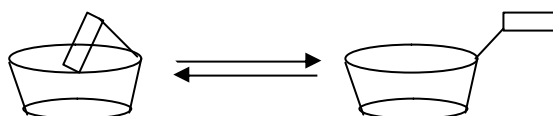


## 8 ESTUDI CONFORMACIONAL DELS DÍMERS DE LA 3-O-(2-METILNAFTIL)- $\beta$ -CICLODEXTRINA

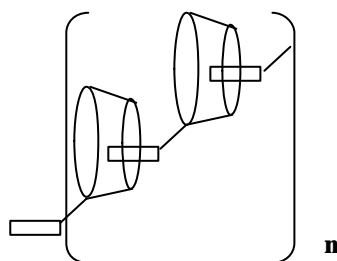
### 8.1 PRECEDENTS I INTRODUCCIÓ AL PROBLEMA

En una publicació de l'any 1996 es descriu la síntesi d'una nova ciclodextrina modificada, la 3-O-(2-metilnaftil)- $\beta$ -ciclodextrina.<sup>65</sup> Dels primers estudis realitzats, els autors arriben a la conclusió que existeix un isomerisme altament depenent de la temperatura entre dues conformacions (Figura 8-1). La primera de les conformacions conté el grup naftil dins de la cavitat (a temperatures baixes), i en la segona conformació (a temperatures altes) el grup naftil es troba fora de la cavitat.



**Figura 8-1** Representació de l'isomerisme entre les estructures amb el grup aromàtic a l'interior i a l'exterior de la cavitat.

Dos anys més tard, els mateixos autors publiquen més anàlisis espectroscòpiques de la mateixa molècula amb l'objectiu de determinar amb més exactitud la posició del grup naftil.<sup>66</sup> Els estudis realitzats per experiments de RMN, NOE i NOESY són consistents amb un tipus d'estructura on el grup naftil està inclòs dins de la cavitat. Però els autors no afirmen que aquesta molècula existeixi en forma de monòmer i postulen estructures polimèriques com les de la Figura 8-2, on un grup naftil podria estar inclòs dins de la cavitat d'una altre ciclodextrina.

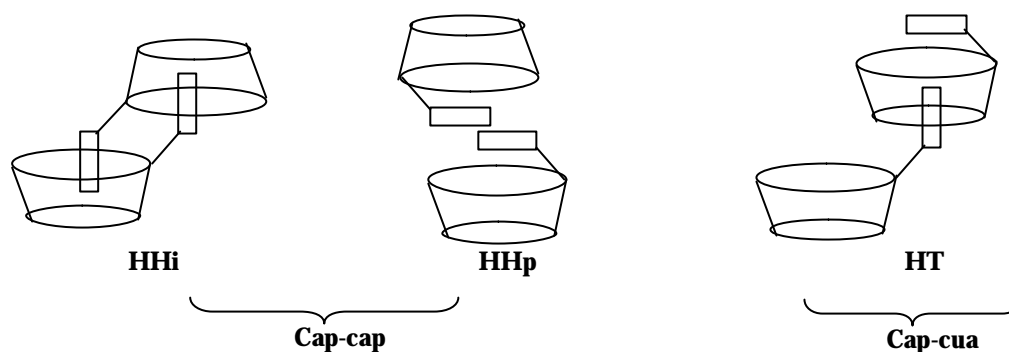


**Figura 8-2** Representació d'una estructura polimèrica per a la 3-O-(2-metilnaftil)- $\beta$ -ciclodextrina.

<sup>65</sup> Mc.Alpine, S.R; Garcia-Garibay, M.A., *J. Am. Chem. Soc.*, **1996**, 118(11), 2750-2751.

<sup>66</sup> Mc.Alpine, S.R; Garcia-Garibay, M.A., *J. Am. Chem. Soc.*, **1998**, 120(18), 4269-4275.

Finalment, per estudis de RMN a diferents concentracions arriben a la conclusió de que aquesta molècula existeix en forma de dímer a concentracions superiors a  $10^{-4}$  M. Així doncs, proposen tres tipus de dímers diferents, dos tipus cap-cap i un tipus cap-cua (Figura 8-3).<sup>67</sup> Els dos dímers cap-cap es diferencien en la posició dels grups aromàtics: en el primer els grups aromàtics estan dins de la cavitat de l'altre ciclodextrina (HHi), i en el segon els grups aromàtics estan paral·lels entre les dues ciclodextrines (HHp).



**Figura 8-3 Representació dels tres tipus de dímers de la 3-O-(2-metilnaftil)-β-ciclodextrina.**

En la mateixa publicació, per mètodes de diroïsmes circulars també arriben a la conclusió que el monòmer existeix a menors concentracions, però no amb la conformació postulada en el seu primer article, sinó que aquest existeix en una estructura tipus gorra en la qual el grup naftil només està parcialment inclòs en la cavitat de la ciclodextrina.

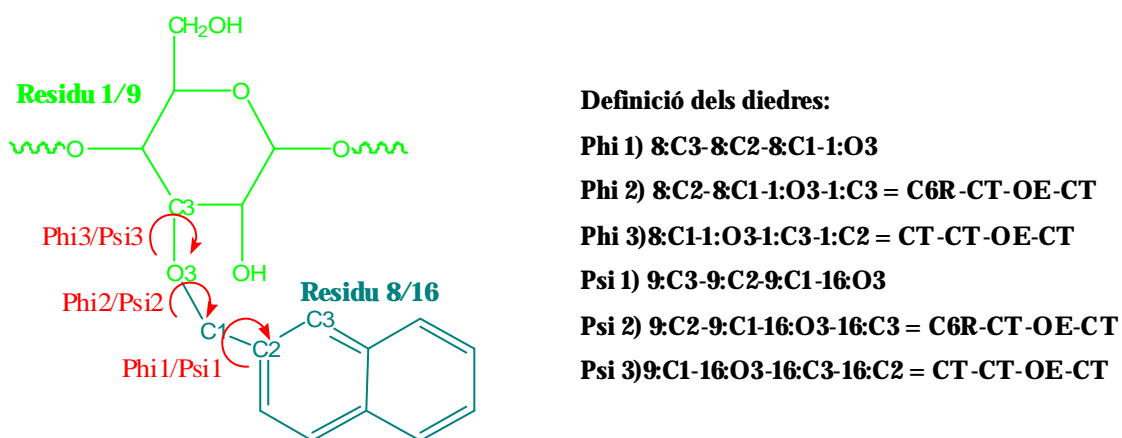
En aquest capítol s'estudien per mètodes de mecànica i dinàmica molecular en fase gas i solució aquosa l'estructura monomèrica i els tres dímers de la 3-O-(2-metilnaftil) β-ciclodextrina. L'estudi s'ha dut a terme en col·laboració amb el *Computational Chemistry Laboratory* de la *University of Cape Town* sota direcció del professor Kevin J. Naidoo. El programa utilitzat ha estat el CHARMM-27 i s'ha utilitzat el camp de forces CSFF, el qual és una modificació dels camps de força originals en CHARMM: el HGFB i el PHLB. No obstant, els paràmetres corresponents als enllaços d'unió entre la glucosa i el grup metilnaftil eren de molt baixa qualitat i s'han tornat a parametritzar. La parametrització i els detalls computacionals estan descrits en l'apèndix.

<sup>67</sup> La orientació cap-cap entre dues ciclodextrines és aquella en què els costats amples del con truncat estan encarats entre si. La orientació cap-cua és aquella en què una part ampla del con truncat està encarada a una part estreta de l'altre con truncat. L'orientació cua-cua seria aquella en què dues parts estretes del con truncat estiguessin encarades entre si.

Un cop obtingudes les trajectòries, s'ha analitzat l'evolució dels diedres que connecten la ciclodextrina amb el grup aromàtic al llarg del temps, amb la finalitat de detectar els diferents confòrmers en cada un dels sistemes estudiats. Posteriorment, s'ha realitzat l'anàlisi dels enllaços d'hidrogen, i finalment, s'han realitzat les anàlisis dels diferents termes energètics que contribueixen a l'energia total del sistema.

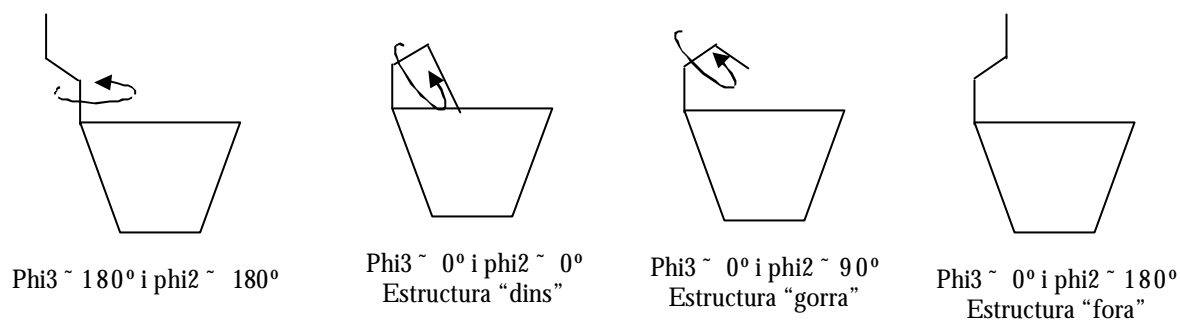
## 8.2 ANÀLISIS DELS DIEDRES

S'han estudiat els diedres indicats a la Figura 8-4 per a cada un dels monòmers. En els dímers s'anomenen phi1, phi2 i phi3 en un dels monòmers i psi1, psi2 i psi3 en l'altre monòmer.



**Figura 8-4 Representació dels diedres estudiats.**

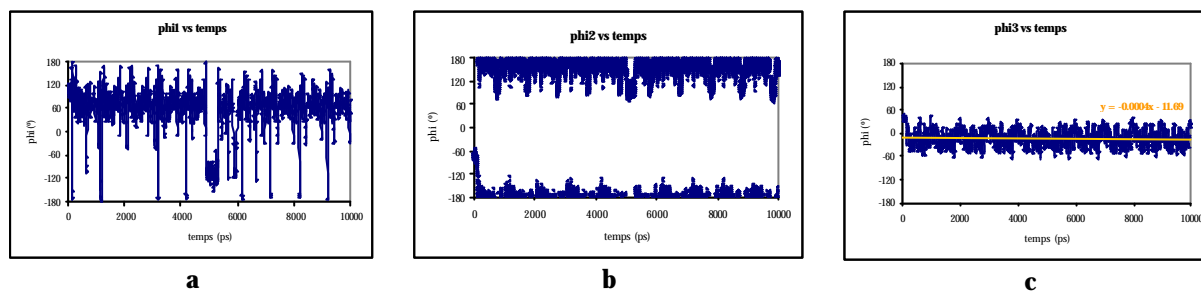
La rotació de phi1/psi1 representa la rotació de l'anell aromàtic sobre si mateix i dóna idea de la llibertat rotacional d'aquest grup segons els impediments estèrics que pugui tenir amb el seu entorn. Per rotació de phi3/psi3 es pot passar de l'estructura on el grup **metilnaftil** es situa clarament a la part externa de la cavitat (valors de 180°) a una estructura on l'esmentat grup es situa a la part interna de la cavitat (valors propers a 0°). A partir d'aquesta darrera estructura, on phi3 té valors pròxims a 0°, per rotació de phi2/psi2 s'obtenen estructures on el grup **naftil** està totalment inclòs dins de la cavitat amb valors de 0°, estructures tipus "gorra" amb valors d'aquest diedre propers als 90° i estructures tipus "fora" amb valors propers als 180°. Així doncs, a partir de la rotació d'aquest darrer diedre es poden detectar totes les conformacions descrites experimentalment (Figura 8-5).



**Figura 8-5** Esquematització de les diferents conformacions obtingudes per rotació dels diedres phi2 i phi3.

### 8.2.1 Anàlisi conformacional del monòmer

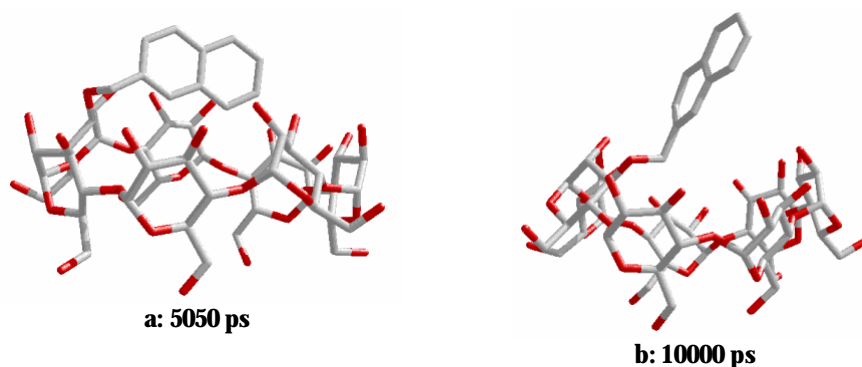
S’han iniciat els estudis dels diedres amb el monòmer. Els resultats obtinguts de l’anàlisi dels diedres al llarg del temps en fase gas es representen a la Gràfica 8-1.



**Gràfica 8-1** Variació de a) diedre phi1, b) diedre phi2 i c) diedre phi3, al llarg del temps en la simulació del monòmer en fase gas.

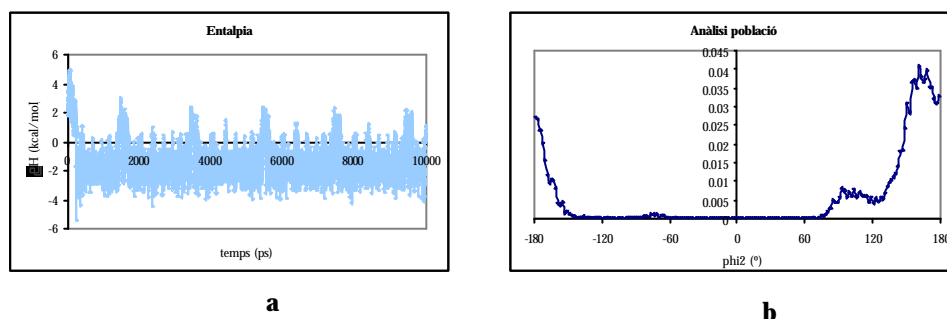
En primer lloc, cal destacar la oscil·lació contínua en els valors de phi2 entre valors propers a 180° i a 90°. Aquesta oscil·lació és una clara indicació que al llarg de la simulació es detecten dues conformacions diferents. D’altra banda, en els primers ps de la simulació es detecta en els valors de phi3 el pas de l’estructura inicial, que té valors d’aquest diedre propers a 50°, a l’estructura que es manté al llarg de la simulació amb valors al voltant dels 0°, per tant, una estructura on el grup metilnaftil clarament es situa a la part interna de la cavitat.

Les dues conformacions detectades al llarg de la simulació en fase gas es representen a la Figura 8-6. La primera conformació detectada en aquesta simulació correspon a l’estructura tipus “gorra” descrita experimentalment (Figura 8-6a) i la segona, correspon a l’estructura tipus “fora” també descrita experimentalment (Figura 8-6b).



**Figura 8-6** Estructures representatives del monòmer amb a) conformació gorra i b) conformació "fora".

Donat que es detecten aquestes dues conformacions al llarg de la simulació s'ha considerat interessant estimar tant l'entalpia com l'entropia del pas d'una a l'altre conformació. Per estimar l'entalpia, s'ha avaluat l'energia de torsió dels tres diedres que connecten l'anell de glucosa amb l'anell aromàtic al llarg del temps (Gràfica 8-2a). D'altra banda, malgrat que l'entropia no es pot calcular exactament a partir de les simulacions de DM, sí que es pot estimar a partir de les poblacions de cada una de les conformacions. Així doncs, s'han calculat les poblacions de cada una de les conformacions emprant una distribució de Boltzman amb l'objectiu d'avaluar l'efecte del terme entròpic (Gràfica 8-2b).



**Gràfica 8-2** a) Estimació de l'entalpia de transició de l'estructura tipus "fora" a l'estructura tipus gorra en fase gas i b) poblacions de les conformacions.

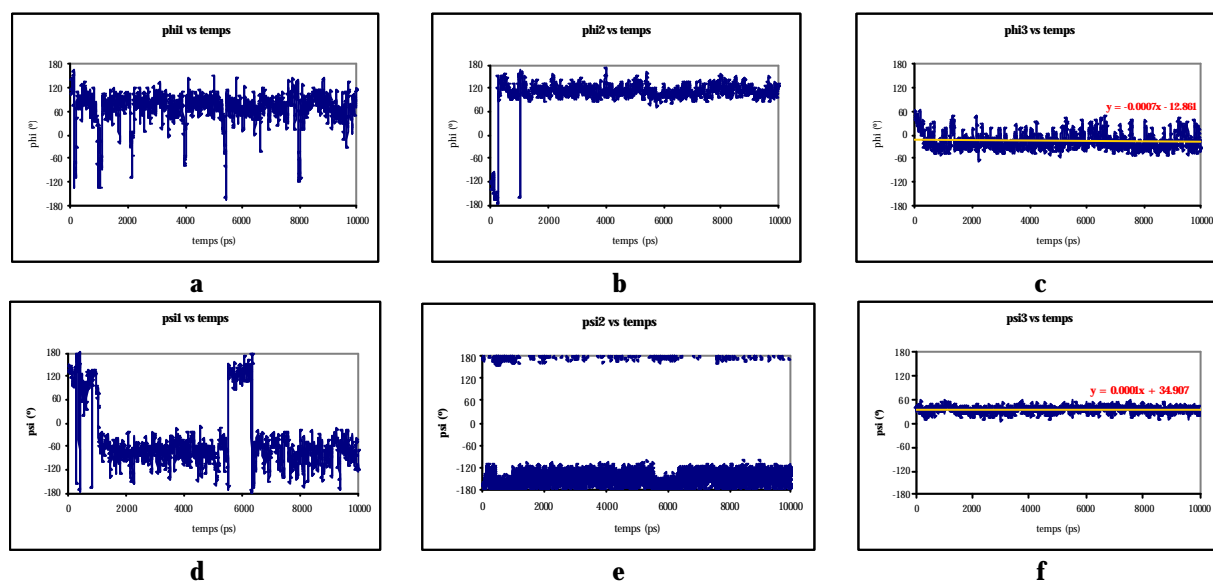
De les gràfiques s'observa en primer lloc, que l'entalpia presenta oscil·lacions equivalents a les observades per al diedre  $\phi_2$  al llarg del temps. Per tant, sembla evident que l'entalpia afavoreix més a l'estructura tipus "fora" que a l'estructura tipus "gorra". D'altra banda, a partir de l'anàlisi de població s'ha calculat l'àrea sota la corba per a cada una de les conformacions, i s'ha estimat que la població relativa de cada un d'aquests conformers és de 1:6,6 a favor de l'estructura "fora" i, per tant, aquesta darrera conformació també està estabilitzada pel terme entròpic.

Els resultats obtinguts de les simulacions en solució aquosa no es mostren aquí, però convé comentar que novament s'han detectat dues conformacions al llarg de la simulació i que

aquestes són equivalents a les detectades en fase gas. A partir de l'anàlisi de població s'ha estimat que en aquest cas la població relativa és 1:27 a favor de l'estructura "fora". Es sospita que aquesta diferència en les poblacions relatives, pot ser deguda a la major àrea de solvatació de l'estructura "fora" que fa que s'estabilitzi més en solució aquosa. De totes maneres, no s'ha d'oblidar que aquestes poblacions relatives tan sols són una aproximació ja que, degut al curt temps de simulació, no es pot tenir cap garantia que aquestes siguin una autèntica representació de la població real del sistema.

### 8.2.2 Anàlisi conformacional del dímer HHi

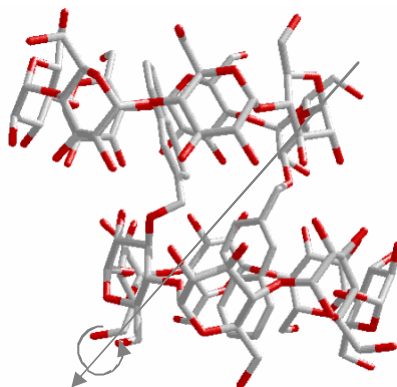
A continuació, s'ha estudiat el primer dels dímers: el HHi. Al llarg de la simulació en fase gas d'aquest sistema no s'observen canvis significatius en els valors del diedres (Gràfica 8-3) i, per tant, es detecta una única conformació. Per a aquesta conformació hi ha valors similars de  $\phi_2/\psi_2$  (propers als  $180^\circ$  en ambdós casos) i igualment, succeeix amb els valors dels diedres  $\phi_3/\psi_3$  (valors propers als  $\pm 45^\circ$ ). Per tant, aquest fet indica que la conformació detectada presenta una certa simetria. Dels valors dels diedres, també s'ha de destacar que són els diedres  $\phi_1$  i  $\psi_1$  els que oscil·len en un rang més ampli al llarg de la simulació.



**Gràfica 8-3 Variació de a) diedre  $\phi_1$ , b) diedre  $\phi_2$ , c) diedre  $\phi_3$ , d) diedre  $\psi_1$ , e) diedre  $\psi_2$  i f) diedre  $\psi_3$ , al llarg del temps en la simulació del dímer HHi en fase gas.**

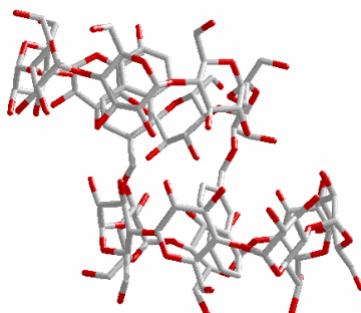
A la Figura 8-7 es mostra l'estructura representativa d'aquesta simulació. Es pot comprovar com efectivament, l'estructura és molt simètrica i existeix un  $C_2$  perpendicular al pla del paper. Si aquest fos el dímer més estable, aquesta simetria s'hauria de veure reflectida en el nombre i el tipus de senyals dels espectres de RMN, ja que davant d'una estructura menys

simètrica, el nombre de senyals hauria de ser aproximadament la meitat. A més, es poden entendre les oscil·lacions dels diedres  $\phi_1/\psi_1$  perquè els anells aromàtics es situen just al mig de la cavitat de la ciclodextrina i, per tant, tenen més facilitat per rotar sobre si mateixos.



**Figura 8-7 Estructura representativa del dímer HHi en fase gas.**

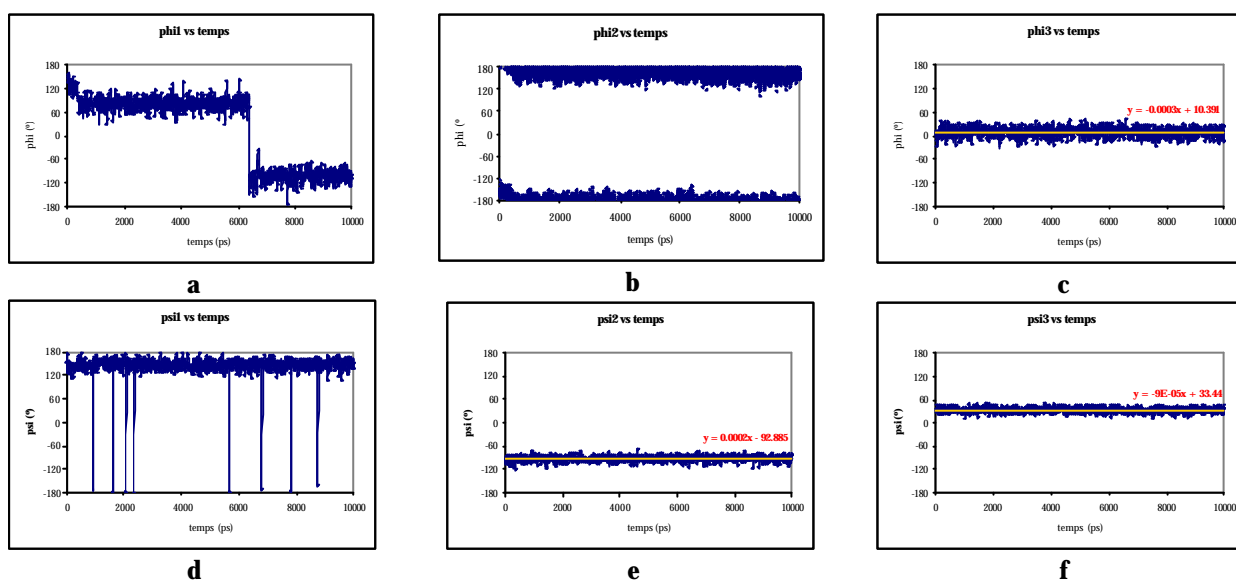
En aquest cas tampoc es mostren els resultats obtinguts en les simulacions en solució aquosa però, igual que en fase gas, s'ha detectat una única conformació per a aquest dímer. Els diedres  $\phi_1$  i  $\psi_1$  són els que més varien al llarg de la simulació i  $\phi_2/\psi_2$  i  $\phi_3/\psi_3$  també tenen valors similars entre ells. No obstant, en aquest cas l'estructura representativa (Figura 8-8) és lleugerament diferent a l'estructura representativa en fase gas. En solució aquosa les dues ciclodextrines continuen estant en plans paral·lels i es continua mantenint l'eix de simetria  $C_2$ , però els plans de les ciclodextrines estan desplaçats cap a costats oposats respecte l'estructura en fase gas (la ciclodextrina superior cap a l'esquerra i la ciclodextrina inferior cap a la dreta). És a dir, degut a la presència de les molècules d'aigua, es trenquen molt enllaços d'hidrogen entre hidroxils secundaris de les dues ciclodextrines i es formen nous enllaços d'hidrogen entre aquests i les molècules d'aigua de l'entorn.



**Figura 8-8 Estructura representativa del dímer HHi en solució aquosa.**

### 8.2.3 Anàlisi conformacional del dímer HHp

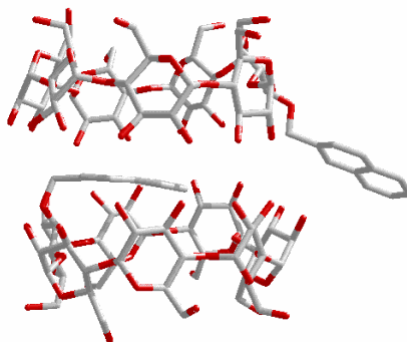
A continuació s'ha estudiat el dímer HHp. Es pot avançar que en aquest cas no es manté la conformació inicial, sinó que ja en les etapes d'escalfament i equilibració es produeixen canvis conformacionals importants. Els canvis més representatius es poden apreciar a la Gràfica 8-4 en els valors del diedre phi2 en els primers ps de la simulació: aquest diedre que inicialment té valors al voltant de  $-150^\circ$ , al cap d'uns ps adquireix valors de  $180^\circ$ . També cal destacar que, a diferència del dímer anterior on phi2/psi2 i phi3/psi3 tenen valors similars, en aquest cas phi3/psi3 continuen tenint valors semblants, però ara phi2 té valors propers als  $\pm 180^\circ$  i psi2 té valors al voltant dels  $-100^\circ$ , per tant, aquest fet suggereix que aquest dímer no té elements de simetria com en tenia el dímer anterior.



**Gràfica 8-4 Variació de a) diedre phi1, b) diedre phi2, c) diedre phi3, d) diedre psi1, e) diedre psi2 i f) diedre psi3, al llarg del temps en el dímer HHp en fase gas.**

A la Figura 8-9 es pot comprovar com efectivament aquest dímer no té elements de simetria; s'observa com un dels anells aromàtics es manté entre les dues ciclodextrines, però l'altre anell aromàtic ha sortit fora de la cavitat. Per tant, sembla que hi ha una forta repulsió electrostàtica entre els dos anells aromàtics que no permet que aquests estiguin propers l'un de l'altre, creant interaccions del tipus *p-stacking*



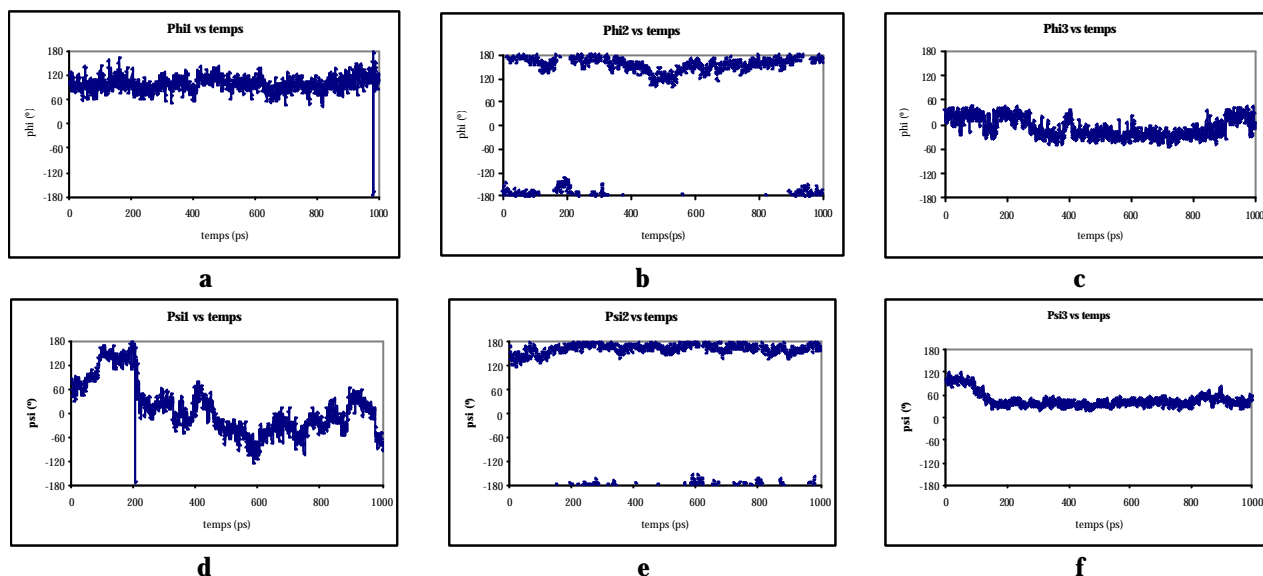


**Figura 8-9 Estructura representativa del dímer HHp en fase gas.**

A diferència dels casos anteriorment descrits (monòmer i dímer HHi), en aquest cas s'observa un comportament molt diferent en solució aquosa que en fase gas (Gràfica 8-5). En primer lloc, cal destacar els canvis que hi ha al voltant dels 100 ps en els valors del diedre  $\psi_3$  i els canvis al voltant dels 300 ps en els valors dels diedres  $\phi_3$ ; aquests indiquen que existeixen diferents conformacions al llarg de la simulació en solució aquosa. A més, l'estructura majoritària té valors d'aquests diedres de  $\pm 45^\circ$ . També s'observa que, a diferència de les variacions observades en fase gas, en aquest cas tant  $\phi_2$  com  $\psi_2$  tenen valors similars i propers als  $\pm 180^\circ$ . En aquest punt és important recordar que valors de  $\phi_2/\psi_2$  propers a  $180^\circ$  i valors de  $\phi_3/\psi_3$  propers a  $\pm 45^\circ$  són característics d'una estructura tipus HHi. Per tant, sembla clar que al llarg de la simulació en solució aquosa es produeix la transició del dímer HHp a un dímer tipus HHi.<sup>68</sup>

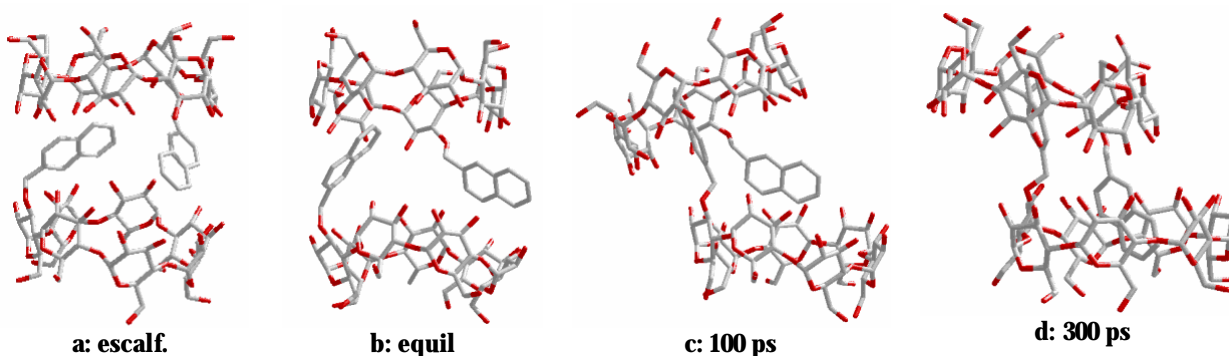
Aquests resultats, semblen indicar que hi ha un fort efecte hidrofòbic dels grups aromàtics que fa que aquests en presència del solvent, de caràcter polar, busquin un entorn més hidrofòbic com la cavitat de la ciclodextrina. Així doncs, a partir d'aquestes observacions ja es poden extreure les primeres conclusions sobre quin pot ser el dímer majoritari i energèticament més estable.

<sup>68</sup> A partir d'aquí per fer referència a aquesta simulació s'utilitzarà l'abreviatura HHpi per diferenciar-lo dels dímers HHp i HHi inicials.

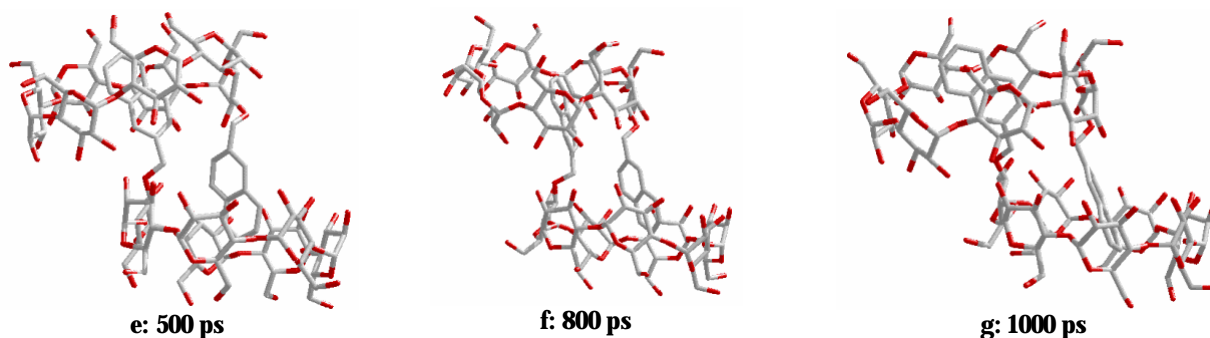


**Gràfica 8-5 Variació a) diedre phi1, b) diedre phi2, c) diedre phi3, d) diedre psi1, e) diere psi2 i f) diedre psi3, al llarg de la simulació del dímer HHp en solució aquosa.**

A la Figura 8-10 i Figura 8-11 es representen diferents estructures al llarg de la simulació en solució aquosa d'aquest dímer, per tal d'obtenir una visió global del procés de passar d'un dels dímers a l'altre dímer. S'observa que ja en les primeres etapes d'escalfament i equilibració els dos anells aromàtics s'allunyen entre si respectivament; als 100 ps de simulació un dels anells aromàtics ja s'ha introduït en la cavitat de l'altre ciclodextrina i als 300 ps l'altre anell aromàtic també s'introdueix en l'altre cavitat. Per tant, als 300 ps ja s'observa un dímer tipus HHi on les dues ciclodextrines estan una sobre de l'altre i en plans paral·lels. Durant la resta de simulació el nou dímer presenta més canvis conformacionals per tal d'estabilitzar-se fins que al final de la simulació les dues ciclodextrines s'han desplaçat lateralment per afavorir els enllaços d'hidrogen amb les molècules d'aigua de la mateixa manera descrita en el dímer HHi en solució aquosa.

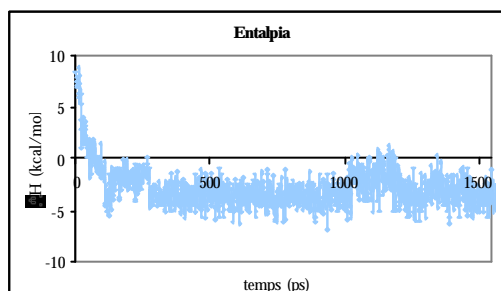


**Figura 8-10 Estructures representatives del pas del dímer HHp al dímer HHpi en solució aquosa. Estructures a) després de l'escalfament b) després de l'equilibració c) als 100 ps d) als 300 ps. Continua pàgina següent.**



**Figura 8-11** Continuació pàgina anterior. Estructures representatives del pas del dímer HHp al dímer HHpi en solució aquosa. Estructures e) als 500 ps f) als 800 ps i g) als 1000 ps.

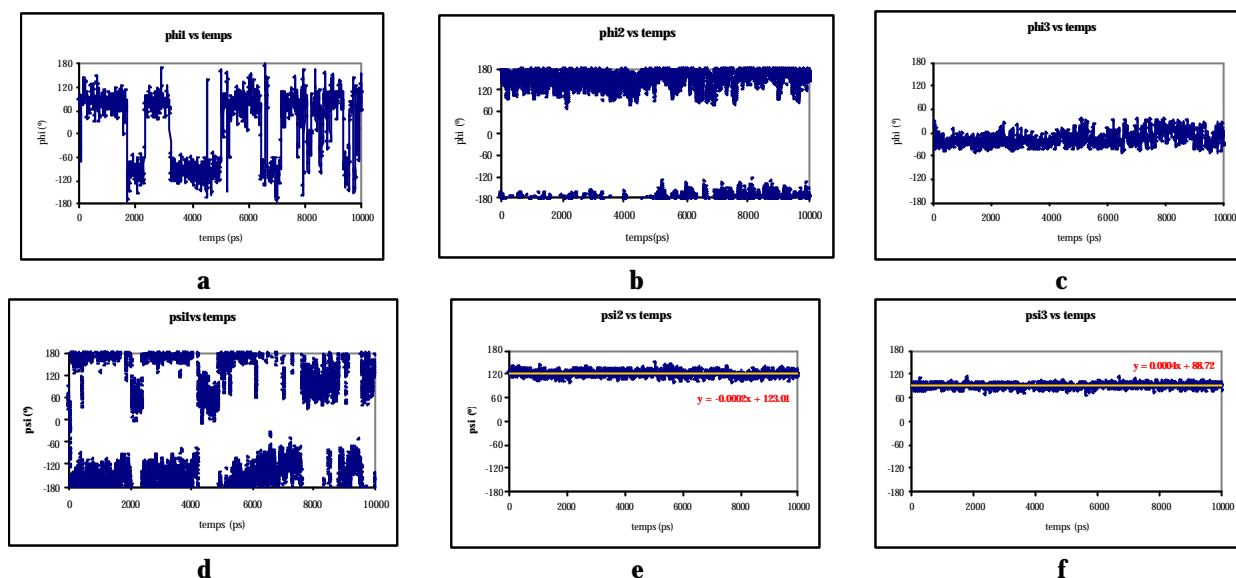
Finalment, en aquest cas també s'ha considerat interessant fer una estimació de l'entalpia requerida per passar d'un dímer tipus HHp a un dímer HHpi. A la Gràfica 8-6 es representa la variació d'entalpia associada a l'energia de torsió al llarg del temps. S'observa que el dímer HHpi és més estable que el dímer HHp i es pot estimar que, segons l'energia de torsió, el pas del dímer HHp al dímer HHpi requereix aproximadament 12 kcal/mol.



**Gràfica 8-6** Estimació de l'entalpia del pas del dímer HHp al dímer HHpi en solució aquosa.

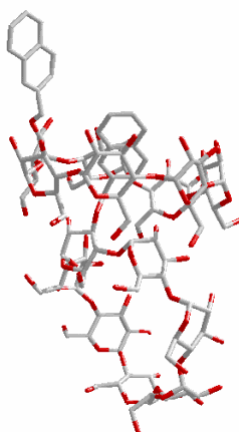
#### 8.2.4 Anàlisi conformacional del dímer HT

Per acabar les anàlisis conformacionals, es detallen els resultats obtinguts per al dímer cap-cua. De l'estudi de la variació dels diedres al llarg del temps en la simulació en fase gas s'observa que no hi ha gaires canvis significatius. Els canvis més importants es produeixen en els valors de  $\phi_{i1}$  i  $\psi_{i1}$ , però com ja s'ha esmentat anteriorment, aquests no proporcionen canvis conformacionals significatius d'interès. Per tant, novament es detecta una única conformació referent a la posició dels anells aromàtics.



**Gràfica 8-7 Variació de a) diedre phi1, b) diedre phi2, c) diedre phi3, d) diedre psi1, e) diedre psi2 i f) diedre psi3, al llarg del temps en la simulació del dímer HT en fase gas.**

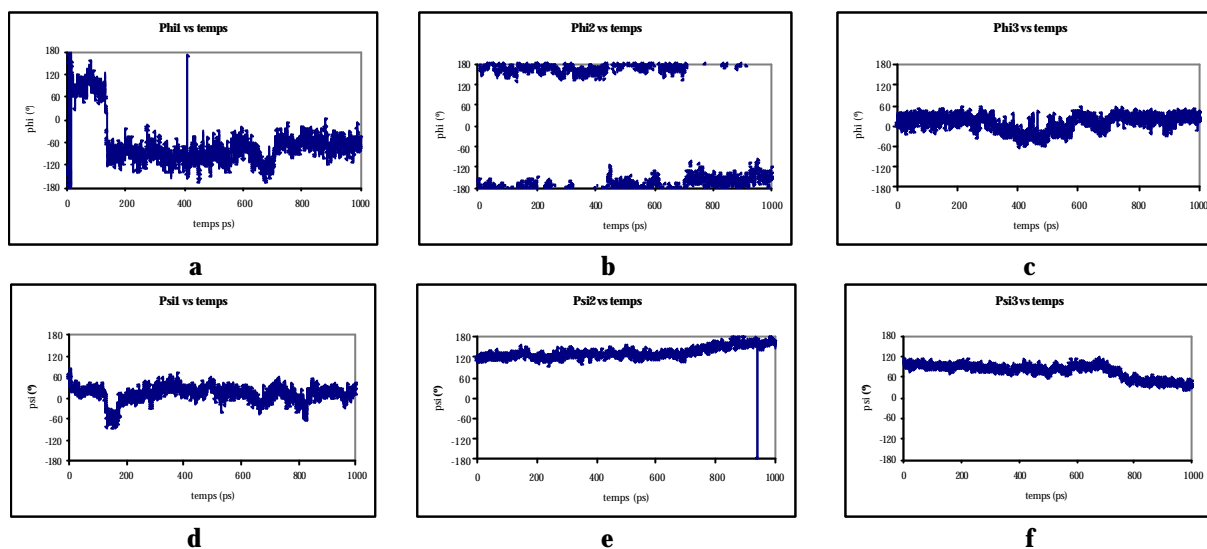
No obstant, analitzant les estructures d'aquesta simulació amb un programa gràfic (Mol view), s'observa que en aquest cas es produeixen altres canvis més importants que fan relació a l'orientació relativa de les dues ciclodextrines entre si. Com a exemple a la Figura 8-12 es mostra l'estructura representativa d'aquesta simulació. En aquest cas, s'observa que les dues unitats de ciclodextrina es mantenen en plans gairebé perpendiculars, és a dir, sembla que hi ha una certa tendència que a partir d'un dímer tipus cap-cua s'obtingui al llarg de la simulació un dímer tipus cap-cap.



