

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tesisenxarxa.net) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tesisenred.net) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tesisenxarxa.net) service has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized neither its spreading and availability from a site foreign to the TDX service. Introducing its content in a window or frame foreign to the TDX service is not authorized (framing). This rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author

**ESCOLA TÈCNICA SUPERIOR D'ENGINYERIA
INDUSTRIAL DE BARCELONA**

Departament d'Enginyeria Química

**ESTUDI DE LA UTILITZACIÓ DELS MAPES DE
POTENCIAL ELECTROSTÀTIC I DE POLARITZACIÓ
COM A DESCRIPTORS MOLECULARS**

Memòria per a optar al títol de Doctor a càrrec de:

Maria Lourdes Roset Calzada

Director:

Dr. Juan Jesús Pérez González

Agraïments :

En primer lloc vull agrair al Dr. Juan Jesús Pérez González, director d'aquesta Tesi, l'oportunitat de treballar en el seu grup en el Laboratori d'Enginyeria Molecular, i el seu suport en tot moment, moltes gràcies per tot l'aprenentatge científic, pel seu suport i ànims en tot moment, i la seva excepcional qualitat humana, tant a nivell de treball com a nivell personal.

Al Dr. Jesús Gómez Catalán, pel seu suport en temes de toxicitat.

A tots els companys del LEM i a tots aquells que han col·laborat en diferents moments, Josep Cantó, Arnau Cordoní, Francesc Corcho, Patricia Gómez, Àlex Martínez, Cecy, Marlet, David, Mónica. A tots els companys del DEQ de l'EPSEVG, en especial a en Miquel, Montse, Nati, Salvador, amb qui he compartit tantes hores de treball.

A la gent del departament de química de la UB, Jaime, Juan Carlos, Laura, Cristian, Oscar, Marta, gràcies per el vostre suport.

Als meus pares, per la seva ajuda incondicional, i al meu germà, en Jaume Roset, que sempre m'ha donat suport i ajuda en temes científics.

Al Carlos i la Laura, per tot.

*A la meva família,
en especial a la meva filla,
la Laura*

1	INTRODUCCIÓ.....	3
2	DESCRIPCIÓ DELS MÈTODES DE CÀLCUL.....	11
	2.1 Mètodes de càlcul de la funció d'ona multieletrònica d'un sistema molecular.....	13
	2.2 Càlcul de les propietats moleculars: moments multipolars permanents	18
	2.3 Càlcul de les propietats moleculars: moments multipolars induïts.....	19
	2.4 Anàlisi de correlació estadística sobre conjunts de punts.....	21
3	EFFECTE DE LA BASE I DEL MÈTODE DE CÀLCUL A L'AVAUACIÓ DE LES PROPIETATS ELÈCTRIQUES DE PRIMER I SEGON ORDRE	25
	3.1 Càlcul de moments dipolar i quadrupolar	28
	3.2 Càlculs de polaritzabilitat i primera hiperpolaritzabilitat moleculars	32
4	EFFECTE DE LA BASE I DEL MÈTODE DE CÀLCUL A L'AVAUACIÓ DELS POTENCIAL ELECTROSTÀTIC I DE POLARITZACIÓ MOLECULARS.....	41
	4.1 Càlcul dels potencials electrostàtic i de polarització	43
	4.2 Anàlisi general dels mapes de polarització calculats al present estudi.....	45
	4.3 Anàlisi comparativa dels valors del MEP i MPP calculats amb diferents bases i metodologies	48
	4.4 Anàlisi espacial dels valors del MEP i MPP calculats amb diferent bases i metodologies	57
5	LA POLARITZABILITAT MOLECULAR COM A DESCRIPTOR DE L'ACTIVITAT BIOLÒGICA DE MOLÈCULES: ESTUDI DE DIOXINES I FURANS.....	73
	5.1 Introducció.....	75
	5.2 Dioxines i furans.....	76
	5.3 Càlcul dels potencials electrostàtic i de polarització	80
	5.4 Mapes de potencial electrostàtic, de polaritzabilitat i interacció.....	80
	5.4.1 Descripció general	80
	5.4.2 Anàlisi de components principals.....	109
	5.5 Càlculs d'ajust parcial per mínims quadrats (Partial Least Square).....	134
	5.5.1 Ajust mitjançant PLS i predicció de toxicitat.....	134
	5.5.2 Estudi de les zones de màxima variabilitat	146
6	CONCLUSIONS.....	151
7	BIBLIOGRAFIA	155
8	ANNEXOS.....	169
	8.1 Annex 1. Anàlisi de correlació. Càlculs MEP i MPP	171
	8.2 Annex 2. Geometries optimitzades de dioxines i furans	195
	8.3 Annex 3. Programes de càlcul emprats.....	203
	8.4 Annex 4. Publicacions	207
	

1 INTRODUCCIÓ

El paradigma fonamental de l'Enginyeria molecular consisteix en l'aprofitament del coneixement de les característiques estereoelectròniques que defineixen el reconeixement molecular. Així doncs, el disseny d'anticossos catalítics o de nous enzims per a la catàlisi de processos d'interès industrial en biotecnologia, el disseny de nous agents terapèutics, així com el disseny de vacunes sintètiques, disseny de materials amb propietats òptiques no-lineals, components electrònics moleculars o el disseny de molècules auto-assemblants, etc, depenen directament del mecanisme de reconeixement entre dues molècules.

A nivell molecular, el reconeixement molecular és el resultat d'una complementarietat, tant geomètrica com electrònica, entre dues entitats moleculars. La fortalesa d'aquest reconeixement es pot quantificar mitjançant l'energia d'interacció, definida com la diferència entre l'energia del sistema global (mètode de la supermolècula) menys l'energia de cadascun dels subsistemes és a dir, de cada molècula aïllada. L'energia d'interacció dona una mesura de la complementarietat entre les dues molècules i per tant de la intensitat del seu reconeixement mutu.

Les interaccions moleculars són interaccions electromagnètiques febles, ja que es produeixen entre sistemes electrònics de capes tancades, i per tant, les energies d'interacció són del ordre de magnitud de les kilocalories per mol, petites en comparació als centenars de kilocalories per mol dels enllaços covalents, presentant tanmateix distàncies d'equilibri més llargues que les dels enllaços químics.

Formalment, les interaccions moleculars o forces intermoleculars es poden interpretar en termes de les interaccions electrostàtiques entre dues distribucions de càrrega, i es poden descriure de forma intuïtiva mitjançant la representació multipolar dels sistemes de càrrega. De fet, a llargues distàncies l'energia d'interacció d'una distribució de càrrega i un camp elèctric F es pot descriure per l'efecte produït per superposició, en ordre decreixent d'importància, dels moments multipolars de la distribució: càrrega neta, dipol, quadrupol, octopol i així successivament (Buckingham, 1959; Hinchliffe, 1985):

$$E(F) = E(0) + q\phi_o - \mu_\alpha F_\alpha - \frac{1}{2!} Q_{\alpha\beta} F'_{\alpha\beta} - \frac{1}{3!} R_{\alpha\beta\gamma} F'''_{\alpha\beta\gamma} - \dots$$

on s'ha utilitzat la notació dels índexs muts d'Einstein, on els índexs repetits impliquen una suma estesa a totes les components del tensor.

El sentit físic d'aquests moments multipolars es pot veure a través de distribucions senzilles de càrrega. Així doncs, una distribució de càrrega formada per dues càrregues iguals i de signe contrari es pot representar per un únic dipol. Igualment, una distribució de càrrega de quatre càrregues iguals dues de signe positiu i dues de signe negatiu, disposades de forma que el moment dipolar sigui nul, es pot representar per un únic quadrupol. De forma semblant, una distribució amb únicament moment octopolar, correspondria a vuit càrregues iguals i amb signes contraris dos a dos, situades als vèrtexs d'un cub. En general, els moments multipolars tenen dimensions $[Q L^n]$ ($n=0,1,2,\dots$): per a $n=0$ s'obté la càrrega neta; per a $n=1$ el dipol $q \cdot r$; per a $n=2$ el quadrupol $q \cdot r^2$, etc...

Així doncs, tal i com hem vist, la interacció d'una distribució de càrregues amb un camp elèctric extern es pot descriure com a una superposició multipolar. Però, de fet, aquest camp elèctric extern pot produir una polarització del sistema de càrregues i com a conseqüència que els moments multipolars depenguin del camp extern. Aquesta polarització es pot expressar en termes dels moments dipolars induïts, que depenen de les polaritzabilitats i hiperpolaritzabilitats dels sistemes de càrrega. Així doncs el moment dipolar total de un sistema de càrregues en funció del camp elèctric extern té l'expressió (Buckingham, 1959; Hinchliffe, 1985):

$$\mu_{\alpha}(F) = \mu_{\alpha}(0) + \alpha_{\alpha\beta} F_{\beta} + \frac{1}{2!} \beta_{\alpha\beta\gamma} F_{\beta} F_{\gamma} + \dots$$

on $\mu(0)$ és el moment dipolar permanent, α és un tensor simètric d'ordre dos que es coneix com a polaritzabilitat molecular i representa la resposta lineal d'una distribució electrònica molecular a l'acció d'un camp elèctric extern F , i β és un tensor d'ordre tres conegut com a primera hiperpolaritzabilitat dipolar, i representa el factor més important de la resposta no-lineal del moment dipolar a un camp elèctric extern.

Una expressió semblant es pot trobar mitjançant l'expressió de l'energia d'interacció d'una distribució de càrregues en presència d'un camp elèctric extern F (Buckingham, 1959; Hinchliffe, 1985):

$$E(F) = E(0) - \mu^{\circ}{}_{\alpha} F_{\alpha} - \frac{1}{2} \alpha_{\alpha\beta} F_{\alpha} F_{\beta} - \frac{1}{6} \beta_{\alpha\beta\gamma} F_{\alpha} F_{\beta} F_{\gamma} + \dots$$

on $E(0)$ és l'energia de la molècula en absència de camp elèctric.

Ara pensem en que qualsevol distribució de càrrega genera un camp elèctric a llarga distància que es pot descriure com el produït per la superposició del camp generat per una

càrrega neta si n'hi ha, el moment dipolar, el moment quadrupolar, el moment octopolar, etc... associats. Com a conseqüència, la interacció entre dues molècules es pot expressar com la interacció del camp elèctric generat per una molècula, expressada en termes dels seus moments multipolars i una altra distribució de càrrega descrita també pels seus moments multipolars. Efectivament, si considerem les molècules com a distribucions de multipols polaritzables, l'energia d'interacció entre dues molècules es pot expressar com una contribució entre els multipols permanents (terme de Coulomb), a més de les contribucions dels moments multipolars induïts (terme de inducció). De fet, si tenim en compte les dimensions de l'energia [$Q^2 L^{-1}$] cada terme està dividit per una funció de la distància amb exponents cada vegada més grans. Així per exemple, mentre que el terme càrrega-càrrega depèn de r^{-1} , el terme dipol-dipol depèn de r^{-3} , etc... (Buckingham, 1967).

Si apliquem la teoria de pertorbacions per a calcular el valor de l'energia d'interacció entre dues molècules utilitzant com a operador de pertorbació l'energia d'interacció de dues distribucions electròniques de càrrega, l'energia d'interacció es pot calcular com un terme de primer ordre que correspon a la energia de Coulomb, seguit d'un terme de segon ordre que conté l'energia d'inducció, més un terme addicional creuat no-clàssic, que s'anomena energia de dispersió i que és degut a la correlació electrònica (Hobza, 1981).

Una eina molt important que avui en dia s'utilitza per dur a terme càlculs en enginyeria molecular és la mecànica molecular (Bowen, 1991). Es tracta d'una eina de càlcul potent que serveix per simular el comportament de sistemes d'un nombre elevat d'àtoms, com proteïnes, on es pot també descriure explícitament el dissolvent a la simulació.

A la mecànica molecular l'energia d'interacció es calcula a partir de dues contribucions: una contribució enllaçant que inclou les desviacions respecte a la geometria estàndard de distàncies d'enllaç i angles d'enllaç, representada per un potencial harmònic, i una contribució no enllaçant que inclou interaccions moleculars.

Els camps de força que s'utilitzen inclouen la contribució de Coulomb (interaccions entre monopols no polaritzables), més un terme de dispersió, representat per un potencial de Lennard-Jones. El terme de polarització, molt més car des de el punt de vista computacional, no s'inclou normalment en aquests càlculs. Tot i així, publicacions recents (Ma, 2004; Engkvist, 2000) han demostrat la importància dels paràmetres corresponents utilitzant tècniques basades en càlculs mecano-quàntics. Per altra banda trobar paràmetres adients per descriure la polarització no és

fàcil. Una forma de fer-ho seria calculant la polarització molecular per mètodes mecano-quàntics i després trobar els paràmetres que millor reproduïen aquests resultats.

La polarització molecular es pot calcular com a diferència entre els mapes d'interacció d'una molècula amb una càrrega puntual, i el mapa corresponent al potencial electrostàtic del sistema molecular (Alkorta, 1993 i 1995). Per això, és necessari definir una xarxa de punts al voltant de la molècula i calcular l'energia del sistema molècula-càrrega a cadascun dels punts de la xarxa amb un càlcul Hartree-Fock. A continuació es calcula el potencial electrostàtic a cadascun dels punts de la xarxa com a valor esperat de l'operador potencial electrostàtic. La diferència entre els dos camps correspon a la contribució de la polarització molecular. Segons la càrrega de prova utilitzada, el mapa de polarització calculat es pot considerar com la resposta lineal de polarització.

L'energia d'interacció d'una molècula amb una càrrega externa 'q' es pot expressar com una expansió pertorbativa en termes de 'q'. Els coeficients de l'expansió són les contribucions de diferents ordres per l'energia d'interacció:

$$\Delta E = qE^{(1)} + q^2 E^{(2)} + q^3 E^{(3)} + \dots$$

La contribució de primer ordre inclou el terme de Coulomb a l'energia d'interacció, i representa l'energia d'interacció amb una càrrega externa 'q'. Els termes de segon ordre i superiors inclouen les contribucions a l'energia degudes a la polarització de la distribució de càrregues per la presència de la càrrega exterior 'q'.

Si 'q' és la càrrega unitària 'e', el primer terme és el potencial electrostàtic molecular, i si es considera que termes de tercer ordre i superiors són menyspreables, l'equació anterior es pot aproximar com :

$$\Delta E = eV(r) + e^2 P(r)$$

on V(r) és el potencial electrostàtic de la molècula i P(r) el potencial de polarització molecular.

La importància de les diferents contribucions electrostàtiques ens permet realitzar un estudi teòric de predicció de la reactivitat i altres propietats moleculars a partir de càlculs dels potencials electrostàtic i de polarització moleculars. Tot i que el terme electrostàtic és el que té més pes a l'energia d'interacció, aquest terme només representa la interacció d'una distribució

estàtica de càrregues amb un camp elèctric extern. El terme de polarització ens dóna una millora en la descripció de les diferents interaccions moleculars.

Els mapes de potencial electrostàtic molecular (MEPs) i potencial de polarització molecular (MPPs) són representacions gràfiques de les diferents contribucions a l'energia d'interacció entre una càrrega unitat i la molècula a diferents punts de l'espai. La polarització d'una molècula deguda a la interacció amb un camp elèctric extern està relacionada amb la seva polaritzabilitat i hiperpolaritzabilitat. Com diferents 'parts' de la molècula tenen diferents polaritzabilitats, la 'polaritzabilitat molecular neta' depèn de la distribució topogràfica de la molècula. Així, regions moleculars amb dobles enllaços, sistemes aromàtics, o parells solitaris presenten polaritzacions més grans.

A partir dels mapes de polarització molecular es poden estudiar les diferents regions de la molècula segons la seva polarització, i predir els llocs més rellevants per un atac nucleofílic, així com estimar els efectes de la contribució de la polarització amb interaccions amb altres molècules. Estudis recents de simulació molecular i càlculs de mecànica quàntica demostren la importància del terme de polarització molecular, obtenint una millora en la descripció de la reactivitat química molecular.

Els resultats del present treball consten de tres parts. A la primera part es realitza un estudi de l'efecte de la base i de la metodologia emprada al càlcul de propietats elèctriques de primer i segon ordre. A la segona part s'estudia la distribució espacial de la polaritzabilitat molecular mitjançant els mapes de potencial de polarització. Aquí es fa també un estudi de l'efecte de la base i de la metodologia emprada pel càlcul. A la tercera part es fa un estudi de l'efecte d'inclusió de la polarització molecular en la predicció de l'activitat biològica de dioxines i furans.

En el capítol 3 s'analitza l'efecte de diferents conjunts de base a la contribució de la polarització a l'energia d'interacció utilitzant diferents conjunts de bases: 6-31G(d) (Harihara, 1973), 6-31G(d,p) (Harihara, 1973) i 6-31G(2d,2p) (Dunning, 1977) i l'efecte de diferents nivells de càlcul: SCF, MP2, MP4 (Møller, 1934), BLYP(Parr, 1989) i B3LYP (Becke, 1988). Aquest estudi es porta a terme sobre tres sistemes moleculars senzills: cianur d'hidrogen, formaldehid i urea. Per a cada sistema es van calcular els moments dipolars, moments quadropolars, polaritzabilitat i hiperpolaritzabilitat moleculars amb les diferents bases i metodologies.

En el capítol 4 es discuteix l'efecte de la base i del nivell de càlcul sobre el potencial de polarització. Com s'ha dit abans, el mapa de polarització d'una molècula es calcula per diferència entre el mapa d'interacció i el mapa corresponent al potencial electrostàtic. Concretament, l'energia d'interacció amb una càrrega q externa en un punt d'una xarxa de punts es pot expressar (Alkorta, 1993 i 1995) com a:

$$\Delta E(xyz) = qMEP(xyz) + q^2MPP(xyz) + q^3MHPP(xyz) + \dots$$

menyspreant el tercer terme i els superiors, tenim una equació per calcular els mapes de polarització molecular a partir dels mapes de potencial electrostàtic (MEP) i d'interacció. Aquests càlculs del potencial de polarització s'han realitzat utilitzant els programes GAUSSIAN 94 (Frisch, 1995) i GAUSSIAN 98 (Frisch, 1998).

Per intentar estalviar el cost computacional que requereixen els càlculs del potencial de polarització sobre tot pensant en molècules grans, s'ha fet també una comparació amb els resultats obtinguts utilitzant la teoria de pertorbacions de Rayleigh-Schrödinger (Sachse, 2002) mitjançant el programa MEPHISTO (Francl, 1985). L'anàlisi d'aquests resultats permetrà realitzar l'estudi de mapes de polarització com a indicadors de l'activitat biològica en dioxines i furans en base a la metodologia que es consideri millor, valorant també el temps computacional necessari.

En el capítol 5 es realitza l'estudi de l'efecte de la inclusió de la polarització molecular com a descriptor en la predicció de l'activitat biològica de dioxines i furans. Per aquest propòsit s'han calculat els potencials electrostàtic, de polarització i l'energia d'interacció per dioxines i furans utilitzant el programa de càlcul Mephisto. Es defineixen les principals zones d'interacció electrostàtica i de polarització molecular a partir del càlcul de components principals (PCA) realitzat amb el programa SIMCA (SIMCA-P. Versió 10.0, UMETRICS AB, 2002) i la posterior representació gràfica del MEP i MPP corresponents. La predicció de la activitat biològica en funció del potencial electrostàtic i del potencial de polarització es basa en l'estudi mitjançant càlculs PLS (Wold, 1966, 1978, 1983) per a un conjunt determinat de dioxines, furans, o ambdós, i la seva correlació amb dades experimentals de l'activitat biològica. En funció dels resultats obtinguts es realitzarà un estudi més detallat dels càlculs PLS realitzats.

2 DESCRIPCIÓ DELS MÈTODES DE CÀLCUL

En el present treball el càlcul de la funció d'ona s'ha realitzat utilitzant els mètodes de Hartree-Fock i Møller-Plesset (Szabo, 1977,1996) a diversos ordres i de la teoria del funcional de la densitat (Parr, 1989). Respecte a les propietats elèctriques, els moments multipolars permanents s'han calculat com valors esperats de la corresponent funció d'ona i els moments multipolars induïts s'han calculat a través de les corresponents derivades de l'energia en funció del camp elèctric. Finalment, també s'ha utilitzat la teoria de pertorbacions per calcular de forma aproximada el potencial de polarització. La majoria dels càlculs s'han portat a terme mitjançant el programa GAUSSIAN 94 (Frisch, 1995) i el càlcul del potencial de polarització utilitzant la teoria de pertorbacions amb el programa MEPHISTO (Francl, 1985). A continuació es presenta una breu descripció dels mètodes de càlcul utilitzats.

2.1 Mètodes de càlcul de la funció d'ona multieletrònica d'un sistema molecular

Per al càlcul de la funció d'ona multieletrònica d'un sistema molecular normalment s'utilitza el mètode de Hartree-Fock, que és un mètode variacional basat en l'aproximació orbital i que descriu el sistema a través d'una única configuració electrònica. La diferència entre l'energia calculada amb el mètode Hartree-Fock i l'energia experimental s'anomena energia de correlació. Per a fer una estimació d'aquesta energia es poden utilitzar mètodes pertorbacionals com el de Møller-Plesset (MP), que ha estat utilitzat en aquest treball, o variacionals com els mètodes multiconfiguracionals.

Alternativament, el mètode del funcional de la densitat (DFT) permet calcular directament de forma aproximada la densitat electrònica del sistema incloent part de l'energia de correlació.

Mètode Hartree-Fock

En l'aproximació Hartree-Fock (HF), el hamiltonià polieletrònic s'aproxima com una suma de N sistemes monoelèctrics, en el que cada electró es troba sotmès a la atracció pels nuclis i a la repulsió per la resta de electrons, que es tracta de forma promig. El mètode consisteix en optimitzar els coeficients dels orbitals que donen l'energia més baixa. L'aplicació del mètode forma un sistema d'equacions no lineals, i l'aplicació del principi variacional a la resolució d'aquest sistema, fa que l'optimització dels coeficients es faci de forma iterativa, donant nom al procediment conegut com camp autoconherent o self-consistent field (SCF).

L'aproximació orbital consisteix en definir una funció multieletrònica com a producte de funcions monoelèctriques (orbitals):

$$\psi = \hat{A} \phi_1(r_1) \phi_2(r_2) \dots \phi_n(r_n)$$

on \hat{A} és l'operador d'antisimetria d'acord amb el principi d'exclusió de Pauli.

L'energia potencial de l'electró i en el camp de l'electró j ve donada per l'expressió:

$$E_{p_i}^j = e \int_V \frac{\phi_j(r_j) \phi_j(r_j)}{r_{ij}} d\tau_j = V_{ij}^{effect}$$

i el potencial efectiu per l'electró i en un camp de varis electrons com

$$V_i^{effect} = \sum_{j \neq i}^n \int \frac{\phi_j \phi_j}{r_{ij}} dr_j = \sum_{j \neq i}^n J_j(\vec{r}_i)$$

aixó dona per a cada electró un hamiltonià efectiu h_i , i el de tot el sistema és la suma dels corresponents a cada electró :

$$h_i^o = -\frac{1}{2} \nabla_i^2 - \sum_k^{n_n} \frac{Z_k}{r_{ik}}$$

$$h_i = h_i^o + V_i^{effect} = h_i^o + \sum_{j \neq i}^n J_j(\vec{r}_i)$$

$$H^{HF} = \sum_{i=1}^n h_i$$

de manera que es transforma el problema de N electrons en N problemes d'un electró:

$$\{h_i \phi_i = \varepsilon_i \phi_i\}$$

que es resol variacionalment.

La implementació del mètode es fa a través del mètode SCF (Self-consistent-field), on partint d'un conjunt d'orbitals $\{\phi_i\}$ amb paràmetres variables, el mètode variacional ens permet arribar al conjunt de valors dels paràmetres que minimitzen l'energia del sistema, arribant a la autoconsistència :

$$h_i^{(n)} \phi_i^{(n+1)} = \varepsilon_i^{(n+1)} \phi_i^{(n+1)} \quad \text{on}$$

$$\phi_i^{(n+1)} \cong \phi_i^{(n)} \quad \text{i} \quad \varepsilon_i^{(n+1)} = \varepsilon_i^{(n)}$$

En el cas que es tracti d'un sistema molecular, els orbitals moleculars normalment s'aproximen com a combinacions lineals d'orbitals atòmics (aproximació MO-LCAO), on els paràmetres a optimitzar són els coeficients de les combinacions lineals.

Mètode pertorbacional de Møller-Plesset

L'energia del sistema que dona una funció Hartree-Fock és en molt casos una bona aproximació de la energia experimental. A la diferència entre aquest dos valors se l'anomena energia de correlació.

El mètode pertorbacional de *Møller-Plesset* és un mètode pertorbacional per estimar l'energia de correlació electrònica, on s'utilitza la funció d'ona HF com a funció del sistema de referència (sense pertorbar). Utilitzarem la nomenclatura MPn per representar els diferents ordres de pertorbació. A les correccions de segon i tercer ordre a l'energia només contribueixen els determinants diexcitats.

Es considera el hamiltonià com a suma de l'operador del sistema de referència, pel que la solució és la funció d'ona Hartree-Fock, i l'operador pertorbació.

$$H_{\text{exacte}} = H_0 + \lambda H_1$$

$$H_0 \psi^{(0)} = E^{(0)} \psi^{(0)}$$

$$E^{(0)} = \langle \psi^{(0)} | H_0 | \psi^{(0)} \rangle$$

Els càlculs MP2 (segon ordre) són els més usuals, i l'energia de correlació es defineix :

$$E_{\text{corr}}^{\text{MP2}} = \sum_{i,j,a,b} \frac{2(ia | jb) - (ib | ja)}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b} (ia | jb)$$

En general tenim :

$$E_{\text{MP0}} = E_0^{(0)} = \sum 2\varepsilon_i \quad \text{on } \varepsilon_i \equiv \text{energies dels OM ocupats HF}$$

$$E_{\text{MP1}} = \sum 2\varepsilon_i + E_0^{(1)} = E_{\text{HF}} \quad \text{on } E_{\text{HF}} \equiv \text{energia de Hartree-Fock}$$

$$E_{\text{MP2}} = E_{\text{HF}} + E_0^{(2)}$$

Pel que fa al temps de càlcul, un càlcul MP2 és 1.5 cops més costós que un càlcul HF, i per MP3 i MP4 els factors són de 3.6 i 5.8 aproximadament.

Teoria del funcional de la densitat

El mètode DFT (density functional theory)(Parr, 1989) no té com a objectiu calcular la funció d'ona, si no calcular la probabilitat electrònica molecular, i d'aquí calcular l'energia electrònica molecular. La densitat electrònica es defineix com la densitat de probabilitat de trobar un electró en un punt de l'espai multiplicada per el nombre d'electrons:

$$\rho(\mathbf{r}) = \rho(\vec{r}_1) = N \int ds_1 dx_2 \dots dx_N \psi_0^* \psi_0$$

La densitat electrònica ens dona la probabilitat de trobar simultàniament dos electrons, de manera que conté informació sobre el moviment correlacionat de dos electrons. Es pot calcular l'energia a partir de la densitat electrònica amb el mètode Kohn-Sham (Kohn, 1965).

El DFT està basat en els teoremes de Hohenberg i Kohn (Hohenberg, 1964):

1. Qualsevol estat estacionari fonamental no degenerat pot ser calculat, de forma exacta, a partir de la densitat electrònica d'aquest estat fonamental, és a dir, pot ser descrit com un funcional de la densitat electrònica de l'estat fonamental.
2. La densitat electrònica d'un estat fonamental no degenerat pot calcular-se, de forma exacta, calculant la densitat que minimitza l'energia de l'estat fonamental.

$$\rho(\vec{r}) \rightarrow V(\vec{r}) \rightarrow \hat{H} \rightarrow \psi$$

$$\left[\frac{\partial E_V(\rho)}{\partial \rho} \right] = 0$$

El principal problema del DFT és que no es coneix la fórmula matemàtica exacta que relaciona l'energia amb la densitat electrònica, i és necessari recórrer a expressions aproximades. Segons la formulació de Kohn-Sham de la teoria (Kohn, 1965), les equacions a resoldre són:

$$\left(-\frac{1}{2} \nabla_i^2 + V_{eff}(\vec{r}) \right) |\psi_i\rangle = \epsilon_i |\psi_i\rangle$$

on V_{eff} és el potencial d'un sistema de electrons no interaccionants, de manera que la densitat construïda a partir dels orbitals de Kohn-Sham és la mateixa que la densitat del sistema real.

El DFT està limitat a les propietats de l'estat fonamental i a la resposta a pertorbacions estàtiques externes. La aproximació TD-DFT (time-dependent density functional theory) inclou els efectes de les pertorbacions dependents del temps.

Teoria de pertorbacions de Rayleigh-Schrödinger

De forma aproximada s'ha calculat el potencial de polarització utilitzant la teoria de pertorbació de Rayleigh-Schrödinger a partir de la funció d'ona HF, avaluant les correccions de primer i segon ordre a l'energia d'una molècula interaccionant amb una càrrega puntual.

El potencial electrostàtic d'una molècula A (V_A) es pot escriure com :

$$V_A(\vec{r}_B, \vec{R}_A) = \int \frac{\rho_A(\vec{r}_A, \vec{R}_A)}{|\vec{r}_A - \vec{r}_B|} d\vec{r}_A$$

on r_A i R_A són les coordenades nuclear i electròniques respectivament, i r_B són les coordenades de la distribució de càrrega. Si considerem B com a un conjunt discret de punts 'k', l'energia d'interacció electrostàtica es pot escriure com :

$$E_{ES,A} = \sum_k q_k V_A(\vec{r}_{k_1}, \vec{R}_A)$$

$$E_{ES,A} = \langle \psi_A | \sum_A q_k \frac{1}{|\vec{r}_A - \vec{r}_k|} | \psi_A \rangle$$

Aquesta expressió és la correcció a l'energia de primer ordre de la teoria de pertorbació de Rayleigh-Schrödinger, on el terme

$$\sum_k q_k \frac{1}{|\vec{r}_A - \vec{r}_k|}$$

és el Hamiltonià de pertorbació.

La correcció de segon ordre ens dona els efectes de polarització :

$$E_{pol,A} \cong E_A^{(2)} \text{ és a dir :}$$

$$E_{pol,A} = \langle \psi^{(0)} | H' | \psi^{(1)} \rangle$$

on H' és al hamiltonià de pertorbació, i $\psi^{(0)}$ i $\psi^{(1)}$ són les correccions d'ordre zero i de primer ordre, respectivament.

Si χ_i són els orbitals de la funció d'ona no pertorbada :

$$E_{pol,A} = \sum_i^{occ} \sum_a^{vir} \left| \langle \chi_i | H' | \chi_a \rangle \right|^2 (\varepsilon_i - \varepsilon_A)^{-1}$$

i en funció del conjunt original de funcions de base ϕ_μ :

$$E_{pol,A} = \sum_i^{occ} \sum_a^{vir} \left| \frac{1}{\varepsilon_i - \varepsilon_A} \left\{ \sum_\mu \sum_\nu C_{\mu i} C_{\nu a} \langle \phi_\mu | H' | \phi_\nu \rangle \right\} \right|^2$$

Una expressió similar pot ser deduïda per la correcció de primer ordre:

$$E_{ES,A} = \sum_i^{occ} \sum_\mu C_{\mu i} C_{\nu a} \langle \phi_\mu | H' | \phi_\nu \rangle$$

El programa MEPHISTO utilitza aquestes dues últimes expressions per calcular els components electrostàtic i de polarització de l'energia d'interacció de Coulomb.

2.2 Càlcul de les propietats moleculars: moments multipolars permanents

Si tenim una pertorbació externa, podem tractar-la per el mètode pertorbatiu, considerant un nou Hamiltonià :

$$H(\lambda) = H^0 + \lambda H'$$

on H' és un operador que mesura alguna propietat de la molècula, i λ és un paràmetre que mesura la força de la interacció amb la pertorbació externa.

Per a petites pertorbacions el valor de la propietat corresponent a l'operador H' és la derivada de l'energia respecte a λ per a $\lambda=0$:

$$\left. \frac{\partial E(\lambda)}{\partial \lambda} \right|_{\lambda=0}$$

i aplicant el teorema de Hellmann-Feynman

$$\left. \frac{\partial E(\lambda)}{\partial \lambda} \right|_{\lambda=0} = \frac{\langle \phi | \left. \frac{\partial H(\lambda)}{\partial \lambda} \right|_{\lambda=0} | \phi \rangle}{\langle \phi | \phi \rangle} = \langle \phi | H' | \phi \rangle$$

El potencial creat per una distribució de càrrega en el punt R es defineix com:

$$V(\mathbf{R}) = \int \rho(\mathbf{r}) |\mathbf{R} - \mathbf{r}|^{-1} d\mathbf{r}$$

i desenvolupant en sèrie de Taylor aquesta expressió:

$$|\mathbf{R} - \mathbf{r}|^{-1} = \frac{1}{R} + \sum_i \left(\frac{\partial}{\partial r_i} \frac{1}{|\mathbf{R} - \mathbf{r}|} \right)_0 r_i + \frac{1}{2} \sum_{ij} \left(\frac{\partial^2}{\partial r_i \partial r_j} \frac{1}{|\mathbf{R} - \mathbf{r}|} \right)_0 r_i r_j + \dots$$

De manera que obtenim:

$$V(\mathbf{R}) = \sum_{l=0}^{\infty} V_l(\mathbf{R})$$

$$V_l(\mathbf{R}) = \frac{(-1)^l}{l!} \sum_{i,j,k,\dots} T_{i,j,k,\dots}^l(\mathbf{R}) \int \rho(\mathbf{r}) r_i r_j r_k \dots d\mathbf{r}$$

La càrrega total es defineix com:

$$q = \int \rho(\mathbf{r}) d\mathbf{r}$$

Les components del *moment dipolar* són:

$$\mu_i = \int \rho(\mathbf{r}) r_i d\mathbf{r}$$

Les components del *moment quadrupolar* venen donades per:

$$\theta_{ij} = \frac{1}{2} \int \rho(\mathbf{r}) (3r_i r_j - r^2 \delta_{ij}) d\mathbf{r}$$

i així successivament pels moments d'ordre superior.

2.3 Càlcul de les propietats moleculars: moments multipolars induïts

Un àtom o molècula en un camp extern es polaritza, amb una intensitat que depèn del moment dipolar intrínsec que presenti i del moment dipolar induït per el camp elèctric aplicat.

L'energia del sistema ve donada per:

$$\hat{H}(\mathbf{F}) = \hat{H}^0 - \hat{\boldsymbol{\mu}} \cdot \mathbf{F}$$

$$\hat{\boldsymbol{\mu}} = - \sum_{i=1}^N q_i \mathbf{r}_i$$

$$E(\mathbf{F}) = \langle \psi(\mathbf{F}) | \hat{H}(\mathbf{F}) | \psi(\mathbf{F}) \rangle$$

i desenvolupant en sèrie de potències de F:

$$E(\mathbf{F}) = E^0 + \sum_i \left(\frac{\partial E}{\partial F_i} \right)_{F=0} F_i + \frac{1}{2} \sum_{i,j} \left(\frac{\partial^2 E}{\partial F_i \partial F_j} \right)_{F=0} F_i F_j + \frac{1}{6} \sum_{i,j,k} \left(\frac{\partial^3 E}{\partial F_i \partial F_j \partial F_k} \right)_{F=0} F_i F_j F_k + \frac{1}{24} \sum_{i,j,k,l} \left(\frac{\partial^4 E}{\partial F_i \partial F_j \partial F_k \partial F_l} \right)_{F=0} F_i F_j F_k F_l + \dots$$

simplificant:

$$E(\mathbf{F}) = E^0 + \Delta E = E^0 + \sum_i \left(\frac{\partial E(\mathbf{F})}{\partial F_i} \right) F_i = E^0 - \sum_i \mu_i(\mathbf{F}) F_i = E^0 - \boldsymbol{\mu}(\mathbf{F}) \cdot \mathbf{F}$$

de manera que obtenim:

$$\mu_i(\mathbf{F}) = \mu_i^0 + \sum_j \alpha_{ij} F_j + \frac{1}{2} \sum_{j,k} \beta_{ijk} F_j F_k + \frac{1}{6} \sum_{j,k,l} \gamma_{ijkl} F_j F_k F_l + \dots \quad \text{on es defineix:}$$

- El moment dipolar permanent com la contribució d'ordre zero a ΔE :

$$\mu_i^0 = - \left(\frac{\partial E(\mathbf{F})}{\partial F_i} \right)_{F=0}$$

- El tensor polaritzabilitat com la contribució d'ordre u:

$$\alpha_{ij} = - \left(\frac{\partial^2 E(\mathbf{F})}{\partial F_i \partial F_j} \right)_{F=0} = \left(\frac{\partial \mu_i(\mathbf{F})}{\partial F_j} \right)_{F=0}$$

- El tensor primera hiperpolaritzabilitat com la contribució d'ordre dos:

$$\beta_{ijk} = - \left(\frac{\partial^3 E(\mathbf{F})}{\partial F_i \partial F_j \partial F_k} \right)_{F=0}$$

- El tensor segona hiperpolaritzabilitat com la contribució d'ordre tres:

$$\gamma_{ijkl} = - \left(\frac{\partial^4 E(\mathbf{F})}{\partial F_i \partial F_j \partial F_k \partial F_l} \right)_{F=0}$$

La polaritzabilitat quantifica la resposta de primer ordre del núvol electrònic amb el camp elèctric aplicat. A més gran és la polaritzabilitat, més gran el moment dipolar induït.

El tensor polaritzabilitat és simètric, i si coincideixen la direcció del camp elèctric aplicat amb l'eix de simetria principal es pot definir la **polaritzabilitat mitjana** com:

$$\bar{\alpha} = \frac{1}{3} (\alpha_{xx} + \alpha_{yy} + \alpha_{zz})$$

magnitud que descriu la resposta al camp extern amitjanada esfèricament.

- L'anisotropia de la polaritzabilitat es defineix com:

$$\eta^2 = \frac{1}{2} [(\alpha_{xx} - \alpha_{yy})^2 + (\alpha_{xx} - \alpha_{zz})^2 + (\alpha_{yy} - \alpha_{zz})^2]$$

Desenvolupant emprant la teoria de pertorbacions :

$$E = E^{(0)} + E^{(1)} + E^{(2)} + \dots = \langle \psi^{(0)} | \hat{H}^0 | \psi^{(0)} \rangle$$

$$E^{(1)} = \langle \psi^0 | \hat{H}^1 | \psi^0 \rangle = - \sum_i \langle \psi^0 | \hat{\mu}_i | \psi^0 \rangle F_i$$

$$E^{(2)} = \sum_n \left\{ \frac{\langle \psi^0 | \hat{H}^1 | \psi^n \rangle \langle \psi^n | \hat{H}^1 | \psi^0 \rangle}{E_0 - E_n} \right\} = \sum_n \sum_i \sum_j \left\{ \frac{\langle \psi^0 | \hat{\mu}_i | \psi^n \rangle \langle \psi^n | \hat{\mu}_j | \psi^0 \rangle}{E_0 - E_n} \right\} F_i F_j$$

I comparant amb les equacions anteriors:

$$\mu_i^0 = \langle \psi^0 | \hat{\mu}_i | \psi^0 \rangle$$

$$\alpha_{ij} = -2 \sum_n \left\{ \frac{\langle \psi^0 | \hat{\mu}_i | \psi^n \rangle \langle \psi^n | \hat{\mu}_j | \psi^0 \rangle}{E_0 - E_n} \right\}$$

2.4 Anàlisi de correlació estadística sobre conjunts de punts

Tant la regressió per components principals (Principal Components Analysis, PCA) (Wold, 1966) com la regressió per mínims quadrats parcials (Partial Least Square, PLS) (Kramer, 1998; Haenlein, 2004; Rosipal, 2006) transformen les variables originals en noves variables ortogonals anomenades components o variables latents que permeten reduir la dimensionalitat de l'espai de les variables originals.

Donat un conjunt de dades consistent en una sèrie de característiques d'un conjunt d'objectes, per exemple el potencial electrostàtic d'un conjunt de molècules, es tracta de trobar una combinació de les variables originals del sistema, que permetin definir noves variables anomenades latents (scores) que revelen informació rellevant del problema que estem tractant. La determinació de les components principals és de fet una rotació dels sistema de coordenades. A l'anàlisi de components principals això es fa mitjançant la variància del conjunt. La variància porta informació potencial del objectes obtinguda per comparació. La primera component (PC1) és aquella que capta la màxima variància; la segona (PC2) és orthogonal a aquesta i captant la

màxima variància del sistema, i així successivament. Aquesta anàlisi permet tenir un coneixement més gran sobre les dades originals, trobar grups d'objectes similars, buscar objectes amb propietats diferents a la resta o buscar característiques rellevants dels objectes (Goldstein; 1989; Kramer, 1998; Shaw, 2003; Patra, 1999).

Suposem que tenim un conjunt de dades que es poden organitzar en forma de matriu X. Per exemple, suposem que disposem del MEP de varies molècules. Les dades es podrien organitzar de forma matricial col·locant a cada fila el valor del MEP per cadascuna de les molècules del conjunt. Per a cada fila (molècula) es pot calcular el valor mitjà i la variància:

$$m_j = \frac{1}{n} \sum x_{ij}$$

$$v_j = \frac{1}{n-1} \sum (x_{ij} - m_j)^2$$

on n és el nombre de punts del MEP. A partir de la matriu original X es construeix una matriu de distàncies de la mitjana i posteriorment una covariància -normalment respecte a la mitjana- que no és més que la multiplicació de $X^T X$:

$$c_{ij} = \frac{1}{n-1} \sum (x_{ij} - m_j)(x_{ik} - m_k)$$

La matriu de correlació té la variància com a elements de la diagonal i s_j , s_k són les desviacions típiques de les variables j i k respectivament. De vegades s'utilitza la matriu de correlació, que es calcula a partir de la covariància on els elements estan normalitzats i és independent de la forma que s'hagi escalat la matriu de covariància:

$$r_{jk} = \frac{c_{jk}}{s_j s_k}$$

A continuació la matriu de covariància es diagonalitza. Els vectors propis són les variables latents o *scores*, que són combinació lineal de las variables originals del sistema. El conjunt de coeficients que defineixen els vectors propis s'anomenen *loadings*. La dimensionalitat del sistema ve donada pel valor dels valors propis. Quan es posen els valors propis en ordre decreixent mostren un esglaó que ens permet determinar la dimensionalitat considerant els més grans abans de l'esglaó. Els valors propis estan relacionats amb la variància descrita del sistema.

La regressió per mínims quadrats parcials (PLS, Partial Least Squares) va ser introduïda per Herman Wold (Wold, 1973, 1978, 1983) per a ser aplicada a ciències econòmiques i socials, tot i que després ha estat aplicada a moltes altres branques científiques. La metodologia PLS generalitza i combina característiques de l'Anàlisi de Components Principals (PCA) i l'Anàlisi de Regressió Múltiple. L'anàlisi PLS és utilitzat quan volem correlacionar un elevat nombre de descriptors (variables independents), amb un determinat nombre de variables dependents. En el cas del PLS partim d'una matriu de dades característiques dels objectes X i una altra Y amb k propietats. Aquest dos espais es poden projectar de forma separada mitjançant PCA. En PLS es busca una projecció de X que estigui influenciada per Y i viceversa. D'aquesta manera la relació entre X i Y es pot extreure de forma més efectiva. Els *loadings* de l'espai X es calculen com vectors propis de la matriu $X^T Y Y^T X$ i els de l'espai Y dels vectors propis de $Y^T X X^T Y$. El criteri utilitzat es màxima variància en X i gran correlació amb Y.

En el present estudi els descriptors seran els valors del potencials electrostàtic, de polarització i de interacció de les nostres molècules, i les variables dependents seran els valors experimentals de l'activitat biològica dels conjunts de dioxines i furans a estudi. Els càlculs PLS ens permeten obtenir, en primer lloc, la representativitat dels nostres càlculs en la determinació de l'activitat biològica experimental, i en segon lloc, validar un model que ens permeti predir els valors experimentals de l'activitat biològica d'altres molècules similars al conjunt de dioxines i furans amb que hem treballat, introduint com a variables independents els nostres càlculs de potencials electrostàtic i de polarització.

La regressió PLS es basa en la transició d'un nombre elevat de descriptors inicial fins a una nova variable basada en un petit nombre de factors ortogonals (variables latents), és a dir els factors són combinacions lineals ortogonals dels descriptors originals. Les variables latents són elegides de manera que donin la màxima correlació amb la variable dependent. D'aquesta manera el model PLS conté el mínim nombre de components necessari per correlacionar els descriptors i la variable dependent. A la figura 2.4 es representen un conjunt de dades hipotètiques amb dues variables independents x_1 i x_2 , i variable dependent y. Es pot observar que les variables originals x_1 i x_2 presenten una correlació molt dolenta, de manera que s'obtenen dos variables latents (factors ortogonals) t_1 i t_2 que són combinacions lineals dels descriptors originals.

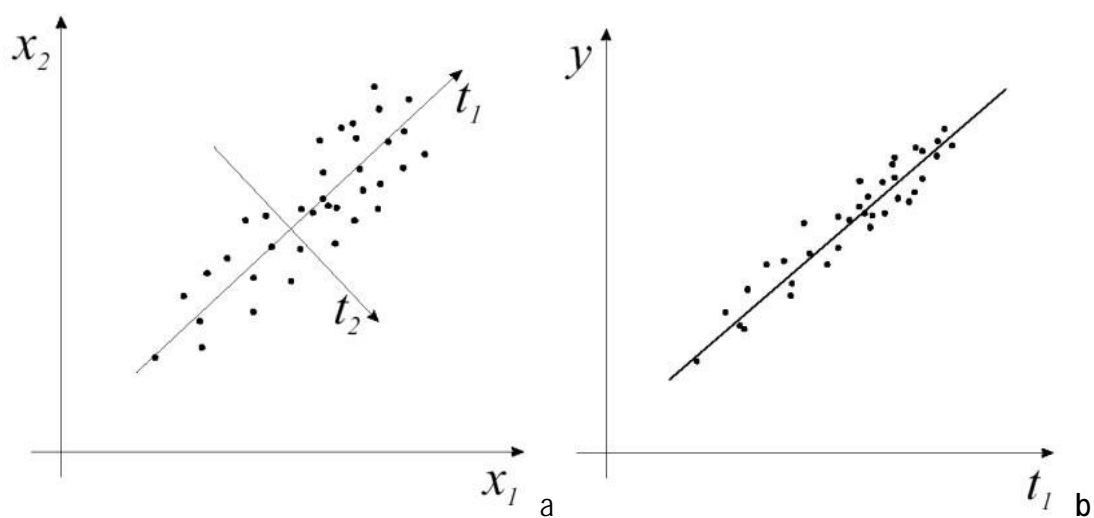


Fig. 2.4. (a) Transformació dels descriptors originals a variables latents; (b) construcció del model PLS

Els càlculs PCA i PLS presentats en aquest treball s'han fet utilitzant el programa SIMCA-P (Versió 10.2, Umetrics AB).

3 EFECTE DE LA BASE I DEL MÈTODE DE CÀLCUL A L'AVAUACIÓ DE LES PROPIETATS ELÈCTRIQUES DE PRIMER I SEGON ORDRE

En aquest capítol s'ha estudiat l'efecte del mètode i del conjunt de base al càlcul de les propietats elèctriques moleculars. En particular s'ha estudiat l'efecte sobre propietats de primer ordre com són els moments dipolar i quadrupolar i sobre propietats de segon ordre, com són la polaritzabilitat i la primera hiperpolaritzabilitat moleculars.

S'han triat per a l'estudi tres sistemes moleculars simples: el formaldehid, el cianur d'hidrogen i la urea . L'estudi s'ha fet a nivell Hartree-Fock i també tenint en compte l'efecte de la correlació electrònica mitjançant la teoria de pertorbacions de Møller-Plesset a diferents ordres: tant, a segon ordre MP2, com a quart ordre MP4. Així mateix s'han fet càlculs dins del marc de la teoria del funcional de la densitat (DFT) utilitzant els funcionals BLYP i B3LYP (veieu la secció de mètodes). Els càlculs s'han portat a terme utilitzant diversos conjunts de base a fi de comparar amb els resultats experimentals i altres prediccions teòriques publicades a la literatura. Les bases utilitzades han estat de tipus doble zeta standard a les quals s'han afegit funcions de polarització i funcions difuses. En particular, s'han utilitzat la base doble zeta 6-31G(d), les bases doble zeta augmentades amb un o dos conjunts de funcions de polarització 6-31G(d,p) i 6-31G(2d,2p) i també s'ha utilitzat la base 6-311++G(2d,2p), que inclou funcions difuses.

Tal i com s'ha explicat a la secció de mètodes, les propietats moleculars s'han calculat de forma analítica mitjançant el càlcul de les derivades de l'energia. Aquest procediment és equivalent al càlcul dels valors esperats en el cas dels moments dipolar i quadrupolar i al mètode de pertorbacions de segon ordre CPHF (coupled perturbed Hartree-Fock method) a nivell Hartree-Fock (Langhoff, 1965) o de doble pertorbació en el cas de la funció d'ona Møller-Plesset pel càlcul de la polaritzabilitat i hiperpolaritzabilitats moleculars.

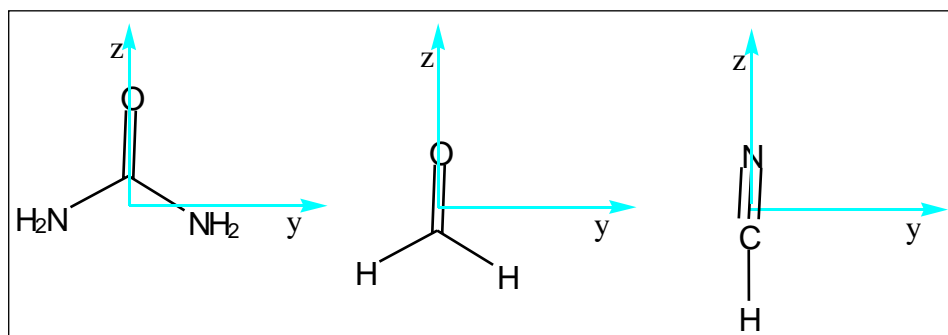


Figura 3.1.1. Orientació de les molècules als eixos Cartesianos.

Els càlculs s'han fet amb les molècules situades sobre el pla YZ, d'acord amb la convenció habitual: l'eix Z és el de més simetria, i si hi ha dos eixos de la mateixa simetria, Z és

el que té un nombre d'àtoms més gran. Si les molècules són planes X es l'eix perpendicular al pla de la molècula. En aquest cas els eixos han estat definits com es pot veure a la Figura 3.1.1.

