins a cert punt reproduir un efecte d'apantallament de les càrregues pel solvent.

$$V_{\text{ele}} = \frac{q_i q_j}{4\pi\epsilon_0 \left(r_{ij} \right)^2} \quad (3.51)$$

3.4.1.7 Forces Especials

Fins aquí s'han definit forces amb sentit físic, però de vegades pot ser útil afegir forces que definim bé de manera arbitrària, bé per tal d'incloure informació obtinguda experimentalment. En conjunt aquestes forces s'anomenen forces de restricció ja que s'utilitzen per fer restriccions en la mobilitat del sistema. Aquests poden ser, entre altres de:

- Distància, si afegim un terme tipus enllaç covalent entre dos àtoms no enllaçats.
- Posició, si apliquem una restricció de distància entre un àtom i un punt de referència en l'espai cartesià.
- Dihedres. A l'angle dihedre format per quatre àtoms se li pot sobreposar una altra funció que tingui un mínim a un cert valor de referència.
- Constants d'acoblament. Es pot afegir un terme que forci el sistema a satisfer observacions obtingudes per ressonància magnètica nuclear (NMR) de les constants d'acoblament *spin-spin* entre dos nuclis.

3.4.1.8 Periodicitat del Sistema

Les forces definides fins aquí ens permeten simular un sistema aïllat, però degut al fet que els àtoms dels extrems no tindran veïns, aquesta regió del sistema es veurà distorsionada respecte a una situació diferent de la fase gas.

Per tal de tenir un sistema més realista podem replicar el sistema de manera poc costosa introduint les anomenades condicions periòdiques de contorn, consistents en

generar una caixa que contingui en el seu interior al sistema i replicar caixa i sistema en totes les direccions de l'espai, tal com s'il·lustra en la figura 12. La caixa pot tenir qualsevol forma que permeti ser replicada infinitament sense deixar espais exclosos, essent les geometries més utilitzades la cúbica i l'octaèdrica. En qualsevol cas el resultat serà que els extrems del sistema ara veuran imatges dels extrems oposats del mateix.

En un sistema bidimensional tal com el dibuixat en la figura 12, cada àtom tindrà 9 posicions (la original més la corresponent a cada una de les 8 rèpliques). Per un sistema tridimensional les posicions possibles seran 27. A l'hora de calcular les forces entre dos àtoms, típicament utilitzarem les coordenades originals per a un d'ells i la posició més propera de les 27 possibles per a l'altre (condició d'imatge mínima). Alternativament, al mètode de Ewald [33], Darden *et al.* [34], i York *et al.* [35] han desenvolupat mètodes que tenen en compte l'efecte sobre la cel·la unitat d'una malla macroscòpica formada per imatges repetides d'aquesta. Mitjançant aquestes tècniques, si el sistema ocupa tot l'espai de la caixa, podem simular sistemes condensats com cristalls o dissolucions.

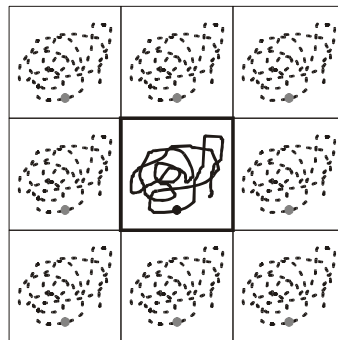


Figura 12. Sistema bidimensional amb condicions periòdiques de contorn.

3.4.2 Mecànica Molecular

La mecànica molecular (a diferència a la dinàmica molecular) considera a un sistema molecular com un camp de forces que pot definir-se per una sèrie de funcions d'energia potencial, aquest camp pot expressar-se per la suma d'aquestes funcions:

$$E_{\text{total}} = E_{\text{extensió}} + E_{\text{plegament}} + E_{\text{dihedral}} + E_{\text{van der Waals}} + E_{\text{electrostàtica}} + E_{\text{pont de H}};$$

L'eina de la mecànica molecular més emprada és la minimització de l'energia que a continuació descrivim.

3.4.2.1 Minimització de l'Energia Potencial

Una vegada tenim la funció energia potencial E_{total} , que té com a variables les coordenades del sistema, podem optimitzar aquesta funció i obtenir per tant una nova geometria a la que li correspon una millor energia. Per fer això disposem de diferents mètodes de minimització, tots ells ben establerts [36].

Els mètodes més senzills per minimitzar són els de cerca directa que troben la geometria de mínima energia a base d'explorar la superfície d'energia potencial generant diferents geometries de prova. Aquest mètodes tenen l'inconvenient que convergeixen molt lentament i per tant gairebé mai són emprats per a sistemes grans.

L'alternativa eficient als mètodes de cerca directa són els mètodes de gradient (o de primer ordre), que requereixen fer la derivada de la funció energia potencial respecte de la posició, obtenint per tant la força. D'entre la varietat de mètodes de gradient, els més utilitzats són:

- *Steepest descents* (descensos de major inclinació). Aquest mètode explora conformacions en la direcció de la força neta; és un mètode força complex però molt eficaç, especialment quan la posició de partida està allunyada del mínim, però que convergeix molt lentament prop del mateix.

- *Conjugate gradients* (gradients conjugats). En aquest mètode la cerca es fa en una direcció que no depèn només del gradient en un punt sinó també de la direcció del pas previ. D'aquesta manera s'arriba a la convergència més ràpidament que en el mètode anterior.

Encara que també existeixen mètodes de segon ordre, aquests són generalment molt poc utilitzats degut a que l'obtenció de la segona derivada pot arribar a ser molt costosa per a sistemes grans com generalment són els tractats per mecànica molecular. La tècnica més habitual consisteix en fer un ús combinat dels mètodes *steepest descents* i *conjugate gradients*.

3.4.3 Dinàmica Molecular

En els darrers anys la visió estàtica de les molècules ha evolucionat considerablement, degut en part a les noves tècniques de dinàmica molecular (*Molecular dynamics*) on cada àtom rep un tractament com si es tractés d'una partícula que apleix les equacions del moviment de Newton: les successives integracions d'aquestes equacions defineixen la trajectòria de cada àtom en termes de velocitat i posició. Els moviments dels àtoms i dels fragments moleculars obtinguts per aquestes simulacions revelen l'organització molecular i fan possible la comprensió dels fenòmens que es produeixen a nivell molecular i que no poden explicar-se sota un punt de vista estàtic.

Durant curts espais de temps de l'ordre de fraccions de *femtosegons* els àtoms semblen oscil·lar, patint fins i tot col·lisions amb grups veïns però mostrant al llarg del temps moviments coherents. L'inconvenient d'aquests mètodes es que requereixen de potents eines informàtiques que incloent tant els programes com les sortides gràfiques.

Les tècniques de dinàmica molecular permeten obtenir gran informació a partir dels seus gràfics moleculars i poden complementar-se amb funcions de correcció per així modular les distàncies d'enllaç, angles d'enllaç i angles dihedres. Poden obtenir-se simulacions de gran qualitat tant de petites molècules com de complexes macromoleculars com proteïnes i àcids nucleics compatibles amb la informació obtinguda en altres tècniques com la RMN o l'espectroscopia atòmica.

Donat que els sistemes tractats mitjançant una formulació clàssica són generalment molt grans, aquests tenen unes hipersuperfícies d'energia potencial terriblement complexes, amb una enorme quantitat de mínims, i resulta completament inabastable el seu estudi quantitatiu mitjançant la mecànica molecular (estàtica). La dinàmica molecular [35-36] permet generar conjunts de configuracions representatives del sistema de manera que podem extreure'n informació estructural i termodinàmica.

Per a una partícula i , les equacions del moviment de Newton prenen la forma 3.52 i 3.53, on r_i és la posició, V_i la velocitat, F_i la força i m_i la massa.

$$\frac{dr_i(t)}{dt} = V_i(t) \quad (3.52)$$

$$\frac{dV_i(t)}{dt} = \frac{F_i(t)}{m_i} \quad (3.53)$$

3.4.3.1 Mètodes d'Integració

Per generar trajectòries de dinàmica molecular s'utilitzen els mètodes de diferència finita, en els quals la integració és un conjunt de petites etapes, cadascuna separada de la següent per un petit increment de temps, Δt . En cada pas (temps t) cal calcular la força total sobre cada partícula del sistema, que serà un vector resultant de sumar la interacció d'aquesta partícula amb cada una de les altres partícules del sistema. De la força, amb la relació $F=m \cdot a$ obtindrem l'acceleració, que combinada amb les posicions i les velocitats ens donaran les noves posicions i velocitats a temps $t + \Delta t$. L'increment de temps utilitzat ha de ser prou petit com per poder fer l'assumpció que la força es manté constant en tot aquest període.

Hi ha diferents algorismes per integrar les equacions del moviment utilitzant els mètodes de diferència finita: Verlet, Beeman, etc., però el l'algorisme que s'utilitza més àmpliament és sens dubte el conegut amb el nom de *leap-frog* (salt de granota), esquematitzat en la figura 13 i que s'explica a continuació.

Leap-frog, com tots els altres algorismes basats en l'aproximació de diferència finita, expressa la posició, velocitat i acceleració com una expansió en sèrie de Taylor:

$$\begin{aligned} r(t + \Delta t) &= r(t) + \Delta t \cdot v(t) + \frac{1}{2} \Delta t^2 \cdot a(t) + \frac{1}{6} \Delta t^3 \cdot b(t) + \frac{1}{24} \Delta t^4 \cdot c(t) + \dots \\ v(t + \Delta t) &= v(t) + \Delta t \cdot a(t) + \frac{1}{2} \Delta t^2 \cdot b(t) + \frac{1}{6} \Delta t^3 \cdot c(t) + \dots \\ a(t + \Delta t) &= a(t) + \Delta t \cdot b(t) + \frac{1}{2} \Delta t^2 \cdot c(t) + \dots \end{aligned} \quad (3.54)$$

on v (velocitat) és la primera derivada respecte del temps de la posició, a (acceleració) és la segona derivada de la posició, b la tercera derivada, c la quarta, etc.

Si expressem la posició $r(t_i + \Delta t)$, a $t = t_i + \Delta t/2$ com una expansió de Taylor, tenim:

$$r(t_i + \Delta t) = r\left(t + \frac{\Delta t}{2}\right) = r(t) + \frac{\Delta t}{2} v(t) + \frac{1}{2} \left(\frac{\Delta t}{2}\right)^2 a(t) + \frac{1}{6} \left(\frac{\Delta t}{2}\right)^3 b(t) + \dots \quad (3.55)$$

A la vegada podem expressar la posició immediatament anterior, $r(t_i)$ també com una expansió de Taylor, també a $t = t_i + \Delta t/2$:

$$r(t_i) = r\left(t - \frac{\Delta t}{2}\right) = r(t) - \frac{\Delta t}{2} v(t) + \frac{1}{2} \left(\frac{\Delta t}{2}\right)^2 a(t) - \frac{1}{6} \left(\frac{\Delta t}{2}\right)^3 b(t) + \dots \quad (3.56)$$

Restant la segona a la primera i menyspreant els termes de tercer ordre, obtenim:

$$r(t_i + \Delta t) = r(t_i) + \Delta t v\left(t_i + \frac{\Delta t}{2}\right) \quad (3.57)$$

Pel que fa a les velocitats, procedim semblantment: fent una expansió de Taylor de $v(t_i + \Delta t/2)$ a temps $t = t_i$:

$$v\left(t + \frac{\Delta t}{2}\right) = v(t) + \frac{\Delta t}{2} a(t) + \frac{1}{2} \left(\frac{\Delta t}{2}\right)^2 b(t) + \frac{1}{6} \left(\frac{\Delta t}{2}\right)^3 c(t) + \dots \quad (3.58)$$

a la que restarem l'expansió de $v(t_i - \Delta t/2)$ també a $t = t_i$

$$v\left(t - \frac{\Delta t}{2}\right) = v(t) - \frac{\Delta t}{2} \cdot a(t) + \frac{1}{2} \left(\frac{\Delta t}{2}\right)^2 \cdot b(t) - \frac{1}{6} \left(\frac{\Delta t}{2}\right)^3 \cdot c(t) + \dots \quad (3.59)$$

una vegada menyspreem els termes de tercer ordre, l'equació 3.59 resulta:

$$v\left(t_i + \frac{\Delta t}{2}\right) = v\left(t_i - \frac{\Delta t}{2}\right) + \Delta t \cdot a(t) \quad (3.60)$$

El primer pas en l'algorisme de *leap-frog* és calcular les velocitats $v(t_i + \Delta t/2)$ a partir de les velocitats a $v(t_i - \Delta t/2)$ i de l'acceleració a temps t_i . A continuació podem trobar les posicions $r(t_i + \Delta t)$ a partir de les posicions $r(t_i)$ i les velocitats v . Tal com s'esquematitza a la figura 10 el càlcul de les velocitats “salta” per sobre de les posicions per obtenir les velocitats a $t_i + \Delta t/2$, a continuació el càlcul de les posicions “salta” per sobre de les velocitats per obtenir les posicions a $t_i + \Delta t$ i així anar fent fins a cobrir el temps desitjat.