**Figura 8-12 Estructura representativa del dímer HT en fase gas.**

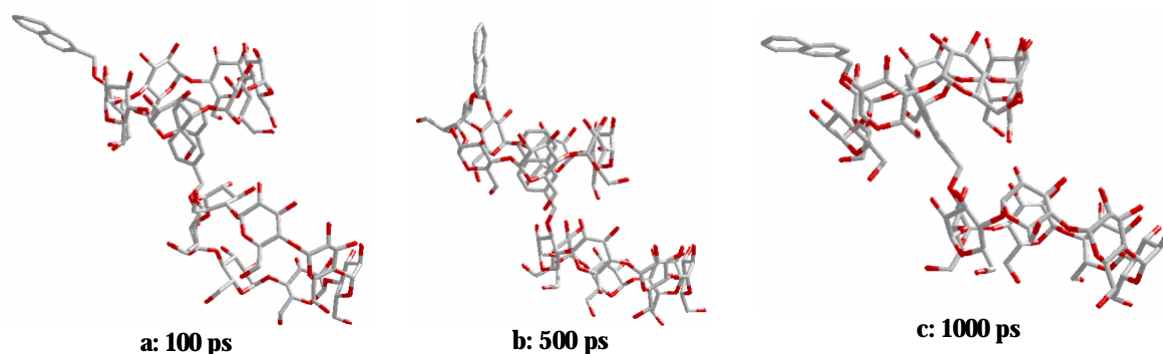
De la mateixa manera que s'ha comentat per al dímer anterior, en aquest cas també s'observa un comportament diferent en solució aquosa (Gràfica 8-8). Al voltant dels 700 ps els valors del diedre phi2 s'aproximen cap als  $100^\circ$ , és a dir, que partint d'un conformació tipus

“fora” en la ciclodextrina superior com la descrita per al monòmer, es tendeix a tenir una estructura tipus “gorra”. D'altra banda, els canvis produïts en la ciclodextrina superior fan que la ciclodextrina inferior també es reajusti, de manera que s'observa que a partir dels 700 ps de la simulació els valors de  $\psi_2$  tendeixen cap a valors propers als  $180^\circ$  i els valors de  $\psi_3$  tendeixen cap a valors propers a  $45^\circ$ .



**Gràfica 8-8 Variació de a) diedre  $\phi_1$ , b) diedre  $\phi_2$ , c) diedre  $\phi_3$ , d) diedre  $\psi_1$ , e) diedre  $\psi_2$  i f) diedre  $\psi_3$ , al llarg del temps en la simulació del dímer HT en solució aquosa.**

En aquest cas, no s'observa el canvi d'orientació relativa entre les dues ciclodextrines sinó que aquestes es mantenen en plans paral·lels igual que en els altres dímers. A la Figura 8-13 es mostren algunes de les estructures més representatives d'aquesta simulació. S'observa com, efectivament, en la major part de la simulació les dues ciclodextrines es mantenen paral·leles l'una respecte de l'altre. Aquesta orientació relativa probablement està estabilitzada pels enllaços d'hidrogen que es formen entre els hidroxils de les ciclodextrines i les molècules d'aigua de l'entorn. També s'observa que la darrera estructura es caracteritza per tenir el naftil de la ciclodextrina superior en conformació tipus gorra i el naftil de la ciclodextrina inferior totalment inclòs dins de la cavitat superior.



**Figura 8-13 Estructures més representatives del dímer HT en solució aquosa a) als 100 ps b) als 500 ps i c) als 1000 ps.**

### 8.3 ANÀLISI DELS ENLLAÇOS D'HIDROGEN

Per a tots els dímers, i tant en fase gas com en solució aquosa, s'han analitzat el nombre d'enllaços d'hidrogen existents entre les dues unitats de ciclodextrina, considerant que la distància entre els àtoms donador i acceptor ha de ser menor o igual a 3.5 Å i l'angle entre els tres àtoms involucrats en l'enllaç d'hidrogen ha de ser major o igual a 120°.

En les estructures en solució aquosa també s'han estudiat els enllaços d'hidrogen existents entre el solut i el solvent. S'han estudiat dos aspectes diferents: en primer lloc el percentatge d'ocupació en què un hidroxil donat del solut està formant enllaços d'hidrogen amb el solvent i en segon lloc, les funcions de distribució radial per tal de conèixer com es disposa el solvent al voltant del solut.

#### 8.3.1 Anàlisi dels enllaços d'hidrogen en fase gas

En primer lloc, es resumeixen el nombre d'enllaços d'hidrogen totals entre les dues unitats de ciclodextrina en cada un dels dímers (Taula 8-1).

**Taula 8-1 Enllaços d'hidrogen entre ciclodextrines en les simulacions en fase gas.**

|         | <b>HHi</b><br><b>Passos: 20000</b> | <b>HHp</b><br><b>Passos: 20000</b> | <b>HT</b><br><b>Passos: 20000</b> |
|---------|------------------------------------|------------------------------------|-----------------------------------|
| Total   | 154491                             | 155090                             | 63098                             |
| Mitjana | 7.725                              | 7.755                              | 3.155                             |

S'observa que la mitjana d'enllaços d'hidrogen per als dos dímers tipus cap-cap és molt similar i superior a la mitjana per al dímer cap-cua. Això és lògic tenint en compte les conformacions descrites per a cada un d'aquests dímers en fase gas: en el dímer HT les dues ciclodextrines es situen en plans gairebé perpendiculars i conseqüentment, el nombre de glucoses

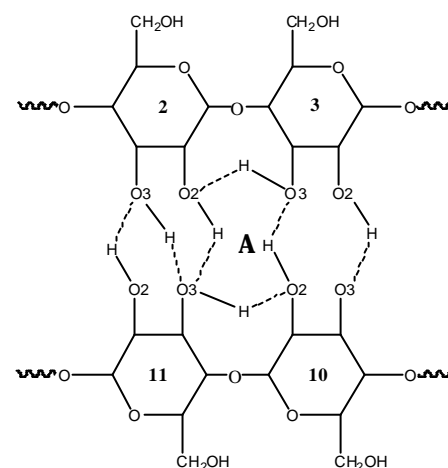
properes entre si per poder formar enllaços d'hidrogen és molt menor que en els dímers cap-cap, on les dues ciclodextrines estan en plans paral·lels i amb un major nombre d'hidroxils encarats que poden formar enllaços d'hidrogen.

A continuació, s'ha aprofundit en l'estudi dels tipus d'enllaços existents. Així doncs, s'han escollit aquells enllaços d'hidrogen entre les dues ciclodextrines que tenen un percentatge d'ocupació més elevat i s'han localitzat els àtoms directament implicats, amb l'objectiu de determinar si existeixen o no conjunts d'enllaços correlacionats entre les dues unitats de ciclodextrina.

El primer dímer estudiat detalladament és el dímer HHi. A la Taula 8-2 es detallen els enllaços d'hidrogen més abundants al llarg de la simulació.

**Taula 8-2 Enllaços d'hidrogen del dímer HHi.**

| DONADOR       |        |              | ACCEPTOR      |        |              | %    |
|---------------|--------|--------------|---------------|--------|--------------|------|
| Número d'àtom | Residu | Tipus d'àtom | Número d'àtom | Residu | Tipus d'àtom |      |
| 29            | 2      | O2           | 220           | 11     | O3           | 60.1 |
| 33            | 2      | O3           | 220           | 11     | O3           | 82.7 |
| 50            | 3      | O2           | 199           | 10     | O3           | 59.0 |
| 195           | 10     | O2           | 54            | 3      | O3           | 69.0 |
| 216           | 11     | O2           | 33            | 2      | O3           | 89.2 |
| 113           | 6      | O2           | 304           | 15     | O3           | 71.4 |
| 134           | 7      | O2           | 283           | 14     | O3           | 70.0 |
| 279           | 14     | O2           | 138           | 7      | O3           | 72.6 |
| 300           | 15     | O2           | 117           | 6      | O3           | 67.7 |



**Figura 8-14 Representació dels enllaços d'hidrogen del dímer HHi.**

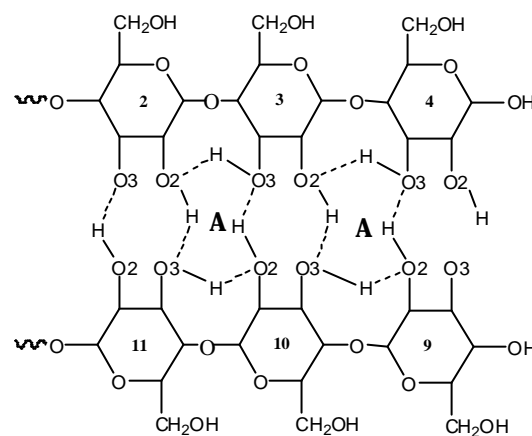
És important observar que la majoria d'enllaços d'hidrogen entre ciclodextrines es donen entre àtoms tipus O2 i O3 (només hi ha un enllaç format per dos àtoms tipus O3). També convé destacar que quan els enllaços d'hidrogen es formen entre les dues ciclodextrines l'àtom donador és un àtom tipus O2 i l'àtom acceptor és un tipus d'àtom O3. No obstant, quan els enllaços es formen dins d'una mateixa ciclodextrina, l'àtom donador és un tipus d'àtom O3 d'una glucosa i l'àtom acceptor és un tipus d'àtom O2 de la glucosa adjacent (no es mostra aquí). Així doncs, combinant aquests dos factors, al final es creen enllaços d'hidrogen correlacionats que formen cicles de diferent nombre d'àtoms. Per exemple, a la Figura 8-14 s'observa que entre les glucoses 2, 3, 10 i 11 es forma un cicle de 8 membres (cicle A) de tipus homodròmic, on tots els enllaços

d'hidrogen van en la mateixa direcció. Però degut a l'existència de l'enllaç d'hidrogen entre els dos àtoms tipus O3 de les glucoses 2 i 11, i degut a què ocasionalment en la simulació també es crea un altre enllaç d'hidrogen entre els àtoms tipus O2 i O3 de la glucosa 2, es forma temporalment un altre cicle de 10 membres també de tipus homodròmic. No obstant, sembla evident que aquest darrer cicle és menys estable que el cicle A de 8 membres.

A continuació, s'ha estudiat el dímer HHp. A la Taula 8-3 es detallen els enllaços d'hidrogen més abundants en aquest dímer.

**Taula 8-3 Enllaços d'hidrogen del dímer HHp.**

| DONADOR       |        |              | ACCEPTOR      |        |              | %    |
|---------------|--------|--------------|---------------|--------|--------------|------|
| Número d'àtom | Residu | Tipus d'àtom | Número d'àtom | Residu | Tipus d'àtom |      |
| 29            | 2      | O2           | 220           | 11     | O3           | 91.7 |
| 50            | 3      | O2           | 199           | 10     | O3           | 79.4 |
| 175           | 9      | O2           | 75            | 4      | O3           | 81.5 |
| 195           | 10     | O2           | 54            | 3      | O3           | 78.4 |
| 216           | 11     | O2           | 33            | 2      | O3           | 71.4 |
| 113           | 6      | O2           | 283           | 14     | O3           | 90.3 |
| 134           | 7      | O2           | 262           | 13     | O3           | 85.8 |
| 258           | 13     | O2           | 138           | 7      | O3           | 86.8 |
| 279           | 14     | O2           | 117           | 6      | O3           | 86.0 |



**Figura 8-15 Representació del enllaços d'hidrogen en el dímer HHp.**

En aquest cas, s'observa que tots els enllaços d'hidrogen entre les dues ciclodextrines es formen entre àtoms tipus O2 i O3 i novament l'àtom tipus O2 és l'àtom donador i l'àtom tipus O3 és l'àtom acceptor. Els enllaços dins d'una mateixa ciclodextrina també tenen com a àtom donador l'àtom tipus O3 i com a àtom acceptor l'àtom tipus O2. És a dir, igual que en el cas anterior, degut a la combinació d'aquests dos factors es tornen a formar cicles de 8 membres d'enllaços d'hidrogen correlacionats del tipus homodròmic. Per exemple, es poden observar a la Figura 8-15 els dos cicles homodròmics entre les glucoses número 2, 3, 10 i 11 i entre les glucoses 3, 4, 9 i 10. En aquest cas, també cal destacar que els enllaços d'hidrogen estan més poblats que en el cas anterior, per tant, sembla que les dues ciclodextrines estan més properes l'una de l'altre en aquest dímer que en el dímer anterior.



Finalment, s'ha realitzat el mateix estudi per al dímer HT. A la Taula 8-4 es detallen els resultats obtinguts.

**Taula 8-4 Enllaços d'hidrogen del dímer HT.**

| DONADOR       |        |              | ACCEPTOR      |        |              | %    | Observacions         |
|---------------|--------|--------------|---------------|--------|--------------|------|----------------------|
| Número d'àtom | Residu | Tipus d'àtom | Número d'àtom | Residu | Tipus d'àtom |      |                      |
| 311           | 15     | O6           | 124           | 6      | O6           | 69.1 | En tota la simulació |
| 300           | 15     | O2           | 65            | 4      | O1           | 4.0  | En els primers ps    |
| 300           | 15     | O2           | 68            | 4      | O5           | 17.4 | En els primers ps    |
| 300           | 15     | O2           | 82            | 4      | O6           | 5.1  | En els primers ps    |
| 199           | 10     | O3           | 61            | 3      | O6           | 47.2 | En els primers ps    |
| 195           | 10     | O2           | 61            | 3      | O6           | 48.8 | En els primers ps    |
| 300           | 15     | O2           | 47            | 3      | O5           | 16.5 | En els últims ps     |
| 300           | 15     | O2           | 44            | 3      | O1           | 7.3  | En els últims ps     |
| 175           | 9      | O2           | 19            | 1      | O6           | 42.3 | En els últims ps     |

En aquest últim cas cal destacar que el nombre i tipus d'enllaços d'hidrogen és força diferent als dos dímers tipus cap-cap. En primer lloc, s'observa que només hi ha un enllaç d'hidrogen que es manté en pràcticament tota la simulació, i aquest és entre dos àtoms tipus O6. La resta d'enllaços d'hidrogen es detecten al principi o al final de la simulació, és a dir, degut al canvi d'orientació de les dues ciclodextrines entre si, canvia el tipus d'enllaços d'hidrogen. Finalment, és important destacar que el residu 15, que és el residu que està parcialment inclòs dins de la cavitat de l'altre ciclodextrina, és el principal responsable dels enllaços d'hidrogen entre les dues ciclodextrines en aquest dímer.

### 8.3.2 Anàlisi dels enllaços d'hidrogen en solució aquosa

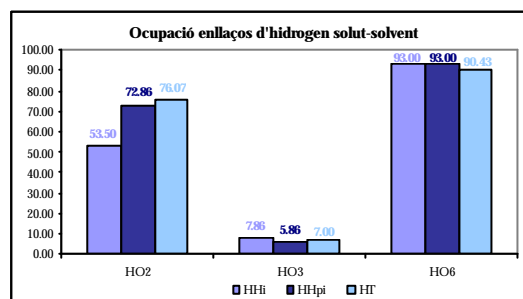
Com ja s'ha comentat, el mateix estudi dels enllaços d'hidrogen s'ha realitzat en les simulacions en solució aquosa. Però, en aquest cas, prèviament a l'anàlisi s'han eliminat les molècules d'aigua i sobre les trajectòries resultants s'han dut a terme les anàlisis dels enllaços d'hidrogen. A la Taula 8-5 es resumeixen els resultats obtinguts.

**Taula 8-5 Enllaços d'hidrogen entre ciclodextrines en les simulacions en solució aquosa.**

|         | HHi<br>Passos: 2000 | HHpi<br>Passos:1- 1000 | HHpi<br>Passos: 1001-2000 | HT<br>Passos: 2000 |
|---------|---------------------|------------------------|---------------------------|--------------------|
| Total   | 9862                | 2248                   | 2183                      | 1301               |
| Mitjana | 4.931               | 1.124                  | 2.183                     | 0.651              |

En primer lloc cal destacar que, en comparació amb les dades obtingudes en fase gas, en aquest cas el nombre d'enllaços d'hidrogen és menor. Aquest fet, és degut a què en solució aquosa també es formen enllaços d'hidrogen entre els hidroxils de les glucoses i les molècules d'aigua i aquí només s'han considerat els enllaços entre les dues ciclodextrines. També és important destacar que en el dímer HHpi durant la primera meitat de la simulació el nombre d'enllaços d'hidrogen entre les dues ciclodextrines és força menor que durant la segona meitat de la simulació, per tant, es confirma que a mesura que avança la simulació les dues ciclodextrines s'apropen entre si. Si es compara el nombre d'enllaços d'hidrogen del dímer HHpi durant els darrers ps de simulació amb el dímer HHi, s'observa que en aquest darrer el nombre d'enllaços d'hidrogen és superior. Per tant, sembla que malgrat tots dos tenen els grups aromàtics dins de la cavitat, la distància entre les dues ciclodextrines és diferent. Finalment, cal destacar que, tal i com ja era previsible, en el dímer cap-cua el nombre d'enllaços d'hidrogen entre les dues ciclodextrines és molt inferior al nombre d'enllaços d'hidrogen en els dímers cap-cap.

A continuació, s'han analitzat els percentatges d'ocupació dels enllaços d'hidrogen entre un àtom donat del solut i el solvent. S'han considerat per una banda els oxígens tipus O2, O3 i O6 com a donadors d'enllaços d'hidrogen i per una altra banda, s'han considerat les molècules d'aigua com a acceptors d'enllaços d'hidrogen. Els resultats es comparen entre els dímers cap-cap i el dímer cap-cua (Taula 8-6).



**Taula 8-6 Ocupació dels enllaços d'hidrogen entre els àtoms tipus O2, O3 i O6 del solut i el solvent.**

En primer lloc, s'observa que els àtoms tipus O2 i O6 tenen percentatges d'ocupació superiors al llarg de la dinàmica i, per tant, aquests actuen preferentment com a donadors d'hidrogen. D'altra banda, l'àtom tipus O3 té un percentatge molt baix d'ocupació i, per tant, aquest actua preferentment com a acceptor d'enllaços d'hidrogen. Si es compara els dos dímers cap-cap amb el dímer cap-cua, s'observa que els oxígens tipus O2 dels dímers cap-cap tenen ocupacions més baixes que el respectiu valor en el dímer cap-cua. Aquest fet no ha de sorprendre donat que el dímer cap-cua té una ciclodextrina amb tots els hidroxils secundaris apuntant cap a

les molècules d'aigua i, en canvi, els dímers cap-cap tenen més tendència a formar enllaços d'hidrogen entre els hidroxils secundaris de les dues ciclodextrines, tal i com s'ha vist en l'estudi dels enllaços d'hidrogen en fase gas. Finalment, es corrobora que el dímer HHpi forma més enllaços d'hidrogen amb l'aigua que el dímer HHi amb el que es confirma que en aquest cas les dues ciclodextrines no es mantenen tan pròximes entre si.

### 8.3.3 Anàlisi de les funcions de distribució radial

Per entendre millor com es disposa el solvent al voltant del solut i per poder determinar quants enllaços d'hidrogen forma cada hidroxil amb el solvent, s'han realitzat les anàlisis de les funcions de distribució radial. S'han estudiat tres tipus de funcions: la  $g_{oo}(r)$  entre O del solut i O del solvent, la  $g_{ho}(r)$  entre H del solut i O del solvent i la  $g_{oh}(r)$  entre O del solut i H del solvent. S'han obtingut les corbes fins a 50 Å, per tal de considerar totes les molècules d'aigua, i s'han analitzat les interaccions en intervals de 0.01 Å.

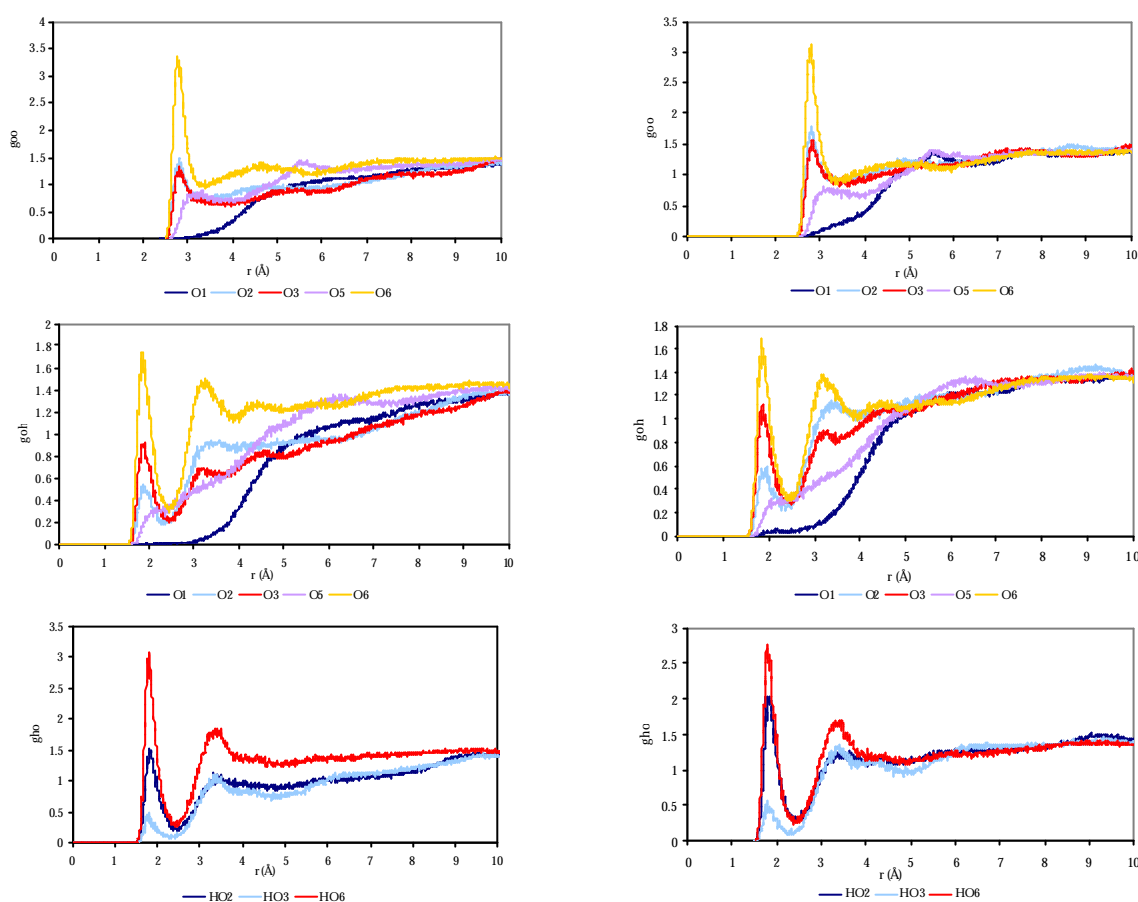
Per integració del primer pic de la funció  $g_{oo}$  fins a 3.5 Å,<sup>69</sup> s'ha obtingut el nombre mitjà de molècules d'aigua que hi ha en la primera capa al voltant d'un hidroxil determinat, és a dir, el nombre de coordinació  $n_{oo}$ . D'altra banda, s'han calculat el nombre d'enllaços d'hidrogen entre un hidroxil determinat i el solvent a partir de les integracions fins a 2.2 Å dels primers pics de les funcions que consideren les interaccions tipus oxigen-hidrogen. El nombre total d'enllaços d'hidrogen,  $n_H$  s'ha calculat com la suma dels nombres de coordinació  $n_{OH} + n_{HO}$ .

Donat que aquesta anàlisi és de caràcter més qualitatiu que les anàlisis anteriors, a la Gràfica 8-9 es representen les funcions de distribució radial estudiades fins a una distància de 10 Å per a un dels dímer cap-cap, el HHi, i per al dímer cap-cua.

Per a les funcions de distribució radial  $g_{oo}(r)$  s'observa que els oxígens tipus O2, O3, i O6 dels dos dímers són els que tenen al seu voltant una capa ben definida de molècules d'aigua, i que aquesta està centrada al voltant de 2.9 Å; aquest fet està d'acord amb l'anàlisi anterior i confirma que aquests àtoms són els que formen enllaços d'hidrogen amb el solvent. D'altra banda, en les funcions corresponents als oxígens tipus O5 s'observa un pic molt més ample centrat al voltant de 3.1 Å i, per tant, en aquest cas no s'aprecia una capa definida de molècules d'aigua al seu voltant. Finalment, s'observa que els oxígens tipus O1 no tenen molècules de solvent properes,

<sup>69</sup> En el cas dels àtoms tipus O6 s'ha integrat fins a 3.6 Å.

fet que és coherent tenint en compte que aquests oxígens de les ciclodextrines apunten cap a l'interior de la cavitat, la qual és de caràcter hidrofòbic.



**Gràfica 8-9** De dalt a baix funcions de distribució radial  $g_{oo}$ ,  $g_{oh}$ , i  $g_{ho}$  per als dímers HHi (esquerra) i HT (dreta).

Respecte a les gràfiques corresponents a les funcions que representen les interaccions oxigen-hidrogen,  $g_{oh}$  i  $g_{ho}$ , s'observen dues capes de molècules de solvent en tots els casos per als oxígens tipus O2, O3 i O6: la primera al voltant de 1.9 Å i la segona al voltant dels 3.2 Å. A més, igual que s'ha esmentat en el cas de les funcions tipus  $g_{oo}$ , ni en l'oxigen tipus O1 ni en l'oxigen tipus O5 s'aprecia una capa definida de molècules d'aigua.

Finalment, a la Taula 8-7 es resumeixen el nombre de molècules d'aigua calculades a partir de les integrals per a cada un dels casos. En les funcions de distribució radial que corresponen a les interaccions tipus oxigen-hidrogen, on s'ha vist que hi ha dos pics, s'han calculat les integrals dels primers pics, ja que només es poden formar enllaços d'hidrogen amb les molècules de solvent de la primera capa.

**Taula 8-7 Nombre de coordinació,  $n_{oo}$ , i nombre d'enllaços d'hidrogen de cada oxigen.**

|                  |    | $n_{oo}$ | $n_{HO}$ | $n_{OH}$ | $n_H$ |
|------------------|----|----------|----------|----------|-------|
| Cap-cap<br>(HHI) | O1 | 0.20     | -        | 0.00     | 0.00  |
|                  | O2 | 3.89     | 2.51     | 0.97     | 3.48  |
|                  | O3 | 3.58     | 0.77     | 1.63     | 2.40  |
|                  | O5 | 2.54     | -        | 0.51     | 0.51  |
|                  | O6 | 7.41     | 4.32     | 2.85     | 7.17  |
|                  | O1 | 0.44     | -        | 0.07     | 0.07  |
| Cap-cua<br>(HT)  | O2 | 4.87     | 3.38     | 1.08     | 4.46  |
|                  | O3 | 4.56     | 0.98     | 1.98     | 2.96  |
|                  | O5 | 2.48     | -        | 0.50     | 0.50  |
|                  | O6 | 7.35     | 4.25     | 2.81     | 7.06  |

De la taula cal destacar en primer lloc el nombre de molècules d'aigua mitjà que té cada hidroxil al seu voltant ( $n_{oo}$ ). S'observa que l'oxigen O1 té major nombre de molècules d'aigua al seu voltant en el cas de dímer cap-cua, que en el cas del dímer cap-cap. A partir d'aquests valors s'ha estimat que el nombre de molècules dins de les cavitats de les ciclodextrines: tres molècules en el cas del dímer cap-cap, i sis molècules en el dímer cap-cua. Aquesta diferència es pot explicar perquè en el dímer cap-cua tan sols és una de les ciclodextrines la que té un grup aromàtic dins de la cavitat i, per tant, l'altre cavitat està lliure per a què hi puguin entrar les molècules del solvent. No obstant, aquests dos valors són força baixos degut al caràcter hidrofòbic de la cavitat. D'altra banda, els hidroxils secundaris tenen quatre molècules d'aigua al seu voltant en el dímer cap-cap, però en el dímer cap-cua en tenen una més. Aquest fet es pot explicar novament perquè aquest darrer dímer té una ciclodextrina amb tots els hidroxils secundaris apuntant cap al solvent. Respecte als oxígens tipus O5 s'observa que en els dos tipus de dímers tenen entre 2 i 3 molècules d'aigua al seu voltant i en tots dos dímers són els àtoms tipus O6 els que, amb diferència, tenen més molècules d'aigua al seu voltant (al voltant de 7).

La darrera columna de la taula mostra el nombre total d'enllaços d'hidrogen que forma cada hidroxil amb les molècules de solvent, que recordem que s'ha calculat a partir de la suma  $n_{OH} + n_{HO}$ . S'observa que en cap dímer els oxígens tipus O1 formen enllaços d'hidrogen amb el solvent. En tots dos dímers els oxígens tipus O2 formen aproximadament un enllaç d'hidrogen més amb l'aigua que els oxígens tipus O3 i s'observa que en promig els hidroxils secundaris formen entre 2 i 4 enllaços d'hidrogen amb el solvent. Els oxígens O5 només ocasionalment formen enllaços d'hidrogen amb el solvent (una mitjana de 0.5 enllaços d'hidrogen al llarg de tota la simulació). I finalment, s'observa que els hidroxils primaris en tots dos casos formen al voltant de 7 enllaços d'hidrogen amb el solvent.

## 8.4 ESTUDI DELS TERMES ENERGÈTICS

Per a cada un dels dímers, i tant en fase gas com en solució aquosa, s'han estudiat els diferents termes energètics que contribueixen a l'energia total del sistema amb l'objectiu de confirmar quin és el dímer més estable. Els diferents termes que contribueixen a l'energia total es detallen a la Taula 8-8 (entre parèntesis s'inclouen les desviacions associades a aquests valors).

**Taula 8-8 Termes energètics de totes les estructures en fase gas i solució aquosa (en kcal/mol).**

|             | Fas gas      |              |              | Solució aquosa |              |              |
|-------------|--------------|--------------|--------------|----------------|--------------|--------------|
|             | HHi          | HHp          | HT           | HHi            | HHpi         | HT           |
| <b>ENER</b> | 799.3 (12.9) | 810.7 (12.7) | 813.1 (12.9) | 771.1 (16.4)   | 769.7 (18.6) | 802.8 (16.3) |
| <b>BOND</b> | 84.4 (7.9)   | 84.7 (7.9)   | 84.6 (8.0)   | 94.1 (7.9)     | 95.4 (8.0)   | 94.7 (8.0)   |
| <b>ANGL</b> | 204.2 (11.0) | 203.7 (10.8) | 208.1 (11.0) | 208.5 (10.9)   | 208.5 (11.6) | 208.5 (11.5) |
| <b>DIHE</b> | -413.4 (4.8) | -409.3 (4.5) | -412.8 (4.9) | -427.8 (5.2)   | -427.4 (6.1) | -425.8 (5.7) |
| <b>IMPR</b> | 3.6 (1.2)    | 3.7 (1.2)    | 3.6 (1.1)    | 3.6 (1.2)      | 3.7 (1.2)    | 3.4 (1.1)    |
| <b>VDW</b>  | 84.4 (8.0)   | 100.3 (7.7)  | 98.9 (7.7)   | 84.0 (8.0)     | 89.8 (8.3)   | 101.9 (8.1)  |
| <b>ELEC</b> | 836.0 (7.9)  | 827.7 (7.8)  | 830.7 (8.1)  | 808.7 (14.4)   | 799.7 (16.2) | 820.0 (13.3) |

De la taula convé destacar en primer lloc que segons les energies totals en fase gas sembla que l'estabilització relativa dels tres dímers segueix l'ordre:

$$\text{HHi} > \text{HHp} > \text{HT}$$

El dímer HHi és 11.5 kcal/mol més estable que el dímer HHp, i 18.5 kcal/mol més estable que el dímer HT. Aquesta millor estabilitat del dímer HHi respecte dels altres dos prové d'un millor terme de l'energia de van der Waals ja que els dos anells es situen just al centre de la cavitat i, per tant, s'optimitzen les interaccions no enllaçants entre els grups aromàtics i les ciclodextrines. L'estabilització d'aquest dímer també prové d'un millor terme de l'energia de torsió respecte dels altres dos. No obstant, cal destacar que les interaccions electrostàtiques són més favorables en el dímer HHp que en els altres dos dímers. Aquest fet pot ser degut en primer lloc, a què en aquest dímer hi ha un major nombre de cicles formats per enllaços d'hidrogen correlacionats i en segon lloc, a unes millors interaccions entre els hidrògens de les ciclodextrines i l'anell aromàtic, que es situa entre ambdues.

D'altra banda en les estructures obtingudes en solució aquosa s'observa que l'estabilització segueix l'ordre:

$$\text{HH} > \text{HT}$$

És a dir, els dímers cap-cap són més estables energèticament que el dímer cap-cua. Aquesta estabilització és d'aproximadament 30 kcal/mol i prové majoritàriament d'una millor energia dels termes corresponents a les interaccions entre àtoms no enllaçats.

Finalment, s'han calculat les diferències entre els diferents termes energètics en fase gas i solució aquosa. Els resultats es resumeixen a la Taula 8-9.

**Taula 8-9 Diferència dels termes energètics en solució aquosa i fase gas (en kcal/mol).**

| Aquo – gas             | ENER  | BOND | ANGL | DIHE  | IMPR | VDW  | ELEC  |
|------------------------|-------|------|------|-------|------|------|-------|
| HHi                    | -28.2 | 9.7  | 4.2  | -14.4 | 0.0  | -0.4 | -27.3 |
| HHpi-HHi <sup>70</sup> | -29.6 | 11.0 | 4.3  | -14.1 | 0.0  | 5.5  | -36.3 |
| HT                     | -10.3 | 10.2 | 0.3  | -13.0 | -0.2 | 3.0  | -10.6 |

S'observa que tots els dímers són més estables energèticament en solució aquosa que en fase gas, degut majoritàriament a un millor terme de les interaccions electrostàtiques. L'estabilització és al voltant de 30 kcal/mol per als dímers amb els grups aromàtics a l'interior de la cavitat i només al voltant d'unes 10 kcal/mol per al dímer cap-cua.

## 8.5 RESUM DELS RESULTATS

En resum, podem dir que el dímer cap-cua és el menys estable tant en fase gas com en dissolució aquosa, i en el cas del dímer HHp en solució aquosa s'ha observat un gran efecte hidrofòbic entre els dos anells aromàtics de manera que, s'ha detectat el pas d'un dímer tipus HHp a un dímer tipus HHi. Així doncs, entre els tres possibles dímers de la 3-O-(2-metilnaftil)- $\beta$ -ciclodextrina el HHi és el més estable degut a un millor terme entròpic i també entàlpic (11.5 kcal/mol més estable en fase gas i 30 kcal/mol més estable en dissolució aquosa), sent els termes energètics corresponents a les torsions i a les interaccions de van der Waals els que més l'afavoreixen. A més, en els dímers cap-cap s'han detectat cicles d'enllaços d'hidrogen correlacionats del tipus homodròmic, que proporcionen estabilitat a aquests respecte del dímer cap-cua, on no s'observen aquests tipus de cicles. Finalment, s'ha estudiat com es disposa el solvent al voltant del solut i les diferències més importants s'han manifestat al voltant dels àtoms tipus O2 i O3, que en general tenen més molècules de solvent al seu voltant en el dímer HT que en els dímers HH.

<sup>70</sup> Aquests valors corresponen a la resta dels valors obtinguts en solució aquosa per al dímer HHpi menys els valors obtinguts en fase gas per al dímer HHi, ja que d'aquesta manera es comparen estructures similars.

## 9 EFECTE D'UNA $\beta$ -CICLODEXTRINA MODIFICADA EN LA DEGRADACIÓ DE L'HSP70 PER LA TRIPSINA

### 9.1 INTRODUCCIÓ AL PROBLEMA

Les ciclodextrines naturals i modificades han estat tradicionalment utilitzades com a amfitrions o *hosts*, incloent una àmplia diversitat de molècules dins de la seva cavitat. No obstant, són escassos els estudis on les ciclodextrines interaccionen amb sistemes més grans i, per tant, on elles actuen com a hostes o *guests*. Per exemple, al 1995<sup>71</sup> es van publicar dades experimentals de les interaccions entre l'antril-(alquilamino)- $\beta$ -ciclodextrina i el DNA. Uns anys més tard, al 1999,<sup>72</sup> es va publicar la inclusió d'una ciclodextrina en el canal format per l' $\alpha$ -hemolisina. I fa només tres anys, al 2001,<sup>73</sup> es van publicar dades experimentals sobre la influència de la 6<sup>A</sup>,6<sup>D</sup>-bis(dansil-Gly)- $\beta$ -ciclodextrina<sup>74</sup> (Figura 9-1) en l'activitat xaperona de dues proteïnes de xoc tèrmic, i en la degradació d'aquestes per part de la tripsina.

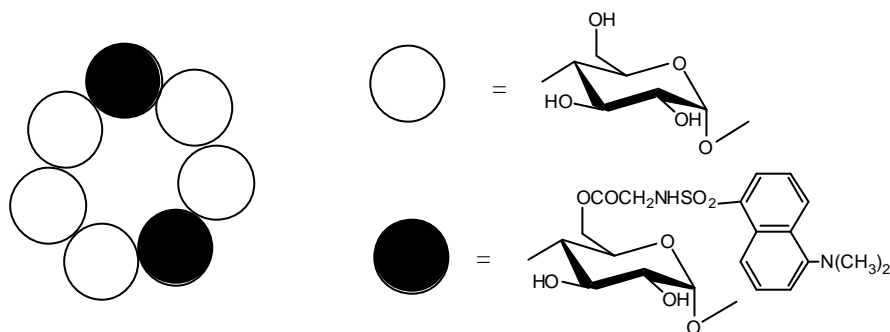


Figura 9-1 Representació de la 6<sup>A</sup>,6<sup>D</sup>-bis(dansil-Gly)- $\beta$ -ciclodextrina.

Més concretament, en aquesta darrera publicació es descriuen estudis de fluorescència per avaluar si existeix complexació o no entre la ciclodextrina modificada amb dos grups dansil i les Hsp70 i Hsp90. El descens en intensitat de l'espectre de fluorescència de la ciclodextrina

<sup>71</sup> Ikeda, T.; Yoshida, K.; Schneider, H.J., *J. Am. Chem. Soc.*, **1995**, 117(4), 1453-1454.

<sup>72</sup> (a) Gu, L.Q.; Braha, O.; Conlan, S.; Cheley, S.; Bayley, H., *Nature*, **1999**, 398(6729), 686-690. (b) Gu, L.Q.; Cheley, S.; Bayley, H., *Science*, **2001**, 291(5504), 636-640.

<sup>73</sup> Hamada, F.; Narita, M.; Makabe, A.; Itoh, H., *Journal of Incl. Phen. and Macro. Chem.*, **2001**, 40, 83-88.

<sup>74</sup> Gly és una abreviatura acceptada per a la glicina, a més, s'estableix que quan està substituïda pel grup COOH es simbolitza: Gly-, quan està substituïda pel grup NH<sub>2</sub> es simbolitza: -Gly, i quan està substituïda per ambdós grups es simbolitza -Gly-. Per a més informació consultar González Alcaraz, F; *Nomenclatura de Química Orgànica*, Universidad de Murcia: Murcia, **1991**.



suggereix que hi ha una complexació *host-guest*, i a més, està descrit que aquest descens és degut a què un dels grups  $-Gly$ -dansil surt fora de la cavitat.<sup>75</sup> D'altra banda, s'estudia la influència d'aquesta ciclodextrina modificada en la degradació de la proteïna per part de la tripsina. Aquest estudi es duu a terme mitjançant electroforesi en gel de SDS-poliacrilamida i s'observa diferent comportament segons si hi ha o no ATP complexat amb la proteïna i segons si hi ha o no la  $\beta$ -ciclodextrina modificada (BCDm) complexada amb la proteïna. En la present tesi doctoral, ens hem centrat principalment en aquest darrer estudi.

Experimentalment està descrit que amb una concentració de 0.01 mg/mL de tripsina s'obtenen diferents fraccions en la degradació de l'Hsp70 segons el substrat que estigui complexat amb la proteïna (Taula 9-1).

**Taula 9-1 Fraccions obtingudes en la degradació de l'Hsp70 bobina amb 0.01 mg/mL de tripsina.**

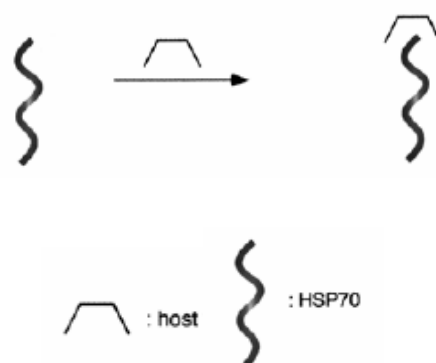
| Condicions         | Fraccions obtingudes (kDa) |
|--------------------|----------------------------|
| Hsp70              | 68, 60, 33, 25             |
| Hsp70 + BCDm       | 68, 60, 33, 25             |
| Hsp70 + ATP        | 68, 60, 45, 43, 38, 33, 25 |
| Hsp70 + BCDm + ATP | 68, 60, 33, 25             |

De la taula es conclou que, en absència d'ATP, l'Hsp70 s'hidrolitza parcialment i, que en les mateixes condicions la ciclodextrina modificada no influeix en la degradació, ja que s'obtenen les mateixes bandes. D'altra banda, s'observa que l'ATP és el que millor promou la degradació de l'Hsp70 i finalment, s'observa que en presència de l'ATP, la  $\beta$ -ciclodextrina inhibeix parcialment la hidròlisi. Per tant, la ciclodextrina modificada afecta a la degradació de l'Hsp70 front a la proteassa en presència d'ATP.