Tots els càlculs van ser realitzats amb el programa Gaussian-94 i Gaussian 98 (Frisch, 1995, 1977, 1984, 1998). Els càlculs es van portar a terme sobre les geometries optimitzades per a cada base i metodologia emprada.

3.1 Càlcul de moments dipolar i quadrupolar

La importància de la metodologia i bases emprades en els càlculs de propietats elèctriques ha estat descrita àmpliament a la literatura (Hinchliffe, 1996, 1997, 1978, 2000a, 2000b; Masia, 2004; Ma, 2004; Haskopoulos, 2006; Risikrishna, 2006; Whitten, 2006).

A la Taula 3.1.1 es mostren els valors obtinguts per al *moment dipolar*. La inspecció de la Taula mostra que tant a nivells RHF, MP2 i MP4 el moment dipolar calculat està per sobre del valor experimental, encara que les diferències es fan més petites a mesura que la base es va fent més estesa. Com es pot veure, la tendència mostrada per les molècules objecte d'aquest estudi és la disminució sistemàtica del valor esperat del moment dipolar, donada una base, quan passem de nivell SCF a nivell MP2, i torna a disminuir a nivell MP4 apropant-se als resultats experimentals. Aquesta tendència ja s'havia observat en estudis previs (Maroulis, 1998a ; Alparone, 2005).

Per altre banda, si fixem el mètode de càlcul, en general els valors decreixen a mesura que la base es va fent més estesa. S'ha de tenir en compte que com a excepció, en el cas de la urea i a nivells MP2 i MP4, els valors augmenten lleugerament al passar de la base 6-31G(d) a la base 6-31G(d,p), encara que tornen a disminuir amb la base 6-31G(2d,2p), tornant a augmentar a mesura que la base es fa més estesa.

Aquestes tendències s'han observat amb les tres molècules estudiades, de manera que deduïm que els millors valors del moment dipolar s'han de calcular a nivell MP4 utilitzant una base que inclogui funcions difuses 6-311++G(2d,2p). S'ha de tenir en compte que en el cas de la urea els valors obtinguts a nivell MP4 amb aquesta base estan per sota del valor experimental publicat, fet que ens fa dubtar sobre aquesta dada experimental.

Taula 3.1.1. Càlculs dels moments dipolar. Valors en unitats atòmiques (μ , $1e a_0 = 8.478358 \times 10^{-30}$ C.m.). Valor experimental formaldehid: $\mu = -0.92$ (Weast, R. C., 1980; Applequist, 1972), Valor experimental HCN: $\mu = -1.17$ (Weast, R. C., 1980; Applequist, 1972); Valor experimental urea: $\mu = -1.80$ (Weast, R. C., 1980; Applequist, 1972).

	μ formaldehid	μ HCN	μ urea
6-31G(d)			
RHF	-1.09	-1.28	-1.86
MP2	-1.08	-1.26	-1.82
MP4	-1.07	-1.24	-1.81
BLYP	-0.79	-1.10	-1.62
B3LYP	-0.86	-1.14	-1.68
6-31G(d,p)			
RHF	-1.07	-1.28	-1.85
MP2	-1.06	-1.27	-1.80
MP4	-1.06	-1.25	-1.79
BLYP	-0.79	-1.09	-1.61
B3LYP	-0.84	-1.13	-1.67
6-31G(2d,2p)			
RHF	-1.01	-1.27	-1.79
MP2	-1.01	-1.25	-1.78
MP4	-0.99	-1.23	-1.74
BLYP	-0.77	-1.10	-1.57
B3LYP	-0.83	-1.14	-1.62
6311++G(2d,2p)			
RHF	-1.00	-1.26	-1.78
MP2	-0.96	-1.22	-1.76
MP4	-0.93	-1.18	-1.72
BLYP	-0.76	-1.09	-1.43
B3LYP	-0.81	-1.12	-1.60

Per altre banda, els càlculs portats a terme utilitzant la teoria del funcional de la densitat donen moments dipolar inferiors als experimentals, fet ja destacat anteriorment a la literatura (Jacquemin, 2004). Tanmateix, es troba també que el moment dipolar calculat és més gran utilitzant el funcional B3LYP que el BLYP.

Respecte al *moment quadrupolar*, a les Taules 3.1.2- 3.1.4 es presenten els resultats obtinguts en aquest treball per a les molècules de formaldehid, urea i cianur d'hidrogen, respectivament. Com es pot observar els càlculs SCF són els més aproximats als valors experimentals, constatant l'allunyament sistemàtic respecte als valors experimentals quan es treballa amb metodologia MP2, fet ja destacat en treballs previs (Maroulis, 1998a). De fet, no només la component zz és més petita quan s'inclou la correlació electrònica, si no que els valors de les components xx i yy sembla que inverteixen el signe.

Taula 3.1.2. Molècula de formaldehid. Càlculs del moment quadrupolar per diferents metodologies i bases. Valors en unitats atòmiques. ($1ea_0^2 = 4.486554 \times 10^{-40} C \cdot m^2$).

Valor experimental: ($\Theta_{zz}=0.245$; $\Theta_{xx}=-0.200$; $\Theta_{yy}=-0.045$) (Kukolich, 1971)

		SCF	MP2	MP4	BLYP	B3LYP
6-31G(d)	Θ_{zz}	0.236	0.134	0.150	0.176	0.179
	Θ_{xx}	-0.048	-0.160	-0.170	-0.196	-0.199
	Θ_{yy}	-0.188	0.026	0.020	0.020	0.020
6-31G(d,p)	Θ_{zz}	0.238	0.137	0.156	0.183	0.180
	Θ_{xx}	-0.051	-0.167	-0.181	-0.201	-0.201
	Θ_{yy}	-0.187	0.030	0.025	0.029	0.021
631G(2d,2p)	Θ_{zz}	0.239	0.138	0.172	0.194	0.196
	Θ_{xx}	-0.054	-0.169	-0.196	-0.261	-0.285
	Θ_{yy}	-0.185	0.031	0.022	0.067	0.089
6311++G(2d,2p)	Θ_{zz}	0.244	0.141	0.179	0.192	0.201
	Θ_{xx}	-0.163	-0.174	-0.198	-0.258	-0.299
	Θ_{yy}	-0.181	0.033	0.019	0.066	0.098

Si ens fixem en els càlculs a nivell Hartree-Fock i analitzem les tendències de variació del moment quadrupolar amb la base, es pot observar que la variació és la mateixa que s'ha trobat per al moment dipolar: Com es pot veure, amb la base 6-311++G(2d,2p) i metodologia Hartree-Fock s'aconsegueixen valors força propers als valors experimentals. Per altre banda, cal destacar els resultats obtinguts amb aquesta base i metodologia B3LYP per sota dels valors que s'obtenen a nivell Hartree-Fock, però millors que els que donen els càlculs amb correlació electrònica.

Taula 3.1.3. Molècula de urea. Càlculs del moment quadrupolar per diferents metodologies i bases. Valors en unitats atòmiques. ($1ea_0^2 = 4.486554 \times 10^{-40} C \cdot m^2$).

Valor experimental: ($\Theta_{zz}=-0.948$; $\Theta_{xx}=1.228$; $\Theta_{yy}=-0.721$) (Gaumann, 1958)

		SCF	MP2	MP4	BLYP	B3LYP
6-31G(d)	Θ_{zz}	-0.971	-0.624	-0.656	-0.715	-0.711
	Θ_{xx}	-0.477	-0.685	-0.640	-0.597	-0.592
	Θ_{yy}	1.448	1.309	1.296	1.312	1.303
6-31G(d,p)	Θ_{zz}	-0.990	-0.772	-0.781	-0.791	-0.782
	Θ_{xx}	-0.530	-0.657	-0.651	-0.657	-0.630
	Θ_{yy}	1.520	1.429	1.432	1.448	1.412
631G(2d,2p)	Θ_{zz}	-1.107	-0.991	-0.994	-1.002	-0.996
	Θ_{xx}	-0.507	-0.510	-0.503	-0.524	-0.472
	Θ_{yy}	1.614	1.501	1.497	1.526	1.468
6311++G(2d,2p)	Θ_{zz}	-1.241	-1.076	-1.082	-1.087	-1.114
	Θ_{xx}	-0.572	-0.560	-0.603	-0.546	-0.517
	Θ_{yy}	1.813	1.636	1.685	1.633	1.631

Resulta interessant comparar els valors obtinguts en aquest treball amb altres càlculs publicats a la literatura. En el cas del formaldehid, càlculs fet amb una base doble zeta a nivell

SCF donen valors per sobre dels experimentals ($\Theta_{zz}=0.453$; $\Theta_{xx}=-0.026$; $\Theta_{yy}=-0.427$) (Dunning, 1968). Al mateix nivell de càlcul, la utilització de bases més esteses milloren els resultats: ($\Theta_{zz}=0.190$; $\Theta_{xx}=-0.004$; $\Theta_{yy}=-0.186$) amb una base doble zeta estesa (Neuman, 1969); ($\Theta_{zz}=0.216$; $\Theta_{xx}=-0.010$; $\Theta_{yy}=-0.119$) utilitzant la base 13.10.4/7.5.2 de Sadlej (Bündgen, 1995) o ($\Theta_{zz}=0.249$; $\Theta_{xx}=-0.056$; $\Theta_{yy}=-0.193$) utilitzant una base 6-311++G(2df,2pd) (Bruna, 1997).

La utilització de funcions d'ona que incloquin la correlació electrònica no millora els resultats. Així per exemple, un càlcul CASSCF utilitzant una base estesa dóna valors semblants al càlcul SCF: ($\Theta_{zz}=0.2$; $\Theta_{xx}=-0.1$; $\Theta_{yy}=-0.1$) (Merchan, 1995). En canvi, el càlcul de la correlació electrònica a nivell MP2 empitjora els resultats obtinguts a nivell SCF, tal i com es mostra al càlcul amb la base 6-311++G(2df,2pd) que dóna valors: ($\Theta_{zz}=0.170$; $\Theta_{xx}=-0.034$; $\Theta_{yy}=-0.136$). Resultat que reproduïm en els nostres càlculs de la Taula 3.1.2.

En el cas de la urea, novament s'observen les mateixes tendències. Amb una base doble zeta s'obtenen valors per sobre dels experimentals: ($\Theta_{zz}=-0.98$; $\Theta_{xx}=0.49$; $\Theta_{yy}=0.46$) (Spackman, 1999); o amb una base 6-21G^{**}: $\Theta_{zz}=-0.93$; $\Theta_{xx}=0.71$; $\Theta_{yy}=0.23$ (Dovesi, 1990); o amb una base 6-31G: $\Theta_{zz}=-1.38$; $\Theta_{xx}=0.23$; $\Theta_{yy}=0.46$ (Duffy, 1993).

Taula 3.1.4. Molècula de HCN. Càlculs del moment quadrupolar per diferents metodologies i bases. Valors en unitats atòmiques. ($1e a_0^2 = 4.486554 \times 10^{40} \text{ C}\cdot\text{m}^2$).
Valor experimental: ($\Theta_{zz}=-2.30$) (Hartford, 1973)

		SCF	MP2	MP4	BLYP	B3LYP
6-31G(d)	Θ_{zz}	2.12	1.73	1.97	2.09	2.10
	Θ_{xx}	2.56	2.61	2.57	2.52	2.51
	Θ_{yy}	2.56	2.61	2.57	2.52	2.51
6-31G(d,p)	Θ_{zz}	2.14	1.79	2.01	2.13	2.11
	Θ_{xx}	2.57	2.62	2.58	2.55	2.54
	Θ_{yy}	2.57	2.62	2.58	2.55	2.54
631G(2d,2p)	Θ_{zz}	2.19	1.81	2.15	2.26	2.33
	Θ_{xx}	2.58	2.64	2.61	2.56	2.55
	Θ_{yy}	2.58	2.64	2.61	2.56	2.55
6311++G(2d,2p)	Θ_{zz}	2.28	1.90	2.17	2.27	2.31
	Θ_{xx}	2.66	2.70	2.68	2.67	2.65
	Θ_{yy}	2.66	2.70	2.68	2.67	2.65

En el cas del cianur d'hidrogen càlculs a nivell SCF amb bases poc esteses donen valors pobres respecte als experimentals. Així doncs Huzinaga (1969) amb una base 4-31G va publicar un valor de 1.44 a.u. Uns anys més tard, utilitzant bases del tipus d'una doble zeta Gready (1978) va publicar una valor de 1.51 a.u. i un any més tard Tyrell (1979) 1.58 a.u., per

sota del valor experimental. Aquestes diferències respecte als valors que es publiquen en aquest treball poden ser degudes a la geometria utilitzada per fer els càlculs, ja que el moment quadrupolar és molt sensible. Novament la utilització de bases més esteses apropen considerablement els resultats teòrics als experimentals. Així doncs Maroulis (1998b) utilitzant una base estesa va trobar un valor pel moment quadrupolar del cianur d'hidrogen de 2.06 a.u. a nivell SCF i molt més petit quan s'inclou la correlació electrònica: 1.60 a.u. a nivell MP2 i 1.70 a nivell MP4.

3.2 Càlculs de polaritzabilitat i primera hiperpolaritzabilitat moleculars

La polaritzabilitat i primera hiperpolaritzabilitat moleculars han estat calculades mitjançant les derivades analítiques del moment dipolar, com s'ha descrit anteriorment a l'apartat dels mètodes. A les Taules 3.2.1-3.2.3 es presenten els resultats obtinguts en el nostre treball corresponents a les components del tensor polaritzabilitat i a les Taules 3.2.4-3.2.6 els resultats de les components corresponents a la primera hiperpolaritzabilitat dipolar.

En quant a la polaritzabilitat molecular, els valors calculats sempre surten per sota dels resultats experimentals i la inclusió de la correlació electrònica millora els resultats obtinguts a nivell SCF. La inclusió de funcions difuses és decisiva, i fa que els valors variïn considerablement i siguin molt similars als experimentals. Els càlculs a nivell DFT amb bases esteses sobreestimen els valors experimentals. Aquests resultats confirmen afirmacions fetes per altres autors prèviament (Hinchliffe, 2000a, 200b; Whitten, 2000; Masia, 2004; Ma, 2004; Haskopoulos, 2006; Risikrishna, 2006; Benkova, 2007).

Com podem observar a les Taules 3.2.4-3.2.6 i comparant amb les dades experimentals, la inclusió de un segon conjunt de funcions de polarització millora tots els components del tensor en aproximadament un 10%, i la utilització de funcions difuses ens dona uns increments d'un 30% respecte a la base 6-31G(2d,2p), aconseguint uns valors força aproximats als valors experimentals (Landolt-Bornstein, 1951; Applequist, 1972).

La polaritzabilitat del formaldehid ha estat àmpliament estudiada a la literatura. A nivell SCF ha estat estudiada amb diverses bases. Els valors obtinguts per diversos autors a nivell SCF estan d'acord amb els obtinguts en aquest treball (Bruna, 1997). Càlculs realitzats per Rice et al. (Rice, 1990) ($\alpha_{xx}=12,42$; $\alpha_{yy}=16,23$; $\alpha_{zz}=21,19$) són propers al límit Hartree-Fock. Aquests resultats milloren a nivell MP2 un 5% en promig ($\alpha_{xx}=12,90$; $\alpha_{yy}=17,70$; $\alpha_{zz}=21,96$). Es pot comprovar que aquests resultats estan d'acord amb els resultats de la Taula 3.2.1.

Taula 3.2.1. Resultats formaldehid.

 Valors en unitats atòmiques ($1e^2 a_0^2 E_h^{-1} = 1.648778 \cdot 10^{41} C^2 \cdot m^2 \cdot J^{-1}$).

 Valor experimental: $\langle \alpha \rangle = 18.9$ (Hurdís, 1943)

		6-31G(d)	6-31G(d,p)	6-31G(2d,2p)	6-311++G(2d,2p)
RHF	xx	6.48	6.77	11.51	12.13
	yy	12.92	13.04	15.23	15.23
	zz	17.64	17.81	20.16	21.51
	$\langle \alpha \rangle$	12.35	12.54	15.63	16.29
MP2	xx	6.45	6.65	11.84	12.36
	yy	13.56	13.66	15.36	15.24
	zz	17.40	17.52	20.62	21.56
	$\langle \alpha \rangle$	12.48	12.61	15.94	16.39
MP4	xx	7.50	6.92	11.99	12.44
	yy	13.71	14.02	15.62	16.54
	zz	17.42	17.62	20.41	22.09
	$\langle \alpha \rangle$	12.88	12.85	16.01	17.02
BLYP	xx	6.76	6.96	12.26	12.68
	yy	14.16	14.31	16.12	17.90
	zz	17.95	18.09	21.47	23.58
	$\langle \alpha \rangle$	12.96	13.12	16.62	18.05
B3LYP	xx	6.69	6.89	12.13	12.57
	yy	13.74	13.84	16.07	16.96
	zz	17.74	17.69	21.28	22.72
	$\langle \alpha \rangle$	12.72	12.89	16.49	17.42

Referent als resultats obtinguts amb la base 6-311++G(2d, 2p) a nivell DFT, en aquest cas els resultats estan sobreestimats al voltant d'un 5-10%, fet que ja s'havia observat en càlculs semblants (Sim, 1992; Colwell, 1993).

L'anàlisi dels resultats de la urea mostra les mateixes tendències. Els resultats obtinguts a nivell SCF amb la base que inclou funcions difuses dona valors similars a la base doble zeta de Dunnig 9s5p/3s augmentada amb funcions de polarització i funcions difuses: ($\alpha_{xx}=24.3$; $\alpha_{yy}=35.0$; $\alpha_{zz}=36.4$) (Pérez, 1991).

Taula 3.2.2. Resultats urea.

 Valors en unitats atòmiques ($1e^2 a_0^2 E_h^{-1} = 1.648778 \cdot 10^{41} C^2 \cdot m^2 \cdot J^{-1}$).

 Valor experimental: $\langle \alpha \rangle = 31.9$ (Applequist, 1972)

		6-31G(d)	6-31G(d,p)	6-31G(2d,2p)	6-311++G(2d,2p)
RHF	xx	11.30	12.16	21.20	22.61
	yy	28.12	28.66	33.82	32.47
	zz	30.90	31.42	34.91	35.37
	$\langle \alpha \rangle$	23.44	24.08	29.98	30.15
MP2	xx	12.06	12.31	21.62	25.14
	yy	29.11	29.52	33.93	33.45
	zz	30.84	31.42	34.99	36.72
	$\langle \alpha \rangle$	24.01	24.36	30.18	31.77
MP4	xx	12.16	12.42	21.86	25.17
	yy	30.24	30.67	34.03	34.50
	zz	30.90	31.82	35.21	36.98
	$\langle \alpha \rangle$	24.41	24.97	30.37	32.22
BLYP	xx	12.25	12.51	21.92	26.45
	yy	31.89	32.30	34.21	40.77
	zz	32.15	32.49	35.68	42.10
	$\langle \alpha \rangle$	25.43	25.77	30.60	36.44
B3LYP	xx	12.16	12.41	21.89	25.13
	yy	30.26	30.67	34.15	38.01
	zz	31.05	31.40	35.48	39.91
	$\langle \alpha \rangle$	24.49	24.83	30.51	34.35

La utilització de la base POL (Sadlej, 1991) especialment dissenyada per al càlcul de polaritzabilitats elèctriques dona valors més propers als experimentals: $\alpha_{xx}=24.4$; $\alpha_{yy}=35.6$; $\alpha_{zz}=37.2$ (Reis, 1998). La inclusió de la correlació electrònica utilitzant la mateixa base millora els resultats un 10-15%: $\alpha_{xx}=27.7$; $\alpha_{yy}=41.3$; $\alpha_{zz}=43.0$. Finalment, observem en aquest treball (Taula 3.2.3) que els càlculs a nivell DFT amb bases esteses sobreestimen els valors experimentals al voltant del 5-10% (Benkova, 2007).

En quant al cianur d'hidrogen hi ha un càlcul de referència a la literatura a nivell Hartree-Fock per a la component zz fet per McLean i Yoshimine (1967) realitzat amb funcions de Slater, amb un valor de 22.4 a.u. Utilitzant una base doble zeta amb funcions de polarització Maroulis (1998b) va obtenir valors per a les components del tensor: $\alpha_{xx}=13.8$; $\alpha_{zz}=22.5$ a nivell Hartree-Fock; $\alpha_{xx}=13.9$; $\alpha_{zz}=21.9$ a nivell MP2 i $\alpha_{xx}=13.9$; $\alpha_{zz}=22.4$ a nivell MP4.

Taula 3.2.3. Resultats HCN.Valors en unitats atòmiques ($1e^2 a_0^2 E_h^{-1} = 1.648778 \cdot 10^{41} C^2 \cdot m^2 \cdot J^{-1}$).Valor experimental: $\langle \alpha \rangle = 17.5$ (Landolt-Bornstein, 1951)

		6-31G(d)	6-31G(d,p)	6-31G(2d,2p)	6-311++G(2d,2p)
RHF	xx	7.47	7.53	12.11	13.80
	yy	7.47	7.53	12.11	13.80
	zz	19.52	19.64	20.87	22.44
	$\langle \alpha \rangle$	10.48	11.56	15.03	16.68
MP2	xx	7.49	7.55	12.26	13.84
	yy	7.49	7.55	12.26	13.84
	zz	18.28	18.41	20.99	22.53
	$\langle \alpha \rangle$	11.09	11.17	15.17	16.74
MP4	xx	7.62	7.70	12.43	13.98
	yy	7.62	7.70	12.43	13.98
	zz	18.81	18.89	21.06	23.70
	$\langle \alpha \rangle$	11.34	11.43	15.31	17.22
BLYP	xx	7.73	7.79	12.57	14.91
	yy	7.73	7.79	12.57	14.91
	zz	19.22	19.36	21.24	24.46
	$\langle \alpha \rangle$	11.56	11.64	15.46	18.09
B3LYP	xx	7.66	7.71	12.50	14.76
	yy	7.66	7.71	12.50	14.76
	zz	19.15	19.28	21.16	24.01
	$\langle \alpha \rangle$	11.49	11.57	15.39	17.84

Valors més recents calculats amb bases de polarització (POL) donen valors de $\alpha_{xx}=13,89$; $\alpha_{zz}=22,43$ a.u. (Baranowska, 2010b). Sorprenentment, en aquest cas la inclusió de la correlació electrònica disminueix la component axial al voltant de un 2%: $\alpha_{xx}=13,92$; $\alpha_{zz}=21,93$ a.u. Resultats que estan d'acord amb els obtinguts en aquest treball amb la base augmentada amb funcions difuses, tal i com es pot veure a la Taula 3.2.2. Càlculs fets amb la metodologia de *coupled cluster* milloren els resultats obtinguts: $\alpha_{xx}=13,97$; $\alpha_{zz}=22,40$ a.u. (Baranowska, 2010a). Els resultats obtinguts en aquest treball han estat publicats recentment (Roset, 2009)

L'anàlisi dels resultats de les Taules 3.2.4-3.2.6 mostra els valors obtinguts a aquest treball per a la **hiperpolaritzabilitat dipolar**. És un fet destacable que la inclusió de la correlació electrònica afecta de forma desigual les diferents components del tensor. Com es pot veure les components β_{xxz} i β_{zzz} es fan més petites, mentre que la component β_{yyz} es fa més gran. També queda palès que la utilització de bases esteses que incloguin funcions de polarització és fonamental per obtenir valors propers als experimentals. Com en el cas de la polaritzabilitat molecular els millors resultats s'obtenen amb els càlculs a nivell MP4 i amb la base que inclou funcions difuses, 6-311++G(2d,2p).

Respecte a la molècula de formaldehid, no hi ha valor experimental amb el qual comparar. En els resultats calculats en aquest treball els valors obtinguts a nivell Hartree-Fock amb la base 6-311++G(2d,2p) són semblants als millors publicats a la literatura (Dixon, 1994;

Bruna, 1997; Alparone, 2005; Benkova, 2007). Així per exemple, Alparone (2005) amb bases de polarització especialment dissenyades pel càlcul de les propietats elèctriques moleculars (Pluta, 1998): amb la base POL s'obté $\beta_z = -45.3$ i amb la base HYPPOL obté $\beta_z = -45.5$ comparables als resultats de la Taula 3.2.4, on $\beta_z = 3/5(\beta_{xxz} + \beta_{yyz} + \beta_{zzz})$. Càlculs més recents amb bases esteses que inclouen funcions difuses donen a nivell Hartree-Fock valors semblants als obtinguts a aquest treball amb la base augmentada amb funcions difuses: $\beta_{xxz} = 10.0$; $\beta_{yyz} = 34.5$; $\beta_{zzz} = 38.6$ (Benkova, 2007).

La inclusió de la correlació electrònica a nivell MP2 fa disminuir les components β_{xxz} i β_{zzz} mentre que la component β_{yyz} augmenta respecte als valors obtinguts a nivell Hartree-Fock. Així doncs, els valors corresponents a les components obtinguts en aquest treball han estat $\beta_{xxz} = 7.3$; $\beta_{yyz} = 49.1$; $\beta_{zzz} = 20.9$, comparables als obtinguts per Benkova (2007) amb la base més estesa: $\beta_{xxz} = 7.9$; $\beta_{yyz} = 50.7$; $\beta_{zzz} = 21.6$. Com a conseqüència el valor mitjà β_z disminueix respecte al càlcul Hartree-Fock. Si ara ens fixem en el càlcul MP4 observem que totes les components augmenten lleugerament donant un valor mitjà pròxim al càlcul Hartree-Fock. Si milloréssim el càlcul amb correlació veuríem que en qualsevol cas hi ha un augment del valor mitjà respecte al càlcul Hartree-Fock, tal i com es pot comprovar fent càlculs Coupled Cluster refinats (Alparone, 2005; Benkova, 2007). Respecte als càlculs mitjançant la teoria del funcional de la densitat, aquest sobreestimen els càlculs fets amb la inclusió de la correlació electrònica al voltant del 20%.

Taula 3.2.4. Valors (en a.u.) de les components del tensor d'hiperpolaritzabilitat per a la molècula de formaldehid amb les diferents bases i mètodes objecte del present estudi. ($e^3 a_0^3 E_h^{-2} = 3.206361 \times 10^{53} C^3 \cdot m^3 \cdot J^{-2}$).

	β_{xxz}	β_{yyz}	β_{zzz}	$\beta_z = 3/5(\beta_{xxz} + \beta_{yyz} + \beta_{zzz})$
6-31G(d)				
RHF	7.26	29.81	34.31	43.08
MP2	6.04	45.52	17.81	41.80
MP4	6.26	45.64	17.98	69.68
BLYP	8.13	51.82	36.32	41.81
B3LYP	8.26	55.81	39.20	61.96
6-31G(d,p)				
RHF	8.04	30.26	35.61	44.35
MP2	6.83	47.42	18.21	43.48
MP4	6.91	47.61	18.83	44.01
BLYP	8.87	52.26	37.41	59.12
B3LYP	8.91	56.14	39.66	62.83
6-31G(2d,2p)				
RHF	9.14	31.89	37.61	47.18
MP2	7.05	48.67	19.23	44.97
MP4	7.13	49.92	21.16	46.93
BLYP	9.06	54.16	38.99	61.33
B3LYP	9.34	57.28	41.09	64.63
6-311++G(2d,2p)				
RHF	9.81	33.26	38.41	49.10
MP2	7.26	49.06	20.86	46.31
MP4	7.54	50.61	22.03	48.11
BLYP	9.14	55.23	40.71	63.05
B3LYP	9.92	58.74	42.26	66.55

Taula 3.2.5. Hiperpolaritzabilitats per les bases 6-31G(d), 6-31G(d,p), 6-31G(2d,2p) i 6-311++G(2d,2p), per la molècula de urea.

Resultats en unitats atòmiques ($e^3 a_0^3 E_h^{-2} = 3.206361 \times 10^{53} C^3 \cdot m^3 \cdot J^{-2}$)

	β_{xyy}	β_{yyy}	β_{zzy}	$\beta_z = 3/5(\beta_{\text{xyy}} + \beta_{\text{yyy}} + \beta_{\text{zzy}})$
6-31G(d)				
RHF	17.21	61.22	-41.11	22.39
MP2	30.71	65.21	-49.26	27.99
MP4	31.23	61.97	-49.44	26.26
BLYP	31.68	77.61	-28.13	48.69
B3LYP	32.49	87.06	-27.76	55.07
6-31G(d,p)				
RHF	18.06	62.34	-42.09	22.98
MP2	31.07	65.97	-50.31	28.04
MP4	31.72	62.99	-51.12	26.15
BLYP	31.94	78.93	-28.70	49.30
B3LYP	32.88	88.72	-28.33	55.96
6-31G(2d,2p)				
RHF	18.94	63.10	-43.21	23.29
MP2	32.16	66.72	-51.23	28.59
MP4	33.24	64.21	-52.68	26.86
BLYP	33.50	81.04	-29.97	50.85
B3LYP	35.11	91.21	-29.46	58.12
6-311++G(2d,2p)				
RHF	19.26	64.06	-44.42	23.34
MP2	33.73	67.49	-52.62	29.16
MP4	34.21	65.68	-53.03	28.12
BLYP	34.86	82.20	-30.84	51.73
B3LYP	36.02	96.82	-30.24	61.56

Dels resultats de la Taula 3.2.5 deduïm que, en incloure correlació electrònica, en conjunt tenim un increment del valor mitjà respecte als valors que s'obtenen utilitzant el mètode d'Hartree-Fock. Respecte als valors obtinguts amb el funcional de la densitat, surten molt desviats respecte als valors de referència del càlcul SCF. També es pot observar que el funcional B3LYP dóna valors una mica per sobre del BLYP.

El valor experimental de β_z és aproximadament de 52 ± 14 a.u. (en aigua) i 33.6 ± 3.5 a.u. (en DMSO) (Ledoux, 1982), molt per sobre dels valors obtinguts a nivell MP4 amb la base que inclou funcions difuses. De fet el solvent influeix de forma important en els valors de la hiperpolaritzabilitat i per tant la comparació no és molt informativa sobre la qualitat dels resultats (Bartkowiak, 2002). Més interessant és la comparació dels resultats d'aquest treball amb altres resultats publicats a la literatura. La molècula de urea ha estat força estudiada amb diferents funcions de base, tant a nivell SCF com amb metodologia perturbacional i funcional de la densitat (Pérez, 1991; Bartkowiak, 2002; Pluta, 2001; Alparone, 2005; Benkova, 2007).

Els resultats obtinguts tant a nivell Hartree-Fock com a nivell MP2 amb la base que inclou funcions difuses són comparables als publicats per Benkova (2007) utilitzant diverses bases esteses amb funcions de polarització. A nivell Hartree-Fock: $\beta_{xxz}=19.2$; $\beta_{yyz}=-44.1$; $\beta_{zzz}=64.1$ i a nivell MP2: $\beta_{xxz}=33.7$; $\beta_{yyz}=-54.6$; $\beta_{zzz}=68.2$. Resultats semblants van ser publicats per en Reis (1998). A nivell Hartree-Fock utilitzant la base POL, especialment dissenyada per al càlcul de propietats elèctriques de segon ordre (Sadlej, 1988): $\beta_{xxz}=17.0$; $\beta_{yyz}=-44.7$; $\beta_{zzz}=67.9$ i a nivell MP2: $\beta_{xxz}=31.7$; $\beta_{yyz}=-56.5$; $\beta_{zzz}=77.0$. Aquests resultats mostren la mateixa tendència que els resultats obtinguts en aquest treball, on s'observa l'increment important de la component xxz al incloure la correlació electrònica en els càlculs.

La inspecció de la Taula 3.2.6 suggereix que en aquest cas la inclusió de la correlació electrònica fa augmentar els valors de les diverses components. Per altre banda l'augment de l'extensió de la base fa disminuir els valors a nivell Hartree-Fock i augmentar a nivell MP2 i MP4. En aquest cas el mètode del funcional de la densitat dona valors molt distants dels de referència.

Els valors obtinguts en aquest treball presenten les mateixes tendències que els publicats per altres autors. Així per exemple, Maroulis (1998b) utilitzant una base doble zeta amb funcions de polarització obté valors de $\beta_{xxz}=-3.0$; $\beta_{zzz}=-6.6$ a nivell Hartree-Fock. Quan s'inclou la correlació electrònica els valors es fan positius, obtenint: a nivell MP2: $\beta_{xxz}=0.2$; $\beta_{zzz}=1.1$; i a nivell MP4: $\beta_{xxz}=-7.6$ i $\beta_{zzz}=0.0$. Com es pot veure els valors són semblants als obtinguts a aquest treball excepte pel cas dels càlculs a nivell MP4. Més recentment, Baranowska (2009) a través d'un procés d'optimització de funcions de base reporta valors de β_{zzz} que oscil·len entre -5.86 i -6.46 a nivell Hartree-Fock i que pugen fins 1.18-1.67 a nivell MP2.

Taula 3.2.6. Valors (en a.u.) de les components del tensor d'hiperpolaritzabilitat per a la molècula de HCN amb les diferents bases i mètodes objecte del present estudi. ($e^3 a_0^3 E_h^{-2} = 3.206361 \times 10^{53} C^3 \cdot m^3 \cdot J^{-2}$).

	β_{xxz}	β_{yyz}	β_{zzz}	$\beta_z = 3/5(\beta_{xxz} + \beta_{yyz} + \beta_{zzz})$
6-31G(d)				
RHF	-2.88	-5.81	-2.88	-6.94
MP2	0.40	1.47	0.40	1.36
MP4	0.46	1.49	0.46	1.45
BLYP	0.33	-3.22	0.33	-1.54
B3LYP	0.42	-3.70	0.42	-1.72
6-31G(d,p)				
RHF	-2.93	-6.04	-2.93	-7.14
MP2	0.47	1.53	0.47	1.48
MP4	0.52	1.57	0.52	1.57
BLYP	0.37	-3.46	0.37	-1.63
B3LYP	0.49	-3.91	0.49	-1.76
6-31G(2d,2p)				
RHF	-3.05	-6.21	-3.05	-7.39
MP2	0.59	1.62	0.59	1.68
MP4	0.61	1.63	0.61	1.71
BLYP	0.40	-3.58	0.40	-1.67
B3LYP	0.52	-3.98	0.52	-1.76
6-311++G(2d,2p)				
RHF	-3.11	-6.58	-3.11	-7.68
MP2	0.63	1.65	0.63	1.75
MP4	0.66	1.68	0.66	1.80
BLYP	0.41	-3.63	0.41	-1.69
B3LYP	0.54	-4.05	0.54	-1.78

4 EFECTE DE LA BASE I DEL MÈTODE DE CÀLCUL A L'AVAUACIÓ DELS POTENCIAL ELECTROSTÀTIC I DE POLARITZACIÓ MOLECULARS

4.1 Càlcul dels potencials electrostàtic i de polarització

L'energia d'interacció d'una molècula amb una càrrega externa 'q' es pot expressar com una expansió pertorbativa en termes de 'q'. Els coeficients de l'expansió són les contribucions de diferents ordres per l'energia d'interacció:

$$\Delta E = qE(1) + q^2E(2) + q^3E(3) + \dots$$

La contribució de primer ordre inclou el terme de Coulomb a l'energia d'interacció, i representa l'energia d'interacció amb una càrrega externa 'q'. Els termes de segon ordre i superiors inclouen les contribucions a l'energia degudes a la polarització de la distribució de càrregues per la presència de la càrrega exterior 'q'. Si 'q' és la càrrega unitària 'e', el primer terme és el potencial electrostàtic molecular, i si es considera que termes de tercer ordre i superiors són menyspreables, l'equació anterior es pot aproximar com :

$$\Delta E = eV(r) + e^2P(r)$$

on $V(r)$ és el potencial electrostàtic de la molècula i $P(r)$ el potencial de polarització molecular.

La importància de les diferents contribucions electrostàtiques ens permet portar a terme un estudi teòric de reactivitat i altres propietats moleculars a partir de càlculs dels potencials electrostàtic i de polarització moleculars. Tot i que el terme electrostàtic és el que té més pes a l'energia d'interacció, aquest terme només representa la interacció d'una distribució estàtica de càrregues amb un camp elèctric extern. El terme de polarització aporta una millora a vegades qualitativa de la descripció de les diferents interaccions moleculars.

L'estudi de la interacció d'una càrrega de prova es pot fer a diferents posicions relatives de la molècula amb el qual podem construir mapes d'interacció i de les seves contribucions: els mapes de potencial electrostàtic i els de polarització.

Els mapes de potencial electrostàtic molecular (MEPs) són representacions espacials de la contribució de primer ordre o terme de Coulomb a l'energia d'interacció entre una càrrega puntual positiva i una distribució continua i estàtica de càrregues (Scrocco, 1973; Bonaccorsi, 1976; Weinstein, 1975; Politzer, 1991; Alkorta, 1996; Dykstra, 1993). L'anàlisi dels MEPs permet tenir coneixement sobre diferents aspectes del reconeixement molecular com: la reactivitat molecular, les interaccions biològiques, els fenòmens de solvatació, de l'estat cristal·lí o dels estudis de densitat electrònica (Poltzer, 1991). Així mateix, la identificació dels mínims de

potencial permet tenir una idea de la "topologia" de la reactivitat de la molècula i de les interaccions amb altres molècules a llargues distàncies, de tal manera que es pot utilitzar com a descriptor molecular.

Els mapes de Potencial de Polarització Molecular (MPPs) (Bonaccorsi,1976; Francl, 1985; Alkorta, 1994, 1993, 1995, 1996) representen la contribució de segon ordre a la l'energia d'interacció d'una càrrega de prova a una molècula, i donen informació sobre la capacitat de la densitat electrònica a moure's quan és sotmesa a un camp elèctric. La polarització d'una molècula està relacionada amb les propietats elèctriques de polaritzabilitat i hiperpolaritzabilitat. Com que diferents 'zones' de la molècula tenen diferents capacitats de polarització, la 'polaritzabilitat molecular neta' depèn de la distribució de densitat electrònica a la molècula. Així, regions moleculars amb dobles enllaços, sistemes aromàtics, o parells solitaris presenten polaritzacions més elevades. Els MPPs es poden considerar com a una correcció als MEPs que milloren la descripció molecular a distàncies curtes, permetent distingir regions moleculars susceptibles d'un atac nucleofílic. Per altre banda, els MPPs ens poden donar una idea de la capacitat descriptiva dels diferents mètodes i bases per a descriure un sistema molecular.

Estudis previs han demostrat l'efecte del conjunt de base en el càlcul de la polarització molecular (Alkorta,1993), així com l'efecte de la metodologia emprada en el càlcul de polaritzacions moleculars (Hinchliffe, 2000a, 2000b; Ma, 2004; Soscun, 2000; 2002; Masia, 2005). L'anàlisi comparativa dels mapes de potencial de polarització molecular a través del càlcul de les diferències de polarització relatives i les desviacions estàndards corresponents, ens permetrà l'estudi comparatiu de les diferents metodologies i bases utilitzades en aquest treball, així com un estudi segons la proximitat o allunyament dels punts de la xarxa a la molècula pels diferents conjunts de base i mètodes.

L'objectiu d'aquest capítol és realitzar l'estudi del càlcul de la polarització molecular en funció del conjunt de base emprat i la metodologia per al càlcul de la densitat electrònica. Per aquest estudi s'han utilitzat les bases 6-31G(d), 6-31G(d,p), 6-31G(2d,2p), 6-311G++(2d,2p) i les metodologies de càlcul de la densitat electrònica: Hartree-Fock, Møller-Plesset de segon ordre (MP2), Møller-Plesset de quart ordre (MP4) i el funcional de la densitat utilitzant els funcionals BLYP i B3LYP. El potencial electrostàtic a cada punt s'ha calculat com valor esperat del operador potencial electrostàtic $\langle r^{-1} \rangle$. En canvi, els mapes de polarització, al ser una propietat de segon ordre, s'han de calcular de forma aproximada mitjançant dos procediments alternatius. Per una banda hem utilitzat el mètode de la supermolècula, on el MPP s'ha calculat com

diferència entre els mapes d'interacció i els de potencial electrostàtic i per altre, mitjançant la teoria de pertorbacions de segon ordre.

Tant el MEP com el MPP s'han calculat sobre una xarxa de punts a 4 bohr sobre el pla de la molècula. Els punts de la xarxa s'han calculat cada bohr i aproximadament amb una extensió de 6 bohr al voltant de la superfície de van Der Waals de les molècules. El MPP a cada punt s'ha calculat com la diferència entre l'energia d'interacció de la molècula amb una càrrega puntual positiva i el potencial electrostàtic en aquell punt. Per evitar que els moments d'ordre superior fossin importants, els càlculs de l'energia d'interacció s'han fet amb una càrrega de 0.1 unitats atòmiques i després s'han escalat a una càrrega unitària.

4.2 Anàlisi general dels mapes de polarització calculats al present estudi

A la figura 4.2.1 es mostren els mapes de polarització calculats amb la base que inclou funcions difuses a nivells MP4 i amb la teoria del funcional de la densitat amb el funcional BLYP. Com es pot veure els mapes qualitativament són semblants. Al contrari dels MEP, els MPP no presenten mínims i s'ajusten molt bé a la superfície de van Der Waals de la molècula.

Per poder fer una comparació *a grosso modo* entre el diferents mapes, es pot calcular el valor mitjà de les diferències mitjanes relatives entre dos mapes així com les desviacions estàndard. La diferència relativa a cada punt de la xarxa es calcula d'acord a la relació:

$$\varepsilon_i = \frac{(P(\text{base A}) - P(\text{base B})) \cdot 100}{P(\text{base A})}$$

on P és el valor de la polarització a cada punt de la xarxa.

L'anàlisi dels resultats suggereix que els MEPs (resultats en els mapes de l'annex 1) que s'obtenen són similars per a tots els nivells de teoria i base emprades, amb dispersions molt petites per les molècules de formaldehid i cianur d'hidrogen, i una mica més grans encara que insignificants per la urea. Aquest efecte negligible de la correlació electrònica al càlcul dels MEPs ja s'havia observat prèviament (Luque, 1991).

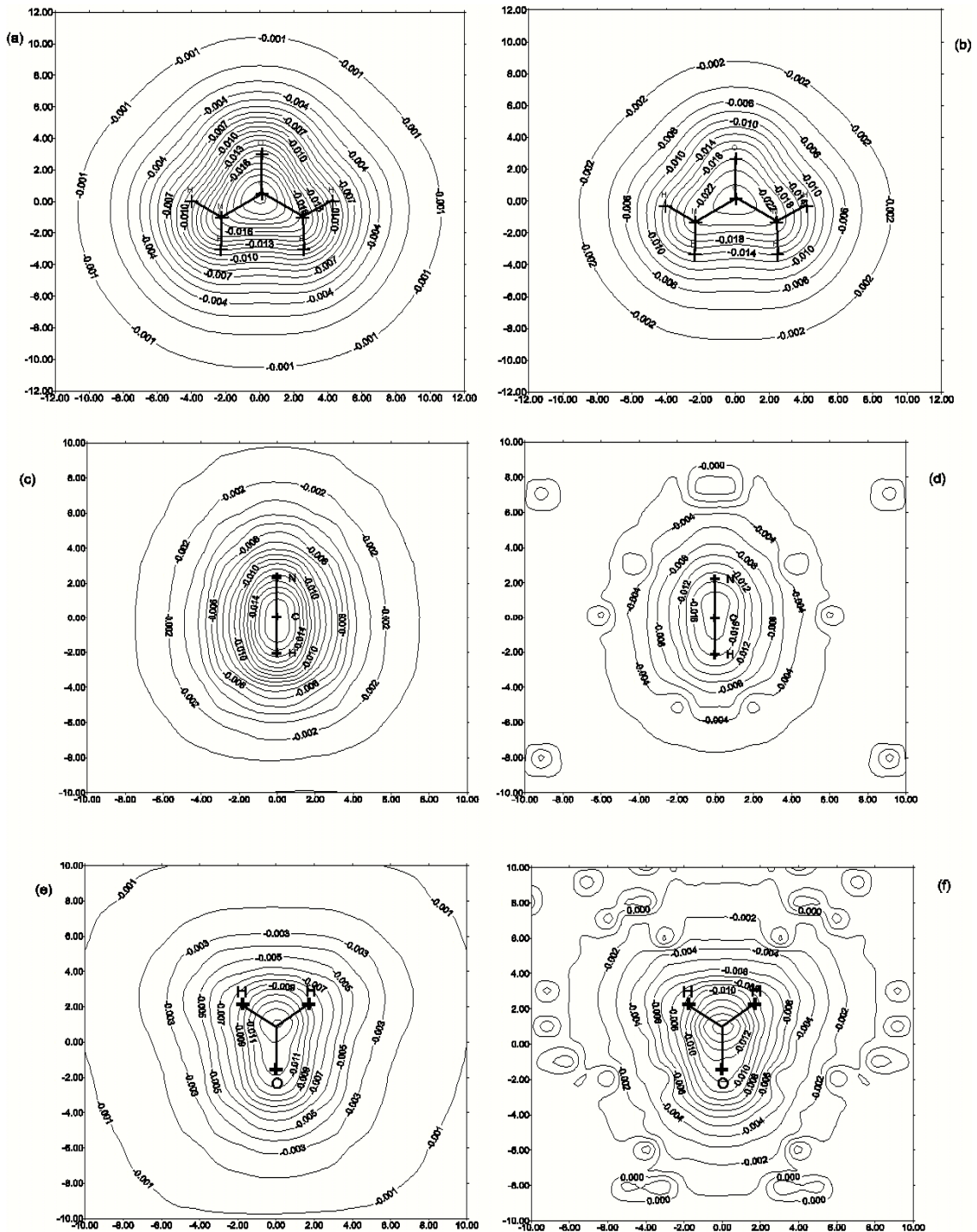


Figura 4.2.1. Mapes del potencial de polarització molecular (MPP). Base 6311++G(2D,2P). (a) Urea. MP4, (b) Urea. BLYP, (c) HCN. MP4, (d) HCN. BLYP, (e) Formaldehid. MP4, (f) Formaldehid. BLYP

A les Taules 4.2.1 - 4.2.3 es presenten els resultats de les diferències i desviacions mitjanes relatives referents a les diferents bases emprades prenent la base 6-31G(d) com a referència per a cadascuna de les molècules objecte d'aquest estudi. Cal destacar que pels

càlculs DFT (BLYP i B3LYP), en el cas del potencial de polarització hi ha punts on les diferències són inferiors a 10^{-4} , de manera que a aquest punts se'ls assigna un valor de zero.

Com es pot veure, a nivell Hartree-Fock l'efecte de l'augment de la base és d'un 5-9% en afegir funcions de polarització als àtoms d'hidrogen (base 6-31G(d,p)). De la mateixa manera al passar a una base amb dos conjunts de funcions de polarització les diferències arriben a ser de l'ordre del 10% i són del 15% quan s'afegeixen funcions de polarització a la base. En el cas dels càlculs a nivell MP2 i MP4 els resultats són semblants: en afegir funcions de polarització als hidrògens s'observen diferències del 10% que arriben fins al 15% amb la base amb funcions difuses. Al cas del funcional de la densitat l'augment no és molt gran en afegir la correlació electrònica, obtenint millores al voltant del 5%, semblants a les observades al càlcul de les polaritzabilitats moleculars.

En el cas dels càlculs fets amb el funcional de la densitat es pot veure que hi ha una millora més petita en augmentar la base, que arriba a ser del 5%. Ara si comparem els valors calculats amb el funcional de la densitat amb els diferents mètodes, veiem que els càlculs són un 15% amb la base menys estesa i de l'ordre d'un 5-10% per les més esteses.

Taula 4.2.1. Diferències i desviacions mitjanes relatives dels diferents MPP calculats pel formaldehid.

	SCF	MP2	MP4	BLYP	B3LYP
6-31G(d)	0.0	1.1±1.3	1.18±1.2	13.8±10.1	15.6±11.3
6-31G(d,p)	5.8±3.3	11.9±5.3	12.4±6.2	15.1±10.6	16.2±12.1
6-31G(2d,2p)	9.1±7.2	12.4±8.7	12.6±10.1	18.7±12.9	20.7±14.3
6311++G(2d,2p)	15±8.2	16.2±9.7	18±8.9	20.5±9.1	21.6±9.8

Taula 4.2.2. Diferències i desviacions mitjanes relatives dels diferents MPP calculats per la urea.

	SCF	MP2	MP4	BLYP	B3LYP
6-31G(d)	0.0	1.0±0.9	1.0±0.7	13.4±9.9	16.8±10.3
6-31G(d,p)	9.3±8.3	11.7±4.2	12.8±5.2	14.1±10.2	17.0±11.1
6-31G(2d,2p)	11.7±5.4	16.1±8.3	16.3±6.5	19.9±11.3	20.1±12.3
6311++G(2d,2p)	14±6.2	18±9.2	20±9.3	22.7±10.8	24.6±11.2

Taula 4.2.3. Diferències i desviacions mitjanes relatives dels diferents MPP calculats pel HCN.

	SCF	MP2	MP4	BLYP	B3LYP
6-31G(d)	0.0	0.4±1.1	0.7±0.6	12.6±11	14.3±11
6-31G(d,p)	5.2±3.1	11.1±9.8	12.4±9.6	14.2±10	16.7±12
6-31G(2d,2p)	8.6±6.2	11.4±5.9	16.3±7.2	19.5±13	21.2±14
6311++G(2d,2P)	14±7.1	18±6.2	20.4±6	21.6±8	22.3±9

4.3 Anàlisi comparativa dels valors del MEP i MPP calculats amb diferents bases i metodologies

Per poder comparar la bondat dels càlculs s'han portat a terme les correlacions entre els diferents punts calculats amb les diverses bases i mètodes de càlcul. En primer lloc es realitza l'anàlisi comparativa dels punts calculats mitjançant diferents mètodes de càlcul amb la base 6-311G++(2d,2p), prenent com referència el càlcul MP4.

Per a la molècula de formaldehid, els resultats de l'anàlisi del MEP es mostren a la Figura 4.3.1. Com es pot observar tant a nivell Hartree-Fock com MP2 els càlculs es poden considerar com equivalents als que s'obtenen a nivell MP4, tal i com hem comentat amb anterioritat. Si ara es comparen els càlculs a nivell MP4 amb els calculats utilitzant el mètode del funcional de la densitat, es pot veure que sistemàticament els valors del MP4 són més petits en valor absolut, tal i com es pot veure, ja que els pendents de la rectes de regressió tenen pendents més grans que la unitat i la correlació entre els valors dels càlculs és molt bona.