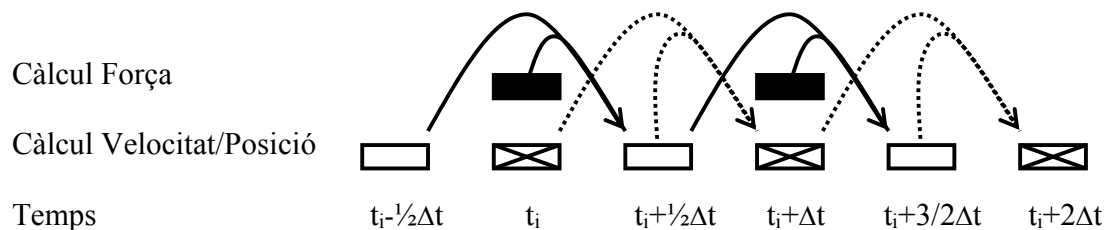


Figura 13. Representació gràfica de l'algorisme de dinàmica molecular *leap-frog*. Els requadres negres representen un càlcul de les forces (acceleracions), de les velocitats els blancs i de les posicions els creuats [29,35].

Un dels problemes d'aquest mètode és que no és possible obtenir els valors de les acceleracions, les velocitats i les posicions per un mateix instant. Les velocitats corresponents a un temps t_i (en el que tenim les posicions) poden ser obtingudes de forma aproximada simplement promitjant:

$$v(t_i) = \frac{1}{2} \left[v\left(t_i + \frac{1}{2} \Delta t\right) + v\left(t_i - \frac{1}{2} \Delta t\right) \right] \quad (3.61)$$

3.4.3.2 Condicions de la Dinàmica

Un dels paràmetres més crucials en tota dinàmica molecular és el pas d'integració. Donat que el càlcul de les forces és el pas més costós, cal escollir un pas d'integració tant gran com sigui possible, de manera que amb el mínim d'integracions simulem el sistema a temps com més llargs millor. El problema consisteix en que per a Δt massa llargs sorgeixen inestabilitats en la dinàmica i no hi hauria conservació de l'energia ni del moment linear. Alguns algorismes tenen millor comportament que altres a aquest respecte, de manera que fins hi tot resultant pas per pas més costosos poden ser escollits si estalvien passos d'integració.

Δt també ve limitat per la freqüència més alta del sistema estudiat, ja que cal que el pas d'integració sigui petit en comparació al temps d'oscil·lació del moviments més ràpid. En una macromolècula, la vibració dels enllaços és el més ràpid dels moviments i permet passos d'integració de 1 fs com a màxim (sempre que disposem d'un algorisme que garanteixi l'estabilitat). Una opció que permet allargar Δt és, doncs, constrenyir artificialment les distàncies dels enllaços a valors fixes, de manera que és possible duplicar el pas d'integració sense introduir inestabilitats a la dinàmica. SHAKE [39] és l'algorisme més popular per mantenir aquestes constriccions.

Seguint l'esquema *leap-frog*, per començar el càlcul requerim unes coordenades inicials ($t=0$) a partir de les quals farem el primer càlcul de les forces i unes velocitats a $t = -\Delta t/2$. El conjunt de coordenades inicials ha de ser raonable i moltes vegades provenen de dades experimentals (raig X o RMN), mentre que les velocitats inicials s'assignen a l'atzar seguint una distribució tipus Maxwell-Boltzmann corresponent a la temperatura inicial de la dinàmica.

Altres factors claus en una dinàmica molecular són les condicions de simulació: temperatura o energia (depenent de si la simulació és isotèrmica o isoenergètica) i pressió o volum (per simulacions isobàriques i isocòriques respectivament). Generalment les dinàmiques moleculars de sistemes d'interès bioquímic realitzades en condicions periòdiques de contorn es fan a temperatura i pressió constants, sovint a 25°C i a 1 atm.

Per tal de mantenir la temperatura del sistema a un valor desitjat, i donat que la temperatura guarda proporcionalitat amb l'energia cinètica, només cal escalar les velocitats per un factor adequat. Això es pot fer directament a cada pas o de manera més suau a través d'un bany tèrmic, seguint el protocol de Berendsen [40].

Pel que fa a la pressió, pot ser ajustada a un valor simplement escalant el volum per tal que aquest es correspongui al volum adequat a la temperatura i al sistema de treball. Alternativament es pot acoblar la pressió a una pressió de referència, anàlogament al que es fa amb la temperatura. Les fluctuacions en la pressió són normalment molt grans, però donada la poca compressibilitat dels sistemes condensats els canvis de volum són en general molt petits.

3.4.3.3 Recollida de Dades

Una dinàmica molecular en general s'allargarà entre pocs ps (picosegons) fins a μs (de l'ordre de 10^3 fins a 10^9 etapes d'integració) i a cada etapa té un conjunt propi de coordenades i velocitats que els correspon uns valors de temperatura, pressió, energia potencial, etc. que s'utilitzaran posteriorment per l'anàlisi de la trajectòria. Donat que la quantitat d'informació generada és molt gran i la capacitat d'emmagatzematge limitada, normalment caldrà fer una tria del material més interessant.

Els anàlisis més corrents que es realitzen a una trajectòria de dinàmica molecular (desviació quadràtica mitja respecte a una estructura de referència, seguiment de distàncies o energies d'interacció, estudis conformacionals, etc.) requereixen només de les coordenades del sistema, així que no cal conservar forces ni velocitats. A més, donat que d'una etapa a una altra les posicions varien molt poc, aquestes només es conserven cada cert temps, depenent de la facilitat amb que canviï el sistema. A tall d'exemple, per a l'estudi d'una proteïna és suficient conservar un conjunt de coordenades cada 2ps.

Tot i que si conservéssim també les velocitats podríem reproduir qualsevol característica del sistema, en general aquestes es conserven amb molta menys freqüència (entre 10 i 100 vegades per a una trajectòria sencera) i es prefereix emmagatzemar valors processats durant la dinàmica com poden ser energies, pressió, temperatura, etc. que ens poden donar una bona idea de com ha anat evolucionant el sistema.

3.4.4 Determinació de l'Energia Lliure

L'energia lliure és segurament la propietat termodinàmica més important, ja que és la que en definitiva marca si un procés determinat, ja sigui una reacció química o la unió d'un lligand a un receptor, és o no viable.

Dues són les funcions que expressen l'energia lliure: A o funció de Helmholtz, vàlida per a un sistema en el que es mantinguin constants el nombre de partícules (n), el volum (V) i la temperatura (T) i G o funció de Gibbs, que s'usa més comunament, ja que és apropiada per sistemes a n , T i pressió (P) constant, condicions en les que s'acostumen a dur a terme la majoria d'experiments.

La determinació computacional de l'energia lliure resulta extremadament complicada per als sistemes que presenten una gran quantitat de mínims separats per barreres energètiques baixes. Això és així degut a la dificultat d'anar testant totes les configuracions que contribueixen significativament a l'energia lliure. Malauradament la majoria dels sistemes d'interès químic, bioquímic i farmacològic (macromolècules, líquids, etc.) cauen dins d'aquesta classificació.

Tal com veurem en el punt 3.4.4.1, el que sí resulta possible computacionalment és calcular el canvi d'energia lliure entre dos estats; això combinat amb la utilització de cicles termodinàmics ens permetrà calcular energies lliures tant importats com la de solvatació o les afinitats relatives d'unió entre una proteïna i dos lligands (punt 3.4.4.2). Quan, malgrat les dificultats, s'arriba a disposar d'un mostreig suficientment bo del sistema és possible estimar les contribucions entàlpiques i entròpiques a l'energia lliure. Per tal d'obtenir aquests termes cal recórrer a l'ús d'aproximacions que es descriuen en el punt 3.4.4.3.

3.4.4.1 Càlcul de diferències en l'Energia Lliure

S'han descrit fins a tres algorismes diferents per a calcular el canvi d'energia lliure: 1) pertorbació de l'energia lliure (*Free Energy Perturbation*; FEP) 2) integració termodinàmica (*Thermodynamic Integration*; TI) i 3) creixement lent (*Slow Growth*; SG).

Tots aquests mètodes es fonamenten en variar el sistema des d'un estat inicial definit per un hamiltonià (X i H_X , respectivament) a un de final definit per un altre

hamiltonià (Y i H_X , respectivament) mitjançant un seguit d'estats intermedis en que el sistema es defineix per una combinació lineal dels dos hamiltonians:

$$H_\lambda = (1 - \lambda) H_X + \lambda H_Y \quad (3.62)$$

on λ és un paràmetre d'acoblament que varia lentament des de 0 fins a 1 al llarg de la simulació, de manera que s'assegura una transició suau des de l'estat X fins a l'estat Y . No cal dir que aquest procés és totalment artificial, però donat que l'energia lliure té naturalesa de funció d'estat la seva variació depèn només dels estats inicial i final i no del camí seguit. En quant al mètode d'integració termodinàmica, aquest es fonamenta en què un sistema a volum constant el canvi en la funció de Helmholtz d'energia lliure ve donat per l'expressió:

$$\Delta A = \int_{\lambda=0}^{\lambda=1} \left\langle \frac{\partial H_\lambda}{\partial \lambda} \right\rangle_\lambda d\lambda \quad (3.63)$$

A la pràctica, aquesta integral es descompon en una sèrie de simulacions a valors discrets de λ i es calcula el promig de la variació a cada un d'aquests valors:

$$\Delta A = \sum_{\lambda=0}^{1-\Delta\lambda} \left[\int_{\lambda}^{\lambda+\Delta\lambda} \left\langle \frac{\partial H_{\lambda'}}{\partial \lambda'} \right\rangle_{\lambda'} d\lambda \right] \quad (3.64)$$

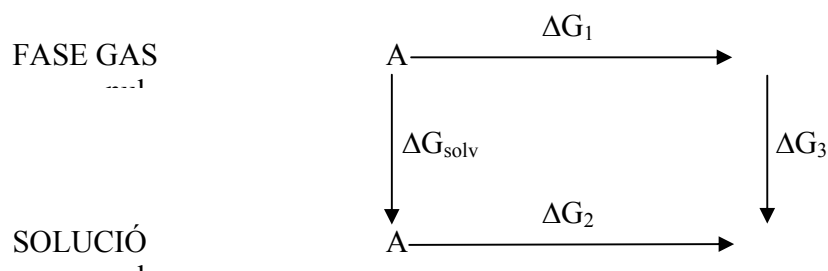
Per al cas més comú de treballar a pressió constant, podem obtenir l'expressió de la funció de Gibbs equivalent:

$$\Delta G = \sum_{\lambda=0}^{1-\Delta\lambda} \left[\int_{\lambda}^{\lambda+\Delta\lambda} \left\langle \frac{\partial E_{\lambda'}}{\partial \lambda'} \right\rangle_{\lambda'} d\lambda \right] \quad (3.65)$$

En aquesta expressió E_λ fa referència a l'energia del sistema quan λ pren un valor concret.