Per explicar aquestes observacions, els autors postulen que hi ha una complexació entre la ciclodextrina modificada i la proteïna com la que es mostra a la Figura 9-2; la primera actua com a *host* i la segona actua com a *guest*. No obstant, considerant les grandàries relatives de la ciclodextrina i de la proteïna, nosaltres proposem com a primera opció que enlloc de formar un complex d'inclusió es podria formar un complex d'associació de manera que, la ciclodextrina

<sup>75</sup> (a) Narita, M.; Hamada, F.; Sato, M.; Suzuki, I.; Osa, T., *J. Incl. Phen. and Macro. Chem.*, **1999**, 34, 421-430. (b) Sato, M.; Narita, M.; Ogawa, N.; Hamada, F., *Anal. Sci.*, **1999**, 15, 1199.

estaria bloquejant alguna de les posicions de la proteïna on podria haver alguna Lys o Arg i, per tant, estaria impedit que la tripsina pogués efectuar la hidròlisi en aquella zona.



**Figura 9-2 Modes de complexació entre la ciclodextrina modificada i l'Hsp70. Figura extreta de *J. Incl. Phen. and Macr. Chem.*, 2001, 40, 83-88.**

En segon lloc, considerant que diverses publicacions evidencien que les ciclodextrines inclouen amb facilitat diversos nucleòtids en la seva cavitat,<sup>76</sup> també podria ser que la ciclodextrina modificada actués com a inhibidora de l'ATP, impedit que aquest pogués dur a terme la seva activitat. No obstant, per iniciar l'estudi ens hem centrat en la primera teoria postulada. Així doncs, tenint en compte aquests precedents, en aquest capítol es presenten, en primer lloc, la cerca dels possibles punts de degradació de l'Hsp70 humana considerant la seva estructura primària i l'estructura secundària del seu domini N-terminal (codi pdb: 1HJO). A continuació, es presenten els resultats de l'anàlisi conformacional de la 6<sup>A</sup>,6<sup>D</sup>-bis(dansil-Gly)- $\beta$ -ciclodextrina, amb l'objectiu de conèixer quina és la conformació més estable d'aquesta quan està aïllada. S'han realitzat simulacions de dinàmica i mecànica molecular amb el programa AMBER i utilitzant el camp de força GAFF. Posteriorment, es detallen els resultats obtinguts de les simulacions de dinàmica molecular en solució aquosa de l'Hsp70 i de l'Hsp70 complexada amb l'ATP, amb l'objectiu de detectar si l'ATP provoca canvis conformacionals importants en la proteïna. Aquestes simulacions també s'han realitzat amb el programa AMBER però s'ha utilitzat el camp de força parm99. Finalment, es presenten les anàlisis de *docking* realitzades amb el

<sup>76</sup> (a) Eliseev, A.V.; Schneider, H.-J., *J. Am. Chem. Soc.*, **1994**, 116(14), 6081-6088. (b) Ikeda, T.; Yoshida, K.; Schneider, H.-J., *J. Am. Chem. Soc.*, **1995**, 117(4), 1453-1454. (c) Suzuki, I.; Miura, T.; Anzai, J-I., *J. Supramolecular Chemistry*, **2001**, 1, 283-288.

programa Autodock 3.0.5,<sup>77</sup> on la ciclodextrina modificada actua com a lligand i la proteïna com a receptor, amb l'objectiu de trobar els possibles punts d'interacció de la proteïna.

## 9.2 CERCA DELS POSSIBLES PUNTS D'HIDRÒLISI DE LA PROTEÏNA

Com ja s'ha esmentat, per trobar els possibles punts d'hidròlisi s'ha agafat com a model l'estructura primària de l'Hsp70 humana i l'estructura secundària del seu domini N-terminal (Figura 9-3).

```

10      20      30      40      50      60
MAKAAAIGID LGTT YSCVGV FQHGKVEIIA NDQGNRTTPS YVAF TDTERL IGDAAKNQVA
70      80      90      100     110     120
LNPNQNTVFD A KRLIGRKFGD PVVQSDMKHW PFQVINDGDK PKVQVSYKGE TKAFYPEEIS
130     140     150     160     170     180
SMVLT KMKEL AEAYL GYPVT NAVITVP AYF NDSQRQAT KD AGVIAGL NVL RIINEPTAAA
190     200     210     220     230     240
IAYGLDRTGK GERNVLI FDL GGGT FDVSL TIDDGIFE VK ATAGDTHLG ED FDNRLVNH
250     260     270     280     290     300
FVEEFKRKHK KDISQNKRAV RRLRTACERA KRTLSSSTQA SLEIDSLFEG IDFYTSITRA
310     320     330     340     350     360
RFEELCSDLF RST LEPVEKA LRDAKLDKKAQ IHDLVLV GGS TRIPKVQKLL QDFFNGRDLN
370     380     390     400     410     420
KSINPDEAVA YGAAVQAAIL MGDKSENVQD LLLLDVAPLS LGLETAGGVM TALIKRNSTI
430     440     450     460     470     480
PTKQTQIFIT YSDNQPGVLI QVYEGERAMT KDNLLGRFE LSGIPPAPRG VPQIEVTFDI
490     500     510     520     530     540
DANGILNVTA TDKSTGKANK ITITNDKGR LSKEEIERMVQ EAEKYKAEDE VQRERVSANK
550     560     570     580     590     600
ALESYAFNMK SAVEDEGLKG KISEADKKKV LDKCQEVISW LDANTLAEKD EFEHKRKELE
610     620     630     640
QVCNPIISGL YQGAGGPGPG GFGAQGPKGG SGSGPTIEEV D

```

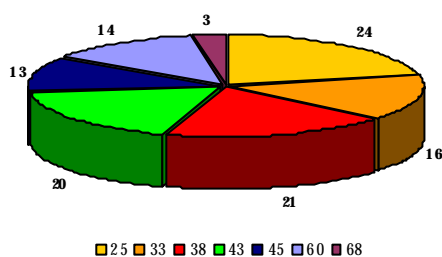
**Figura 9-3 Estructura primària de l'Hsp70 humana i estructura secundària del domini N-terminal. En groc estan marcades les fulles  $\beta$ , en verd les hèlixs  $\alpha$  i en blau els girs. Els aminoàcids dels quals no es coneix la seva estructura secundària no estan colorats. Els residus d'arginina (R) i lisina (K) es ressalten en negreta.**

S'han buscat tots els possibles punts d'hidròlisi de l'Hsp70 amb el programa Peptide Cutter,<sup>78</sup> i tenint en compte que hi pot haver un o més punts d'hidròlisi dins de la proteïna, s'han calculat totes les combinacions que proporcionen fraccions de 25, 33, 38, 43, 45, 60, i 68 kDa. No obstant, amb les fraccions que provenen pel trencament en un i dos punts ja es cobreixen totes les altres. A continuació, s'han imposat restriccions ja que el nombre de combinacions és molt elevat. Així doncs, considerant que la tripsina és un enzim que hidrolitza estructures secundàries poc organitzades, s'han seleccionat tots aquells residus d'arginina i lisina que la seva estructura secundària o bé sigui un gir (color blau), o bé no estigui identificada (sense marcar), i també

<sup>77</sup> Morris, G.M.; Goodsell, D.S.; Halliday, R.S.; Huey, R.; Hart, W.E.; Belew, R.K.; Olson, A.J., *J. Comput. Chem.*, **1998**, *19*(14), 1639-1662.

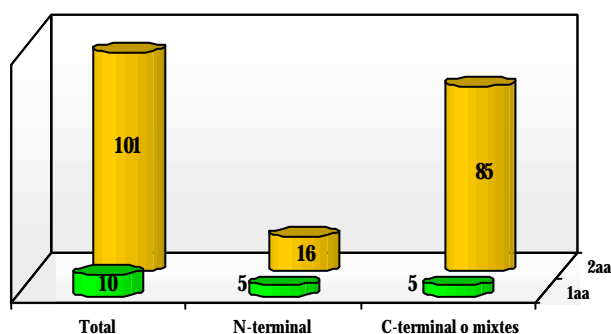
<sup>78</sup> <http://www.expasy.org/tools/peptidecutter>

aquelles arginines o lisines que són el primer aminoàcid d'una hèlix  $\alpha$  o d'una fulla  $\beta$ . Per calcular totes les possibles combinacions i per seleccionar aquelles que ens interessin, hem emprat un programa de combinatòria agafat d'Internet<sup>79,80</sup> (comb.f), i dins del nostre grup d'investigació hem preparat un nou programa fortran (sumafrags.f) amb els que, partint de les dades obtingudes amb el programa Peptide Cutter, s'obté un fitxer de sortida on s'especifica: la massa de la fracció, si aquesta prové de la hidròlisi en un únic punt de la proteïna o de la hidròlisi en dos punts diferents, quina de les dues o de les tres fraccions és la que compleix tots els requisits, i l'aminoàcid o aminoàcids concrets pels quals s'ha produït la hidròlisi. Ambdós programes es detallen en l'Apèndix. En resum, considerant aquestes característiques s'han detectat 111 combinacions possibles que proporcionen les fraccions de la Gràfica 9-1.



**Gràfica 9-1 Nombre i massa de totes les fraccions de la hidròlisi de l'Hsp70 humana.**

A partir d'aquestes dades, s'ha aprofundit en l'anàlisi i s'han considerat, per una banda, les combinacions que provenen del trencament per un únic aminoàcid i, per una altra banda, les combinacions que provenen de la hidròlisi per dos aminoàcids. Els resultats obtinguts es representen a la Gràfica 9-2.



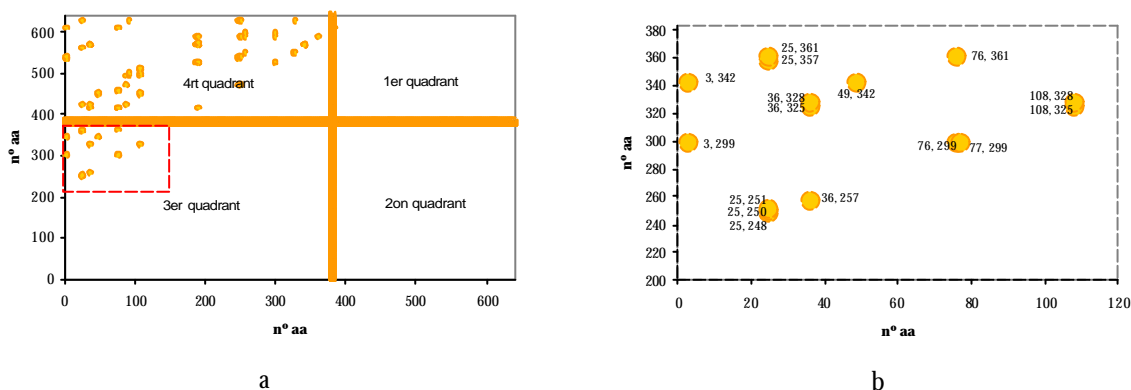
**Gràfica 9-2 Classificació de les combinacions obtingudes.**

<sup>79</sup> <http://lib.stat.cmu.edu/apstat>.

<sup>80</sup> Algorithm AS 88: *Applied Statistics Algorithms Editors*, 1975, 24(3).

En primer lloc, és important observar que hi ha 10 combinacions on la hidròlisi es dona per un únic punt o aminoàcid, i 101 combinacions on la hidròlisi es dona per dos punts diferents. Entre les primeres, hi ha 5 aminoàcids que estan dins del domini N-terminal i els altres 5 estan dins del domini C-terminal. D'altra banda, quan la fracció prové de la hidròlisi en dos punts diferents, hi ha 16 combinacions en què els dos aminoàcids estan dins del domini N-terminal, i en la resta de combinacions, que s'han anomenat mixtes, un dels aminoàcids està dins d'un domini i l'altre aminoàcid està dins de l'altre domini. A més, no hi ha cap combinació en què els dos aminoàcids estiguin dins del domini C-terminal perquè aquest té una massa de 25 kDa, i les fraccions que s'han cercat com a mínim són d'aquesta massa.

A la Gràfica 9-3a es representen totes aquestes combinacions de dos aminoàcids. En primer lloc és important esmentar que el domini N-terminal està comprès fins a l'aminoàcid 381 i el domini C-terminal a partir d'aquest en endavant. Així doncs, no es detecten combinacions de dos aminoàcids que estiguin dins del domini C-terminal (1er quadrant), però sí que s'observen algunes combinacions en què els dos aminoàcids estan dins del domini N-terminal (3er quadrant). A més, al quart quadrant, que corresponen a les combinacions mixtes, és on hi ha major nombre de punts. Aprofitant el mateix gràfic, s'ha ampliat la zona del tercer quadrant per poder caracteritzar millor totes les combinacions de dos aminoàcids que estan dins del domini N-terminal (Gràfica 9-3b).



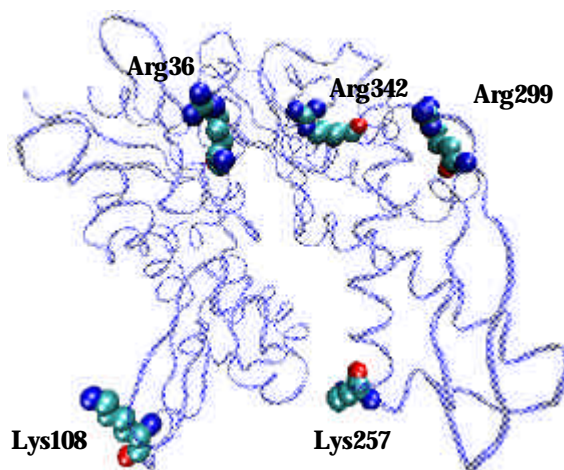
**Gràfica 9-3 a) Combinacions d'hidròlisi per dos punts de l'Hsp70 humana i b) Ampliació de les combinacions dins del domini N-terminal.**

Així doncs, un cop conegudes totes les combinacions possibles, per iniciar l'estudi ens hem centrat en primer lloc en aquelles fraccions que provenen del trencament per un únic punt. A la Taula 9-2 es detallen els punts d'hidròlisi, així com les fraccions en kDa que proporciona el trencament per cada punt concret.

**Taula 9-2 Fraccions provinents de la hidròlisi per un únic punt.**

| Domini     | Aminoàcid | Estructura Secundària | Fragment (kDa) |
|------------|-----------|-----------------------|----------------|
| N-terminal | Arg36     | no definit            | 3.9 + 67.9     |
|            | Lys108    | gir                   | 12.2 + 59.6    |
|            | Lys257    | 1er aa hèlix          | 28.8 + 43.0    |
|            | Arg299    | 1er aa hèlix          | 33.7 + 38.1    |
|            | Arg342    | 1er aa hèlix          | 38.7 + 33.1    |
| C-terminal | Lys384    | -                     | 43.3 + 28.5    |
|            | Lys415    | -                     | 46.5 + 25.3    |
|            | Arg416    | -                     | 46.7 + 25.1    |
|            | Arg533    | -                     | 59.9 + 11.9    |
|            | Arg535    | -                     | 60.3 + 11.5    |

S'observa que només en el domini N-terminal hi ha una fracció de 68 kDa. D'altra banda les fraccions de 43 kDa es troben en ambdós dominis i a més, en el domini N-terminal es troben les fraccions de 33 i 38 kDa. La fracció de 25 kDa molt probablement correspon al domini C-terminal i finalment, és important observar que en cap dels casos es detecta la fracció de 45 kDa. Així doncs, en resum podem dir que considerant un únic punt d'hidròlisi dins del domini N-terminal hi ha cinc punts claus de possibles degradacions que podrien correspondre a les bandes observades experimentalment. Aquests punts estan definit per l'Arg36, la Lys108, la Lys257, l'Arg299, i l'Arg342 (Figura 9-4) i són els que s'utilitzaran per als càlculs de *docking*.



**Figura 9-4 Punts d'hidròlisi dins del domini N-terminal per obtenir les fraccions de 68 (Arg36), 60 (Lys108), 43 (Lys257), 38 (Arg299) i 33 (Arg342) kDa.**

La mateixa caracterització s'ha realitzat per a les combinacions de dos aminoàcids dins del domini N-terminal (Taula 9-3). En aquest cas, convé destacar que totes les fraccions obtingudes són entre 25 i 38 kDa, per tant, novament s'observa l'absència de la fracció de 45 kDa dins

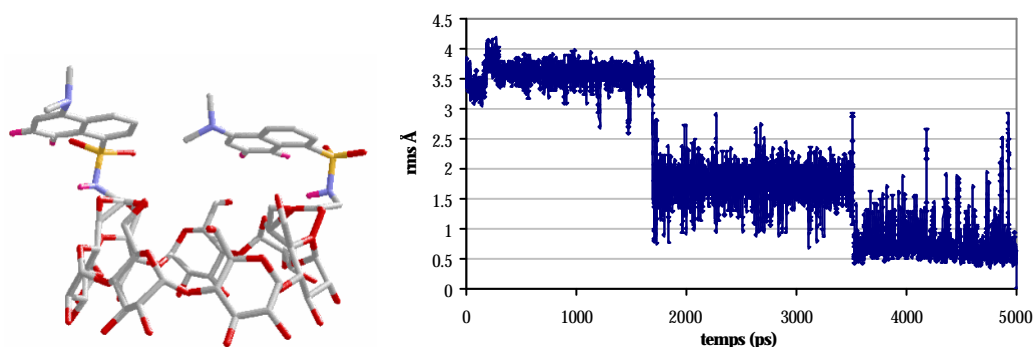
d'aquest domini. Així doncs, postulem que molt probablement aquesta fracció prové d'una de les combinacions que hem anomenat mixtes.

**Taula 9-3 Fraccions provinents de la hidròlisi per dos aminoàcids dins del domini N-terminal.**

| 1er aa | 2on aa | Fragment (kDa)    | 1er aa | 2on aa | Fragment (kDa)     |
|--------|--------|-------------------|--------|--------|--------------------|
| Lys3   | Arg299 | 0.3 + 33.4 + 38.1 | Arg36  | Lys325 | 3.9 + 32.9 + 35.0  |
| Lys3   | Arg342 | 0.3 + 38.3 + 33.1 | Arg36  | Lys328 | 3.9 + 33.3 + 34.6  |
| Lys25  | Lys248 | 2.7 + 25.0 + 44.1 | Arg49  | Arg342 | 5.4 + 33.3 + 33.1  |
| Lys25  | Lys250 | 2.7 + 25.3 + 43.9 | Arg76  | Arg299 | 8.4 + 25.3 + 38.1  |
| Lys25  | Lys251 | 2.7 + 25.4 + 43.7 | Arg76  | Lys361 | 8.4 + 32.6 + 30.8  |
| Lys25  | Arg357 | 2.7 + 37.8 + 31.3 | Lys77  | Arg299 | 8.5 + 25.2 + 38.1  |
| Lys25  | Lys361 | 2.7 + 38.3 + 30.8 | Lys108 | Lys325 | 12.2 + 24.6 + 35.0 |
| Arg36  | Lys257 | 3.9 + 24.9 + 43.0 | Lys108 | Lys328 | 12.2 + 25.0 + 34.6 |

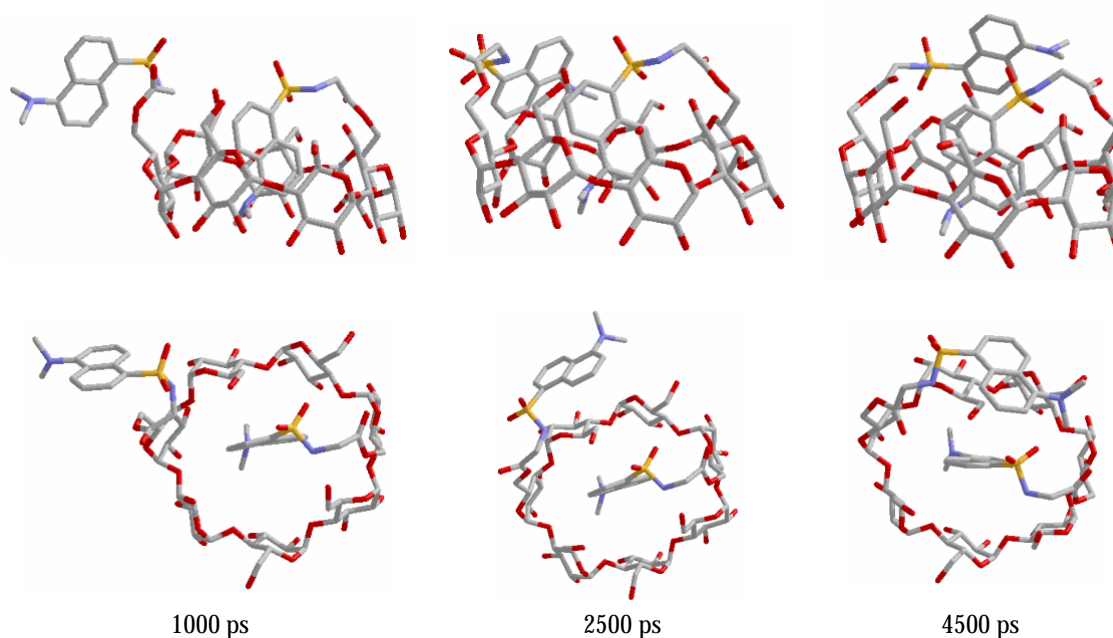
### 9.3 ANÀLISI CONFORMACIONAL DE LA 6<sup>A</sup>,6<sup>D</sup>-BIS(DANSIL-GLY)- $\beta$ -CICLODEXTRINA

S'ha construït la 6<sup>A</sup>,6<sup>D</sup>-bis(dansil-Gly)- $\beta$ -ciclodextrina amb la interfase gràfica del programa LeAP i, després de minimitzar l'estructura inicial, s'ha realitzat la simulació de dinàmica molecular en fase gas durant 5 ns (prèvies etapes d'escalfament i d'equilibració). Un cop obtinguda la trajectòria s'ha realitzat l'anàlisi del RMS considerant l'última estructura com a referència i fent una primera superposició dels carbonis de les glucoses i una segona superposició amb tots els àtoms pesats del sistema (C, O, N i S). La variació del RMS al llarg del temps es representa a la Figura 9-5 Estructura de partida de la 6<sup>A</sup>,6<sup>D</sup>-bis(dansil-Gly)- $\beta$ -ciclodextrina per a les simulacions de dinàmica molecular Gràfica 9-4.



**Figura 9-5 Estructura de partida de la 6<sup>A</sup>,6<sup>D</sup>-bis(dansil-Gly)- $\beta$ -ciclodextrina per a les simulacions de dinàmica molecular Gràfica 9-4 Variació del RMS al llarg del temps en la simulació de dinàmica molecular de la 6<sup>A</sup>,6<sup>D</sup>-bis(dansil-Gly)- $\beta$ -ciclodextrina.**

S'aprecia durant els primers piconsegons un primer conjunt d'estructures minoritari, i a continuació tres conjunts d'estructures que corresponen a tres conformacions diferents de la ciclodextrina. Per avaluar les diferències s'han escollit com a model les estructures als 1000, 2500, i 4500 ps. S'observa a la Figura 9-6 que la diferència entre les conformacions està en la posició del primer grup dansil-Gly-. En la primera estructura aquest es situa perpendicular a l'eix de la cavitat i a la part externa d'aquesta. En la segona conformació el grup dansil-Gly- es manté a la part externa però més proper a l'exterior del con truncat. Finalment, en la darrera estructura aquest es situa es forma de gorra a sobre del con truncat. D'altra banda, s'observa a la figura que en les tres estructures el segon grup dansil-Gly- es situa a l'interior de la cavitat i la seva posició no varia significativament d'un conformer a l'altre.



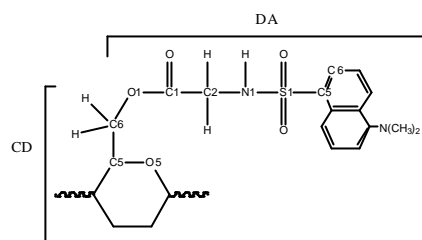
**Figura 9-6 Vista lateral i vista superior de les estructures als 1000, 2500 i 4500 ps de la simulació de dinàmica molecular de la 6<sup>A</sup>,6<sup>D</sup>-bis(dansil-Gly)- $\beta$ -ciclodextrina.**

S'han caracteritzat les tres conformacions estudiant els diedres pertanyents a la cadena del grup ancorat, amb l'objectiu d'utilitzar els valors d'aquests diedres com a model per avaluar les diferències conformacionals que es produeixen quan aquesta ciclodextrina interacciona amb l'Hsp70. A Taula 9-4 es resumeixen els valors.



**Taula 9-4 Valors mitjans dels diedres de la 6<sup>A</sup>,6<sup>D</sup>-bis(dansil-Gly)- $\beta$ -ciclodextrina.**

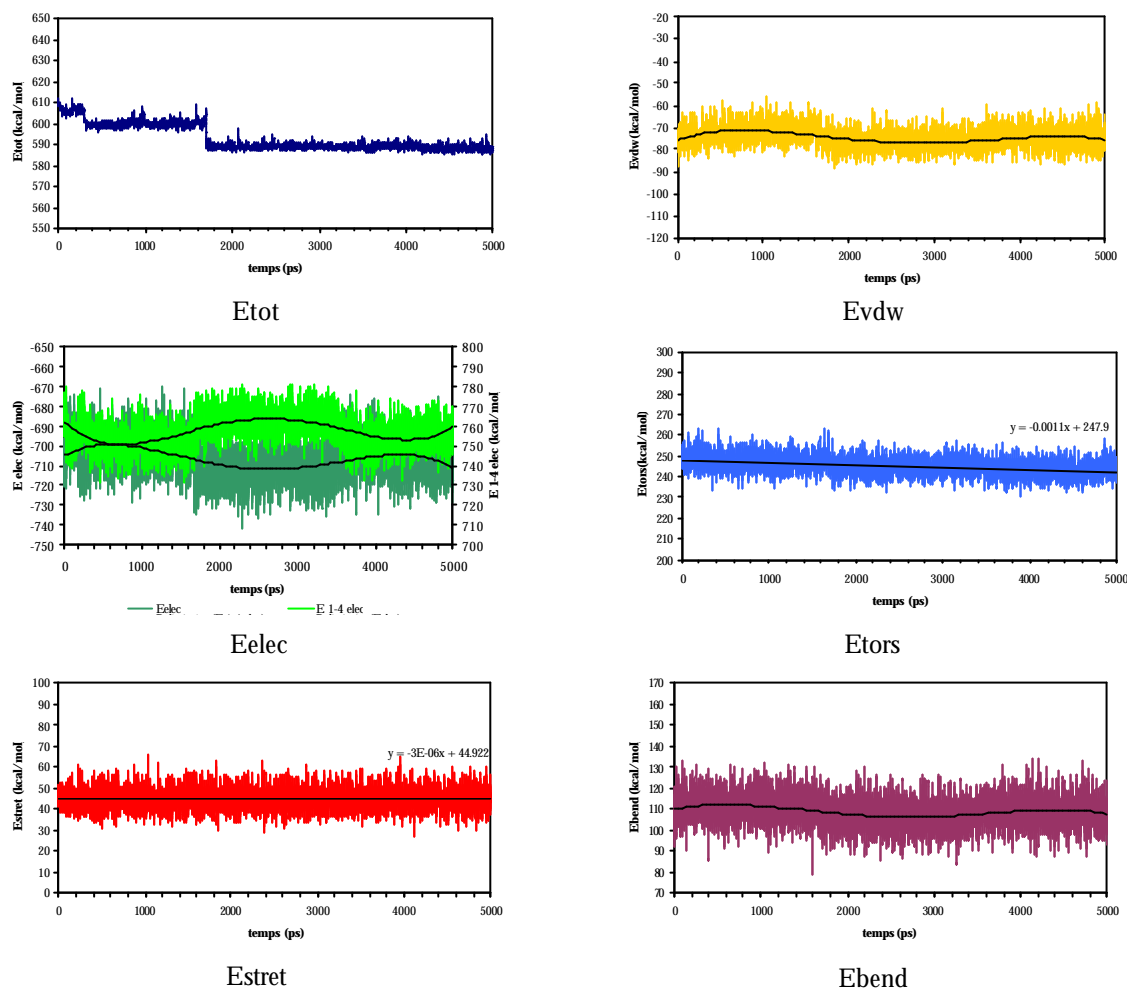
|                             | DA1    |       | DA2    |       |
|-----------------------------|--------|-------|--------|-------|
|                             | diedre | desv. | diedre | desv. |
| C6(DA)-C5(DA)-S1(DA)-N1(DA) | 119.9  | 8.6   | 106.9  | 9.9   |
| C5(DA)-S1(DA)-N1(DA)-C2(DA) | -150.4 | 42.7  | 166.5  | 18.5  |
| S1(DA)-N1(DA)-C2(DA)-C1(DA) | 87.2   | 48.8  | -83.4  | 29.6  |
| N1(DA)-C2(DA)-C1(DA)-O1(DA) | -93.2  | 56.5  | 120.4  | 18.7  |
| C2(DA)-C1(DA)-O1(DA)-C6     | 179.9  | 9.5   | 178.5  | 6.9   |
| C1(DA)-O1(DA)-C6(CD)-C5(CD) | 78.3   | 44.3  | 107.2  | 13.4  |
| O1(DA)-C6(CD)-C5(CD)-O5(CD) | 118.5  | 28.9  | 41.2   | 11.0  |



S'observa com, efectivament, les desviacions associades als valors dels diedres del segon grup dansil-Gly- (DA2) són força més baixes que les respectives desviacions per al primer grup dansil-Gly- (DA1). A més, sembla que els cavis conformacionals es produeixen majoritàriament per rotació del enllaços S1(DA)-N1(DA), N1(DA)-C2(DA), C2(DA)-CA(DA), i O1(DA)-C6(CD).

A continuació, s'ha avaluat com varien els diferents termes energètics al llarg de la simulació (Gràfica 9-5). S'observa en l'energia total, que les dues últimes conformacions tenen una energia pràcticament igual, no obstant, la primera conformació descrita té una energia de gairebé 10 kcal/mol superior a les altres dues. Aquesta diferència prové principalment del terme de l'energia de van der Waals i de l'energia de torsió, que disminueix progressivament al llarg de la simulació. A més, és important observar que la segona i la tercera conformació es diferencien principalment en les interaccions electrostàtiques: la segona conformació té interaccions electrostàtiques entre els àtoms en posició relativa 1,4 força superiors a la tercera conformació, i en canvi l'energia electrostàtica total és inferior en aquesta que en la darrera conformació. Per últim, també es poden apreciar diferències en l'energia de *bending*, que és lleugerament superior en la primera conformació que en les altres dues.

Finalment, s'ha minimitzat l'energia de cada una de les estructures per mètodes de mecànica molecular i s'han obtingut valors de les energies de 232.7, 322.8 i 185.6 kJ/mol per a les estructures als 1000, 2500, i 4500 ps respectivament. Per tant, tenint en compte aquests valors s'ha escollit la darrera conformació com a lligand per realitzar els càlculs de *docking*.



**Gràfica 9-5 Evolució dels diferents termes energètics al llarg del temps per a la 6A,6D-bis(dansil-Gly)-β-ciclodextrina.**

## 9.4 ANÀLISI CONFORMACIONAL DE L'HSP70 SENSE I AMB ATP COMPLEXAT

S'han realitzat simulacions de dinàmica molecular de l'Hsp70 sola i complexada amb l'ATP en dissolució aquosa emprant condicions periòdiques de contorn a pressió constant. Les simulacions s'han dut a terme durant 1000 ps amb un *timestep* de 2 fs, i amb les prèvies etapes d'escalfament i equilibració del sistema.<sup>81</sup> Per seguir les explicacions del següents apartats és important recordar que l'Hsp70 està formada per dos subdominis globulars (I i II) separats per una gran escletxa central.

<sup>81</sup> Per a més detalls consultar l'Apèndix

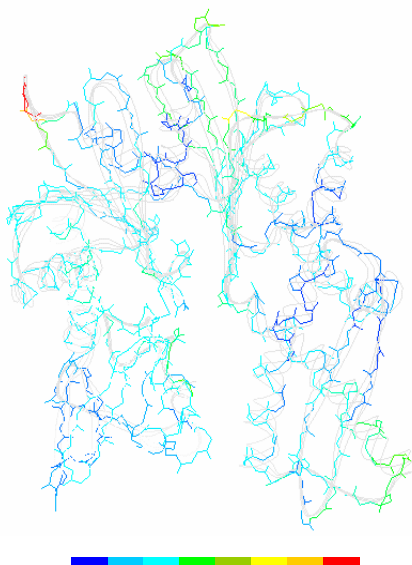
### 9.4.1 Anàlisi estructural de l'Hsp70 sense i amb ATP complexat

En primer lloc, s'ha calculat el valor del RMS respecte la primera estructura de la simulació superposant els àtoms de la cadena principal (o *backbone*) (Taula 9-5). Dels valors de la taula s'aprecia una diferència de 0.034 Å en el valor del RMS segons si l'ATP està o no complexat, per tant, sembla que no es produeixen canvis conformacionals significatius en presència de l'ATP.

**Taula 9-5 RMS per a l'Hsp70 sense i amb ATP complexat.**

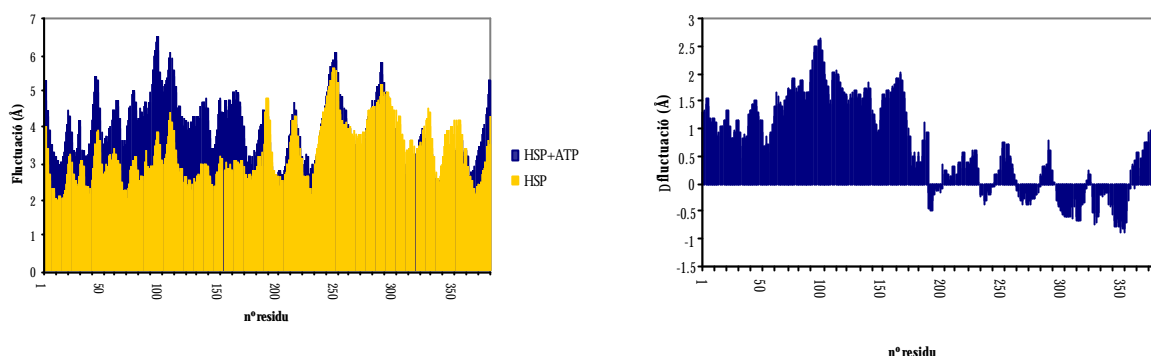
|             | <b>Avg.</b> | <b>Dev.</b> | <b>Max.</b> | <b>Min.</b> |
|-------------|-------------|-------------|-------------|-------------|
| Hsp70       | 1.553       | 0.145       | 2.062       | 1.196       |
| Hsp70 + ATP | 1.519       | 0.112       | 1.951       | 1.127       |

No obstant, per confirmar aquest fet, s'han superposat les dues estructures obtingudes a partir de la mitjana de les coordenades. A la Figura 9-7 es representa l'esquelet de l'Hsp70 complexada amb l'ATP, on el color de cada aminoàcid depèn del seu RMS respecte de la mateixa estructura de l'Hsp70 sola (els aminoàcids de color vermell són els que més es diferencien i els de color blau fosc els que menys). S'observa que les desviacions més importants estan en els primers i últims aminoàcids de la proteïna, i en alguns dels *loops*. Però és conegut que la modelització d'aquestes estructures secundàries per simulacions de dinàmica molecular és molt difícil, i normalment aquests esdevenen molt flexibles al llarg de la simulació. Per tant, es conclou que no es detecten canvis conformacionals significatius entre una estructura i l'altre.



**Figura 9-7 RMS del *backbone* de l'Hsp70 complexada amb l'ATP.**

Finalment, s'han estudiat les fluctuacions atòmiques de la cadena principal, i s'ha representat la mitjana dels resultats per a cada residu (Gràfica 9-6).

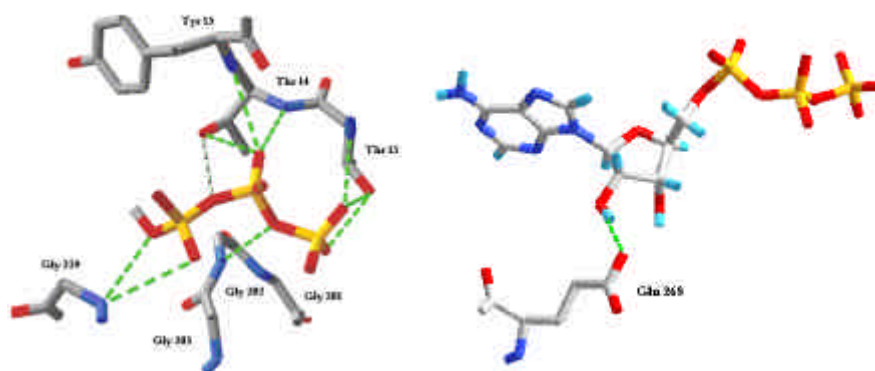


**Gràfica 9-6 Esquerra: Fluctuacions atòmiques per residu en l'Hsp70 i l'Hsp70 + ATP. Dreta: Diferències entre les fluctuacions atòmiques.**

En aquest cas, s'observa que els perfils de les fluctuacions són pràcticament idèntics en els dos casos. Això és degut a què el perfil depèn de l'estructura secundària, i aquesta no varia d'una simulació a l'altra. No obstant, el que diferencia les dues simulacions és la magnitud de la fluctuació. Així doncs, es pot apreciar que la presència de l'ATP provoca canvis conformacionals més importants en l'estructura, i més concretament, dins del primer subdomini. D'altra banda, s'ha realitzat la mateixa anàlisi considerant tots els àtoms de l'estructura (no es mostra aquí), i en ambdós casos s'han observat majors fluctuacions en les cadenes laterals que en la cadena principal.

#### 9.4.2 Anàlisi geomètrica de l'Hsp70 sense i amb ATP complexat

En primer lloc, s'ha avaluat com interaccionen l'Hsp70 i l'ATP (Figura 9-8). S'han detectat deu enllaços d'hidrogen, on la proteïna actua com a donadora i l'ATP com a acceptor i només un, on l'ATP actua com a donador i la proteïna com a acceptora. A la figura es pot apreciar que al voltant de l'ATP hi ha dos conjunts de tres aminoàcids que són els que estan directament implicats en la formació dels enllaços d'hidrogen i a part, la Gly339 també actua com a donadora.

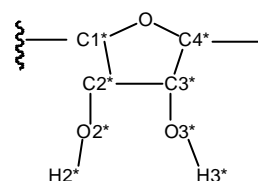
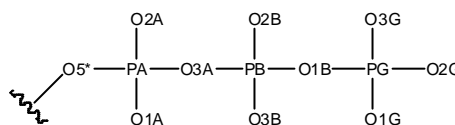


**Figura 9-8 Representació dels enllaços d'hidrogen entre l'Hsp70 i l'ATP.**

Aquests mateixos enllaços d'hidrogen s'han caracteritzat emprant el mòdul Carnal d'AMBER (Taula 9-6).

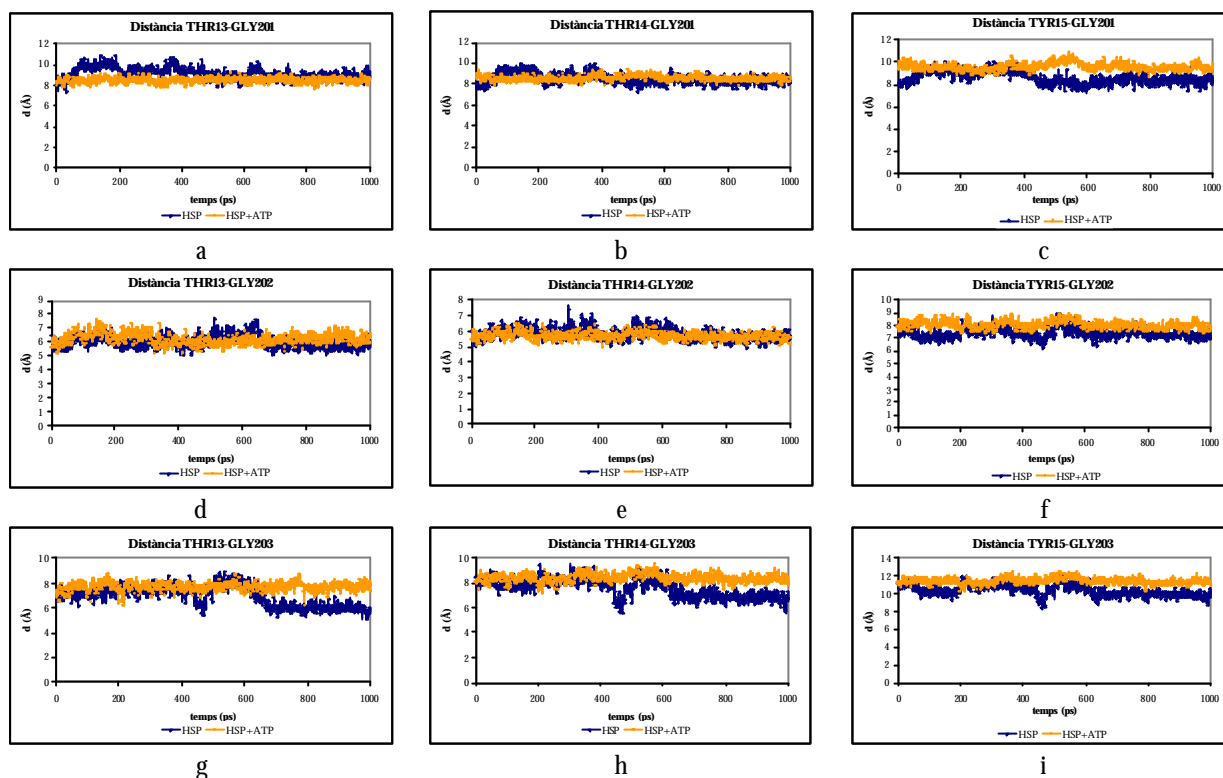
**Taula 9-6 Enllaços d'hidrogen entre l'Hsp70 i l'ATP.**

| Àtom donador | Àtom acceptor | % ocupació |       |
|--------------|---------------|------------|-------|
| Thr13 OG1    | Thr13 HG1     | atp O2G    | 99.1  |
| Thr13 OG1    | Thr13 HG1     | atp O3G    | 81.8  |
| Thr13 N      | Thr13 H       | atp O3G    | 98.7  |
| Thr14 OG1    | Thr14 HG1     | atp O2B    | 100.0 |
| Thr14 OG1    | Thr14 HG1     | atp O3A    | 95.1  |
| Thr14 N      | Thr14 H       | atp O2B    | 100.0 |
| Thr15 N      | Thr15 H       | atp O2B    | 100.0 |
| Gly 202 N    | Gly 202 H     | atp O3B    | 95.0  |
| Gly 339 N    | Gly 339 H     | atp O2A    | 100.0 |
| Gly 339 N    | Gly 339 H     | atp O5*    | 97.3  |
| atp O2*      | atp H2'       | Glu268 OE1 | 89.8  |



S'observa que en tots els casos hi ha percentatges d'ocupació molt elevats (superiors al 80%), per tant, es pot dir que gràcies a aquestes interaccions es manté la geometria de l'esqueletxa al llarg de la simulació. D'altra banda, s'observa que els àtoms de l'ATP implicats en la formació d'enllaços d'hidrogen són principalment els oxígens dels grups fosfat en la posició  $\beta$  (tipus B), i ? (tipus G).