La correlació entre els resultats del MPP a nivell MP4 i la resta de metodologies es mostren a la Figura 4.3.2. Com es pot veure, a nivell Hartree-Fock i MP2 la linealitat és bona per a valors petits de la polarització, i s'observen desviaments de la linealitat per a punts de polaritzabilitat alta. Donat que hi ha més punts amb valors petits, la correlació obtinguda està lleugerament esbiaixada cap aquests, observant-se una bona correlació, que a nivell Hartree-Fock és de 0.96 amb un $r^2=0.98$ i a nivell MP2 és de 0.98 amb $r^2=0.98$. Aquest resultats evidencien unes diferències sistemàtiques entre 5-10% com ja hem comentat abans. Si ara comparem amb els càlculs proporcionats mitjançant la teoria del funcional de la densitat, s'observa que la correlació no és tan bona com abans i que el DFT no dona valors prou acurats quan el MPP és petit. Hi ha una bona correlació amb el funcional BLYP amb un pendent de 0.99 i un $r^2=0.97$ que no és tan bo amb el funcional B3LYP amb un pendent de 0.93 i un $r^2=0.96$.

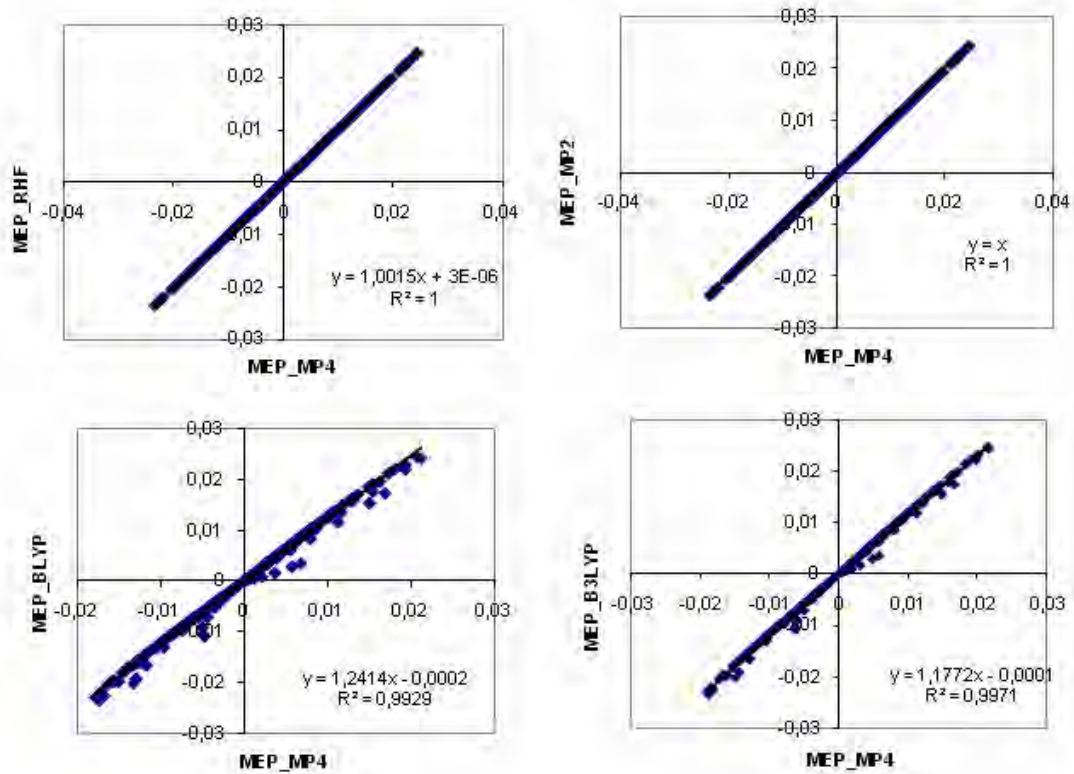


Figura 4.3.1. MEP del formaldehid. Anàlisi de correlació amb la base 6-311++G(2d,2p)

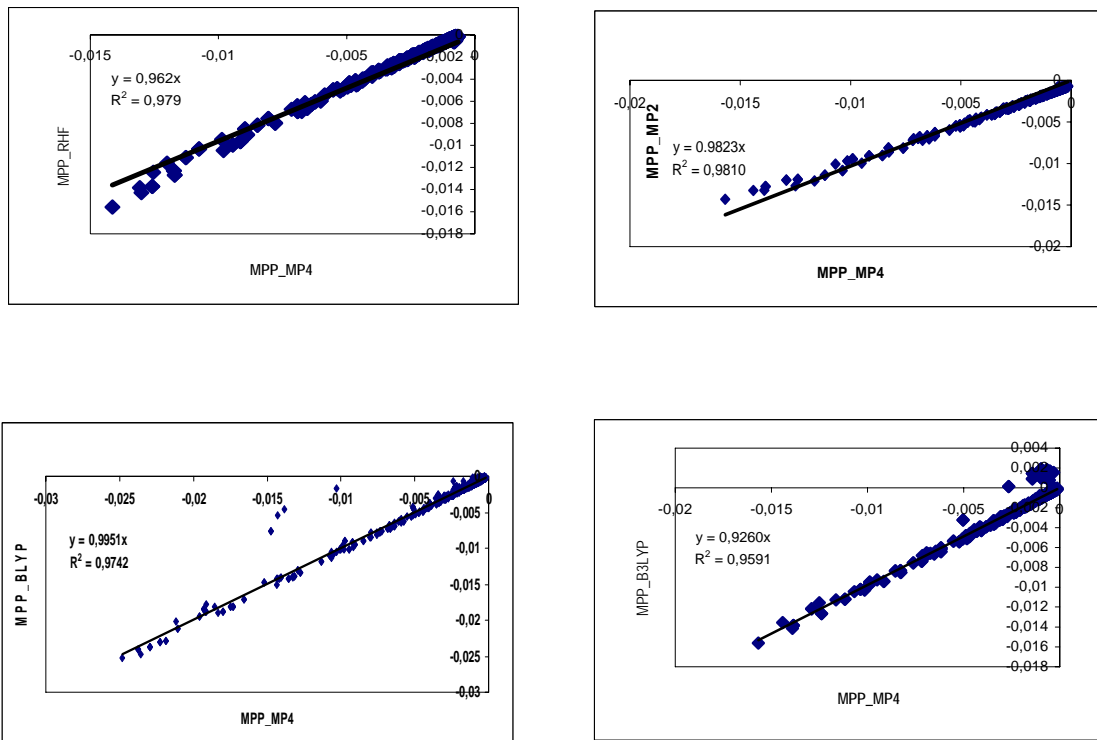


Figura 4.3.2. MPP del formaldehid. Anàlisi de correlació amb la base 6-311++G(2d,2p)

Per a la molècula d'urea, la correlació dels punts del MEP calculats amb diferents metodologies i els calculats a nivell MP4 es poden veure a la Figura 4.3.3. Com es pot veure els càlculs a nivell Hartree-Fock i MP2 són molt semblants als obtinguts a nivell MP4 al igual que havíem observat amb el formaldehid. També en aquest cas al comparar amb els càlculs obtinguts amb el funcional de la densitat s'observa que el DFT sobreestima els valors respecte als càlculs a nivell MP4.

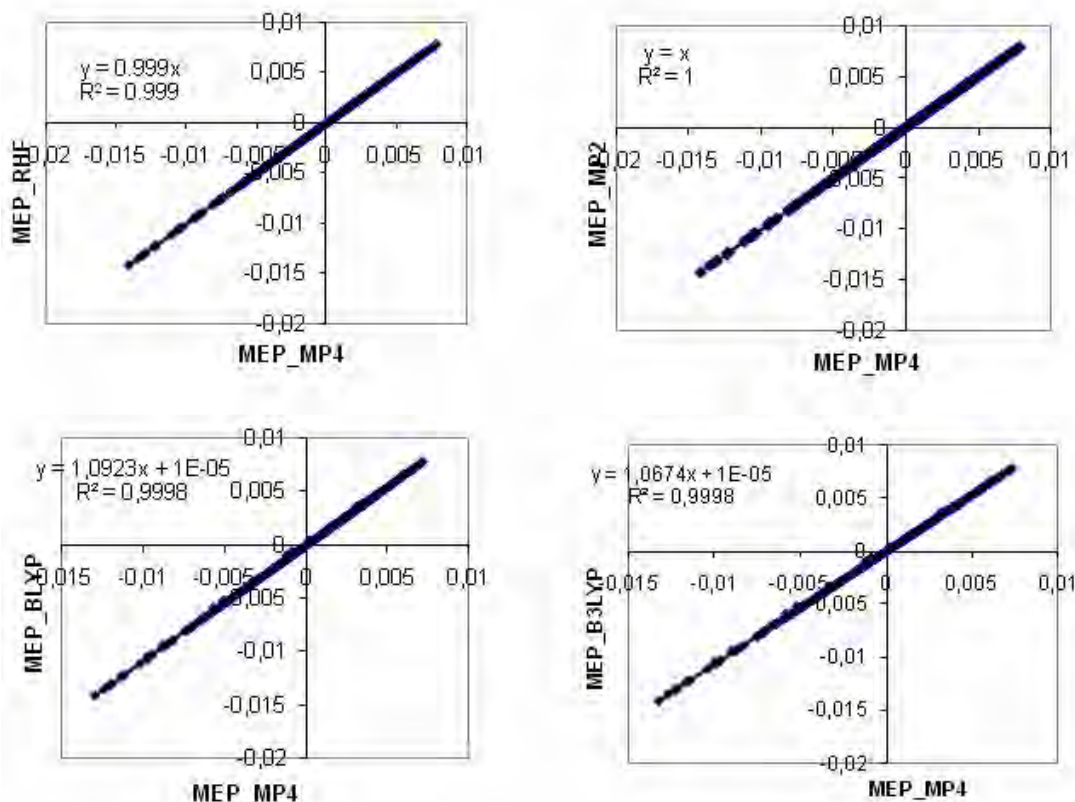


Figura 4.3.3. MEP de la urea. Anàlisi de correlació amb la base 6-311++G(2d,2p)

La comparació del valors del MPP es pot veure a la Figura 4.3.4. Com es pot veure els càlculs a nivell Hartree-Fock i MP2 són molt semblants als obtinguts a nivell MP4. En el cas dels càlculs del funcional de la densitat s'observa, tal i com ja ho havíem vist al cas del formaldehid, desviacions per a valors petits. Novament es pot veure que el càlcul amb el funcional BLYP resulta més semblant al realitzat a nivell MP4.

Finalment en el cas del cianur d'hidrogen la comparació de resultats corresponent al MEP es troba a la Figura 4.3.5. Novament els càlculs a nivell Hartree-Fock i MP2 són molt semblants als obtinguts a nivell MP4. En canvi novament veien que el funcional de la densitat sobreestima el valor del potencial electrostàtic.

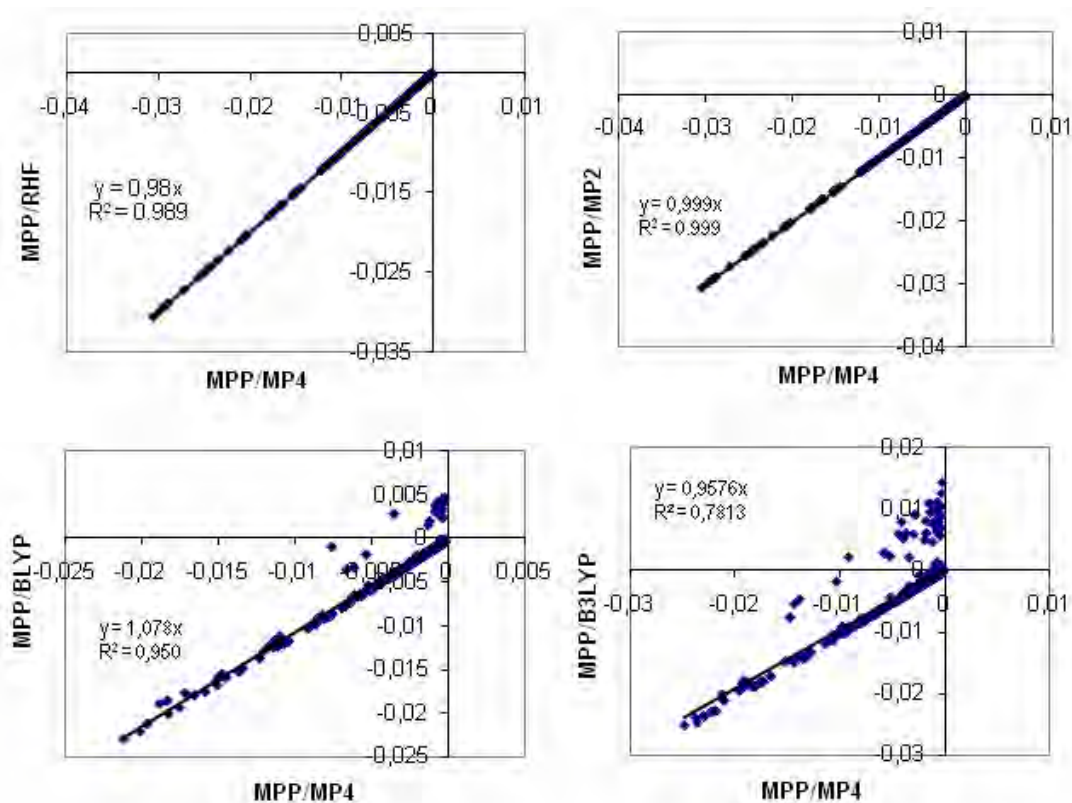


Figura 4.3.4. MPP de la urea . Anàlisi de correlació amb la base 6-311++G(2d,2p)

La comparació dels resultats pel MPP es troben a la Figura 4.3.6. Tal i com havíem vist en els casos anteriors la polarització a nivell Hartree-Fock i MP2 és més petita que la calculada a nivell MP4 però, la correlació es força bona: pels càlculs Hartree-Fock el coeficient de correlació és 0.91 i $r^2=0.99$; pels càlculs MP2 el coeficient és 0.95 i $r^2=0.99$. Quan comparem els càlculs del funcional de densitat veiem que hi ha molt punts que surten dels valors esperats, principalment per a valors petits de la polarització.

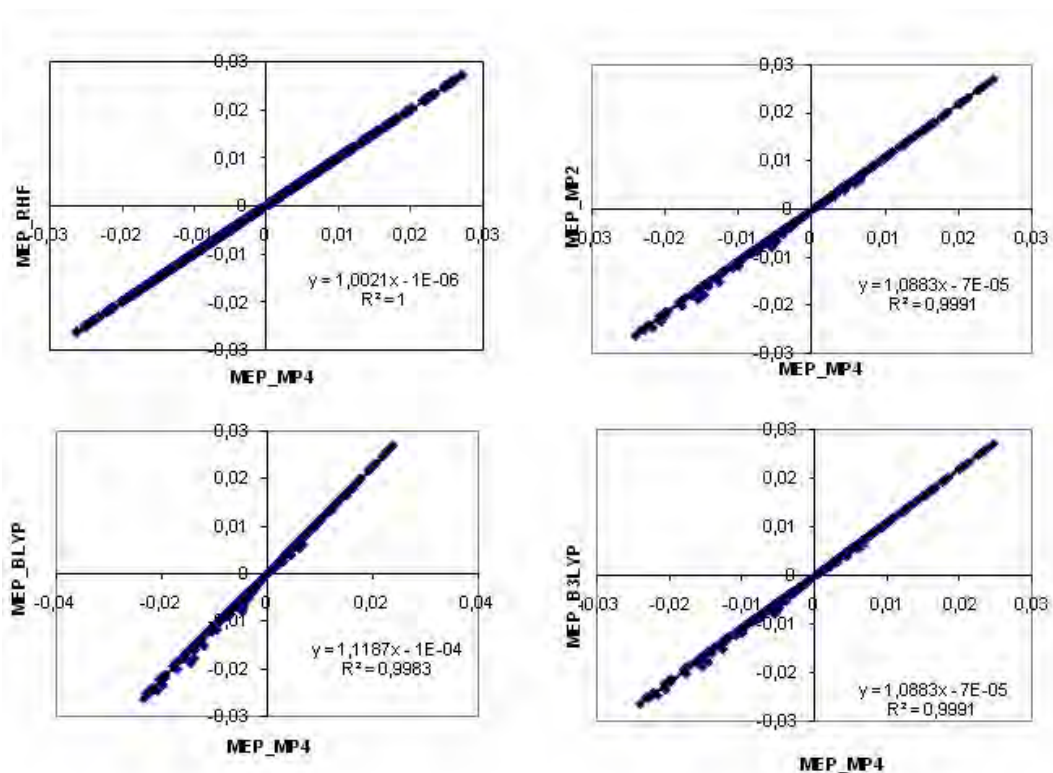


Figura 4.3.5. MEP del cianur d'hidrogen. Anàlisi de correlació amb la base 6-311++G(2d,2p)

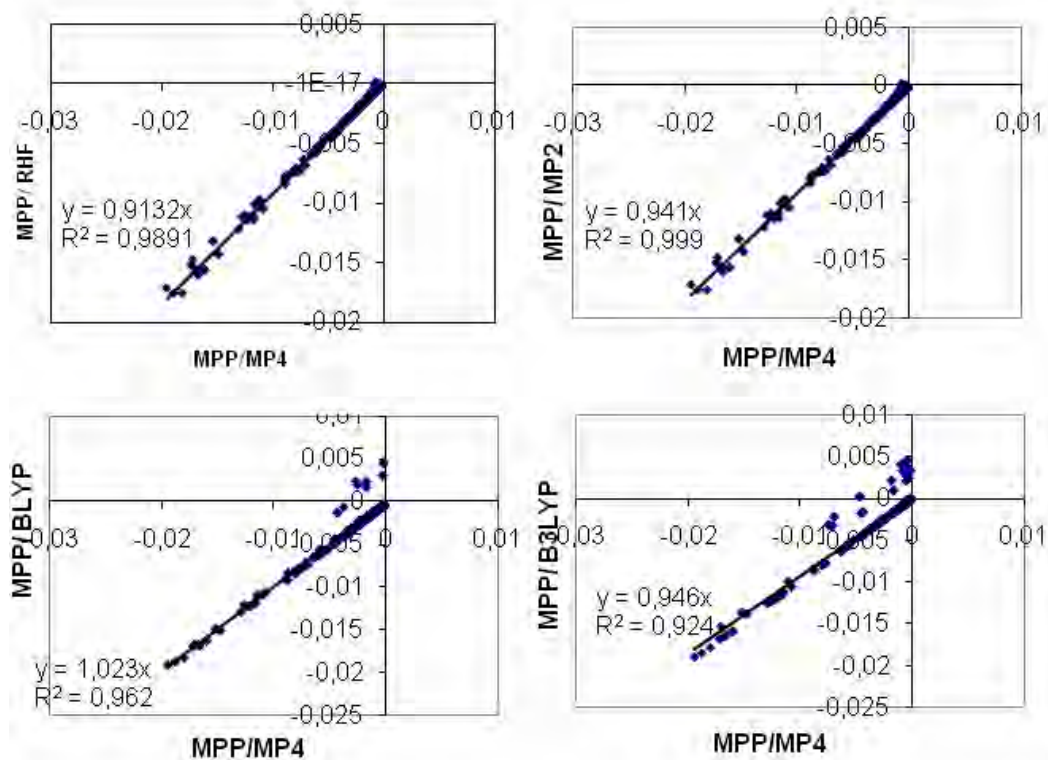


Figura 4.3.6. MPP del cianur d'hidrogen. Anàlisi de correlació amb la base 6-311++G(2d,2p)

Si fem una anàlisi de l'efecte de la base, haurem de comparar per a cada nivell de càlcul els valors obtinguts amb cada base. Aquí presentem la comparació dels resultats obtinguts quan s'augmenta la base utilitzant MP4 com mètode de càlcul. Per això hem pres com a referència els càlculs fets a nivell MP4 amb la base que inclou funcions difuses 6-311++(2d,2p). La resta de gràfiques comparatives es poden trobar a l'annex 1.

A les Figures 4.3.7-4.3.9 es mostren els resultats de la comparació dels valors de potencial de polarització per a les tres molècules objecte d'aquest estudi calculats a nivell MP4 amb diferents bases en comparació amb la base que porta funcions difuses. Com es pot veure la dispersió dels valors disminueix a mesura que fem la base més estesa. En el cas del formaldehid i del cianur d'hidrogen els coeficients de regressió són força bons de $r^2=0.98$. En canvi en el cas de la urea els resultats presenten una dispersió més gran com es pot veure per la davallada dels valors del coeficient de regressió. El pas de la base 6-31G (d) a 6-31G(d,p) no millora significativament la correlació. En canvi la utilització de dos conjunts de polarització permet una millora substancial.

Com es podrà veure més endavant la dispersió dels punts a la recta de regressió es pot analitzar millor amb l'estudi de la distribució espacial dels mapes de diferències de polarització que es presenta a la secció següent.

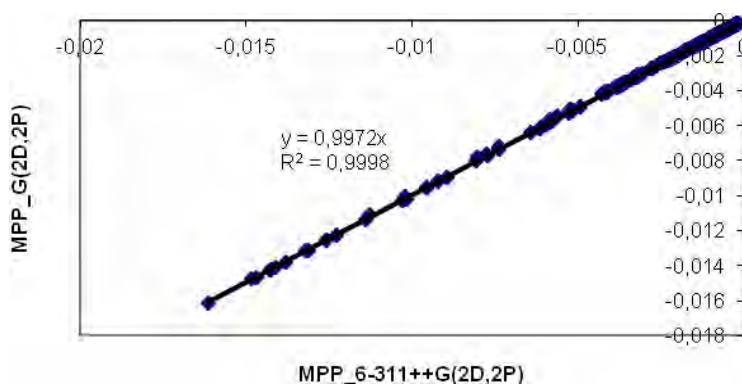
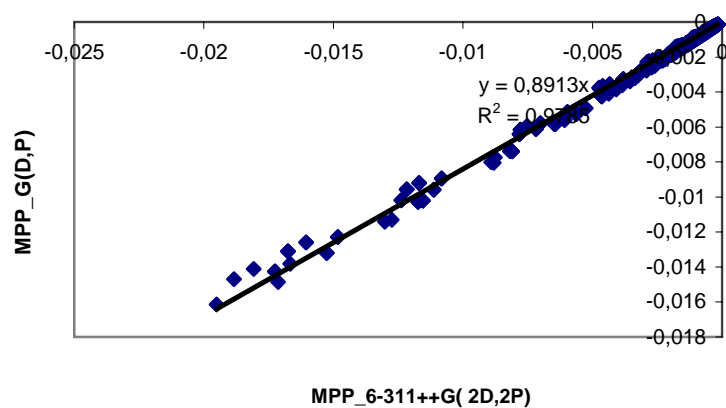
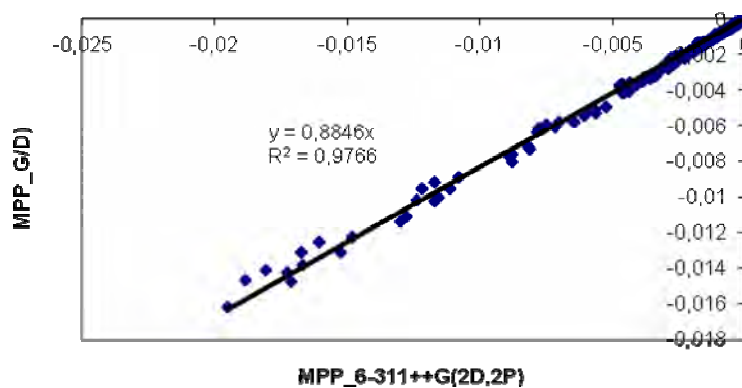


Figura 4.3.7. Comparació del càlculs del MPP per el formaldehid. Comparació dels càlculs a nivell MP4 fets amb la base 6-311++G(2d,2p) i les diferents bases estudiades en aquest treball

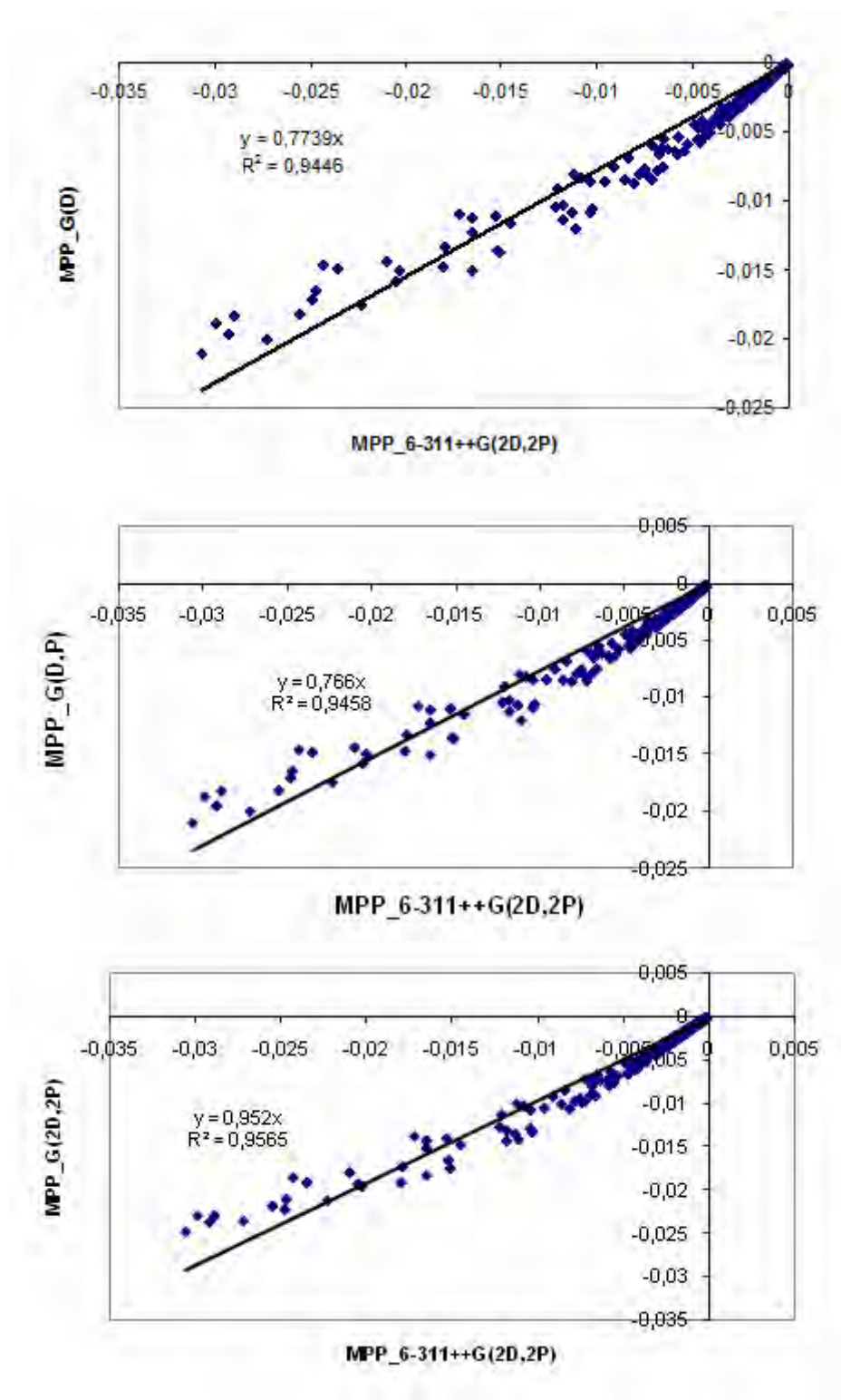


Figura 4.3.8. Comparació del càlculs del MPP per la urea . Comparació dels càlculs a nivell MP4 fets amb la base 6-311++G(2d,2p) i les diferents bases estudiades en aquest treball

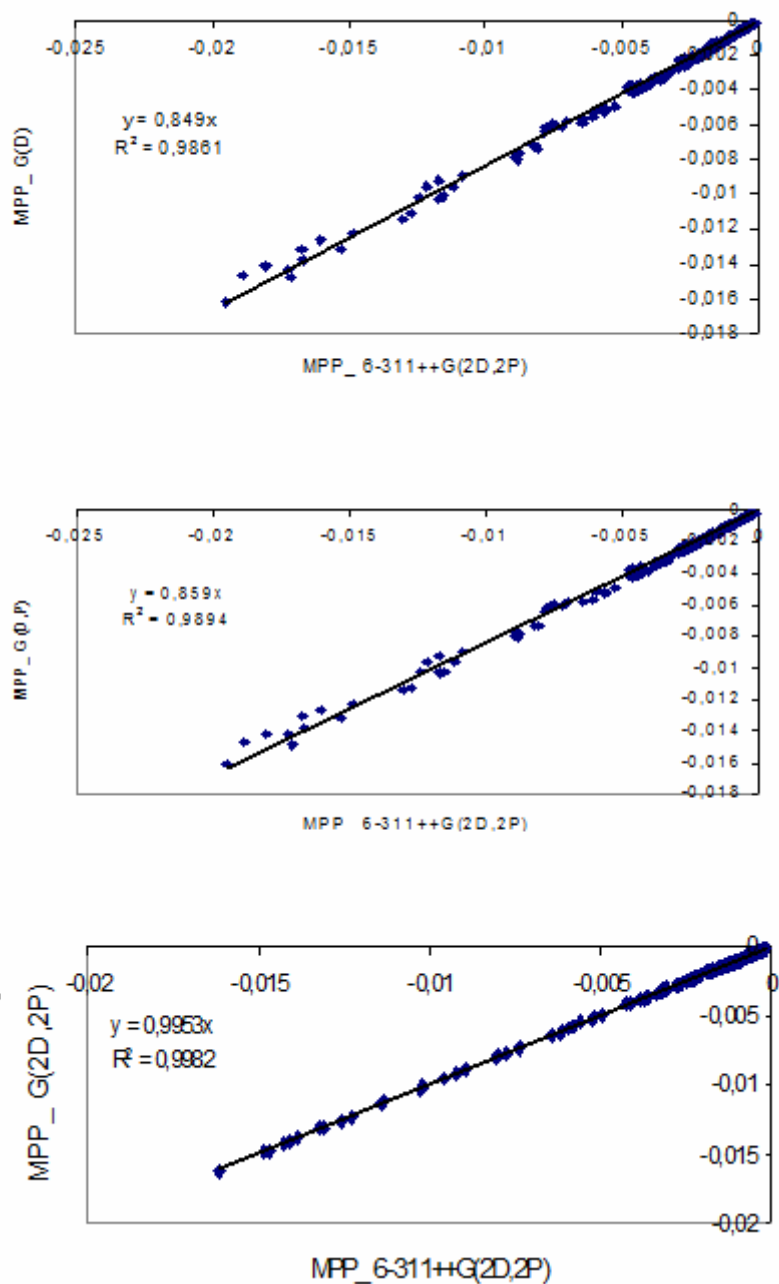


Figura 4.3.9. Comparació del càlculs del MPP per al cianur d'hidrogen. Comparació dels càlculs a nivell MP4 fets amb la base 6-311++G(2d,2p) i les diferents bases estudiades en aquest treball

4.4 Anàlisi espacial dels valors del MEP i MPP calculats amb diferents bases i metodologies

L'anàlisi de la distribució espacial de les diferències entre els MPPs calculats amb diferents metodologies i bases ens pot proporcionar informació sobre les limitacions del mètode de càlcul en relació a la distribució electrònica molecular. En aquest apartat presentem una selecció dels mapes per poder fer una anàlisi completa.

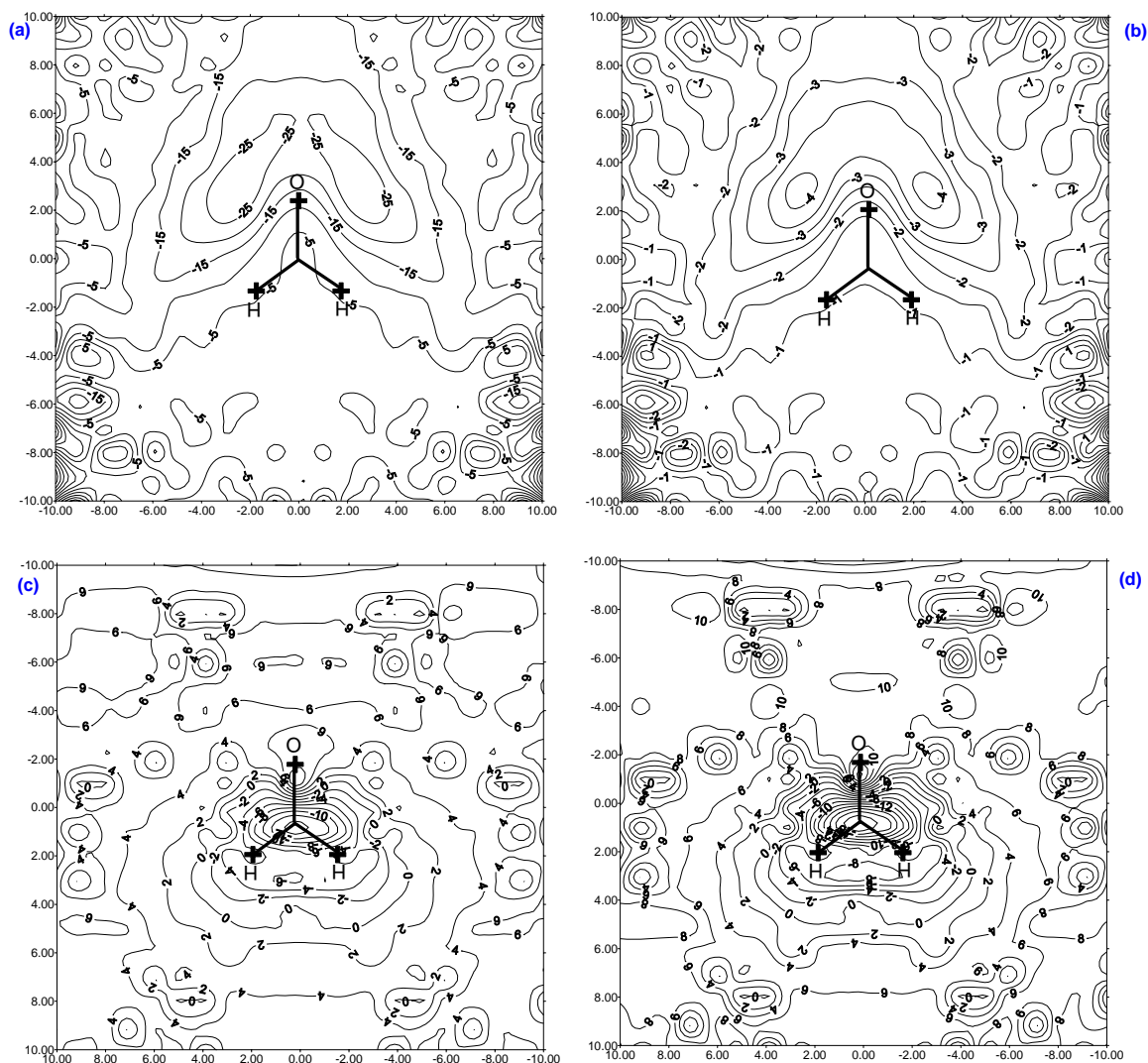


Figura 4.4.1. Mapes de diferències del potencial de polarització molecular (en percentatge) per al formaldehid a diferents nivells de teoria emprant com a referència els càlculs a nivell MP4 amb la base 6-311++G(2d,2p)(a) RHF/MP4 (b) MP2/MP4 (c) BLYP/MP4 (d) B3LYP/MP4

En primer lloc analitzarem l'efecte del mètode de càlcul. A la Figura 4.4.1 es mostren els mapes de diferències en percentatge per al formaldehid fets amb la base 6-311++G(2d,2p) prenent com a referència els càlculs fets a nivell MP4. Com es pot veure a la Figura 4.4.1 la inclusió de la correlació electrònica dona diferències de l'ordre del 25% en àrees properes als

parells no-enllaçants de l'oxigen i per tant més polaritzables. En altres direccions aquest efecte no és tan gran i dona en promig valors de diferències de l'ordre del 10%, que correlacionen bé amb els resultats obtinguts per a la polarizabilitat molecular descrits en el capítol anterior. Si ara analitzem el cas del càlcul MP2 (Figura 4.4.2) les diferències són més petites. Com es pot veure les variacions més grans són també al voltant del parells no-enllaçants de l'oxigen però, amb diferències del 4%. Això vol dir que la major part de l'efecte sobre la polarització de la inclusió de la correlació electrònica queda recollida a nivell MP2.

La comparació entre el càlcul fet a nivell MP4 i els càlculs fets amb el funcional de la densitat és molt interessant (Figures 4.4.1c i 4.4.1d). Com es pot veure el comportament dels funcionals BLYP i B3LYP són semblants. A diferència de les comparacions anteriors es pot veure que les variacions són petites però la superfície presenta bonys amb valors positius i negatius. Clarament a llargues distàncies de la molècula el funcional de la densitat sobreestima la polarització un 10% respecte al mètode de referència, MP4. Aquest fet es pot correlacionar amb la sobreestimació dels valors de la polarizabilitat que s'han mostrat al capítol anterior.

A la Figura 4.4.2 es mostren els mapes de diferències del comportament de la metodologia (en percentatge) per a la molècula d'urea. En aquest cas la comparació entre el càlcul MP4 i el càlcul Hartree-Fock mostra unes diferències d'un 20% al voltant de tota la molècula. Si ara fem la comparació amb els càlculs MP2 es pot veure també una distribució més uniforme al voltant de la molècula amb diferències del 3%. En la comparació amb els càlculs fets amb el funcional de la densitat s'observa el mateix que s'havia estat observant amb el formaldehid: el funcional de la densitat sobreestima la polarització a llargues distàncies de la molècula. També s'observa (Figures 4.4.2c i 4.4.2d) que els dos funcionals estudiats tenen un comportament semblant.

Per últim, a la Figura 4.4.3 es mostren els mapes corresponents a la molècula de cianur d'hidrogen. En aquest cas la diferència entre els càlculs MP4 i Hartree-Fock són més petites, encara que de l'ordre del 15% al voltant del parell no-enllaçant del nitrogen (Figura 4.4.3a). Novament es pot veure que les diferències amb el càlcul MP2 no són molt grans, amb valors del 3%. En el cas dels càlculs efectuats amb el funcional de densitat es torna a veure que el mapa de diferències té bonys amb valors positius i negatius. En aquest cas són les direccions perpendiculars a la molècula on el DFT sobreestima els valors de la polarització, en canvi a la direcció de la molècula la variació de la polarització és oscil·lant. També es pot observar que els dos funcionals presenten un comportament semblant.

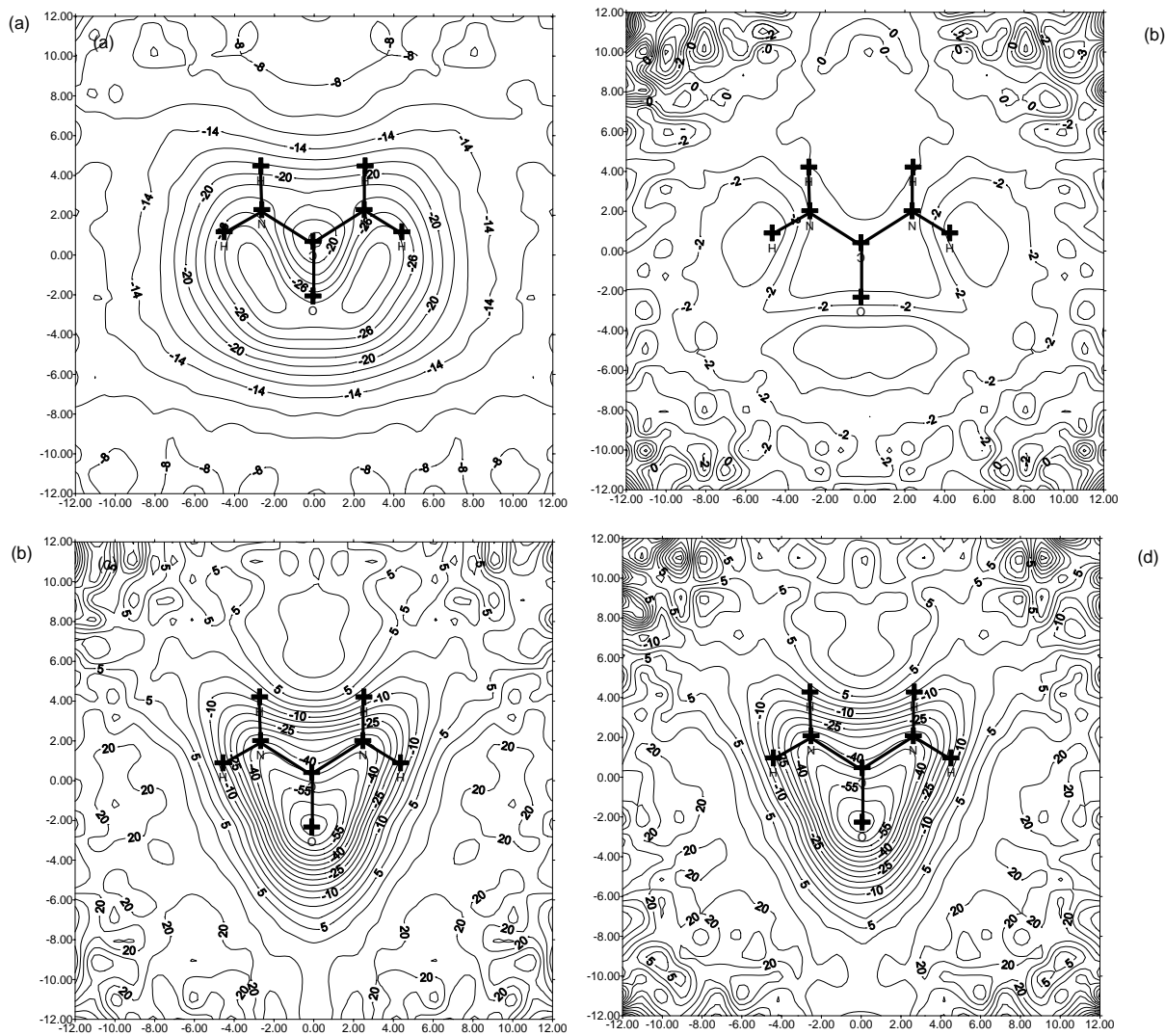


Figura 4.4.2. Mapes de diferències del potencial de polarització molecular (en percentatge) per la urea a diferents nivells de teoria emprant com a referència els càlculs a nivell MP4 amb la base 6-311++G(2d,2p)(a) RHF/MP4 (b) MP2/MP4 (c) BLYP/MP4 (d) B3LYP/MP4

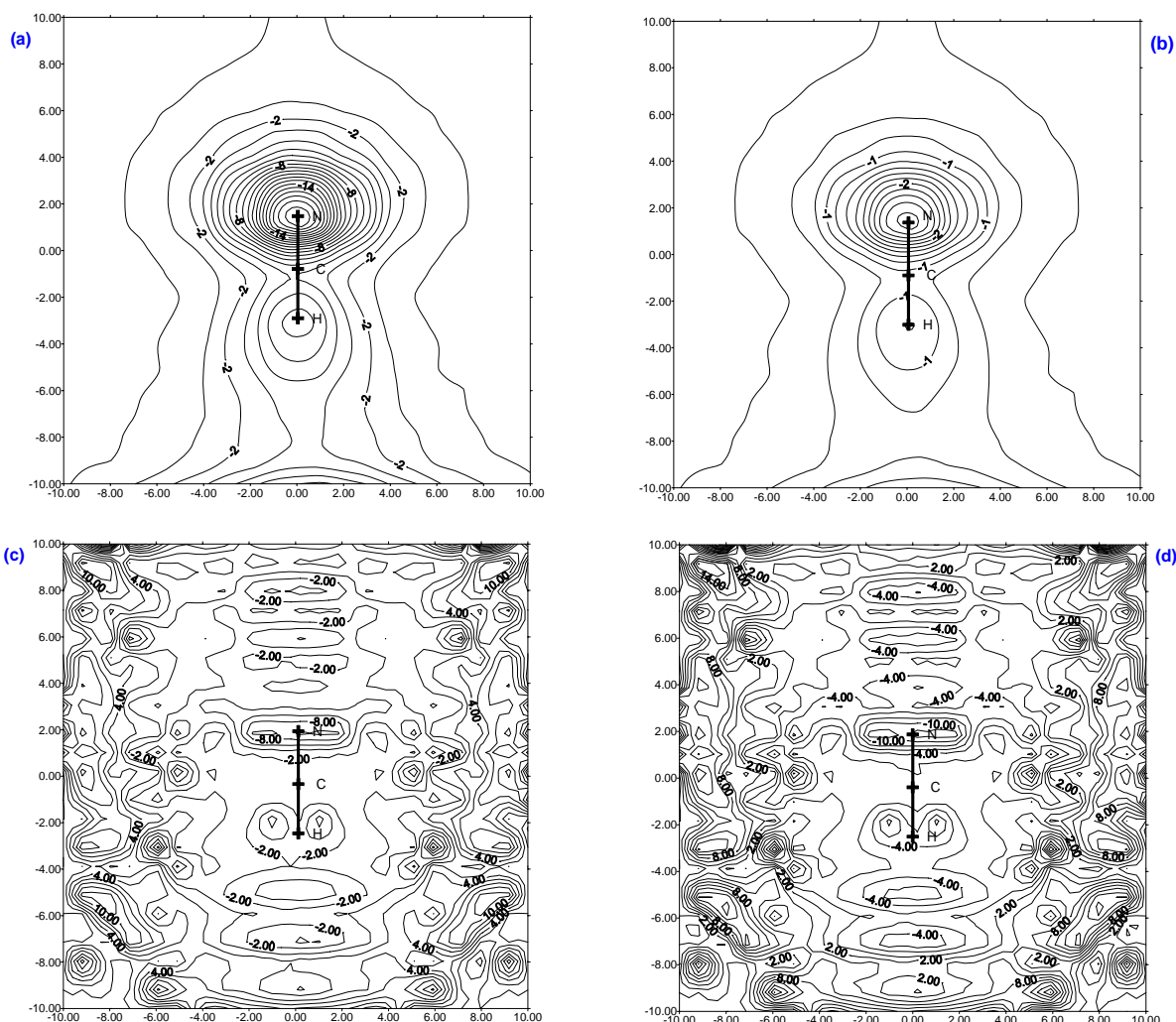


Figura 4.4.3. Mapes de diferències del potencial de polarització molecular (en percentatge) per al formaldehid a diferents nivells de teoria emprant com a referència els càlculs a nivell MP4 amb la base 6-311++G(2d,2p)(a) RHF/MP4 (b) MP2/MP4 (c) BLYP/MP4 (d) B3LYP/MP4

La comparació de l'efecte de la base es pot analitzar a les gràfiques 4.4.4-4.4.6. En aquestes gràfiques s'han representat els mapes de diferències de MPP a nivell MP4, prenent com a referència el càlcul MP4 amb la base més estesa 6-311++G(2d,2p).

El cas del formaldehid es mostra a la gràfica 4.4.4. Es pot veure en general, que l'efecte d'afegir les funcions difuses permet millorar els valors de la polarització considerablement. Si comparem amb la base menys estesa els valors promig són del ordre del 12% repartides de forma semblant per tot l'espai (Figura 4.4.4c).

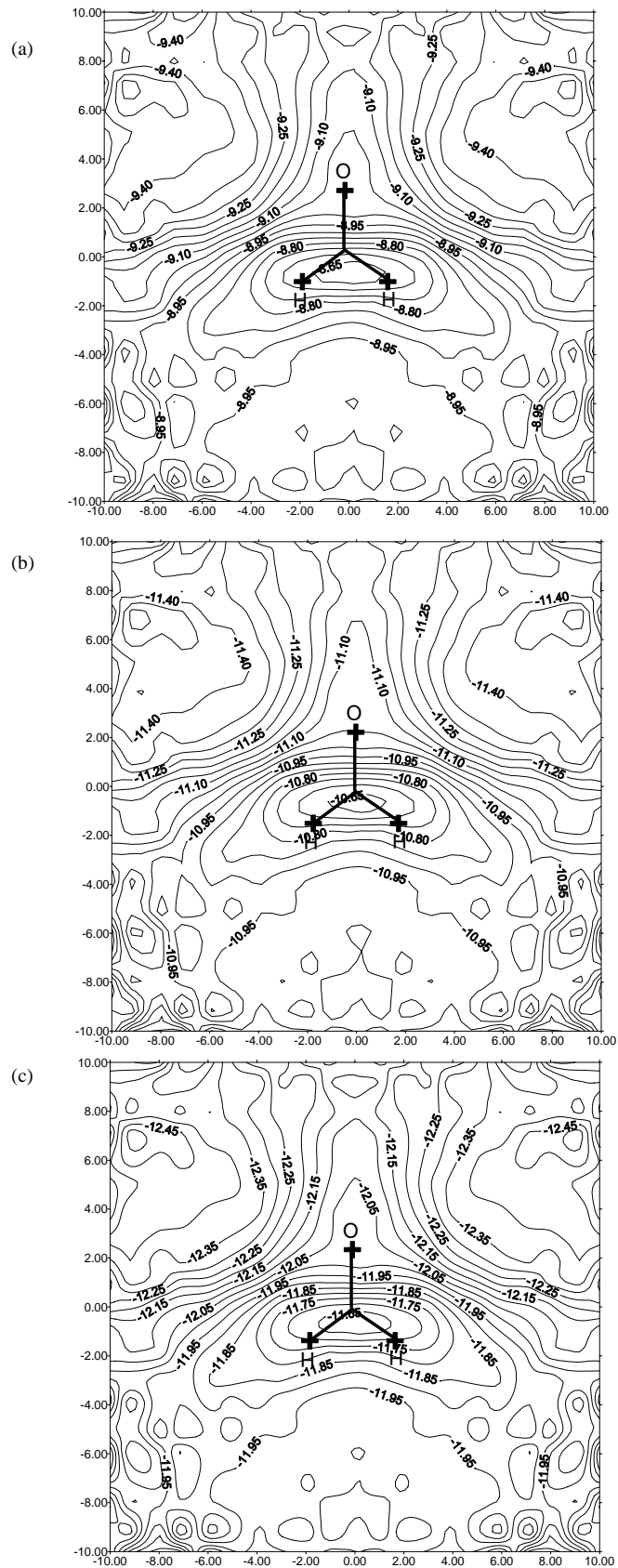


Figura 4.4.4. Mapes de diferències del potencial de polarització molecular (en percentatge) per al formaldehid amb diferents bases emprant com a referència els càlculs a nivell MP4 amb la base 6-311++G(2d,2p) (a) 6-31G(2d,2p); (b) 6-31G(d,p); (c) 6-31G(d)

S'observa que aquest fet es corregeix lleugerament en afegir funcions de polarització als hidrògens, passant a una diferència de l'ordre del 10% (Figura 4.4.4b), distribuïdes de forma regular per tot el mapa. Si s'analitza ara l'efecte d'afegir un segon conjunt de funcions de polarització es veu que el mapa millora sobre tot al voltant dels nuclis, però encara hi ha diferències de l'ordre del 9% a llargues distàncies de la molècula.