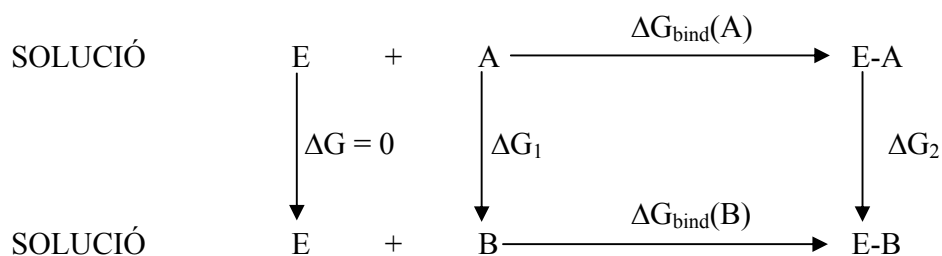
3.4.4.2 Ús de Cicles Termodinàmics

La tècnica descrita en el punt anterior, així com els altres mètodes per calcular diferències d'energia lliure, esdevenen realment útils quan les utilitzem combinades amb cicles termodinàmics dissenyats expressament. Si, per exemple, volem obtenir l'energia lliure de solvatació (ΔG_{solv}) d'un compost A, podem dibuixar el següent cicle:



Donat que l'energia lliure és una funció d'estat, el seu valor al voltant del cicle termodinàmic ha de ser 0, de manera que $\Delta G_1 + \Delta G_3 = \Delta G_{\text{solv}} + \Delta G_2$, com que $\Delta G_3 = 0$, obtenim que $\Delta G_{\text{solv}} = \Delta G_1 - \Delta G_2$. Així doncs, l'energia lliure de solvatació d'un compost es pot obtenir com la diferència d'energia lliure d'anul·lar la molècula d'interès en fase gas i solució. Les tècniques de perturbació de l'energia lliure ens permeten obtenir ΔG_1 i ΔG_2 i per tant ΔG_{solv} .

Un altre cicle termodinàmic molt més emprat en el camp del disseny de fàrmacs per comparar les capacitats d'unió (inhibició) de dues molècules (A i B) que s'uneixen de forma no covalent al centre actiu d'un enzim (E) és:



Com en el cas anterior: $\Delta G_{\text{bind}}(A) + \Delta G_2 = \Delta G_{\text{bind}}(B) + \Delta G_1$, de manera que podem conèixer la diferència d'energia lliure d'unió entre els compostos A i B ($\Delta\Delta G_{\text{bind}}$) pertorbant A en B tant quan està lliure en solució com quan està unit a l'enzim i restant aquestes quantitats: $\Delta\Delta G_{\text{bind}} = \Delta G_1 - \Delta G_2$. Donat que típicament existeix una relació directa entre capacitat d'unió i capacitat d'inhibició, l'ús d'aquest cicle termodinàmic combinat amb les tècniques de pertorbació de l'energia lliure ens permeten predir quantitativament quin dels dos compostos és l'inhibidor més actiu.

3.4.4.3 Càlcul d'Entropies

Tal com s'ha comentat anteriorment, una de les principals dificultats a l'hora de calcular l'entropia d'un sistema d'interès bioquímic o farmacològic és disposar d'un mostreig suficientment bo del seu espai conformacional. Si aquest problema ha estat resolt, es disposa de diferents mètodes que permeten obtenir un estimat d'aquesta propietat.

El mètode més emprat per calcular la contribució entròpica en sistemes de pocs àtoms consisteix en tractar tots els modes vibracionals (excepte rotació i translació) com oscil·ladors harmònics i emprar aquestes freqüències vibracionals per fer un anàlisi termoquímic clàssic [41]. El problema d'aquest mètode és que només pot ser aplicat sobre mínims harmònics i, per tant, fer-lo servir per estudiar sistemes amb una gran quantitat de mínims pot ser compromès.

Un mètode alternatiu, proposat originalment per Karplus i Kushik [42] i desenvolupat per Schlitter [43], permet calcular les entropies absolutes a partir del conjunt de coordenades cartesianes generades al llarg d'una trajectòria de dinàmica molecular. D'acord amb aquest mètode, per a un sistema quantificat en estats discrets $|n\rangle$, $n=0,1,2,\dots$, l'energia de cada un dels quals és ε_n , es pot escriure la seva funció de partició per a un conjunt canònic a temperatura T com:

$$Z = \sum_n \exp(-\beta\varepsilon_n) \quad (3.66)$$

on k_B és la constant de Boltzmann. Per simplificar introduïm el paràmetre β ($\beta=1/k_B T$) i podem expressar la probabilitat de trobar el sistema en un estat n segons l'equació 3.67, i obtenir l'entropia d'acord amb l'equació 3.68:

$$p_n = \frac{\exp(-\beta \epsilon_n)}{Z} \quad (3.67)$$

$$S = -k_B \sum_n p_n \ln p_n \quad (3.68)$$

Per a un sistema de massa m que es mou en una única dimensió x , que el seu promig $\langle x \rangle$ és 0 i que la seva variança $\langle x^2 \rangle$ pot ser expressada en funció dels valors esperats corresponents a cada estat (equacions 3.69 i 3.70), es pot demostrar que el valor màxim d'entropia es troba quan es compleix la condició 3.71, és a dir, que per a cada estat l'energia i la variança posicional que li correspon són proporcionals.

$$x_n^2 = \langle n | x^2 | n \rangle \quad (3.69)$$

$$\langle x^2 \rangle = \sum_n p_n \langle n | x^2 | n \rangle \quad (3.70)$$

$$\epsilon_n \propto \langle n | x^2 | n \rangle \quad (3.71)$$

Un sistema per al que es compleixen les condicions imposades més amunt és l'oscil·lador harmònic. L'entropia de l'oscil·lador harmònic monodimensional (S_{HO}) és un límit màxim del valor de l'entropia real del sistema i es pot expressar analíticament com:

$$S \leq S_{HO} = \frac{k_B \alpha}{e^\alpha - 1} - k_B \ln[1 - e^{-\alpha}] \quad (3.72)$$

Donat que la freqüència depèn de la variança mecano-quàntica, cal connectar-la a la variança clàssica, $\langle x^2 \rangle_C$, que és el que es pot mesurar en simulacions clàssiques. Per fer-ho s'utilitza l'expressió 3.73 on ω és la freqüència de l'oscil·lador, sent una bona aproximació per a qualsevol valor de $\langle x^2 \rangle_C$ excepte per a les més altes freqüències, que de totes maneres contribueixen molt poc a l'entropia.

$$m\omega^2 \langle x^2 \rangle_C = k_B T \quad (3.73)$$

Abans de generalitzar el mètode a un sistema amb molts graus de llibertat, s'introdueix una aproximació al càlcul de l'entropia S' (equació 3.74), per tal d'obtenir una expressió més senzilla i pràctica d'usar que 3.72:

$$S \leq S_{HO} < S' = \frac{1}{2} k_B \ln \left(1 + \frac{e^2}{\alpha^2} \right) = \frac{1}{2} k_B \ln \left(\frac{k_B T e^2}{\hbar^2} m \langle x^2 \rangle_C \right) \quad (3.74)$$

La generalització a molts graus de llibertat es basa en estimar l'entropia a partir de la matriu de covariances. Per un sistema de N àtoms situats en les coordenades x_1, x_2, \dots, x_{3N} , aquesta és una matriu de $3N \times 3N$ elements, cada un dels quals es defineix com:

$$\sigma_{ij} = \langle (x_i - \langle x_i \rangle) (x_j - \langle x_j \rangle) \rangle \quad (3.75)$$

La matriu de covariança també pot ser expressada amb coordenades ponderades per la massa de l'àtom que li correspon: $x_i' = x_i m_i^{1/2}$ convertint-se en:

$$\sigma' = M^{1/2} \sigma M^{1/2} = M \sigma \quad (3.76)$$

on $M = M^{1/2} M^{1/2}$ és la matriu de masses de rang $3N$ en la que els elements de la diagonal conté la massa dels àtoms i la resta d'elements són nuls.

La matriu de covariança que inclou les masses atòmiques (σ') pot ser diagonalitzada de manera que l'entropia pot ser calculada mitjançant l'equació 3.74 i els elements de la diagonal, que són les covariances $\langle q_{ii}^2 \rangle_C$:

$$S < S' = \frac{1}{2} k_B \sum_{i=1}^{3N} \ln \left(\frac{k_B T e^2}{\hbar^2} \langle q_{ii}^2 \rangle_C \right) = \frac{1}{2} k_B \ln \left[\sum_{i=1}^{3N} \left(\frac{k_B T e^2}{\hbar^2} \langle q_{ii}^2 \rangle_C \right) \right] \quad (3.77)$$

Mitjançant una darrera transformació podem escriure l'equació 3.77 com:

$$S < S' = \frac{1}{2} k_B \ln \det \left(1 + \frac{k_B T e^2}{\hbar^2} M \sigma \right) \quad (3.78)$$

Mitjançant aquesta nova expressió, es pot calcular l'entropia simplement diagonalitzant la matriu de covariança a partir d'una trajectòria de dinàmica molecular o un conjunt de coordenades obtingudes mitjançant la tècnica de Monte Carlo. Generalment s'eliminen els graus de llibertat de rotació i translació abans de generar la matriu de covariança per tal que l'entropia convergeixi més ràpidament.

3.4.5 Monte Carlo

Com la dinàmica o la mecànica molecular, per tal de poder usar la tècnica de Monte Carlo abans cal definir un sistema molecular amb un conjunt de paràmetres (*force field*) que el descriguin correctament. En aquesta tècnica no minimitzem ni veiem l'evolució temporal del sistema, sinó que simplement generem multitud de conformacions a l'atzar i construïm amb elles una distribució de Boltzmann. Es tracta, per tant, d'una visió estàtica però estadística del sistema, que ens ha de permetre explorar les zones de més baixa energia de l'hipersuperfície de potencial del sistema d'interès.

3.4.5.1 Algoritme de Metropolis

Tot i que existeixen altres algorismes que permeten obtenir un conjunt de Boltzmann de configuracions d'un sistema, en realitat quan es parla del mètode de Monte Carlo quasi sempre es fa referència al mètode tal i com va ser implementat per Metropolis i col·laboradors [44].

L'algoritme genera el que es coneix com una cadena d'estats de Markov. Això implica que: 1) cada configuració generada depèn només de la immediatament anterior i no de les precedents a aquesta i, 2) cada estat generat pertany a un conjunt finit de resultats possibles. Per fer això es segueixen els passos següents (veure figura 14):

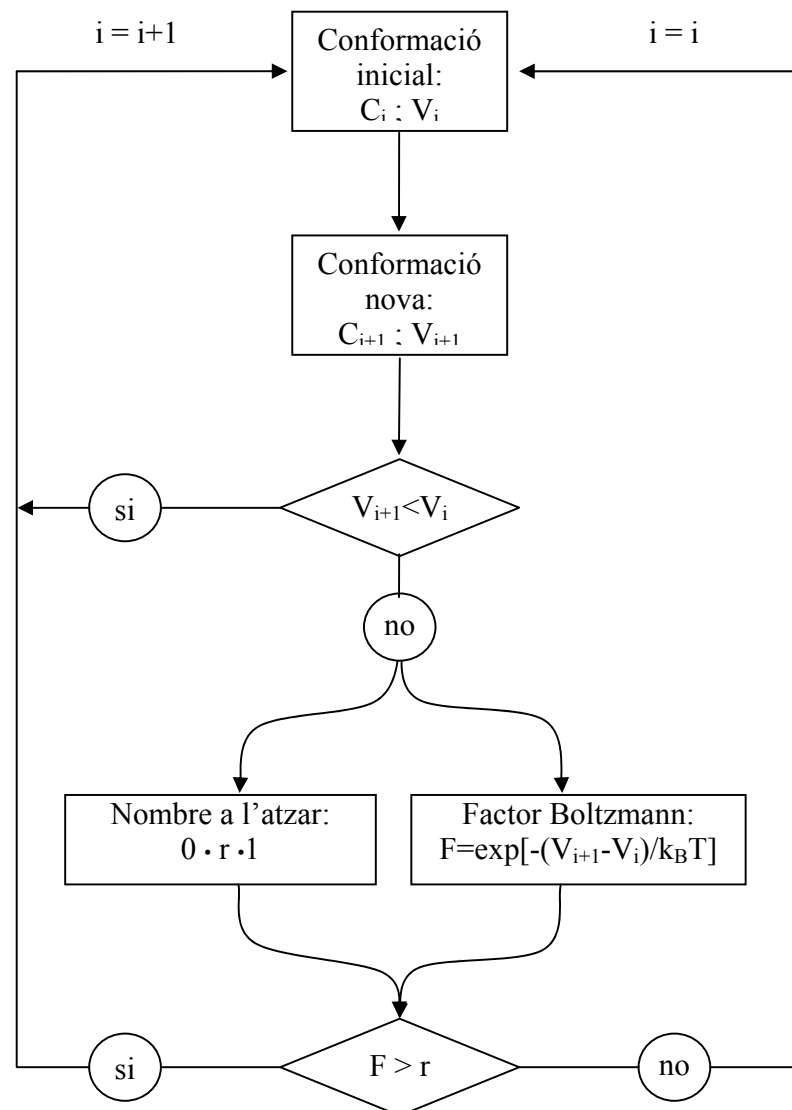


Figura 14. Diagrama de flux del mètode de Monte Carlo

Aquestes etapes del mètode de Monte Carlo representats en la figura 14 es resumeixen en:

- A partir d'una conformació en generem una altra movent aleatòriament les posicions dels àtoms que formen el sistema o bé modificant a l'atzar els valors de variables internes (angles, dihedres, distàncies d'enllaç, volum del sistema,...)
- Calculem el valor de l'energia potencial (V) de la nova configuració del sistema que, tal com s'ha vist, és funció de les seves coordenades (C).
Si l'energia de la nova conformació és menor que la de la conformació de partida, aquesta s'accepta i es torna a iniciar el procés. Si no és així, es calcula el factor de Boltzmann $F = \exp(-\Delta V/k_B T)$. Si F és major que un nombre entre 0 i 1 generat aleatòriament, aleshores la nova configuració s'accepta; en cas contrari, es descarta la nova configuració i s'accepta novament l'original, a partir de la qual es continuarà la cadena de Markov.

Per tal d'assegurar-nos un correcte mostreig de l'espai conformacional, la magnitud de les modificacions geomètriques que apliquem al sistema han de ser tals que s'acceptin al voltant del 40% de les estructures generades. La raó d'això és que si l'amplitud dels moviments fos molt petita s'acceptarien moltes conformacions, però s'exploraria una petita part de l'espai conformacional i si, contràriament, es realitzessin modificacions molt grans, aleshores el grau d'acceptació seria molt baix i el procediment seria molt poc eficient.

Cada una de les conformacions acceptades és emmagatzemada per tal de poder fer els anàlisis posteriors, que poden ser de tipus estructural, energètic o poblacional. Per a sistemes en que hi ha coexistència de dos o més tipus d'estructures ben diferenciades (diferents confòrmers, per exemple), si el mostreig és complet, es podrà derivar els valors d'energia lliure del procés d'interconversió a partir de les poblacions respectives.

3.4.5.2 Comparació entre Monte Carlo i Dinàmica Molecular

Si bé en principi tant Monte Carlo com Dinàmica Molecular permeten generar conjunts de configuracions representatives d'un sistema de manera que podem explorar-ne l'hipersuperfície de potencial i extreure'n informació estructural i termodinàmica, ambdós mètodes presenten notables diferències. El més evident és que mentre que la dinàmica molecular és funció del temps i per tant donades les coordenades i velocitats en un instant concret podrem arribar a trobar les configuracions del sistema en qualsevol altre temps, en una simulació per Monte Carlo cada conformació depèn exclusivament de l'anterior. Aquesta diferència ens obligarà a escollir la dinàmica molecular quan vulguem calcular propietats que siguin dependents del temps com ara la difusió.

Com a conseqüència de la dependència temporal, en la tècnica de dinàmica molecular sorgeixen fluctuacions de pressió i temperatura que cal controlar. En el cas de Monte Carlo es guanya en simplicitat ja que aquests fenòmens no tenen lloc.

Una altra diferència entre ambdós mètodes és que en dinàmica molecular es considera tant l'energia potencial com la cinètica mentre que Monte Carlo només inclou la primera.

A nivell pràctic, Monte Carlo permet un control més elevat del sistema ja que pot treballar amb coordenades internes, mentre que la dinàmica molecular simula sistemes més lliures ja que generalment treballa amb coordenades cartesianes.

Els dos mètodes també poden divergir en la respectiva capacitat per explorar l'espai conformacional. El mètode de Monte Carlo és molt poc eficient testant conformacions de molècules grans, donat que cal fer canvis estructurals molt petits per tal de tenir un grau d'acceptació raonable. Per contra, la capacitat d'aquest mètode per fer moviments sense sentit físic li pot permetre salvar barreres d'energia molt grans que una dinàmica molecular normal no podria superar si no fos a temps molt llargs. L'estudi dels isòmers *cis*- i *trans*- del 2-butè seria un exemple prototípic.

De fet les dues tècniques en certa manera es complementen i si bé en general caldrà triar-ne una o l'altra, hi ha casos en que l'ús combinat de les dues tècniques pot donar resultats òptims.

3.4.6 Docking

La tècnica de *Molecular Docking* (encaixat molecular) té com a finalitat predir el mode i l'afinitat d'unió entre dues molècules, generalment complementàries com és el cas d'un enzim i el seu substrate, un antigen i el seu anticòs, una hormona i el seu receptor específic o entre una proteïna i un lligand petit com podria ser un inhibidor (fàrmac). És aquesta una tècnica molt utilitzada en el disseny de fàrmacs per la seva rapidesa generant hipotètics modes d'unió i avaluant-ne la bondat dels mateixos [64].

Per al cas d'un lligand rígid encaixant-se a un receptor també rígid, hi ha 3 graus de llibertat de translació i 3 de rotació, i generalment no és pas fàcil trobar el mode d'unió adoptat. De totes maneres, poden fer-se prediccions manuals del mode d'unió molt raonables si es coneix com s'uneix una altra molècula molt semblant o si hi ha gran complementaritat estèrica entre lligand i receptor. No cal dir que els algorismes de *docking* automàtics són molt més sistemàtics i fan una exploració més completa de l'espai accessible, raó per la qual és generalment desaconsellable fer un model d'unió basant-se només en un modelat manual.

Els programes de *docking* es poden dividir en dues parts ben diferenciades: generació del model (punt 3.4.6.1) i avaluació del mateix (punt 3.4.6.2). Donada l'abundància i diversitat de programes existents, aquí es descriurà només el funcionament del programa AUTODOCK [45], que és el que ha estat utilitzat en aquesta tesi.

Si a més de rotació i translació volem considerar la llibertat conformacional del lligand i/o del receptor, el problema adquireix dimensions desproporcionades. Actualment molts programes de *docking* poden tractar la llibertat conformacional del lligand (punt 3.4.6.3), però incloure graus de llibertat al receptor és encara un objectiu a llarg termini, tot i que s'han descrit mètodes per incloure-la implícitament [46].

3.4.6.1 Posicionament del lligand

Si es vol explorar sistemàticament la unió d'un lligand al centre actiu d'una proteïna (fent totes les rotacions i translacions que calgui) la quantitat d'orientacions a considerar és enorme i fa que, per a cada lligand, el temps de càlcul sigui considerable, impossibilitant el crivellat (*screening*) d'una gran quantitat de lligands tal i com seria desitjable. Per resoldre aquest problema, el programa AUTODOCK té un algoritme basat en la complementarietat de forma entre lligand i receptor que genera moltes menys orientacions de les teòricament possibles però amb més sentit.

Per tal de poder generar un model d'unió amb AUTODOCK, abans que res cal descriure el centre d'ancoratge dels lligands amb un conjunt d'esferes que omplim l'espai que no és ocupat per la proteïna. Això s'aconsegueix mitjançant l'ús del programa SPHGEN [47] que, a partir de la superfície de la proteïna, genera conjunts (*clusters*) d'esferes de diferents radis que es superposen entre elles, formant imatges negatives de cada un dels solcs o cavitats de la mateixa.

El conjunt d'esferes associat al centre actiu (que coneixerem per dades experimentals o per ser el *cluster* més poblat) serà utilitzat pel programa AUTODOCK per col·locar els lligands. El nombre ideal d'esferes està entre 20 i 50, depenent de la mida del *locus* d'unió. El procediment en el que es basa consisteix en aparellar àtoms del lligand amb els centres d'algunes de les esferes que defineixen el centre actiu; aquests parells són usats per generar una orientació del lligand en la que aquests parells àtom-esfera es superposin el millor possible. Com a mínim hi haurà d'haver 3 aparellaments per a cada superposició (*match*) i es fa un gran nombre de combinacions de manera que es generen moltes orientacions que posteriorment seran avaluades [64].

3.4.6.2 Avaluació del Model

El programa AUTODOCK pot avaluar de fins a tres maneres independents quan bona és la interacció entre lligand i receptor per a cada estructura generada: basant-se en una classificació química, de contacte o energètica. Aquestes classificacions són independents i es poden combinar entre elles per obtenir informació útil de les possibles interaccions.

Per a cada criteri d'avaluació necessitarem una xarxa tridimensional (*grid*) de punts que conté informació sobre l'entorn: en cada punt de la *grid* està precalculat els

valors que prenen el potencial electrostàtic generat per la proteïna, els termes de van der Waals corresponents al receptor, el tipus d'entorn químic, etc.

De fet, doncs, el valor de la interacció en realitat no s'obté del càlcul directe entre lligand i proteïna sinó entre el lligand i la *grid*. Això obliga a adoptar algunes aproximacions però que fa que cada càlcul sigui molt més ràpid. A més, AUTODOCK permet utilitzar un filtre de contacte (*Bump Filter*) que exclou aquelles orientacions que presenten col·lapses estèrics abans de ser avaluades.

3.4.6.2.1 Filtre de Contacte

Una part de les estructures generades poden presentar una significant superposició del lligand amb els àtoms de la proteïna. Per aquesta raó és important disposar d'un filtre que elimini aquestes orientacions sense sentit abans d'avaluar-les, evitant així la part més costosa del càlcul. Per defecte AUTODOCK considera que existeix un col·lapse estèric (*bump*) quan la distància entre un àtom del lligand i un del receptor és menor que 0.75 vegades la suma dels radis de van der Waals d'aquests àtoms. L'usuari pot fixar el nombre de *bumps* permesos per a cada orientació; un model d'interacció serà descartat quan tingui un nombre de col·lapses superior a aquest valor.

3.4.6.2.2 Classificació per contacte

La funció d'interacció més senzilla que permet classificar de més a menys òptimes les orientacions generades és la de contacte: per a cada sobreposició que es doni d'un àtom del lligand amb l'espai ocupat per la proteïna es penalitza el posicionament amb un valor fix. Aquesta classificació permet només avaluar la complementarietat de forma entre lligand i receptor, però és molt ràpida i pot ser útil per estudiar interaccions no polars.

3.4.6.2.3 Classificació per Energia

La classificació més rigorosa i utilitzada que permet fer AUTODOCK es basa en criteris energètics. Per fer això, primerament cal assignar paràmetres de càrrega electrostàtica i termes de van der Waals a cada àtom de lligand i proteïna. El *force field*

emprat és una versió una mica simplificada d'AMBER, podent optar entre models *all atom* o *united atom*.

Per assignar les càrregues dels lligands s'utilitza el mètode de Gasteiger [48] que genera càrregues aproximades emprant criteris simples d'electronegativitat o de potencial d'ionització de forma molt ràpida.

Una vegada tenim tots aquests termes, per a cada conformació podem calcular l'energia del sistema i fins hi tot optimitzar la geometria. Convé remarcar però, que tal com s'ha esmentat anteriorment, la proteïna no està representada explícitament sinó per una xarxa tridimensional de punts en la que s'extrapolen les propietats dels àtoms.

3.4.6.2.4 Classificació Química

Una altra possibilitat que ofereix AUTODOCK és avaluar cada interacció en funció de la complementarietat química en els contactes entre lligand i receptor. Per fer això cal primer definir químicament els àtoms, per exemple: apolar, polar, donador de pont d'hidrogen, acceptor de pont d'hidrogen, etc. Segon cal disposar de criteris de compatibilitat entre parells d'interaccions, per exemple: donador de pont d'hidrogen amb acceptor de pont d'hidrogen és compatible, polar amb apolar no és compatible.