Finalment, tenint en compte els residus directament implicats en la formació dels enllaços d'hidrogen amb l'ATP, s'han estudiat com varien les distàncies entre ambdós grups d'aminoàcids al llarg de la simulació. Els resultats es representen a la Gràfica 9-7.



**Gràfica 9-7** Evolució de les distàncies al llarg del temps a) Thr13-Gly201, b) Thr14-Gly201, c) Tyr15-Gly201, d) Thr13-Gly202, e) Thr14-Gly202, f) Tyr15-Gly202, g) Thr13-Gly203, h) Thr14-Gly203, i) Tyr15-Gly203.

L'evolució de les distàncies al llarg del temps mostra que, en general, quan l'ATP està present les distàncies són una mica superiors que quan aquest no hi és. A més, és important observar que al voltant dels 600 ps de la simulació es produeixen els canvis conformacionals més destacats. En aquest moment, les distàncies Thr13-Gly203 i Thr14-Gly203 disminueixen lleugerament en la simulació de la proteïna sola. Per avaluar les diferències numèricament a la Taula 9-7 es mostren els valors mitjans de les distàncies mitjanes al llarg de les simulacions.

**Taula 9-7** Distàncies en Å entre carbonis de la cadena principal.

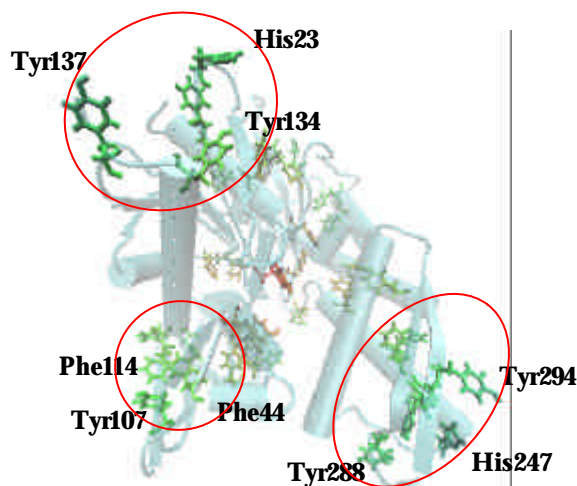
|              | Hsp70 |      |       |      | Hsp70+ATP |      |       |       | ? d   |
|--------------|-------|------|-------|------|-----------|------|-------|-------|-------|
|              | avg   | dev  | max   | min  | avg       | dev  | max   | min   |       |
| Thr13-Gly201 | 9.17  | 0.57 | 10.83 | 7.35 | 8.47      | 0.24 | 9.19  | 7.71  | -0.70 |
| Thr14-Gly201 | 8.51  | 0.48 | 9.96  | 7.29 | 8.52      | 0.23 | 9.29  | 7.90  | 0.01  |
| Tyr15-Gly201 | 8.63  | 0.55 | 10.05 | 7.36 | 9.57      | 0.32 | 10.84 | 8.67  | 0.94  |
| Thr13-Gly202 | 5.99  | 0.45 | 7.69  | 5.10 | 6.21      | 0.39 | 7.62  | 5.16  | 0.22  |
| Thr14-Gly202 | 5.82  | 0.38 | 7.60  | 4.50 | 5.68      | 0.24 | 6.47  | 4.92  | -0.14 |
| Tyr15-Gly202 | 7.40  | 0.39 | 8.97  | 6.28 | 8.02      | 0.29 | 8.95  | 7.39  | 0.62  |
| Thr13-Gly203 | 6.95  | 0.83 | 8.89  | 5.13 | 7.67      | 0.33 | 8.78  | 6.19  | 0.72  |
| Thr14-Gly203 | 7.59  | 0.74 | 9.44  | 5.58 | 8.37      | 0.32 | 9.48  | 7.17  | 0.78  |
| Tyr15-Gly203 | 10.37 | 0.63 | 12.20 | 8.44 | 11.37     | 0.33 | 12.44 | 10.19 | 1.00  |

És important destacar la diferència dels valors de les distàncies entre les dues estructures: aquesta mostra que l'amplada de la cavitat varia en un màxim 1.0 Å segons si hi ha o no l'ATP complexat, per tant, la diferència és suficientment petita com per confirmar que l'ATP tampoc influeix en la grandària i forma de l'esclatxa.

## **9.5 PUNTS DE COMPLEXACIÓ ENTRE LA CICLODEXTRINA I LA PROTEÏNA**

En aquest darrer apartat és pretén trobar el complex ciclodextrina-proteïna de mínima energia. Donat que no s'han trobat canvis conformacionals significatius entre l'Hsp70 sola i aquesta complexada amb l'ATP, s'ha escollit la darrera estructura de la primera com a receptor i el confòrmer de mínima energia de la ciclodextrina modificada com a lligand.

S'ha emprat el programa Autodock, s'ha mantingut el receptor fix, i s'han buscat les diverses posicions, orientacions, i conformacions del lligand al voltant d'un "centre actiu". S'han girat un màxim de 16 enllaços del lligand (els corresponents a les cadenes dels grups ancorats) en intervals de 5°, les traslacions s'han realitzat en increments de 0.2 Å, i per modificar l'orientació del lligand s'ha considerat aquest com un cos rígid. Per a l'elecció del centre actiu, s'han tingut en compte dos criteris diferents: s'han seleccionat com a centres actius els aminoàcids que proporcionen les fraccions per un únic punt de trencament descrit al principi d'aquest capítol. I en segon lloc, considerant que les ciclodextrines inclouen amb facilitat dins de la seva cavitat grups aromàtics, s'han buscat zones de la proteïna on hi hagi un elevat nombre de grups aromàtics (Figura 9-9). S'han localitzat tres zones concretes que s'han definit pels tres triangles següents: Phe114-Tyr107-Phe44 (arom1), Tyr137-His23-Tyr134 (arom2), i Tyr294-His249-Tyr288 (arom3).

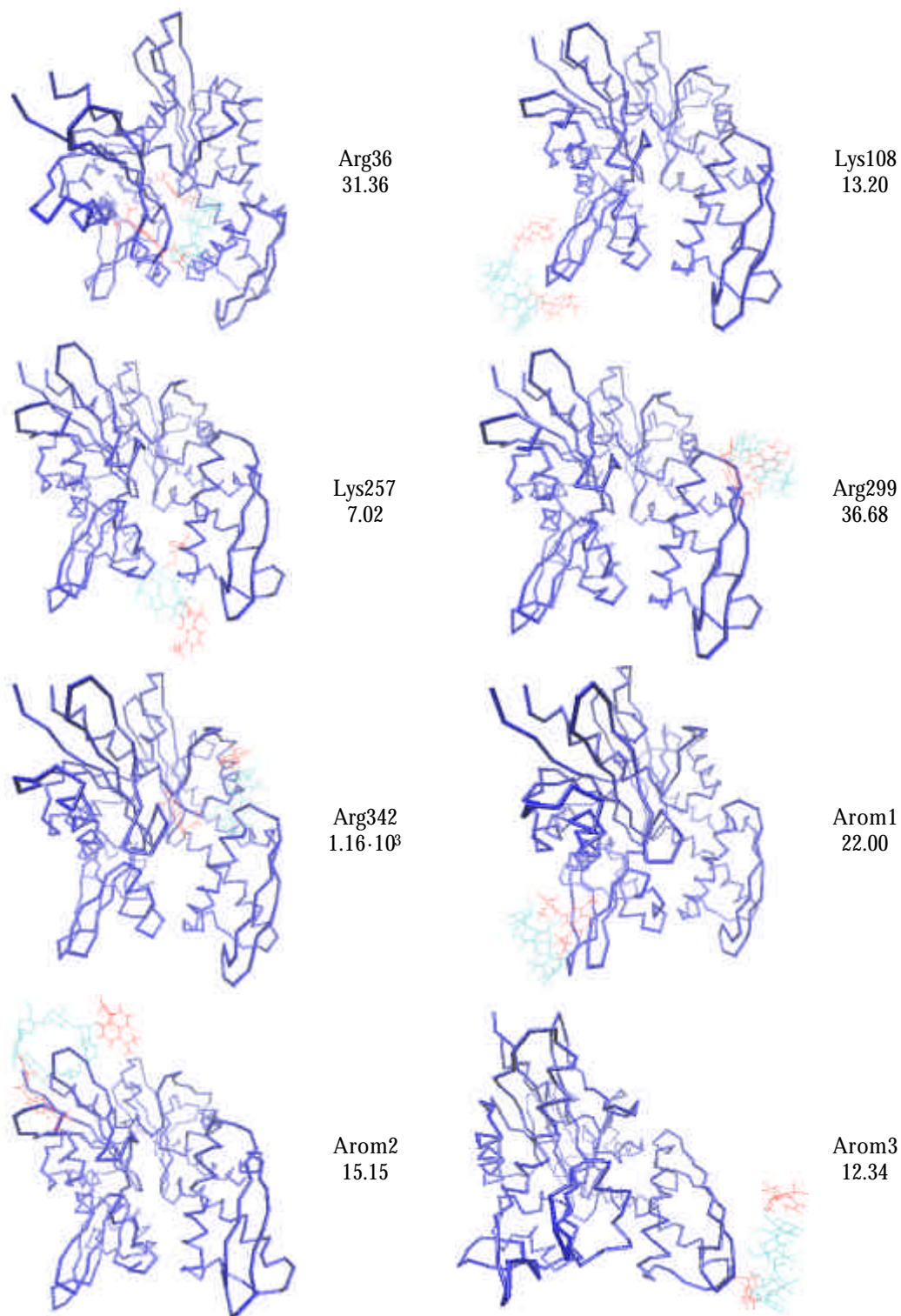


**Figura 9-9 Grups aromàtics de l'Hsp70, colorats segons la seva accessibilitat.**

Per tant, en total s'han considerat vuit “centres actius” de la proteïna, al voltant dels quals s'ha realitzat l'anàlisi de *docking*. Els càlculs al voltant de cinc d'aquests centres actius permetran confirmar si la ciclodextrina bloqueja alguna de les posicions de la proteïna o no, i els altres tres permetran rebutjar la hipòtesi postulada pels autors que hi ha una part de la proteïna que s'inclou dins de la cavitat de la ciclodextrina.

A la Figura 9-10 es representa els complexos de mínima energia obtinguts al voltant de cada un dels centres actius. L'esquelet de la ciclodextrina es representa en color blau clar, i els grups ancorats es representen de color vermell, per identificar amb claredat la seva posició. És important destacar que el complex de mínima energia és aquell en què la ciclodextrina està al voltant de la Lys257. Cal recordar, que per hidròlisi en aquest punt s'obté una fracció de 43 kDa, però només s'obté en presència de l'ATP, ja que quan s'afegeix la ciclodextrina al complex Hsp70-ATP s'inhibeix la hidròlisi i no s'obté aquesta fracció. Per tant, aquest fet confirma que la ciclodextrina podria estar bloquejant una determinada posició de la proteïna i no permetria que la tripsina hidrolitzés per aquest punt, i a més, no sembla que la ciclodextrina inclogui cap part de la proteïna.





**Figura 9-10 Complexos de mínima energia al voltant de cada centre actiu, a la dreta especificació del “centre actiu” i energia en kcal/mol.**

Finalment, és important observar, que en la majoria dels casos s’han produït canvis conformacionals importants en la ciclodextrina i que els dos grups –Gly-dansil es troben fora de la cavitat (excepte al voltant de l’Arg299 on hi ha un dels grups a l’interior de la cavitat). Per tant,

aquests grups ancorats interaccionen directament amb diversos grups de la proteïna. A la Taula 9-8 es detallen els valors dels diedres de la cadena del grup ancorat per a cada un dels complexos. Comparant aquests valors amb els valors de la ciclodextrina aïllada es confirmen les observacions experimentals: la ciclodextrina aïllada té un dels grups ancorats dins de la cavitat, però quan interacciona amb la proteïna ambdós surten d'aquesta.

**Taula 9-8 Valors dels diedres del grup – Gly-dansil en els diferents complexos.**

|                             | Arg36  | Lys108 | Lys257 | Arg299 | Arg342 | Arom1  | Arom2  | Arom3  |
|-----------------------------|--------|--------|--------|--------|--------|--------|--------|--------|
| DA1                         |        |        |        |        |        |        |        |        |
| C6(DA)-C5(DA)-S1(DA)-N1(DA) | -68.8  | -25.4  | 30.0   | 98.9   | -60.1  | -146.1 | 36.7   | -109.9 |
| C5(DA)-S1(DA)-N1(DA)-C2(DA) | 162.7  | -155.6 | 86.4   | 97.4   | -49.0  | 47.3   | 55.2   | -27.5  |
| S1(DA)-N1(DA)-C2(DA)-C1(DA) | 36.7   | -55.8  | 76.3   | 131.1  | 118.6  | 132.8  | 11.9   | 163.8  |
| N1(DA)-C2(DA)-C1(DA)-O1(DA) | 9.3    | 134.9  | 52.0   | 165.0  | 153.3  | 146.5  | 74.4   | -136.6 |
| C2(DA)-C1(DA)-O1(DA)-C6     | 74.8   | 109.9  | -177.3 | -58.5  | 16.2   | 160.3  | 32.6   | 132.9  |
| C1(DA)-O1(DA)-C6(CD)-C5(CD) | 15.8   | 106.7  | -40.7  | -136.4 | 126.4  | 86.5   | 163.0  | -112.1 |
| O1(DA)-C6(CD)-C5(CD)-O5(CD) | -55.2  | 140.7  | 163.6  | 10.7   | 122.2  | -150.5 | -172.0 | 6.1    |
| DA2                         |        |        |        |        |        |        |        |        |
| C6(DA)-C5(DA)-S1(DA)-N1(DA) | -72.7  | -27.8  | -95.1  | 164.8  | 123.0  | 115.5  | -117.2 | -101.9 |
| C5(DA)-S1(DA)-N1(DA)-C2(DA) | 113.4  | 152.7  | 35.7   | 84.7   | -44.4  | -3.8   | -109.9 | -162.8 |
| S1(DA)-N1(DA)-C2(DA)-C1(DA) | -14.0  | 105.1  | 67.5   | 123.9  | 115.0  | 72.1   | -110.0 | 131.8  |
| N1(DA)-C2(DA)-C1(DA)-O1(DA) | 99.9   | -83.4  | 162.4  | 39.7   | 168.7  | -111.5 | -8.1   | -156.2 |
| C2(DA)-C1(DA)-O1(DA)-C6     | -140.4 | 88.3   | -106.0 | 64.8   | -35.3  | -33.9  | -103.3 | -61.9  |
| C1(DA)-O1(DA)-C6(CD)-C5(CD) | -154.3 | 136.2  | 100.4  | -6.1   | -105.1 | -4.3   | 101.5  | 161.6  |
| O1(DA)-C6(CD)-C5(CD)-O5(CD) | -142.6 | -30.1  | -79.7  | 81.1   | 20.1   | 100.7  | -60.6  | 169.6  |

Finalment, a la Taula 9-9 es desglossa l'energia de *docking* en els diferents termes energètics (Energia intermolecular i energia interna del lligand) i a més, s'inclou el nombre d'enllaços d'hidrogen que hi ha entre el receptor i el lligand.

**Taula 9-9 Energia de *docking*, energia intermolecular i energia interna del lligand en kcal/mol per a cada complex. Nombre d'enllaços d'hidrogen receptor-lligand.**

| Centre actiu | E <sub>docking</sub> | E <sub>intramolecular</sub> | E <sub>int</sub> (ligand) | Nº enllaços H |
|--------------|----------------------|-----------------------------|---------------------------|---------------|
| Arg36        | 31.36                | 7.51                        | 23.86                     | 5             |
| Lys108       | 13.20                | 3.35                        | 9.84                      | 0             |
| Lys257       | 7.02                 | -3.43                       | 10.45                     | 1             |
| Arg299       | 36.68                | 14.55                       | 22.13                     | 3             |
| Arg342       | 1.16·10 <sup>3</sup> | 987.08                      | 175.27                    | 13            |
| Aromàtics 1  | 22.00                | 5.33                        | 16.67                     | 1             |
| Aromàtics 2  | 15.15                | 6.41                        | 8.74                      | 0             |
| Aromàtics 3  | 16.49                | 7.36                        | 9.13                      | 3             |

De la taula es conclou que el complex al voltant de la Lys257 és el més estable energèticament gràcies a un millor terme de l'energia intermolecular entre els dos substrats. No obstant, per a aquest complex, la conformació del lligand no és la de mínima energia, sinó que aquesta és menor en el cas del complex format al voltant de les zones d'aromàtics denominades arom2 i arom3. A més, dels valors d'energia també es conclou que al voltant de l'Arg342 hi ha

fortes repulsions estèriques i per aquest motiu, el terme de l'energia intramolecular és molt elevat comparat amb els altres casos. Finalment, és important destacar que la formació d'enllaços d'hidrogen no proporciona informació rellevant ja que el nombre d'enllaços d'hidrogen és molt superior en el complex d'energia més elevada.

## 9.6 RESUM DELS RESULTATS

Com a resum dels estudis realitzats s'ha vist que les estructures de l'Hsp70 obtingudes a partir de la mitjana de les coordenades al llarg de la simulació amb i sense ATP són pràcticament idèntiques, per tant, no s'aprecien canvis conformacionals destacables. No obstant, a partir de l'anàlisi de les fluctuacions atòmiques sí que s'han detectat canvis conformacionals més importants dins del primer subdomini de l'Hsp70 en presència de l'ATP.

D'altra banda, s'han trobat totes les combinacions possibles que proporcionen les fraccions de 25, 33, 38, 43, 45, 60 i 68 kDa, però donat que només es disposa de l'estructura cristal·lina del domini N-terminal de l'Hsp70, s'han escollit les combinacions que estan dins d'aquest domini. A més, per iniciar l'estudi s'han considerat aquelles que provenen del trencament per un únic aminoàcid. De totes les anàlisis de *docking* s'ha trobat que el complex de mínima energia és aquell que està al voltant de la Lys257, per unes millors interaccions intermoleculares receptor-ligand. Per tant, podem dir que aquests resultats estan d'acord amb la primera teoria postulada, en què suposem que la ciclodextrina podria estar impeding algun dels possibles punts d'hidròlisi de l'Hsp70.

No obstant, convé destacar que, per calcular les energies de *docking*, només s'han emprat les interaccions no enllaçants, i s'han omès tots els altres termes energètics que hi ha dins d'un camp de força, és a dir, aquests resultats s'han de considerar com un primer pas, a partir del qual s'hauran de continuar els estudis. A més, el fet de que el complex de mínima energia que s'ha trobat conté la ciclodextrina en l'entrada de la cavitat també podria explicar la segona teoria postulada, en què la ciclodextrina podria interaccionar amb l'ATP, i inhibir la seva activitat.



## **IV. CONCLUSIONS**

---



## 10 CONCLUSIONS

Les conclusions d'aquest treball d'investigació es resumeixen en els següents punts:

- S'ha desenvolupat una nomenclatura per caracteritzar cadenes de polisacàrids que permet avaluar els canvis conformacionals i es poden determinar seqüències concretes que caracteritzen elements estructurals determinats: molècules estirades tenen glucoses consecutives en conformació  ${}^4C_1$ , seqüències del tipus  ${}^4C_1, {}^4C_1, {}^4C_1$  implica que la glucosa del mig obstaculitzi l'entrada de la cavitat, i glucoses en conformació  ${}^1C_4$  formen hèlixs o voltes d'hèlix.
- La modelització de les ciclodextrines comunes no es veu excessivament influenciada pel camp de força, ja que aquestes *per se* són suficientment rígides. D'altra banda, la modelització de les ciclodextrines més grans, es veu altament influenciada pel camp de força.
- Els camps de força parm94, parm99 i MM3\* no són adequats per a la modelització de carbohidrats flexibles degut a la manca de paràmetres adients. El camp de força glycam 2000a conté un bon conjunt de paràmetres per realitzar estudis en fase gas, però s'haurien de reparametritzar les interaccions amb el solvent.
- Entre els dímers de la 3-O-(metilnaftil)- $\beta$ -ciclodextrina el HHi és el més estable entròpicament i entàlpicament.
- L'ATP promou canvis conformacionals en el primer subdomini de l'Hsp70 humana.
- El punt de complexació més estable energèticament entre l'Hsp70 i la 6<sup>A</sup>,6<sup>D</sup>-bis(dansil-Gly)- $\beta$ -ciclodextrina es situa a l'entrada de l'esclatxa, al voltant de la Lys257.





## **V. METODOLOGIA**

---



## 11 MÈTODES DE CÀLCUL

Tradicionalment, els mètodes de càlcul havien estat dividits en dos grans grups: els mètodes basats en la mecànica quàntica i els mètodes basats en la mecànica clàssica o mecànica molecular. Els primers consideren els electrons del sistema amb el que el nombre d'àtoms amb què tracten és limitat. D'altra banda, els segons ignoren el moviment dels electrons i, per tant, el nombre d'àtoms amb els que poden treballar és molt superior.

No obstant, degut als avenços tecnològics i a l'aparició dels supercomputadors i de nou programari, aquesta divisió no permet classificar bé els nous mètodes de càlcul existents avui en dia. Així doncs, actualment és recomanable considerar una nova classificació dels mètodes de càlcul:

- a. Mètodes basats en la mecànica quàntica (QM)
- b. Mètodes basats en la mecànica clàssica (MM)
- c. Mètodes híbrids (QM/MM)

Com ja s'ha esmentat, els **mètodes basats en la mecànica quàntica (QM)** representen explícitament els electrons del sistema. Per tant, aquests mètodes permeten investigar propietats que depenen de la distribució electrònica i en particular, estudiar reaccions químiques en les quals es produeix el trencament i la formació d'enllaços. Amb aquests mètodes es fa ús de l'aproximació de Born-Oppenheimer, la qual separa el moviment dels electrons i el dels nuclis ja que, la baixa massa dels electrons comparada amb la massa dels nuclis permet ajustar molt més ràpidament la posició dels electrons per a qualsevol canvi en les posicions nuclears.

Hi ha actualment dues metodologies quàntiques per tractar els sistemes moleculars:

- a. Mètodes basats en la funció d'ona
- b. Mètodes basats en el funcional de densitat (DFT)

Els mètodes basats en la funció d'ona resolen l'equació d'Schrödinger electrònica independent del temps. Dintre d'aquest grup es troben els mètodes *ab initio* i els mètodes semi-empírics. En els mètodes *ab initio*, on s'inclouen els mètodes Hartree-Fock i post Hartree-Fock, la funció d'ona multieletrònica s'expressa com una sèrie de determinants d'Slater que es construeix a partir d'un conjunt d' $N$  funcions simples electròniques, on  $N$  és el nombre d'electrons del

sistema de tal manera que, calcula la funció d'ona completa per als N electrons. Per tant, aquests mètodes poden requerir d'un elevat cost computacional ja que al considerar tots els electrons del sistema el nombre d'integrals a resoldre és molt elevat. D'altra banda, els mètodes semi-empírics són conceptualment iguals als anteriors però aproximen algunes integrals, només consideren els electrons de valència del sistema i engloba els electrons més interns conjuntament amb el nucli. Entre aquest darrer grup es troben els mètodes MINDO/3, MNDO, i AM1 entre d'altres.

D'altra banda, els mètodes basats en el funcional de densitat també consideren funcions electròniques simples però, a diferència dels mètodes basats en la funció d'ona, calcula l'energia electrònica total mitjançant la distribució de la densitat electrònica en el sistema complet. És a dir, l'energia és funció de la densitat electrònica.

Els **mètodes basats en la mecànica molecular (MM)**,<sup>82</sup> com ja s'ha esmentat, ignoren els electrons del sistema i calculen l'energia del sistema únicament com a funció de les posicions nuclears. Aquests són els mètodes que s'han utilitzat majoritàriament en la present tesi doctoral, ja que són idonis per realitzar anàlisis conformacionals de sistemes amb elevat nombre d'àtoms. Per aquest motiu, es comentaran amb més detall en els següents apartats. No obstant, cal destacar que, malgrat l'aparent senzillesa d'aquests mètodes, en molts casos poden proporcionar resultats tan acurats com els càlculs quàntics de més alta qualitat i en només una fracció del temps computacional.

Finalment, els **mètodes híbrids QM/MM** tracten la part reactiva del sistema amb mecànica quàntica i la part restant del model es representa mitjançant la mecànica molecular. En aquest cas, l'energia del sistema es pot escriure com:

$$E_{TOT} = E_{QM} + E_{MM} + E_{QM/MM} \quad \text{Equació 11-1}$$

On  $E_{QM}$  és l'energia de la part tractada quànticament,  $E_{MM}$  és l'energia de la part tractada clàssicament i  $E_{QM/MM}$  és l'energia d'interacció entre ambdues parts.

---

<sup>82</sup> Leach, A.R., Molecular Modelling: Principles and applications; Pearson: Harlow, 2002.

S'han descrit diferents mètodes QM/MM, aquests es diferencien en la teoria quàntica utilitzada (semi-empíric,<sup>83</sup> *ab initio*,<sup>84</sup> enllaç de valència<sup>85</sup> o DFT<sup>86</sup>), en el model de mecànica molecular, i en la manera en què es representa el solvent (explícit o model simplificat). A més, també es poden diferenciar en com tracten la frontera entre les regions QM i MM.

Relacionats amb aquests mètodes se n'han implementat d'altres que descriuen una aproximació multicapa. Aquest són els mètodes ONIOM.<sup>87</sup> A diferència dels QM/MM, en aquest cas es defineixen major nombre de capes i cadascuna es tracta amb un nivell de teoria diferent, *ab initio* amb bases grans per a la capa més interna, i en les següents capes es va reduint la qualitat del nivell fins a les més externes, que es tracten amb mètodes semi-empírics o de mecànica molecular.

A continuació, s'expliquen els fonaments de la mecànica molecular amb detall, ja que aquest mètode de càlcul s'ha utilitzat àmpliament en aquesta tesi doctoral.

## 11.1 MECÀNICA MOLECULAR

Els mètodes de mecànica molecular també es basen en l'aproximació de Born-Oppenheimer, sense la qual seria impossible calcular l'energia en funció de les posicions nuclears. Aquests mètodes consideren els àtoms com a masses de certa grandària i deformabilitat, i els enllaços com a molles de diferent elasticitat. Així doncs, l'energia d'un sistema es calcula a partir de contribucions tan simples com l'estirament i arronsament dels enllaços (*stretching*), la deformació dels angles d'enllaç (*bending*), la rotació al voltant d'un enllaç donat (*torsion*), i les interaccions entre àtoms no enllaçats (*non bonded*).

Les equacions que representen els canvis energètics d'aquests processos i els paràmetres formen els **camp de forces**. Aquests, estan dissenyats per reproduir i predir propietats estructurals i existeixen camps de forces molt diversos, els quals han estat parametritzats per

---

<sup>83</sup> Warshel, A.; Levitt, M., *Journal of Molecular Biology*, **1976**, 103(2), 227-249.

<sup>84</sup> Singh, U.C.; Kollman, P.A., *J. Chem. Phys.*, **1985**, 83(8), 4033-4040.

<sup>85</sup> Field, M.J.; Bash, P.A.; Karplus, M., *J. Comput. Chem.*, **1990**, 11(6), 700-733.

<sup>86</sup> Maseras, F.; Morokuma, K., *J. Comput. Chem.*, **1995**, 16(9), 1170-1179.

<sup>87</sup> Svensson, M.; Humbel, S.; Froese, R.D.J.; Matsubara, T.; Sieber, S.; Morokuma, K., *J. Phys. Chem.*, **1996**, 100(50), 19357-19363.

modelitzar diferents famílies de molècules. Per exemple alguns dels camps de forces més altament utilitzats i el seu domini d'aplicació són:

MM2,<sup>88</sup> MM3,<sup>89</sup> MM4<sup>90</sup> (Allinger): molècules orgàniques

OPLS<sup>91</sup> (Jorgensen): simulacions líquids

ECEPP<sup>92</sup> (Scheraga): pèptids

AMBER<sup>93</sup> (Kollman): proteïnes i àcids nucleics

CHARMM<sup>94</sup> (Karplus): proteïnes

Glycam<sup>95</sup> (Woods): carbohidrats i glicoproteïnes

AMB99C<sup>96</sup> (Momany): carbohidrats

CSFF<sup>97</sup> (Brady): carbohidrats

Per calcular l'energia d'un sistema donat, tots els camps de forces actuals utilitzen un conjunt de termes energètics que fan referència a les interaccions entre àtoms enllaçats i un altre conjunt de termes que fan referència a les interaccions entre àtoms no enllaçats. L'Equació 11-2 representa els termes que generalment contenen els camps de forces: es tenen en compte les interaccions entre àtoms enllaçats de deformacions dels enllaços ( $E_{str}$ ), de deformació dels angles ( $E_{bnd}$ ) i de rotació al voltant dels enllaços ( $E_{tors}$ ), i les interaccions entre àtoms no enllaçats d'atracció i repulsió de masses ( $E_{vdw}$ ) i d'atracció i repulsió de càrregues ( $E_{elec}$ ).

---

<sup>88</sup> Allinger, N.L., *J. Am. Chem. Soc.*, **1977**, 99(25), 8127-8134.

<sup>89</sup> Allinger, N.L.; Yuh, Y.H.; Lii, J-H., *J. Am. Chem. Soc.*, **1989**, 111(23), 8551-8582.

<sup>90</sup> Allinger, N.L.; Chen, K.; Lii, J-H., *J. Comput. Chem.*, **1996**, 17(5 i 6), 642-668.

<sup>91</sup> Jorgensen, W.L.; Tirado-Rives, J., *J. Am. Chem. Soc.*, **1988**, 110(6), 1657-1666.

<sup>92</sup> Momany, F.A.; Carruthers, L.M., Mc.Guire, R.F.; Scheraga, H.A., *J. Phys. Chem.*, **1974**, 78(16), 1595-1620.

<sup>93</sup> Weiner, P.K.; Kollman, P.A., *J. Comput. Chem.*, **1981**, 2(3), 287-303.

<sup>94</sup> Brooks, B.R.; Bruccoleri, R.E.; Olasfon, B.D.; States, D.J.; Swaminathan, S.; Karplus, M., *J. Comput. Chem.*, **1983**, 4(2), 187-217.

<sup>95</sup> Woods, R.J.; Dwek, R.A.; Edge, C.J.; Fraser-Reid, B., *J. Phys. Chem.*, **1995**, 99(11), 3832-3846.

<sup>96</sup> Momany, F.A.; Willet, J.L., *Carbohydrate Research*, **2000**, 326(3), 194-209.

<sup>97</sup> Kuttel, M.; Brady, J.W.; Naidoo, K.J., *J. Comput. Chem.*, **2002**, 23(13), 1236-1243.

$$E_{TOT} = \sum E_{str}(l) + \sum E_{bnd}(\mathbf{q}) + \sum E_{tors}(\mathbf{f}) + \sum E_{vdW}(r) + \sum E_{elec}(r) \quad \text{Equació 11-2}$$

L'energia total és un sumatori de sumatoris per tant, depèn del nombre d'àtoms del sistema: com més àtoms hi hagi al sistema més elevada serà l'energia total. Aquest fet implica que sempre s'ha de comparar valors energètics de molècules que tinguin igual nombre d'àtoms i amb la mateixa connectivitat.

És habitual trobar dos camps de forces que utilitzen un mateix conjunt de funcions però que contenen diferents paràmetres. Per exemple, per reproduir un mateix fenomen un camp de forces donat pot emprar un conjunt adient de paràmetres de torsió i un altre camp de forces pot reproduir-lo mitjançant interaccions entre àtoms no enllaçats. Així doncs, **la parametrització** esdevé un punt clau a l'hora de crear un nou camp de forces. Aquesta es duu a terme a partir de dades experimentals, o a partir de càlculs quàntics del tipus *ab initio* o DFT. La transferència dels paràmetres dins d'un camp de forces és un punt clau, ja que per parametritzar s'utilitza un nombre relativament petit de molècules, però posteriorment els paràmetres s'apliquen a un ampli conjunt de molècules. A més, s'ha de tenir en compte que els paràmetres no es poden transferir d'un camp de forces a l'altre perquè aquests es comporten com una entitat única. No obstant, alguns termes són suficientment independents (distàncies i angles d'enllaç) i en alguns casos l'intercanvi d'aquests paràmetres entre camps de forces pot ser vàlid.

A diferència de la mecànica quàntica, on és necessari definir el número atòmic dels nuclis presents, la geometria del sistema, la càrrega total i la multiplicitat d'espín, en un camp de forces només s'ha de definir el tipus d'àtom i la geometria. El tipus d'àtom representa més que el número atòmic, i habitualment també conté informació sobre la hibridació i l'entorn de l'àtom. Per exemple, en un camp de forces hi ha diferents tipus d'àtom per carbonis amb hibridació  $sp^2$  i per carbonis amb hibridació  $sp^3$ .

Ja s'ha vist en la secció precedent que en aquesta tesi doctoral s'han emprat diferents camps de forces per realitzar els estudis. És per aquest motiu, que en els següents apartats es detallen les equacions de diferents camps de forces, i es comenten molt breument algunes de les seves característiques. Així doncs, primer es detallen les equacions del camp de forces MM3, les quals són la base per entendre les equacions del camp de forces MM3\*. A continuació es detallen les equacions dels camps de forces d'AMBER. I finalment, es detallen aquelles equacions del camp de forces CHARMM que són diferents a les del camp de forces AMBER.

### 11.1.1 Camp de forces MM3

#### 11.1.1.1 Energia de deformació de l'esquelet molecular

La deformació de les distàncies d'enllaç es representa estrictament per un potencial de Morse. No obstant, s'han d'utilitzar tres paràmetres per ajustar la funció i per aquest motiu, habitualment s'utilitzen altres tipus de funcions. En MM3 la deformació de les distàncies està representada per una funció quàrtica (Equació 11-3), on  $k_s$  és la constant de força,  $l_0$  el valor de referència de la distància, i  $l$  el valor real de la distància. Aquesta funció evita que els enllaços es deformin excessivament quan la geometria de partida és dolenta i s'assembla a un potencial de Morse en un rang força ampli de distàncies.

$$E_s(l) = 71.94k_s(l-l_0)^2\left[1 - 2.55(l-l_0) + \frac{7}{12}2.55(l-l_0)^2\right] \quad \text{Equació 11-3}$$

En el cas de la deformació dels angles (Equació 11-4),  $k_q$  és la constant de força,  $q_0$  és l'angle de referència i  $q$  l'angle real. Amb el terme quàrtic de l'equació ja es reproduïen correctament els sistemes més comuns. No obstant, s'inclouen termes fins a la sexta potència per descriure millor les dades experimentals de sistemes altament tensionats com per exemple el biciclo[1.1.1]pentà.

$$E_q(q) = 0.021914k_q(q-q_0)^2\left[1 - 0.014(q-q_0) + 5.6 \cdot 10^{-5}(q-q_0)^2 - 7 \cdot 10^{-7}(q-q_0)^3 + 9 \cdot 10^{-10}(q-q_0)^4\right] \quad \text{Equació 11-4}$$

En les funcions que descriuen la deformació de la distància i l'angle d'enllaç, el mínim d'energia correspon al valor de la distància o de l'angle parametritzat ( $l_0$  i  $q_0$ ), i qualsevol desviació d'aquesta situació d'equilibri provoca un augment en l'energia, que és proporcional a la magnitud de la desviació. Si el sistema s'allunya de l'equilibri, les constants de força  $k_s$  i  $k_q$  són les responsables de retornar el sistema a la seva posició.

#### 11.1.1.2 Energia dels diedres propis i impropis

Les barreres de rotació al voltant d'un enllaç donat són fonamentals per entendre les propietats estructurals i l'anàlisi conformacional, per tant, és d'elevada importància que el camp de forces pugui reproduir correctament el perfil energètic d'aquests processos. En MM3 les torsions estan representades per una sèrie de Fourier de tres termes, cada un d'ells caracteritzat per una funció cosinus (Equació 11-5), on  $f$  és l'angle diedre i  $V_1$ ,  $V_2$  i  $V_3$  són les constants que



depenen de la naturalesa de l'enllaç i donen una idea de la barrera energètica: com més grans siguin més alta és la barrera energètica.

$$E_f(\mathbf{f}) = \frac{V_1}{2}(1 + \cos \mathbf{f}) + \frac{V_2}{2}(1 - \cos 2\mathbf{f}) + \frac{V_3}{2}(1 + \cos 3\mathbf{f}) \quad \text{Equació 11-5}$$

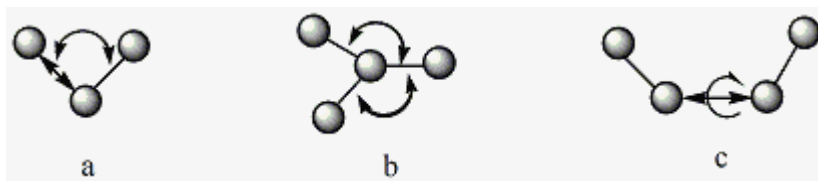
En el camp de forces MM3, cada un dels termes  $V_1$ ,  $V_2$  i  $V_3$  té un significat característic:  $V_1$  representa la diferència entre dipòls deguda a les diferències d'electronegativitat dels àtoms directament units a l'enllaç que gira,  $V_2$  representa la hiperconjugació i la torsió al voltant de dobles enllaços, i  $V_3$  representa la rotació al voltant d'enllaços  $sp^3-sp^3$  en la qual hi ha interaccions estèriques entre els àtoms en posició relativa 1,4.

D'altra banda les torsions impròpies estan associades a la deformació dels àtoms amb hibridació  $sp^2$  i s'inclouen per mantenir aquest àtom amb hibridació  $sp^2$  i als tres àtoms units a ell en un mateix pla. En el camp de forces MM3 es representen mitjançant una funció harmònica (Equació 11-6) on  $\mathbf{x}$  és l'angle de desviació del pla i  $k_x$  la constant de força característica.

$$E_x(\mathbf{x}) = \frac{1}{2}k_x(\mathbf{x} - \mathbf{x}_0)^2 \quad \text{Equació 11-6}$$

### 11.1.1.3 Termes creuats

Els termes creuats reflecteixen la correlació existent entre les coordenades internes, com per exemple l'augment de la distància d'enllaç al tancar-se l'angle d'enllaç per tal de minimitzar les interaccions entre els àtoms en posició relativa 1,3. En MM3 s'inclouen tres termes creuats: *stretching-bending*, *stretching-torsion* i *bending-bending* aquests es representen a la Figura 11-1.



**Figura 11-1 Representació de la correlació a) *stretching-bending* b) *bending-bending* i c) *stretching-torsion*.**

Les equacions que modelitzen aquests termes són les següents:

$$E_{sq}(l, \mathbf{q}) = 2.51118k_{sq} [(l - l_0) + (l' - l_0')](\mathbf{q} - \mathbf{q}_0) \quad \text{Equació 11-7}$$

$$E_{qq}(\mathbf{q}) = -0.021914k_{qq}(\mathbf{q} - \mathbf{q}_0)(\mathbf{q}' - \mathbf{q}_0') \quad \text{Equació 11-8}$$

$$E_{sf}(l, \mathbf{f}) = 11.995 \frac{k_{sf}}{2} (l - l_0)(1 + \cos 3\mathbf{f}) \quad \text{Equació 11-9}$$

On  $k_{sq}$ ,  $k_{qq}$  i  $k_{sf}$  representen les constants de força de cada terme creuat,  $l$  i  $l'$  són les distàncies d'enllaç reals i  $l_0$  i  $l_0'$  les distàncies d'enllaç ideals,  $\mathbf{f}$  és l'angle de torsió real, i  $\mathbf{q}$  i  $\mathbf{q}'$  són els angles reals, i  $\mathbf{q}_0$  i  $\mathbf{q}_0'$  els angles d'enllaç ideals.

#### 11.1.1.4 Interaccions entre àtoms no enllaçats

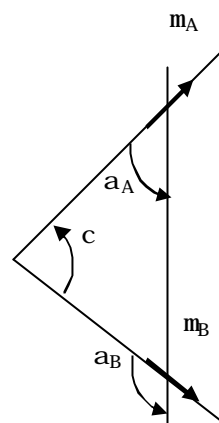
Les interaccions entre àtoms no enllaçats són interaccions a través de l'espai i es centren en les interaccions d'atracció i repulsió de masses (van der Waals) i les interaccions d'atracció i repulsió de càrregues (electrostàtiques). En MM3 l'energia de van der Waals (Equació 11-10) està representada per una funció tipus Buckingham on  $R$  és la suma dels radis de van der Waals dels àtoms per separat,  $r$  és la distància real entre els dos àtoms, i  $e$  el paràmetre de l'energia per a la interacció de dos àtoms. El terme exponencial té efecte repulsiu al apropar-se massa dos àtoms, i el terme  $(R/r)^6$  té efecte atractiu i reproduceix l'atracció de masses.

$$E_{vdW}(r) = e \left[ -2.25 \left( \frac{R}{r} \right)^6 + 1.84 \cdot 10^5 \exp \left[ -12.0 \frac{R}{r} \right] \right] \quad \text{Equació 11-10}$$

D'altra banda, per tractar les interaccions electrostàtiques, en el cas de que la molècula sigui neutra, MM3 associa moments dipolars i tracta aquestes interaccions com a dipol-dipol mitjançant l'equació de Jeans (Equació 11-11) on  $\mathbf{m}_A$  i  $\mathbf{m}_B$  són els moments dipolars,  $e$  és la constant dielèctrica i la resta de paràmetres es representen a la Figura 11-2. No obstant, en el cas de que la molècula estigui carregada, existeixen altres subrutines per avaluar les interaccions dipol-càrrega i càrrega-càrrega.

$$E_{dd} = 14.39418 \frac{m_A m_B (\cos c - 3 \cos a_A \cos a_B)}{R^3 e}$$

Equació 11-11

Figura 11-2 Representació dels paràmetres per avaluar l'energia  $E_{dd}$ .