A la figura 4.4.5 es mostra en els mapes de diferències corresponents a la molècula d'urea. Com es pot veure en aquest cas les diferències amb la base menys estesa són de l'ordre del 10% (Figura 4.4.5c). Quan s'afegeix la polarització sobre els hidrògens disminueixen les diferències al voltant del hidrògens, però es mantenen les diferències en promig al voltant del 10%. Finalment, quan es fa la comparació amb la base que conté dos conjunts de funcions de polarització les diferències baixen al voltant dels nuclis, però es mantenen a llargues distàncies de la molècula.

Finalment, a la Figura 4.4.6 es mostra l'efecte de la base en el cas de la molècula de cianur d'hidrogen. Es poden observar les mateixes tendències que les que hem vist amb la resta de molècules. Hi ha diferències de l'ordre del 10% (Figura 4.4.6a) encara que una mica més altes al voltant de l'hidrogen. Aquestes darreres milloren al passar a la base 6-31G(d,p) i baixen un 2% al voltant dels nuclis quan es posen dos conjunts de funcions de polarització.

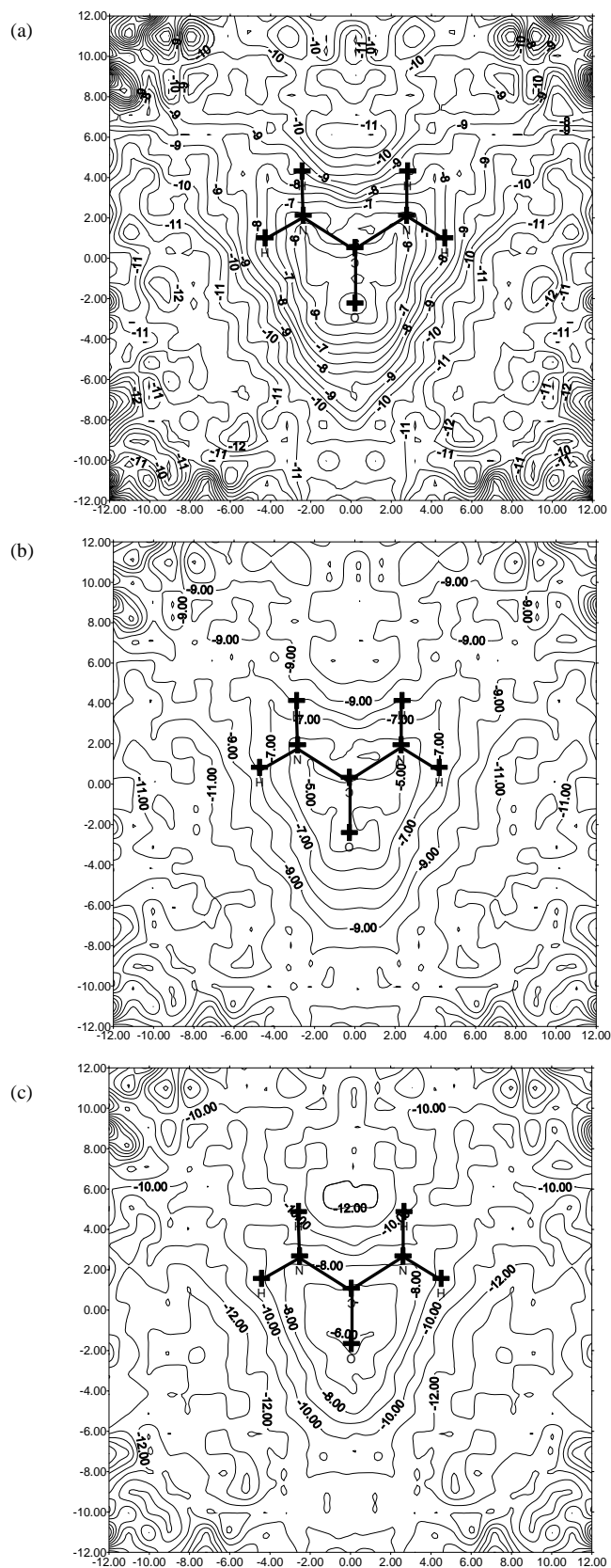


Figura 4.4.5. Mapes de diferències del potencial de polarització molecular (en percentatge) per la urea amb diferents bases emprant com a referència els càlculs a nivell MP4 amb la base 6-311++G(2d,2p) (a) 6-31G(2d,2p); (b) 6-31G(d,p); (c) 6-31G(d)

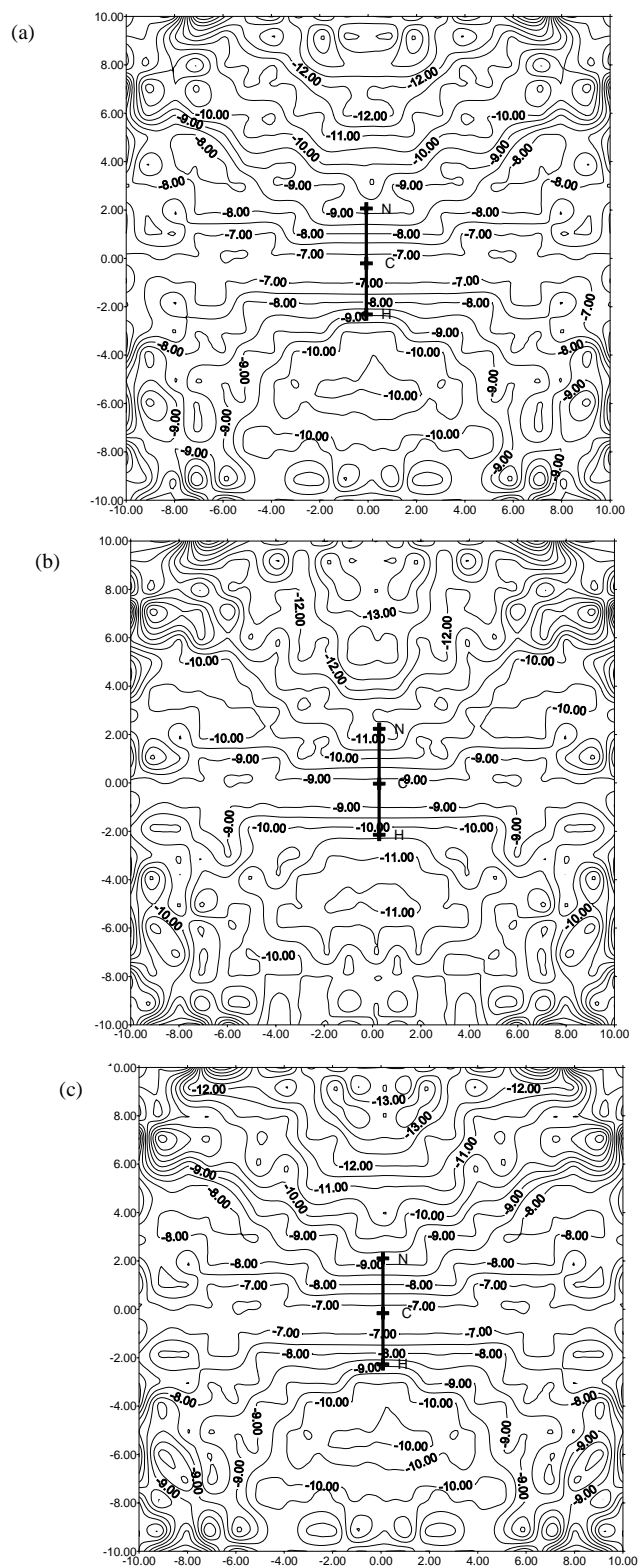


Figura 4.4.6. Mapes de diferències del potencial de polarització molecular (en percentatge) per al cianur d'hidrogen amb diferents bases emprant com a referència els càlculs a nivell MP4 amb la base 6-311++G(2d,2p) (a) 6-31G(2d,2p); (b) 6-31G(d,p); (c) 6-31G(d)

Si ara fem l'estudi de l'efecte de la base sobre els càlculs fets amb el funcional de la densitat podem veure que aquest efecte no es tan dramàtic com a nivell MP4. A la Figura 4.4.7 es podem veure els mapes de diferències calculats amb el funcional BLYP amb les diferents bases, prenent com a referència la corresponent a la base amb funcions difuses.

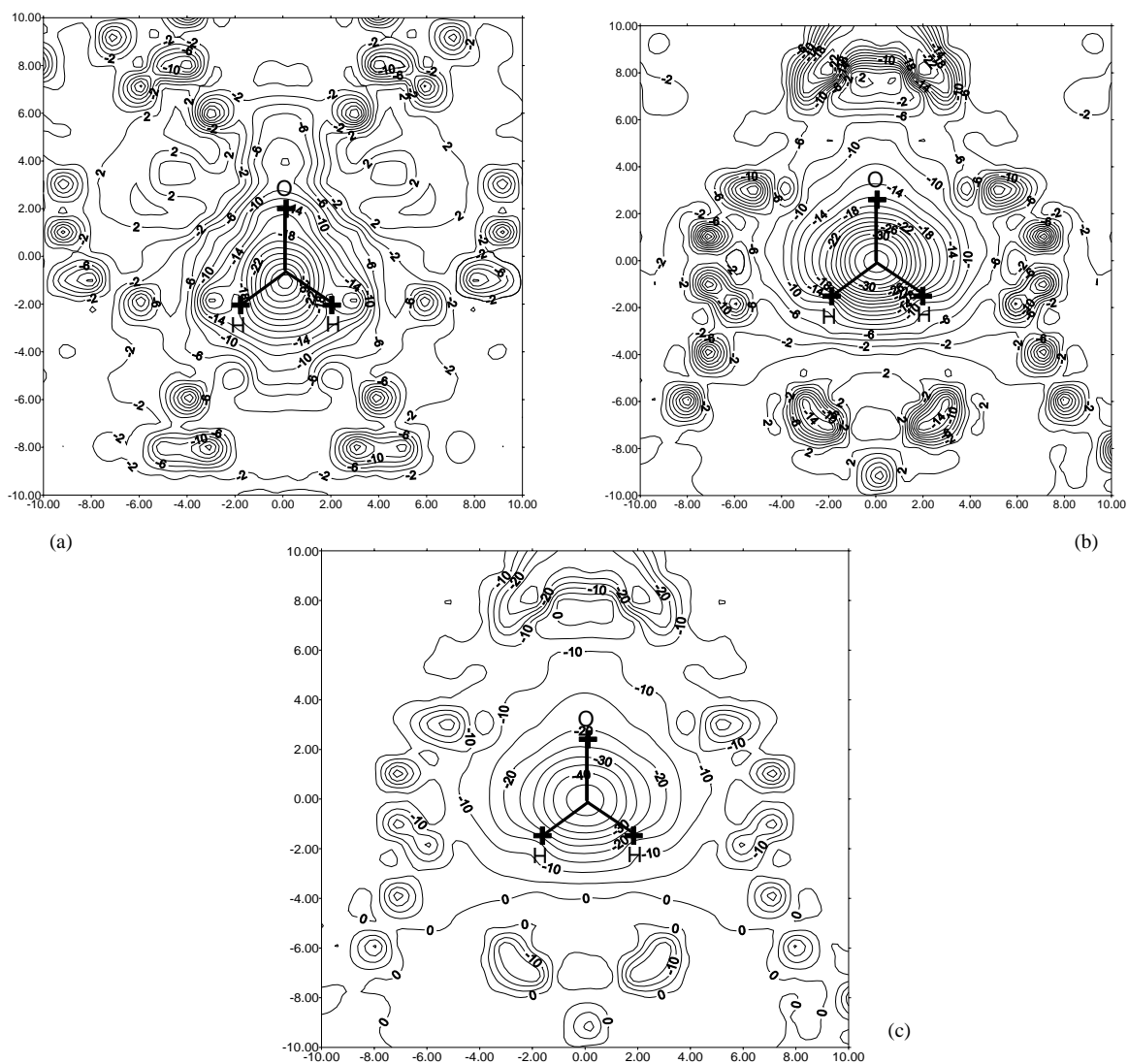


Figura 4.4.7. Mapes de diferències del potencial de polarització molecular (en percentatge) per al formaldehid amb diferents bases emprant com a referència els càlculs a nivell DFT amb el funcional BLYP amb la base 6-311++G(2d,2p) (a) 6-31G(2d,2p); (b) 6-31G(d,p); (c) 6-31G(d)

Com es pot veure al voltant dels nuclis els resultats dels tres càlculs són semblants amb polaritzacions de l'ordre de 20-25% més petites que respecte als càlculs amb la base amb difuses. En canvi a llargues distàncies de la molècula aquestes diferències disminueixen fins un 10% per el cas de las bases amb un conjunt de funcions de polarització. És un cas notable que a punts allunyats de la molècula aquestes diferències disminueixen fins al 2%, en quant al càlcul s'afegeix un segon conjunt de funcions de polarització, encara que puntualment es mantenen.

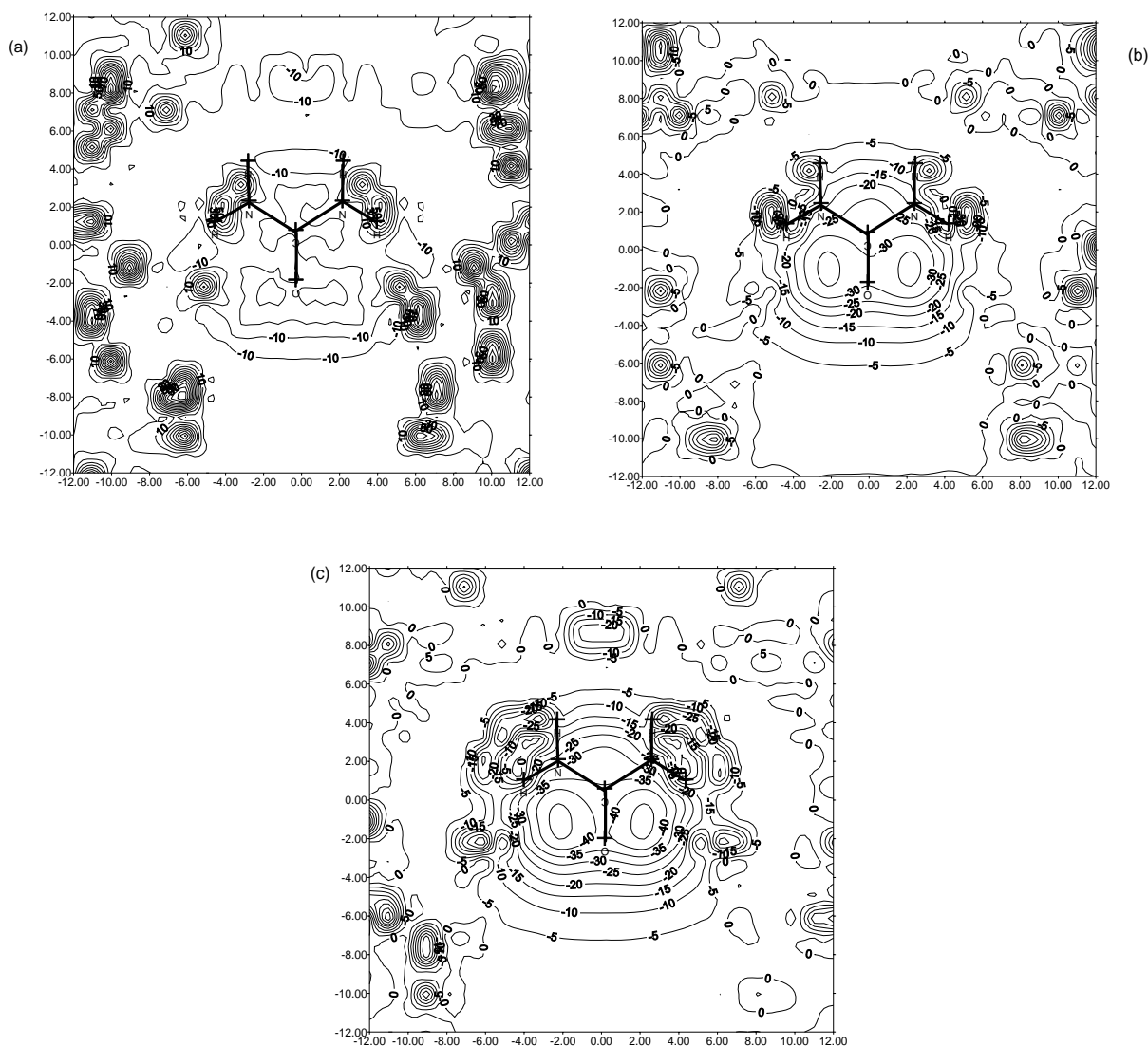


Figura 4.4.8. Mapes de diferències del potencial de polarització molecular (en percentatge) per la urea amb diferents bases emprant com a referència els càlculs a nivell DFT amb el funcional BLYP amb la base 6-311++G (2d,2p) (a) 6-31G(2d,2p); (b) 6-31G(d,p); (c) 6-31G(d)

Aquest comportament s'estén a la molècula d'urea tal i com es pot observar a la figura 4.4.8. Al voltant dels nuclis les diferències amb el càlcul fet amb la base que conté funcions difuses són de l'ordre del 25%. En canvi lluny de la molècula aquestes diferències pràcticament desapareixen. En el cas del càlcul amb dos conjunts de funcions de polarització aquestes diferències són més petites, al voltant del 10%. En punts llunyans de la molècula les diferències estan en l'interval 5-10%.

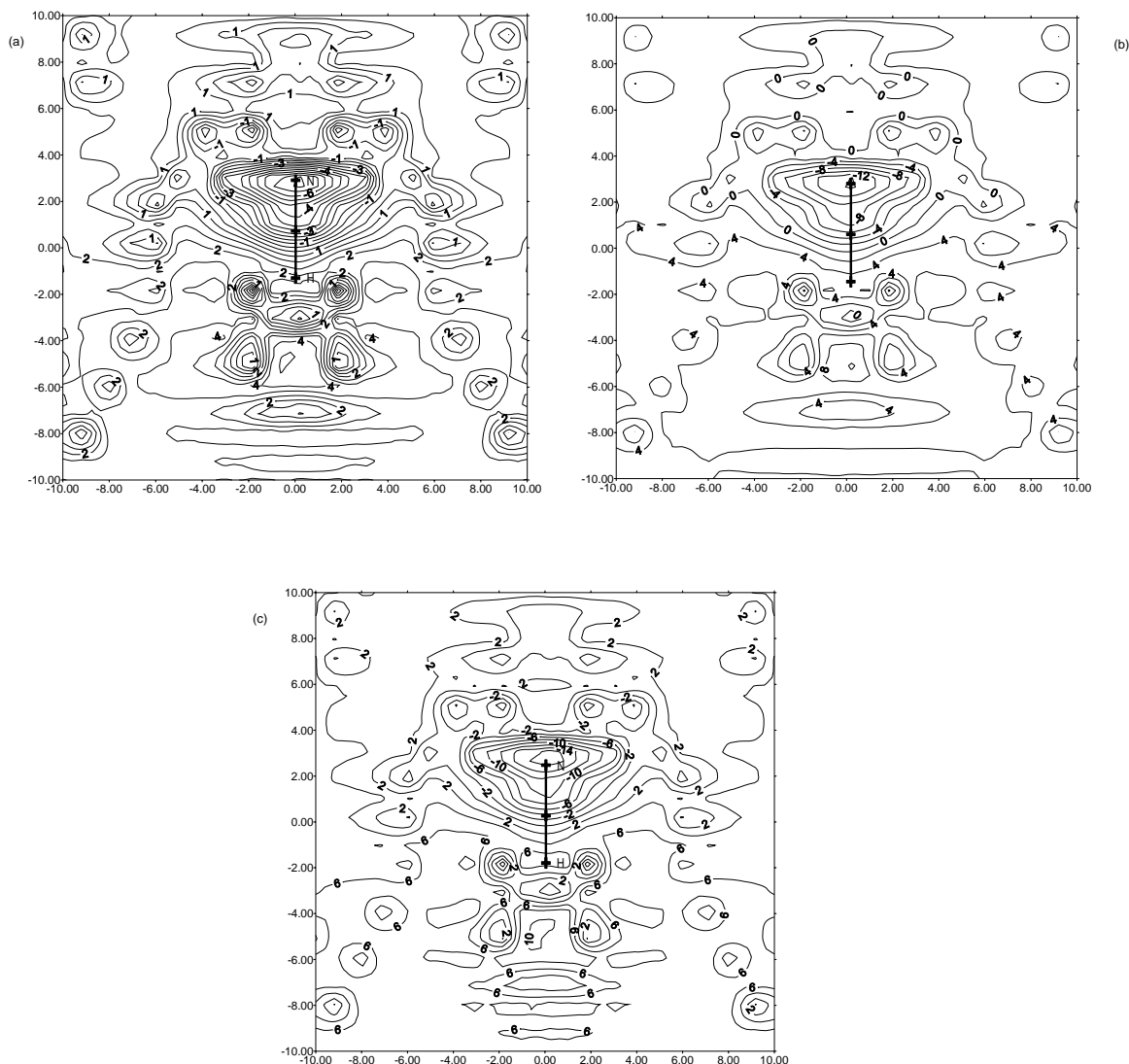


Figura 4.4.9. Mapes de diferències del potencial de polarització molecular (en percentatge) per al cianur d'hidrogen amb diferents bases emprant com a referència els càlculs a nivell DFT amb el funcional BLYP amb la base 6-311++G(2d,2p) (a) 6-31G(2d,2p); (b) 6-31G(d,p); (c) 6-31G(d)

A la Figura 4.4.9 es presenten gràficament els resultats per el cianur d'hidrogen. En aquest cas les diferències al voltant dels nuclis són del 10% o més petites, 5% pel cas dels càlculs amb dos conjunts de funcions de polarització. Les diferències a distàncies llunyanes són de l'ordre del 5% pels càlculs amb un conjunt de funcions de polarització que disminueixen fins a

2% quan s'estén la base amb la utilització de dos conjunts de polarització. Els resultats presentats han estat publicats recentment (Roset, 2010).

Finalment, podem fer la comparació dels càlculs fets amb el mètode de la supermolècula, és a dir el mètode on el MPP es calcula com a diferència entre el mapa d'interacció i el de potencial electrostàtic, i els càlculs fets utilitzant la teoria de pertorbacions de Rayleigh-Schrödinger. A les Figures 4.4.10- 4.4.12 es mostren les diferències percentuals dels càlculs fets amb la teoria de pertorbacions amb les diferents bases utilitzades en aquest treball prenent com a referència el càlcul Hartree-Fock fet amb el mètode de la supermolècula amb la base que conté funcions difuses. Com es pot observar, la polarització calculada al voltant dels nuclis a través del mètode de la supermolècula és més gran que amb la teoria de pertorbacions. A distàncies més llargues aquesta tendència s'inverteix i a llargues distàncies les diferències són nul·les. Aquests efectes es fan menys petits amb la grandària de la base.

A la Figura 4.4.10 es presenten els resultats per a la molècula de formaldehid. Amb la base 6-31G(d) s'observen variacions de l'ordre del 14% al voltant del parell d'electrons no-enllaçants de l'oxigen que disminueixen a un 10% en afegir funcions de polarització als hidrògens i que baixen fins a un 4% en utilitzar una base amb dos conjunts de funcions de polarització. A llargues distàncies de la molècula es veu que les diferències no són molt grans, oscil·lant entre un 6% fins un 1% per a la base més estesa.

A la Figura 4.4.11 es mostren els resultats per a la molècula d'urea. De manera molt semblant al cas del formaldehid, les variacions per zones properes als nuclis van del 30% amb la base menys estesa fins al 10% amb la base amb funcions difuses. El màxim positiu va des del 10% en el cas de la base menys estesa fins el 4% al cas de la base amb funcions difuses.

A la Figura 4.4.12 es mostren els resultats per a la molècula de cianur d'hidrogen. De manera molt semblant als casos del formaldehid i la urea, les variacions per zones properes als nuclis van del 15% amb la base menys estesa fins al 10% amb la base amb funcions difuses.

La comparació entre els càlculs de la supermolècula i els efectuats mitjançant la teoria de pertorbacions ens permet deduir que aquesta última és efectiva per al càlcul de potencials de polarització molecular amb un cost computacional més petit.

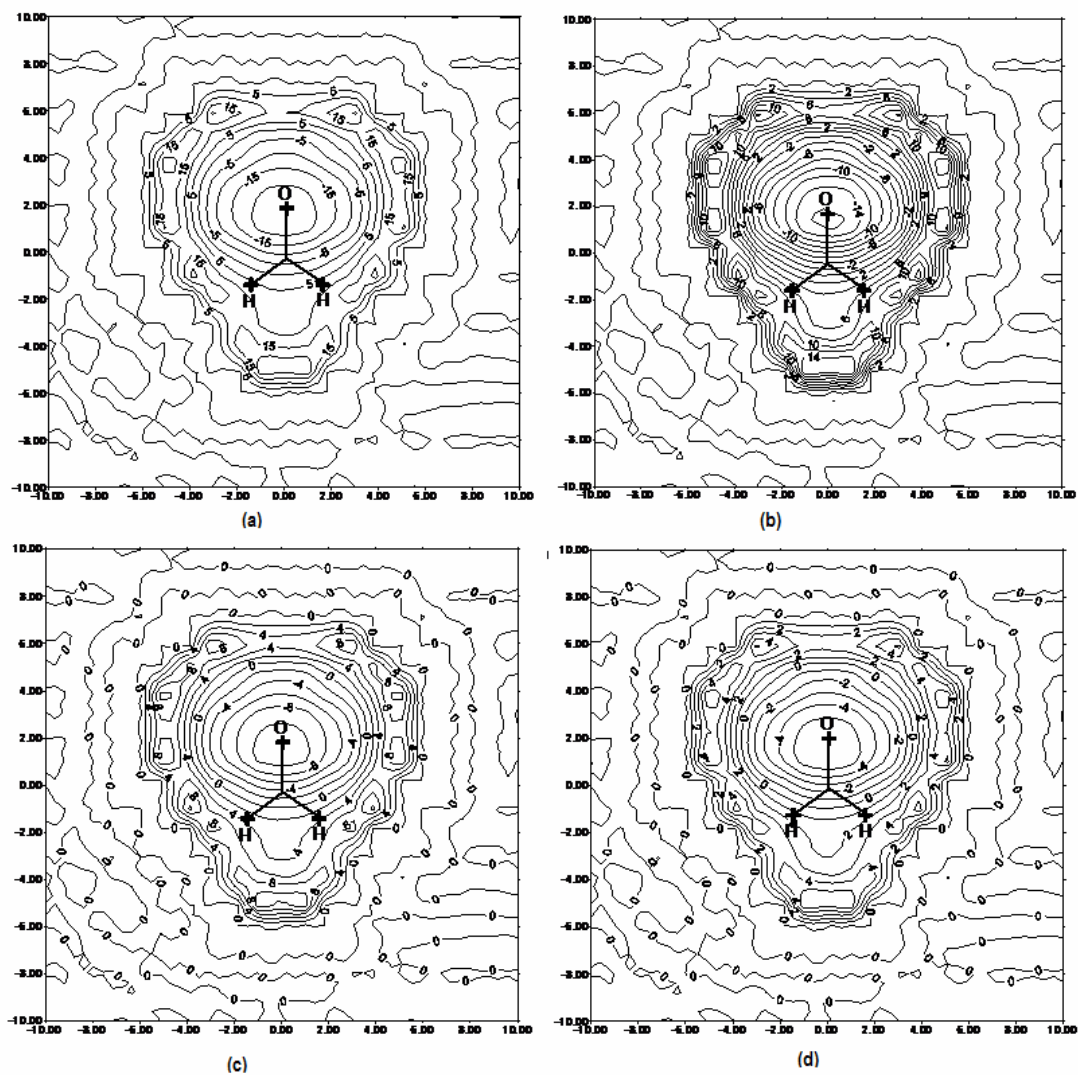


Figura 4.4.10. Mapes de diferències del potencial de polarització molecular (en percentatge) per al formaldehid calculat amb la teoria de pertorbacions amb diferents bases emprant com a referència els càlculs a nivell Hartree-Fock amb la base 6-311++G(2d,2p). a) 6-31G(d); b) 6-31G(d,p); c) 6-31G(2d,2p); d) 6-311++G(2d,2p)

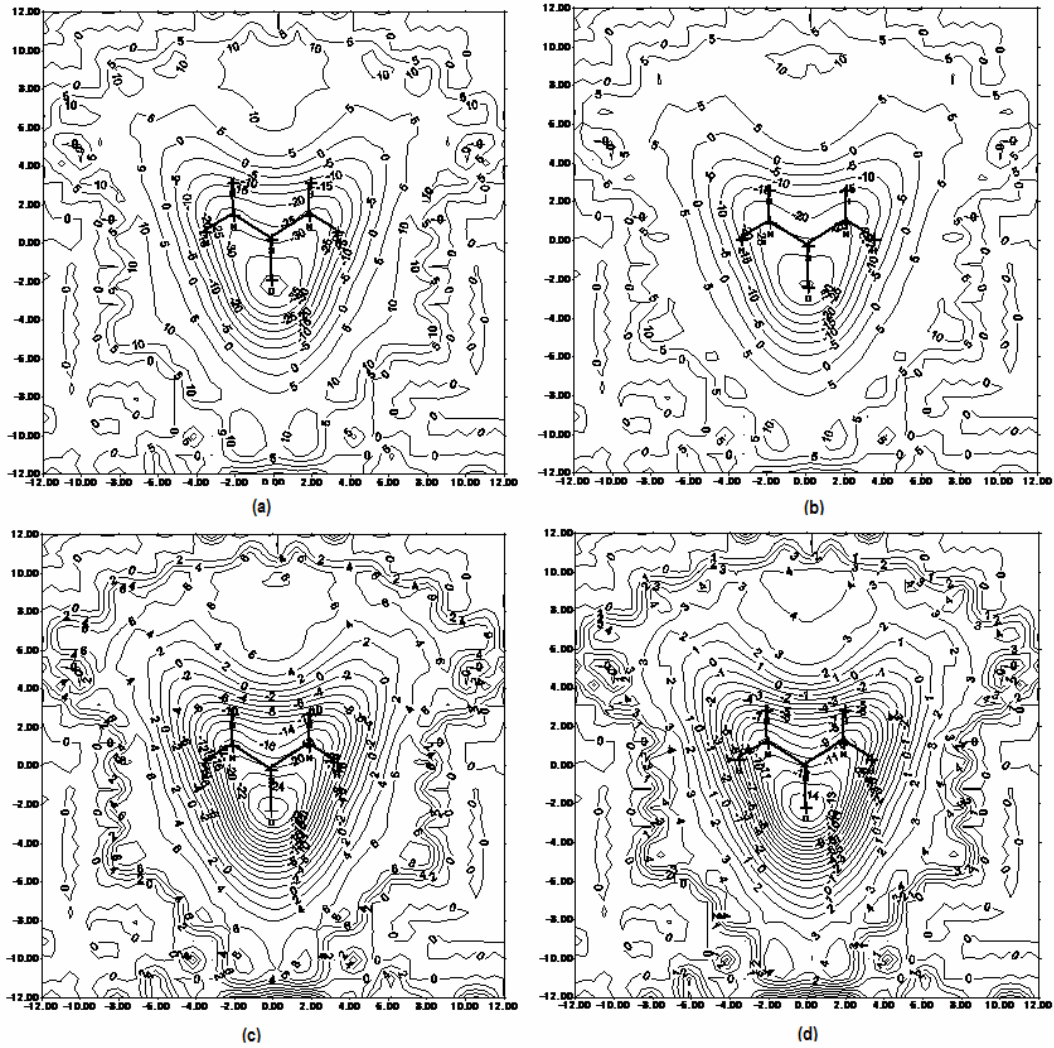


Figura 4.4.11. Mapes de diferències del potencial de polarització molecular (en percentatge) per la urea calculat amb la teoria de pertorbacions amb diferents bases emprant com a referència els càlculs a nivell Hartree-Fock amb la base 6-311++G(2d,2p) a) 6-31G(d); (b) 6-31G(d,p); (c) 6-31G(2d,2p); (d) 6-311++G(2d,2p)

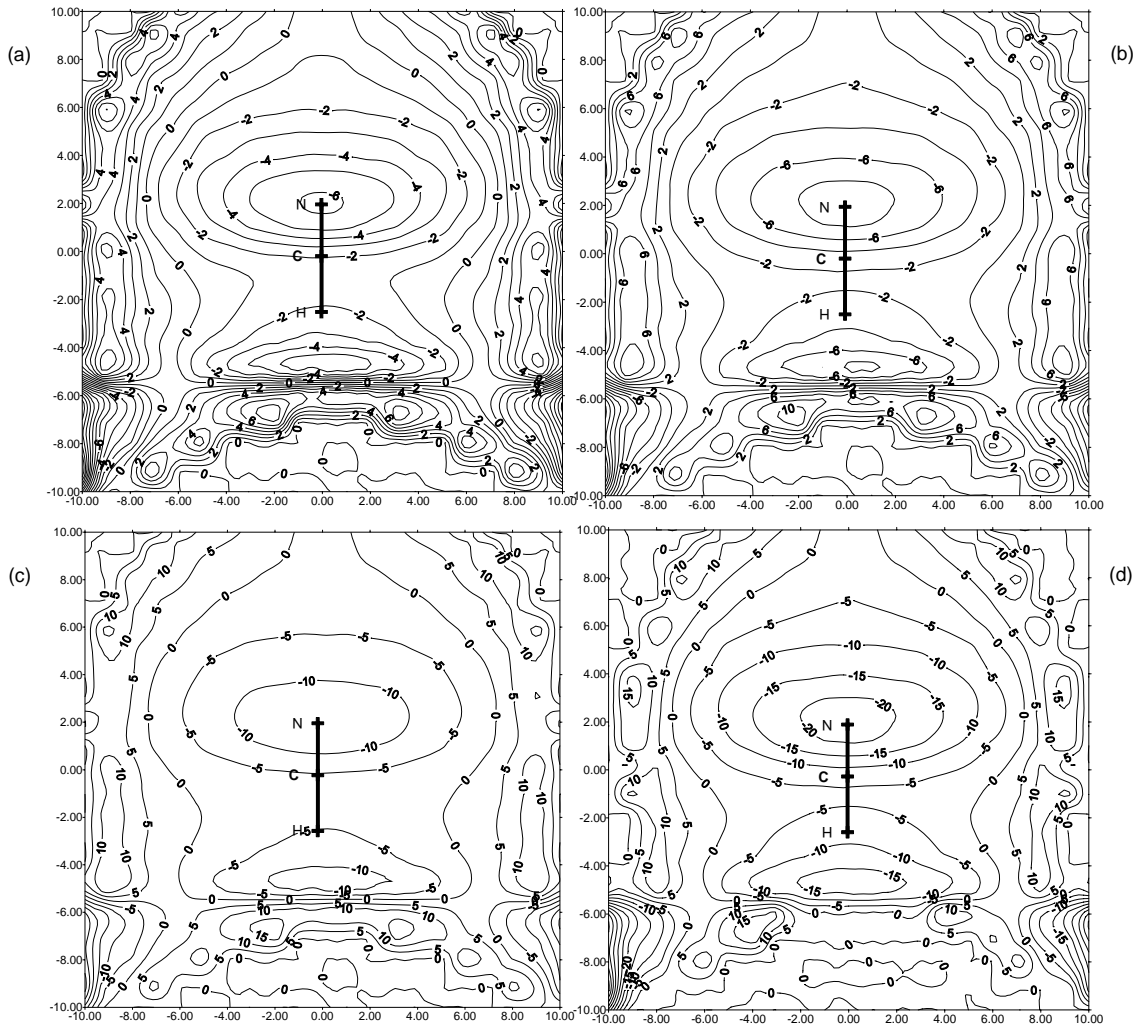


Figura 4.4.12. Mapes de diferències del potencial de polarització molecular (en percentatge) per al cianur d'hidrogen calculat amb la teoria de pertorbacions amb diferents bases emprant com a referència els càlculs a nivell Hartree-Fock amb la base 6-311++G(2d,2p) a) 6-31G(d); (b) 6-31G(d,p); (c) 6-31G(2d,2p); (d) 6-311++G(2d,2p)

5 LA POLARITZABILITAT MOLECULAR COM A
DESCRIPTOR DE L'ACTIVITAT BIOLÒGICA DE
MOLÈCULES: ESTUDI DE DIOXINES I FURANS

5.1 Introducció

Com ja hem comentat en capítols anteriors, la interacció entre dos sistemes moleculars de capes tancades a llargues distàncies es pot descriure com una interacció electrostàtica amb dues contribucions: la de Coulomb i la d'inducció. A curtes distàncies s'ha d'afegir una contribució atractiva addicional que prové del comportament quàntic de la matèria i que es coneix com energia de dispersió i un terme de repulsió que correspon al solapament dels radis de van der Waals. Les interaccions de Coulomb descriuen la interacció entre dos densitats de càrrega estàtiques i per tant, depenen de la distribució de densitat de càrrega molecular i es poden descriure mitjançant els diferents moments multipolars: càrrega, dipol, quadrupol, etc... Com ja hem vist al capítol 4 una forma pràctica de visualitzar-les és a través dels mapes de potencial electrostàtic (Politzer, 1981; Naray-Szabo, 1995; Barnett, 1978; Murray, 2000; Hussein, 2001; Gálvez, 2001). Per una altra banda, la inducció és una contribució de naturalesa també electrostàtica i que descriu la polarització d'una densitat electrònica de càrrega quan està sotmesa a un camp elèctric exterior, inclòs el que és generat per una molècula pròxima. Per paral·lelisme amb els mapes de potencial electrostàtic, aquesta informació es pot veure de forma gràfica als mapes de potencial de polarització, que donen idea de les regions més deformables per la presència d'un camp elèctric extern. Finalment, la dispersió és una contribució quàntica a l'energia d'interacció que únicament es pot calcular quan s'inclou la correlació electrònica.

Al present capítol s'utilitzen els mapes de potencial electrostàtic, de polarització i d'interacció per a l'anàlisi de les característiques del reconeixement molecular d'un conjunt de compostos bioactius, amb la intenció d'analitzar la importància de la contribució de la polarització a la descripció de les seves activitats. Per aquest motiu hem triat molècules amb una alta polaritzabilitat i en concret, hem triat dues famílies de compostos amb abundants àtoms de clor i que presenten activitat tòxica dins del grup de molècules conegudes com a dioxines i del conjunt de molècules conegudes com a dibenzofurans.

Després dels resultats analitzats al capítol anterior, hem triat per fer aquest estudi una metodologia que fos un compromís entre la qualitat dels resultats i el temps de càlcul necessari per tal de tenir a l'abast, de forma raonable, els mapes d'energia potencial i de polarització i com a conseqüència els d'interacció -com a suma dels dos primers- d'aquest grup de molècules. Els càlculs s'han fet a nivell Hartree-Fock utilitzant una base doble zeta amb un conjunt de funcions de polarització amb exponents especialment optimitzats per donar bons resultats de

polaritzabilitat. El càlcul de la polarització molecular s'ha efectuat utilitzant la teoria de pertorbacions de Rayleigh-Schrödinger.

Per tal de poder analitzar la informació del conjunt de molècules es va portar a terme un estudi exploratori dels mapes mitjançant l'anàlisi per components principals. L'estudi de la variància dels mapes de potencial electrostàtic, de polarització i d'interacció del conjunt de molècules objecte del present estudi permet una anàlisi dels factors comuns i de diversitat de resposta a la interacció amb altres molècules mitjançant les diferents contribucions que s'obtenen de l'anàlisi de les components principals. Així mateix, es va portar a terme un estudi de la predicció de l'activitat biològica de les dioxines i furans mitjançant càlculs d'ajust parcial de mínims quadrats (partial least squares, PLS) dels potencials electrostàtic, de polarització i d'interacció per a l'estudi de les zones de màxima variabilitat, per tal de correlacionar l'activitat biològica amb les zones més actives de les molècules.

5.2 Dioxines i furans

Les dibenzo-p-dioxines (PCDDs o dioxines) i els dibenzofurans (PCDFs) són part d'un ampli conjunt de derivats halogenats amb anells aromàtics condensats o no, coneguts com compostos aromàtics polihalogenats (PHAs) que estan fortament substituïts amb clor i/o brom en diferents posicions. La toxicitat d'aquest compostos ha estat estudiada en diferents models: carcinogènesis, hepatotoxicitat, lesions gàstriques o hiperplàsia del tracte urinari (Poland, 1982) i està mediada a través del receptor d'hidrocarburs d'aril (*aryl hydrocarbon receptor, AhR*) (Safe, 1986, 1990), de forma que la toxicitat d'aquests compostos està directament relacionada amb la complementarietat amb la cavitat d'unió. La importància mediambiental de l'estudi de l'activitat biològica i de toxicitat d'aquests compostos ha estat descrita en nombroses ocasions (Keith, 1983; Hansch, 2003). El perfil de toxicitat d'aquests compostos és molt sensible al nombre i a la distribució dels halògens a la molècula (Keith, 1983; Mason, 1986; Poland, 1982), amb la qual cosa el llocs de substitució del clors a les estructures bàsiques fan que els compostos presentin una toxicitat diferenciada.

La relació estructura-toxicitat ha estat estudiada en treballs previs (McKinney, 1984; Murray, 1987; Poso, 1993; Bonati, 1991, 1994, 1995; Wagener, 1995; Waller, 1992, 1995; Fraschini, 1995), en un intent de comprendre l'especificitat dels congèneres de PCDD. Diversos models s'han proposat per descriure l'afinitat d'aquest compostos amb Ahr, inclosos els de transferència d'electrons (Arulmozhiraja, 2000; Arulmozhiraja, 2004), model de dispersió

(McKinney, 1984), i models basats en la interacció electrostàtica (Murray, 1990; Bonati, 1994, 1995; Mhin, 2001, 2002).

La utilització del potencial electrostàtic per explicar l'activitat de diferents dibenzo-*p*-dioxines ha estat descrita a la literatura en estudis previs (Politzer, 1988; Murray, 1987; Sjöberg, 1990; Murray, 1990; Bonati, 1994), però mai s'ha estudiat la inclusió dels efectes de polarització, i com a conseqüència, el potencial d'interacció per explicar l'activitat biològica dels compostos. En aquest capítol s'analitza el paper de la polarització molecular en la descripció de l'activitat biològica d'un conjunt de dioxines i furans llistats a la Taula 5.2.1, l'estructura de les quals es mostra a les Figures 5.2.1 i 5.2.2.

Taula 5.2.1 Nomenclatura i activitat biològica de les dioxines i furans objecte d'aquest estudi.

ID	Dibenzo- <i>p</i> -dioxines	pEC_{50}	ID	Dibenzofurans	pEC_{50}
Tcdd (1)	2,3,7,8- Tetrachlorodibenzo- <i>p</i> -dioxin	8.00	23478df (1)	2,3,4,8- Tetrachlorodibenzofuran	6.70
12378dd (2)	1,2,3,7,8- Pentachlorodibenzo- <i>p</i> -dioxin	7.10	2347df (2)	2,3,7,8- Tetrachlorodibenzofuran	7.39
2367dd (3)	2,3,6,7- Tetrachlorodibenzo- <i>p</i> -dioxin	6.80	2378df (3)	1,2,3,7,8- Pentachlorodibenzofuran	7.13
236dd (4)	2,3,6- Trichlorodibenzo- <i>p</i> -dioxin	6.66	234678df (4)	1,2,4,7,8- Pentachlorodibenzofuran	5.89
123478dd (5)	1,2,3,4,7,8- Hexachlorodibenzo- <i>p</i> -dioxin	6.55	12378df (5)	2,3,4,7,8- Pentachlorodibenzofuran	7.82
1378dd (6)	1,3,7,8- Tetrachlorodibenzo- <i>p</i> -dioxin	6.10	1237df (6)	1,2,3,4,7,8- Hexachlorodibenzofuran	6.64
12478dd (7)	1,2,4,7,8- Pentachlorodibenzo- <i>p</i> -dioxin	5.96	13478df (7)	1,2,3,6,7,8- Hexachlorodibenzofuran	6.57
1234dd (8)	1,2,3,4- Tetrachlorodibenzo- <i>p</i> -dioxin	5.89	23479df (8)	2,3,4,6,7,8- Hexachlorodibenzofuran	7.33
237dd (9)	2,3,7- Trichlorodibenzo- <i>p</i> -dioxin	7.15	23478df (9)	1,2,3,7- Tetrachlorodibenzofuran	6.96
28dd (10)	2,8- Dichlorodibenzo- <i>p</i> -dioxin	5.50	23478df (10)	1,3,4,7,8- Pentachlorodibenzofuran	6.70
12347dd (11)	1,2,3,4,7- Pentachlorodibenzo- <i>p</i> -dioxin	5.19	2348df (11)	2,3,4,7,9- Pentachlorodibenzofuran	6.70
124dd (12)	1,2,4- Trichlorodibenzo- <i>p</i> -dioxin	4.89	12379df (12)	2,3,4,7- Tetrachlorodibenzofuran	7.60
Ocdd (13)	1,2,3,4,6,7,8,9- Octachlorodibenzo- <i>p</i> -dioxin	4.30	12478df (13)	1,2,4,6,8- Pentachlorodibenzofuran	5.51
1mcdd (14)	1- Chlorodibenzo- <i>p</i> -dioxin	3.30	12468df (14)	1,2,3,7,9- Pentachlorodibenzofuran	6.40

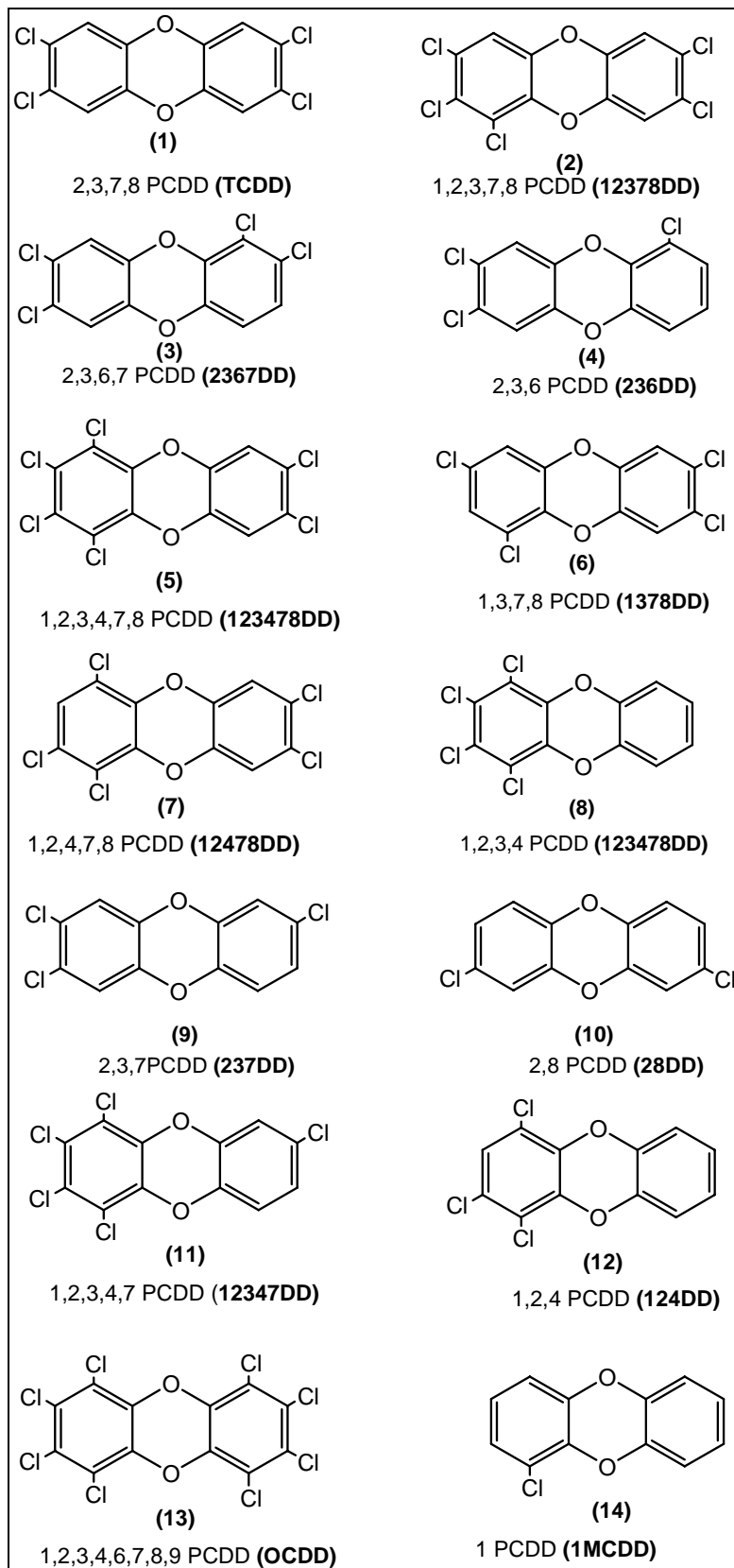


Figura 5.2.1. Derivats de dibenzo-p-dioxines utilitzats en el present treball

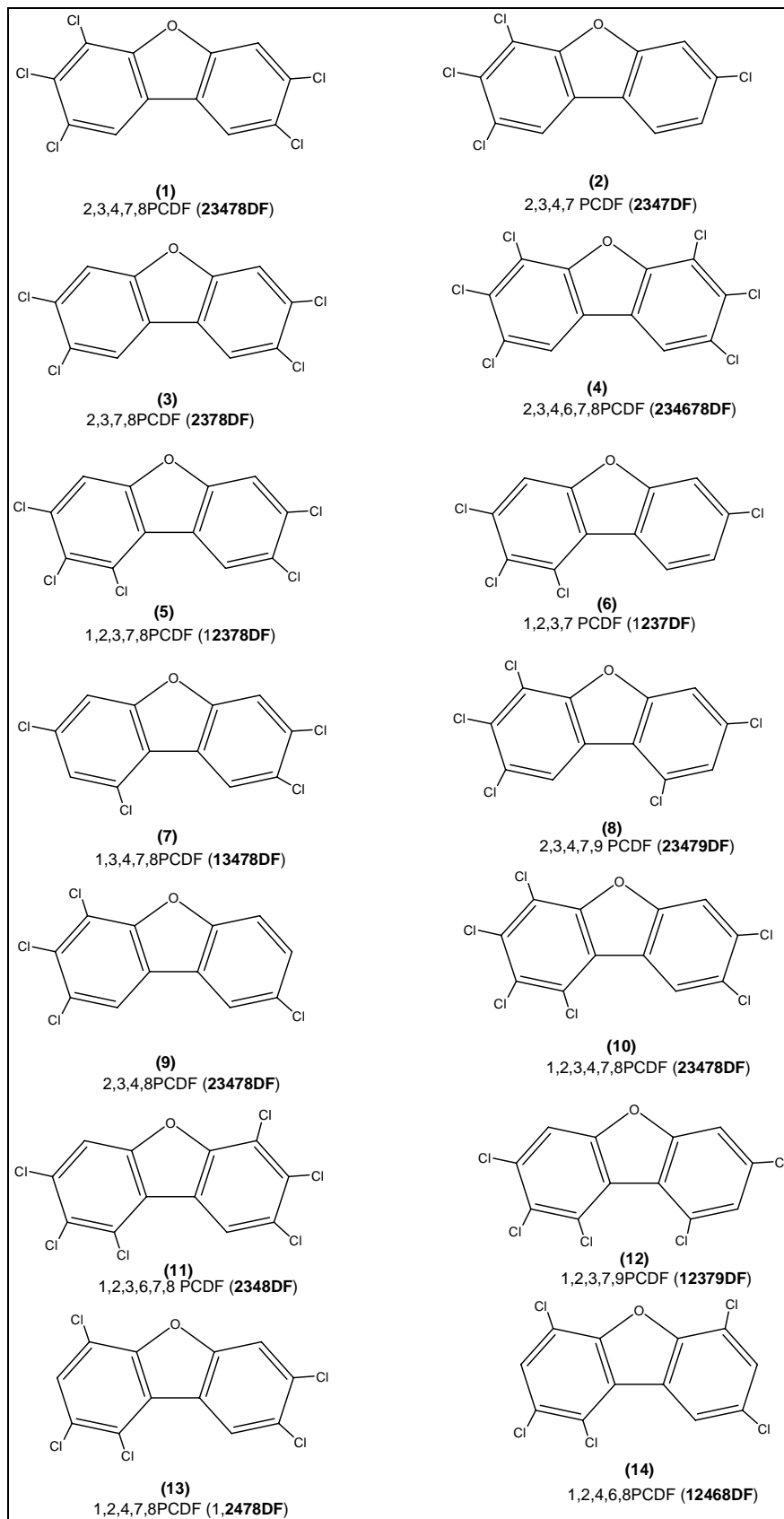


Figura 5.2.2. Derivats del furà utilitzats en el present treball

5.3 Càlcul dels potencials electrostàtic i de polarització

Com ja s'ha dit anteriorment, per calcular els diferents potencials s'ha seguit un procediment que implica triar un mètode de càlcul de la funció d'ona; un mètode de càlcul de l'energia d'inducció i un conjunt de funcions de base. En aquest cas s'ha emprat el mètode de Hartree-Fock per el càlcul de la funció d'ona utilitzant la base 9s5p de Dunning (Dunning, 1970), amb contracció a una base *split-valence* (DZ) (Dunning, 1976) a la que s'ha afegit un conjunt de funcions de polarització optimitzades per donar bons valors de polaritzabilitats (DZP') (Duijneveldt, 1982). Els càlculs dels potencials electrostàtic i de polarització han estat portats a terme sobre les estructures moleculars optimitzades sobre una superfície plana a 4 Å del pla molecular, utilitzant el programa Gaussian 94 (Frisch, 1995). El potencial de polarització s'ha calculat mitjançant la teoria de pertorbacions de Rayleigh-Schrödinger utilitzant el programa Mephisto (Francl, 1985).