Aquests senzills criteris no poden proporcionar informació quantitativa però si qualitativa i poden, a més, complementar les dades provinents dels altres mètodes d'avaluació.

3.4.6.3 Llibertat Conformacional del Lligand

Donat que els lligands no acostumen a ser gaire grans i que sovint una part importat de la seva estructura és rígida, és possible incloure els seus graus de llibertat interns mitjançant diverses metodologies.

Una manera molt senzilla de tractar la variabilitat conformacional de les molècules que s'han d'unir és utilitzar algun altre programa capaç d'explorar l'espai conformacional per generar un conjunt de conformacions per a cada lligand i fer *docking* de cada una de les estructures com si en realitat es tractessin de diferents molècules. Aquest mètode té com a principals inconvenients que augmenta molt els requeriments de disc (cal emmagatzemar moltes vegades cada compost), no permet ajustar la conformació a la forma del receptor i no és un procediment especialment eficient.

Una alternativa és variar la conformació de la molècula durant el procés de *docking*. El programa AUTODOCK inclou flexibilitat conformacional del lligand a nivell

de torsions pròpies. A continuació es descriuen els passos que segueix el programa per incloure la rotació dels dihedres:

- Identificació de fragments rígids: tots aquells enllaços que formen part d'anells o que tenen caràcter de doble enllaç (total o parcial) són considerats rígids; la resta són flexibles. Es divideix la molècula en fragments que no tenen cap enllaç flexible.
- El fragment rígid que agrupa més àtoms és considerat el "punt d'ancoratge" que s'utilitzarà per fer un *docking* normal, com si en realitat es tractés d'una molècula completa.
- Per a cada orientació que pren aquest fragment se li uneixen els fragments rígids en etapes successives de nivells de veïnatge. En cada etapa s'exploren els dihedres dels enllaços flexibles que uneixen els fragments rígids, generant molts possibles conformers dels quals només alguns seran seleccionats per la següent fase. La selecció de les conformacions es basa en dos criteris: que tinguin la millor puntuació possible i que representin el màxim de diversitat conformacional.
- Quan s'arriba a l'últim nivell de veïnatge es fa una classificació definitiva dels conformers. Si s'està estudiant més d'una molècula de lligand, cada conformer competeix independentment en la classificació final.

Per últim assenyalarem la importància que tenen els mètodes de com una eina d'estudi directe del mecanisme d'inhibició fàrmac-receptor que a més permeten elucidar les característiques del farmacòfor per al futur disseny de nous i més potents fàrmacs.

3.5 Tractament de l'Efecte del Solvent

La majoria de processos bioquímics que es simulen al laboratori o que tenen en investigació industrial tenen lloc en fases condensades, on el sistema d'interès està rodejat d'un gran nombre de molècules de solvent amb les que interacciona. I que poden modificar significativament les propietats del solut en aquest cas l'interacció fàrmac-receptor.

Hi ha una gran varietat de mètodes computacionals disponibles per avaluar l'efecte del solvent sobre el solut, cada un d'ells amb els seus propis defectes i virtuts [64]. Aquesta diversitat metodològica permet disposar d'un model apropiat per a cada sistema [49].

3.5.1 L'efecte del solvent

A nivell microscòpic podem veure la relació entre solut i solvent com un conjunt dinàmic d'interaccions no enllaçants entre una molècula (que anomenarem solut) i un conjunt pràcticament infinit d'altres molècules, generalment idèntiques, que l'envolten (solvent). Cal aclarir que la distinció entre solut i solvent és fins a cert punt arbitrària, ja que cada molècula del solvent pot ser estudiada individualment com a solut.

La magnitud fonamental quan estudiem un fenomen de solvatació és l'energia lliure (ΔG_{sol}) implicada en aquest procés. Seguint la definició de Ben-Naim [50] aquesta energia lliure és el treball reversible necessari per transferir una molècula de solut des d'una posició fixa en fase gas a una posició fixa de l'interior d'un solvent a temperatura, pressió i composició constants.

Conceptualment la solvatació pot ser dividida en tres etapes ben diferenciades: cavitació, interaccions de dispersió-repulsió i un terme electrostàtic. El primer pas és la formació d'una cavitat en l'interior del solvent on pugui ser encabit el solut i a l'energia lliure implicada en aquest procés s'anomena de cavitació (ΔG_{cav}).

Donat que per crear aquesta cavitat cal trencar interaccions i desestructurar el solvent, aquest serà sempre un procés penalitzat energèticament, tot i que la magnitud d'aquesta penalització variarà molt en funció del tipus de solvent: com més robustes siguin les interaccions entre les molècules del dissolvent, major serà el cost de crear una cavitat en el seu interior.

El segon terme en que dividim el procés de solvatació és aquell en que entren en joc les interaccions de dispersió-repulsió (forces de van der Waals) entre solut i solvent. Aquestes forces contribueixen favorablement a la solvatació, de manera que l'energia

lliure implicada (ΔG_{vw}) sempre és negativa. La magnitud d'aquest terme també pot variar en funció del tipus de dissolvent, a la vegada que la natura del solut juga un paper important. Els termes de cavitació i de van der Waals poden ser considerats conjuntament amb el nom de terme estèric o no-electrostàtic i típicament existeix una compensació entre ambdós. Aquesta compensació acostuma a saldar-se favorablement per al cas de soluts dissolts en medis apolars, ja que el terme favorable de dispersió-repulsió és per norma general major en termes absoluts que el que de cavitació. Per contra, en el cas de soluts en dissolvents polars, βG_{vw} no pot arribar a compensar βG_{cav} que és força gran en aquest tipus de medis.

El darrer dels termes implicats en la solvatació és l'electrostàtic (βG_{ele}), que considera el treball implicat en crear la distribució de càrrega del solut en la dissolució. Aquest treball pot, a la seva vegada, subdividir-se en dues parts: 1) generar la distribució de càrrega que el solut té en fase gas a l'interior del solvent i 2) polaritzar aquesta distribució de càrrega per acomodar-la al nou medi. Cal tenir present que l'energia implicada no es deu tan sols a les interaccions electrostàtiques solut-solvent, sinó també al treball que cal realitzar per redistribuir les càrregues del solvent en resposta a l'aparició del solut (generant el que s'anomena camp de reacció), i per distorsionar la distribució electrònica en fase gas del solut. Per soluts i solvents apolars la magnitud d'aquest terme serà més aviat petita ja que la redistribució electrònica en aquest tipus de molècules és molt petita i no es generen càrregues netes importants.

Per contra en soluts i solvents polars aquest és el terme predominant ja que es generen fortes interaccions electrostàtiques. En definitiva, l'energia lliure de solvatació pot expressar-se d'aquesta manera:

$$\Delta G_{sol} = \Delta G_{ele} + \Delta G_{vw} + \Delta G_{cav} \quad (3.79)$$

El conjunt d'aquestes tres interaccions pot provocar importants modificacions en les propietats dels soluts a molts nivells:

- Propietats electròniques: el pas d'una molècula en fase gas a dissolució causa una redistribució electrònica que altera les propietats químiques d'aquesta

molècula, com poden ser el moment dipolar, el potencial electrostàtic molecular, la densitat de càrrega o el volum molecular.

- Propietats estructurals: distàncies d'enllaç, angles i torsions poden variar sensiblement per efecte del solvent, donant lloc a diferents preferències conformacionals en funció del medi.
- Propietats espectroscòpiques: quan la solvatació estabilitza de manera diferent una molècula en el seu estat fonamental que en el seu estat excitat, pot provocar modificacions en les transicions electròniques i per tant en el seu espectre d'absorció, així com en els seus modes vibracionals.
- Propietats àcid/base: les propietats àcid/base intrínseques del solut es veuen modificades en solució en funció de la capacitat del solvent per acomodar espècies neutres o carregades.
- Propietats tautomèriques: com en el cas de les conformacions, la solvatació pot afavorir uns tautòmers més que a altres, de manera que les poblacions respectives poden veure's afectades.
- Reactivitat: la reactivitat dels soluts pot veure's modificada per tres factors diferents: 1) canvi en les propietats intrínseques dels reactius (distribució electrònica, per exemple), 2) diferent grau d'estabilització de reactius, productes i estat de transició i 3) efecte de la fricció.
- Associació molecular: mentre que els dissolvents apolars afavoreixen la formació d'interaccions polars entre soluts, els dissolvents polars tendeixen a trencar aquestes interaccions solut-solut per tal de formar-ne solut-solvent. Variacions en la mida de la cavitat i en l'entropia de dimerització poden jugar també un paper en l'associació de soluts en el si del solvent.

Els sistemes bioquímics tals com pèptids, proteïnes, àcids nucleics o membranes lipídiques són casos en els que és especialment important tenir en compte aquests efectes, ja que l'aigua, el dissolvent per excel·lència en els éssers vius, genera un entorn particularment pertorbador de les propietats dels soluts.

3.5.2 Models discrets

Sens dubte, la millor i més òbvia manera d'incloure l'efecte del solvent seria simplement considerar explícitament cada una de les molècules de dissolvent. No obstant, això és impossible d'abordar mitjançant mètodes mecano-quàntics. Fins i tot, si només volguéssim incloure una part del solvent (la primera capa de solvatació, per exemple), difícilment una sola geometria seria representativa del sistema ja que aquests tipus de supramolècules presenten moltes ordenacions estructurals possibles.

L'alternativa viable prové dels mètodes QM/MM [51], que permeten dividir el sistema en dues parts, una que serà tractada quànticament (tot o part del solut i, opcionalment, una petita part del solvent) i una altra de manera clàssica. Malgrat els avantatges que aporten els mètodes QM/MM, el seu cost computacional és encara enorme i fa que només siguin aplicats allà on els mètodes basats purament en formulismes clàssics no poden arribar, com és l'estudi de reaccions químiques. Tot i l'interès d'aquests mètodes, i donat que no han estat emprats en aquesta tesi, no seran descrits.

Els mètodes purament clàssics, descrits en el capítol anterior, són els mètodes més emprats quan es tracta de representar explícitament el solvent. Per fer això només cal definir el solut i rodejar-lo de molècules de dissolvent fins a ocupar un volum que pugui ser replicat en totes les direccions de l'espai i aplicar condicions periòdiques de contorn a aquest volum. Una forma més aproximada de fer un tractament explícit del solvent és el de situar una esfera d'aigües (*cap*) al voltant del solut. Aquesta tècnica té l'inconvenient que part del solvent està molt a prop del buit, i no està per tant ben representat, però pot ser de molta utilitat quan s'està interessat en estudiar una petita part d'un solut molt gran (típicament el centre actiu d'un enzim) ja que ens permet incloure l'efecte del solvent a un preu computacional molt menor que les condicions periòdiques de contorn.

L'ús d'aquests formalismes clàssics ha reportat molts bons resultats, però cal no perdre de vista que venen limitats per la bondat de la parametrització del camp de forces que s'estigui fent servir. El costat positiu d'un tractament clàssic del sistema és que la rapidesa del càlcul ens permet realitzar estudis de dinàmica molecular o Monte Carlo incloent l'efecte del solvent.

3.5.3 Models continus

En aquests models com per exemple Tomasi & Persico [52]; Cramer & Truhlar [53]; i Orozco *et al.* [54]; entre altres el dissolvent es considera com un medi continu dielèctric i polaritzable a dins del qual s'ha generat una cavitat per encabir-hi el solut. El continu es caracteritza per la constant dielèctrica del solvent (ϵ_s) al qual representa. La distribució de càrrega del solut polaritza el dielèctric, creant així interaccions de tipus electrostàtic (ΔG_{ele}) entre el solut i el dielèctric. Tanmateix s'han desenvolupat diversos mètodes, tots basats en parametritzacions, que permeten obtenir una estimació més o menys acurada dels termes estèrics ($\Delta G_{\text{vw}} + \Delta G_{\text{cav}}$).