Fins ara s'han descrit les equacions del camp de forces MM3 no obstant, en la present tesi doctoral s'ha utilitzat el camp de forces MM3\*, que és una adaptació del camp de forces original MM3 per treballar amb Macromodel.<sup>98</sup> Així doncs, cal destacar que, malgrat són molt similars, hi ha algunes diferències entre ambdós camps de forces, les quals es resumeixen en els següents punts:

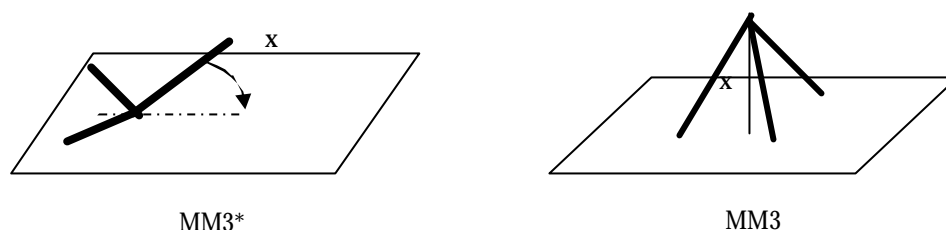
- MM3 tracta les interaccions electrostàtiques com a interacció dipol-dipol, en canvi MM3\* utilitza càrregues parcials i fa servir la llei de Coulomb per modelitzar aquestes interaccions. Per calcular les càrregues, el camp de forces MM3\* agafa els paràmetres originals de MM3 i mitjançant l'Equació 11-12 calcula el valor de les càrregues:  $m$  representa el moment dipolar,  $L$  la distància d'enllaç i  $Q$  la càrrega.

$$m = \frac{L \cdot Q}{0.2082}$$

Equació 11-12

- Per descriure el bending fora del pla, MM3 utilitza una distància de piramidalització mentre MM3\* fa ús d'una torsió impròpia (Figura 11-3).

<sup>98</sup> Mohamadi, F.; Richards, N.G.J.; Guida, W.C.; Liskam, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W.C., *J. Comput. Chem.*, **1990**, 11(4), 440-467.



**Figura 11-3 Representació de les torsions impròpies.**

- c. Per tractar sistemes conjugats MM3 realitza un càlcul SCF quan aquests són poc habituals, en canvi MM3\* utilitza paràmetres de torsió  $V_1$ - $V_3$  específics per a aquests sistemes.

### 11.1.2 Camp de forces AMBER

El camp de forces AMBER (*Assised Model Building with Energy Refinement*) ha estat dissenyat per a la modelització d'àcids nucleics i proteïnes. No obstant, existeixen altres camps de forces basats en aquest que utilitzen el mateix conjunt de funcions però diferents paràmetres i que estan dissenyats per a l'estudi d'altres famílies de molècules, com per exemple, les diferents versions de Glycam i l'AMB99C que són específics per a carbohidrats. Les equacions en el camp de forces AMBER són molt diferents a les equacions del camp de forces MM3, per aquest motiu, a continuació es descriuen detalladament.

#### 11.1.2.1 Energia de deformació de l'esquelet molecular

L'energia d'*stretching* (Equació 11-13) i de *bending* (Equació 11-14) es representen mitjançant una funció harmònica, on  $l_0$  i  $q_0$  són la distància i l'angle d'enllaç d'equilibri respectivament, i  $k_s$  i  $k_q$  les respectives constants de força. Aquesta funció permet descriure correctament l'estructura i l'energia de proteïnes que no estiguin excessivament tensionades.

$$E_s(l) = k_s(l - l_0)^2 \quad \text{Equació 11-13}$$

$$E_q(q) = k_q(q - q_0)^2 \quad \text{Equació 11-14}$$

#### 11.1.2.2 Energia de torsió

En aquest cas, l'energia de torsió també es descriu amb una sèrie de Fourier (Equació 11-15) on cada un dels termes és una funció cosinus però, a diferència del camp de forces MM3, en aquest cas el nombre de termes cosinus pot ser variable (normalment entre 1 i 3). En l'equació

$V_n$  és el paràmetre de torsió i novament dóna idea de la barrera energètica,  $n$  és la multiplicitat, és a dir, el nombre de punts de mínima energia,  $f$  és l'angle diedre i en aquest cas, també s'inclou el factor de fase  $g$  que indica quan l'energia de torsió passa pel seu mínim.

$$E_f(\mathbf{f}) = \frac{V_n}{2}(1 + \cos(n\mathbf{f} - \mathbf{g})) \quad \text{Equació 11-15}$$

D'altra banda, cal remarcar que en aquest cas, les torsions impròpies es modelitzen amb la mateixa equació que les torsions pròpies.

### 11.1.2.3 Interaccions entre àtoms no enllaçats

Les interaccions de van der Waals es representen mitjançant un potencial de Lennard-Jones (Equació 11-16), on hi ha un primer terme repulsiu:  $r^{-12}$ , i un segon terme atractiu:  $r^{-6}$ . Les constants  $A_{ij}$  i  $B_{ij}$  es calculen com  $\mathbf{e}_{ij} \cdot R_{ij}^{12}$  i  $2 \cdot \mathbf{e}_{ij} \cdot R_{ij}^6$  respectivament, on  $R_{ij}$  és la suma dels radis de van der Waals dels àtoms per separat, i  $\mathbf{e}_{ij}$  és la profunditat del pou.

$$E_{vdW}(r) = \frac{A_{ij}}{r^{12}} - \frac{B_{ij}}{r^6} \quad \text{Equació 11-16}$$

Per tractar les interaccions electrostàtiques s'utilitzen l'equació de Coulomb (Equació 11-17) on  $q_i$  i  $q_j$  són càrregues puntuals centrades en l'àtom  $i$  i  $j$  respectivament,  $\epsilon$  la constant dielèctrica i  $r$  la distància entre àtoms.

$$E_{elec}(r) = \frac{q_i \cdot q_j}{\epsilon \cdot r} \quad \text{Equació 11-17}$$

### 11.1.3 Camp de forces CHARMM

Finalment, el camp de forces CHARMM (*Chemistry at HARvard Macromolecular Mechanics*) també és específic per a proteïnes però, igual que s'ha esmenat en el cas d'AMBER, també hi ha altres camps de forces que fan servir el mateix conjunt de funcions però diferents paràmetres específics per a altres famílies de molècules com per exemple, el CSFF (*Carbohydrate Solution Force Field*) específic per a carbohidrats. En aquest cas, les equacions són més semblants a les equacions d'AMBER i, per tant, només es farà èmfasi a aquells termes energètics que estan representats amb altres funcions.

## 11.1.3.1 Energia de torsió

Les torsions pròpies també es descriuen mitjançant una sèrie de Fourier (Equació 11-18) on cada un dels termes és una funció cosinus però, en aquest cas, el nombre de termes pot ser superior que en el camp de forces AMBER ( $n = 1,2,3,4,5,6$ ).

$$E_f(\mathbf{f}) = k_f(1 - \cos(n\mathbf{f})) \quad \text{Equació 11-18}$$

D'altra banda, les torsions impròpies es descriuen segons una funció harmònica (Equació 11-19) i s'introdueix per tal de mantenir la quiralitat al voltant de centres tetraèdrics i la planaritat al voltant d'àtoms amb hibridació  $sp^2$ .

$$E_w(\mathbf{w}) = k_w(\mathbf{w} - \mathbf{w}_0)^2 \quad \text{Equació 11-19}$$

## 11.1.3.2 Interaccions entre àtoms no enllaçats

Les interaccions no enllaçants d'atracció i repulsió de masses es descriuen novament segons un potencial de Lennard-Jones (Equació 11-20), però cal destacar que també s'inclou una funció de "switching" ( $S_w$ ) que permet que, entre dues distàncies predeterminades, la corba d'energia tendeixi gradualment fins a zero de manera que, a partir del límit superior predeterminat, no es consideren aquestes interaccions a llarga distància.

$$E_{vdW}(r) = \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) \cdot S_w(r_{ij}^2, r_{on}^2, r_{off}^2) \quad \text{Equació 11-20}$$

---


$$* \text{ Si } x \leq x_{on} \quad S_w(x_{ij}^2, x_{on}^2, x_{off}^2) = 1$$

$$\text{Si } x_{on} < x < x_{off} \quad S_w(x_{ij}^2, x_{on}^2, x_{off}^2) = \frac{(x_{off} - x)^2(x_{off} + 2x - 3x_{on})}{(x_{off} - x_{on})^3}$$

$$\text{Si } x > x_{off} \quad S_w(x_{ij}^2, x_{on}^2, x_{off}^2) = 0$$

Finalment, per avaluar les interaccions electrostàtiques existeixen diferents equacions i segons les necessitats i interessos de l'investigador es pot recórrer a una o a una altra tipus de funció. No obstant, totes elles estan basades en la llei de Coulomb. A continuació es resumeixen els diferents tipus de funcions:

Dielèctric constant: 
$$E_{el} = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}$$

Dielèctric depenent de la distancia: 
$$E_{el} = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} S_w(r_{ij}^2, r_{on}^2, r_{off}^2)$$

Dielèctric ajustat: 
$$E_{el} = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \left(1 - \frac{2r_{ij}^2}{r_{cut}^2} + \frac{r_{ij}^4}{r_{cut}^4}\right)$$

Interaccions electrostàtiques entre grups : És igual que la segona equació però en la funció de “switching” les distàncies no són entre àtoms sinó que son entre els centres geomètrics dels grups.

Funció extesa:  $E_{el} = E_{el}(dir) + E_{el}(ext)$  la primera part es tracta de manera habitual segons la llei de Coulomb i la segona part es tracta mitjançant multipols.

## 12 MÈTODES DE CERCA CONFORMACIONAL

Les propietats físiques, químiques i biològiques de les molècules depenen en la majoria dels casos de la seva estructura tridimensional o conformació. Per tant, trobar el confòrmer o confòrmers més estables energèticament d'un sistema donat és un estudi habitual dins de la química computacional. El problema més habitual amb el que es troba el químic teòric és que, per a sistemes d'una certa grandària, no es coneix la hipersuperfície de potencial i, per tant, mai es té la certesa de trobar el mínim absolut.

Hi ha diferents metodologies que es poden emprar per realitzar cerques conformacionals, i que es diferencien en l'algorisme que utilitzen per crear noves conformacions del sistema. A continuació es detallen els fonaments teòrics de diverses metodologies que s'han emprat en aquesta tesi doctoral.

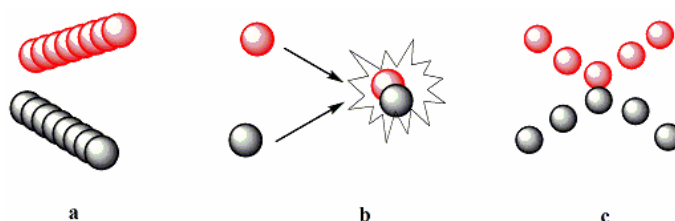
## 12.1 DINÀMICA MOLECULAR

Les simulacions de dinàmica molecular, igual que la mecànica molecular, també tenen el seu fonament en la mecànica clàssica i generen conformacions del sistema integrant la segona llei del moviment de Newton (Equació 12-1). Aquesta descriu el moviment d'una partícula  $i$  de massa  $m$  en la posició definida pel vector  $r$  degut a l'acció d'una força  $F_r$ . Per tant, l'única incògnita de l'equació és el temps i per integració d'aquesta s'obté una trajectòria que especifica com varien les posicions i les velocitats de les partícules al llarg del temps.

$$F = m \cdot a \quad ; \quad \frac{\partial^2 r_i}{\partial t^2} = \frac{F_{ri}}{m_i} \quad \text{Equació 12-1}$$

La força sobre una partícula es pot calcular a partir del gradient negatiu del potencial (calculat amb les mateixes equacions de la mecànica molecular), no obstant, aquesta varia quan la partícula o quan les partícules que interaccionen amb ella varien la seva posició. Per tant, el moviment de les partícules està correlacionat, i el problema a resoldre és un problema de molts cossos, el qual no es pot resoldre analíticament. Així doncs, les equacions del moviment es resolen mitjançant mètodes de diferències finites, els quals divideixen la integral en molts passos petits definits per un  $\delta t$  o **time step**. Existeixen molts algorismes basats en aquests mètodes entre els quals cal destacar el de Verlet i els de *Leap-frog velocity-Verlet* i Beeman, que són simples modificacions del primer. Tots ells consideren que la posició i les propietats dinàmiques es poden desenvolupar en sèrie de Taylor.

Així doncs, per calcular la trajectòria és necessari definir el *time step*. Aquest no pot ser gaire petit (Figura 12-1a) perquè aleshores s'explora molt lentament l'espai conformacional, però tampoc pot ser gaire gran (Figura 12-1b) perquè les partícules podrien col·lapsar en un mateix punt de l'espai, per tant, s'ha d'escollir aquell *time step* que permeti explorar una àmplia part de l'espai conformacional i que permeti que les col·lisions entre partícules siguin suaus (Figura 12-1c).



**Figura 12-1** Efecte de la magnitud del *time step* a) petit: s'explora molt lentament l'espai conformacional b) gran: les partícules col·lapsen c) adequat: s'explora un ampli espai conformacional i amb col·lisions suaus.



Quan es treballa amb molècules flexibles una bona orientació és considerar que el *time step* ha de ser aproximadament una desena part del període més curt de moviment. Així doncs, tenint en compte que la freqüència de vibració més elevada correspon als enllaços C-H, i que aquesta és aproximadament 10 fs, per a aquests sistemes un bon *time step* seria d'1 fs. No obstant, és habitual que en les simulacions de dinàmica molecular s'utilitzi la metodologia *SHAKE*,<sup>99</sup> la qual restringeix les vibracions dels enllaços aplicant forces addicionals i permet que les distàncies es mantinguin en el seu valor d'equilibri. El *SHAKE* s'aplica habitualment als enllaços de major freqüència de vibració C(sp<sup>3</sup>)-H, per tant, si s'utilitza aquesta metodologia és possible augmentar el *time step* fins a 2 fs.

Per iniciar la simulació de dinàmica molecular també és necessari definir una geometria inicial del sistema i una velocitat inicial per a cada àtom, aquesta es selecciona aleatòriament d'una distribució Maxwell-Boltzmann (Equació 12-2) on s'expressa la probabilitat  $p$  de que un àtom  $i$  de massa  $m$  tingui una velocitat  $v$  en la direcció  $x$  a la temperatura  $T$ .

$$p(v_x) = \left( \frac{m_i}{2\pi k_B T} \right)^{1/2} \exp \left[ -\frac{1}{2} \frac{m_i v_{ix}^2}{k_B T} \right] \quad \text{Equació 12-2}$$

Un cop obtingudes les velocitats inicials, es pot iniciar la simulació. A cada pas s'avalua la força sobre cada àtom diferenciant l'energia potencial. A diferència de la MM, en la DM aquesta energia es calcula a partir de les coordenades cartesianes i no a partir de les coordenades internes, però el tipus d'interaccions que es tenen en compte són les mateixes. En els primers passos de la simulació s'incrementa la temperatura fins a arribar al valor final desitjat, al qual es durà a terme la resta de la simulació. Donat que la temperatura està directament relacionada amb l'energia cinètica, i que aquesta està relacionada amb la velocitat, realitzant un escalat de velocitats és possible augmentar progressivament la temperatura.<sup>100</sup> Un cop assolida la temperatura desitjada, el sistema s'equilibra fins que totes les propietats termodinàmiques es mantenen estables i finalment, es realitza l'etapa de producció de la simulació. En aquesta darrera etapa, la temperatura es manté al voltant d'un valor donat mitjançant un esquema d'acoblament feble, en

<sup>99</sup> Ryckaert, J.P.; Ciccotti, G.; Berendsen, H.J.C., *J. Comput. Phys.*, **1977**, 23(3), 327-241.

<sup>100</sup> Woodcock, L.V., *Chem. Phys. Lett.*, **1971**, 10(3), 257-261.

el qual el sistema s'acobra a un bany extern a temperatura fixa que actua com a font d'energia tèrmica, subministrant o extraient calor al sistema segons sigui necessari.<sup>101</sup>

## 12.2 MÈTODE MONTE CARLO

En el mètode Monte Carlo,<sup>102</sup> mitjançant variacions de les coordenades d'una estructura de partida, es localitzen noves geometries del sistema dins del seu espai conformacional. En la present tesi s'ha utilitzat el programa Macromodel per realitzar aquests càlculs i, per tant, les explicacions que venen a continuació estan basades en les opcions i possibilitats que permet aquest programa.

Existeixen diferents modalitats per realitzar la cerca conformacional i que intenten evitar centrar la cerca en una zona limitada de la hipersuperfície de potencial. Aquests mètodes es diferencien en la manera de generar les geometries de partida. Es poden classificar o bé pel tipus de coordenades que utilitzen, o bé per la manera en què modifiquen aquestes coordenades.

Els tipus de coordenades més àmpliament utilitzats són:

- a. Coordenades internes (angles de torsió).
- b. Coordenades externes (Coordenades cartesianes).
- c. Matriu de distàncies internuclears.

I la manera de modificar les coordenades pot ser:

- a. Variació sistemàtica.
- b. Variació aleatòria.

Les matrius de distàncies internuclears generen geometries molt lentament i a més, són necessàries moltes restriccions en les distàncies per obtenir resultats comparables als obtinguts amb les altres coordenades. L'avantatge de les coordenades internes, respecte les coordenades externes, és que petites variacions de les coordenades externes d'un sistema no generen nous confòrmers, en canvi, la diferència més gran entre dos estereoisòmers d'un mateix sistema està en els valors dels seus angles de torsió, és a dir, en les coordenades internes.

---

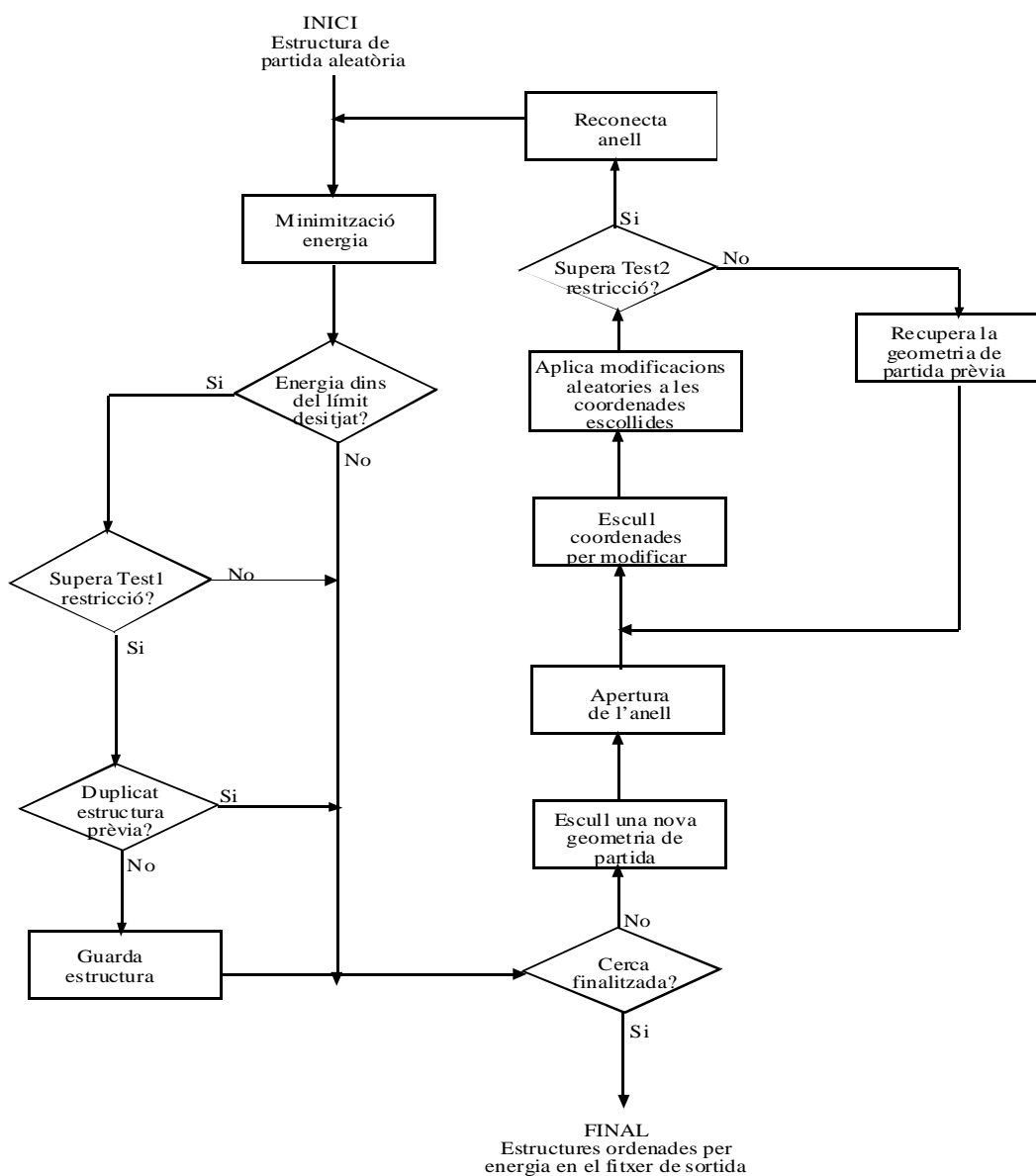
<sup>101</sup> Berendsen, H.J.C.; Postma, J.P.M.; van Gunsteren, W.I.; Di Nola, A.; Haak, J.R., *J. Chem. Phys.*, **1984**, *81*(8), 3648-3690.

<sup>102</sup> Chang, G., Guida, W.C., Still, W.C., *J. Am. Chem. Soc.*, **1989**, *111*(12), 4379-4386.

La cerca sistemàtica escombra tot l'espai configuracional del sistema molecular. És possible trobar tots els confòrmers més poblats de baixa energia però de sistemes que continguin com a màxim 9 angles de torsió cíclics per modificar o 6 acíclics. D'altra banda, la cerca aleatòria només genera un conjunt de configuracions representatives però en sistemes que continguin fins a 14 angles de torsió cíclics per modificar o 10 acíclics.

### 12.2.1 Algorisme de MC

A continuació, en l'Esquema 12-1 es presenta l'algorisme de Monte Carlo que utilitza el programa Macromodel, que com s'ha esmentat, és el programa que s'ha emprat en aquesta tesi doctoral per realitzar aquests càlculs.



Esquema 12-1 Algorisme MC emprat en Macromodel.

El càlcul parteix d'una estructura generada, la qual es minimitza, i sobre aquesta estructura s'aplica el primer test. Aquest té com a objectiu eliminar estructures minimitzades anteriorment, i aquestes s'eliminen segons tres criteris diferents:

- a. Si l'energia està fora del límit superior relatiu al mínim energètic que es tingui en aquest pas.
- b. Si s'han invertit els centres quirals.
- c. Si les distàncies interatòmiques o els angles de torsió no compleixen les restriccions imposades.

Si l'estructura supera aquest test es guarda, i es comença el proper pas. Per començar un pas del càlcul s'ha d'escollir una estructura de partida, hi ha dues opcions per escollir-la:

- a. L'estructura de partida pot ser l'última estructura obtinguda.
- b. L'estructura de partida pot ser qualsevol de les estructures obtingudes anteriorment.

Si el sistema que estem estudiant és cíclic, abans de realitzar qualsevol modificació sobre les coordenades, es trenca temporalment un enllaç del sistema amb la finalitat d'evitar tensions per realitzar les posteriors modificacions en les coordenades. A continuació, es seleccionen les coordenades que es volen modificar, i s'apliquen les modificacions aleatòriament. En aquest punt s'ha obtingut una nova estructura, i abans de minimitzar la seva energia s'aplica el segon test. Aquest test elimina totes aquelles estructures que tinguin:

- a. Geometries que no permeten tancar l'anell, denominades "*poor geometries*".
- b. Valors molt alts de les interaccions internuclears.

Si es supera el test es torna a tancar l'anell cíclic i finalment es minimitza l'energia. És necessari valorar novament l'energia de l'estructura minimitzada, per aquest motiu es torna a aplicar el Test 1. Si es supera aquest test, es té una nova estructura, però abans de guardar-la s'ha de comprovar que no sigui una estructura duplicada d'alguna anterior. Els confòrmers duplicats s'identifiquen per la diferència de mínims quadrats al superposar dues estructures; el programa considera les estructures equivalents excepte si al realitzar la superposició dels àtoms, troba una o més parelles d'àtoms equivalents separats per un valor de  $RMS > 0.25 \text{ \AA}$ .

Si no és un confòrmer duplicat es guarda, i la cerca de nous confòrmers continua de la manera descrita. Quan s'han realitzat tots els passos del càlcul, i un cop guardada l'última estructura, en el fitxer de sortida s'ordenen totes les estructures per ordre creixent d'energia.

### 12.3 ALGORITMES GENÈTICS

Els algoritmes genètics generen poblacions d'estructures, a diferència dels mètodes de dinàmica molecular i Monte Carlo, que generen una única estructura.<sup>103</sup> Aquests mètodes tenen els seus fonaments en l'evolució d'una població i imiten la majoria de les característiques de l'evolució de Darwin. Així doncs, sobre una població inicial s'apliquen operadors genètics per obtenir una nova població.

Els graus de llibertat s'expressen com a gens, i la col·lecció de gens formen els cromosomes, a més, es defineix la salut com l'energia total. Els operadors genètics poden ser de tres tipus: mutacions, creuaments, i migracions. Les mutacions varien el valor d'un gen aleatòriament, els creuaments intercanvien un conjunt de gens d'un cromosoma a l'altre, i les migracions mouen individus d'una sub-població a l'altra. Així doncs, a cada cromosoma se li diagnostica la salut mitjançant la funció de puntuació. En el cas concret del programa Autodock 3.0,<sup>104</sup> que és el que s'ha emprat en la present tesi doctoral, la funció de puntuació conté cinc termes: terme 6-12 de Lennard-Jones de dispersió/repulsió, terme 12-10 d'enllaç d'hidrogen, potencial electrostàtic de Coulomb, terme proporcional al nombre d'enllaços  $sp^3$  per representar l'entropia desfavorable del lligand complexat al receptor degut a les restriccions en els graus de llibertat conformacionals, i terme de dessolvatació. Cada un d'aquests termes s'avalua mitjançant factors d'escalatge empírics, determinats per regressió lineal de 30 complexos proteïna-lligand amb constants de complexació conegudes.

En resum, en la cerca conformacional mitjançant algoritmes genètics, parelles aleatòries d'individus s'aparellen en un procés de creuament, en el qual l'individu resultant hereta gens del seu parent, a més, alguns descendents es muten aleatòriament. La selecció dels descendents de la generació actual es dona segons la salut de cada un d'ells: els que millor s'adapten a l'entorn es podran reproduir, i els que no s'adapten, moren.

---

<sup>103</sup> Taylor, R.D.; Jewsbury, P.J.; Essex, J.W., *J. Comput.-Aided Molec. Design*, **2002**, 16(3), 151-166.

<sup>104</sup> Morris, G.M.; Goodsell, D.S.; Halliday, R.S.; Huey, R.; Hart, W.E.; Belew, R.K.; Olson, A.J., *J. Comput. Chem.*, **1998**, 19(14), 1639-1662.

## 13 MÈTODES DE SOLVATACIÓ

La majoria de processos químics es donen en solució i, per tant, és important considerar com influeix el solvent en el comportament del sistema a estudiar. Normalment el solvent utilitzat és aigua i, per tant, els models de solvatació que es descriuran a continuació seran per a aquest dissolvent. Existeixen diferents tipus de mètodes de solvatació, els quals es poden classificar en dos grans grups segons si tracten l'aigua implícita o explícitament.

### 13.1 MÈTODES IMPLÍCITS DE SOLVATACIÓ

Hi ha diferents mètodes de solvatació que tracten al solvent implícitament; alguns d'ells modifiquen la constant dielèctrica de l'equació de Coulomb, i altres també inclouen l'efecte de la distància en la mateixa equació. No obstant, la polarització és un factor important a tenir en compte, i els mètodes de solvatació implícita més altament emprats tracten al solvent com un continu i estan basats en l'equació generalitzada de Born (GB, Equació 13-1), on  $\epsilon$  és la constant dielèctrica,  $q_i$  i  $q_j$  la càrrega dels àtoms,  $r_{ij}$  la distància entre ells, i  $a_i$  i  $a_j$  el radi dels àtoms.

$$G_{GB} = -\frac{1}{2} \left(1 - \frac{1}{\epsilon}\right) \sum_{i=1}^N \frac{q_i q_j}{r_{ij}^{1/2} + (a_i a_j)^{1/2} \exp(-r_{ij}^2 / 4a_i a_j)} \quad \text{Equació 13-1}$$

En molts casos també es considera l'àrea del solut accessible al solvent (SA), donant lloc als mètodes de solvatació GB/SA.<sup>105</sup> La importància de considerar també aquest terme (Equació 13-2), recau en el fet que es consideren les interaccions de caràcter no polar entre el solut i el solvent ( $G_{vdW}$ ), així com l'entropia associada a la reorganització de les molècules del solvent per crear una cavitat per al solut ( $G_{cav}$ ).

$$G_{SA} = G_{vdW} + G_{cav} = g_k SA_k \quad \text{Equació 13-2}$$

### 13.2 MÈTODES EXPLÍCITS DE SOLVATACIÓ

Quan les simulacions es duen a terme amb molècules d'aigua explícites, el solut es pot envoltar d'una caixa periòdica d'aigua (condicions periòdiques de contorn) o d'una única esfera d'aigua (condicions estocàstiques de contorn). En qualsevol dels casos, les molècules d'aigua poden representar un 80% de tots els àtoms del sistema i per aquest motiu, les interaccions entre

<sup>105</sup> Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T., *J. Am. Chem. Soc.*, **1990**, 112(16), 6127-6129.

les molècules d'aigua ocupen el major cost de computació i, per tant, el model utilitzat per descriure l'aigua ha de ser acurat però també ràpid.

Hi ha diferents models per definir les molècules d'aigua els quals es diferencien principalment en si la molècula d'aigua té una geometria flexible o no, i en el nombre de punts d'interacció. Un dels models més altament emprat, el TIP3P,<sup>106</sup> es caracteritza per presentar una geometria rígida i tres punts d'interacció electrostàtica. En aquest model, els àtom d'hidrogen tenen una càrrega parcial positiva i l'àtom d'oxigen té una càrrega parcial negativa, que compensa les altres dues. La interacció de van der Waals entre dues molècules d'aigua es representa mitjançant un potencial de Lennard-Jones amb un únic punt d'interacció centrat en l'àtom d'oxigen. Les interaccions de van der Waals que inclouen els àtoms d'hidrogen s'ometen. D'altra banda hi ha altres models que inclouen major nombre de punts d'interacció com el TIP4P<sup>106</sup> i el TIP5P<sup>107</sup> que inclouen 4 i 5 punts d'interacció respectivament. En els darrers anys han sorgit modificacions per incloure els efectes de polarització TIP4P-Pol,<sup>108</sup> i modificacions per millorar les interaccions coulòmbiques a llarga distància TIP4P-Ew.<sup>109</sup>

### 13.2.1 Condicions periòdiques de contorn

Les condicions periòdiques es descriuen amb una caixa cúbica, la qual està envoltada en les tres direccions de l'espai de caixes idèntiques a ella. En dues dimensions es pot visualitzar una caixa central envoltada de vuit caixes més (Figura 13-1). Les coordenades de les partícules en les caixes imatges s'obtenen sumant o restant múltiples del costat de la caixa central.

Les molècules d'aigua també es mouen al llarg de la dinàmica molecular. No obstant, aquest mètode permet que el nombre de partícules en la caixa central sempre es mantingui constant, ja que si en qualsevol moment de la simulació una partícula abandona la caixa aquesta és reemplaçada per la seva imatge del costat oposat.

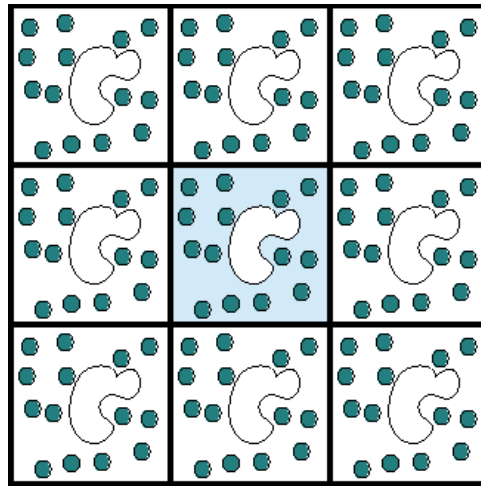
---

<sup>106</sup> Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., *J. Chem. Phys.*, **1983**, 79(2), 926-935.

<sup>107</sup> Mahoney, M. W.; Jorgensen, W. L., *J. Chem. Phys.*, **2000**, 112(20), 8910-8922.

<sup>108</sup> Chen, B.; Xing, J.; Siepmann, J.I., *J. Phys. Chem. B*, **2000**, 104(10), 2391-2401.

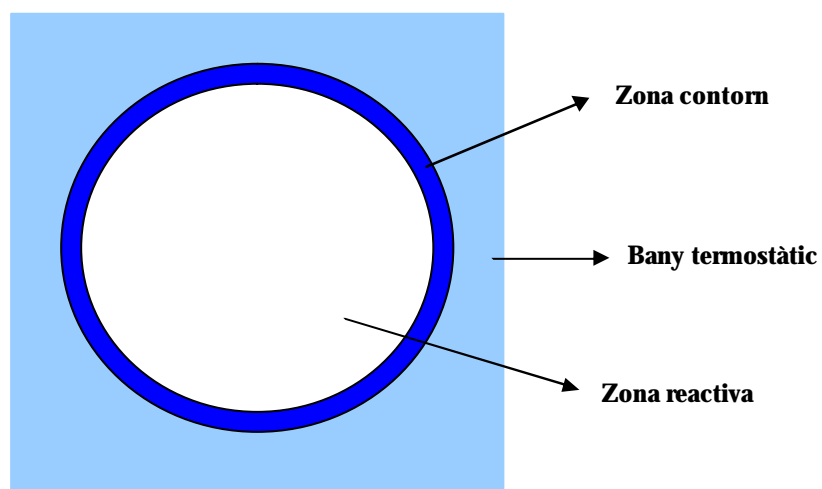
<sup>109</sup> Horn, H.W.; Swope, W.C.; Pitera, J.W.; Madura, J.D.; Dick, T.J.; Hura, G.L.; Head-Gordon, T., *J. Chem. Phys.*, **2004**, 120(20), 9665-9678.



**Figura 13-1 Representació de les caixes de solvatació en dues dimensions en un model de condicions periòdiques.**

### 13.2.2 Condicions estocàstiques

En el model estocàstic de solvatació el solut s'envolta total o parcialment d'una esfera de molècules d'aigua. L'esfera es divideix en dues zones (Figura 13-2): una primera zona (blanc), denominada zona reactiva, on la simulació es duu a terme de la manera habitual utilitzant les equacions de moviment de Newton, i una segona zona (blau fosc), denominada zona de contorn, on la simulació es duu a terme seguint les equacions de moviment de Langevin. Aquest sistema híbrid s'introdueix en una bany termostàtic (blau clar) que manté el sistema en equilibri tèrmic i, a més, s'utilitza un potencial esfèric de contorn per mantenir la correcta distribució mitjana de les molècules d'aigua i evitar que s'escapin cap al buit.



**Figura 13-2 Representació del model estocàstic de solvatació.**



En l'equació de Langevin, la força sobre una partícula es divideix en tres termes (Equació 13-3). El primer terme és degut a la interacció interatòmica entre els àtoms en el sistema (equivalent a la força utilitzada en les equacions de Newton), el segon terme és una força de fricció que descriu l'arrossegament d'una partícula per part del solvent, i el tercer terme és una força aleatòria deguda a les fluctuacions tèrmiques del solvent.

$$m_i a = F\{x_i(t)\} - \mathbf{g} \cdot \mathbf{n}_i m_i + R_i(t) \quad \text{Equació 13-3}$$

Aquesta equació genera un moviment brownià que descriu el moviment d'una partícula sota la influència de les col·lisions aleatòries amb el seu entorn. L'efecte del coeficient de fricció i de les forces aleatòries és una manera aproximada de considerar l'efecte de les molècules d'aigua que estarien fora de l'esfera.

## 14 AMBER

El programa AMBER és un paquet que conté diversos programes i camps de forces (parm91,<sup>110</sup> parm94,<sup>111</sup> parm98,<sup>112</sup> parm99,<sup>113</sup> GAFF<sup>114</sup>). Inicialment va estar dissenyat per realitzar simulacions de dinàmica molecular, encara que, actualment permet realitzar altres tipus de càlculs com per exemple, les pertorbacions d'energia lliure (FEP) o anàlisis vibracionals. El terme "amber" sovint s'utilitza per fer referència als camps de forces implementats en aquest programa, no obstant, el codi i els camps de forces són independents. Altres programes han implementat el camp de forces amber i alguns dels programes d'AMBER permeten treballar amb altres camps de forces.

---

<sup>110</sup> Weiner, S.J.; Kollman, P.A.; Case, D.A.; Singh, U.C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner Jr & P., *J. Chem. Am. Soc.*, **1984**, 106(3), 765-784.

<sup>111</sup> Cornell, W.D.; Cieplak, P.; Bayly, C.; Gould, I.R.; Merz, K.M.; Ferguson, D.M.; Spellmeyer, D.C.; Fox, T.; Caldwell, J.W.; Kollman, P.A., *J. Am. Chem. Soc.*, **1995**, 117(19), 5179-5197.

<sup>112</sup> Cheatham III, T E.; Cieplak, P.; Kollman, P.A., *J. Biomol. Struct. Dyn.*, **1999**, 16(4), 845-862.

<sup>113</sup> Wang, J.; Cieplak, P. Kollman, P.A., *J. Comput. Chem.*, **2000**, 21(12), 1049-1074.

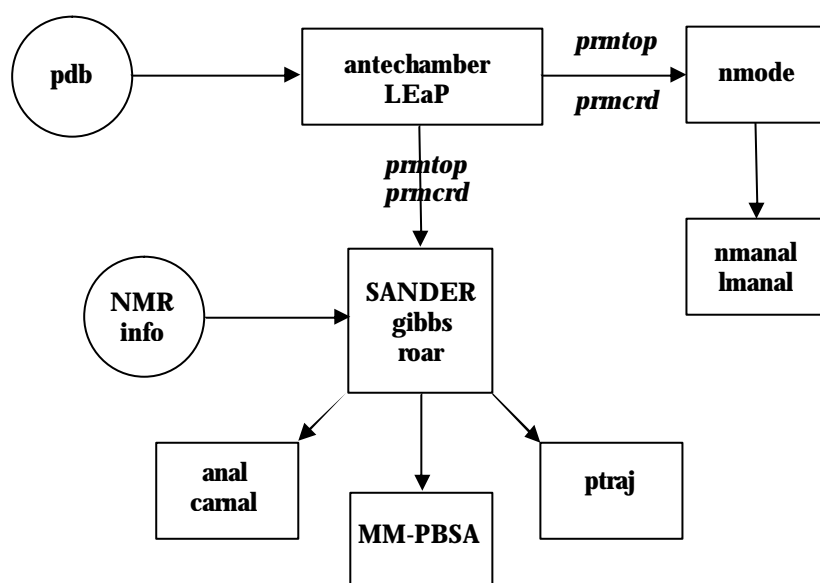
<sup>114</sup> Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A., *J. Comput. Chem.*, **2004**, 25(9), 1157-1174.

En la present tesi doctoral s'han utilitzat les versions v5<sup>115</sup> i v7<sup>116</sup> del programa, però la darrera versió ha estat la més emprada i, per tant, la que es comentarà amb més detall.

Els diferents programes inclosos en Amber es poden classificar en tres grups diferents:

- a. Programes preparatoris: LEaP, antechamber.
- b. Programes energètics: SANDER, gibbs, nmode, roar.
- c. Programes d'anàlisi: anal, ptraj, carnal, nmanal/lmanal, MM-PBSA.

A l'Esquema 14-1 es representa l'esquema general d'AMBER, on es pot observar la relació existent entre els diferents programes. Per iniciar qualsevol simulació es requereixen les coordenades del sistema, que habitualment provenen de l'estructura cristal·lina i estan en format pdb. Amb els programes preparatoris s'obtenen els fitxers de coordenades d'amber (*prmcrd*) i de topologies (*prmtop*), els quals serveixen d'entrada per als programes energètics.



**Esquema 14-1 Esquema general d'AMBER.**

<sup>115</sup> Case, D.A.; Pearlman, D.A.; Caldwell, J.W.; Cheatham III, T.E.; Ross, W.S.; Simmerling, C.L.; Darden, T.A.; Merz, K.M.; Stanton, R.V.; Cheng, A.L.; Vincent, J.J.; Crowley, M.; Ferguson, D.M.; Radmer, R.J.; Seibel, G.L.; Singh, U.C.; Weiner, P.K.; Kollman, P.A., **1997**, AMBER 5, University of California, San Francisco

<sup>116</sup> Case, D.A.; Pearlman, D.A.; Caldwell, J.W.; Cheatham III, T.E.; Wang, J.; Ross, W.S.; Simmerling, C.L.; Darden, T.A.; Merz, K.M.; Stanton, R.V.; Cheng, A.L.; Vincent, J.J.; Crowley, M.; Tsui, V.; Gohlke, H.; Radmer, R.J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G.L.; Singh, U.C.; Weiner, P.K.; Kollman, P.A., **2002**, AMBER 7, University of California, San Francisco.