5.4 Mapes de potencial electrostàtic, de polaritzabilitat i interacció.

5.4.1 *Descripció general*

Els mapes de potencial electrostàtic, de polarització i d'interacció es mostren per cada molècula a les Figures 5.4.1.-5.4.28. Els mapes s'han dibuixat amb un codi de colors de forma que les regions positives estan dibuixades en blau i les negatives en vermell. La representació tridimensional d'aquests mapes ens permet observar més clarament les distribucions espaials. Cal indicar que les gràfiques dels potencials de polarització, com els valors són sempre negatius, s'acolorixen en vermell.

Els mapes de potencial electrostàtic presenten mínims al voltant dels clorurs, de les àrees de densitat electrònica negativa dels anells aromàtics i a sobre dels oxígens, encara que aquests últims són una mica més febles. És interessant observar que depenent de la distribució de clorurs a la molècula es modula l'aparició de mínims sobre la distribució electrònica aromàtica. Per exemple, si comparem dues molècules molt diferents como el octaclorodibenzo-*p*-dioxina a la Figura 5.4.1 i la 1-clorodibenzo-*p*-dioxina (Figura 5.4.3) es pot veure que al primer cas la contribució de la densitat electrònica dels anells aromàtics als mínims és petita, mentre que al cas del derivat monoclorat s'observen mínims al voltant de la densitat electrònica dels anells aromàtics. Si bé els mínims apareixen al mateix lloc als mapes d'interacció, l'alta polaritzabilitat del clorur en comparació de l'oxigen fa que els primers siguin molt més importants. Observant la representació gràfica dels diferents mapes de potencial electrostàtic, potencial de polarització, i energia d'interacció, podem comprovar que la inclusió de la polarització molecular

permet definir més clarament les diferents zones d'interacció. Deduïm doncs la importància del potencial de polarització en la descripció dels MEPs i per tant, a la predicció de la reactivitat de les nostres molècules.

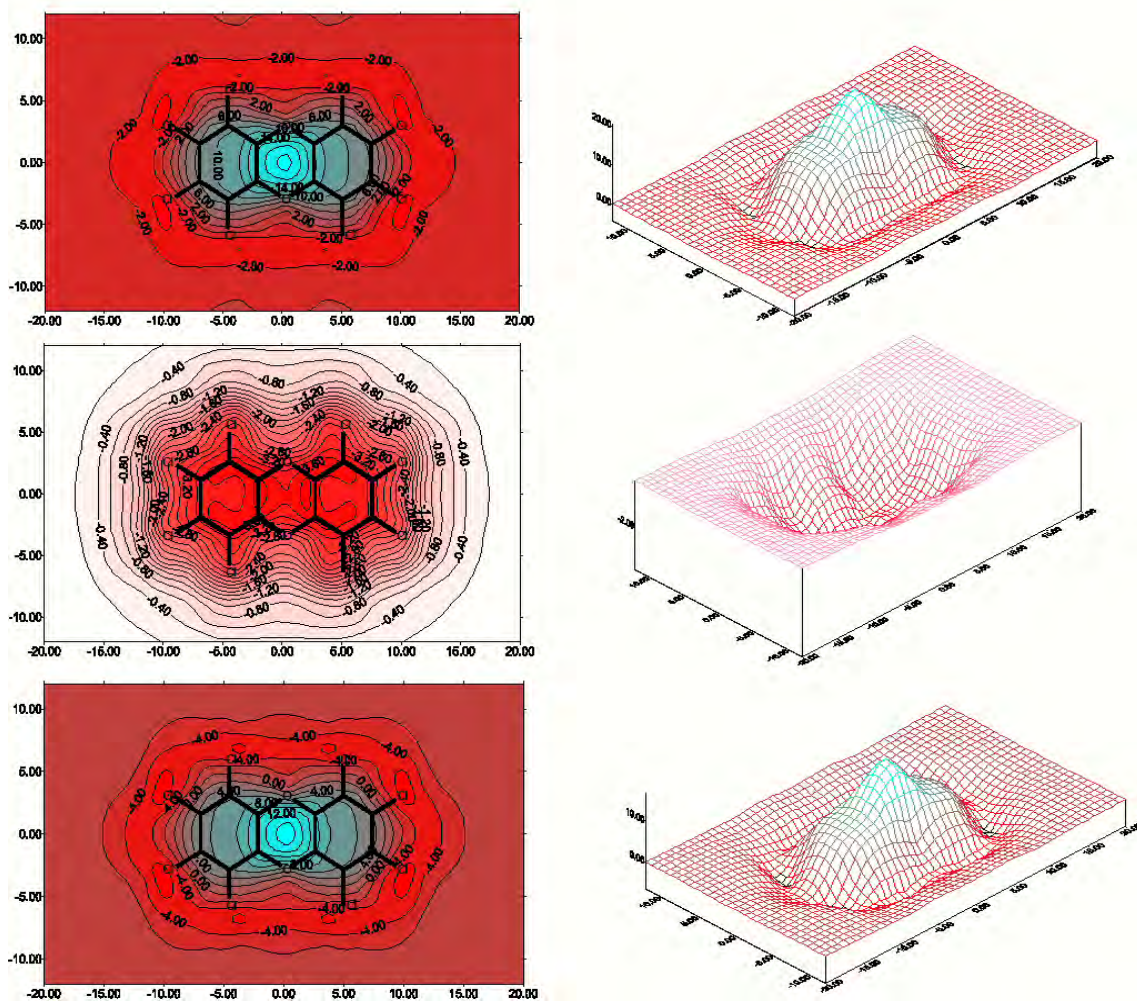


Figura 5.4.1. Mapes de Potencial electrostàtic, de polarització i d'interacció per la dioxina OCDD

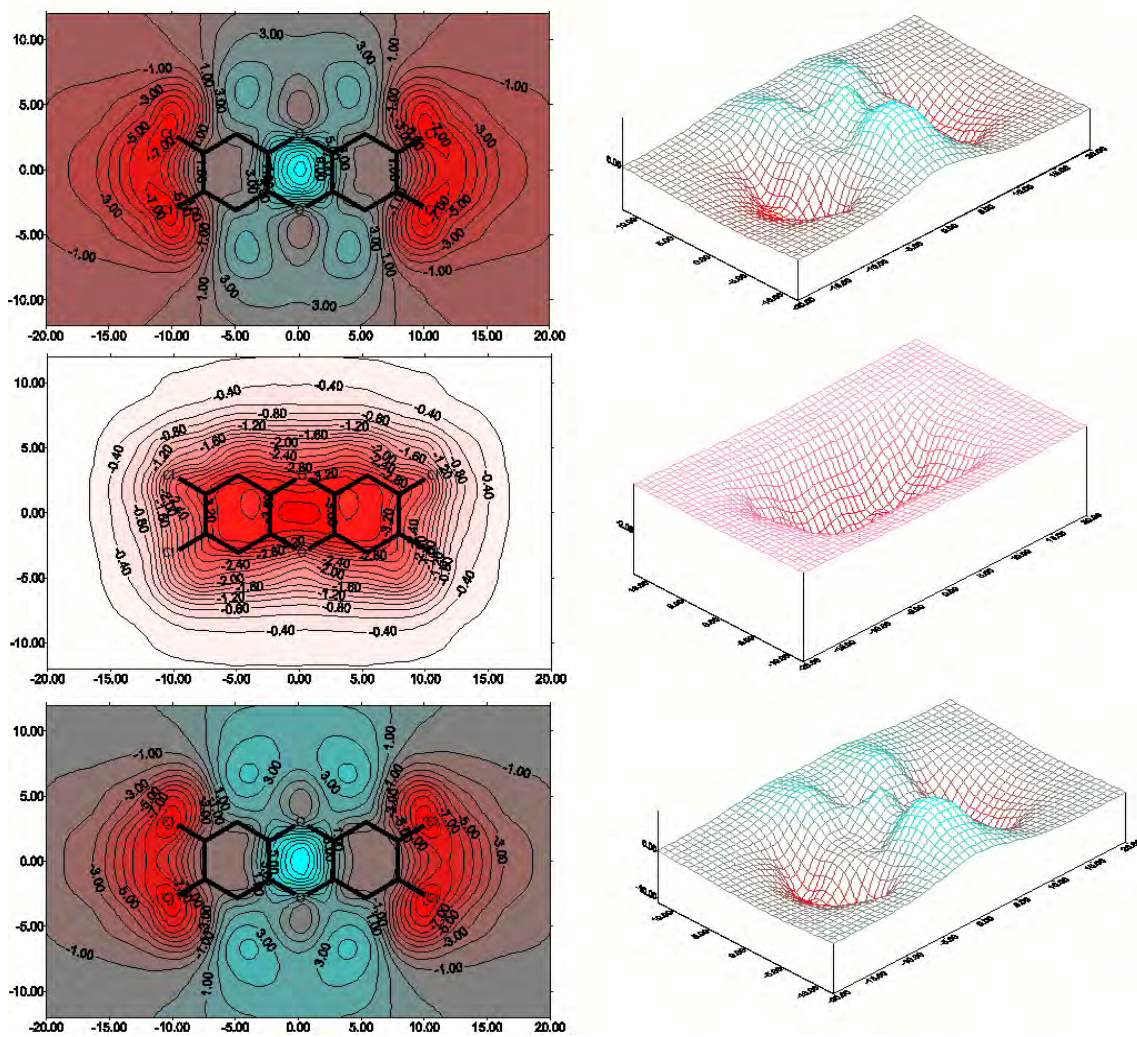


Figura 5.4.2. Mapes de Potencial electrostàtic, de polarització i d'interacció per la dioxina TCDD

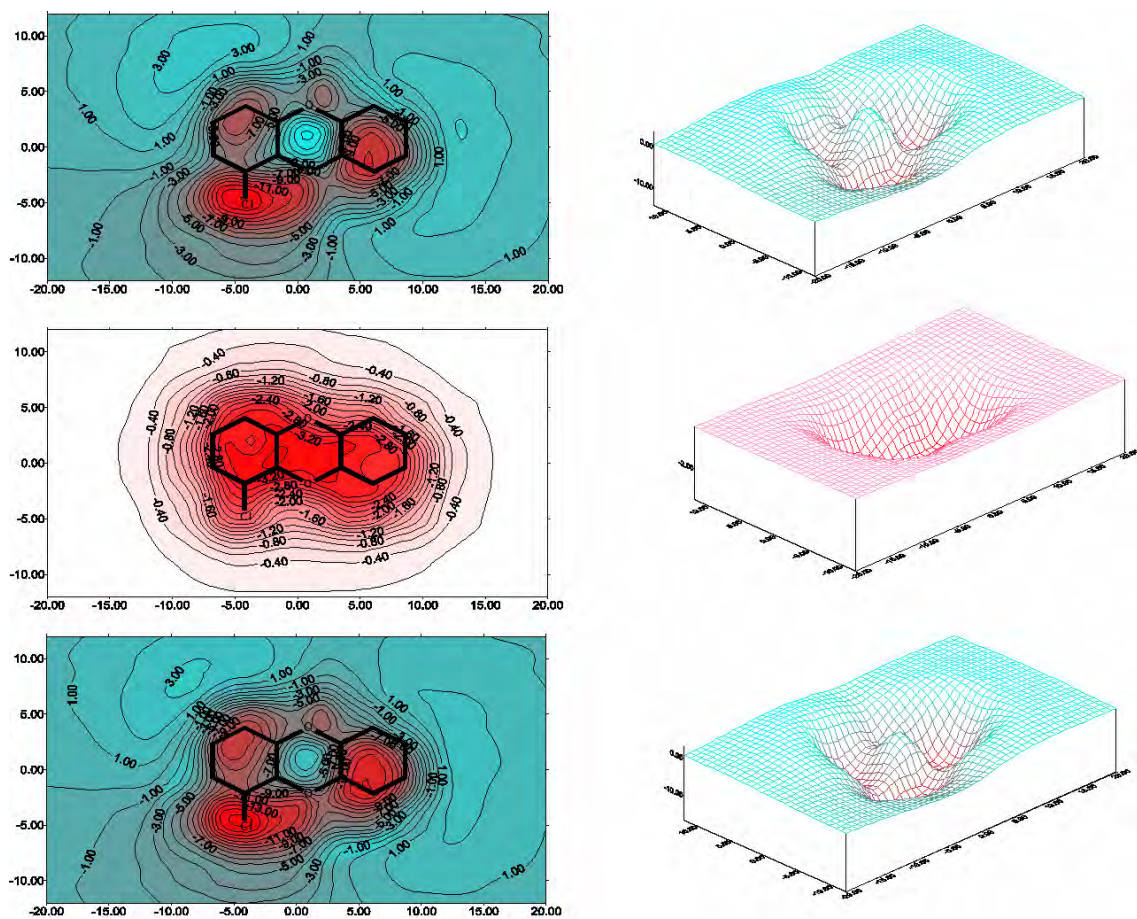


Figura 5.4.3. Mapes de Potencial electrostàtic, de polarització i d'interacció per la dioxina 1MCDD

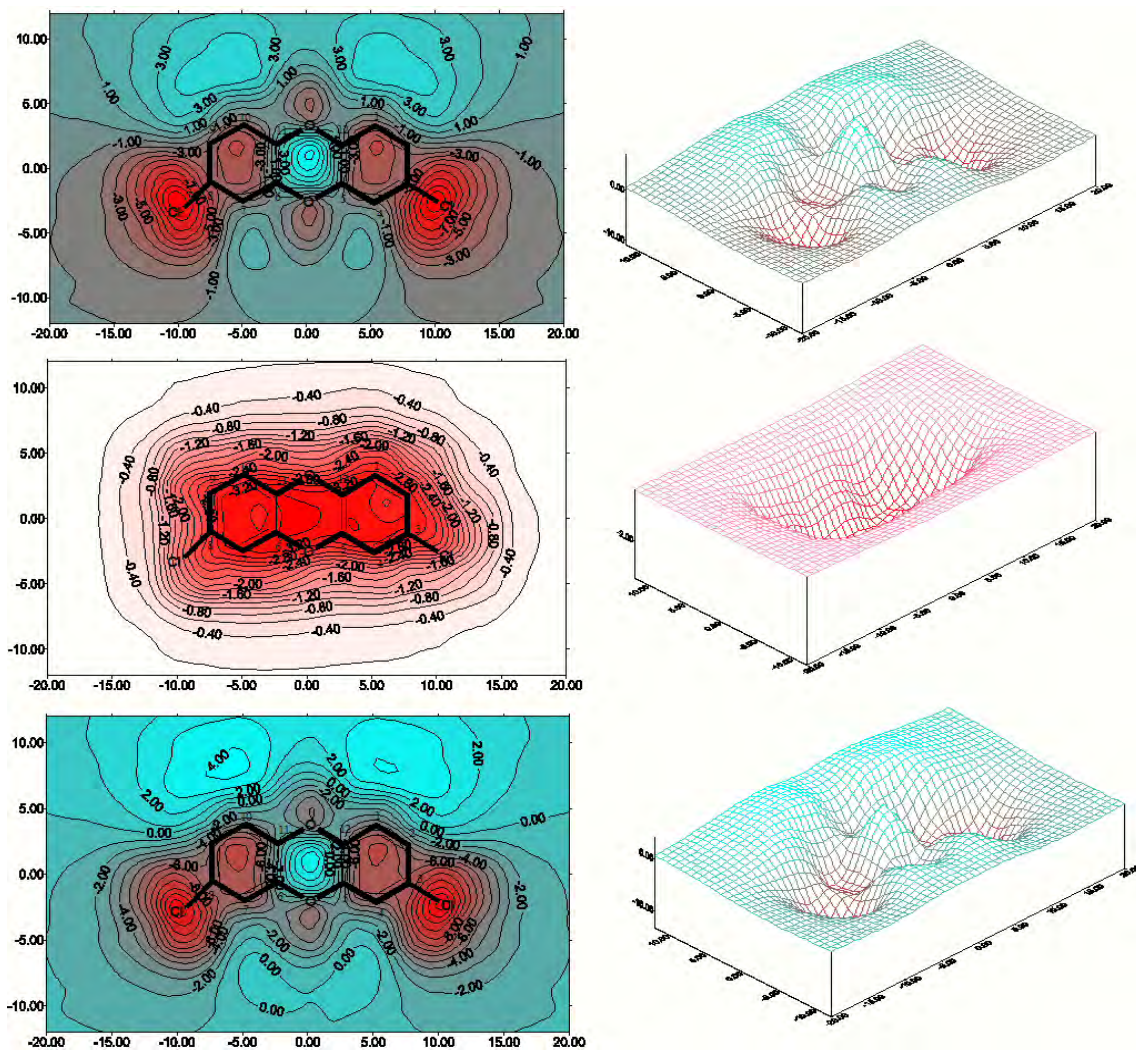


Figura 5.4.4. Mapes de Potencial electrostàtic, de polarització i d'interacció per la dioxina 28DD

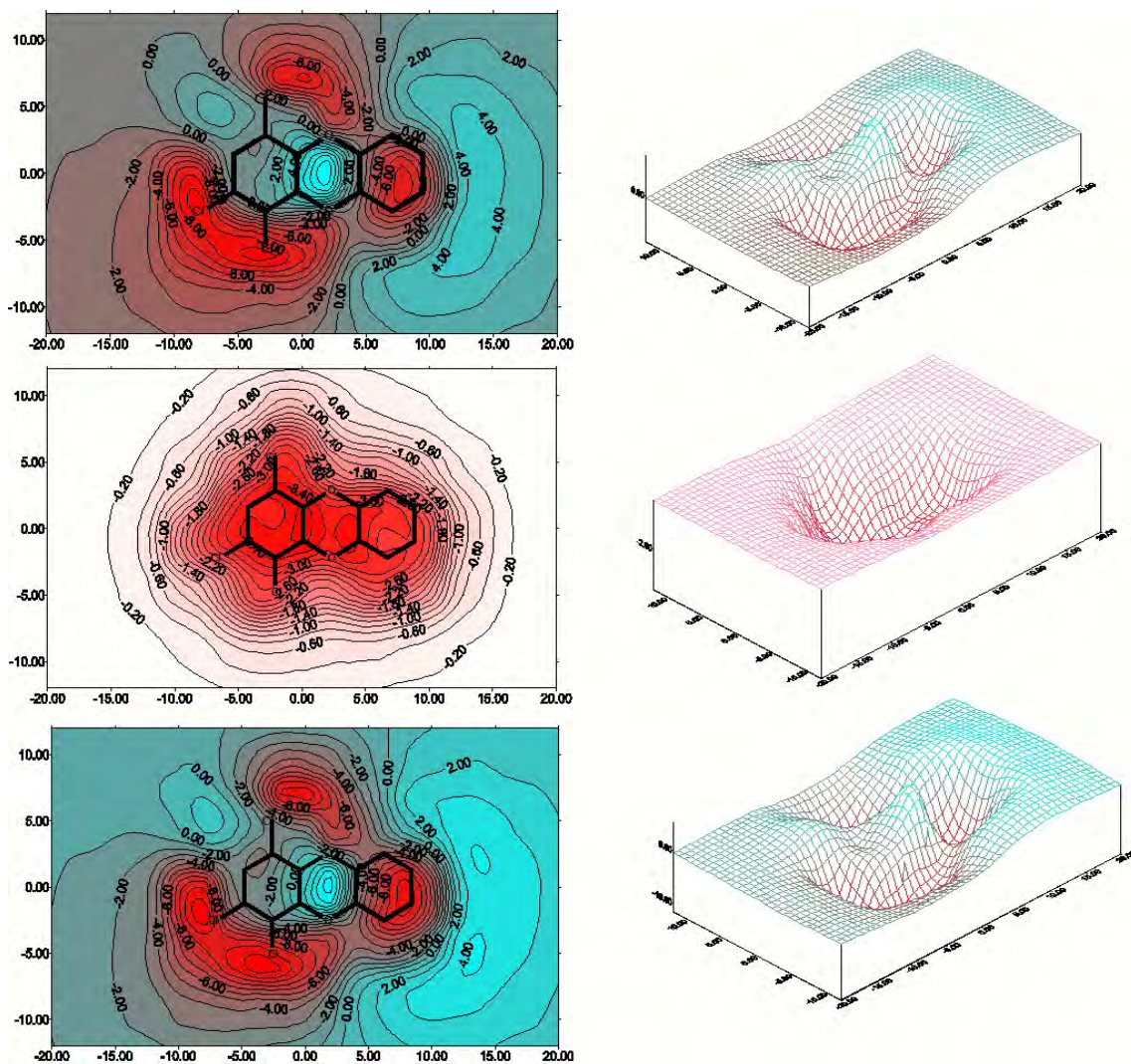


Figura 5.4.5. Mapes de Potencial electrostàtic, de polarització i d'interacció per la dioxina 124DD

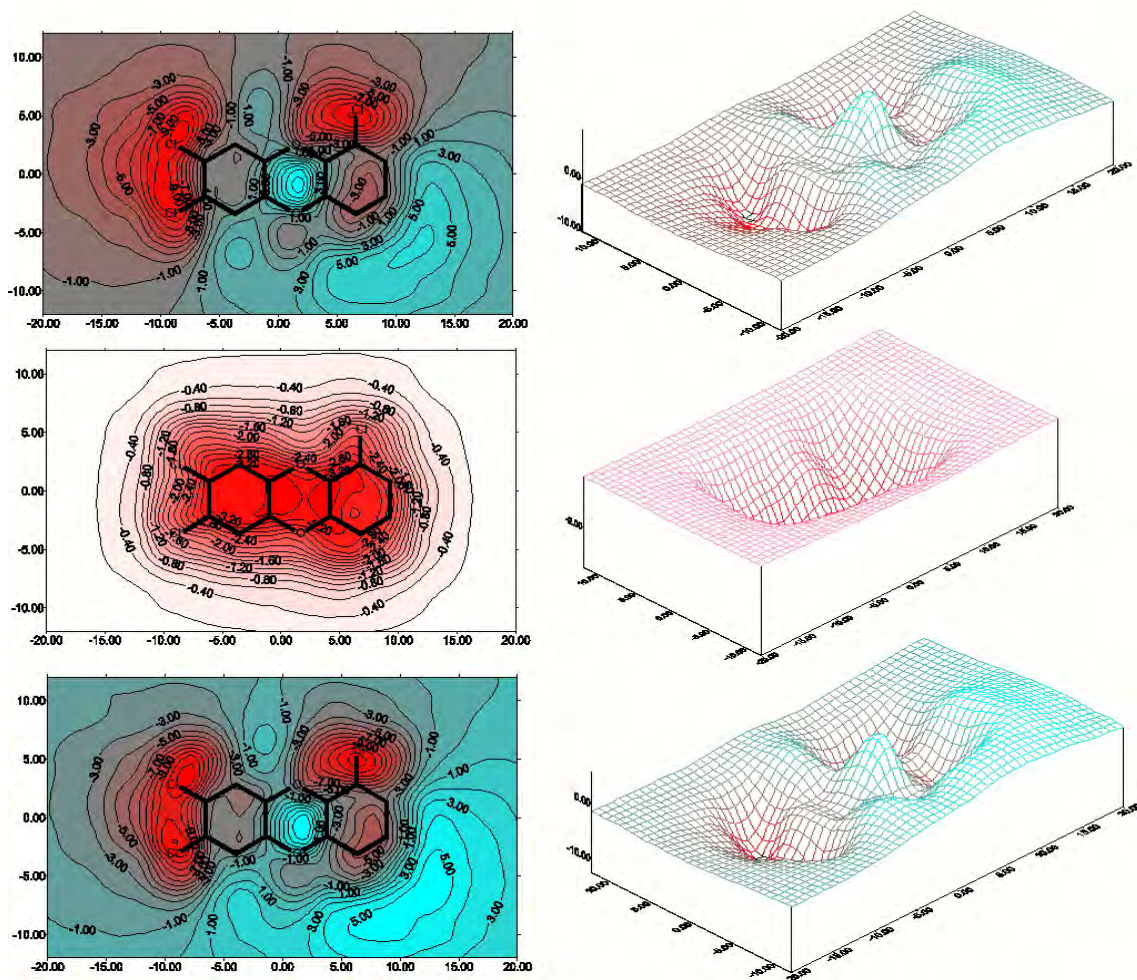


Figura 5.4.6. Mapes de Potencial electrostàtic, de polarització i d'interacció per la dioxina 236DD

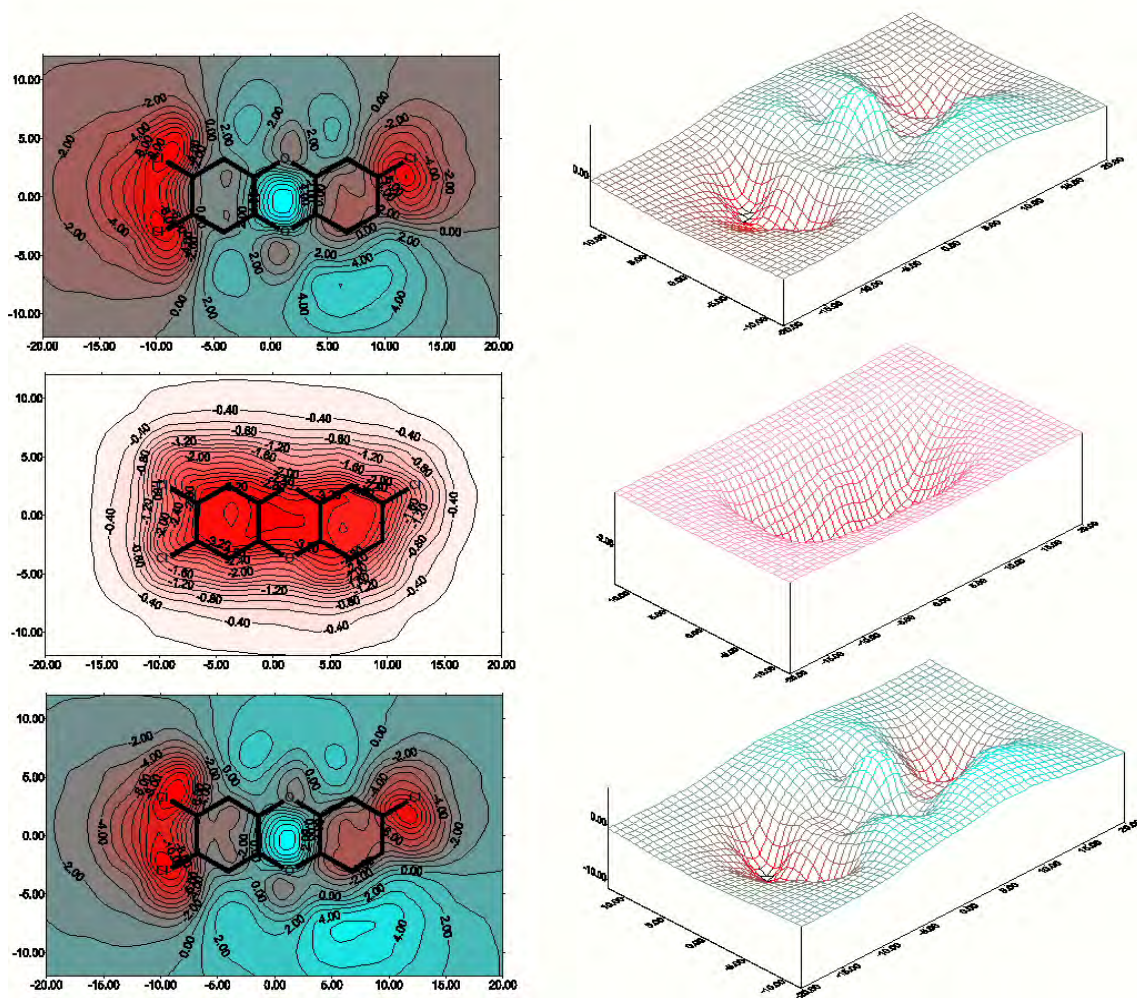


Figura 5.4.7. Mapes de Potencial electrostàtic, de polarització i d'interacció per la dioxina 237DD

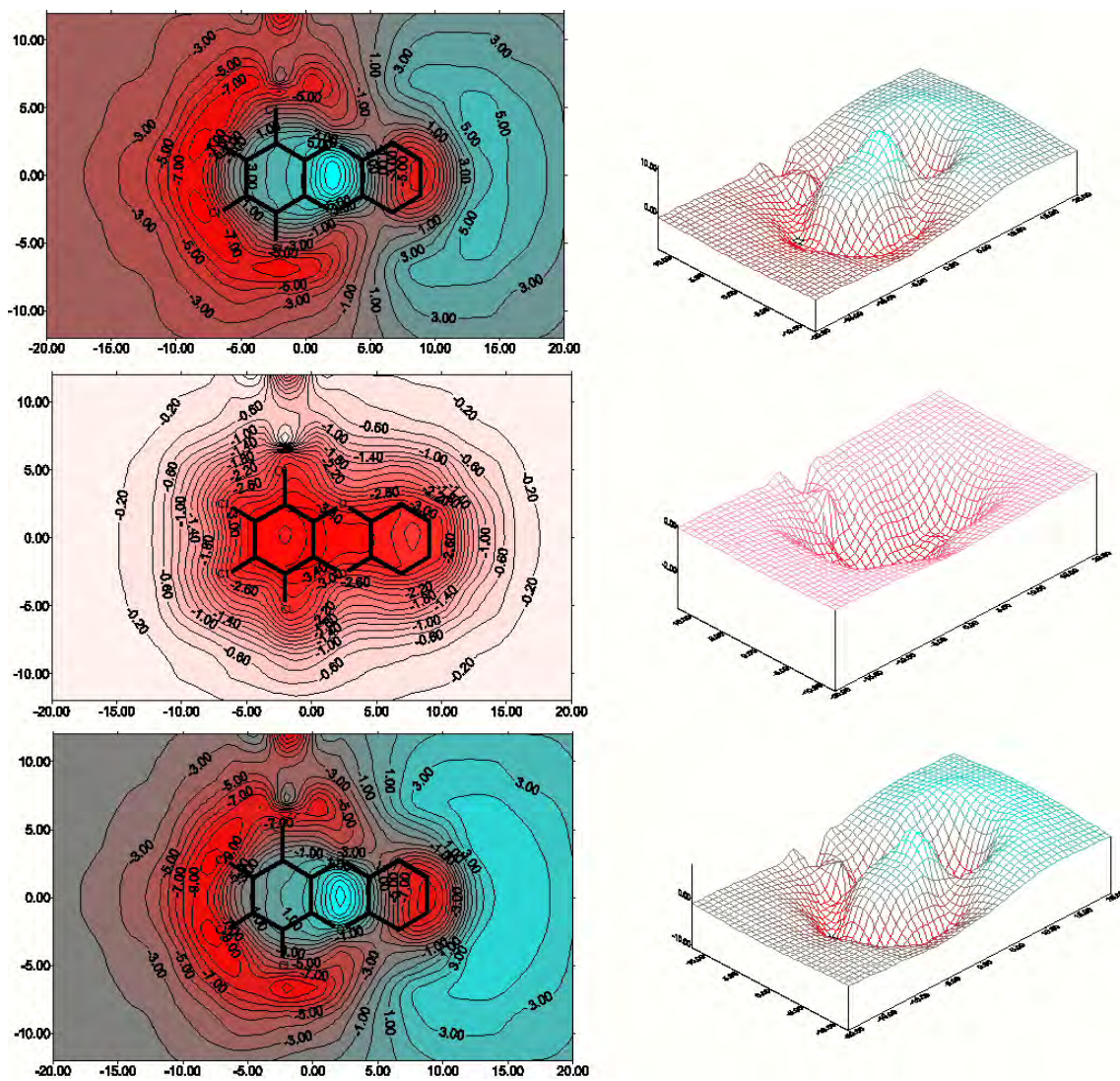


Figura 5.4.8. Mapes de Potencial electrostàtic, de polarització i d'interacció per la dioxina 1234DD

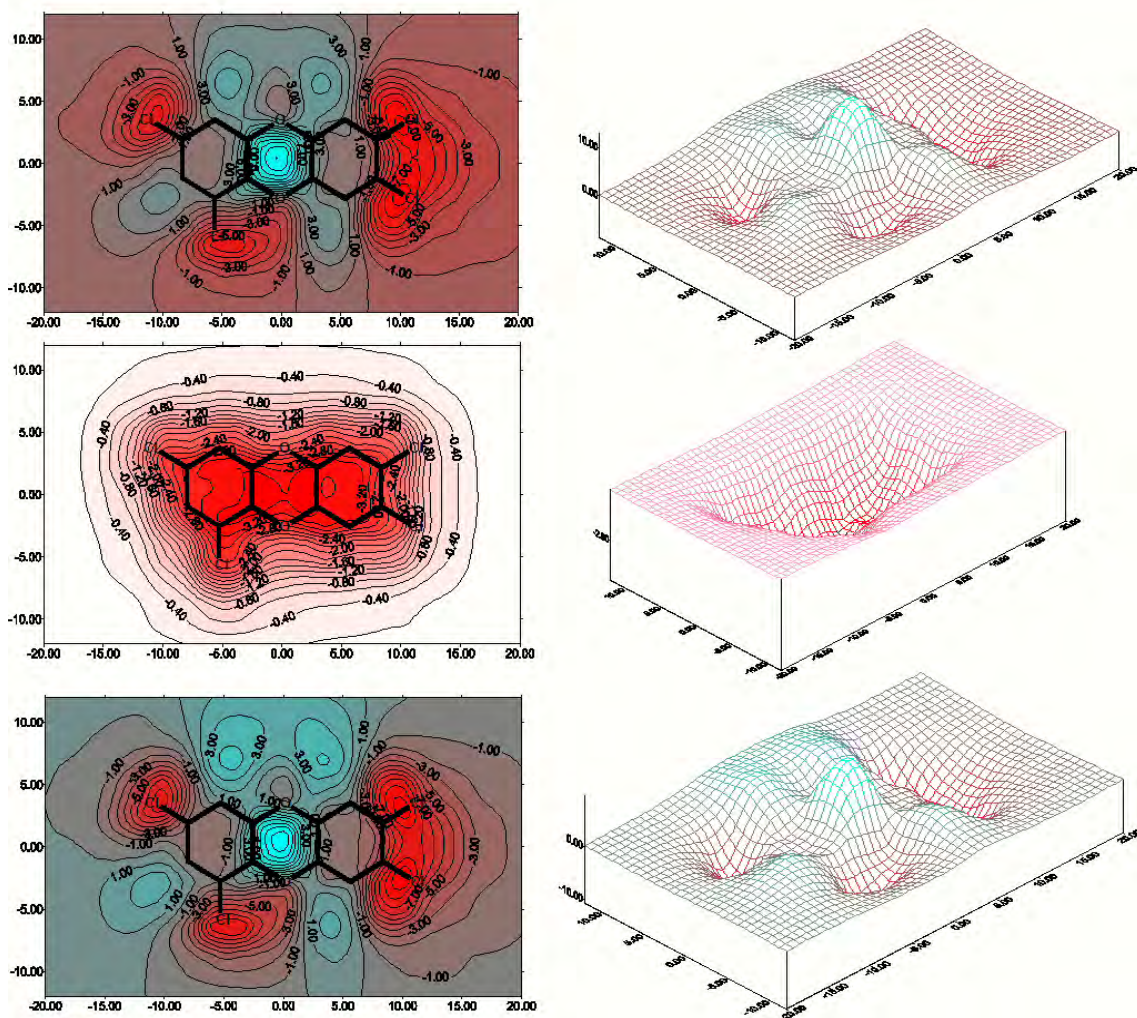


Figura 5.4.9. Mapes de Potencial electrostàtic, de polarització i d'interacció per la dioxina 1378DD

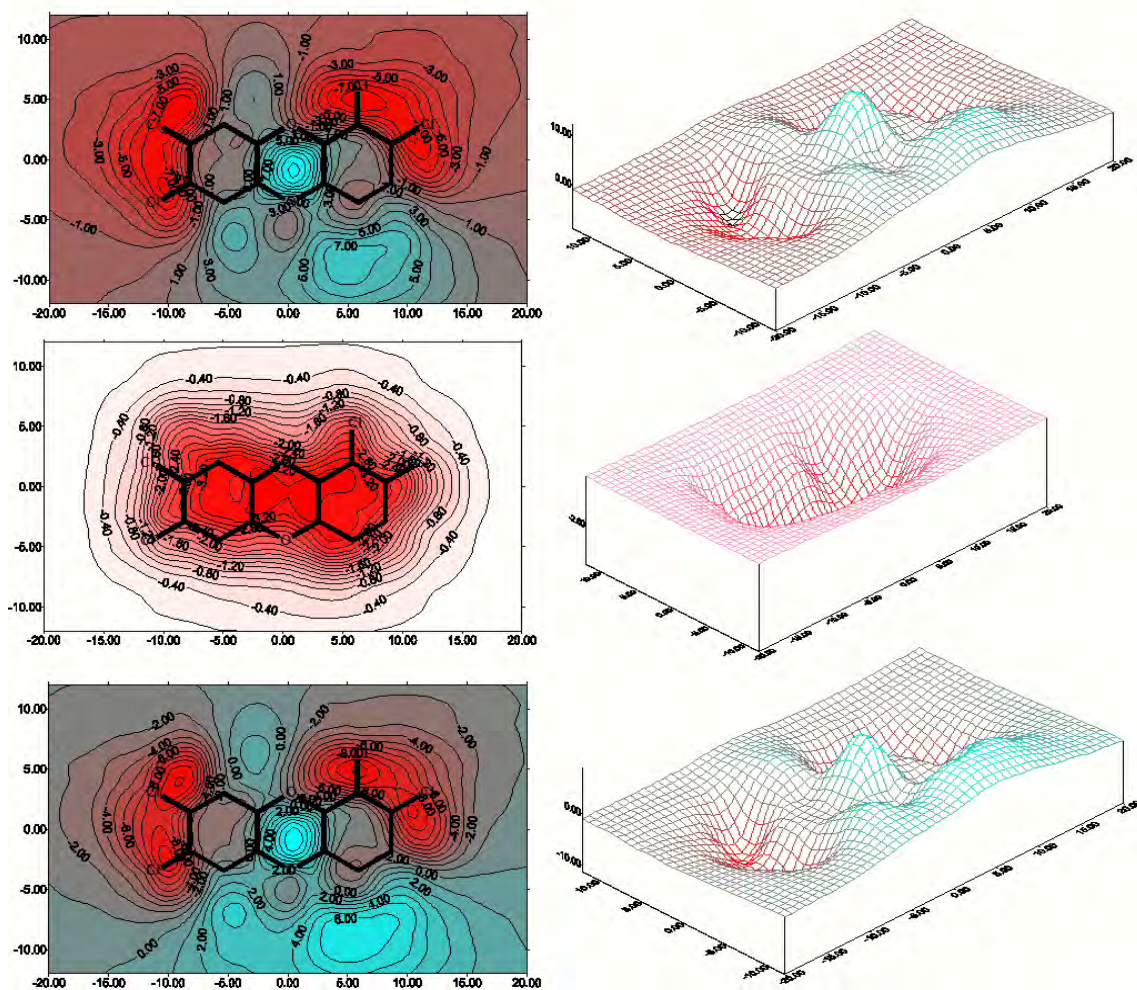


Figura 5.4.10. Mapes de Potencial electrostàtic, de polarització i d'interacció per la dioxina 2367DD

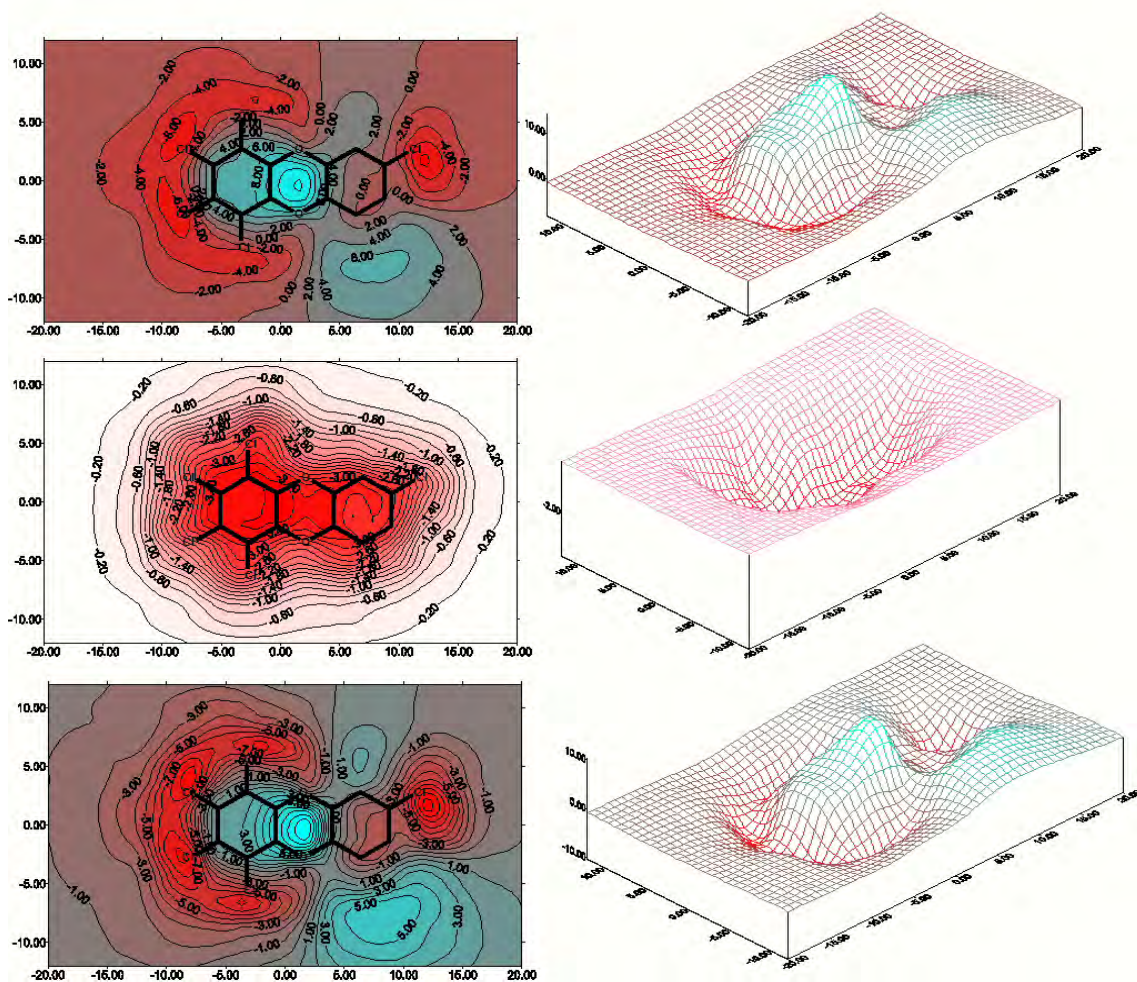


Figura 5.4.11. Mapes de Potencial electrostàtic, de polarització i d'interacció per la dioxina 12347DD

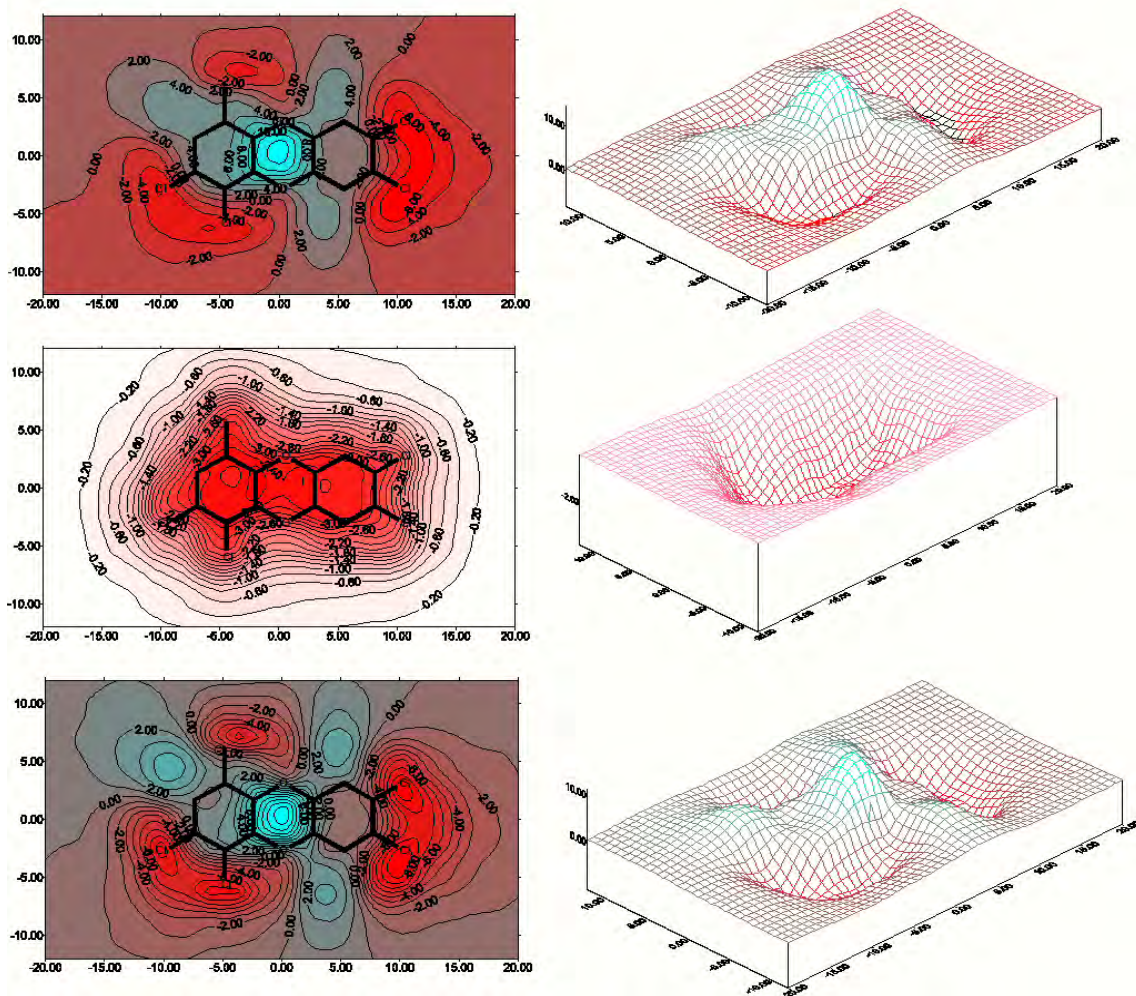


Figura 5.4.12. Mapes de Potencial electrostàtic, de polarització i d'interacció per la dioxina 12478DD