3.6 Càlcul de les propietats moleculars

En el capítol 3.2 hem vist que els mètodes mecanoquàntics permetien obtenir la funció d'ona aproximada, aquest càlcul és molt important ja que a partir d'ella podran parametritzar-se les diferents propietats moleculars que com a objectiu d'aquesta tesi han servit per relacionar-les amb l'activitat farmacològica per tal d'establir quines propietats són les que afavoreixen l'activitat i en definitiva diferencien als compostos estudiats en actius i pocs actius. Entre aquestes propietats moleculars destaquen:

- 3.6.1 Geometria molecular optimitzada.
- 3.6.2 Moment dipolar.
- 3.6.3 Potencial d'ionització.
- 3.6.4 Anàlisi conformacional.
- 3.6.5 Reactivitat molecular.
- 3.6.6 Densitat de probabilitat electrònica.
- 3.6.7 Potencial electrostàtic molecular.
- 3.6.8 Densitat electrònica en el punt crític.
- 3.6.9 Paràmetres termodinàmics

3.6.1 Geometria molecular optimitzada

La geometria molecular pot calcular-se mitjançant l'obtenció prèvia de la funció d'ona molecular per a diferents configuracions espaials dels nuclis de la molècula que corresponen a petites variacions entre distàncies, angles d'enllaç i angles diedres, obtenint aquella configuració espaiial dels nuclis per a la qual l'energia electrònica (E_e) és mínima. El tractament matemàtic seria similar al descrit en el capítol anterior sent en aquest cas E_e la funció a optimitzar.

La configuració amb menor energia electrònica és la geometria molecular optimitzada a la qual correspon una calor de formació ΔH_f mínima.

La calor de formació d'una molècula ΔH_f es pot calcular a partir de la diferència entre l'energia total d'una molècula E_T i la suma de les energies electròniques E_e dels àtoms per finalment sumar les calors experimentals dels àtoms que componen la molècula ΔH_f^a :

$$\Delta H_f = E_T - \sum_A E_e^A + \sum_A \Delta H_f^A \quad (3.80)$$

Sent E_T l'energia total de la molècula que es pot definir com la suma de l'energia electrònica i la de repulsió dels nuclis (cores) dels àtoms.

3.6.2 Moment dipolar

El valor del moment dipolar també pot obtenir-se a partir de la funció d'ona electrònica ψ mitjançant l'equació del moment dipolar elèctric:

$$p = \sum_i Q_i r_i \quad (3.81)$$

Sent r_i el vector definit des de l'origen fins a la càrrega Q_i , introduint l'expressió de ψ funció d'ona amb el seu respectiu operador ψ^* arribarem a obtenir l'equació 3.82 que defineix el càlcul del moment dipolar a partir de la funció d'ona molecular.

$$p = \int \psi^* \sum_i Q_i r_i \psi \, dr = \sum_i Q_i \int |\psi|^2 r_i \, dr \quad (3.82)$$

La suma s'estén per a tots els electrons i nuclis obtenint el valor del moment dipolar p per a la geometria d'equilibri.

3.6.3 Potencial d'ionització

Es correspon amb el valor energètic de l'orbital ocupat de major energia (HOMO), sent l'energia requerida per arrencar l'electró de la capa electrònica més externa.

3.6.4 Anàlisi conformacional

A partir de la geometria d'equilibri variant suaument la configuració espacial dels nuclis a través de les lliures rotacions permeses en enllaços simple s'obtenen diferents energies electròniques i calors de formació. Representant el valor de la calor de formació (o entalpia de formació) front a cada rotació d'un determinat enllaç s'obtindrà la representació dels mapes de barreres conformacionals, a partir de l'estudi dels quals poden determinar-se les configuracions més estables i per tant possiblement bioactives.

3.6.5 Reactivitat molecular

La facilitat de què s'iniciï una reacció química ve condicionada per la probabilitat de trobar l'electró en una zona determinada i per l'energia d'aquest electró. L'orbital ocupat de major energia (HOMO) correspon a una regió de la molècula rica en electrons i per tant de caràcter nucleòfil.

De la mateixa forma, l'orbital desocupat de menor energia es correspon amb una regió pobre en electrons i per tant és de caràcter electròfil, és a dir té apetència perquè els electrons l'ocupin. La densitat electrònica d'aquests orbitals així com els seus valors energètics són diferents per a cada molècula i ens ajudaran a localitzar aquestes regions moleculars susceptibles de ser punt d'unió amb un receptor. Així segons el teorema de Koopman, el potencial d'ionització d'un compost, o energia necessària per arrancar-li un electró es correspon amb l'energia de l'HOMO, quan més petita sigui aquesta energia amb més facilitat el compost cedirà un electró. En quant a l'afinitat electrònica, o energia necessària per incorporar un electró a una molècula es correspon amb l'energia de l'orbital LUMO, quan més petita sigui aquesta energia amb més facilitat el compost captarà un electró.

Tanmateix la diferència de nivells energètics entre aquests orbitals (HOMO-LUMO) quantificarà la major o menor reactivitat molecular.

3.6.6 Densitat de probabilitat electrònica

Es calcula per integració de la funció d'ona electrònica Ψ_i sobre les coordenades de spin de tots els electrons de la molècula i sobre les coordenades espaials x, y, z de tots els electrons menys un i multiplicant el resultat obtingut pel número total d'electrons de la molècula [53]. De la mateixa forma, els mètodes mecanoquàntics permetran conèixer la distribució del potencial electrostàtic en l'entorn molecular PEM.

3.6.7 Potencial Electrostàtic Molecular

El potencial electrostàtic $V(\mathbf{r})$ generat per una molècula (MEP) en un punt \mathbf{r} de l'espai ve determinat per la distribució de la seva càrrega total que inclourà tant a nuclis com a electrons, sent \mathbf{r}_1 i \mathbf{R}_A els vectors de posició de l'electró 1 i del nucli A, tal i com es representa en l'equació 3.83:

$$V(\mathbf{r}) = \int d\mathbf{r}_1 \frac{\rho_T(\mathbf{r}_1)}{|\mathbf{r} - \mathbf{r}_1|} \quad (3.83)$$

El conjunt de nuclis que formen la molècula i fixats dins de l'aproximació de Born-Oppenheimer, són equivalents a una distribució discreta de càrrega.

Així mateix, si es coneix la funció d'ona electrònica de la molècula: Ψ_0 , la funció densitat de càrrega electrònica $\rho_{el}(\mathbf{r}_1)$ que vindrà determinada per $|\Psi_0|^2$, expressarà la distribució contínua dels electrons en l'espai. Per tant, la distribució de càrrega molecular total $\rho_T(\mathbf{r}_1, \mathbf{R}_A)$ pot expressar-se en unitats atòmiques com:

$$\rho_T(\mathbf{r}_1) = -\rho_{el}(\mathbf{r}_1) + \sum_A Z_A \delta(\mathbf{r}_1 - \mathbf{R}_A) \quad (3.84)$$

sent: Z_A el número atòmic de l'àtom A i $\delta(\mathbf{r}_1 - \mathbf{R}_A)$ la funció delta de Dirac.

Dins de l'aproximació SCF-RHF, la distribució de càrrega electrònica està relacionada amb el conjunt d'orbitals moleculars $\{\phi_i\}$ per l'equació 3.85:

$$\rho_{el}(\mathbf{r}_1) = 2 \sum_i \phi_i^*(\mathbf{r}_1) \phi_i(\mathbf{r}_1) \quad (3.85)$$

Si descrivim cada orbital molecular com una combinació lineal d'orbitals (μ, λ) , l'expressió de la densitat de càrrega electrònica adoptaria la forma de l'equació 3.86:

$$\rho_{el}(\mathbf{r}_1) = \sum_{\mu} \sum_{\lambda} P_{\mu\lambda} \phi_{\mu}^*(\mathbf{r}_1) \phi_{\lambda}(\mathbf{r}_1) \quad (3.86)$$

Per tant, el potencial electrostàtic molecular generat en un punt \mathbf{r} per la distribució de la càrrega molecular (equació 3.83) adquirirà ara la següent expressió:

$$V(\mathbf{r}) = \sum_{\mu} \sum_{\lambda} P_{\mu\lambda} \int d\mathbf{r}_1 \frac{\phi_{\mu}^*(\mathbf{r}_1) \cdot \phi_{\lambda}(\mathbf{r}_1)}{|\mathbf{r} - \mathbf{r}_1|} + \sum_{\mathbf{A}} \frac{Z_{\mathbf{A}}}{|\mathbf{r} - \mathbf{R}_{\mathbf{A}}|} \quad (3.87)$$

D'acord amb l'equació 3.87, el potencial electrostàtic molecular (MEP) queda definit per el resultat de l'operador monoelèctric $1/r$ i aquest el resultat és independent del mètode utilitzat per calcular la funció d'ona, així com del nivell d'aproximació seguit. Així mateix, representa el valor aproximat fins a primer ordre de pertorbació d'energia d'interacció de la molècula amb la unitat de càrrega positiva.

Per últim les característiques qualitatives dels mapes de PEM estan correctament reproduïdes i poden ser comparables tant si es determinen a partir de la funció d'ona semiempírica (AM1-MNDO-PM3) com *ab initio* reflectint ambdós mètodes les característiques topològiques de les superfícies d'isopotencial de l'entorn molecular [56-58].

Una variant del MEP és el MIP (Molecular Interaction Potential) [59] que té en compte la contribució de les forces de repulsió-dispersió a l'energia total d'interacció, de forma que proporciona una millor descripció de la reactivitat molecular a l'incorporar la interacció de van der Waals tal i com es recull en l'equació 3.88:

$$V_{\text{MIP}}(\mathbf{r}) = V_{\text{REP}}(\mathbf{r}) + V_{\text{DISP}}(\mathbf{r}) + V_{\text{MEP}}(\mathbf{r}) \quad (3.88)$$

Tant els mapes de MEP com de MIP constitueixen importants eines de treball que permeten localitzar aquelles regions de la molècula més susceptibles d'una possible interacció.

3.6.8 Densitat en el punt crític

El punt crític d'enllaç és la mínima densitat electrònica al llarg de la coordenada que defineix l'enllaç químic entre dos àtoms [60], en general quan major és la densitat electrònica en el punt crític l'enllaç serà més làbil, ja que l'electró no està compartit sinó s'apropa més a un dels àtoms, generalment a aquell que sigui el més electronegatiu., per exemple l'enllaç O-H; en cas contrari quan la densitat electrònica en el punt crític és baixa, aleshores ens indica que l'enllaç és més estable, per exemple un enllaç covalent C-C.

Matemàticament, el punt crític es fonamenta en considerar a la densitat de càrrega electrònica ρ com a un camp escalar on cada punt està definit pel vector gradient $\rho(\mathbf{r})$.

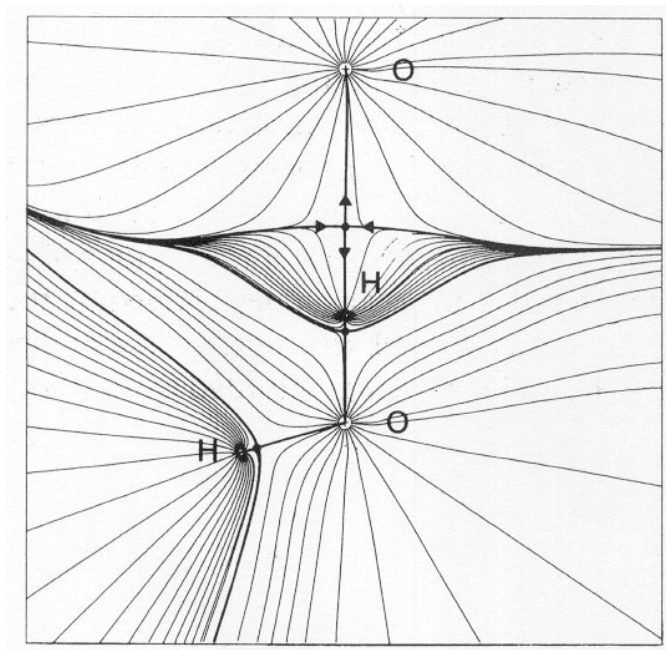


Figura 15. Representació del mapa de densitat electrònica en el punt crític per un dímer format per dues molècules d'aigua. Cada línia representa la trajectòria de $\nabla\rho$.