## 14.1 PROGRAMES PREPARATORIS

### 14.1.1 Obtenció de les càrregues. Metodologia RESP

La filosofia d'AMBER es basa en l'acurada representació de les interaccions electrostàtiques, la qual és crucial per a l'aplicació amb èxit d'aquest camp de forces. Les càrregues atòmiques de la molècula s'obtenen mitjançant la metodologia RESP<sup>117</sup> (*Restrained ElectroStatic Potential*), que és una modificació de la metodologia original ESP.<sup>118</sup> Aquesta metodologia es basa en ajustar les càrregues centrades en els nuclis atòmics, per un mètode de mínims quadrats, que reproduïx el potencial electrostàtic de la molècula. A més, s'utilitzen diferents conformacions amb l'objectiu de que les càrregues obtingudes siguin independents de la conformació adoptada per la molècula.

En primer lloc, s'optimitza la geometria de la molècula amb mètodes de mecànica quàntica amb un nivell de teoria HF i fent servir una base de baixa qualitat STO-3G, i a continuació, s'avalua el potencial electrostàtic fent servir una base de major grandària 6-31G\* (veure fitxers d'exemple en l'apèndix). Un cop obtingut el potencial electrostàtic es calculen les càrregues en dues etapes diferents. En la primera etapa es calcula la càrrega de tots els àtoms de la molècula i en la segona, es mantenen fixes les càrregues dels àtoms interns (àtoms pesats) i s'ajusta la càrrega dels àtoms externs, especialment la dels hidrògens. Un fet important a tenir en compte en aquesta segona etapa és que al reajustar la càrrega dels tres hidrògens d'un grup metil també s'ha de reajustar la càrrega del carboni metílic i igualment, s'hauria de fer amb el carboni d'un grup *tert*-butil. En totes dues etapes es poden imposar restriccions dins d'una mateixa molècula deguda a l'equivalència per rotació, com per exemple els tres hidrògens d'un grup metil, o restriccions entre àtoms de les diferents conformacions per a què àtoms equivalents tinguin la mateixa càrrega final. També és possible assignar una càrrega determinada en una part concreta de la molècula (veure exemple de l'apèndix).

### 14.1.2 LEaP

LEaP és un acrònim construït a partir de les inicials dels programes que substitueix de versions més antigues d'AMBER: *Link*, *Edit and Prep* i és el programa per crear un nou sistema.

---

<sup>117</sup> (a) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A., *J. Phys. Chem.*, **1993**, 97(40), 10269-10280. (b) Cornell, W.D.; Cieplak, P.; Bayly, C.I.; Kollman, P.A., *J. Am. Chem. Soc.*, **1993**, 115(21), 9620-9631.

<sup>118</sup> (a) Momany, F.A., *J. Phys. Chem.*, **1978**, 82(5) 592-601. (b) Cox, S.R.; Williams, D.E., *J. Comput. Chem.*, **1981**, 2(3), 304-323. (c) Singh, U.C.; Kollman, P.A., *J. Comput. Chem.*, **1984**, 5(2), 129-145.

L'entitat fonamental d'aquest programa són els objectes, els quals poden ser simples (números i cadenes) o complexos (àtoms, residus i unitats). Els objectes complexos tenen propietats que es poden manipular i poden contenir altres objectes dintre seu. Els àtoms són similars als àtoms químics i són objectes que es poden unir entre ells per formar molècules. Tenen diverses propietats com per exemple el nom, el tipus, la càrrega, l'element i la posició. Els residus estan formats per un conjunt d'àtoms i o bé poden ser una molècula, o bé es poden unir entre ells per formar molècules. Per exemple, la glucosa és un residu per formar les ciclodextrines i els aminoàcids són residus per formar les proteïnes. Les propietats dels residus són les connexions (àtoms que connecten un residu amb l'altre), el tipus (indefinit, solvent, proteïna, nucleic o sacàrid) i el nom. Finalment, les unitats són els objectes més complexos i els més importats. Aquestes representen al sistema que es vol estudiar com per exemple, la ciclodextrina o la proteïna.

Així doncs, amb LEaP es crea la unitat a partir dels fitxers *prep* d'entrada, i s'assignen les coordenades atòmiques al sistema a partir d'un fitxer amb format *pdb*. A continuació, si es necessari, es solvata el sistema i s'assignen les pertorbacions si es vol realitzar un càlcul de pertorbació d'energia. Finalment, s'assignen els paràmetres del camp de forces al sistema a estudiar. D'aquest procés s'obtenen els fitxers de paràmetres i topologies necessaris per als programes energètics.

### 14.1.3 Antechamber

Antechamber és un nou mòdul implementat en la versió 7 d'AMBER i està dissenyat per ser utilitzat conjuntament amb el camp de forces GAFF (*General Amber Force Field*). Antechamber permet preparar els fitxers *prep* d'entrada per a LEaP començant per una estructura tridimensional, la qual pot tenir diversos formats. D'altra banda, permet assignar les càrregues atòmiques mitjançant diverses metodologies i els tipus d'àtoms amb el programa *atmtype* inclòs en aquest mòdul. Finalment, en cas de ser necessari, també es poden assignar els paràmetres que manquen amb un altre programa inclòs dins d'aquest mòdul, el *prmchk*

Entre la diversitat de formats que accepta antechamber cal destacar alguns dels formats més emprats com per exemple *mol2*, *pdb*, *prep* (coordenades internes i cartesianes), *z-matrix*, *mopac* (coordenades internes i cartesianes) i *output* de Gaussian, entre d'altres. I els mètodes per assignar càrregues atòmiques també són diversos, entre els quals cal destacar els tipus de càrrega RESP, AM1-BCC, ESP, Mulliken i Gasteiger, entre d'altres.

La combinació d'aquest programa amb el nou camp de forces, permet que AMBER pugui ampliar les seues camps d'aplicació i fer-lo apte per estudiar problemes de complexació de lligands a proteïnes i de disseny de nous fàrmacs.

## 14.2 PROGRAMES ENERGÈTICS

### 14.2.1 SANDER

El programa SANDER (*Simulated Annealing with NMR-Derived Energy Restraints*) és el principal programa per dur a terme minimitzacions i simulacions de dinàmica molecular. En la dinàmica molecular es guarden conformacions a intervals regulars de temps, i en la minimització es guarden les coordenades finals. Mitjançant una gran diversitat de comandes es poden dur a terme cerques conformacionals i simulacions de dinàmica molecular més elaborades, on, per exemple, es poden incloure diversos tipus de restriccions. En l'apèndix es mostren fitxers d'exemple per realitzar una minimització i una simulació de dinàmica molecular amb aquest programa.

### 14.2.2 Gibbs, nmode i roar

Gibbs és el programa per realitzar perturbacions d'energia lliure. En aquesta metodologia es transforma un sistema inicial en un sistema final, similar al primer, mitjançant petits passos que connecten l'un amb l'altre a través de sistemes intermedis que no necessàriament han de ser intermedis físics. Gibbs utilitza el conjunt de conformacions generades per calcular la diferència d'energia lliure entre els dos estats inicial i final els quals s'han de definir amb LEaP.

Nmode és un programa que permet realitzar càlculs de mecànica molecular utilitzant mètodes de minimització de primeres i segones derivades per trobar punts de mínima energia i estats de transició. Nmode també permet realitzar anàlisis vibracionals per determinar els modes normals de vibració.

Roar és molt similar a SANDER, però la diferencia més rellevant és que permet realitzar càlculs del tipus QM/MM.

## 14.3 PROGRAMES D'ANÀLISI

### 14.3.1 Anal

Anal és un programa que permet analitzar l'energia per mecànica molecular d'una única estructura o conformació del sistema. Una característica important és que permet descompondre

els termes energètics entre diferents grups d'àtoms, amb l'objectiu de determinar les energies d'interacció entre diferents parts de la molècula. Si es realitza l'anàlisi sobre el sistema construït amb LEaP, es poden detectar errors en l'estructura de partida que podrien conduir a augments energètics indesitjables en els posteriors càlculs o simulacions.

### **14.3.2 Ptraj**

Ptraj és el programa per analitzar i manipular les trajectòries obtingudes de les simulacions de dinàmica molecular. Són moltes les opcions que permet fer aquest programa entre les quals cal destacar la possibilitat d'extreure coordenades d'una trajectòria, la superposició d'estructures i detectar fluctuacions en les posicions atòmiques. En l'apèndix es mostren fitxers d'exemple de la primera opció.

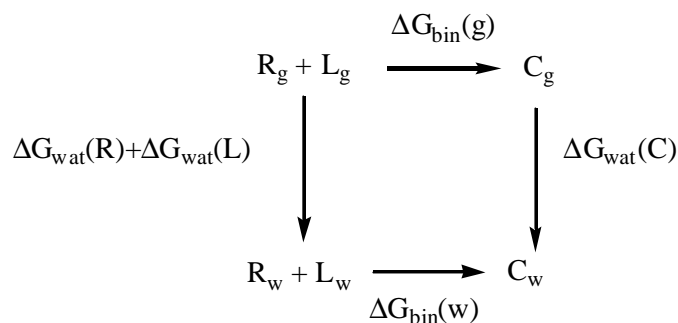
### **14.3.3 Carnal**

Carnal és un altre programa que permet analitzar les trajectòries obtingudes de les simulacions de dinàmica molecular. Entre totes les anàlisis que permet fer cal destacar la dels paràmetres geomètrics com distàncies, angles d'enllaç, i de diedres, i també la possibilitat d'analitzar enllaços d'hidrogen. Aquest programa ha estat àmpliament utilitzat en aquesta tesi per realitzar les anàlisis de les simulacions de DM. En l'apèndix es detallen exemples de fitxers per realitzar anàlisis de les distàncies, enllaços d'hidrogen i del RMS.

### **14.3.4 MM/PBSA**

MM/PBSA permet avaluar energies lliure d'interacció o calcular energies lliures absolutes de les molècules en solució aquosa. Les estructures a analitzar es prenen habitualment de la trajectòria obtinguda d'una simulació de dinàmica molecular. Aquesta metodologia combina els càlculs de mecànica molecular amb l'aproximació continua del solvent que es pot calcular a partir de l'equació Poisson-Boltzman (PB) donant lloc a la metodologia MM/PBSA, o a partir de l'equació generalitzada de Born (GB) donant lloc a la metodologia MM/GBSA .

Per calcular l'energia de complexació (*binding*) s'ha de tenir en compte el cicle termodinàmic representat a la Figura 14-1, on R fa referència al receptor, L fa referència al lligand i C fa referència al complex. El subíndex g indica que les espècies estan en fase gas, i el subíndex w indica que les espècies estan en solució aquosa.



**Figura 14-1** Cicle termodinàmic per calcular energies de complexació.

A partir del cicle termodinàmic es pot calcular la diferència en l'energia lliure d'interacció d'un complex en solució aquosa i en fase gas a partir de l'expressió:

$$\Delta\Delta G_{bin} = \Delta G_w(C) - (\Delta G_w(R) + \Delta G_w(L))$$

On cada component  $\Delta G_w$  representa l'energia absoluta de l'espècie en solució aquosa i conté tres termes energètics:

$$\Delta G_w(i) = E_{gas} + \Delta G_{solv} - TS$$

$E_{gas}$  fa referència a l'energia calculada a partir de la mecànica molecular i es calcula amb SANDER. Aquesta inclou el terme de l'energia interna ( $E_{int}$ ), el corresponent a l'energia de van der Waals ( $E_{vdW}$ ) i de l'energia electrostàtica ( $E_{el}$ ).  $\Delta G_{solv}$  és l'energia lliure de solvatació, i aquest terme es pot expressar com:

$$\Delta G_{solv} = G_{GB} + G_{nopolar}$$

On  $G_{GB}$  és la contribució electrònica a l'energia lliure de solvatació (es calcula amb SANDER fent servir l'equació generalitzada de Born) i  $G_{nopolar}$  és la contribució de les interaccions no polars a l'energia lliure de solvatació que depèn de la superfície accessible al solvent (s'avalua amb el programa *molsurf* també implementat en AMBER). Finalment, el terme entròpic  $TS$  s'avalua amb el programa Nmode, tot i que només es pot obtenir la contribució del solut ja que el càlcul del solvent és computacionalment molt costós.

## 15 CHARMM

L'objectiu principal del programa CHARMM és la modelització de sistemes macromoleculars. Igual que AMBER, es basa en la funció empírica d'energia la qual s'utilitza per realitzar minimitzacions de l'energia, simulacions de dinàmica molecular, anàlisis vibracionals i càlculs termodinàmics. A més a més, existeix una interfase cap a diversos programes de mecànica quàntica que permet realitzar càlculs híbrids del tipus QM/MM. D'altra banda, el programa inclou funcions que permeten realitzar diverses anàlisis com per exemple comparar estructures, avaluar l'energia d'un sistema, estudiar l'evolució d'un paràmetre al llarg del temps i determinar les funcions de correlació.

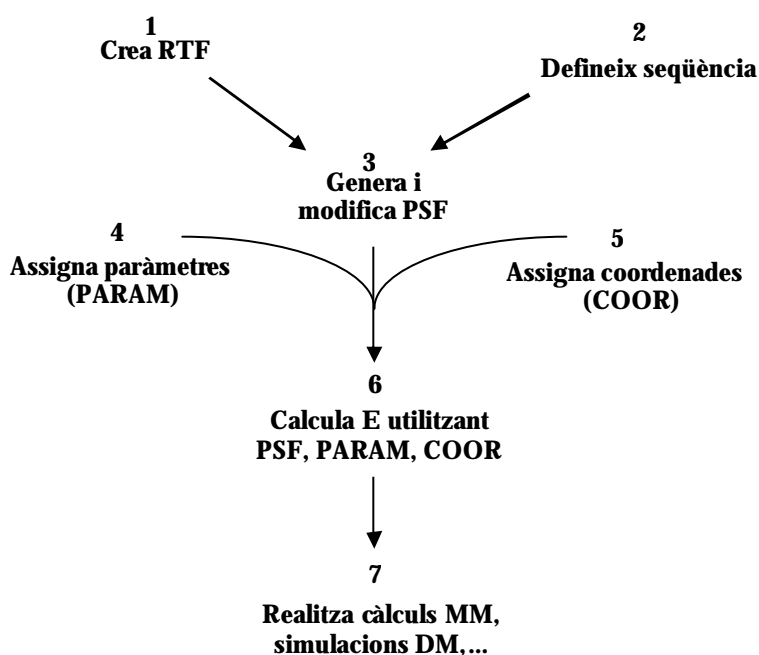
CHARMM és un programa que també es controla mitjançant comandes. Habitualment en el fitxer de comandes o *input* s'inclouen les comandes que guiaran al programa a efectuar una sèrie d'operacions seqüencials. Normalment l'ordre de les comandes ve limitat per les dades que requereix l'operació, així doncs, no es podrà calcular l'energia si la informació sobre les coordenades i els paràmetres no s'han llegit prèviament en el programa. Un dels punts més importants per obtenir resultats satisfactoris en els posteriors càlculs de mecànica molecular i simulacions de dinàmica molecular, és la construcció del sistema de partida. A continuació s'explica breument els conceptes i passos més importants que s'han de tenir en compte per generar nous sistemes amb aquesta programa.

### 15.1 GENERACIÓ DE NOUS SISTEMES

En sistemes macromoleculars és habitual trobar subgrups d'àtoms que es repeteixen al llarg del sistema. CHARMM crea aquests subgrups i, igual que AMBER, també els anomena residus. La informació local sobre àtoms, enllaços, angles i demés paràmetres de cada residu es troben en el fitxer de topologies del residu (RTF). Per tant, en el cas de treballar amb residus poc habituals, com per exemple la glucosa en el cas de les ciclodextrines, el primer pas és generar el RTF per a aquests residus (veure exemple en l'Apèndix). Així doncs, totes les topologies dels residus que siguin necessàries per construir el sistema final s'inclouen en un mateix RTF.



Un cop definits els residus (Esquema 15-1), es defineix la seqüència de residus dins l'estructura final i posteriorment, es crea el fitxer d'estructura de la proteïna (PSF).<sup>119</sup> Aquest fitxer té una organització jeràrquica; els àtoms formen grups, els grups formen residus i els residus segments, la unió de tots els segments formen l'estructura final. A més, aquest fitxer es pot modificar mitjançant els residus d'ajust (*patch residues*), els quals també estan definits en el fitxer de topologies. El tipus de modificacions que es poden dur a terme són molt diverses i depenen del sistema a estudiar. En el cas de la present tesi doctoral s'han utilitzat *patches* per unir els diferents residus de glucosa entre si, i també per unir el residu metilnaftil a una de les glucoses de la ciclodextrina. En l'apèndix es detalla un exemple d'utilització d'aquests *patches*. Per tant, és important remarcar que el PSF conté informació del que serà el model final i és pot guardar en un fitxer extern i utilitzar per a posteriors càlculs i simulacions. Finalment, s'assignen les coordenades dels àtoms i els paràmetres al PSF i s'obté el sistema a estudiar sobre el que es podran realitzar els càlculs MM i simulacions de dinàmica molecular.



**Esquema 15-1 Passos a seguir per construir un nou sistema amb CHARMM.**

<sup>119</sup> El nom d'aquest fitxer, *Protein Structure File*, es pot entendre perquè inicialment el programa CHARMM va estar dissenyat per a treballar amb sistemes d'interès biològic. Malgrat actualment existeixen camps de forces que permeten treballar amb altres famílies de molècules, com per exemple els carbohidrats, s'ha mantingut el nom del fitxer.



## **VI. APÈNDIX**

---



## APÈNDIX A: TAULES DE NOMENCLATURA

### A.1 NOMENCLATURA DE LES ESTRUCTURES OBTINGUDES AMB EL CAMP DE FORCES PARM94

**Taula A1 Nomenclatura de les estructures de la CA14 amb el camp de forces parm94.**

|         |      | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---------|------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|
| Inicial |      | s | s | s | s | s | s | a | s | s | s  | s  | s  | s  | a  |
| F. gas  | 300  | - | s | s | - | - | - | a | - | s | s  | -  | -  | -  | a  |
|         | 3000 | - | s | s | - | - | - | a | - | s | s  | -  | -  | -  | a  |
| Dis.    | 500  | - | s | s | - | - | s | + | - | s | s  | -  | s  | -  | +  |
|         | 3500 | - | s | s | - | - | - | + | - | s | +  | -  | +  | -  | +  |

**Taula A2 Nomenclatura de les estructures de la CA21 amb el camp de forces parm94.**

|             |      | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 |
|-------------|------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|
| Inicial     |      | s | s | s | a | a | a | - | s | + | -  | s  | s  | -  | -  | -  | s  | -  | s  | s  | a  | -  |
| F. gas      | 1000 | s | - | s | - | s | + | - | s | + | -  | -  | s  | a  | a  | -  | s  | -  | s  | a  | -  | s  |
|             | 4000 | s | - | - | - | s | + | - | s | a | -  | -  | -  | -  | -  | -  | s  | -  | s  | a  | s  | -  |
| Dis. aquosa |      | s | - | s | - | - | + | - | s | + | -  | s  | -  | -  | -  | -  | s  | -  | s  | -  | -  | -  |

**Taula A3 Nomenclatura de les estructures de la CA26 amb el camp de forces parm94.**

|             |      | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 |   |
|-------------|------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|---|
| Inicial     |      | s | s | s | s | s | s | s | s | s | s  | s  | s  | a  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | a |
| F. gas      | 500  | s | s | a | s | s | a | - | s | - | s  | -  | s  | +  | -  | s  | a  | s  | s  | s  | -  | s  | s  | -  | s  | -  | a  |   |
|             | 4000 | s | s | a | s | a | s | - | - | - | s  | -  | -  | +  | -  | s  | a  | s  | -  | -  | a  | -  | s  | -  | -  | -  | -  |   |
| Dis. aquosa |      | s | - | s | s | - | s | - | s | s | -  | +  | -  | +  | s  | -  | -  | s  | -  | -  | s  | -  | s  | s  | -  | s  | a  |   |

**Taula A4 Nomenclatura de les estructures de la CA28 amb el camp de forces parm94.**

|             |      | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 |   |
|-------------|------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|---|
| Inicial     |      | s | s | s | s | s | s | s | s | s | s  | s  | s  | a  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | -  | s  | s  | a |
| F. gas      | 1000 | s | - | - | + | - | s | s | - | s | -  | -  | s  | +  | s  | -  | -  | s  | s  | -  | s  | s  | -  | s  | s  | -  | s  | s  | +  |   |
|             | 2000 | s | - | - | + | - | s | s | - | s | -  | -  | s  | -  | s  | -  | -  | +  | s  | -  | s  | s  | -  | s  | s  | s  | -  | s  | -  |   |
|             | 4000 | s | - | - | + | - | s | s | - | s | -  | -  | s  | -  | s  | -  | -  | +  | s  | -  | s  | s  | -  | -  | s  | -  | -  | s  | -  |   |
| Dis. aquosa | 50   | + | - | s | - | s | - | s | - | s | s  | -  | s  | a  | s  | -  | s  | s  | -  | s  | s  | -  | s  | -  | s  | s  | -  | s  | +  |   |
|             | 3000 | + | - | s | - | s | - | s | - | s | s  | -  | s  | a  | s  | -  | s  | s  | -  | s  | s  | -  | s  | -  | s  | s  | -  | s  | +  |   |

## A.2 NOMENCLATURA DE LES ESTRUCTURES OBTINGUDES AMB EL CAMP DE FORCES PARM99

**Taula A5 Nomenclatura de les estructures de la CA14 amb el camp de forces parm99.**

|             | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|-------------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|
| Inicial     | s | s | s | s | s | s | a | s | s | s  | s  | s  | s  | a  |
| Fase gas    | a | s | s | s | s | s | a | - | - | s  | s  | s  | s  | -  |
| 158         | - | s | - | s | - | - | + | s | s | s  | -  | s  | -  | a  |
| 381         | - | s | s | s | - | - | + | - | s | s  | -  | s  | -  | a  |
| Dis. 676    | - | s | - | s | - | s | + | - | s | -  | -  | s  | s  | a  |
| aquosa 1125 | - | - | - | s | s | s | a | - | s | -  | -  | s  | s  | +  |
| 2160        | - | s | - | s | s | - | a | - | s | -  | s  | s  | -  | +  |
| 3367        | - | s | s | s | a | - | s | s | - | s  | s  | a  | -  | s  |

**Taula A6 Nomenclatura de les estructures de la CA21 amb el camp de forces parm99.**

|             | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 |
|-------------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|
| Inicial     | s | s | s | a | a | a | - | s | + | -  | s  | s  | -  | -  | -  | s  | -  | s  | s  | a  | -  |
| F. gas      | s | s | s | s | s | a | s | s | a | -  | s  | s  | s  | s  | a  | s  | s  | s  | s  | -  | -  |
| 150         | s | s | s | - | s | a | - | s | + | -  | -  | s  | -  | -  | a  | s  | -  | s  | -  | s  | -  |
| 434         | s | - | s | - | - | a | s | s | a | -  | -  | s  | s  | -  | -  | -  | s  | -  | -  | -  | -  |
| 1150        | s | - | s | - | - | - | - | s | + | -  | s  | -  | -  | -  | -  | s  | s  | s  | -  | -  | s  |
| Dis. 1864   | s | s | - | - | s | - | s | - | a | -  | -  | s  | -  | -  | s  | -  | s  | -  | -  | -  | -  |
| aquosa 2285 | s | s | - | - | s | s | s | - | + | -  | s  | -  | -  | -  | s  | s  | s  | -  | -  | -  | -  |
| 2492        | s | s | - | - | - | s | - | - | a | -  | s  | s  | -  | -  | -  | -  | s  | -  | -  | -  | -  |
| 2840        | - | s | - | - | - | - | s | - | a | -  | s  | -  | s  | -  | -  | -  | -  | s  | s  | -  | -  |
| 3881        | s | s | - | - | - | - | s | - | a | s  | s  | s  | s  | s  | -  | -  | -  | s  | -  | -  | -  |

**Taula A7 Nomenclatura de les estructures de la CA26 amb el camp de forces parm99.**

|             | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 |
|-------------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Inicial     | s | s | s | s | s | s | s | s | s | s  | s  | s  | a  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | a  |
| F. gas      | - | s | s | s | a | s | - | s | s | -  | s  | s  | a  | s  | s  | s  | s  | s  | s  | -  | s  | s  | s  | s  | s  | -  |
| 650         | - | s | - | s | s | - | s | s | - | s  | -  | s  | a  | s  | -  | -  | s  | s  | s  | s  | s  | s  | s  | s  | s  | a  |
| Dis. 2157   | - | s | s | s | s | - | s | s | - | s  | -  | s  | a  | -  | -  | -  | s  | s  | s  | s  | s  | s  | s  | -  | s  | a  |
| aquosa 3451 | - | s | s | - | s | - | s | s | s | s  | s  | a  | s  | -  | -  | -  | -  | s  | -  | s  | s  | s  | s  | s  | s  | a  |
| 4046        | - | s | s | - | s | s | s | - | s | s  | s  | a  | s  | -  | -  | s  | -  | s  | s  | s  | s  | s  | s  | s  | s  | a  |
| 4952        | - | s | s | - | - | s | s | - | s | s  | +  | s  | -  | -  | -  | -  | s  | -  | s  | s  | s  | s  | s  | s  | s  | a  |

**Taula A8 Nomenclatura de les estructures de la CA28 amb el camp de forces parm99.**

|             | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 |
|-------------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Inicial     | s | s | s | s | s | s | s | s | s | s  | s  | s  | a  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | -  | s  | s  | a  |
| F. gas      | s | s | - | s | s | s | - | s | s | s  | s  | s  | a  | s  | -  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | -  | a  |
| 365         | s | s | s | s | s | - | s | s | s | s  | s  | s  | a  | s  | s  | -  | s  | -  | s  | s  | -  | s  | s  | -  | s  | -  | s  | a  |
| Dis. 2507   | - | s | - | - | s | s | s | s | s | -  | s  | s  | a  | s  | -  | -  | s  | s  | -  | s  | s  | s  | s  | s  | s  | s  | s  | a  |
| aquosa 3067 | s | s | s | - | s | - | s | - | s | s  | -  | s  | a  | s  | s  | s  | s  | s  | -  | s  | s  | -  | s  | s  | s  | s  | s  | a  |
| 3917        | - | s | - | s | - | - | s | s | s | s  | -  | s  | +  | s  | s  | -  | s  | +  | -  | s  | s  | -  | s  | s  | -  | s  | a  |    |
| 4659        | - | s | - | s | - | s | - | s | s | s  | -  | s  | a  | s  | -  | s  | s  | +  | -  | s  | s  | s  | -  | s  | s  | s  | a  |    |
| 4950        | - | s | - | s | s | - | s | s | - | s  | -  | s  | a  | s  | -  | -  | s  | s  | -  | s  | s  | s  | s  | s  | s  | s  | a  |    |

### A.3 NOMENCLATURA DE LES ESTRUCTURES OBTINGUDES AMB EL CAMP DE FORCES GLYCAM 2000A

**Taula A9 Nomenclatura de les estructures de la CA14 amb el camp de forces glycam 2000a.**

|             |         | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|-------------|---------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|
|             | Inicial | s | s | s | s | s | s | a | s | s | s  | s  | s  | s  | a  |
| F. gas      | 700     | s | s | s | s | s | s | + | - | s | s  | -  | -  | s  | s  |
|             | 4000    | s | s | s | s | s | s | + | - | s | s  | -  | -  | s  | s  |
| Dis. aquosa | 186     | s | s | s | - | - | s | + | s | s | s  | s  | s  | -  | +  |
|             | 755     | s | s | s | - | s | - | + | s | s | s  | -  | -  | s  | +  |
|             | 1725    | - | s | s | - | s | s | s | s | s | s  | s  | -  | s  | +  |
|             | 3820    | - | - | - | s | s | s | s | s | s | s  | -  | -  | s  | +  |

**Taula A10 Nomenclatura de les estructures de la CA21 amb el camp de forces glycam 2000a.**

|             |         | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 |   |
|-------------|---------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|---|
|             | Inicial | s | s | s | a | a | a | - | s | + | -  | s  | s  | -  | -  | -  | s  | -  | s  | s  | a  | -  |   |
| F. gas      | 2000    | s | s | s | s | s | s | s | s | s | -  | s  | s  | s  | -  | s  | s  | s  | s  | s  | s  | a  | - |
|             | 4500    | s | s | s | s | s | s | s | s | s | -  | -  | s  | s  | -  | -  | s  | s  | s  | s  | s  | a  | - |
| Dis. aquosa | 60      | s | s | - | - | - | - | s | + | s | -  | s  | -  | -  | -  | -  | s  | s  | -  | -  | s  | -  |   |
|             | 334     | s | s | s | - | - | - | s | - | a | s  | s  | -  | s  | -  | -  | s  | s  | -  | -  | -  | -  |   |
|             | 2025    | s | s | s | - | - | - | s | s | + | s  | -  | -  | -  | -  | -  | s  | s  | -  | -  | s  | -  |   |
|             | 2570    | s | - | s | s | - | s | - | - | + | -  | s  | -  | a  | s  | -  | s  | s  | -  | -  | -  | s  |   |
|             | 3930    | s | s | s | - | s | - | s | - | + | -  | -  | -  | -  | -  | -  | s  | s  | s  | -  | s  | -  | s |

**Taula A11 Nomenclatura de les estructures de la CA26 amb el camp de forces glycam 2000a.**

|             |         | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 |
|-------------|---------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
|             | Inicial | s | s | s | s | s | s | s | s | s | s  | s  | s  | a  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | a  |
| F. gas      | 100     | s | s | s | s | s | s | s | s | s | s  | s  | s  | a  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | -  |
|             | 1200    | s | s | s | s | s | s | - | s | s | s  | s  | s  | a  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  |
|             | 3500    | s | s | s | s | s | s | s | s | s | s  | s  | s  | a  | s  | s  | s  | -  | s  | s  | s  | s  | s  | s  | s  | -  | s  |
| Dis. aquosa | 80      | s | s | s | s | s | - | s | s | - | s  | s  | s  | a  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | a  |
|             | 421     | s | s | s | s | - | s | s | s | s | s  | s  | a  | -  | s  | -  | s  | s  | -  | -  | s  | s  | -  | s  | s  | a  |    |
|             | 1138    | s | - | - | s | s | s | s | s | s | s  | s  | s  | a  | -  | s  | s  | -  | s  | s  | -  | s  | s  | s  | s  | s  | +  |
|             | 1469    | s | - | s | - | s | s | - | s | s | s  | s  | s  | a  | s  | s  | s  | s  | -  | -  | -  | s  | s  | s  | s  | s  | +  |
|             | 2693    | s | - | s | s | s | s | - | - | s | -  | s  | +  | -  | s  | -  | -  | s  | -  | -  | -  | -  | s  | s  | s  | s  | a  |
|             | 4250    | s | - | - | s | s | s | s | - | - | -  | s  | +  | s  | -  | -  | s  | s  | -  | -  | -  | -  | s  | s  | s  | s  | a  |

**Taula A12 Nomenclatura de les estructures de la CA28 amb el camp de forces glycam 2000a.**

|         | 1    | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 |   |
|---------|------|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|---|
| Inicial | s    | s | s | s | s | s | s | s | s | s  | s  | s  | a  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | -  | s  | s  | a |
| F. gas  | 50   | s | - | s | s | s | s | s | s | s  | s  | s  | a  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | -  | s  | a |
|         | 500  | s | - | s | s | s | s | s | s | s  | s  | s  | a  | -  | s  | s  | s  | s  | s  | s  | s  | -  | s  | s  | s  | s  | s  | -  | a |
|         | 3500 | - | - | s | s | - | + | s | s | s  | s  | s  | a  | -  | s  | s  | s  | s  | s  | s  | s  | -  | s  | s  | -  | s  | s  | s  | a |
| Dis.    | 145  | s | s | - | s | s | s | s | s | s  | s  | s  | a  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | a |
|         | 671  | s | s | - | - | s | s | s | s | s  | s  | s  | a  | -  | s  | -  | s  | -  | s  | -  | s  | s  | s  | s  | s  | s  | s  | s  | a |
| aquosa  | 3291 | - | - | - | - | - | s | s | s | s  | s  | s  | a  | -  | s  | s  | s  | -  | s  | -  | s  | -  | s  | -  | s  | -  | s  | -  | + |
|         | 4930 | - | - | - | s | - | - | s | s | s  | s  | s  | a  | -  | s  | s  | -  | s  | s  | -  | s  | -  | -  | s  | s  | s  | s  | a  |   |

#### A.4 NOMENCLATURA DE LES ESTRUCTURES OBTINGUDES AMB EL CAMP DE FORCES MM3\*

**Taula A13 Nomenclatura de les conformacions de les cicloamiloses obtingudes amb el camp de forces MM3\*.**

|      | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 |   |
|------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|---|
| CA14 | s | s | s | s | s | s | a | s | s | s  | s  | s  | a  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |   |
| CA21 | s | s | s | - | - | a | - | - | a | -  | s  | -  | s  | -  | -  | s  | s  | s  | s  | s  | -  |    |    |    |    |    |    |    |   |
| CA26 | s | s | s | s | s | s | s | s | s | s  | s  | s  | a  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | -  | a  |    |    |   |
| CA28 | s | s | s | s | s | s | s | s | s | s  | s  | s  | a  | -  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | s  | -  | s  | - |



## APÈNDIX B: PROGRAMES FORTRAN COMB.F I SUMAFRAGS.F

### B.1 COMB.F

```

C
integer r, j(10)
C
OPEN(UNIT=2,FILE='OUTPUT',STATUS='UNKNOWN')
C
write(6,*) 'Calcul de combinacions de n elements agafats'
write(6,*) 'de 1 en 1 fins a de r en r'
write(6,*) 'Escriu n='
read(5,*) n
write(6,*) 'Escriu r='
read(5,*) k
do r=1,k
call allnr(n,r,j,ifault)
end do
C
stop
end
C
C
C
subroutine allnr(n, r, j, ifault)
c
c   Algorithm AS 88 Appl. Statist. (1975) Vol.24, No. 3
c
c   When called once, generates all possible combinations
c   from a group of N items. Each combination (represented in j as
c   r ordered integers between 1 and n) is processed within allnr.
c
c   Parameters:-
c
c   n      integer          input:  The size of the group from which
c                                   the combinations are selected.
c
c   r      integer          input:  The size of each combination.
c
c   j      integer array(r) workspace: Used by allnr to store
c                                   combinations.
c
c   ifault integer          output: Fault indicator, equal to:
c                                   0 if 1 ≤ R ≤ N;
c                                   1 otherwise.
c
integer r, j(r)
c
ifault = 1
if (r .lt.1 .or. r .gt. n) return
ifault = 0
kount = 0
nmr = n - r
c
c   Initialize J(1) to lower limit separately, since lower limit for
c   each index depends on lower limit for previous index
c
i = 1
j(1) = 1
c
c   Initialize indices for loops i=1,...,r to lower limits
c
1 if (i .eq. r) goto 3
ipl = i + 1
do 2 l = ipl, r
2 j(l) = j(l - 1) + 1
c
c   Update the count (kount) of combinations and process the current
c   combination. The call to Subroutine job may be replaced by
c   statements to process the current combination.
c
3 kount = kount + 1

```

```

        call job(n, r, j, kount)
c
c      Increment the first possible index (of loop i) among indices of
c      loops R, R-1,...,1
c
        i = r
4      if (j(i) .lt. nmr + i) goto 5
        i = i - 1
c
c      Return after all indices have achieved their upper limits
c
        if (i .le. 0) return
        goto 4
5      j(i) = j(i) + 1
        goto 1
        end
c
c
c
        subroutine job (n, r, j, kount)
        integer r, j(r)
        write(2,10) j
10      format(10I5)
        return
        end

```

## B.2 SUMAFRAGS.F

```

C23456789012345678901234567890123456789012345678901234567890123456789012
C
        COMMON SUM(6), NAA(1000),FRAG(1000),NSEC(1000)
C
        OPEN(UNIT=1,FILE='INPUT1',STATUS='OLD')
        OPEN(UNIT=2,FILE='INPUT2',STATUS='OLD')
        OPEN(UNIT=6,FILE='OUTPUT',STATUS='UNKNOWN')
C
C INPUT1:ARXIU QUE CONTE LES COMBINACIONS DE N ELEMENTS AGAFATS
C      DE 1 EN 1 FINS AL MAXIM DE 5 EN 5 (format 5I5)
C Example:
C   82 (de 1 en 1)
C   24  32 (de 2 en 2)
C   12  56  74 (de 3 en 3)
C   etc...
C
C INPUT2:ARXIU QUR CONTE EL PUNT DE TALL, LA MASSA DE CADA FRAGMENT
C      I UNA MARCA QUE INDICA LA SEVA ESTRUCT SECUNDARIA
C      (1 SI ES POT TRENCA I RES SI NO POT)
C      FORMAT:I5,F10.5,I5
C Example:
C   3  349.1904  1
C  25 2311.1209  1
C  36  1228.628  1
C  49 1487.7012  1
C  56  687.4035
C  71 1658.8496
C OUTPUT: ARXIU DE SORTIDA I NO HI HA MES!
C
C
        DO NFRAG=1,1000
        READ(2,20,END=99) NAA(NFRAG),FRAG(NFRAG),NSEC(NFRAG)
20      FORMAT(I5,F10.5,I5)
        END DO
99      NFRAG=NFRAG-1
C
        WRITE(6,11)
11      FORMAT(' Masa (Da) Fragment           Aminoacids', /,
1         ' -----' )
C
1      READ(1,10,END=999) I,J,K,L,M
10      FORMAT(5I5)
C
        DO KK=1,6

```

```

        SUM(KK)=0.
    END DO
C
    IF (M.EQ.0) THEN
        GO TO 40
    ELSE
        DO KK=1,I
            SUM(1)=SUM(1)+FRAG(KK)
        END DO
        DO KKK=I+1,J
            SUM(2)=SUM(2)+FRAG(KKK)
        END DO
        DO KKKK=J+1,K
            SUM(3)=SUM(3)+FRAG(KKKK)
        END DO
        DO KKKKK=K+1,L
            SUM(4)=SUM(4)+FRAG(KKKKK)
        END DO
        DO KKKKKK=L+1,M
            SUM(5)=SUM(5)+FRAG(KKKKKK)
        END DO
        DO MMM=M+1,NFRAG
            SUM(6)=SUM(6)+FRAG(MMM)
        END DO
    ENDIF
    NIN=2
    NFI=5
    CALL COMPARE(NIN,NFI,I,J,K,L,M)
    GO TO 100
C
40 IF(L.EQ.0) THEN
    GO TO 50
ELSE
    DO KK=1,I
        SUM(1)=SUM(1)+FRAG(KK)
    END DO
    DO KKK=I+1,J
        SUM(2)=SUM(2)+FRAG(KKK)
    END DO
    DO KKKK=J+1,K
        SUM(3)=SUM(3)+FRAG(KKKK)
    END DO
    DO KKKKK=K+1,L
        SUM(4)=SUM(4)+FRAG(KKKKK)
    END DO
    DO KKKKKK=L+1,NFRAG
        SUM(5)=SUM(5)+FRAG(KKKKKK)
    END DO
ENDIF
NIN=2
NFI=4
CALL COMPARE(NIN,NFI,I,J,K,L,M)
GO TO 100
C
50 IF(K.EQ.0) THEN
    GO TO 60
ELSE
    DO KK=1,I
        SUM(1)=SUM(1)+FRAG(KK)
    END DO
    DO KKK=I+1,J
        SUM(2)=SUM(2)+FRAG(KKK)
    END DO
    DO KKKK=J+1,K
        SUM(3)=SUM(3)+FRAG(KKKK)
    END DO
    DO KKKKK=K+1,NFRAG
        SUM(4)=SUM(4)+FRAG(KKKKK)
    END DO
ENDIF
NIN=2
NFI=3
CALL COMPARE(NIN,NFI,I,J,K,L,M)
GO TO 100
C
60 IF(J.EQ.0) THEN
    GO TO 30

```

## VI. Apèndix

---

```
ELSE
  DO KK=1,I
    SUM(1)=SUM(1)+FRAG(KK)
  END DO
  DO KKK=I+1,J
    SUM(2)=SUM(2)+FRAG(KKK)
  END DO
  DO KKKK=J+1,NFRAG
    SUM(3)=SUM(3)+FRAG(KKKK)
  END DO
ENDIF
NIN=2
NFI=2
CALL COMPARE(NIN,NFI,I,J,K,L,M)
GO TO 100
C
30 DO KK=1,I
  SUM(1)=SUM(1)+FRAG(KK)
END DO
DO KKK=I+1,NFRAG
  SUM(2)=SUM(2)+FRAG(KKK)
END DO
NIN=1
NFI=2
CALL COMPARE(NIN,NFI,I,J,K,L,M)
C
C
100 GO TO 1
999 STOP
END
C
C+++++
SUBROUTINE COMPARE(NIN,NFI,I,J,K,L,M)
C
COMMON SUM(6),NAA(1000),FRAG(1000)
C
CONTROL DE PESOS.
C
100 DO KK=NIN,NFI
  KSUM=SUM(KK)/100
C
  IF(KSUM.LT.245 .OR. KSUM.GT.685) THEN
    GO TO 200
  ELSE
    IF(KSUM.LE.255 .OR.KSUM.GE.675) THEN
      CALL SWRITE(KK)
    ELSE
      IF(KSUM.GE.245 .AND. KSUM.LE.255) CALL SWRITE(I,J,K,L,M,KK)
      IF(KSUM.GE.325 .AND. KSUM.LE.335) CALL SWRITE(I,J,K,L,M,KK)
      IF(KSUM.GE.375 .AND. KSUM.LE.385) CALL SWRITE(I,J,K,L,M,KK)
      IF(KSUM.GE.425 .AND. KSUM.LE.435) CALL SWRITE(I,J,K,L,M,KK)
      IF(KSUM.GE.445 .AND. KSUM.LE.455) CALL SWRITE(I,J,K,L,M,KK)
      IF(KSUM.GE.595 .AND. KSUM.LE.605) CALL SWRITE(I,J,K,L,M,KK)
      IF(KSUM.GE.675 .AND. KSUM.LE.685) CALL SWRITE(I,J,K,L,M,KK)
    ENDIF
  END DO
C
RETURN
END
C
C
Subroutine SWRITE(I,J,K,L,M,KK)
C
COMMON SUM(6),NAA(1000),FRAG(1000),NSEC(1000)
C
IF(J.NE.0) GO TO 40
IF (NSEC(I).EQ.1)
1WRITE(6,20) SUM(KK),KK,NAA(I),NAA(J),NAA(K),NAA(L),NAA(M)
20 FORMAT(F10.2,x,I5,4x,5I5)
GO TO 30
C
40 IF(NSEC(I).EQ.0 .OR. NSEC(J).EQ.0) GO TO 30
WRITE(6,20) SUM(KK),KK,NAA(I),NAA(J),NAA(K),NAA(L),NAA(M)
30 RETURN
END
```

## APÈNDIX C: CAPÍTOLS REALITZATS AMB AMBER<sup>120</sup>

### C.1 CÀLCUL DE LES CÀRREGUES: OPTIMITZACIÓ I CÀLCUL DEL POTENCIAL ELECTROSTÀTIC AMB GAUSSIAN

A continuació es detallen els fitxers per optimitzar l'estructura amb Gaussian (Fitxer C1) i per calcular el potencial electrostàtic un cop realitzada l'optimització (Fitxer C2).

```
%nproc=2
$Mem=200Mb
%chk=/cescascratch/itziar/dansyl-a.chk
#p STO-3G guess=read Opt=(restart) scf=(save,maxcyc=500)
pop=none geom(nodistance,noangle)

Optimitzacio dansyl-a

0 1
Z-MATRIX
```

**Fitxer C1 Model per optimitzar una estructura amb Gaussian.**

```
%nproc=2
$Mem=200Mb
%chk=/cescascratch/itziar/dansyl-a.chk
#N HF/6-31G* pop=mk iop(6/33=2) guess=read geom(checkpoint,nodistance,noangle)

potencial electrostatic dansyl-a

0 1
```

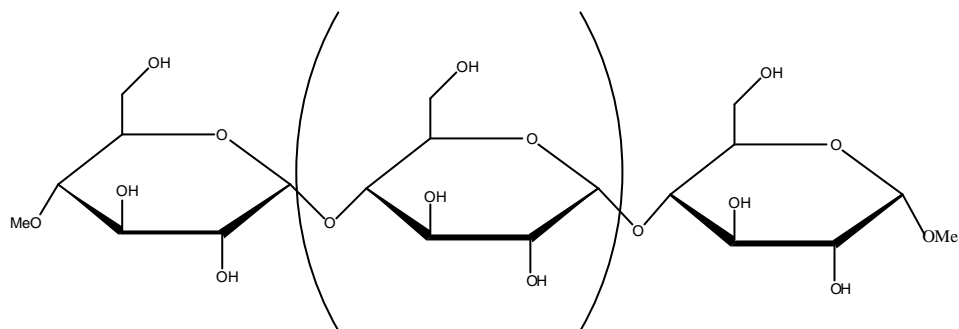
**Fitxer C2 Model per calcular el potencial electrostàtic amb Gaussian.**

#### C.1.1 Càrregues atòmiques de les molècules estudiades amb AMBER

##### C.1.1.1 Càrregues de la glucosa

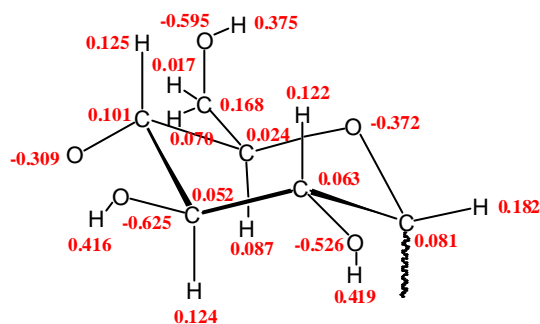
S'han utilitzat les mateixes càrregues de la glucosa calculades prèviament en el nostre grup de recerca (veure tesi doctoral Ivan Beà). Les càrregues van ser determinades utilitzant tres conformacions d'un trímer de glucoses (Figura C1).