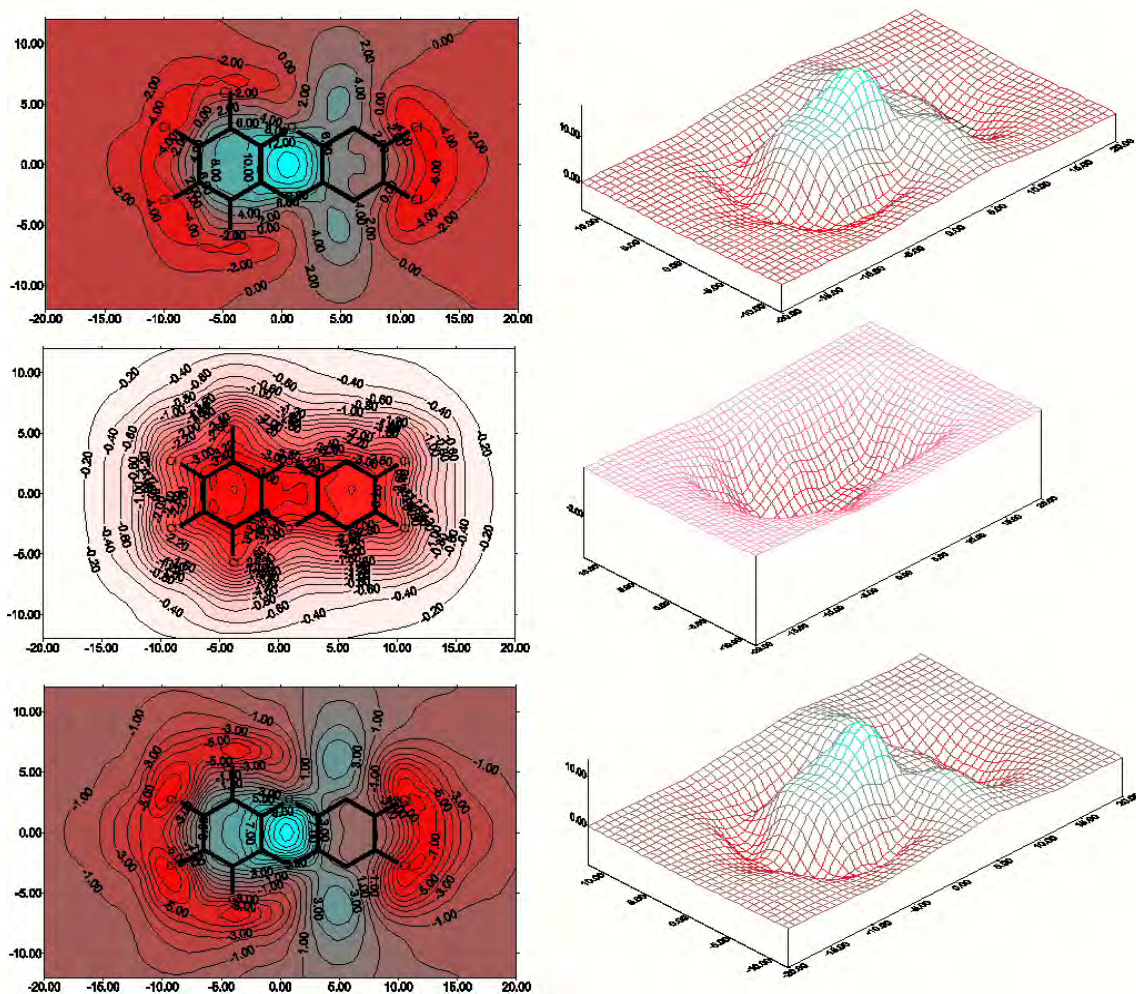


Figura 5.4.13. Mapes de Potencial electrostàtic, de polarització i d'interacció per la dioxina 123478DD

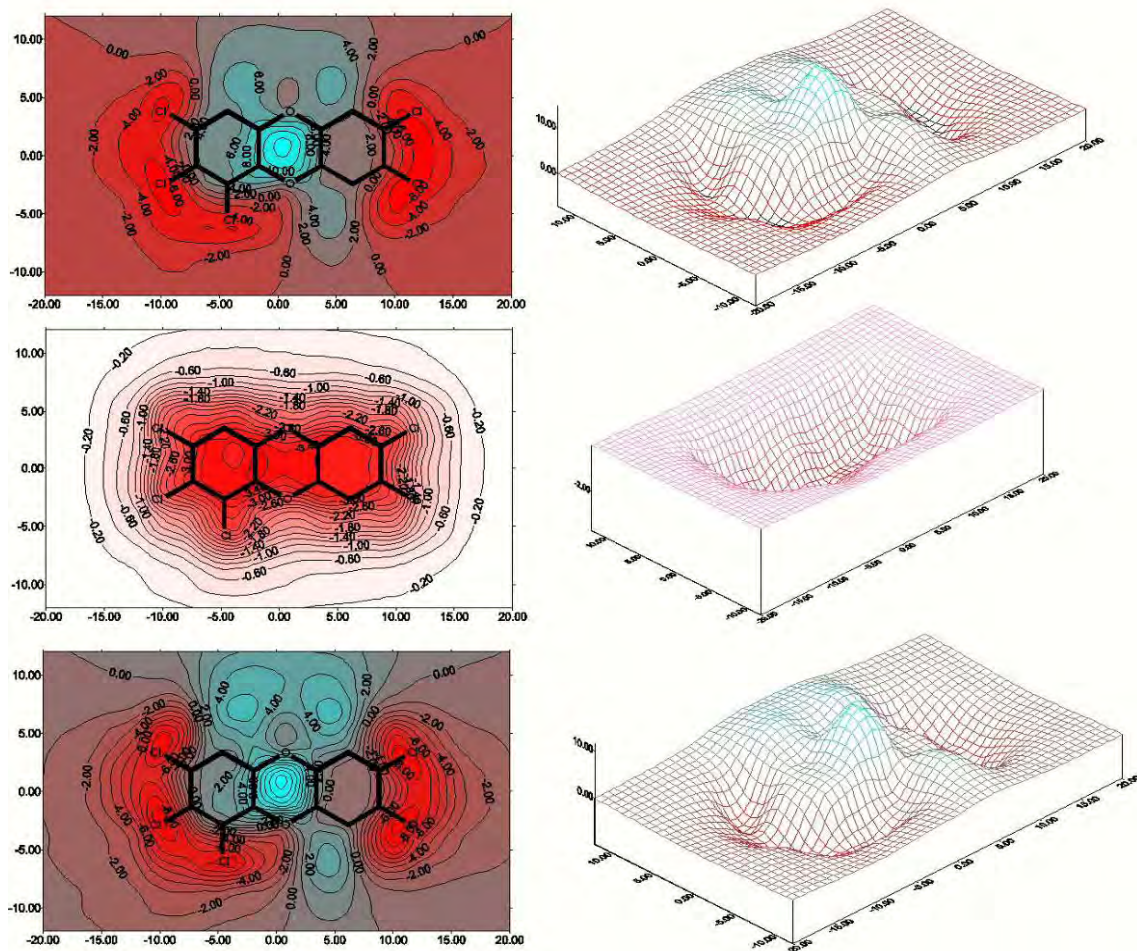


Figura 5.4.14. Mapes de Potencial electrostàtic, de polarització i d'interacció per la dioxina 12378DD

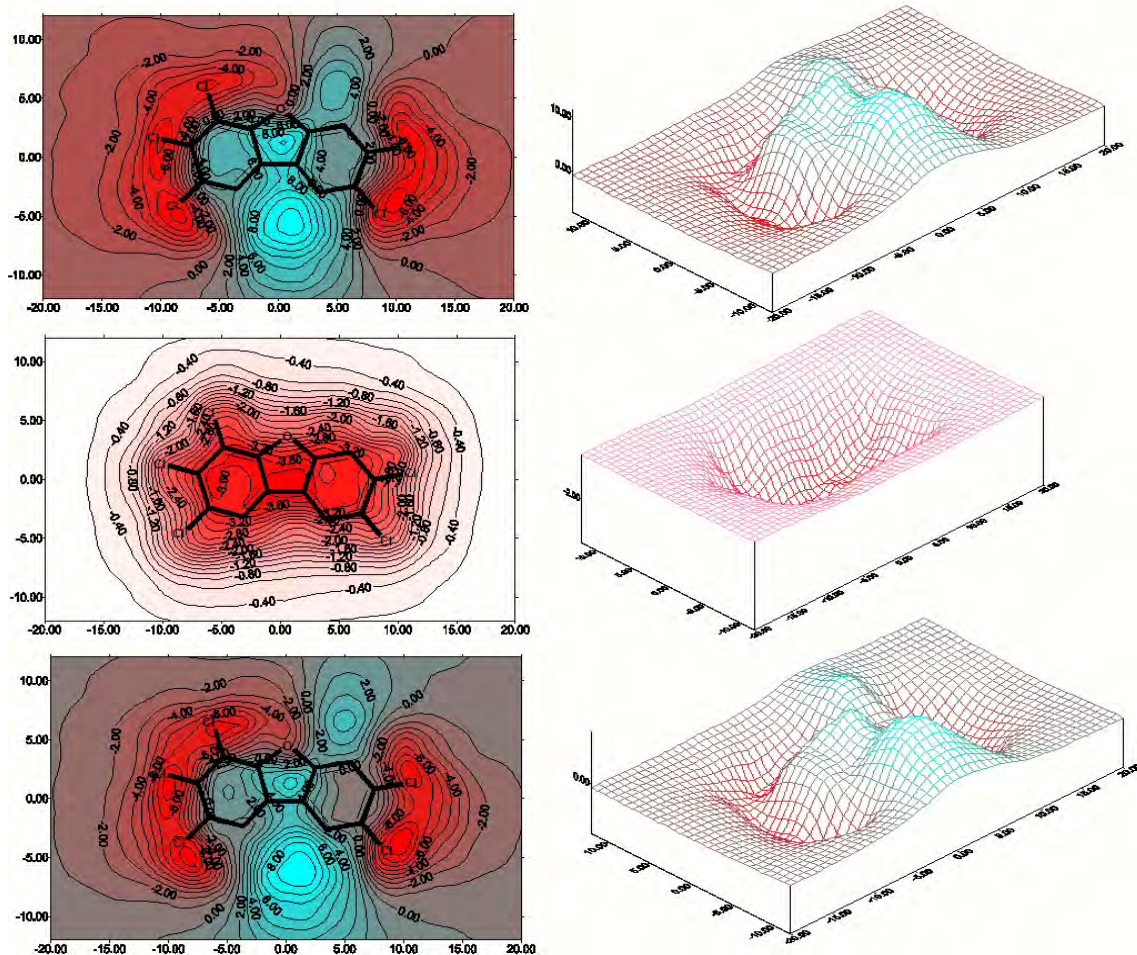


Figura 5.4.15. Mapes de Potencial electrostàtic, de polarització i d'interacció per el furà 23478df

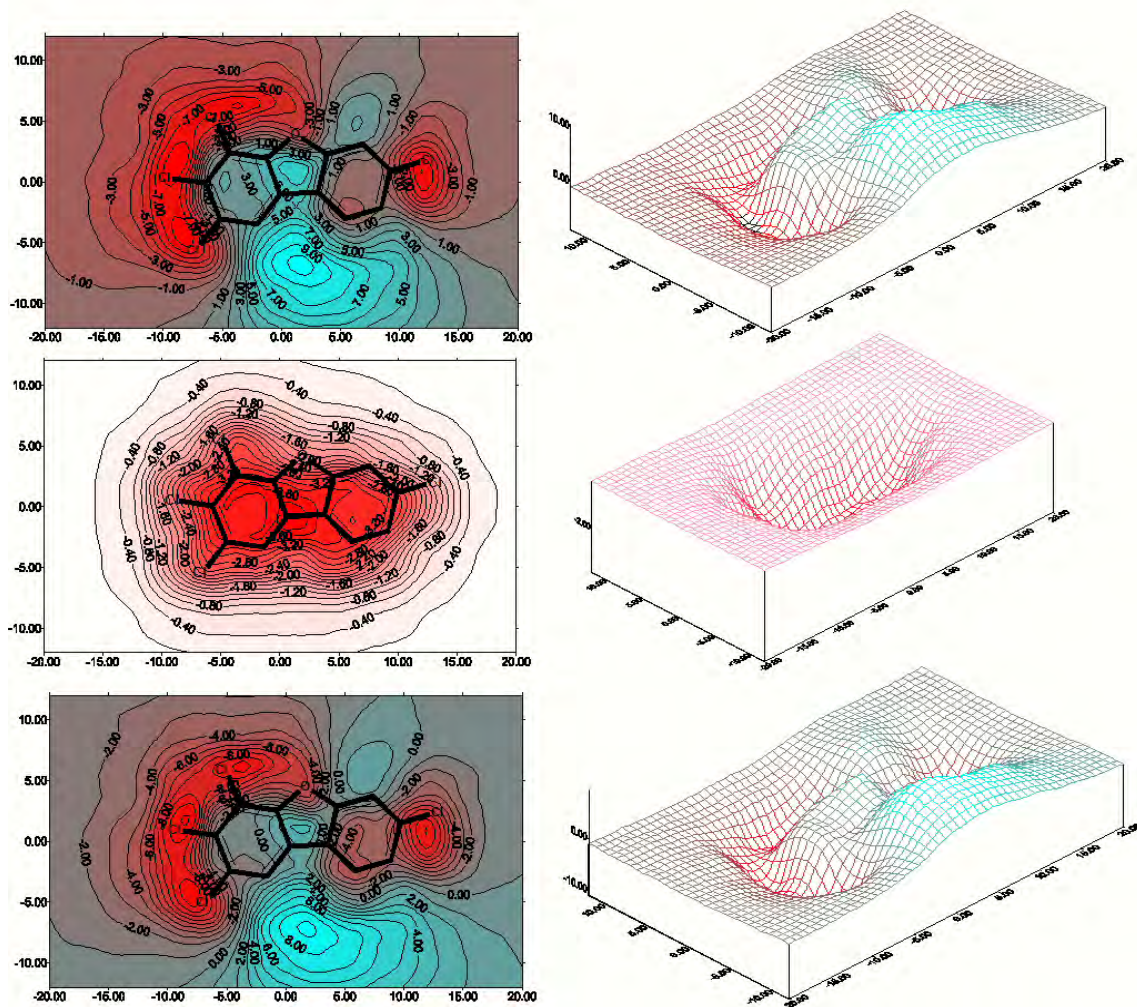


Figura 5.4.16. Mapes de Potencial electrostàtic, de polarització i d'interacció per el furà 2347df

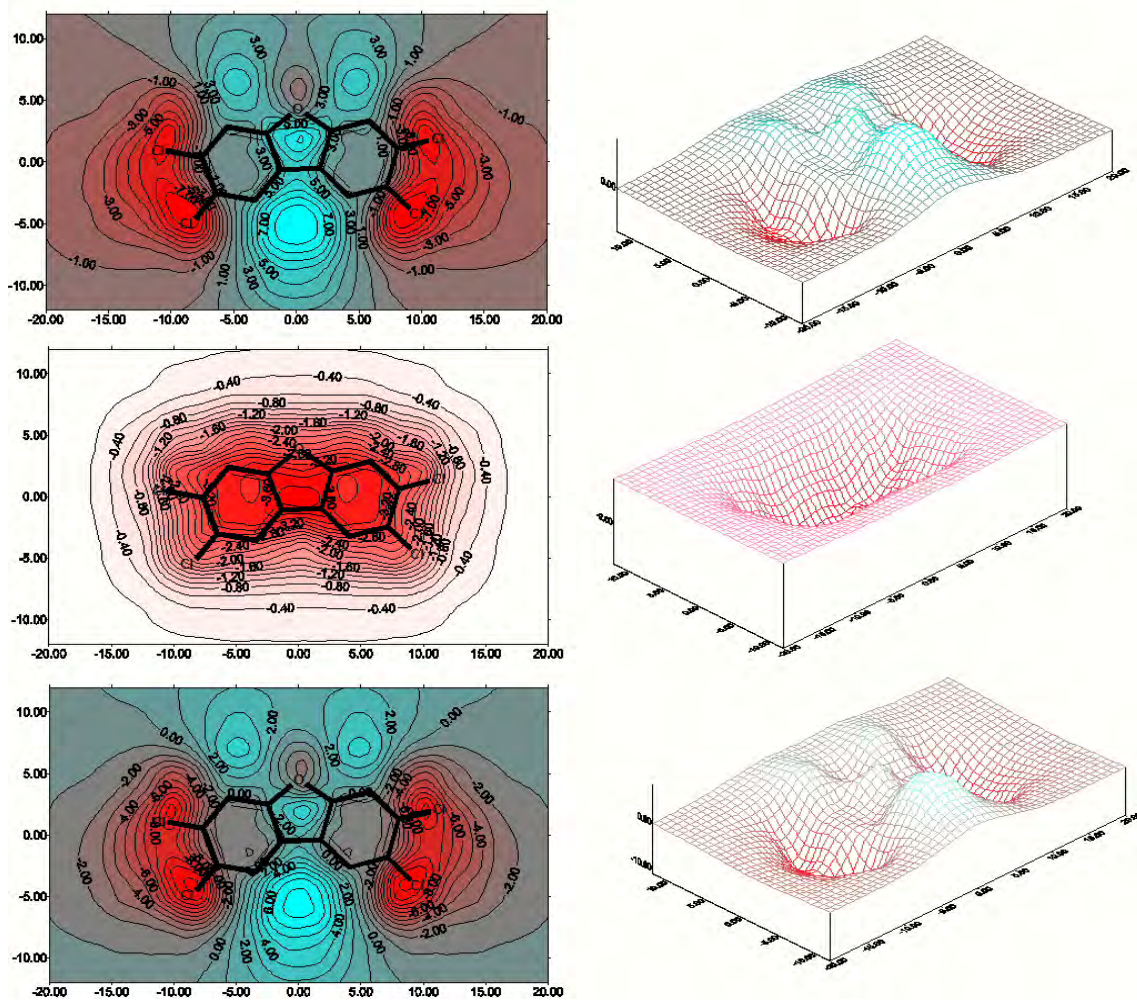


Figura 5.4.17. Mapes de Potencial electrostàtic, de polarització i d'interacció per el furà 2378df

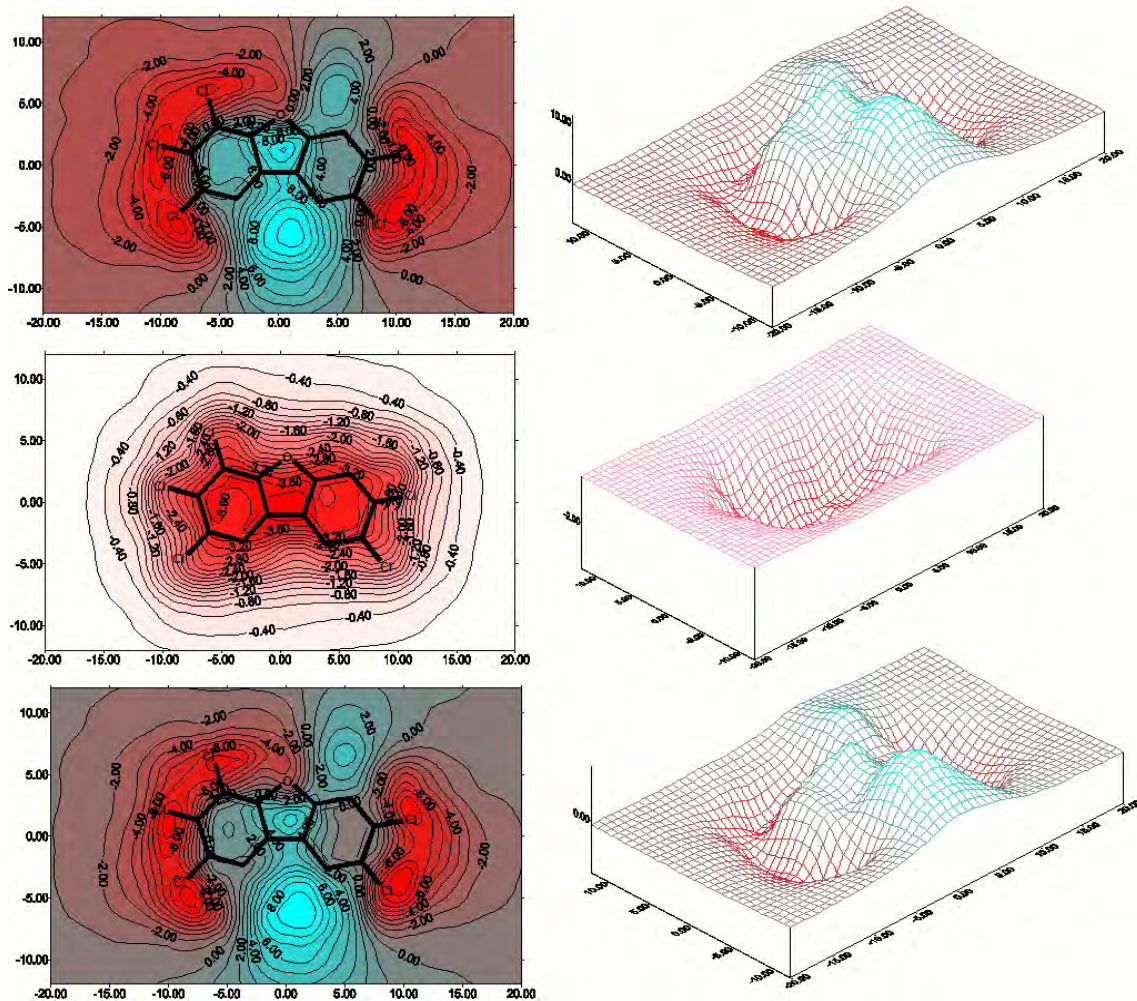


Figura 5.4.18. Mapes de Potencial electrostàtic, de polarització i d'interacció per el furà 23478df

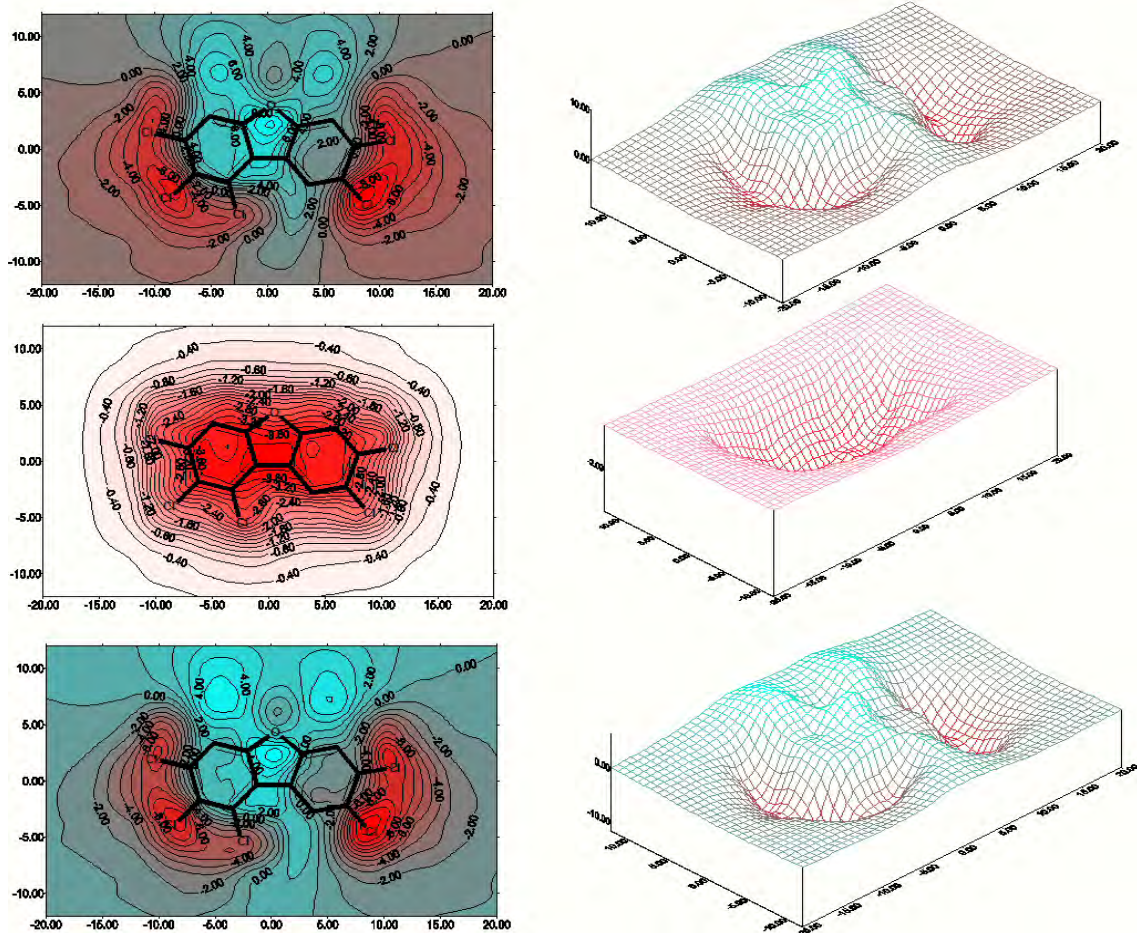


Figura 5.4.19. Mapes de Potencial electrostàtic, de polarització i d'interacció per el furà 12378df

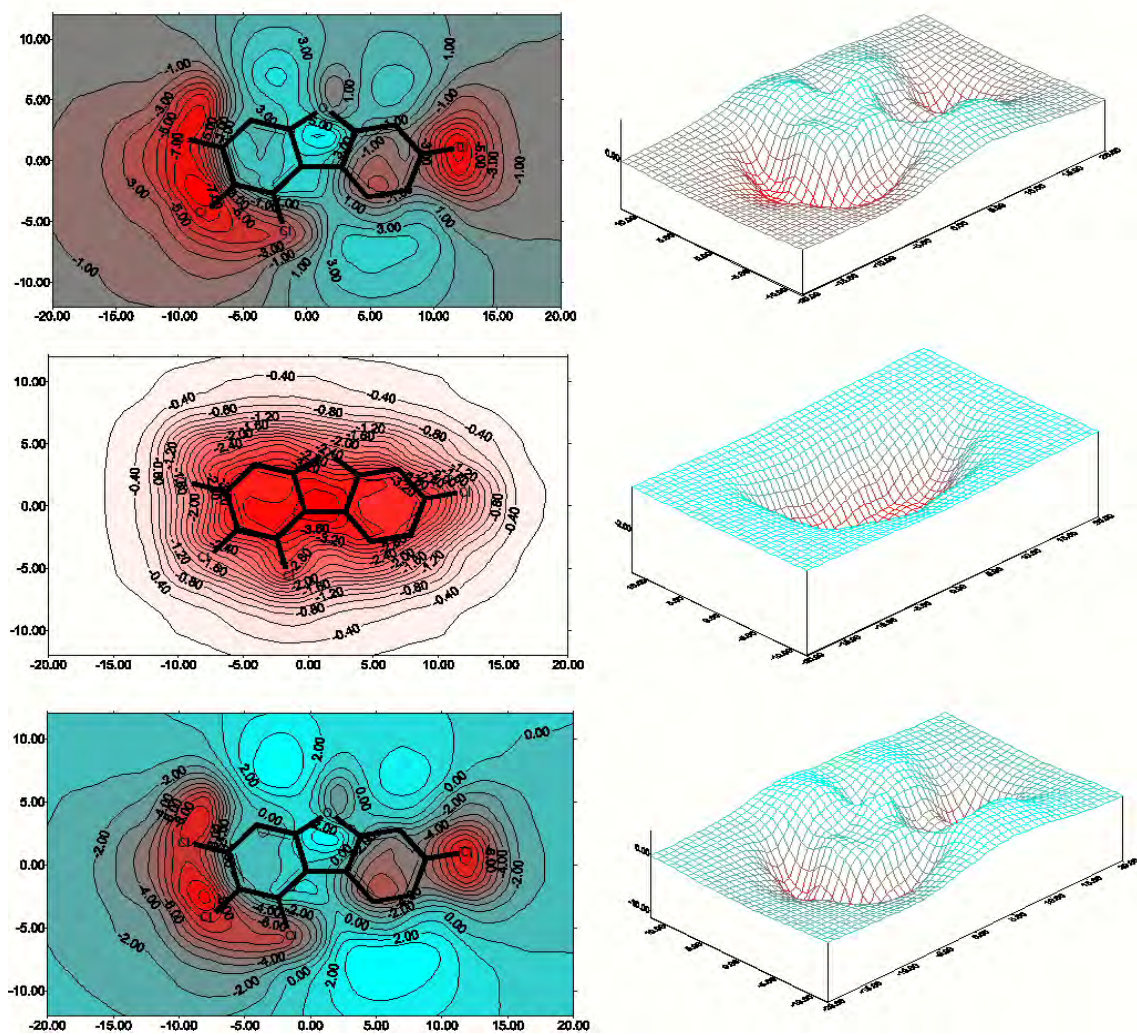


Figura 5.4.20. Mapes de Potencial electrostàtic, de polarització i d'interacció per el furà 1237df

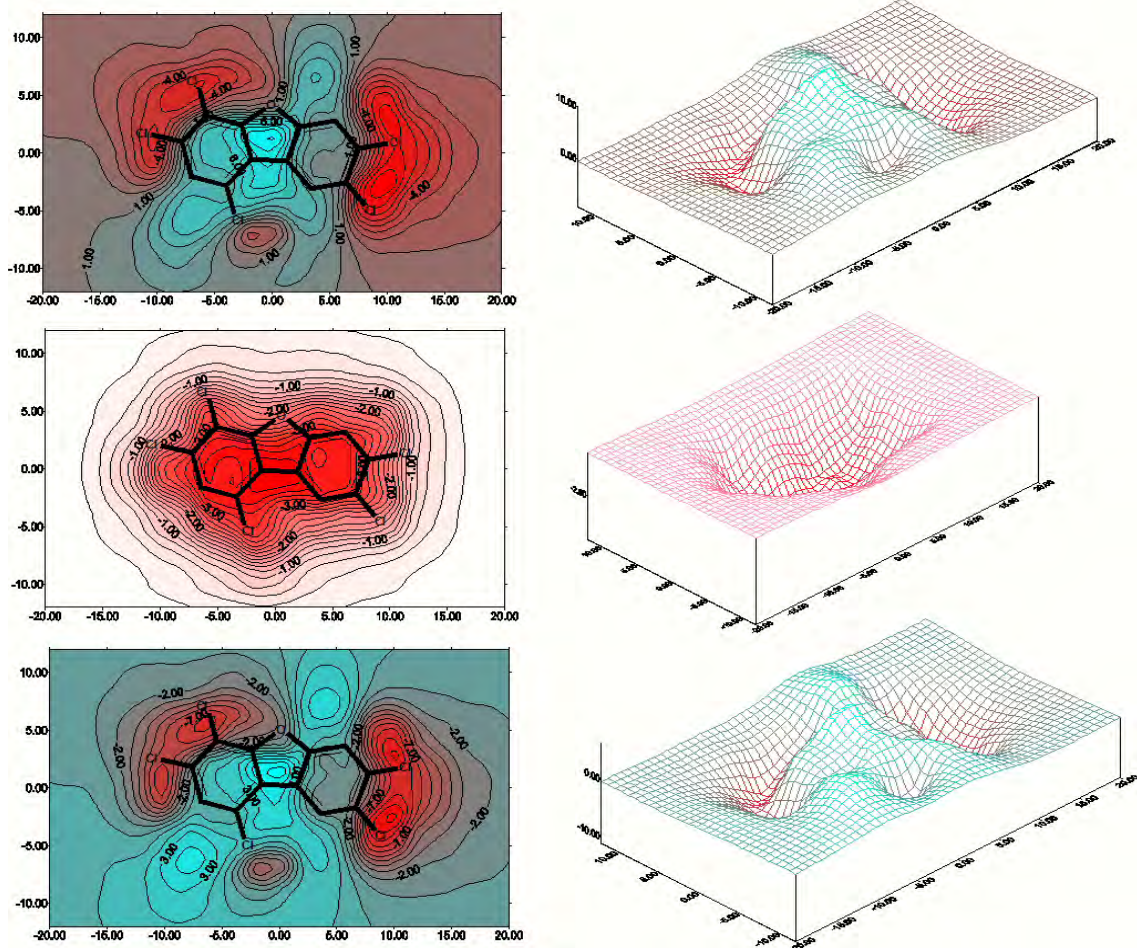


Figura 5.4.21. Mapes de Potencial electrostàtic, de polarització i d'interacció per el furà 13478df

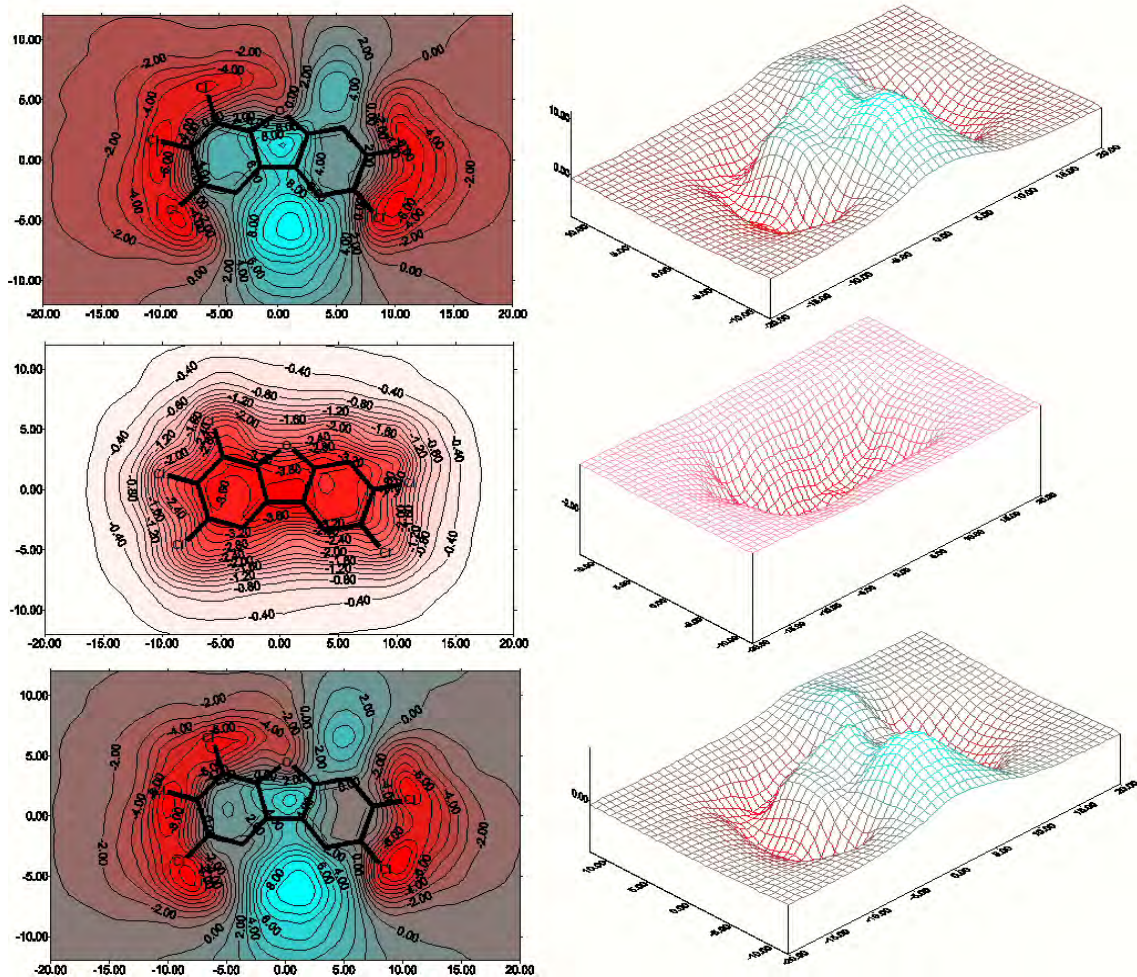


Figura 5.4.22. Mapes de Potencial electrostàtic, de polarització i d'interacció per el furà 23478df

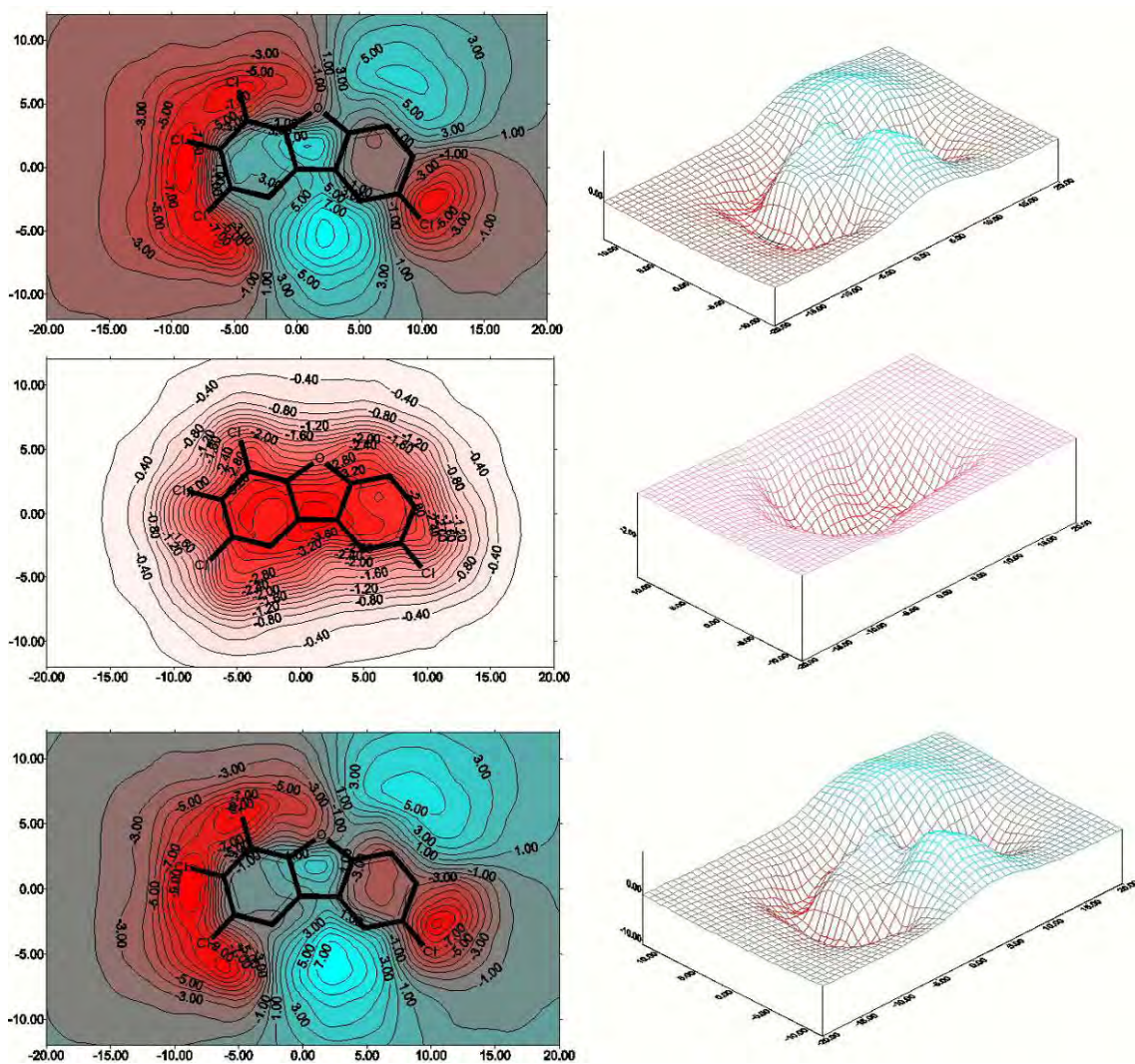


Figura 5.4.23. Mapes de Potencial electrostàtic, de polarització i d'interacció per el furà 2348df

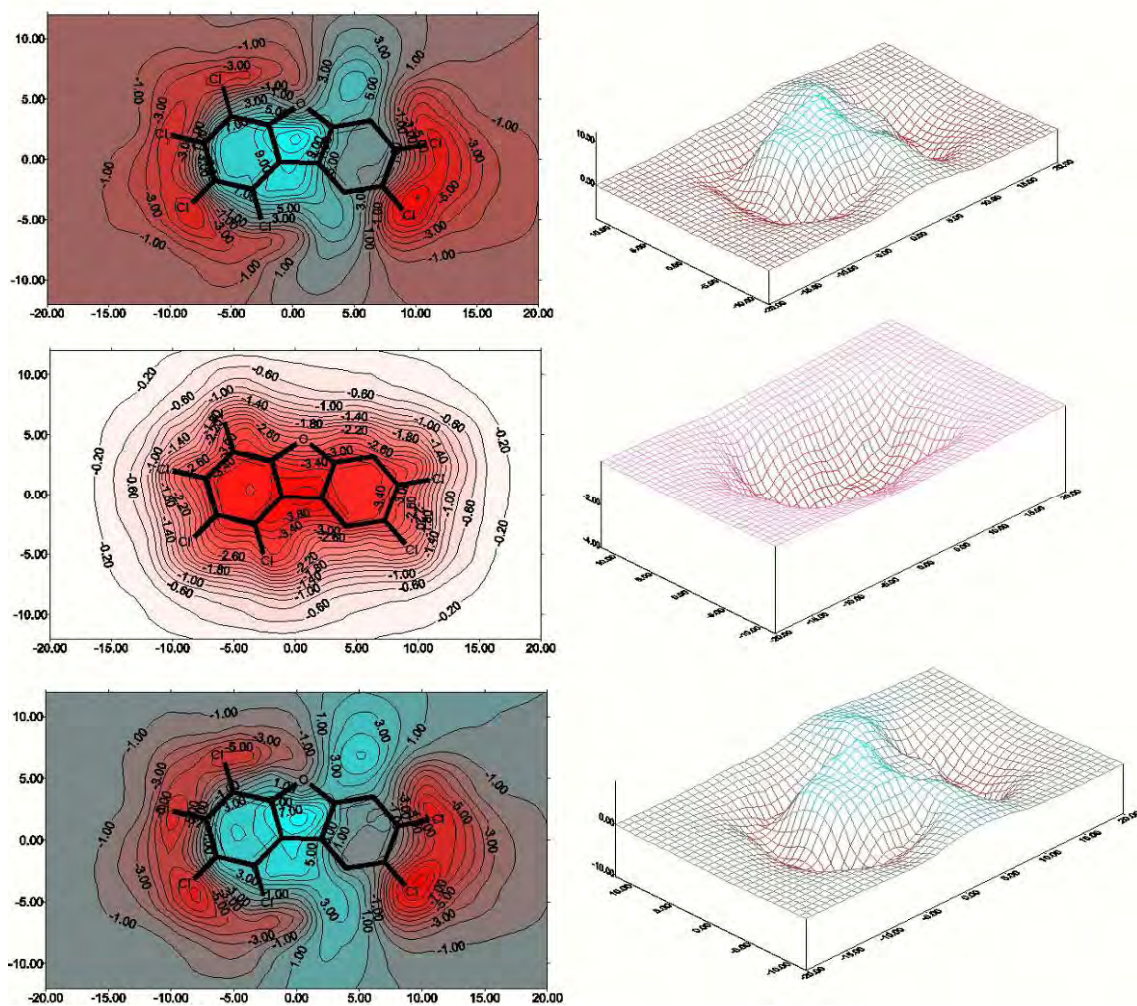


Figura 5.4.24. Mapes de Potencial electrostàtic, de polarització i d'interacció per el furà 123478df

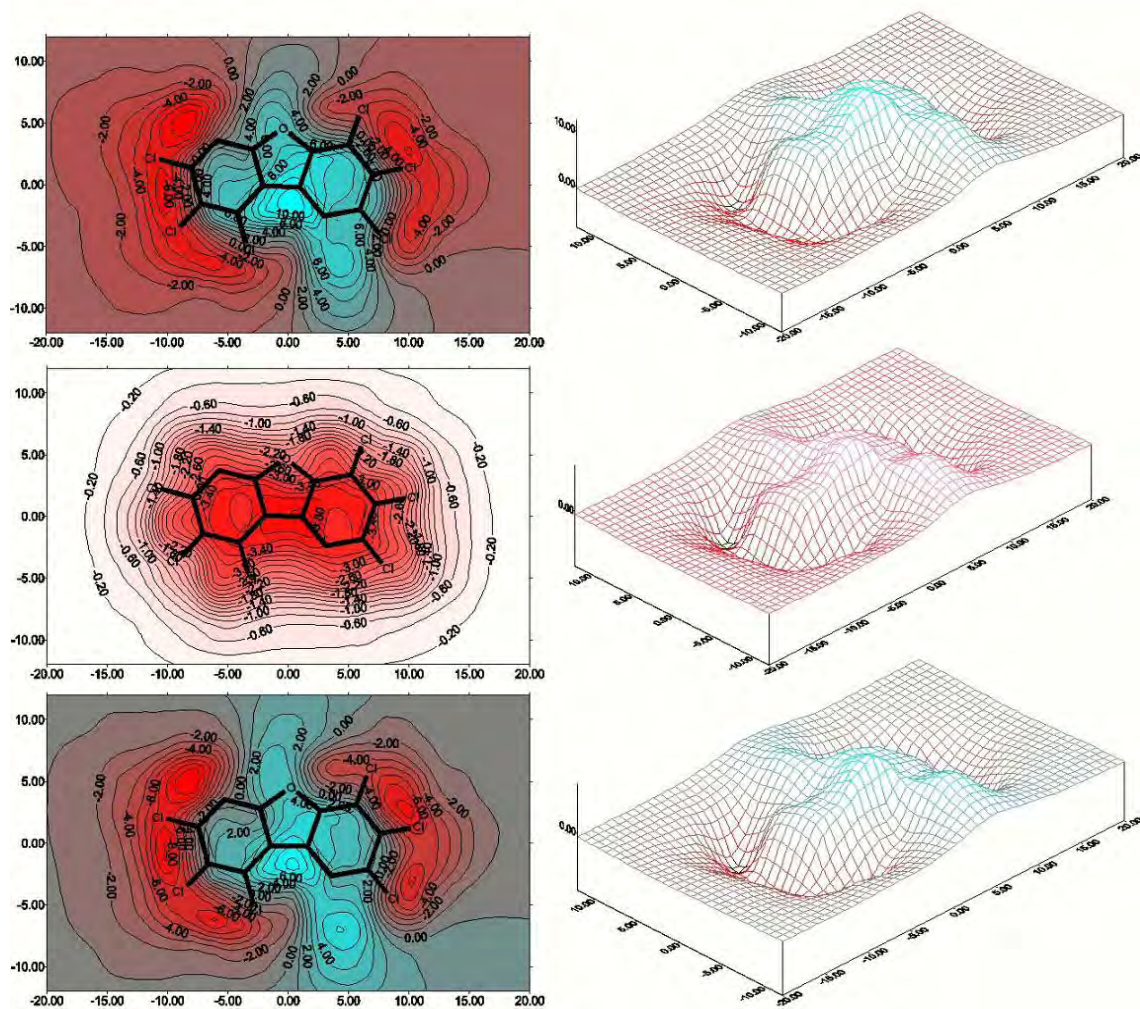


Figura 5.4.25. Mapes de Potencial electrostàtic, de polarització i d'interacció per el furà 123678df

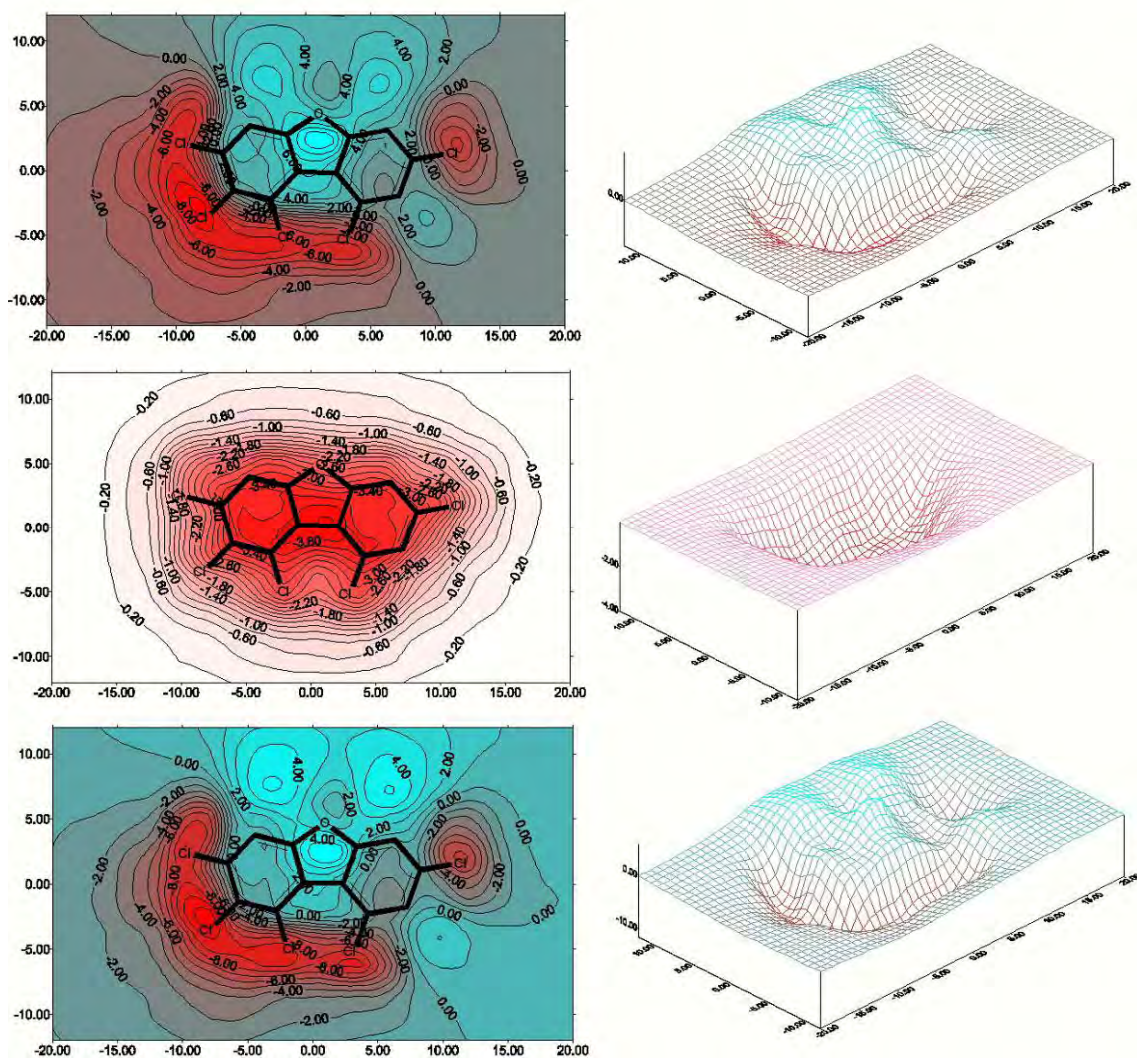


Figura 5.4.26. Mapes de Potencial electrostàtic, de polarització i d'interacció per el furà 12379df