Considerant un enllaç químic com un tensor, aquest tindrà unes energies cinètica i potencial característiques en funció de les interaccions existents entre els àtoms que formen l'enllaç. Sigui $L(r)$ la densitat electrònica fonamental (també anomenada Lagrangiana) corresponent a una partícula elemental i definida per l'equació 3.89:

$$L(r) = \int dr \left\{ (i/2) (\psi^* \psi) - (h/2m) \sum_i \nabla_i \psi^* \cdot \nabla_i \psi - \nabla \psi^* \psi \right\} \quad (3.89)$$

Si ψ compleix amb l'equació de Schrödinger's depenent del temps, aleshores l'equació 3.89 es simplifica com:

$$L(r) = - (h/4m) \nabla^2 \rho(r) \quad (3.90)$$

On $\rho(r)$ és la densitat de càrrega electrònica, $h/4m$ és un operador matemàtic de l'equació d'Heisenberg i $\nabla^2 \rho(r)$ la segona derivada de la funció densitat de càrrega electrònica també anomenada Laplaciana. $\nabla^2 \rho(r)$ és capaç de mesurar la concentració de la densitat de càrrega electrònica al llarg d'un enllaç químic i pot localitzar el punt on

es concentra la densitat de càrrega electrònica, així en el cas d'un enllaç (interacció) covalent aquesta segona derivada es negativa : $\nabla^2\rho < 0$ i en el cas d'enllaç (interacció) més dèbil com el pont d'hidrogen $\nabla^2\rho > 0$ que es representa en la figura 15 per els dímers d'unió entre molècules d'aigua, en la figura 16 per l'enllaç entre el fluor i l'hidrogen en la molècula d'àcid fluorhídric, com es veu els nuclis atòmics atrauen la densitat de càrrega i el punt crític seria el punt on la densitat de càrrega disminueix indicant el punt on seria més fàcil la ruptura de l'enllaç.

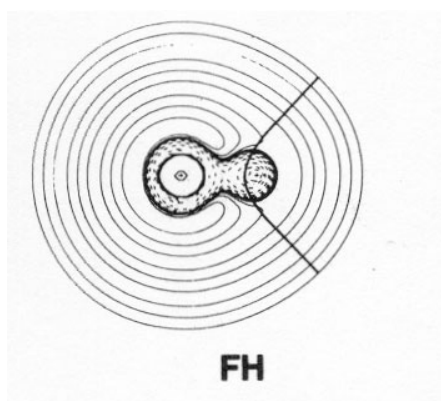


Figura 16. Representació del mapa de densitat electrònica en el punt crític la molècula d'àcid fluorhídric, la línia contínua indica on es situaria el punt crític.

A l'igual que els mapes de MEP's la representació dels mapes de punt crític en molècules són interessants per conèixer aquelles àrees que serien susceptibles d'un possible atac nucleòfil o electròfil, també poden predir amb molta fiabilitat la capacitat donadora o acceptora de H és a dir l'acidesa i la basicitat [61-62], i per últim són de gran ajuda per establir el mecanisme de reacció d'un procés biològic o bé d'un grup de molècules amb interès farmacològic.

3.6.9 Paràmetres termodinàmics

En l'apartat 3.3 hem descrit els principals mètodes d'estudi dels estats de transició que són estructures intermèdies entre la transformació dels reactius i la formació dels productes. L'aplicació de la Química Quàntica ha permès l'estudi en detall dels

processos bioquímics, augmentant així els coneixements d'aquests sistemes moleculars.

Les reaccions químiques que es produeixen en la natura i en els éssers vius estan relacionades amb l'energia. Per interpretar, descobrir i predir els processos biològics s'utilitzen les funcions termodinàmiques, les més importants són:

- l'Energia Interna
- L'Entalpia
- L'Entropia
- L'Entalpia Lliure

En general totes les funcions termodinàmiques són funcions d'estat, és a dir són magnituds que el seu valor no depèn del camí de la reacció, si no dels estats inicial i final. Matemàticament, sigui X una funció d'estat, X_1 : el valor de la funció en l'estat inicial i X_2 el valor de la funció en l'estat final, la variació de la funció d'estat ΔX és igual a la diferència dels valors de la funció en l'estat final i inicial, tal com mostra l'equació 3.91.

$$\Delta X = X_2 - X_1 \quad (3.91)$$

Les funcions termodinàmiques estan relacionades entre sí mitjançant equacions d'estat.

- **Energia Interna**

L'Energia interna (E_i) és la suma de totes les energies del sistema. Entre aquests components de l'energia interna es troben l'energia de rotació (E_r), l'energia de vibració (E_v), l'energia de translació (E_t), de Van der Waals (E_w), energia de repulsió nuclear (E_n) i altres amb una contribució total de resultat final igual a l'Energia interna del sistema, tal i com mostra l'equació 3.92.

$$E_i = E_r + E_v + E_t + E_w + E_n + \dots = \sum E_i \quad (3.92)$$

Si el sistema allibera calor o es produeix una reacció exotèrmica E_i disminueix, de la mateixa forma si en el sistema es produeix una reacció endotèrmica, el sistema guanya calor i E_i augmenta.

- Entalpia

És una funció d'estat que correspon a l'energia interna del sistema més la variació de pressió i volum que experimenta davant qualsevol canvi provocat per causes externes o internes. La variació de l'entalpia del sistema és igual a la calor absorbida o despesa pel propi sistema quan la transformació es realitza a pressió constant.

L'entalpia pot ser estàndard, de formació i de reacció.

Cada element atòmic o substància pura té una entalpia que correspon a la quantitat d'energia calorífica d'un mol de substància quan aquest es troba a pressió constant, aquesta entalpia s'anomena estàndard ΔH^0 i és la que utilitzen molts mètodes semiempírics en el càlcul de diverses propietats moleculars.

Les entalpies estàndard de formació es troben tabulades i s'indiquen a la temperatura de 25°C a la pressió d'una atmosfera.

Una altra entalpia és la de formació ΔH_f i correspon a l'energia calorífica absorbida o cedida al formar-se un mol de la substància a partir dels elements que la constitueixen quan aquests es troben en el seu estat estable o estàndard. És a dir l'entalpia de formació d'una molècula es pot calcular a partir de les entalpies estàndards dels àtoms que la constitueixen.

Quan les molècules es combinen entre sí o bé es transformen per donar a lloc a altres ens trobem davant d'una reacció química, aquesta reacció o procés químic té una entalpia associada que depèn de les entalpies de les molècules inicials (reactius) que han originat les dels productes, aquesta entalpia és l'entalpia de reacció ΔH_r .

L'entalpia de reacció ΔH_r correspon a la variació d'entalpia dels reactius quan aquests es transformen en productes trobant-se tots ells en condicions estàndard. Matemàticament es pot definir per l'equació 3.93 on la variació d'entalpia del sistema després d'haver-se produït la reacció (entalpia de reacció) és igual a la suma de les

entalpies de formació dels productes menys la suma de les entalpies de formació dels reactius.

$$\Delta H_r = \sum \Delta H_f (\text{productes}) - \sum \Delta H_f (\text{reactius}) \quad (3.93)$$

La variació de l'entalpia associada a una reacció o procés químic és la mateixa si la reacció es verifica directament en una sola etapa o si té lloc en varies (Llei d'Hess), segons aquest principi, si relacionem les entalpies de formació de tres espècies químiques: A (reactiu), TS (estat de transició) i C (producte) que intervenen en una reacció molt simple constituïda per una sola etapa, podríem definir tres paràmetres termodinàmics [63] que es representen en la figura 17:

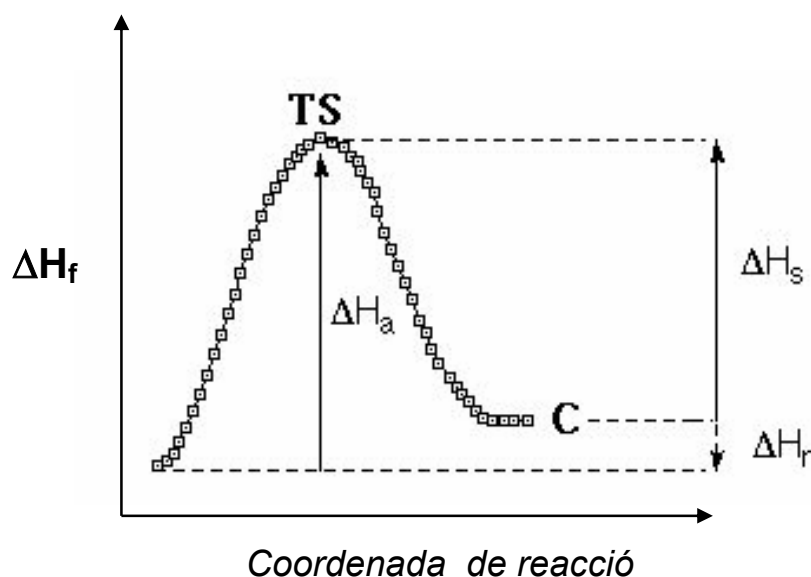


Figura 17. Diagrama energètic per tres paràmetres termodinàmics que permeten estudiar una reacció química: entalpies d'activació (ΔH_a), d'estabilització (ΔH_s) i de reacció (ΔH_r).

- l'energia o entalpia d'activació ΔH_a , que correspon a la diferència entre l'entalpia de formació de l'estat de transició i el reactiu ($\Delta H_{TS} - \Delta H_A$);
- l'energia o entalpia d'estabilització ΔH_s , que és la diferència entre les entalpies de formació de l'estat de transició i el producte ($\Delta H_{TS} - \Delta H_C$)

- l'energia o entalpia de reacció ΔH_r , que és la diferència entre les entalpies de formació de producte i reactiu inicial ($\Delta H_C - \Delta H_A$).

L'entalpia de reacció (ΔH_r) no té en compte l'estat de transició i és un paràmetre termodinàmic molt interessant, ja que el seu significat té a veure amb la facilitat perquè es produeixi una determinada reacció o procés biològic, en general a menor energia de reacció més facilitat té la reacció perquè es produeixi.

Com veurem aquest paràmetre s'ha utilitzat en aquesta tesi per la quantificació de la facilitat de formació de radicals, avaluant així la viabilitat de determinats fàrmacs antiinflamatoris per seguir un cert mecanisme antioxidant.

- Entropia

És una funció d'estat que mesura el desordre del sistema. Matemàticament queda definida per l'equació 3.94:

$$dS = \frac{dE_{q \text{ rev}}}{T} \quad (3.94)$$

L'entropia es manté constant en tots els processos reversibles i augmenta en els irreversibles. Les substàncies pures amb una geometria perfecta tenen una entropia igual a zero. El significat energètic de l'entropia és aquella quantitat d'energia que pertany al sistema i que per tant no pot cedir. Com altres funcions termodinàmiques l'entropia del sistema és la diferència entre la suma d'entropies dels productes menys la suma d'entropies dels reactius:

$$\Delta S_{\text{reacció}} = \sum \Delta S_{\text{(productes)}} - \sum \Delta S_{\text{(reactius)}} \quad (3.95)$$

- Entalpia lliure

Es representa per ΔG , aquesta funció termodinàmica determina la quantitat d'energia que el sistema pot cedir. Permet interpretar i predir un procés químic si aquest transcorre a pressió i temperatura constants.

Matemàticament ve definida per l'equació 3.96:

$$\Delta G = \Delta H - T\Delta S \quad (3.96)$$

Si un procés químic presenta una entalpia lliure negativa, el procés serà espontani i per tant es produirà lliurement. Per exemple com la dissolució d'un solut cristal·lí en aigua, encara que l'entalpia (ΔH) és positiva (procés endotèrmic), perquè absorbeix calor, l'increment d'entropia és elevat, perquè els cristalls del solut passen d'un estat ordenat a un altre de màxim desordre i ΔG resulta < 0 sent un procés irreversible i espontani. En canvi els processos amb una $\Delta G > 0$ resulten irreversibles però no espontanis.