<sup>120</sup> <http://amber.scripps.edu>



**Figura C1 Representació del trimer de glucoses.**

Les càrregues finals (Figura C2) són aquelles de la glucosa central, ja que és la que millor reproduceix l'entorn de la ciclodextrina.

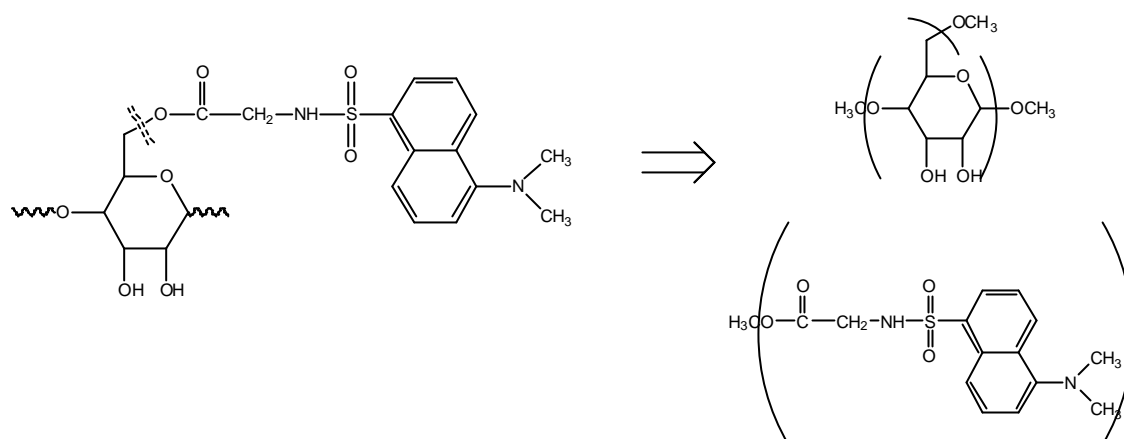


**Figura C2 Càrregues atòmiques de la glucosa.**

#### C.1.1.2 Càrregues atòmiques de la glucosa modificada i del grup -Gly-dansil

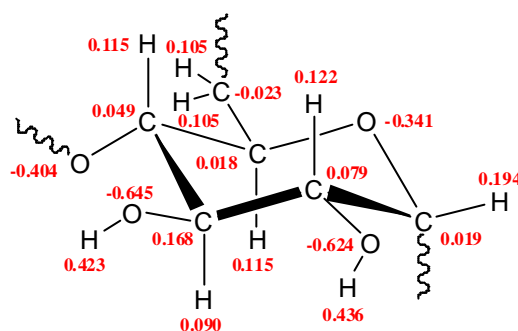
Per calcular les càrregues de les glucoses modificades amb el grup -Gly-dansil s'han escollit dues molècules model: una per representar la glucosa modificada, i l'altre per representar al grup -Gly-dansil (Figura C3). Amb cada una de les molècules s'ha realitzat una cerca conformacional amb càlculs de Monte Carlo (MC) i amb el camp de forces MM3\* implementat en Macromodel. En el cas de la glucosa, s'han girat aleatòriament els enllaços C5-C6 i C6-O6, i en el cas del grup -Gly-dansil, s'han girat aleatòriament els enllaços simples de la cadena compresa entre l'ester i el grup sulfòxid. Les dues conformacions de mínima energia de la glucosa amb energies de 187.0 i 189.5 kJ/mol, representen un 83% de la població, i les dues conformacions de mínima energia del grup -Gly-dansil amb energies de 12.20 i 14.49 kJ/mol, representen un 70% de la població. Per tant, són aquestes les que s'han emprat per a la metodologia RESP.

En ambdós casos, en la metodologia RESP s'ha restringit la càrrega a zero d'aquells àtoms compresos entre parèntesis de la Figura C3 amb l'objectiu de garantir la neutralitat del sistema final.

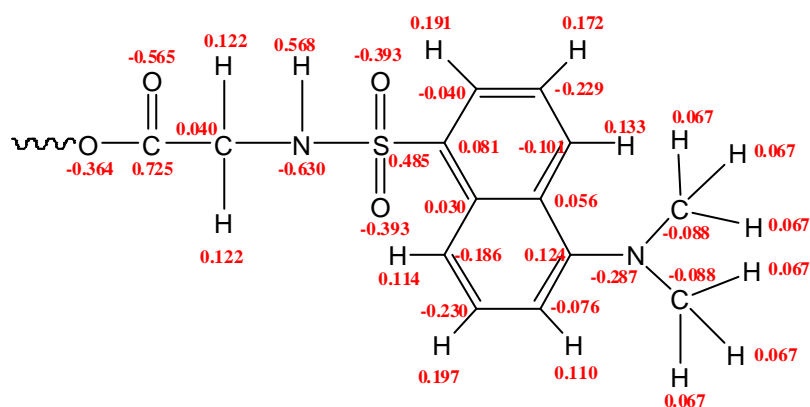


**Figura C3** Mòl·cules model per calcular les càrregues de la glucosa modificada i del grup -Gly-dansil.

Les càrregues finals obtingudes per a la glucosa i per al grup -Gly-dansil es representen a la Figura C4 i a la Figura C5, respectivament.



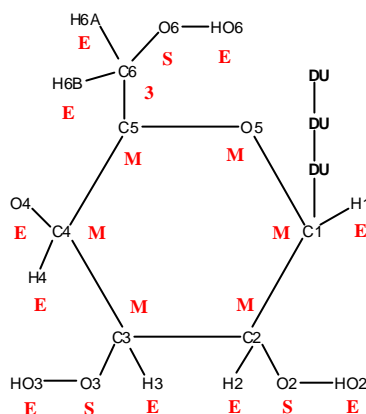
**Figura C4** Càrregues atòmiques de la glucosa modificada.



**Figura C5** Càrregues atòmiques del grup -Gly-dansil.

## C.2 TOPOLOGIA I FITXER *PREP* DEL RESIDU DE GLUCOSA

En AMBER, els residus tenen una estructura d'arbre i els àtoms es classifiquen en cinc tipus diferent. Els àtoms tipus M (*main*) són els àtoms de la cadena principal, i els àtoms tipus E (*end*) són els àtoms finals de les branques. Tots els altres tipus d'àtom surten de la cadena principal i segons el nombre de connexions que tinguin amb altres àtoms poden ser S (*side*) per als que tenen dues connexions amb altres àtoms, B (*branch*) els que tenen tres connexions amb altres àtoms, "3" els que tenen 4 connexions amb altres àtoms i així successivament es podrien definir àtoms tipus "4", tipus "5", i tipus "6". A la Figura C6 es representa la topologia per a un residu de glucosa, es marquen els noms dels àtom i a sota de cada un d'ells la seva topologia. També es representen tres àtoms fantasmes o *dummies* necessaris per tal de definir els eixos espacials del residu.



**Figura C6 Esquema de topologia de la glucosa amb AMBER.**

A continuació, es detalla el fitxer *prep* d'una glucosa (Fitxer C3) que s'ha utilitzat per crear totes les ciclodextrines gegants. La primera línia correspon al títol del fitxer i en la segona línia es detalla la identificació del residu (CD1) i el tipus de coordenades que s'escriuran al fitxer de sortida o *output* (internes, INT). A continuació, en la tercera línia, amb la comanda *CHANGE* s'indica que es realitzarà el canvi de coordenades cartesianes de l'estructura d'entrada a coordenades internes, i les altres tres comandes indiquen que els tres àtom fantasmes del principi s'ometran després de generar tota la informació del sistema. Quan el residu és cíclic, també és necessari definir el punt de tancament mitjançant la comanda LOOP. Per exemple, en aquest cas s'ha definit el LOOP entre els àtoms C3 i C4.



```

glul.res
CD1 INT 0
CHANGE OMIT DU BEG
0.000000
1 DUMM DU M 8.000 1.500 2.000 0.000000
2 DUMM DU M 8.000 1.500 0.500 0.000000
3 DUMM DU M 7.000 1.500 0.500 0.000000
4 C1 CT M 4.627 2.667 4.970 0.080575
5 H1 H2 E 4.602 3.635 5.471 0.181549
6 C2 CT 3 5.668 1.801 5.683 0.063000
7 H2 H1 E 5.446 1.764 6.751 0.122104
8 O2 OH S 6.988 2.307 5.484 -0.525616
9 H2O HO E 7.581 1.607 5.782 0.419530
10 C3 CT B 5.644 0.391 5.109 0.052478
11 H3 H1 E 5.939 0.409 4.059 0.124366
12 O3 OH S 6.567 -0.425 5.831 -0.625007
13 H3O HO E 6.407 -1.349 5.584 0.416007
14 O1 OS M 3.325 2.080 5.062 -0.372447
15 C5 CT M 3.240 0.757 4.506 0.024495
16 H5 H1 E 3.459 0.781 3.429 0.086982
17 C6 CT 3 1.807 0.249 4.695 0.167900
18 H6A H1 E 1.716 -0.728 4.232 0.017186
19 H6B H1 E 1.580 0.167 5.758 0.069678
20 O6 OH S 0.858 1.124 4.085 -0.594882
21 H6O HO E 0.857 0.981 3.126 0.374554
22 C4 CT M 4.234 -0.180 5.221 0.101322
23 H4 H1 E 3.966 -0.245 6.284 0.124923
24 O4 OS M 4.315 -1.452 4.566 -0.308669
    
```

LOOP C3 C4  
 DONE  
 STOP

Diagrama de etiquetatge:  
 - Nom àtom: apunta a la columna 1.  
 - Tipus d'àtom: apunta a la columna 2.  
 - Topologia: apunta a la columna 3.  
 - Coordenades: apunta a les columnes 4, 5 i 6.  
 - Càrrega: apunta a la columna 7.

**Fitxer C3 Exemple de fitxer *prep* per a la glucosa.**

### C.3 EXEMPLE DE FITXER D'ENTRADA DE LEAP

A continuació, es detalla el fitxer exemple per crear els fitxers de topologies i coordenades d'una  $\gamma$ -ciclodextrina amb el programa LEaP (Fitxer C4). En primer lloc, es carreguen tots els fitxers de paràmetres i topologies associats al camp de forces (*source leaprc.ff99*), en aquest cas s'ha escollit com a exemple el camp de forces parm99. A continuació, es llegeixen els fitxers *prep* de cada una dels residus (*loadamberprep nom.prp*) que en el cas de la  $\gamma$ -ciclodextrina són 8. Quan es llegeixen els fitxers *prep* es crea una unitat per a cada residu amb la identificació del fitxer *prep* (CD1-CD8). Amb aquestes noves unitats es defineix la seqüència de la ciclodextrina (*sequence { CD1 CD2 CD3 CD4 CD5 CD6 CD7 CD8}*) i es crea una nova unitat (CA8 ⇒). Donat que la ciclodextrina és una molècula cíclica es crea un nou enllaç entre els dos extrems de la seqüència anterior (*bond CA8.1.C1 CA8.8.O4*) i a continuació, s'assignen les coordenades cartesianes (*loadpdb gcd.pdb*) i es crea una nova unitat (*pdb* ⇒). En aquest punt es solvata el sistema (*solvatebox pdb WATBOX216 10.0*) amb una caixa de molècules d'aigua i finalment, es creen els fitxers de topologies i coordenades (*saveamberparm*).

```

source leaprc.ff99
loadamberprep glc1.prp
loadamberprep glc2.prp
loadamberprep glc3.prp
loadamberprep glc4.prp
loadamberprep glc5.prp
loadamberprep glc6.prp
loadamberprep glc7.prp
loadamberprep glc8.prp
CA8 = sequence { CD1 CD2 CD3 CD4 CD5 CD6 CD7 CD8 }
bond CA8.1.C1 CA8.8.O4
pdb = loadpdb gcd.pdb
solvatebox pdb WATBOX216 10.0
saveamberparm pdb CA8w.top CA8w.crd
saveoff pdb CA8w.lib

```

**Fitxer C4 Exemple de fitxer d'entrada de LEaP per crear topologies i paràmetres per a la  $\alpha$ -ciclodextrina.**

#### C.4 EXEMPLE DE FITXERS PER MINIMITZAR L'ENERGIA

En primer lloc, es detalla el fitxer reduït per dur a terme una minimització de l'energia en fase gas (Fitxer C5). Les explicacions de les comandes estan detallades al costat del fitxer. Convé esmentar que, amb SANDER, la diferenciació d'una minimització i d'una simulació de dinàmica molecular està en el valor de la comanda `imin` (0 per a la minimització i 1 per a la simulació de dinàmica molecular).

```

#Minimització en fase gas
&cntrl
  imin = 1,
  maxcyc = 10000, nycyc = 5000,
  ntx = 1,
  scee = 2.0,   scnb= 1.2,
  ntb = 0,
&end

```

**Les coordenades es llegeixen sense informació de la velocitat.**

**Realitza 10000 passos de minimització i al cap de 5000 passos canvia el mètode de minimització de *Steepest Descent* a *Conjugate Gradient*.**

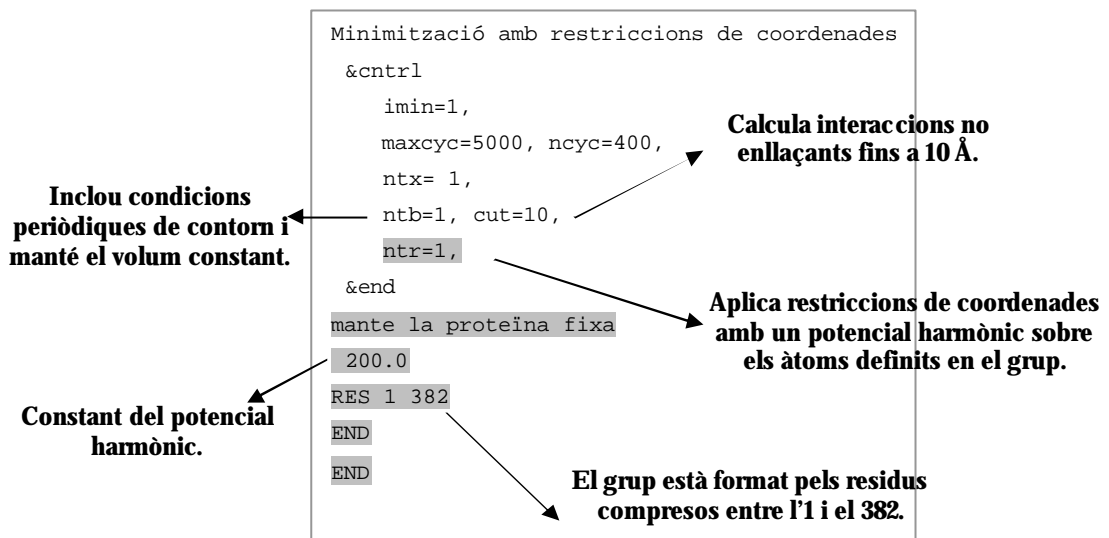
**Les interaccions electrostàtiques i de van der Waals es divideixen per 2.0 i 1.2 respectivament.**

**No s'inclouen condicions periòdiques de contorn.**

**Fitxer C5 Exemple de fitxer d'entrada de SANDER per a una minimització en fase gas.**

A continuació es detalla el fitxer d'entrada per realitzar una minimització de l'energia en solució aquosa (Fitxer C6), on a més, s'inclouen les comandes necessàries per imposar restriccions en les coordenades del solut. Aquest és el fitxer utilitzat per minimitzar l'energia de

l'Hsp70 solvatada: aquesta es manté fixa i només es minimitzen les molècules d'aigua. En aquest cas, només s'expliquen les comandes que difereixen de les respectives opcions en fase gas, i convé destacar que, si es volgués minimitzar l'energia de tot el sistema (també seria aplicable a les minimitzacions realitzades en les ciclodextrines), aquelles comandes marcades amb gris s'haurien d'ometre.



**Fitxer C6 Exemple de fitxer d'entrada de SANDER per a una minimització en solució aquosa amb restriccions de coordenades.**

## C.5 EXEMPLE DE FITXERS PER REALITZAR LES SIMULACIONS DE DINÀMICA MOLECULAR

### C.5.1 Simulació de dinàmica molecular de les ciclodextrines

A continuació es detallen els fitxers que s'han utilitzat per realitzar les simulacions de dinàmica molecular en solució aquosa de les ciclodextrines gegants. En la simulació de dinàmica molecular hi ha una primera etapa d'escalfament del sistema on s'augmenta la temperatura gradualment fins a arribar a la temperatura desitjada. Aquesta primera etapa es duu a terme a volum constant (`ntb=1`). A continuació, es realitza una primera equilibració del sistema a la temperatura desitjada i també a volum constant per tal d'assegurar que la temperatura és manté constant al voltant del valor desitjat.

Les comandes d'aquestes dues etapes es detallen al Fitxer C7. En aquest cas, a diferència dels fitxers descrits per realitzar minimitzacions, es mostren els fitxers complets amb totes les comandes. No obstant, en color verd estan marcades aquelles de major importància per realitzar

una bona simulació i al seu costat s'expliquen els seus significats. Per entendre el significat de les altres comandes es recomana consultar el manual.

```

escalfament i lra equilibració a V ctant
&cntrl

imin=0, nmropt=1,
ntx=1, irect=0, ntrx=1,

ntxo=1, ntp=250, ntw=500,
iwrap=1, ntwx=500, ntwv=0, ntwe=500,
ioutfm=0, ntwprt=0, idecomp=0,

ntf=2, ntb=1, igb=0, nsnb=15,
ipol=0, gbsa=0,
dielc=1.0, cut=8.0, intdiel=1.0,
scnb=2.0, scee=1.2,

ibelly=0, ntr=0,

nsthlim=20000, nscm=500, nrespa=1,

ntt=1, t=0.0, dt=0.001, vlimit=20.0,

ntp=0, pres0=1., comp=44.6, taup=0.2, npscal=1,
ntc=2, jfastw=0,
tol=0.00001,

&end
&wt type='TEMP0', istep1=0, istep2=1000,
value1=0., value2=30., &end
&wt type='TEMP0', istep1=1001, istep2=2000,
value1=30., value2=60., &end
...
...

&wt type='TEMP0', istep1=9001, istep2=10000,
value1=270., value2=300., &end
&wt type='TEMP0', istep1=10001, istep2=20000,
value1=300., value2=300., &end
&wt type='END', &end

```

Per a cada molècula la imatge més propera al centre de la caixa primària és la que s'escriu a l'output.

Les interaccions amb els àtoms d'H s'ometen.

Nombre de passos de la DM.

Time step (1 fs).

SHAKE aplicat als enllaços que contenen H.

Defineix l'etapa d'escalfament. Cada 1000 passos (1 ps) s'augmenta la temperatura 30 K fins arribar als 300 K.

Defineix l'etapa d'equilibració a V constant. Durant 10000 passos (10 ps) la temperatura es manté a 300 K.

Es llegiran canvis progressius en les variables (&wt).

Cada 15 passos s'actualitza la llista de les interaccions no enllaçants.

Si algun component de la velocitat supera |20| es redueix a 20 per evitar inestabilitats.

**Fitxer C7 Exemple de fitxer d'entrada de SANDER per realitzar l'escalfament i la primera equilibració a V constant d'una dinàmica molecular.**

Un cop que la temperatura està equilibrada es fa una segona equilibració a pressió constant, ja que és l'única manera d'obtenir una densitat adequada. A més, el mètode de solvatació emprat en LEaP pot crear bombolles a l'interior de la caixa si la simulació es duu a terme a volum constant i també pot deixar petits forats als vèrtexs de la caixa. Les comandes d'aquesta segona etapa es detallen al Fitxer C8 però només es comenta el significat d'aquelles comandes que són diferents a les descrites fins ara. En aquest cas, degut a les pròpies característiques del sistema les condicions són més "suaus" per evitar problemes associats al SHAKE i que les glucoses adoptin una conformació nau. Així doncs el *time step* és menor i igualment el temps de relaxació de la pressió.

```

2na equilibracio a P ctant
&cntrl
  imin=0, nmropt=1,
  ntx=5,  irest=1,  ntrx=1,
  ntxo=1, ntp=1, ntwr=500,
  iwrap=1, ntwx=500, ntwv=0, ntwe=500,
  ioutfm=0, ntwprt=0, idecomp=0,
  ntf=2,  ntb=2,  igb=0,  nsnb=15,
  ipol=0, gbsa=0,
  dielc=1.0, cut=8.0, intdiel=1.0,
  scnb=2.0, scee=1.2,
  ibelly=0, ntr=0,
  nstlim=60000, nscm=500, nrespa=1,
  ntt=1, t=0.0, dt=0.0005, vlimit=20.0,
  ntp=1,  pres0=1,  comp=44.6,  taup=0.01,  npscal=1,
  ntc=2, jfastw=0,
  tol=0.00001,
&end
&wt type='TEMP0', istep1=0, istep2=60000,
                                value1=300., value2=300., &end
&wt type='END', &end
  
```

**Es leeixen les coordenades amb informació de la velocitat i de la caixa de solvatació.** ←

**Es reinicia la simulació.** →

**Inclou condicions periòdiques de contorn i manté P constant.** ←

**Escalat isotròpic.** ←

**Valor de referència de la pressió (1 bar~1 atm).** ←

**Temps de relaxació de la pressió.** →

**Escalat segons el centre de masses de les submolècules.** →

**Fitxer C8 Exemple de fitxer d'entrada de SANDER per realitzar l'equilibració a P constant d'una dinàmica molecular.**

Per últim es mostra el fitxer per dur a terme l'etapa de producció de la simulació (Fitxer C9). En aquest cas, cal destacar que el *time step* torna a augmentar fins a 1 fs perquè el sistema ja està equilibrat i pel mateix motiu, el temps de relaxació de la pressió augmenta fins a 0.2.

```

produccio a P ctant
&cntrl

  imin=0, nmropt=1,

  ntx=5, irect=1, ntrx=1,

  ntxo=1, ntp=250, ntwr=500,
  iwrap=1, ntwx=500, ntwv=0, ntwe=500,
  ioutfm=0, ntwprt=0, idecomp=0,

  ntf=2, ntb=2, igb=0, nsnb=15,
  ipol=0, gbsa=0,
  dielc=1.0, cut=8.0, intdiel=1.0,
  scnb=2.0, scee=1.2,

  ibelly=0, ntr=0,

  nstlim=1000000, nscm=500, nrespa=1,
  ntt=1, t=0.0, dt=0.001, vlimit=20.0,

  ntp=1, pres0=1., comp=44.6, taup=0.2, npscal=1,

  ntc=2, jfastw=0,
  tol=0.00001,

&end
&wt type='TEMP0', istep1=0, istep2=1000000,
                                value1=300., value2=300., &end

&wt type='END', &end

```

### Fitxer C9 Exemple de fitxer d'entrada de SANDER per realitzar l'etapa de producció d'una dinàmica molecular.

#### C.5.2 Simulació de dinàmica molecular de l'HSP70

En les simulacions de dinàmica molecular de l'Hsp70 i d'aquesta complexada amb una molècula d'ATP s'ha equilibrat el sistema lentament, introduint restriccions en la posició dels àtoms i eliminant aquestes progressivament al llarg de les etapes d'equilibració. A més, les etapes d'equilibració s'han iniciat amb un *time step* d'1 fs. Així doncs, en l'etapa d'escalfament de 5 ps i la primera equilibració de 40 ps només es permet el moviment de les molècules d'aigua, la restricció imposada és:

```

fixa la proteina
  100.0
RES 1 382
END
END

```

Per tant, en aquestes primeres etapes es fixen totes els àtoms de la proteïna amb una constant de força de  $100 \text{ kcal/mol} \cdot \text{Å}^2$ . A continuació, en la següent etapa d'equilibració de 120 ps, es manté l'esquelet de la proteïna fixa i es mouen les cadenes laterals d'aquesta i també la molècula d'ATP en el cas de que aquest estigui present. S'entén com a esquelet de la proteïna la cadena formada pels àtoms tipus CA, N i C de cada aminoàcid. En aquest cas, la constant de la restricció és de  $10 \text{ kcal/mol} \cdot \text{Å}^2$ :

```
mante cadena principal fixa
10.0
FIND
CA * * *
N * * *
C * * *
SEARCH
RES 1 382
END
END
```

A continuació, es realitza una equilibració de 100 ps de tot el sistema però mantenint el *time step* d'1 fs de les etapes anteriors i finalment, es realitza l'equilibració de tot el sistema augmentant el *time step* fins a 2 fs, que és el mateix utilitzat en l'etapa de producció. Aquesta seqüència d'etapes es duu a terme amb l'objectiu d'evitar deformacions en l'estructura de la proteïna.

## C.6 EXEMPLE DE FITXERS PER REALITZAR LES ANÀLISIS DE LES SIMULACIONS DE DINÀMICA MOLECULAR

### C.6.1 Exemples d'anàlisis realitzades amb ptraj

#### C.6.1.1 Extracció de les molècules d'aigua d'una trajectòria

En primer lloc es mostra el fitxer utilitzat per eliminar les molècules d'aigua de les trajectòries de les ciclodextrines gegants i escriure una nova trajectòria, on únicament s'inclouen els àtoms del solut i s'ometen les coordenades de les molècules d'aigua, així com la informació corresponent a la caixa de solvatació. Cal observar que es llegeixen diverses trajectòries inicials però al final només s'escriu una, això és degut a que, per evitar treballar amb fitxers molt grans l'etapa de producció de la dinàmica molecular s'ha dividit en quatre dinàmiques de 1000 ps cada una. Així doncs, un cop extretes les molècules del solvent, es poden continuar les anàlisis a partir

de la nova trajectòria obtinguda, sempre i quan es vulguin realitzar anàlisis que no impliquin les molècules del solvent.

```

trajin ./CA14/CA14_mdw3.trj.Z }
trajin ./CA14/CA14_mdw4.trj.Z } Llegeix les trajectòries.
trajin ./CA14/CA14_mdw5.trj.Z }
trajin ./CA14/CA14_mdw6.trj.Z }
strip :WAT → Elimina les molècules d'aigua.
trajout ./CA14/CA14_mdw.trj.nosolv nobox → Escriu una nova trajectòria
go amb les coordenades del
solut i sense la informació
de la caixa de solvatació.

```

**Fitxer C10 Exemple de fitxer d'entrada de ptraj per eliminar les molècules del solvent.**

## C.6.2 Exemples d'anàlisis realitzades amb carnal

### C.6.2.1 Anàlisi del RMS

Al Fitxer C11 es detallen les comandes per realitzar l'anàlisi del RMS de les ciclodextrines respecte de la darrera estructura de la simulació de dinàmica molecular. Aquesta anàlisi dóna idea de com varia l'estructura al llarg de la simulació. Només s'escull l'esquelet de la molècula (definit pels àtoms C1, C2, C3, C4 i O4) i no s'agafen tots els àtom de la ciclodextrina ja que les desviacions serien molt grans.

```

FILES_IN
  PARM p1 ../CA14.top; → Llegeix topologia de la molècula.
  STREAM s1 ../CA14_mdw.trj.nosolv; → Llegeix trajectòria.
  STATIC ref_set ../CA14_mdw6.rst.Z; → Defineix estructura de referència
FILES_OUT
  TABLE tab1 CA14w.RMS; → Defineix fitxer sortida amb format de taula
(cada mesura ocupa una columna).
DECLARE
  GROUP grp1 ((ATOM NAME C1 C4 O4)&(RES NAME CD?)); → Defineix un primer grup d'àtoms i
calcula el RMS respecte l'estructura
de referència.
  RMS fit1 FIT grp1 s1 ref_set;
  GROUP grp2 ((ATOM NAME C1 C2 C3 C4 O4)&(RES NAME CD?)); → Defineix el segon grup d'àtoms i
calcula el RMS respecte el primer
grup definit.
  RMS fit2 grp2 fit1;
OUTPUT
  TABLE tab1 fit1 fit2; → Escriu resultats.
END

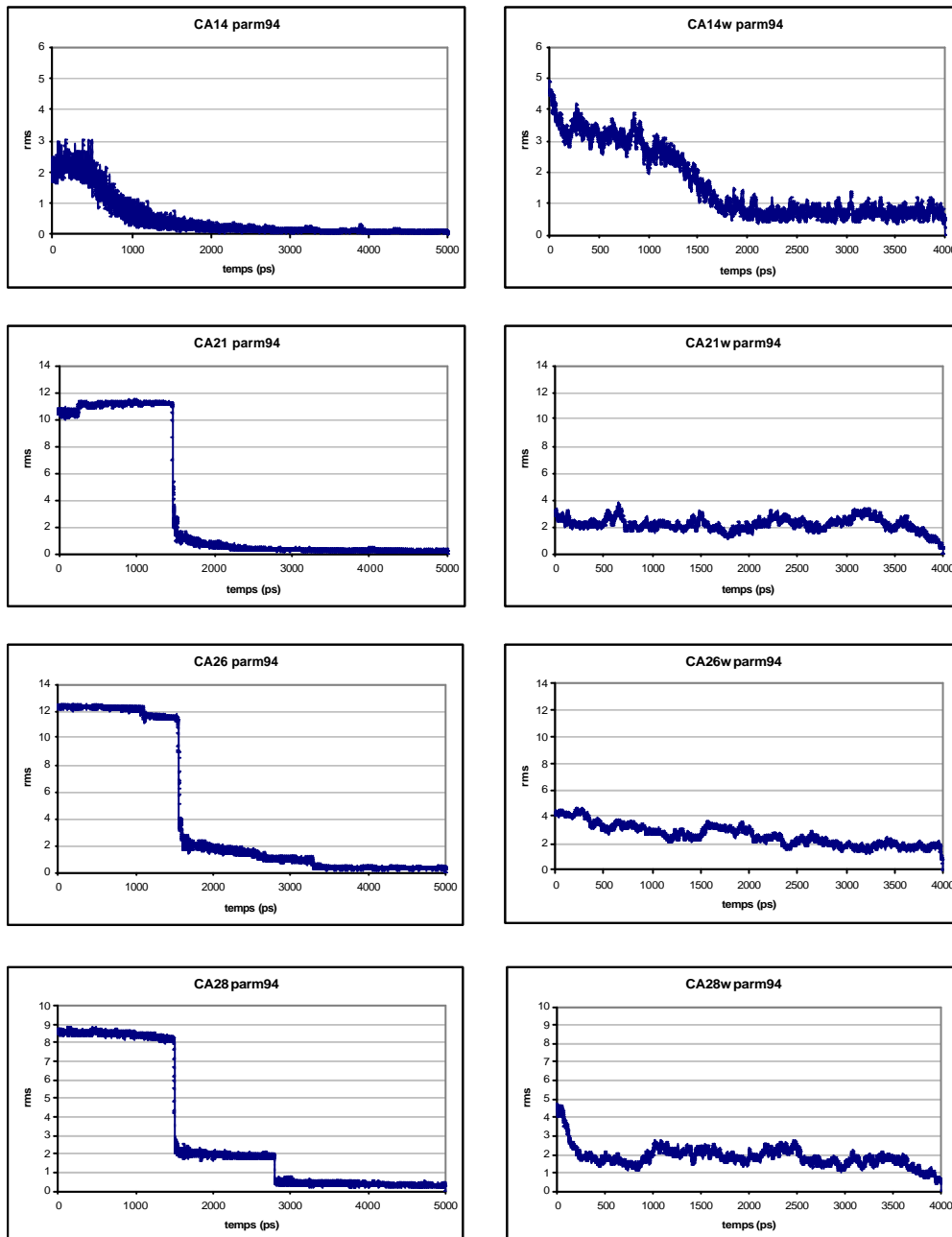
```

**Fitxer C11 Exemple de fitxer d'entrada de carnal per realitzar l'anàlisi del RMS.**

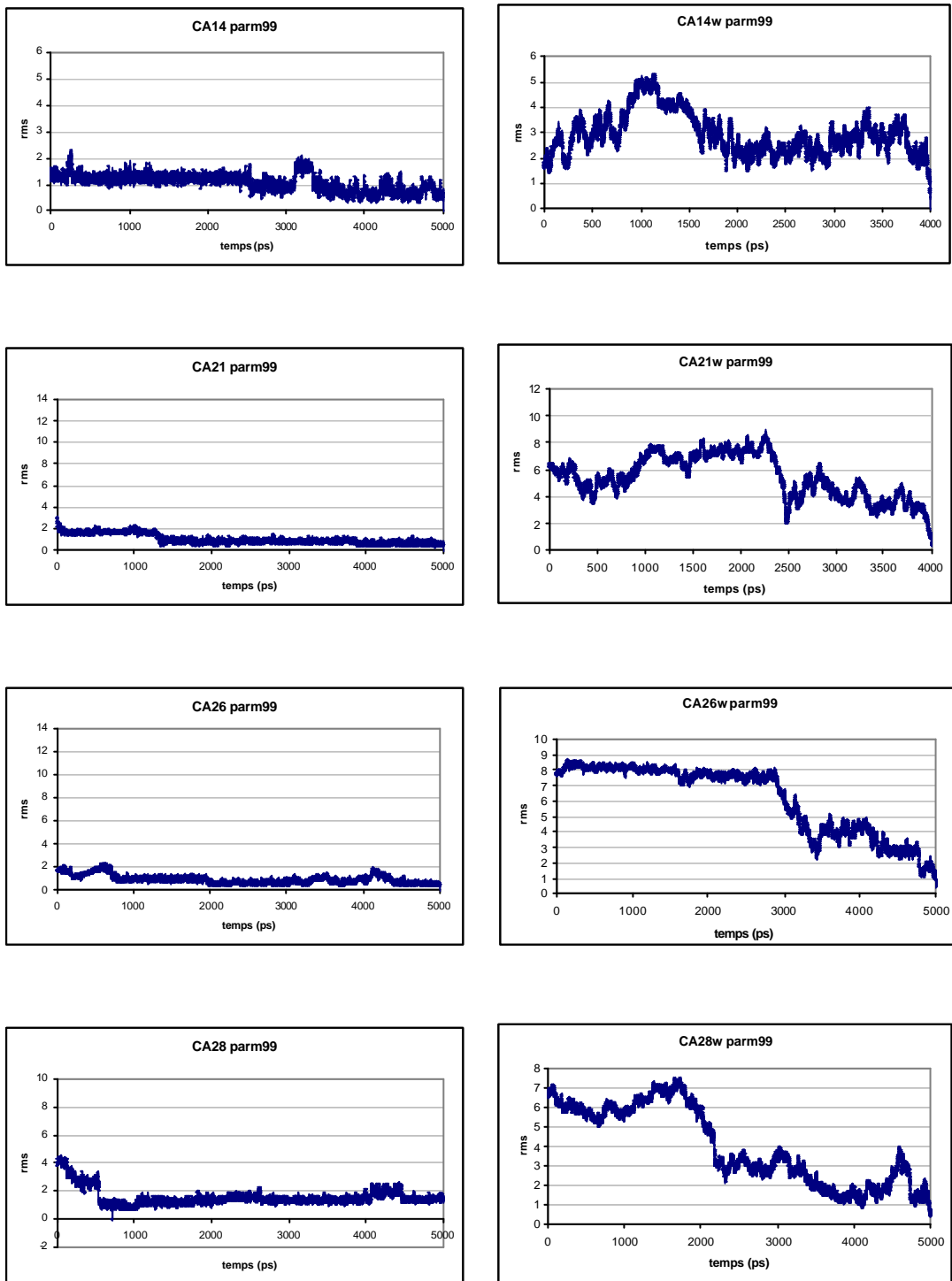


C.6.2.2 Gràfiques del RMS per a les ciclodextrines gegants

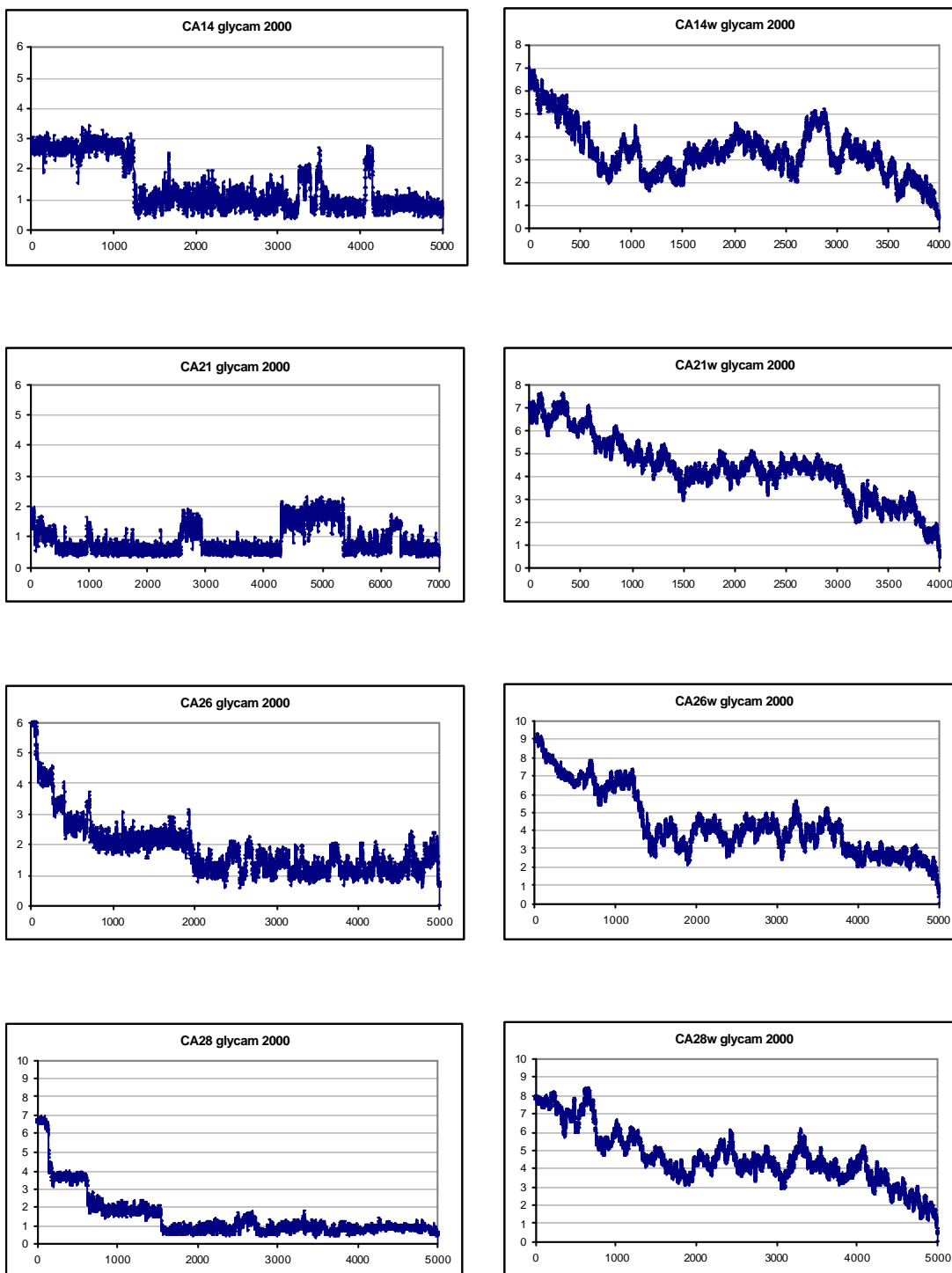
En aquest subapartat s'inclouen les gràfiques que representen la variació del RMS al llarg del temps per a les ciclodextrines gegants, i amb els diferents camps de forces implementats en AMBER. Les gràfiques de la columna de l'esquerra corresponen a les simulacions en fase gas, i les gràfiques de la columna de la dreta corresponen a les simulacions en solució aquosa.