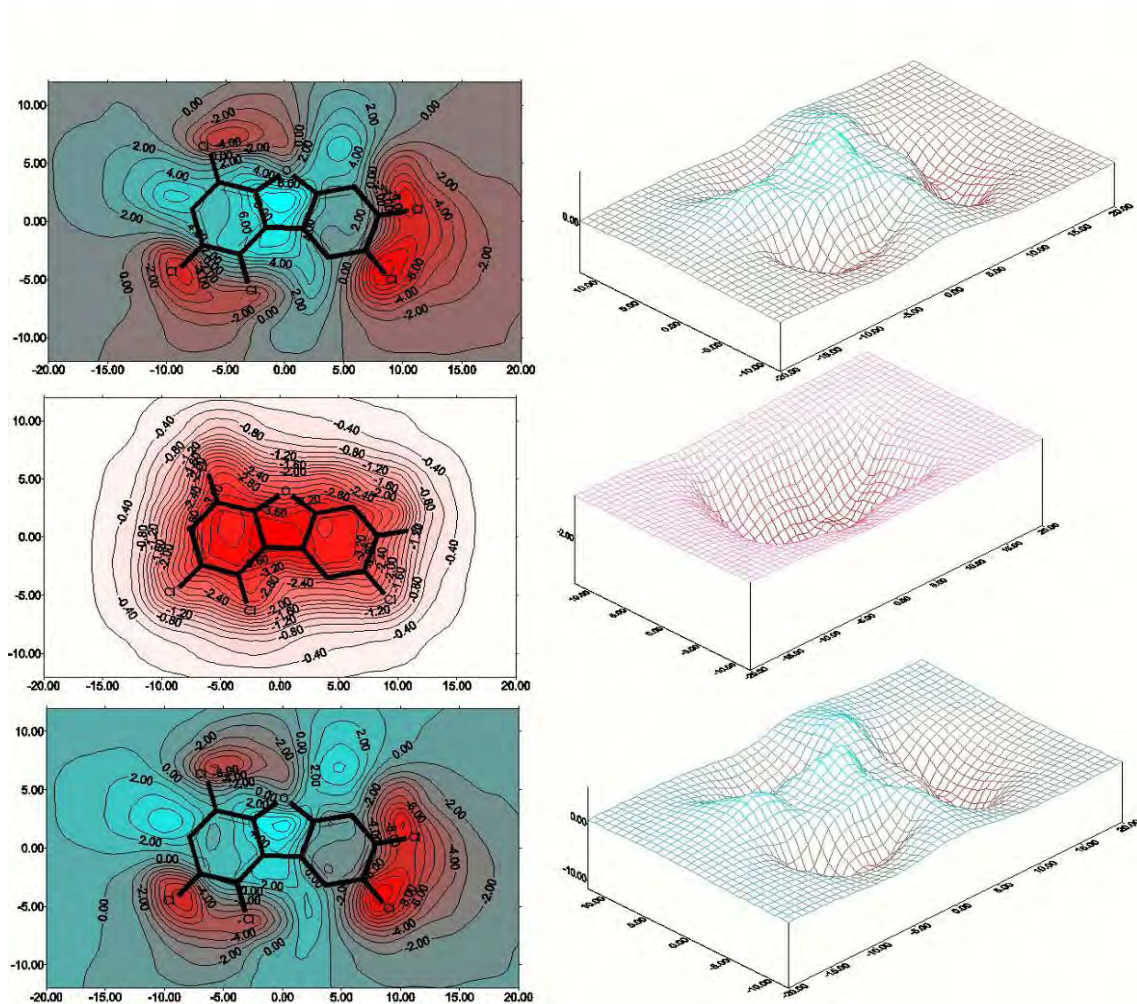


Figura 5.4.27. Mapes de Potencial electrostàtic, de polarització i d'interacció per el furà 12478df

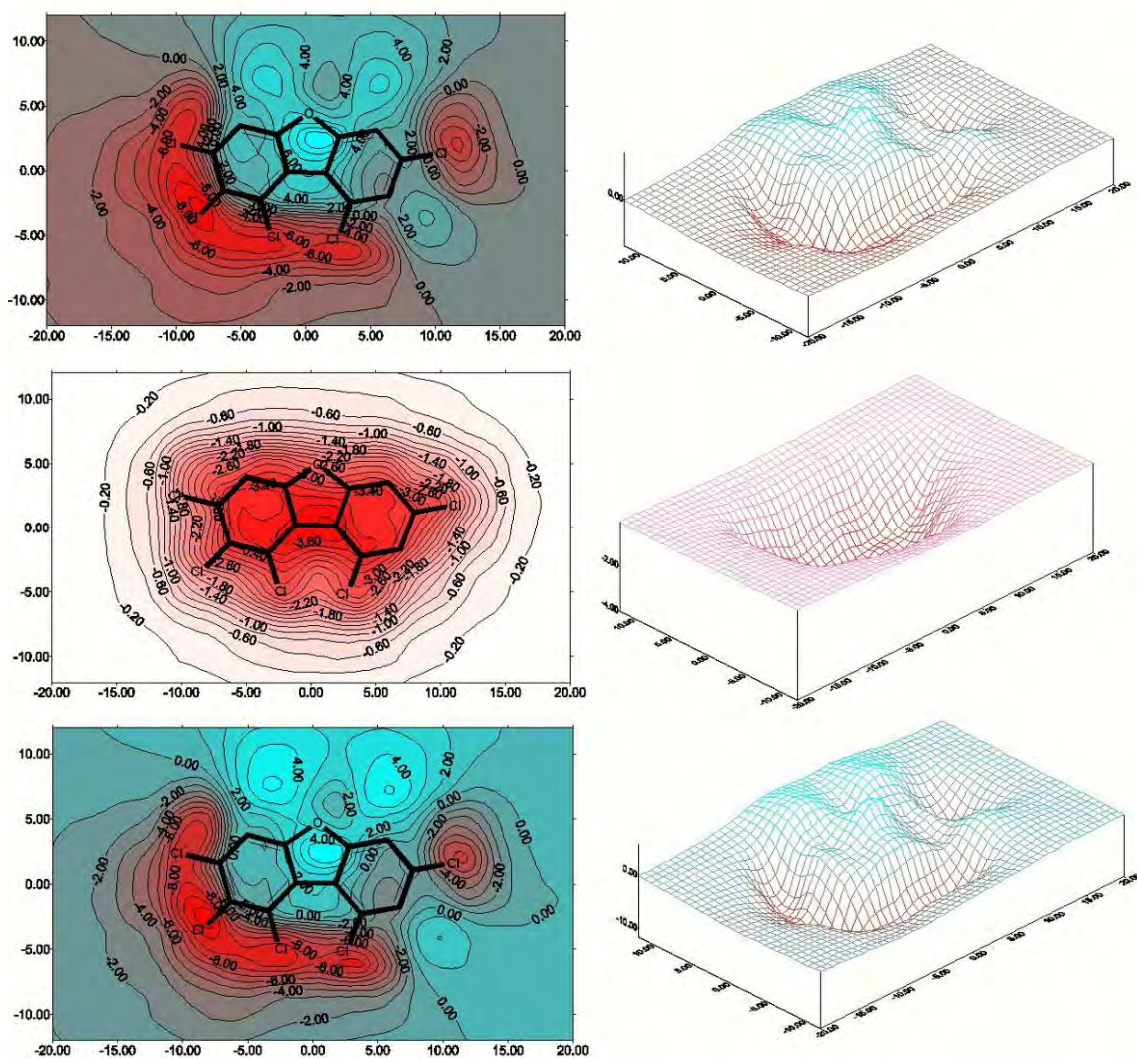


Figura 5.4.28. Mapes de Potencial electrostàtic, de polarització i d'interacció per el furà 12379df

5.4.2 Anàlisi de components principals

Per poder extreure informació rellevant de forma exploratòria en quant a tendències, elements caracteritzadors comuns i fins i tot trobar trets diferencials observables entre els diferents mapes, es va portar a terme una anàlisi de les components principals dels diferents mapes en conjunt. Hem fet aquest estudi independentment per les dioxines, pels furans i per al conjunt complet.

Dioxines

Els valors dels valors propis de les diferents components, així com la variància acumulada es mostren a la Taula 5.4.1. Com es pot veure, amb tres components pels mapes de potencial electrostàtic, un pel potencial de polarització i tres pel potencial d'interacció, queda força ben explicada la variabilitat del conjunt de molècules amb un 95% de variància acumulada.

Taula 5.4.1. Aportació de les diverses components principals a la variància del conjunt per als MEP, MPP i MEI pel cas de les dioxines.

DIOXINES	N	Comp	Variància Acumulada (%)
MEP	6	Comp[1]	86,2
		Comp[2]	91,9
		Comp[3]	95,4
		Comp[4]	95,6
		Comp[5]	95,9
		Comp[6]	96,2
		MPP	2
Comp[2]	98,0		
Comp[1]	87,2		
MEI	6	Comp[2]	91,3
		Comp[3]	95,6
		Comp[4]	96,8
		Comp[5]	97,2
		Comp[6]	97,8

La representació gràfica dels coeficients (*loadings*) del vectors propis o variables latents, de cada una de les components principals corresponents als resultats de la Taula 5.4.1 es mostra a les Figures 5.4.29-5.4.31.

Els mapes de les components principals ens donen informació de trets comuns al grup de molècules estudiades. Com que la estructura química d'aquest grup de molècules és la mateixa, els resultats obtinguts es poden explicar en termes del nombre dels clors que hi ha a cada posició de l'anell fusionat.

En el cas del MEP s'observa que les àrees que donen la variabilitat més grans són les que corresponen a les posicions més freqüents dels clors al conjunt de les molècules situats a l'esquerra de la molècula (posicions 2 i 3), amb igual importància les posicions 7, a la dreta de l'anell fusionat, encara que s'estén amb menys intensitat cap a la posició 8. També s'observa una zona d'alta variabilitat situada a la zona de l'anell central, que explicaria el potencial positiu observat en totes les molècules del grup. Aquesta component representa trets comuns del grup de molècules estudiat: la importància dels clors en posicions 2 (dotze molècules de catorze tenen un clor a aquesta posició), 3 (deu de catorze), 7 (nou de catorze) i 8 (set de catorze). De la mateixa manera, el valor positiu del potencial al centre de l'anell el mostren totes les molècules del grup.

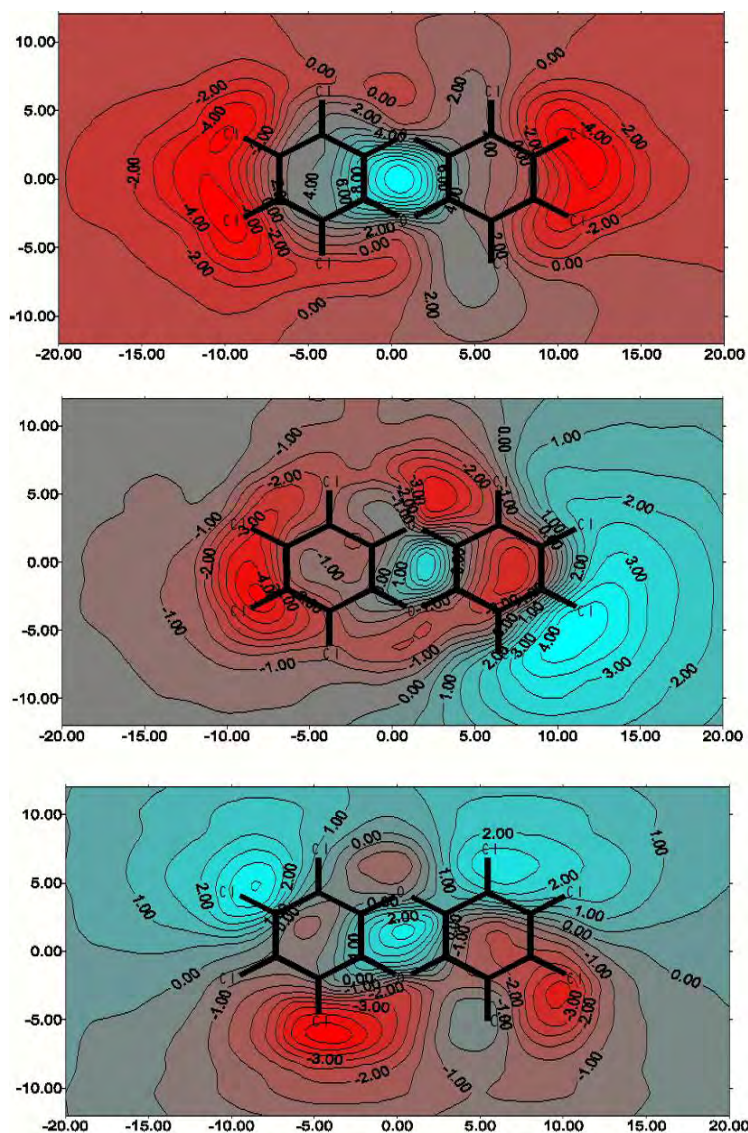


Figura 5.4.29. Representació gràfica dels components principals (loadings) corresponent al MEP de les dioxines.

Si observem la segona component, veiem que les zones d'alta variabilitat són trets comuns diferenciats amb la primera component. Per una banda torna a aparèixer el clor a l'esquerra de l'anell fusionat en posició 2 i també el clor en posició 8, que sent molt comú en aquest grup de molècules (set de catorze) no havia sortit a la primera component. També s'observa alta variabilitat a la densitat electrònica de l'anell aromàtic de la dreta i als oxígens, encara que més important al situat a la part superior de l'anell de dioxina. Finalment, la tercera component encara que és de més difícil interpretació, es podria considerar com una component que posa de relleu els elements diferenciats més importants entre les diferents dioxines d'aquest estudi: la substitució en la posició 1 que apareix en 9 molècules; la substitució en posició 3 que apareix en 10 molècules; la substitució en 6 apareix 3 vegades i finalment, la substitució en 8 apareix 7 vegades. Aquest caràcter diferenciador es posarà de relleu quan fem una projecció dels *scorings* de cada molècula sobre els eixos de les components principals.

Si ara analitzem els mapes de potencial de polarització, de fet veiem que amb una sola component obtenim un percentatge molt alt de la variància del conjunt. Aquest vector propi descriu la tendència comú per a totes les molècules d'un augment de l'energia d'interacció al voltant de la molècula de forma simètrica.

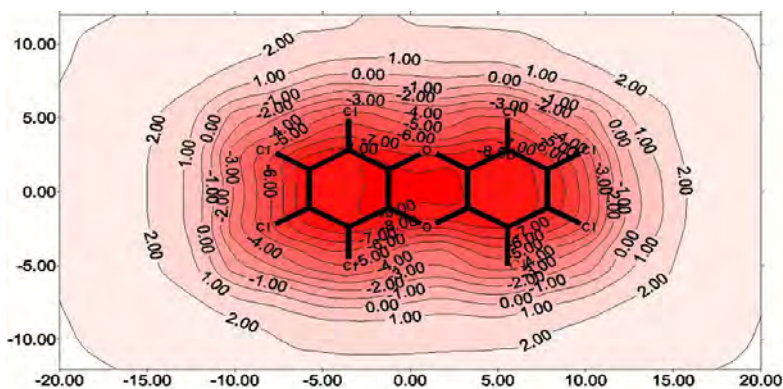


Figura 5.4.30. Representació gràfica de la component principal (*loadings*) corresponent al MPP de les dioxines.

Finalment, en analitzar els mapes de interacció veiem que amb tres components es descriu un 95% de la variabilitat del sistema. Les components se semblen molt a les components obtingudes al MEP però, degut a la distribució uniforme, característica del potencial de polarització, els mapes es veuen com més suaus amb menys zones de variabilitat. La primera component té molt elements comuns amb la component corresponent del MEP, encara que hi ha diferències significatives. De fet, desapareixen les taques de variabilitat sobre els oxígens de l'anell de dioxina i sobre el clor en posició 1 i apareix una àrea de variabilitat a sobre del clor en

posició 9. La segona component reflexa trets comuns entre el grup de molècules, encara que és diferent de la segona component del MEP. Desapareixen l'anell aromàtic de la dreta i per damunt dels clors en posicions 3 i 8. En canvi, apareixen les àrees intermèdies entre les posicions 1 i 2; 4 i 5 o 8 i 9 a sobre de tots els clors i oxígens i sobre la posició 7. Per últim, la tercera component al igual que en el cas del MEP, reflexa trets diferencials entre molècules. Apareixen com importants les posicions 1, 4, 6 i 8. Apareix en aquest cas també l'àrea sobre la densitat electrònica de l'anell aromàtic de la dreta.

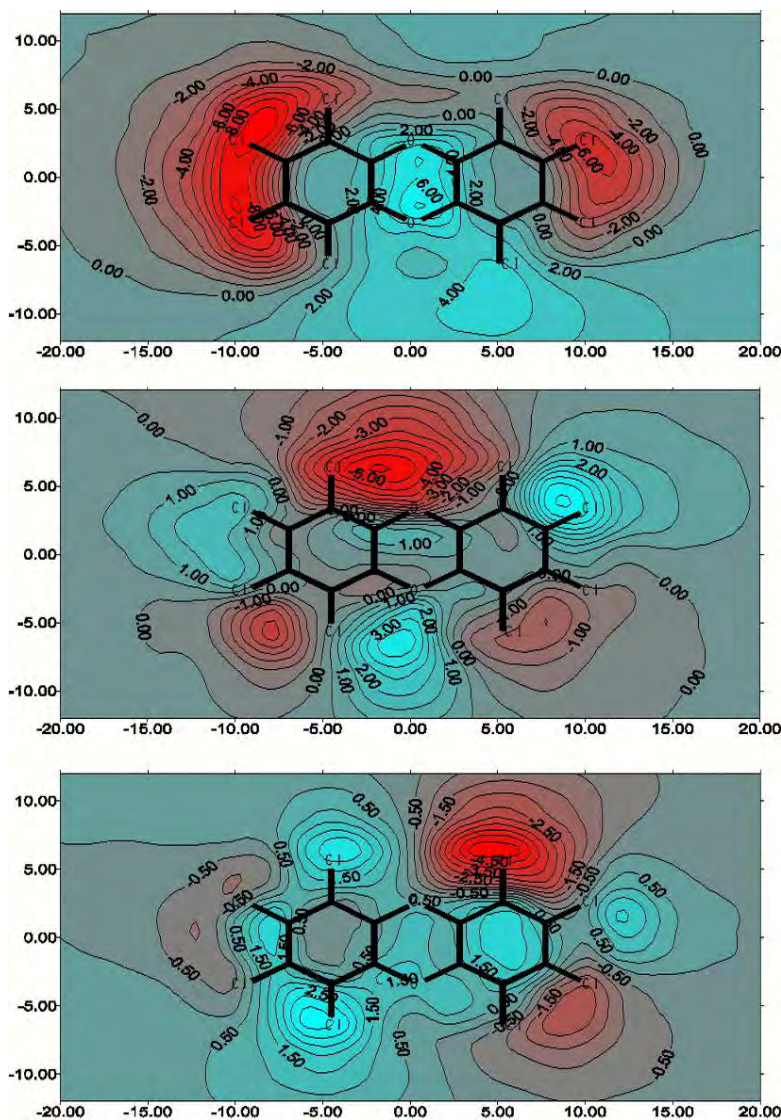


Figura 5.4.31. Representació gràfica dels components principals (loadings) corresponent al MEP de les dioxines

A les Taules 5.4.2 i 5.4.3 es mostren els *scorings* o projeccions de cada molècula sobre les components principals (*scorings*) corresponents al MEP i MEI respectivament, calculats com el producte de la primera fila de la matriu de dades inicial per les components del vector propi corresponent:

$$S_i = \sum A_{ik} c_k$$

Taula 5.4.2. Valors de pes (*scorings*) per el MEP de les dioxines

Anàleg	M1.p[1]	M1.p[2]	M1.p[3]
Tcdd	0,299562	-0,248431	0,181627
12378dd	0,326017	-0,141378	0,042015
2367dd	0,323916	0,067155	0,291580
236dd	0,243725	0,336833	0,129209
123478dd	0,303046	-0,214605	-0,252574
1378dd	0,303988	-0,066407	0,012633
12478dd	0,313629	-0,068780	-0,035667
1234dd	0,169783	0,373727	-0,431935
237dd	0,322875	0,096115	0,241892
28dd	0,203441	0,209592	0,487129
12347dd	0,315777	0,092943	-0,230015
124dd	0,112853	0,465738	-0,328290
ocdd	0,257712	-0,249229	-0,383296
1mcdd	-0,060936	0,514419	0,091299

Aquests valors permeten representar gràficament les posicions de les diferents molècules sobre l'espai definit per dues components principals. A la Figura 5.4.32 es mostren les projeccions dels *scorings* sobre les components p1 vs p2; p1 vs p3 i p2 vs p3 respectivament.

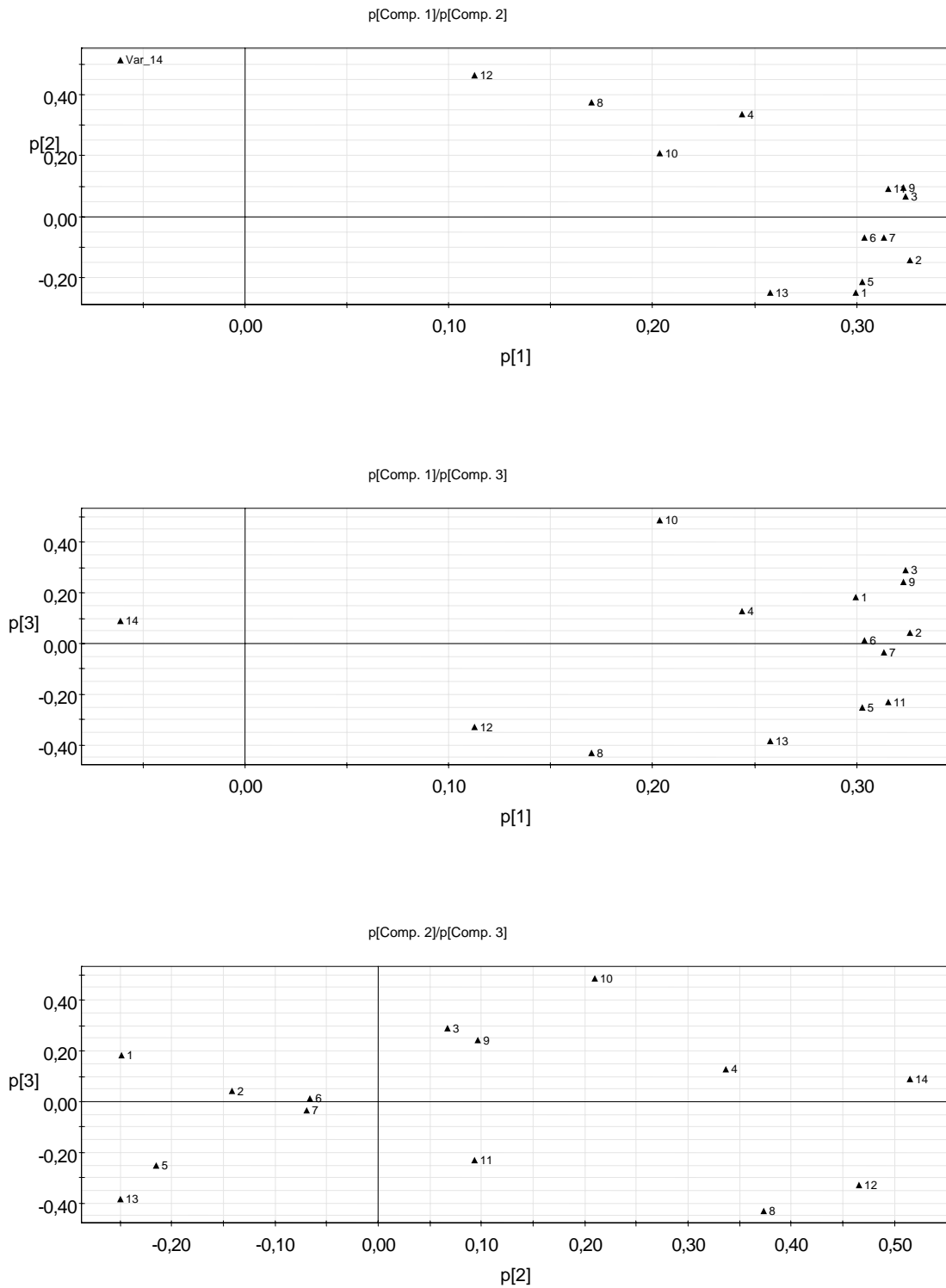


Figura 5.4.32 Projecció dels scorings del MEP de cada molècula sobre les diferents components principals corresponent al conjunt de dioxines: a) projecció sobre les components p1 i p2; b) projecció sobre les components p1 i p3; c) projecció sobre les components p2 i p3.

Una anàlisi de la Figura 5.4.32a mostra que la dioxina 14 surt aïllada. Aquesta estructura és la que té menys substitucions de clor, i també és la que té la toxicitat més petita. Les dioxines 4, 8, 10 i 12, encara que separades entre elles, apareixen agrupades. Aquestes molècules es caracteritzen estructuralment (veure Figura 5.2.1) per tenir un nombre petit de clors (molècula 10) i cap clor a l'anell de la dreta excepte la molècula 10. La resta de molècules apareixen més o menys properes encara que es pot fer una anàlisi més fina. La molècula 13 sembla ser més diferent, això està relacionat amb la seva estructura amb totes les posicions possibles saturades amb àtoms de clor. Les molècules 1, 2, 5, 6 i 7 apareixen properes, totes amb clors a les posicions 7 i 8 i amb un nombre més gran o més petit de clors a l'anell de la esquerra. És un fet interessant que aquestes molècules presenten una toxicitat elevada. Finalment, surten agrupades les molècules 3, 9 i 11 amb un nombre elevat de clors i amb un d'ells en posició 7. Aquestes molècules també presenten toxicitats elevades. Les altres projeccions donen resultats semblants: la molècula 14 apareix també aïllada; les molècules 4, 8, 10 i 12, encara que separades de la resta apareixen en dos grups: 8 i 12 (amb cap clor a l'anell de la dreta) per una banda i 4 i 10 per un altre (amb una substitució de clor a l'anell de la dreta). És interessant veure com a la Figura 5.4.32b la molècula 13 es troba a prop de la resta de molècules, mentre que a la Figura 5.4.32c es troba propera a la molècula 5. Les molècules 1, 2, 5, 6 i 7 apareixen agrupades a la Figura 5.4.32c i en canvi a la Figura 5.4.32b apareixen agrupades amb la resta: 3, 9 i 11, que a la Figura 5.4.32c apareixen formant un grup. Aquesta anàlisi ens permet concloure que la projecció p2/p3 és la que ens dóna més informació sobre aquest grup de molècules i que per tant, les podríem classificar en els grups: i) 5 i 13; ii) 1, 2, 6 i 7; iii) 3, 9, 11; iv) 4 i 10; v) 8 i 12; vi) 14. La mateixa anàlisi que hem fet pel MEP el farem a continuació pel MEI. A la Taula 5.4.3 es troben els valors dels *scorings* pels mapes d'interacció (MEI).

Taula 5.4.3. Valors de pes (*scorings*) dels MEIs per les dioxines

Anàleg	MEI.p[1]	MEI.p[2]	MEI.p[3]
Tcdd	0,285732	-0,262575	-0,246101
12378dd	0,313281	-0,211944	-0,0372082
2367dd	0,342217	0,0194802	-0,257665
236dd	0,285545	0,279223	-0,102373
123478dd	0,265643	-0,289852	0,293251
1378dd	0,289586	-0,0894444	-0,0924753
12478dd	0,290718	-0,0997317	0,00116769
1234dd	0,202997	0,321687	0,435111
237dd	0,343048	0,0582054	-0,217974
28dd	0,249903	0,221563	-0,375971
12347dd	0,310927	0,0299846	0,310715
124dd	0,161104	0,423871	0,326395
ocdd	0,187685	-0,326666	0,429753
1mcdd	0,0585072	0,512437	-0,0259915

Seguint els mateix procediment que hem realitzat anteriorment, les projeccions dels valors del *scorings* sobre les components principals es troben a la Figura 5.4.33.

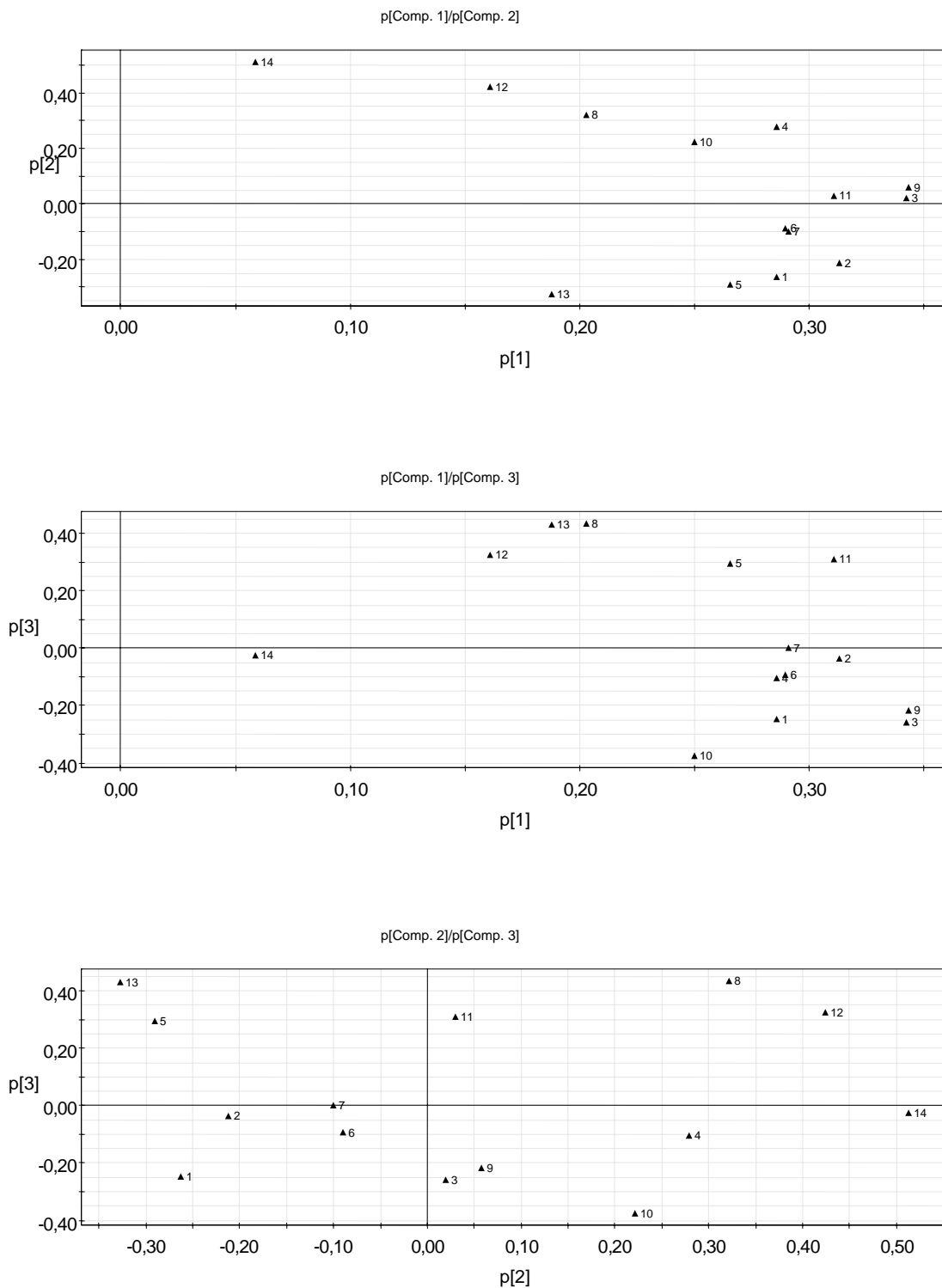


Figura 5.4.33 Projecció dels scorings dels MEI de cada molècula sobre les diferents components principals corresponent al conjunt de dioxines: a) projecció sobre les components p1 i p2; b) projecció sobre les components p1 i p3; c) projecció sobre les components p2 i p3.

En aquest cas també sembla que és la projecció p2/p3 la que més informació dona. Per una banda, ara apareixen clarament agrupades les molècules 5 i 13. Novament les molècules 1, 2, 6 i 7 apareixen agrupades, així com les 3, 9 i 11 on es fa més evident la diferència entre les dos primeres i la darrera. Les molècules 8 i 12 apareixen agrupades, així com la 4 i 10, i finalment la 14 apareix aïllada.

Si comparem els resultats per al potencial electrostàtic i d'interacció, observem que la inclusió del potencial de polarització ens permet agrupar de forma més nítida les diferents molècules, depenent de la seva distribució de clors. A grans trets sembla com si la classificació a través del MEP estigués més relacionada amb la distribució dels clors substituents i en canvi en el MEI les molècules es classifiquen de forma semblant, però la distribució queda modulada pel nombre total de clors.

Benzofurans

Com al cas de las dioxines, l'anàlisi de les components principals dels furans ens permet explicar un 95% de la variabilitat amb tres components pels mapes de potencial electrostàtic, una pel potencial de polarització i tres pel potencial d'interacció. A la Taula 5.4.4 es poden veure els valors propis i *scorings*, així com el percentatge de variància que representen.

Taula 5.4.4. Aportació de les diverses components principals a la variància del conjunt per als MEP, MPP i MEI pel cas de les dioxines.

BENZOFURANS	N	Comp	Variància Acumulada (%)
MEP	6		
		Comp[1]	92,6
		Comp[2]	95,2
		Comp[3]	95,9
		Comp[4]	96,1
		Comp[5]	97,3
		Comp[6]	97,6
MPP	2		
		Comp[1]	97,7
		Comp[2]	98,5
MEI	6		
		Comp[1]	85,2
		Comp[2]	90,1
		Comp[3]	96,3
		Comp[4]	96,9
		Comp[5]	97,1
		Comp[6]	97,7

La representació gràfica dels coeficients del vectors propis *loadings* de cada una de les components principals es mostra a les Figures 5.4.34-5.4.36. A la Figura 5.4.34 es mostren els resultats per al cas del MEP, i com es podrà observar, els resultats són molt semblants als obtinguts amb les dioxines. A la primera component veiem que les àrees de variabilitat més gran

són aquelles que corresponen a les posicions 2, 3 i 4, i que corresponen a les substitucions dels clors comuns situats a l'esquerra de la molècula, i amb menor importància la posició 7. També s'observa una zona d'alta variabilitat situada a la zona de l'anell central que explicaria el potencial positiu observat en totes les molècules del grup. Aquests resultats expliquen la importància dins d'aquest grup de molècules dels clors en posicions 2, 3 i 4 (tretze, dotze i deu molècules d'un total de catorze, respectivament) i dels més comuns a la dreta en posicions 7 i 8 (dotze i onze molècules de catorze). De la mateixa manera s'observa la zona d'alta variabilitat que es localitza al centre de l'anell (és mostrat per totes les molècules del grup). Aquesta component ens aporta trets comuns entre el grup de molècules objecte d'aquest estudi. La component segona es semblant a la que havíem obtingut a l'estudi de les dioxines tret de les diferències estructurals corresponents. Els valors de variabilitat més gran apareixen a les posicions corresponents a les substitucions 2 i 3 a la esquerra dels anells fusionats i a la posició 8 a la dreta. També s'observa una àrea d'alta variabilitat al voltant de l'oxigen furànic i a la zona de densitat electrònica corresponent a l'anell aromàtic de la dreta. Com en el cas de les dioxines, aquesta segona component descriu elements comuns del conjunt que no necessàriament surten a la primera component. Finalment, la tercera component és de més difícil interpretació, però està relacionada amb clors que apareixen un nombre mitjà de vegades. També apareixen punts estacionaris dins dels anells aromàtics que no s'observaven a la tercera component de les dioxines.

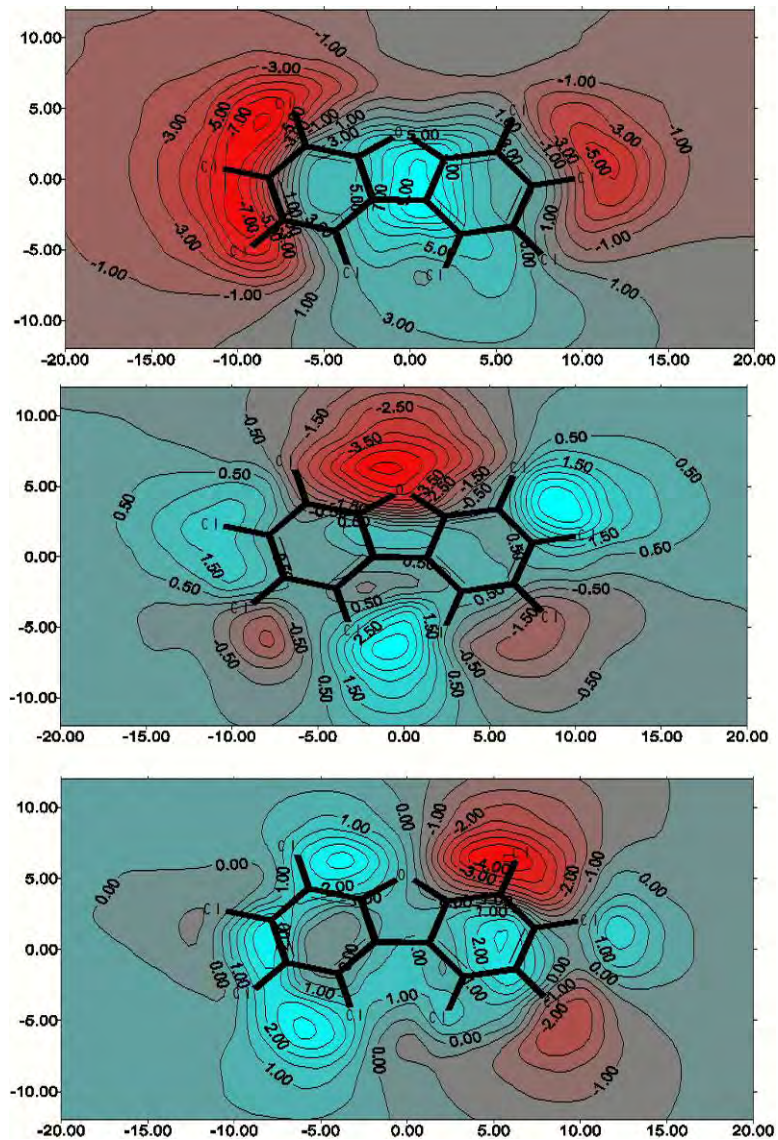


Figura 5.4.34. Representació gràfica dels components principals (loadings) corresponent al MEP dels furans

Els mapes de potencial de polarització són semblants als descrits de les dioxines, de fet veiem que amb una sola component obtenim un percentatge molt alt de la variància del conjunt. Aquest vector propi és una tendència comuna: l'augment de l'energia d'interacció (negativa) al voltant de la molècula de forma simètrica.

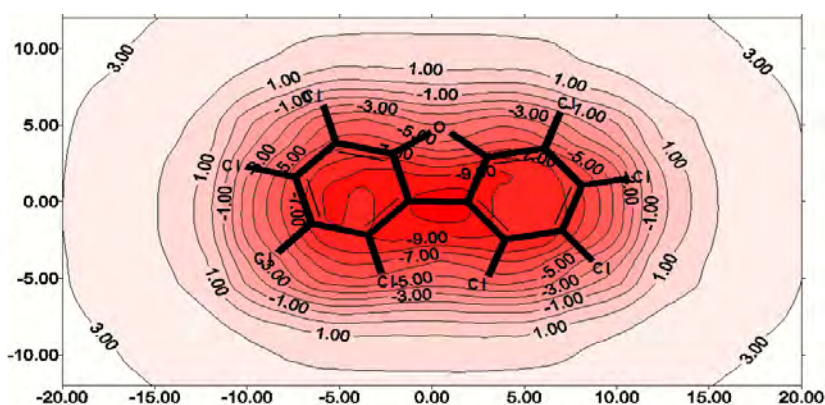


Figura 5.4.35. Representació gràfica dels components principals (loadings) corresponent al MPP dels derivat dels furans.

Finalment, si passem als mapes d'interacció veiem que amb tres components es pot explicar un 96% de la variabilitat del sistema. La primera component torna a ser la primera component del MEP però amb més èmfasi respecte a les zones comunes negatives i positives. La segona component reflexa trets comuns entre el grup de molècules encara que és diferent de la segona component del MEP. Desapareixen l'anell aromàtic de la dreta i sobre els clors en posicions 3 i 8. En canvi, apareixen les àrees intermèdies entre les posicions 1 i 2; 4 i 5 o 8 i 9 per sobre de tots els clors i oxígens i sobre la posició 7. Per últim, la tercera component al igual que en la cas del MEP reflecteix trets diferencials entre molècules. Aparèixen com importants les posicions 1, 4, 6 i 8. Apareix en aquest cas també l'àrea sobre la densitat electrònica de l'anell aromàtic de la dreta.

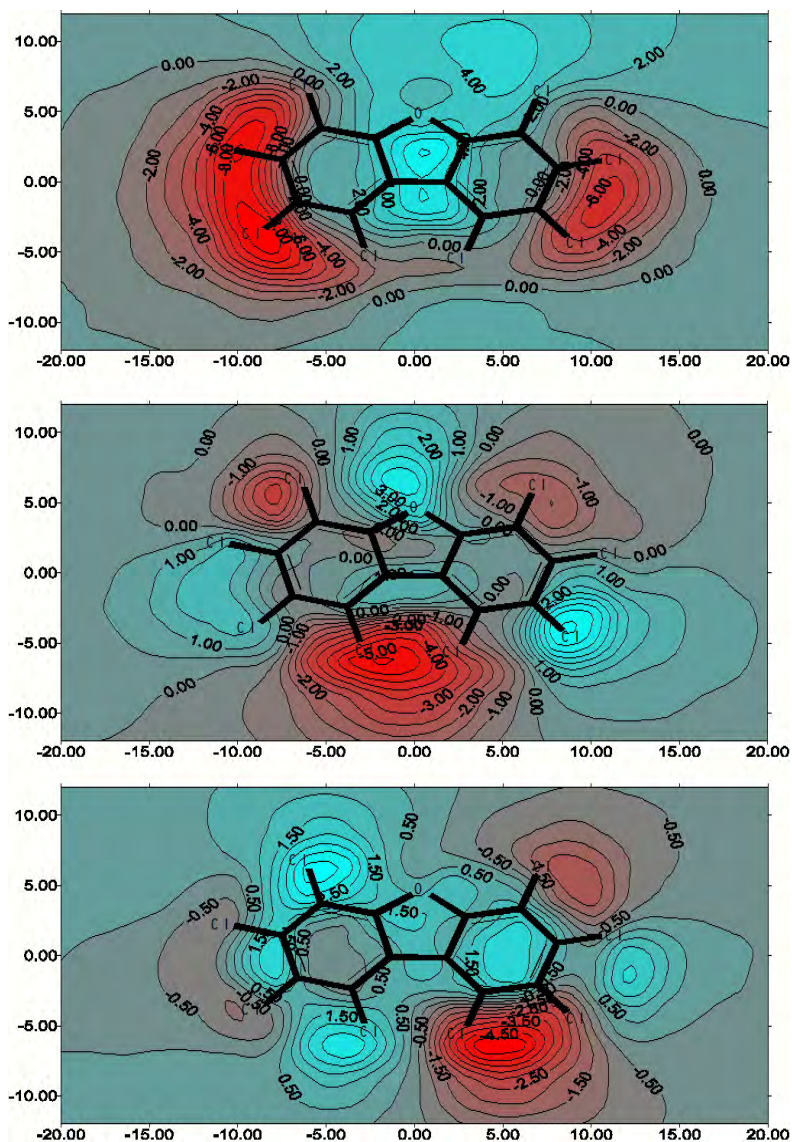


Figura 5.4.36. Representació gràfica dels components principals (loadings) corresponent al MEI dels derivat dels furans.

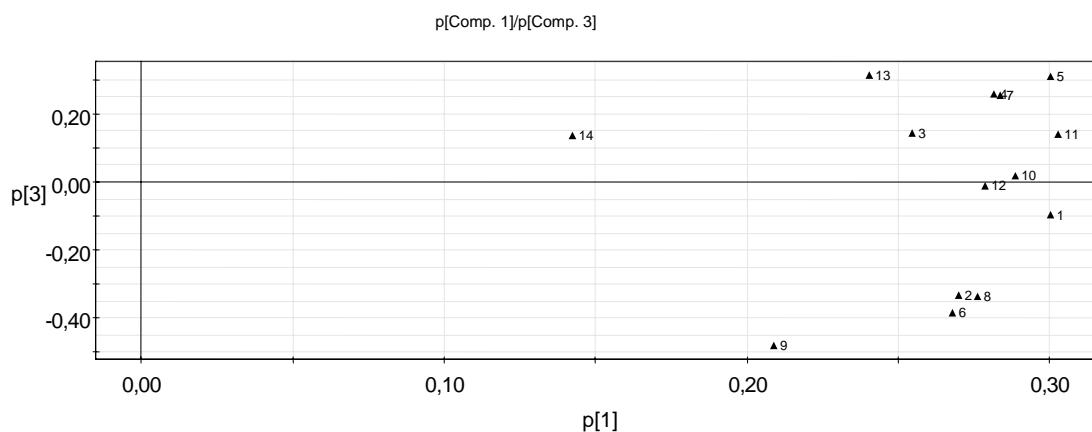
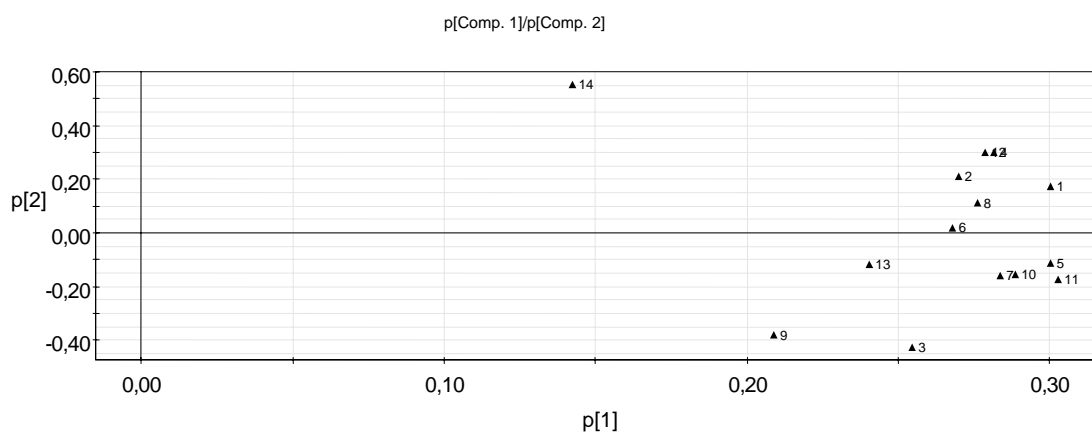
A la Taula 5.4.5 es mostren els valors de pes *scorings* pel cas del MEP de cada molècula del conjunt de furans per a cadascun dels components identificats, calculats com el producte de la primera fila de la matriu de dades inicial per les components del vector propi corresponent:

$$S_i = \sum A_{ik} c_k$$

Taula 5.4.5. Valors de pes (scorings) per el conjunt dels mapes MEP dels furans

Var ID (Primary)	M1.p[1]	M1.p[2]	M1.p[3]
1	0,300431	0,17169	-0,095795
2	0,269944	0,211845	-0,333919
3	0,254407	-0,424666	0,142696
4	0,281721	0,29901	0,25914
5	0,300206	-0,111434	0,311751
6	0,267895	0,0196518	-0,382775
7	0,283781	-0,159769	0,254215
8	0,276179	0,112995	-0,336767
9	0,208721	-0,37961	-0,481136
10	0,288597	-0,155584	0,0188203
11	0,302582	-0,173028	0,141005
12	0,278657	0,298323	-0,0127675
13	0,240323	-0,115709	0,314816
14	0,142374	0,551511	0,138069

Aquests valors permeten representar gràficament les projeccions de les diferents molècules sobre l'espai definit per dues components principals. A la Figura 5.4.37 es mostren les projeccions dels scorings sobre les components p1 vs p2; p1 vs p3 i p2 vs p3 respectivament.



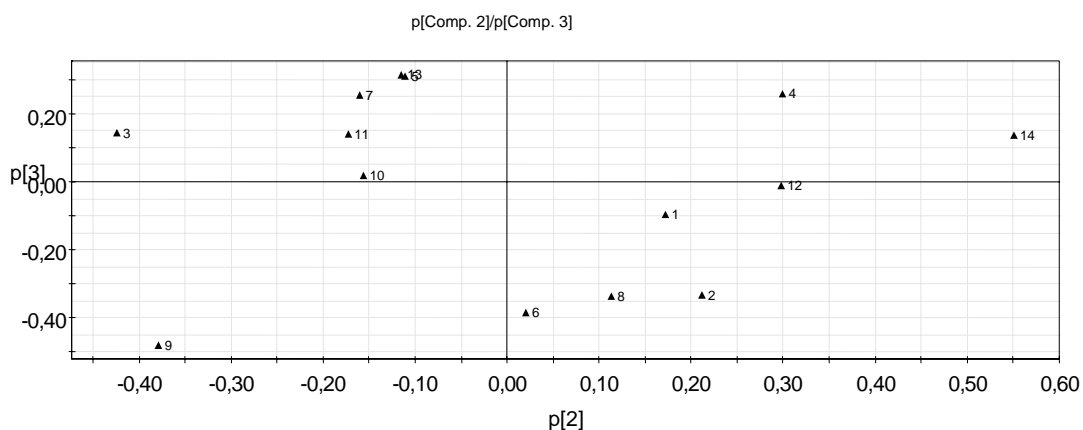


Figura 5.4.37 *Projecció dels scorings dels MEP de cada molècula sobre les diferents components principals corresponent al conjunt de furans: a) projecció sobre les components p1 i p2; b) projecció sobre les components p1 i p3; c) projecció sobre les components p2 i p3.*

Com en el cas de les dioxines, en aquest cas la projecció p2/p3 és també la que dona més informació. Per una banda veiem les molècules 3, 9 i 14, es pot considerar que no tenen cap altra molècula amb trets comuns. Les molècules 5, 7, 10, 11 i 13 comparteixen una estructura molt semblant amb clors a les posicions 7 i 8, i surten agrupades al mapa de projecció. Les molècules 1, 2, 6 i 8, encara que més disperses, poden també considerarse dins del mateix grup, totes amb clors en les posició 2,3 i 7. Finalment, les molècules 4 i 12 que se les pot considerar dins del mateix grup.