3.7 Referències bibliogràfiques del capítol

- [1] Parr R.G. i Yang W. Density-functional theory of atoms and molecules. Oxford University Press: New York (1989)
- [2] Kollman, P.A. i Stewart, J.J.P. Calculation of Electrostatic field using MOPAC.-6 ESP version in atoms and molecules. *J. Comput. Chem.*, 11, 431-439 (1990).
- [3] Liotard, D.A., Healy, E.F., Ruiz, J.M. I Dewar, M.J.S. AMPAC- vs. 2.1. Quantum Chemistry Program Exchange program 506. QCPE Bull., 9, 123 (1989).
- [4] Frisch, M.J Trucks, G.W, Schlegel, H.B., Gill, P.M.W., Johnson, B.G., Robb, M.A., Cheeseman, J.R., Keith, T., Petersson, G.A., Montgomery, J.A., Raghavachari, K., Al-Laham, M.A., Zakrzewski, V.G., Ortiz, J.V., Foresman, J.B., Cioslowski, J., Stefanov, B.B., Nanayakkara, A., Challacombe, M., Peng, C.Y., Ayala, P.Y., Chen, W., Wong, M.W., Andres, J.L., Replogle, E.S., Gomperts, R., Martin, R.L., Fox, D.J., Binkley, J.S., Defrees, D.J., Baker, J., Stewart, J.P., Head-Gordon, M., Gonzalez, C., People, J.A. GAUSSIAN 94 Rev. E.1 Gaussian Inc. Pittsburgh, P.A. 1995.
- [5] Dupuis, M., Spangler, D. i Wendoloski, J.J. NRCC Software catalog. Program RG01-GAMESS (1980).
- [6] Dupuis, M., Watts, J.D., Hurst, G.J.B. i Villar, H.O. HONDO 7.0 QCPE program 544 (1987).
- [7] Szabo A. i Ostlund NS. Modern Quantum chemistry. *Introduction to advanced electronic structure theory*. McGraw-Hill: New York (1982)
- [8] Levine I.R. Quantum chemistry. Englewood Cliffs Eds. (New Jersey): Prentice-Hall (1991).
- [9] Turner, A.G. *Methods in molecular Orbital Theory*. Academic Press. Inc. (London) Ltd. (1985).
- [10] Dewar, M.J.S., Zoebisch, E.G., Healy, E.F. i Stewart, J.J.P. Modified Neglect of Diatomic Overlap (MNDO): A Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.*, 99, 4899-4909 (1977).
- [11] Dewar, M.J.S., Zoebisch, E.G., Healy, E.F. i Stewart, J.J.P. AM1: A new general purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.*, 107, 3902-3909 (1985).

-
- [12] Stewart, J.J.P. Optimization of Parameters for semiempirical methods. Applications. *J. Comp. Chem.*, 10, 221-264 (1989).
- [13] Lipkowitz, K.B. *Reviews in Computational Chemistry*. D.B. Boyd (Ed.) V.C.H. Publishers, (1990).
- [14] McIver, J.W. i Komornicki, A.. Structure of Transition States in Organic Reactions. General Theory and Applications using a Semiempirical Molecular Orbital Method. *J. Am. Chem. Soc.*, 94, 2625-2633 (1972).
- [15] Poppinger, D. On the Calculation of Transition states. *Chem. Phys. Letters*. 35, 550-554 (1975).
- [16] Komornicki, A. Efficient Determination and characterization of transition states using *ab initio* methods. *Chem. Phys. Letters*. 45, 595-603 (1977).
- [17] Bell, S., Crighton, J.S. i Fletcher, R. A New Efficient Method for locating Saddle Points. *Chem. Phys. Letters*. 82, 122-126 (1981).
- [18] Dewar, M.S.J. i Kirscher, L. MINDO/3 Study of Thermal Conversion of Cyclobutene to 1,3-Butadiene. *J. Am. Chem. Soc.*, 96, 6809-6817 (1974).
- [19] Schlegel, H.B. Optimization of Equilibrium Geometries and Transition Structures. *J. Comput. Chem.*, 3, 214-218 (1982).
- [20] Murrell, J.N. i Laidler, K.J. Determination and characterization of transition states using quantum chemical methods. *Trans. Faraday Soc.*, 64, 371-380 (1968).
- [21] Komornicki, A. i McIver, J.W. Rapid Geometry Optimization for Semi-Empirical Molecular Orbital Methods. *Chem. Phys. Lett.*, 10, 303-311 (1971).
- [22] Fletcher, R. i Powell M.J.D. Characterization of transition states *Comput. J.*, 6, 163-171 (1963).
- [23] Davidson, W.C., Fletcher, R. i Powell M.J.D. Determination and characterization of transition structures. *Comput. J.*, 10, 406-411 (1968).
- [24] Fletcher, R. Algorithmic tools in the study of transition structures. *Comput. J.*, 13, 317-325 (1970).
- [25] Murtagh, B.A.. i Sargent, R.W.H. A new strategy for the location of Transition States using molecular orbital methods. *Comput. J.*, 13, 185-199 (1970).
- [26] Liotard, D.A. Algorithmic tools in the study of semiempirical potential surfaces. *Int. J. Quantum. Chem.*, 44, 723-741 (1992).
- [27] Baker, J. An algorithm for the location of transition structures. *J. Comput. Chem.*, 7, 385-392 (1986).
- [28] Dewar, M.S.J., Healy, E.F. i Stewart, J.J.P. Location of Transition States in Reaction Mechanisms. *J. Am. Chem. Soc. Faraday Trans.*, 80, 227-235 (1984).
- [29] Boyd, D.B., Smith, D.W., Stewart, J.J.P i Wimmer, E. Numerical Sensitivity of Trajectories across Conformational Energy Hypersurfaces from Geometry Optimized Molecular Orbital Calculations. *J. Comput. Chem.*, 9, 387-394 (1988).

-
- [30] Case, D.A., Pearlman, D.A., Caldwell, J.W., Cheatham, T.E., Ross, W.S., Simmerling, C.L., Dearden, T.A., Merz, K.M., Stanton, R.V., Cheng, A.L., Vicent, J.J., Crowley, M., Ferguson, D.M., Radmer, R.J., Seibel, G.L., Singh, U.C., Weiner, P.K. i Kollman, P.A. AMBER vs. 5. (1997). University of California, San Francisco.
- [31] Van Gunsteren, W.F., Billeter SR, Eising, A.A., Hünenberger, P.H., Krüger, P., Mark, A.E., Scott, W.R.P., Tironi, I.G. Biomolecular Simulation: The GROMOS-96 Manual and User Guide, Vdf Hochschulverlag, Zürich, Switzerland (1996)
- [32] CHARMM Brooks, B.R., Brucoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. I Karplus, M. CHARMM: A program for macromolecular energy, minimization and dynamics calculations. *J. Comput. Chem.*, 9, 387-394 (1988).
- [33] Ewald, P.P. Die Berechnung Optischer und Elektrostatischer Gitterpotentiale. *Ann. Phys. (Leipzig)* 64, 253-266 (1921).
- [34] Darden T., York D. i Pedersen L.. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.*, 98, 10089-10104 (1993).
- [35] York, D.M., Darden, T.A. i Pedersen, L.G. The effect of long-range electrostatic interactions in simulations of macromolecular crystals: A comparison of the Ewald and truncated list methods. *J. Chem. Phys.*, 99, 8345-8356 (1993).
- [36] Leach, A.R. Molecular Modelling. Principles and Applications. Addison Wesley Logman Limited. (1996)
- [37] Van Gunsteren, W.F. In Computer Simulation of Biomolecular Systems, Vol. 2. van Gunsteren WF, Weiner P.K, Wilkinson A.J. (Eds.) ESCOM: Leiden (1993).
- [38] McCammon, J.A. i Harvey, S.C. Dynamics of proteins and nucleic acids. Cambridge University Press. (1987).
- [39] Ryckaert, J.P., Ciccotti G. i Berendsen H.J.C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.*, 23,327-339 (1977).
- [40] Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., Di Nola, A. i Haak, J.R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, 81,3684-3697 (1984).
- [41] McQuarrie, D.A. Statistical Thermodynamics, Harper and Row: Weiner P.K, Frautschy, S.A i Wilkinson A.J. (Eds.) New York, (1973).
- [42] Karplus, M. i Kushick, J.N. Method for estimating the configurational entropy of macromolecules. *Macromolecules*, 14, 325-339 (1981).
- [43] Schlitter, J. Estimation of absolute and relative entropies of macromolecules using the covariance matrix. *Chem. Phys. Lett.*, 215, 617-623 (1993).
- [44] Metropolis, N., Rosenbluth A.W., Rosenbluth M.N., Teller A.H. i Teller E. Equation of state calculation by fast computing machines. *J. Chem. Phys.*, 21, 1087-1099 (1953).
- [45] Kuntz, I.D., Blaney, J.M., Oatley, S.J., Langridge R. i Ferrin, T.E. A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.*, 161, 269-281 (1982)
- [46] Knegtel, R.M., Kuntz, I.D. i Oshiro, C.M. Molecular docking to ensembles of protein structures. *J. Mol. Biol.*, 266, 424-435 (1997).
- [47] DesJarlais, R.L., Sheridan, R.P., Seibel, G.L., Dixon, J.S., Kuntz, I.D. i

- Venkataraghavan R. Using shape complementarity as an initial screen in designing ligands for a receptor binding site of known three-dimensional structure. *J. Med. Chem.*, 31, 722-735 (1998).
- [48] Gasteiger J. I Marsili M. Iterative partial equalization of orbital electronegativity: A rapid access to atomic charges. *Tetrahedron*, 36, 3219-3227 (1980).
- [49] Orozco, M. i Luque, F.J. Theoretical Methods for the description of the solvent effect in biomolecular systems. *Chem. Rev.*, 100,4187-4199 (2000).
- [50] Ben-Naim, A. Solvation Thermodynamics. Plenum: New York (1987).
- [51] Gao J. Hybrid quantum and molecular mechanical simulations: an alternative avenue to solvent effects in organic chemistry. *Acc. Chem. Res.*, 29, 298-315 (1996).
- [52] Tomasi, J. i Perisco, M. Molecular interactions in solution: an overview of methods based on continuous distributions of the solvent. *Chem. Rev.*, 94, 2027-2094 (1994).
- [53] Cramer, J. i Truhlar, D.G. , in *Solvent Effects and Chemical Reactivity*. Tapia, O. I Bertán, J. (Eds.) Kluwer Academic Publishers, Dordrecht (1996).
- [54] Orozco M., Alhambra C., Barril X., Lopez J., Busquets M. i Luque F.J. Theoretical Methods for the representation of solvent. *J. Mol. Mod.*, 2, 1-19 (1996).
- [55] Hinchliffe, A.i Dobson, J.C. chemical interpretations of molecular wavefunctions *J. Chem. Soc. Rev.*, 5, 79-81 (1976).
- [56] Luque, F.J., Illas, F. i Orozco. M. Comparative study of the molecular electrostatic potential obtained from different wavefunctions. Reability of the semiempirical MNDO wavefunction. *J. Comput. Chem.*, 11, 416-419 (1990).
- [57] Francisco Javier Luque Garriga. Estudio del mecanismo de activación del receptor de la histamina H₂ a través de los métodos de la Química Cuántica. *Tesis Doctoral* (1989).
- [58] Orozco, M. i Luque, F.J. On the use of AM1 and MNDO wave functions to compute accurate electrostatic charges. *J. Compt. Chem.*, 11, 909-923 (1990).
- [59] Orozco, M. i Luque, F.J. Molecular Interaction Potential: anew tool for the theoretical study of molecular reactivity. *J. Compt. Chem.*, 14, 587-602 (1993).
- [60] Bader, R.F.W. i Essén, H. The characterization of atomic interactions. *Chem. Phys.*, 80, 1943-1957 (1984).
- [61] Platts, J.A. Theoretical prediction of hydrogen bond donor capacity. *Phys. Chem. Chem. Phys.*, 2, 973-980 (2000).
- [62] Platts, J.A. Theoretical prediction of hydrogen bond basicity. *Phys. Chem. Chem. Phys.*, 2, 3115-3120 (2000).
- [63] Dickerson, R.E. Molecular Thermodynamics. Benjamin, W.A. Ed. Menlo Park, California, USA. (1969).

- [64] Xavier Barril Alonso. Tesi doctoral: Estudis teòrics de mecanismes i especificitat d'interacció lligand-receptor. 2001.