**Gràfica C1 Variació del RMS de les ciclodextrines gegants en les simulacions realitzades amb el camp de forces parm94.**



**Gràfica C2 Variació del RMS de les ciclodextrines gegants en les simulacions realitzades amb el camp de forces parm99.**



**Gràfica C3 Variació del RMS de les ciclodextrines gegants en les simulacions realitzades amb el camp de forces glycam 2000a.**

## C.6.2.3 Anàlisi de les distàncies entre àtoms

A continuació, es mostra un fitxer exemple d'anàlisi de les distàncies (Fitxer C12). Les anàlisis dels angles i dels diedres són anàlogues a aquesta.

```

FILES_IN
  PARM p1 ../CA14.top;
  STREAM s1 ../CA14_mdw.trj.nosolv;
FILES_OUT
  TABLE tab1 CA14w_o1.tab;
  ...
  TABLE tab15 CA14w_sec1.tab;
  ....
DECLARE
  DIST o1 O4 1 O4 2;
  ...
  DIST sec1 O3 1 O2 2;
  ....
OUTPUT
  TABLE tab1 o1;
  ...
  TABLE tab15 sec1;
  ....

```

**Llegeix topologia de la molècula.**  
**Llegeix trajectòria.**  
**Defineix fitxers sortida amb format de taula.**  
**Calcula distància entre oxígens glicosídics O4(n)...O4(n+1).**  
**Calcula distància entre oxígens hidroxílics O3(n)...O2(n+1).**  
**Escriu els resultats de cada anàlisi en una taula.**

**Fitxer C12 Exemple de fitxer d'entrada de carnal per realitzar l'anàlisi de les distàncies.**

## C.6.2.4 Anàlisi dels enllaços d'hidrogen

Finalment, es mostra el fitxer per calcular els enllaços d'hidrogen intramoleculars per a ciclodextrines (Fitxer C13). OH és el tipus d'àtom corresponent als oxígens hidroxílics (OH2, OH3 OH6), i OS és el tipus d'àtom corresponent als oxígens tipus èter (O4 i O5).

```

FILES_IN
  PARM p1 ../CA14.top;
  STREAM s1 ../CA14_mdw.trj.nosolv;
FILES_OUT
  HBOND h1 CA14w_hb LIST;
DECLARE
  GROUP g1 (ATOM TYPE OH);
  GROUP g2 (ATOM TYPE OH OS);
OUTPUT
  HBOND h1 DONOR g1 ACCEPTOR g2 STATS;

```

**Es defineix el fitxer de sortida amb format de llista on s'especifica l'àtom donador, l'acceptor, la distància i l'angle.**  
**Definició del grup donador i del grup acceptor de protons.**  
**Escriu resultat al fitxer de sortida i a més, inclou resum al output estàndard.**

**Fitxer C13 Exemple de fitxer d'entrada de carnal per realitzar l'anàlisi dels enllaços d'hidrogen.**

És important remarcar que, en aquest cas, s'han agafat les dades obtingudes del fitxer de sortida estàndard, ja que en aquest és on s'inclou el % d'ocupació d'un enllaç d'hidrogen donat.

D'altra banda, en el cas de les anàlisis realitzades amb el camp de forces glycam s'ha afegit un tipus d'àtom més en el segon grup (els àtoms acceptors) ja que el dos oxígens tipus èter tenen tipus d'àtom diferent en aquest camp de forces: O5 té un tipus d'àtom OS i O4 té un tipus d'àtom OG. La definició d'aquest grup és:

```
GROUP g2 (ATOM TYPE OH OS OG);
```

Si també volguéssim estudiar els enllaços d'hidrogen entre el solut i el solvent aleshores s'hauria de definir diferents grups. Per exemple amb la comanda:

```
GROUP g1 (ATOM TYPE OH);
GROUP g2 (ATOM TYPE OH OS OW);
```

s'analitzarien tots els enllaços d'hidrogen en què el donador és el solut i l'acceptor pot ser el mateix solut i també el solvent. O amb la comanda:

```
GROUP g1 (ATOM TYPE OH OW);
GROUP g2 (ATOM TYPE OH OS);
```

s'analitzarien els enllaços d'hidrogen on el solut i el solvent serien els donadors de protons i només el solut seria l'acceptor.

#### 15.1.1.1 Càlcul de l'estructura obtinguda a partir de la mitjana de les coordenades

Per obtenir l'estructura a partir de la mitjana de les coordenades, primer es realitza un ajust del RMS i sobre el resultat obtingut es calcula la mitjana. Així doncs el fitxer d'entrada és idèntic al Fitxer C11 descrit per realitzar l'anàlisi del RMS però en el fitxer de sortida s'especifica l'opció:

```
OUTPUT
COORD c1 fit1 AVERAGE;
```

## APÈNDIX D: CAPÍTOLS REALITZATS AMB CHARMM<sup>121</sup>

### D.1 PARAMETRITZACIÓ

En el camp de forces CSFF, els diedres entre els enllaços pertanyents a la connexió entre el sucre i l'anell aromàtic estan representats de forma genèrica (diedres tipus X-CT-OE-X). Donat que la rotació d'aquests enllaços permet passar d'un dels dímers tipus cap-cap a l'altre, s'ha trobat necessari el fet de disposar d'uns bons paràmetres que ens puguin representar correctament aquestes rotacions. Així doncs, s'ha escollit com a molècula model l'isopropilfenilmetilèter i s'han parametritzat els diedres C6R-CT-OE-CT i HA-CT-OE-CT (Figura D1).

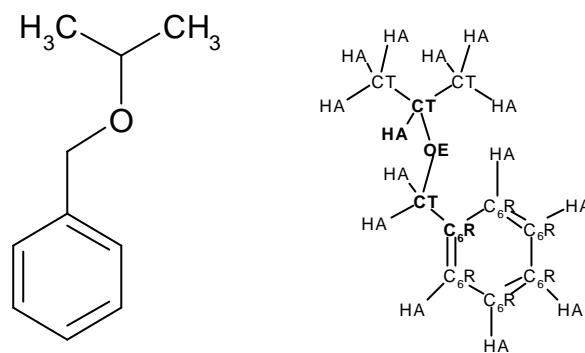
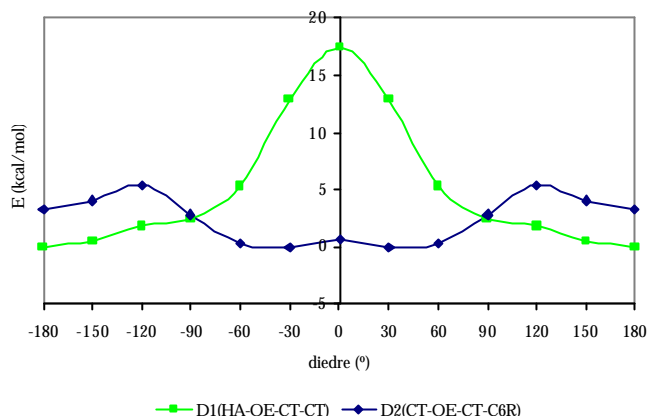


Figura D1

En primer lloc, s'han optimitzat diferents estructures amb valors dels angles diedres a parametritzar compresos entre 0 i 180° en intervals de 30°. S'han realitzat càlculs tipus *ab-initio* amb una base 6-31g(d) i el programa Gaussian 98.<sup>122</sup> En totes les estructures un dels diedres C6R-C6R-CT-OE s'ha fixat a 90° per tal d'evitar que en la rotació de l'enllaç C6R-CT-OE-CT els protons dels metils col·lapsin amb els protons aromàtics. A partir de les optimitzacions d'aquestes estructures i considerant la simetria de la molècula, s'han obtingut les corbes energètiques corresponents a aquestes rotacions representades en la Gràfica D1.

<sup>121</sup> <http://www.charmm.org/document/Charmm/>

<sup>122</sup> *Gaussian 98*, Firsch, M.J.; Trucos, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Zakrzewski, V.G.; Montgomery, J.A.; Stratmann, R.E.; Burant, J.C.; Dapprich, S.; Millam, J.M.; Daniels, A.D.; Kudin, K.N.; Strain, M.C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G.A.; Ayala, P.Y.; Cui, Q.; Morokuma, K.; Malick, D.K.; Rabuck, A.D.; Raghavachari, K.; Foresman, J.B.; Cioslowski, J.; Ortiz, J.V.; Stefanov, B.B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R.L.; Fox, D.J.; Keith, T.; Al-Laham, M.A.; Peng, C.Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P.M.W.; Johnson, B.G.; Chen, W.; Wong, M.W.; Andres, J.L.; Head-Gordon, M.; Replogle, E.S.; Pople, J.A., Gaussian, Inc, Pittsburg PA, **1998**.



**Gràfica D1 Corbes energètiques de rotació dels diedres HA-CT-OE-CT i C6R-CT-OE-CT amb Gaussian.**

El segon pas de la parametrització, ha estat ajustar a aquestes corbes energètiques els valors corresponents a les rotacions d'aquests mateixos enllaços obtinguts a partir dels càlculs de mecànica molecular amb el programa CHARMM. Per això, s'ha de tenir en compte que en el camp de forces l'energia de torsió es calcula a partir de l'Equació D1 on  $k_1$ ,  $k_2$  i  $k_3$  són els paràmetres de torsió i  $x$  el valor de l'angle diedre.

$$E = k_1 \times (1 + \cos x) + k_2 \times (1 + \cos 2x) + k_3 \times (1 + \cos 3x)$$

#### Equació D1

Si es considera la molècula model per realitzar la parametrització s'ha de tenir en compte que per al diedre C6R-CT-OE-CT hi ha tres contribucions diferents en l'energia de torsió; una contribució del diedre tipus C6R-CT-OE-CT, i dues del diedre tipus HA-CT-OE-CT. De la mateixa manera, per a la parametrització del diedre HA-CT-OE-CT també hi ha tres contribucions; una del diedre tipus HA-CT-OE-CT, i dues del diedre tipus CT-CT-OE-CT.

Per tant, dins de cada un dels termes de l'equació general del camp de forces, s'ha de considerar en cada cas el tipus de contribucions que hi ha en l'energia total de torsió, i també s'ha de considerar la degeneració dels angles. Per exemple, en el cas de la parametrització del diedre C6R-CT-OE-CT els dos àtoms d'hidrogen de les dues contribucions tipus HA-CT-OE-CT estan amb un desfase de  $\pm 120^\circ$  respectivament amb l'àtom de C del diedre que volem parametritzar. I el mateix raonament s'ha de fer en el cas de l'altre diedre (Figura D2).

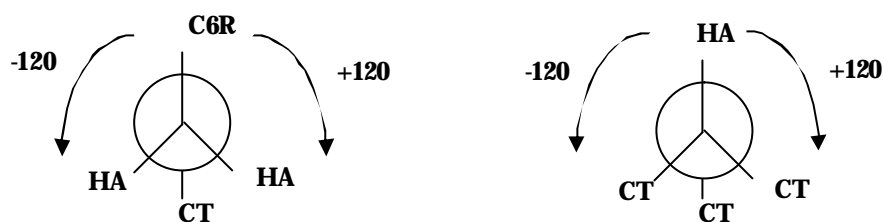


Figura D2

Així doncs, l'expressió general de l'energia que s'ha d'ajustar per al diedre C6R-CT-OE-CT és:

$$E = a1 \times \left[ \left( 1 + \cos \left( x - \frac{2p}{3} \right) \right) + \left( 1 + \cos \left( x + \frac{2p}{3} \right) \right) \right] + a2 \times \left[ \left( 1 + \cos \left( 2x - \frac{2p}{3} \right) \right) + \left( 1 + \cos \left( 2x + \frac{2p}{3} \right) \right) \right] + a3 \times \left[ \left( 1 + \cos \left( 3x - \frac{2p}{3} \right) \right) + \left( 1 + \cos \left( 3x + \frac{2p}{3} \right) \right) \right] + a4 \times \left[ \left( 1 + \cos \left( 4x - \frac{2p}{3} \right) \right) + \left( 1 + \cos \left( 4x + \frac{2p}{3} \right) \right) \right] + c1 \times (1 + \cos x) + c2 \times (1 + \cos 2x) + c3 \times (1 + \cos 3x) + c4 \times (1 + \cos 4x)$$

Equació D2

on a1, a2, a3 i a4 són els paràmetres per al diedre HA-CT-OE-CT, c1, c2, c3 i c4 són els paràmetres per al diedre C6R-CT-OE-CT i x és el valor de l'angle diedre.

I l'expressió que s'ha d'ajustar per al diedre HA-CT-OE-CT és:

$$E = a1 \times (1 + \cos x) + a2 \times (1 + \cos 2x) + a3 \times (1 + \cos 3x) + a4 \times (1 + \cos 4x) + b1 \times \left[ \left( 1 + \cos \left( x - \frac{2p}{3} \right) \right) + \left( 1 + \cos \left( x + \frac{2p}{3} \right) \right) \right] + b2 \times \left[ \left( 1 + \cos \left( 2x - \frac{2p}{3} \right) \right) + \left( 1 + \cos \left( 2x + \frac{2p}{3} \right) \right) \right] + b3 \times \left[ \left( 1 + \cos \left( 3x - \frac{2p}{3} \right) \right) + \left( 1 + \cos \left( 3x + \frac{2p}{3} \right) \right) \right] + b4 \times \left[ \left( 1 + \cos \left( 4x - \frac{2p}{3} \right) \right) + \left( 1 + \cos \left( 4x + \frac{2p}{3} \right) \right) \right]$$

Equació D3

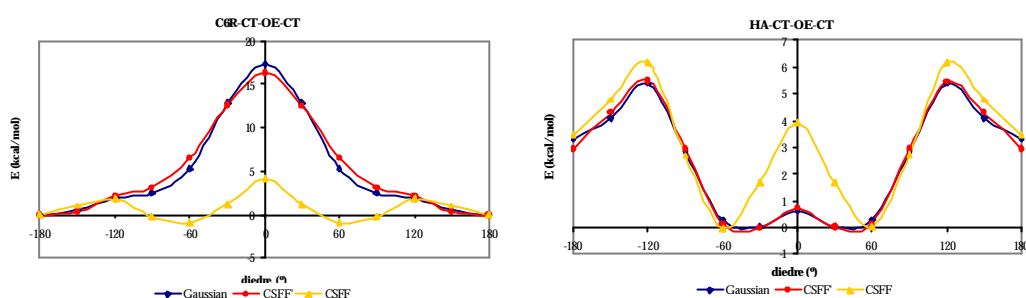
on a1, a2, a3 i a4 també són els paràmetre per al diedre HA-CT-OE-CT i b1, b2, b3 i b4 són els paràmetres per al diedre CT-CT-OE-CT i x el valor de l'angle diedre.

A la Taula D1 es resumeixen dels paràmetres obtinguts al final de la parametrització i a la Gràfica D2 es representen les corbes finals ajustades (CSFF') en comparació amb les corbes energètiques calculades amb el Gaussian (Gaussian) i les corbes amb els paràmetres inicials (CSFF).



**Taula D1 Valors del nou conjunt de paràmetres obtingut.**

|          | a     | b    | c    |
|----------|-------|------|------|
| <b>1</b> | -2.5  | 0.20 | 3.8  |
| <b>2</b> | -0.20 | 0.20 | 2.0  |
| <b>3</b> | 1.0   | 0.10 | -0.8 |
| <b>4</b> | 0.60  | 0.60 | 0.8  |



**Gràfica D2 Comparació entre les corbes energètiques de les rotacions del diedre tipus C6R-CT-OE-CT (esquerra) i del diedres tipus HA-CT-OE-CT (dreta) amb el Gaussian, paràmetres originals del camp de forces CSFF i paràmetres nous per al camp de forces CSFF.**

Es pot comprovar a les gràfiques que hi ha una bon ajust de les corbes amb els nous paràmetres per al camp de forces CSFF (color vermell) a les corbes de Gaussian (color blau fosc) i a més, i molt més important, que s'ha millorat molt respecte les corbes obtingudes amb els paràmetres originals del camp de forces CSFF (color groc).

Per acabar de validar els paràmetres obtinguts, s'ha calculat la corba energètica de rotació del diedre CT-OE-CT-CT amb el nou conjunt de paràmetres i s'ha observat que també hi ha un bon ajust per a aquest diedre. Per tant, es pot concloure que aquest conjunt de paràmetres ens representen correctament el nostre sistema.

## D.2 FITXER DE TOPOLOGIA DE RESIDUS (RTF)

### D.2.1 Topologia de la glucosa

Els fitxers de topologies dels residus (RTF) tenen un format molt diferent en CHARMM i en AMBER. A continuació es mostra el fitxer de topologia per a la unitat de glucosa (Fitxer D1). La comanda RESI de la primera línia indica que aquest residu no s'utilitzarà com a un residu d'ajust (*patch residue*) i les comandes de la darrera línia fan referència als residus d'ajust i, per tant, en aquest cas no són operatives.

|  |      | Identificació del residu |         |        |                                   | Càrrega            |        |         |        |        |     |
|--|------|--------------------------|---------|--------|-----------------------------------|--------------------|--------|---------|--------|--------|-----|
| <b>Marca la definició de grups.</b>                  | RESI | AGLC                     | 0.00000 | !      | 4C1 alpha-D-glucopyranose monòmer |                    |        |         |        |        |     |
|  | GROU |                          |         | !      |                                   |                    |        |         |        |        |     |
|  | ATOM | C1                       | CTS     | 0.200  | !                                 | H61                |        |         |        |        |     |
|  | ATOM | H1                       | HAS     | 0.090  | !                                 |                    |        |         |        |        |     |
|  | ATOM | O1                       | OHS     | -0.660 | !                                 | H62-C6-O6-HO6      |        |         |        |        |     |
|  | ATOM | HO1                      | HOS     | 0.430  | !                                 |                    |        |         |        |        |     |
|  | ATOM | C5                       | CTS     | 0.250  | !                                 | H4 C5--O5 O1-HO1   |        |         |        |        |     |
|  | ATOM | H5                       | HAS     | 0.090  | !                                 | / \ / \ /          |        |         |        |        |     |
|  | ATOM | O5                       | OES     | -0.400 | !                                 | HO4-O4-C4 H5 C1-H1 |        |         |        |        |     |
|  | GROU |                          |         | !      |                                   |                    |        |         |        |        |     |
|  | ATOM | C2                       | CTS     | 0.140  | !                                 | H3-C3--C2-O2-HO2   |        |         |        |        |     |
|  | ATOM | H2                       | HAS     | 0.090  | !                                 |                    |        |         |        |        |     |
|  | ATOM | O2                       | OHS     | -0.660 | !                                 | O3 H2              |        |         |        |        |     |
|  | ATOM | HO2                      | HOS     | 0.430  | !                                 |                    |        |         |        |        |     |
|  | GROU |                          |         | !      | HO3                               |                    |        |         |        |        |     |
|  | ATOM | C3                       | CTS     | 0.140  | !                                 |                    |        |         |        |        |     |
|  | ATOM | H3                       | HAS     | 0.090  | !                                 |                    |        |         |        |        |     |
|  | ATOM | O3                       | OHS     | -0.660 | !                                 |                    |        |         |        |        |     |
|  | ATOM | HO3                      | HOS     | 0.430  | !                                 |                    |        |         |        |        |     |
|  | GROU |                          |         | !      |                                   |                    |        |         |        |        |     |
|  | ATOM | C4                       | CTS     | 0.140  | !                                 |                    |        |         |        |        |     |
|  | ATOM | H4                       | HAS     | 0.090  | !                                 |                    |        |         |        |        |     |
|  | ATOM | O4                       | OHS     | -0.660 | !                                 |                    |        |         |        |        |     |
|  | ATOM | HO4                      | HOS     | 0.430  | !                                 |                    |        |         |        |        |     |
|  | GROU |                          |         | !      |                                   |                    |        |         |        |        |     |
|  | ATOM | C6                       | CPS     | 0.050  | !                                 |                    |        |         |        |        |     |
|  | ATOM | H61                      | HAS     | 0.090  | !                                 |                    |        |         |        |        |     |
|  | ATOM | H62                      | HAS     | 0.090  | !                                 |                    |        |         |        |        |     |
|  | ATOM | O6                       | OHS     | -0.660 | !                                 |                    |        |         |        |        |     |
|  | ATOM | HO6                      | HOS     | 0.430  | !                                 |                    |        |         |        |        |     |
| <b>Defineix tots els enllaços entre àtoms.</b>       | BOND | C1                       | O1      | C1     | H1                                | O1                 | HO1    | C1      | O5     | C1     | C2  |
|  | BOND | C2                       | H2      | C2     | O2                                | O2                 | HO2    | C2      | C3     | C3     | H3  |
|  | BOND | C3                       | O3      | O3     | HO3                               | C3                 | C4     | C4      | H4     | C4     | O4  |
|  | BOND | O4                       | HO4     | C4     | C5                                | C5                 | H5     | C5      | C6     | C6     | H61 |
|  | BOND | C6                       | H62     | C6     | O6                                | O6                 | HO6    | C5      | O5     |        |     |
| <b>Defineix àtoms donadors d'enllaç d'hidrogen.</b>  | DONO | HO1                      | O1      |        |                                   |                    |        |         |        |        |     |
|  | DONO | HO2                      | O2      |        |                                   |                    |        |         |        |        |     |
|  | DONO | HO3                      | O3      |        |                                   |                    |        |         |        |        |     |
|  | DONO | HO4                      | O4      |        |                                   |                    |        |         |        |        |     |
|  | DONO | HO6                      | O6      |        |                                   |                    |        |         |        |        |     |
| <b>Defineix àtoms acceptors d'enllaç d'hidrogen.</b> | ACCE | O1                       |         |        |                                   |                    |        |         |        |        |     |
|  | ACCE | O2                       |         |        |                                   |                    |        |         |        |        |     |
|  | ACCE | O3                       |         |        |                                   |                    |        |         |        |        |     |
|  | ACCE | O4                       |         |        |                                   |                    |        |         |        |        |     |
|  | ACCE | O5                       |         |        |                                   |                    |        |         |        |        |     |
|  | ACCE | O6                       |         |        |                                   |                    |        |         |        |        |     |
| <b>Taula de coordenades internes.</b>                | !    | I                        | J       | K      | L                                 | R(IK)              | T(IKJ) | PHI     | T(JKL) | R(KL)  |     |
|  | IC   | O1                       | C2      | *C1    | H1                                | 1.3889             | 109.35 | -122.69 | 108.98 | 1.0950 |     |
|  | IC   | O1                       | O5      | *C1    | C2                                | 1.3889             | 111.55 | -121.57 | 110.06 | 1.5340 |     |
|  | IC   | O2                       | C3      | *C2    | H2                                | 1.4154             | 112.27 | -118.21 | 108.23 | 1.0919 |     |
|  | IC   | O2                       | C1      | *C2    | C3                                | 1.4154             | 110.87 | -125.56 | 111.08 | 1.5253 |     |
|  | IC   | O3                       | C4      | *C3    | H3                                | 1.4157             | 110.61 | 120.65  | 108.81 | 1.1068 |     |
|  | IC   | O3                       | C2      | *C3    | C4                                | 1.4157             | 108.09 | 120.77  | 109.86 | 1.5177 |     |
|  | IC   | O4                       | C5      | *C4    | H4                                | 1.4252             | 110.90 | -120.61 | 108.35 | 1.1024 |     |
|  | IC   | O4                       | C3      | *C4    | C5                                | 1.4252             | 108.31 | -122.08 | 111.17 | 1.5287 |     |
|  | IC   | O6                       | O5      | *C5    | H5                                | 1.5099             | 108.10 | 118.69  | 109.65 | 1.1042 |     |
|  | IC   | O6                       | C4      | *C5    | O5                                | 1.5099             | 111.57 | 119.10  | 108.69 | 1.4274 |     |
|  | IC   | O6                       | H62     | *C6    | H61                               | 1.4132             | 110.47 | -120.32 | 107.85 | 1.0945 |     |
|  | IC   | O6                       | C5      | *C6    | H62                               | 1.4132             | 110.45 | -121.53 | 108.99 | 1.0959 |     |
|  | IC   | O5                       | C1      | C2     | C3                                | 1.4254             | 110.06 | 54.09   | 111.08 | 1.5253 |     |
|  | IC   | C1                       | C2      | C3     | C4                                | 1.5340             | 111.08 | -51.23  | 109.86 | 1.5177 |     |
|  | IC   | C2                       | C3      | C4     | C5                                | 1.5253             | 109.86 | 53.25   | 111.17 | 1.5288 |     |
|  | IC   | C3                       | C4      | C5     | O5                                | 1.5177             | 111.17 | -57.46  | 108.69 | 1.4274 |     |
|  | IC   | C4                       | C5      | O5     | C1                                | 1.5288             | 108.69 | 62.25   | 113.77 | 1.4254 |     |
|  | IC   | C5                       | O5      | C1     | C2                                | 1.4274             | 113.77 | -60.97  | 110.06 | 1.5340 |     |
|  | IC   | C4                       | C5      | C6     | O6                                | 1.5287             | 111.57 | -170.28 | 110.45 | 1.4132 |     |
|  | IC   | O5                       | C1      | O1     | HO1                               | 1.4254             | 111.55 | 74.87   | 107.83 | 0.9684 |     |
|  | IC   | C1                       | C2      | O2     | HO2                               | 1.5340             | 110.87 | -100.51 | 112.13 | 0.9638 |     |
|  | IC   | C2                       | C3      | O3     | HO3                               | 1.5253             | 108.09 | -165.88 | 112.08 | 0.9730 |     |
|  | IC   | C3                       | C4      | O4     | HO4                               | 1.5177             | 108.31 | 134.18  | 106.97 | 0.9713 |     |
|  | IC   | C5                       | C6      | O6     | HO6                               | 1.5099             | 110.44 | -143.88 | 107.72 | 0.9641 |     |
|  | PATC | FIRS                     | NONE    | LAST   | NONE                              |                    |        |         |        |        |     |

Fitxer D1 Exemple de fitxer de topologia de la glucosa amb CHARMM

### D.2.2 Topologia de residus d'ajust (*patches*)

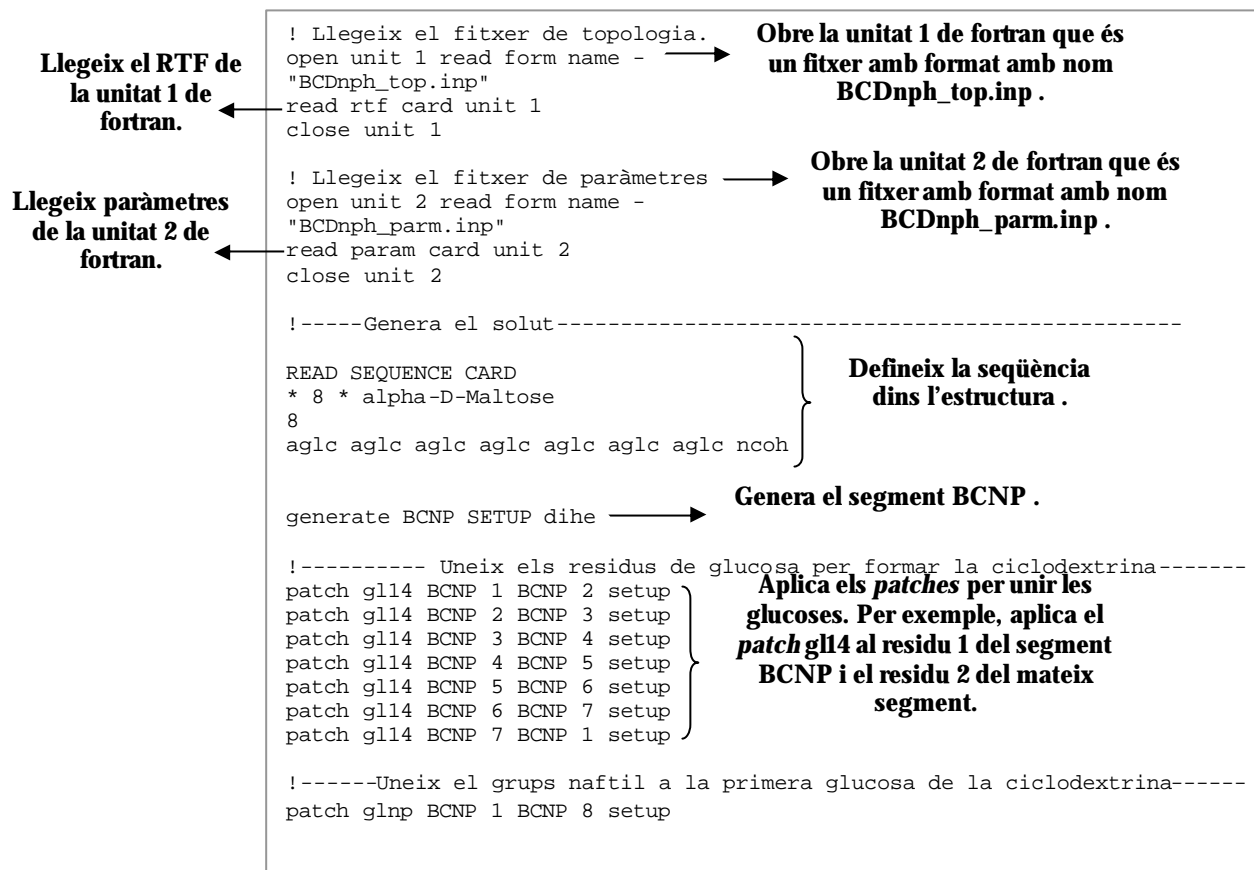
A continuació, es mostra el residu d'ajust per unir les diferents glucoses entre si (Fitxer ). La comanada PRES de la primera línia indica que aquest és un residu d'ajust. En aquest cas s'eliminen tres àtoms de la glucosa (HO1, O4 i HO4) i es crea un nou enllaç entre els àtoms O1 i C4. Un cop realitzades aquestes operacions es defineixen els nous angles, diedres i coordenades internes.

|   |   |                                   |
|---|---|-----------------------------------|
| <p><b>Elimina àtoms</b><br/>(opció vàlida només en els <i>patches</i>).</p> | DELETE ATOM 1HO1<br>DELETE ATOM 2O4<br>DELETE ATOM 2HO4   | <p><b>Crea un enllaç nou.</b></p> |
|   | BOND 1O1 2C4  |                                   |
|   | ANGL 1C1 1O1 2C4 1O1 2C4 2C3<br>ANGL 1O1 2C4 2C5 1O1 2C4 2H4  |                                   |
|   | DIHE 1O5 1C1 1O1 2C4 1C1 1O1 2C4 2c5<br>DIHE 1C2 1C1 1O1 2C4 1C1 1O1 2C4 2c3<br>DIHE 1h1 1C1 1O1 2C4 1C1 1O1 2C4 2h4<br>DIHE 1O1 2C4 2C5 2C6 1O1 2C4 2C5 2H5<br>DIHE 1O1 2C4 2C5 2O5 1O1 2C4 2C3 2C2<br>DIHE 1O1 2C4 2C3 2H3 1O1 2C4 2C3 2O3  |                                   |
|   | IC 1H1 1C1 1O1 2C4 1.0882 110.64 4.85 117.87 1.4195<br>IC 1C1 1O1 2C4 2H4 1.4138 117.87 13.26 110.59 1.0977<br>IC 2C3 2C5 *2C4 1O1 1.5265 109.15 -118.85 110.91 1.4195<br>IC 2C5 1O1 *2C4 2H4 1.5340 110.91 121.00 110.59 1.0977<br>IC 1O1 2C4 2C5 2O5 1.4195 110.91 -178.47 109.01 1.4229<br>IC 1O1 2C4 2C3 2C2 1.4195 107.98 174.59 110.07 1.5280 |                                   |

#### Fitxer D2 Exemple de fitxer d'ajust per unir dues glucoses amb CHARMM

### D.3 CREACIÓ D'UN NOU SEGMENT

Al Fitxer D3 es mostra la seqüència de comandes necessàries per crear un nou segment, que en aquest cas, és el sistema a estudiar. En l'exemple es genera un segment per a la ciclodextrina modificada amb un grup metilnaftil anomenat BCNP.



Fitxer D3 Exemple per crear un nou segment amb CHARMM

#### D.4 FITXER PER SOLVATAR EL SOLUT

La seqüència de comandes necessàries per solvatar el solut es mostren al Fitxer D4. En aquest cas, el primer pas és centrar el solut en l'origen de coordenades, i després s'afegeix una primera esfera d'aigua i s'eliminen tots aquells àtoms del solvent que o bé estan massa a prop del solut (a un radi menor de 2.5 Å), o bé estan més enllà del diàmetre de l'esfera (24.3 Å). En aquest procés queden buits, especialment al voltant del solut i, per aquest motiu, es torna a afegir una altra esfera d'aigua, es fa rotar, i es tornen a eliminar els àtoms que col·lapsen, i així successivament cinc vegades fins que el solut queda uniformement solvatarat.

```

LLEGEIX EL FITXER DE TOPOLOGIES
LLEGEIX EL FITXER DE PARAMETRES
GENERA EL SOLUT
LLEGEIX LES COORDENADES DE LA MOLECULA

Set 1 21.0
Set 2 24.3
Set 3 300.
Set 4 2.61
Set 5 -2.81
Set 6 12.13
    } Defineix variables.

Scalar x statistics sele resid 8 .or. resid 16 end
Scalar y statistics sele resid 8 .or. resid 16 end
Scalar z statistics sele resid 8 .or. resid 16 end
Coor translate xdir -@4 ydir -@5 zdir -@6
    } Centra el solut a l'origen de
    coordenades.

ESCRIU LES COORDENADES DEL SOLUT CENTRAT A L'ORIGEN
GENERA EL SOLVENT
LLEGEIX COORDENADES DEL SOLVENT

Delete atom sele .byres. (segid sph .and. type OH2 .and. (.not. segid sph .and. .not.
(hidrogen .or. lone)) .around. 2.5)) end
Delete atom sele .byres. (.not. (point 0.0 0.0 0.0 cut @2) .and. (segid sph .and. type oh2)
end
    } Elimina les molècules d'aigua que
    estan a una distància menor de 2.5 Å
    del solut o a una distància major de
    24.3 Å del centre de coordenades.

Delete atom sele .byres. (segid sph .and. type OH2 .and. (.not. segid sph .and. .not.
(hidrogen .or. lone)) .around. 2.5)) end
Delete atom sele .byres. (.not. (point 0.0 0.0 0.0 cut @2) .and. (segid sph .and. type oh2)
end
    } Gira la segona esfera.

Torna a LLEGIR COORDENADES DEL SOLVENT
Coor rotate phi -65.0 ydir 1.0 sele segid SPH end
Coor rotate phi -60.0 xdir 1.0 sele segid SPH end
    } Torna a eliminar les molècules d'aigua
    que col·lapsen amb altre àtoms.

Delete atom sele .byres. (segid SPH .and. type oh2 .and. ((.not. segid SPH .and. .not.
(hydrogen .or. lone) .and. (.not. segid |refs)) .around. 2.5 )) end
Dele atom sele .byres. (.not. (point 0.0 0.0 0.0 cut 24.0) .and. (segid SPH .and. type
oh2)) end

ESCRIU LES COORDENADES DEL SISTEMA SOLVATAT
    
```

**x 5** {

**Fitxer D4 Exemple de fitxer per solvatar el solut amb CHARMM**

**D.5 EXEMPLE DE FITXER PER REALITZAR UNA SIMULACIÓ DE DINÀMICA MOLECULAR**

A continuació, es mostra un exemple del fitxer d'entrada utilitzat per realitzar les simulació de DM en solució aquosa sota condicions estocàstiques de contorn (Fitxer D5). En aquest cas, les tres etapes de la simulació (escalfament, equilibració i producció) estan detallades en el mateix fitxer d'entrada. Al principi del fitxer (no es mostra aquí) es llegeixen els fitxers de topologies i paràmetres així com les coordenades del sistema. En aquest cas, les molècules d'aigua es mantenen en l'esfera per l'efecte d'un potencial esfèric. Aquest, es pre-calcula per a esferes d'una grandària determinada, per exemple, en aquest cas s'utilitza un potencial per a una esfera de 24 Å de diàmetre (wat24.pot). Del fitxer cal destacar que, quan s'ajusta el potencial, és necessari definir l'origen d'aquest, normalment coincideix amb l'origen de coordenades i, a continuació, s'estableix que el potencial actüi sobre tots els oxígens pertanyents als residus d'aigua (TIP3). Aquest potencial és una funció de la distància entre l'origen i els àtoms d'oxigen, i és pràcticament

igual a zero fins que es supera el límit de l'esfera. Un cop definit el potencial ja s'inicia la simulació de dinàmica molecular.

En la simulació de dinàmica molecular s'utilitza per defecte l'algoritme de *Lepfrog* per integrar les equacions del moviment. El tipus de simulació que es duu a terme està basada en les equacions de Langevin. En les tres etapes el *time step* és d'1 fs i el nombre de passos per a cada etapa està definit pel valor de la comanda *nstep*. La simulació es duu a terme a 300 K (*finalt*) i el bany termostàtic està a la mateixa temperatura (*tbath*). Les comandes *twindh* i *twindl* defineixen el límit superior i inferior de desviació de la temperatura respecte la temperatura final de cada etapa de la simulació. Les comandes *teminci* i *ihfrq* de l'etapa d'escalfament indiquen que la temperatura s'incrementa en 10 K cada 50 passos de la simulació. Finament, són importants les comandes *ilbfrq* i *nbfrq*. La primera defineix la freqüència per comprovar si un àtom està o no dins de la regió de Langevin (definida per *rbuf*) i la segona determina la freqüència amb la que es regenera la llista d'interaccions no enllaçants.

Les característiques de les interaccions no enllaçants entren dins de la subrutina *nbonds*. En aquest cas, es consideren les interaccions electrostàtiques i de van der Waals (*elec* i *vdw*) entre àtoms (*atom*). Les primeres es tracten mitjançant una funció de *shifting* (*shift*) i les segones mitjançant una funció de *switching* (*vswitch*). Les funcions de *shifting* modifiquen tota la superfície de potencial, per tal que en la distància de *cutoff* el potencial d'interacció sigui nul. D'altra banda, en les funcions de *switching* la corba de potencial s'estreta al llarg d'un rang predeterminat de distàncies de manera que tendeix gradualment a zero.

```

set b 21.0 !buffer boundary
set l 24.3 !Langevin boundary
set t 300 !bath temperature

define dímer sele (atom bcnp * *) end
define sphere sele point 0.0 0.0 0.0 cut @l end

open unit 1 form read name -
"~/SETUP_MDwat/wat24.pot"
sbound read unit 1
close unit 1

sbound set xref 0.0 yref 0.0 zref 0.0 -
assign 1 sele resname TIP3 .and. type O* .and. -
(.byres. (point 0.0 0.0 0.0 cut @l)) end

scalar fbeta set 62.0 sele resname TIP3 .and. type OH2 end

! SIMULACIO DE DM
shake bonh tolerance 1.0e-06

! ETAPA D'ESCALFAMENT
open unit 20 write form name nom_heat_mdw.rst"
open unit 22 write uniform name nom_heat_mdw.dcd"
open unit 24 write form name nom_heat_mdw.ene"

DYNALANGEVIN langevin strt nstep 20000 timest 0.001 rbuf @b iseed 27041994 -
firststt 50. finalt 300. tbath @t.
twindh 2.0 twindl -2.0 -
ieqfrq 100 iasvel 1 ISVFRQ 5000 -
IHTFRQ 50 TEMINC 10 -
ilbfrq 5 inbfrq 10 -

IUNREAD 7 IUNWRI 20 IUNCRD 22 KUNIT 24 -
nprint 50 iprfrq 20 nsavc 10 -

elec vdw atom shift vswitch vdistance e14fac 1.0 -
cdie eps 1.0 ctonnb 12.0 ctofnb 13.0 cutnb 14.0 ctexnb 999.0
ESCRIU LES COORDENADES FINALS D'AQUESTA ETAPA

! ETAPA D'EQUILIBRACIÓ
LLEGEIX COORDENADES
OBRE LES UNITATS DE FORTRAN CORRESPONENTS ALS FITXERS DE SORTIDA

DYNALANGEVIN langevin rest nstep 30000 timest 0.001 rbuf @b iseed 27041994 -
firststt 300. finalt 300. tstruct 300. tbath @t -
twindh 0.0 twindl 0.0 icheckw 0 -
ieqfrq 100 iasvel 1 ISVFRQ 5000 -
ilbfrq 5 inbfrq 10 -
ESCRIU RESULTATS ALS FITXERS DE SORTIDA I COORDENADES FINALS ETAPA
MATEIXOS PARAMETRES INTERACCIONS NO ENLLAÇANTS QUE L'ETAPA ANTERIOR

!ETAPA DE PRODUCCIÓ (10*100ps)
set p 0
set n 100
label loop

if n eq 100 open unit 7 read form name nom_equi_mdw.rst"
if n gt 100 open unit 7 read form name nom_prod_mdw_@P.rst"
OBRE LES UNITATS DE FORTRAN CORRESPONENTS ALS FITXERS DE SORTIDA

DYNALANGEVIN langevin rest nstep 100000 timest 0.001 rbuf @b iseed 27041994 -
firststt 300. finalt 300. tstruct 300. tbath @t -
twindh 0.0 twindl 0.0 icheckw 0 -
ieqfrq 100 iasvel 1 ISVFRQ 5000 -
ilbfrq 5 inbfrq 20 -
ESCRIU RESULTATS ALS FITXERS DE SORTIDA I COORDENADES FINALS ETAPA
MATEIXOS PARAMETRES INTERACCIONS NO ENLLAÇANTS QUE ETAPES ANTERIORS

set p @n
incr n by 100
if n le 1000 goto loop

```

**Defineix la zona reactiva i la zona de contorn.**

**Llegeix el potencial esfèric.**

**Ajusta el potencial esfèric.**

**Ajusta els coeficients de fricció.**

**Obre les unitats de fortran que corresponen als fitxers de sortida de la simulació: coordenades finals .rst, trajectòria .dcd i energies .ene.**

**Escriu resultats als fitxers de sortida.**

**Defineix paràmetres interaccions no enllaçants.**

**Fitxer D5 Exemple de fitxer per realitzar la simulació de dinàmica molecular amb CHARMM**