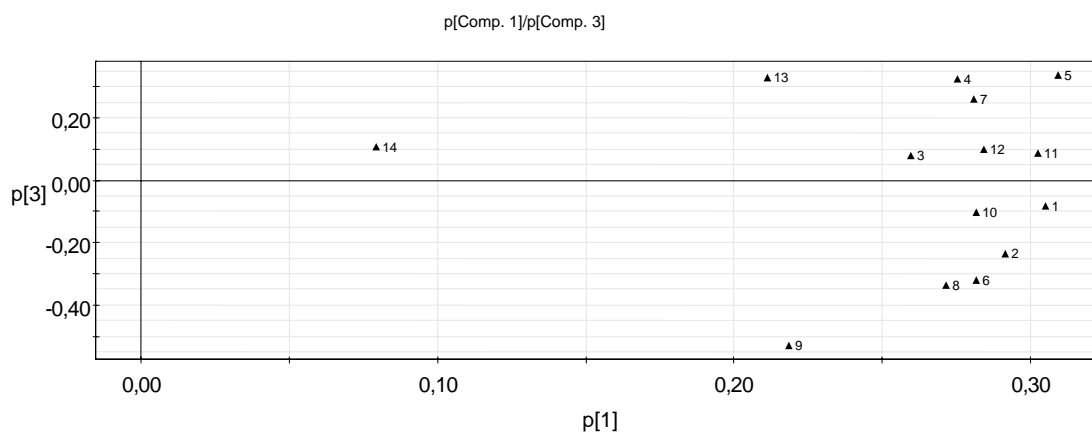
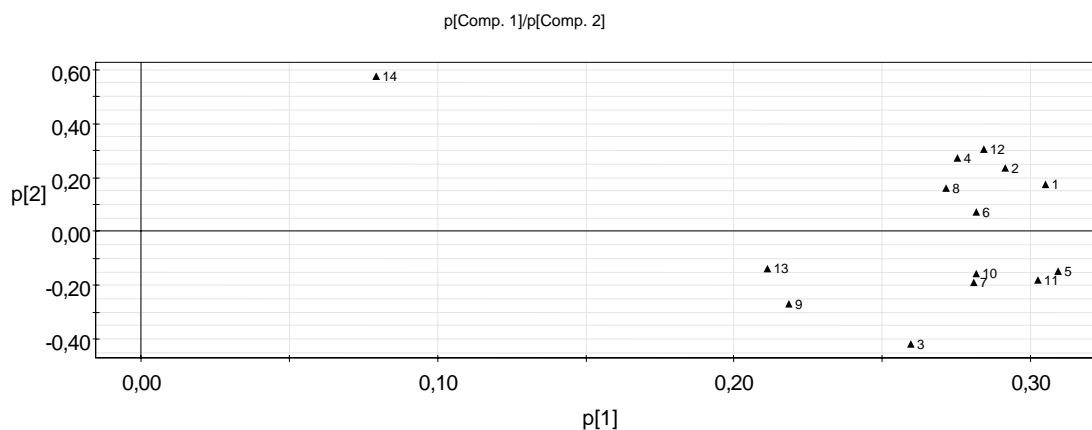
A la Taula 5.4.6 es mostren els valors de pes *scorings* per a cadascun dels components identificats pel cas del MEI de cada molècula del conjunt de furans, calculats com el producte de la primera fila de la matriu de dades inicial per les components del vector propi corresponent:

$$S_i = \sum A_{ik} c_k$$

Taula 5.4.6. Valors de pes (scorings) per el conjunt dels mapes MEI dels furans

Anàleg	M1.p[1]	M1.p[2]	M1.p[3]
23478df	0,305082	0,174601	-0,0825741
2347df	0,291519	0,23702	-0,235076
2378df	0,259901	-0,419178	0,0785395
234678df	0,275566	0,270263	0,324244
12378df	0,309253	-0,145418	0,338493
1237df	0,281879	0,0733891	-0,319018
13478df	0,281159	-0,189264	0,259754
23479df	0,271616	0,161231	-0,336353
2348df	0,218362	-0,269651	-0,530898
123478df	0,281783	-0,156619	-0,102284
123678df	0,302824	-0,181681	0,0863835
12379df	0,284246	0,306809	0,0981409
12478df	0,211397	-0,137418	0,329093
12468df	0,0791675	0,577037	0,109325

Aquest valors permeten representar gràficament les projeccions de les diferents molècules sobre l'espai definit per dues components principals. A la Figura 5.4.38 es mostren les projeccions dels scorings sobre les components p1 vs p2; p1 vs p3 i p2 vs p3 respectivament.



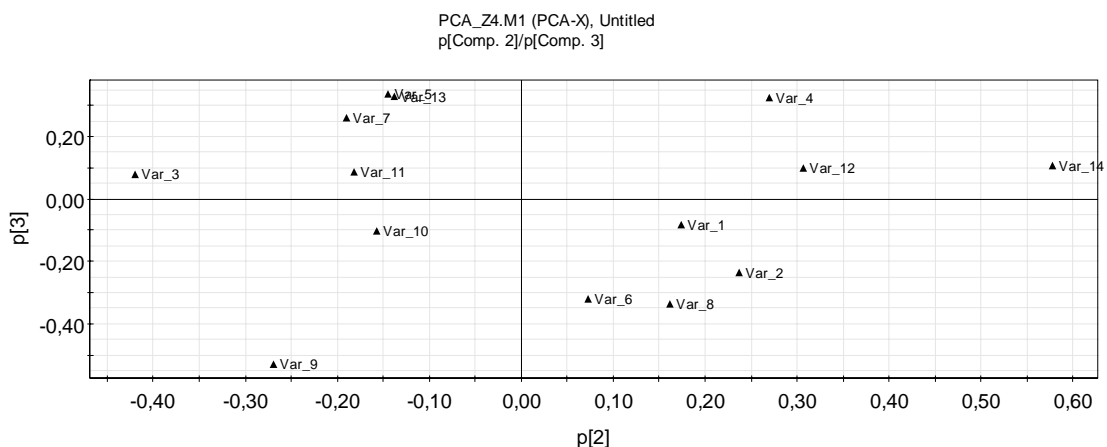


Figura 5.4.38 *Projecció dels scorings dels MEI de cada molècula sobre les diferents components principals corresponent al conjunt de furans: a) projecció sobre les components p1 i p2; b) projecció sobre les components p1 i p3; c) projecció sobre les components p2 i p3.*

En aquest cas també sembla que és la projecció p2/p3 la que dóna més informació. Com es pot observar la Figura 5.4.38c dóna una distribució de punts molt semblant a la que havíem obtingut de l'anàlisi dels mapes de potencial electrostàtic (Figura 5.4.37c) però molt millor definides. Per una banda observant les molècules 3, 9 i 14 es pot considerar que no tenen cap altra molècula amb trets comuns. Les molècules 5, 7, 10, 11 i 13 comparteixen una estructura molt semblant amb clors a les posicions 7 i 8, i surten agrupades al mapa de projecció. Les molècules 1, 2, 6 i 8, encara que més disperses, poden també considerarse dins del mateix grup, totes amb clors en les posicions 2,3 i 7. Finalment, les molècules 4 i 12 que es poden considerar dins d'un mateix grup.

La comparació de les projeccions dels valors de *scoring* calculats segons el MEP o el MEI mostren, com en el cas de les dioxines, que el MEI permet una definició més clara, i de la mateixa manera que el primer classifica al voltant de la distribució dels substituents clors, mentre que el segon modula d'acord al nombre total de clors.

Conjunt de dioxines i furans

A continuació s'analitza el resultat de l'anàlisi de components principals dels mapes MEP, MPP i MEI del conjunt format per dioxines i furans. Fem aquest estudi amb l'objectiu de millorar el nombre de mostres i millorar les descripcions de regions comunes entre aquest conjunt de molècules. Es pot veure a la Taula 5.4.7 l'abast de variabilitat explicada per les diferents components. Com en els casos anteriors es poden agafar tres components per al MEP, una component per al MPP i tres components per al MEI amb una variabilitat del 95%.

Taula 5.4.7. Aportació de les diverses components principals a la variància del conjunt per als MEP, MPP i MEI pel cas del conjunt de dioxines i furans

DIOX_FUR	N	Comp	Variància Acumulada (%)
MEP	6	Comp[1]	91,1
		Comp[2]	93,3
		Comp[3]	94,9
		Comp[4]	95,2
		Comp[5]	95,8
		Comp[6]	96,0
		MPP	2
Comp[2]	98,4		
Comp[1]	89,6		
MEI	6	Comp[2]	91,8
		Comp[3]	95,1
		Comp[4]	96,7
		Comp[5]	96,9
		Comp[6]	97,2

La representació gràfica dels coeficients, *loadings*, del vectors propis o variables latents, de cada una de les components principals corresponents als resultats de la Taula 5.4.7 es mostra a les Figures 5.4.39-5.4.41. Com es veu a continuació, els resultats són idèntics als obtinguts estudiant els dos conjunts de molècules per separat.

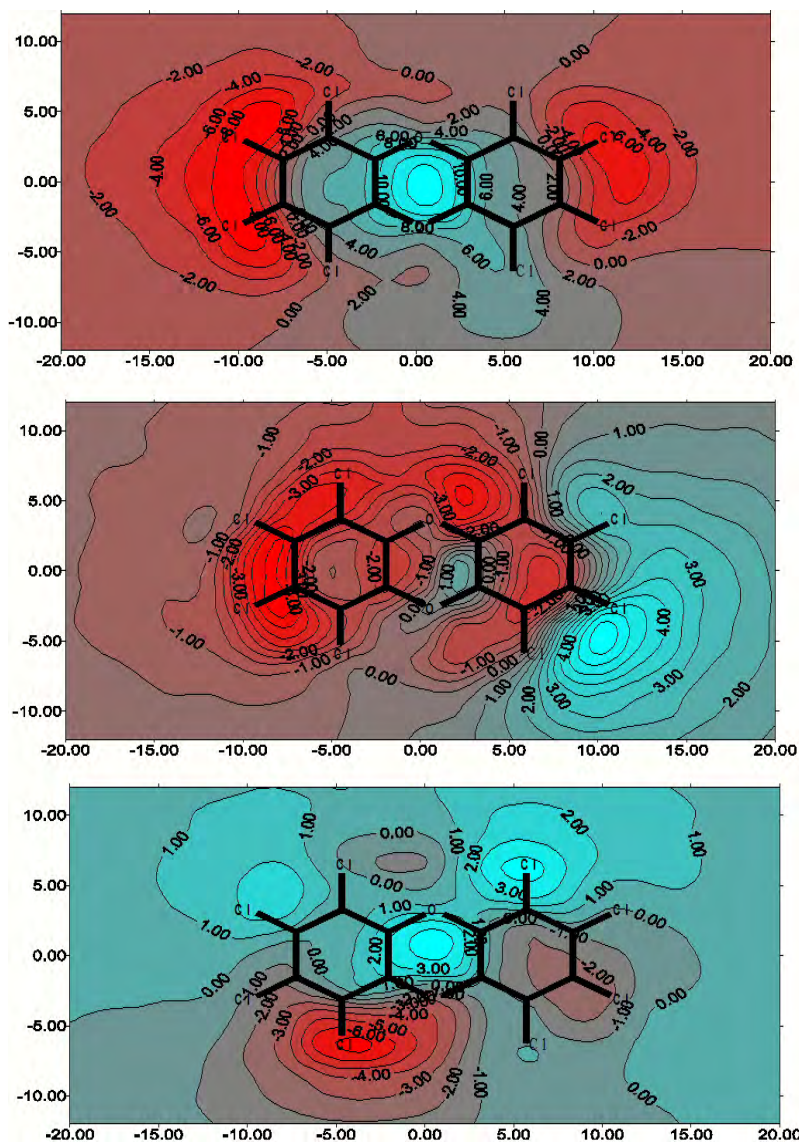


Figura 5.4.39 . Representació gràfica dels components principals (loadings) corresponent al MEP del conjunt total barrejat de dioxines i derivats dels furans.

En el cas del MEP s'observa que les àrees que donen la variabilitat més gran són les que corresponen a les posicions dels clors més freqüents al conjunt de les molècules situats a l'esquerra de la molècula (posicions 2 i 3), amb igual importància les posicions 7, a la dreta de l'anell fusionat, encara que s'estén amb menys intensitat cap a la posició 8. També s'observa una zona d'alta variabilitat situada a la zona de l'anell central, que explicaria el potencial positiu observat en totes les molècules del grup. Aquesta component dóna informació de trets comuns del grup de molècules estudiat: la importància dels clors en posicions 2 (dotze molècules de catorze tenen un clor a aquesta posició), 3 (deu de catorze), 7 (nou de catorze) i 8 (set de catorze). Així mateix, el potencial positiu al centre de l'anell s'observa per a totes les molècules del grup. Si ara mirem la segona component, veiem que les zones d'alta variabilitat són també trets comuns diferenciats amb la primera component. Per una banda torna a aparèixer el clor a

l'esquerra de l'anell fusionat en posicions 2, i també el clor en posició 8 que encara que molt comú en aquest grup de molècules (set de catorze) no s'observa a la primera component. També s'observa una alta variabilitat a la densitat electrònica de l'anell aromàtic de la dreta i als oxigens, encara que més important al situat a la part superior de l'anell de dioxina. Finalment, la tercera component tot i que és de més difícil interpretació, es podria considerar com una component que posa de relleu els elements diferencials més importants entre les diferents dioxines d'aquest estudi: la substitució en la posició 1 que apareix en 9 molècules; la substitució en posició 3 que apareix en 10 molècules; la substitució en 6 apareix 3 vegades i finalment, la substitució en 8 apareix 7 vegades. Aquest caràcter diferenciador es posarà de relleu quan fem una projecció dels *scorings* de cada molècula sobre els eixos de les components principals.

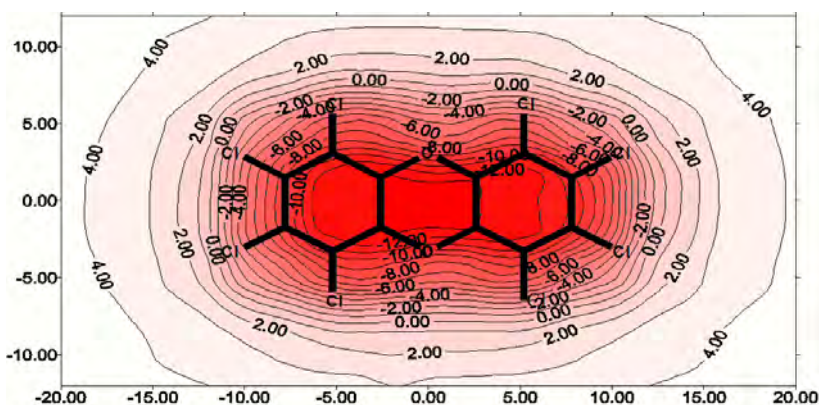


Figura 5.4.40. Representació gràfica dels components principals (loadings) corresponent al MPP del conjunt total barrejat de dioxines i derivats dels furans.

Si ara analitzem els mapes de potencial de polarització, de fet veiem que amb una sola component obtenim un percentatge molt alt de la variància del conjunt. Aquest vector propi descriu la tendència comuna a totes les molècules d'augment de l'energia d'interacció al voltant de la molècula de forma simètrica.

Finalment, en analitzar els mapes de interacció veiem que amb tres components es descriu un 95% de la variabilitat del sistema. Les components se semblen molt a les components obtingudes al MEP però, degut a la distribució uniforme característica del potencial de polarització, els mapes es veuen com més suaus amb menys zones de variabilitat. La primera component té molts elements comuns amb la component corresponent del MEP, encara que hi ha diferències significatives. De fet, desapareixen les taques de variabilitat sobre els oxígens de l'anell de dioxina i sobre el clor en posició 1 i apareix una àrea de variabilitat a sobre del clor en posició 9. La segona component reflexa trets comuns entre el grup de molècules encara que és diferent de la segona component del MEP. Desapareixen l'anell aromàtic de la dreta i sobre els

clors en posicions 3 i 8. En canvi, apareixen les àrees intermèdies entre les posicions 1 i 2; 4 i 5 o 8 i 9 a sobre de tots els clors i oxígens i sobre la posició 7. Per últim, la tercera component al igual que en la cas del MEP, reflecteix trets diferencials entre molècules. Apareixen com importants les posicions 1, 4, 6 i 8. Apareix en aquest cas també l'àrea sobre la densitat electrònica de l'anell aromàtic de la dreta.

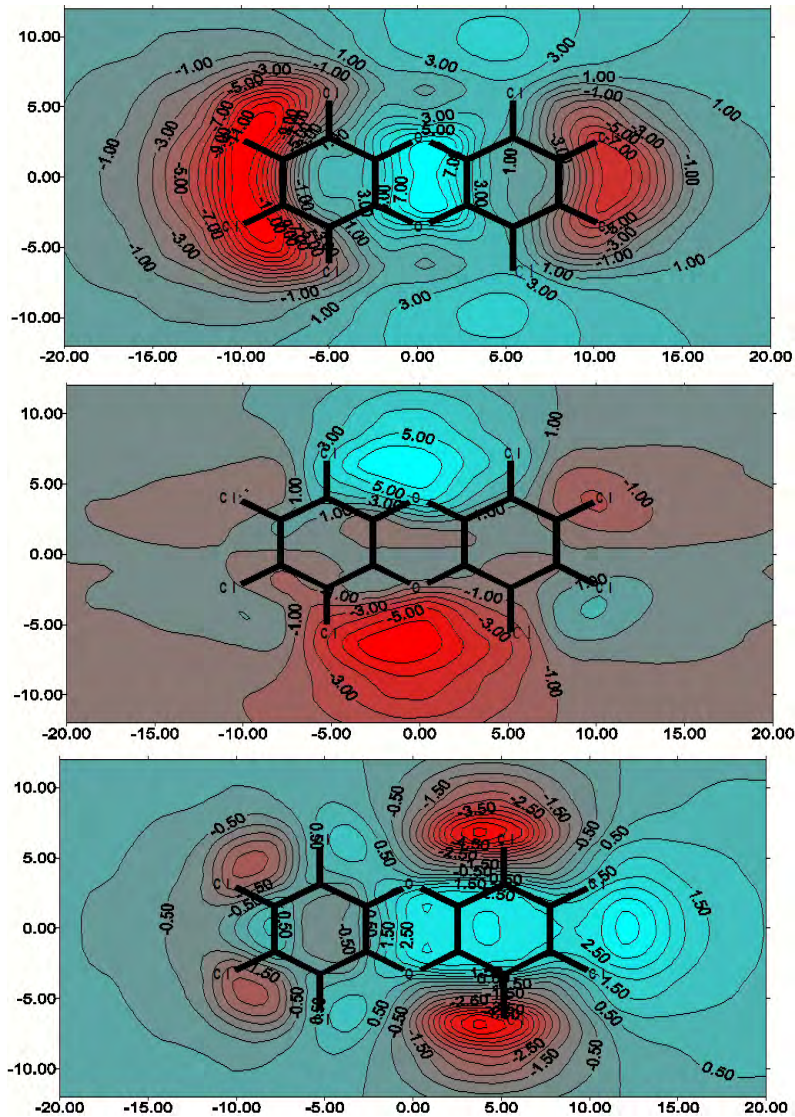


Figura 5.4.41. Representació gràfica dels components principals (loadings) corresponent al MEI del conjunt total barrejat de dioxines i derivats dels furans

A la Taula 5.4.8 es mostren els valors de pes *-scorings-* pel cas del MEP de cada molècula del conjunt de dioxines i furans per a cadascun dels components identificats, calculats com el producte de la primera fila de la matriu de dades inicial per les components del vector propi corresponent:

$$S_i = \sum A_{ik} c_k$$

Taula 5.4.8. Valors de pes (*scorings*) per el conjunt de MEPs de dioxines i furans

Anàleg	M1.p[1]	M1.p[2]	M1.p[3]
Tcdd	0,20742	-0,201493	-0,167731
12378dd	0,232295	-0,0957658	-0,00634691
2367dd	0,213764	0,0693734	-0,235644
236dd	0,155383	0,307192	-0,106535
123478dd	0,211262	-0,156834	0,176578
1378dd	0,207046	-0,042781	0,0156941
12478dd	0,20154	-0,0485486	0,00127355
1234dd	0,107156	0,360356	0,295067
237dd	0,214603	0,129796	-0,220861
28dd	0,124086	0,188616	-0,416841
12347dd	0,210043	0,131403	0,157367
124dd	0,0624224	0,425078	0,199617
ocdd	0,17598	-0,21202	0,299255
1mcdd	-0,0517952	0,44997	-0,0445557
23478df	0,212748	-0,10035	0,0438325
2347df	0,19564	0,169131	-0,0174877
2378df	0,191545	-0,139624	-0,284377
234678df	0,203286	-0,11651	0,0418187
12378df	0,226959	-0,061523	-0,174825
1237df	0,193573	0,117458	-0,0254577
13478df	0,221294	-0,0422669	-0,0745642
23479df	0,203232	0,123914	0,213308
2348df	0,158831	0,150137	-0,0155489
123478df	0,215983	-0,158986	0,0699311
123678df	0,226857	-0,131343	0,0398927
12379df	0,203989	0,133629	0,0631693
12478df	0,187096	0,0202984	0,0295302
12468df	0,0984639	-0,0692678	0,478608

Aquests valors permeten representar gràficament les projeccions de les diferents molècules sobre l'espai definit per dues components principals. A la Figura 5.4.42 es mostren les projeccions dels *scorings* més representatives, sobre les components p1 vs p2, p1 vs p3 i p2 vs p3 respectivament.

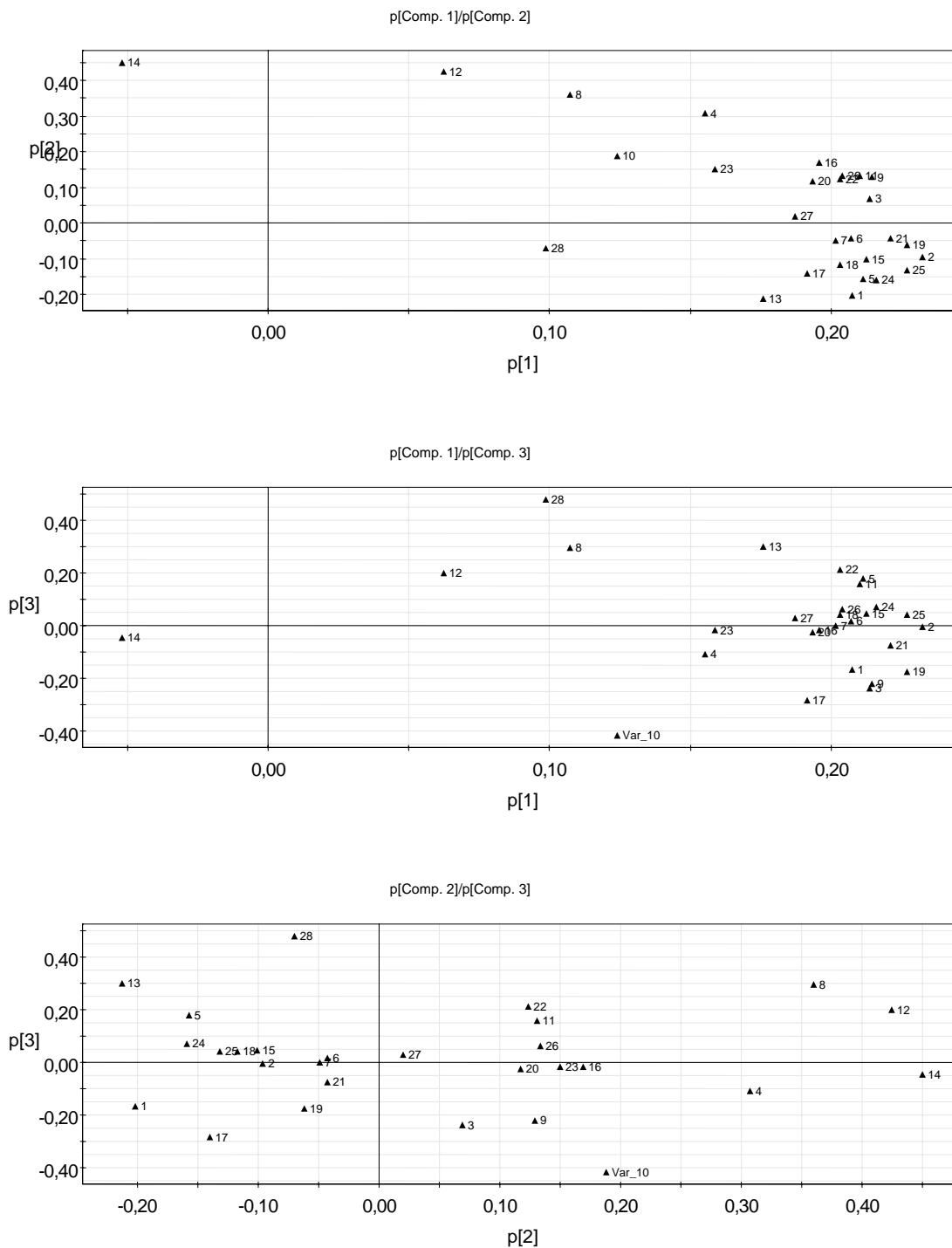


Figura 5.4.42 *Projecció dels scorings dels MEP de cada molècula sobre les diferents components principals corresponent al conjunt de dioxines i furans: a) projecció sobre les components p1 i p2; b) projecció sobre les components p1 i p3; c) projecció sobre les components p2 i p3.*

Una anàlisi de la Figura 5.4.42 mostra que la projecció p2/p3 és la que ens dóna més informació sobre aquest grup de molècules. Per una banda, apareixen juntes les molècules 2, 5, 15, 18, 24 i 25, no molt lluny d'aquest conjunt es troben les molècules 6,7,19, 21 i 27. Més allunyades d'aquest dos grups es troben les molècules 1, 13, 17 i 28. En un altre grup es poden

distingir les molècules 3, 9, 11, 16, 20, 22, 23 i 26, i un altre grup el podrien definir les molècules 4, 8, 12 i 14.

Seguint el mateix procediment hem calculat els *scorings* dels mapes MEI per a cada una de les molècules del conjunt de dioxines i furans, es troben llistats a la Taula 5.4.9.

Taula 5.4.9. Valors de pes (scorings) per el conjunt dels mapes MEI del conjunt de dioxines i furans

Anàleg	M1.p[1]	M1.p[2]	M1.p[3]
Tcdd	0,205527	-0,218536	-0,171823
12378dd	0,232113	-0,148645	0,0496753
2367dd	0,229611	0,0271849	-0,215379
236dd	0,182808	0,264071	-0,120504
123478dd	0,193811	-0,200278	0,232879
1378dd	0,201177	-0,05259	-0,0181443
12478dd	0,185545	-0,050343	-0,0831971
1234dd	0,125201	0,342506	0,251209
237dd	0,231898	0,0929673	-0,193198
28dd	0,154754	0,185831	-0,352595
12347dd	0,208813	0,101511	0,205155
124dd	0,0872966	0,409145	0,128785
ocdd	0,132513	-0,241884	0,321747
1mcdd	0,0208139	0,44918	-0,0366914
23478df	0,210576	-0,128401	0,113281
2347df	0,210403	0,131612	0,0632308
2378df	0,194428	-0,135465	-0,294519
234678df	0,194077	-0,151627	0,0922289
12378df	0,233473	-0,0828837	-0,17946
1237df	0,204182	0,123349	0,0112004
13478df	0,220396	-0,0342587	-0,128581
23479df	0,198961	0,134178	0,223053
2348df	0,166336	0,147476	-0,00240931
123478df	0,206781	-0,182523	0,118137
123678df	0,222977	-0,160542	0,0518003
12379df	0,207937	0,119938	0,0969969
12478df	0,168393	0,0497672	-0,0708613
12468df	0,04985	0,0411186	0,464273

Igualment, com en els casos anteriors podem fer les projeccions dels valors dels *scorings* sobre les components principals, es troben a la Figura 5.4.43.

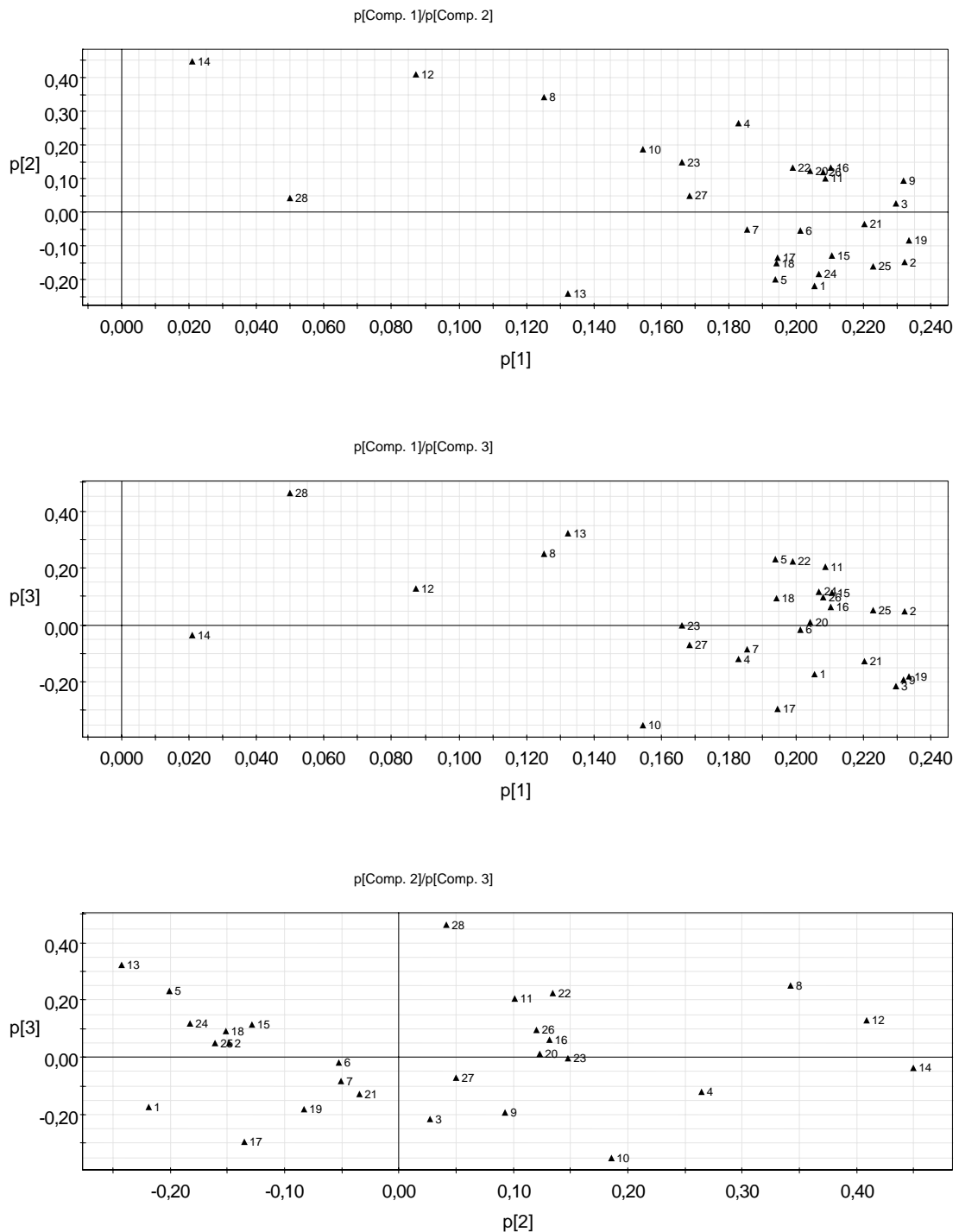


Figura 5.4.43 Projecció dels scorings dels MEI de cada molècula sobre les diferents components principals corresponent al conjunt de furans: a) projecció sobre les components p1 i p2; b) projecció sobre les components p1 i p3; c) projecció sobre les components p2 i p3.

L'anàlisi de la Figura 5.4.43 mostra que la projecció p2/p3 és la que permet classificar millor les molècules objecte d'estudi. Com es pot veure la Figura 5.4.43c és molt semblant a la 5.4.42c, encara que les agrupacions queden molt més clares. Podem trobar en el mateix grup les molècules 1, 2, 5, 13, 15, 18, 24 i 25, totes elles caracteritzades per un nombre elevat de clors al voltant de les posicions 2, 3, 7 i 8. Un altre grup inclou les molècules 6, 7, 17, 19 i 21, que tenen una distribució de clors semblant però menys simètrica que al grup anterior. Un altre grup conté les molècules 3, 9, 10, 11, 16, 20, 22, 23, 26, 27 i 28, caracteritzades per tenir pocs àtoms de clor a l'anell de la dreta. També es pot observar el grup format per les molècules 8, 12 i 14, sense àtoms de clor al anell de la dreta, i finalment la molècula 4, que apareix com aïllada.

5.5 Càlculs d'ajust parcial per mínims quadrats (Partial Least Square)

Com ja vam comentar a l'introducció d'aquest capítol es va utilitzar el mètode ajust parcial per mínims quadrats (*Partial Least Squares*) (PLS) per tal de correlacionar l'activitat biològica de dioxines i furans amb els mapes de potencial electrostàtic, de polarització molecular i d'interacció. Com en el cas de l'anàlisi de components principals exploratori, la tècnica PLS consisteix en portar a terme una anàlisi de components principals, on les components s'ajusten per tal de descriure els valors d'activitat biològica (en aquest cas la toxicitat) del conjunt de molècules objecte d'estudi. En aquest cas per determinar la dimensió del model, és a dir, el nombre de components significants, s'utilitza el mètode de validació creuada (*cross-validation*) en el que es treuen punts de les dades de partida per poder analitzar la precisió de la predicció que es fa d'aquests amb la resta de dades.

5.5.1 Ajust mitjançant PLS i predicció de toxicitat

A la Taula 5.5.1. es presenten els valors del coeficient de regressió de predicció dels valors experimentals assolits amb les diferents components (R^2), considerant els MEPs, MPPs i els MEIs. Com es pot veure, amb tres components s'obté una predicció dels valors experimentals molt acurada, de manera que podem deduir que la utilització de l'ajust parcial per mínims quadrats realitzada amb els descriptors indicats és totalment representativa en la determinació de l'activitat biològica. Com s'ha comentat abans, la fiabilitat predictiva del model es pot estimar en base al valor del coeficient de correlació de validació creuada després de treure punts del conjunt amb qual es fa la correlació, Q^2 , definit com:

$$Q^2 = 1 - \frac{\sum (y_{i,obs} - y_{i,pred})^2}{\sum (y_{i,obs} - \overline{y_{i,pred}})^2}$$

Aquest resultat coincideix amb els resultats obtinguts a l'apartat anterior relatiu a l'anàlisi de components principal exploratori, on veiem que amb tres components es podia explicar un 95% de la variabilitat del sistema. A les Figures 5.5.1-5.5.3 es mostren les gràfiques de correlació entre els valors del pEC_{50} obtinguts i els experimentals per a cada grup de molècules i considerant el MEP, MPP, o bé el MEI.

Taula 5.5.1 Valors del coeficient de regressió de pEC_{exp}/pEC_{calc} utilitzant diferents components de l'anàlisi de components principals

	MEP		MPP		MEI	
	R ²	Q ²	R ²	Q ²	R ²	Q ²
DIOXINES						
C1	0.8633	0.575	0.7691	0.502	0.9026	0.521
C2	0.9392	0.600	0.8880	0.516	0.9472	0.593
C3	0.9438	0.653	0.9775	0.545	0.9559	0.609
FURANS						
C1	0.9533	0.506	0.9181	0.482	0.9424	0.517
C2	0.9780	0.658	0.9637	0.509	0.9801	0.662
C3	0.9741	0.745	0.9900	0.643	0.9919	0.783
DIOXINES I FURANS						
C1	0.8206	0.514	0.6198	0.481	0.8246	0.546
C2	0.8789	0.532	0.8457	0.506	0.8751	0.550
C3	0.9500	0.672	0.9590	0.618	0.9710	0.689

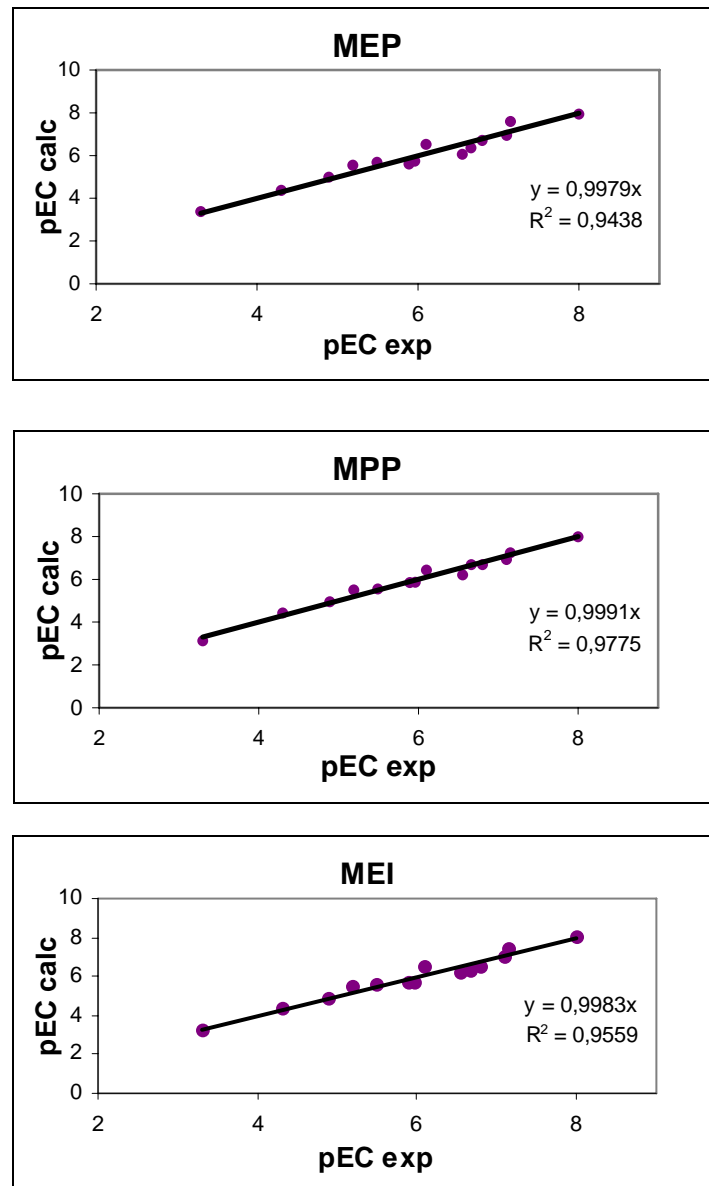


Figura 5.5.1. Representació gràfica de la relació pECcalculat en funció del pECexperimental pell MEI, MPP i MEP, per al conjunt de dioxines.

Si comparem els càlculs PLS realitzats amb el potencial electrostàtic i el potencial de interacció, s'observa, per a tots els conjunts observats, una millora lleugera en la correlació obtinguda entre els valors experimentals i calculats de la l'activitat biològica, el que ens permet deduir la importància de la inclusió del potencial de polarització en els càlculs (Roset, 2007).

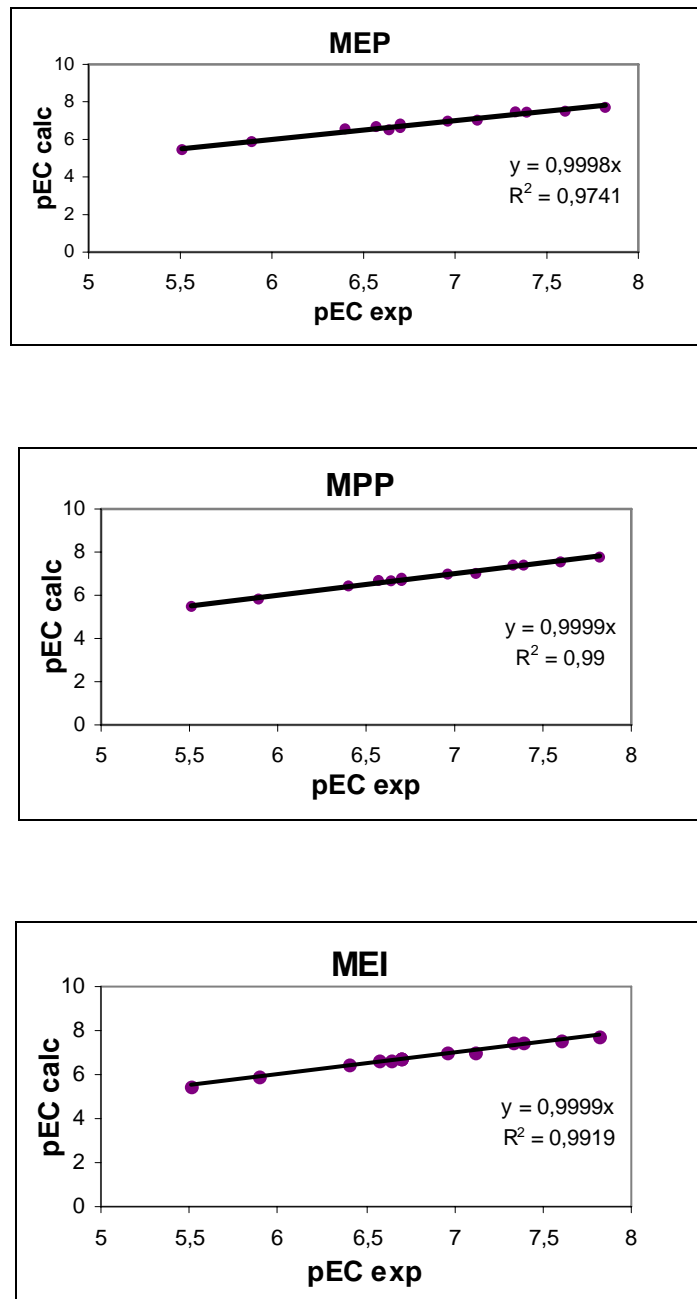


Figura 5.5.2. Representació gràfica de la relació pEC calculat en funció del pEC experimental per al cas del MEI, MPP i MEP per al conjunt de furans.

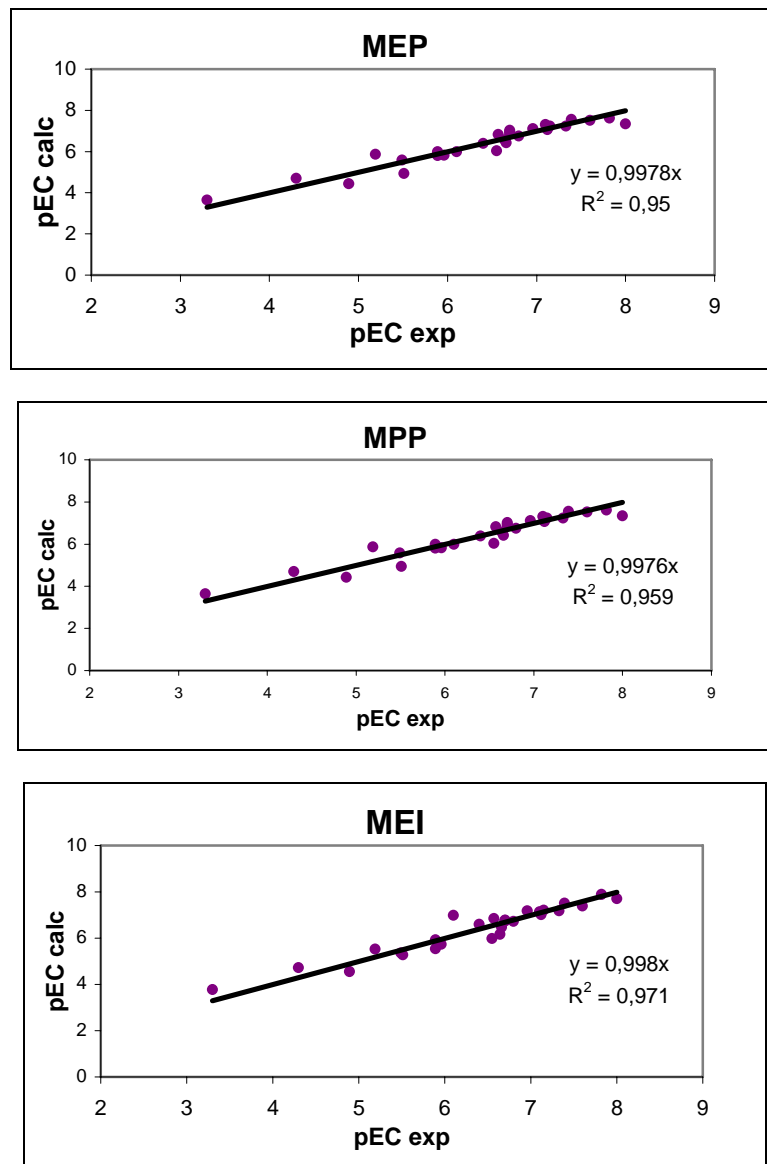


Figura 5.5.3. Representació gràfica de la relació pECcalculat en funció del pECexperimental pel MEI, MPP i MEP per al conjunt de dioxines i furans.

A continuació fem una anàlisi més detallada per a cada component i grup, i es representen els valors de toxicitat esperats amb el model obtinguts per càlculs PLS i els valors de l'activitat biològica experimental. Com és pot veure l'error comès a l'avaluació del valors del pEC_{50} és petit, de l'ordre de 0.2 unitats logarítmiques.

Taula 5.5.2. Contribució als valors pEC₅₀ per a cada una de les components principals. Resultats per al MEP de les dioxines.

Dioxines	pEC_C1	pEC_C2	pEC_C3	pECexp
tcdd	7,35058	7,29381	7,91575	8
12378ddxy	6,625	6,4701	6,84276	7,10
2367ddxy	6,90688	7,00725	6,58241	6,80
236ddy	6,20589	6,81231	6,28836	6,66
123478ddx	6,21703	5,81035	6,13806	6,55
1378ddxy	6,59603	6,61534	6,60779	6,10
12478dd	6,42448	6,18635	5,75513	5,96
1234dd	4,93859	5,64528	5,75358	5,89
237dd	6,97874	7,40097	7,58071	7,15
28dd	5,76268	5,54677	5,49683	5,49
12347ddy	5,92498	5,93194	5,69087	5,19
124dd	4,34896	4,72744	4,99162	4,89
ocdd	5,64698	4,76795	4,48664	4,30
1mcddx	3,45319	3,16415	3,24949	3,30

Taula 5.5.3. Contribució als valors pEC₅₀ per a cada una de les components principals. Resultats per al MPP de les dioxines.

Dioxines	pEC_C1	pEC_C2	pEC_C3	pECexp
tcdd	6,85339	7,65237	7,71842	8
12378ddxy	6,7819	6,96375	7,14499	7,10
2367ddxy	6,33357	6,80214	6,59076	6,80
236ddy	5,70339	6,19598	6,2378	6,66
123478ddx	6,61762	6,09981	6,51376	6,55
1378ddxy	6,31487	6,57311	6,30486	6,10
12478dd	6,29092	5,90648	5,36478	5,96
1234dd	5,16819	4,80784	5,58134	5,89
237dd	6,19641	7,17748	7,39608	7,15
28dd	5,55207	6,28454	5,80362	5,49
12347ddy	5,96467	5,68758	6,18043	5,19
124dd	4,59416	4,10891	4,45126	4,89
ocdd	6,73776	4,87053	4,45598	4,30
1mcddx	4,27108	4,24947	3,6359	3,30

Taula 5.5.4. Contribució als valors pEC₅₀ per a cada una de les components principals. Resultats per al MEI de les dioxines.

Dioxines	pEC_C1	pEC_C2	pEC_C3	pECexp
tcdd	6,5708	6,52357	6,53451	8
12378ddxy	4,83804	5,60947	5,61780	7,10
2367ddxy	6,8765	7,4722	7,42368	6,80
236ddy	6,10402	6,74486	6,69385	6,66
123478ddx	4,38233	4,77143	5,86308	6,55
1378ddxy	5,89528	5,77501	5,93778	6,10
12478dd	7,38162	7,47401	6,94620	5,96
1234dd	6,74574	6,66705	6,13705	5,89
237dd	6,83041	6,98732	6,99365	7,15
28dd	6,21608	5,79299	5,60452	5,49
12347ddy	6,44179	6,01616	5,92714	5,19
124dd	5,77934	5,80341	5,82466	4,89
ocdd	5,7863	4,52133	4,40723	4,30
1mcddx	3,53175	3,22119	3,21965	3,30

Taula 5.5.5. Contribució als valors pEC₅₀ per a cada una de les components principals. Resultats per al MEP dels furans.

Furans	pEC_C1	pEC_C2	pEC_C3	pECexp
23478df	7,29702	7,63195	7,64057	7,82
2347dfxy	7,44455	7,61286	7,60926	7,60
2378dfy	7,43777	7,28386	7,28991	7,39
234678df	7,18873	7,37336	7,36920	7,33
12378dfxy	7,29128	7,06491	7,06475	7,12
1237dfy	6,96541	7,05993	7,06824	6,96
13478dfx	6,76259	6,58955	6,59147	6,70
23479dfx	6,45407	6,5469	6,49520	6,70
2348dfxy	6,7695	6,6853	6,68924	6,70
123478dfxy	6,74842	6,79766	6,79838	6,64
123678dfy	6,77952	6,72305	6,72364	6,57
12379dfy	6,77741	6,69196	6,69196	6,40
12478df	6,16956	5,78476	5,82530	5,89
12468dfy	5,24417	5,48396	5,49713	5,51

Taula 5.5.6. Contribució als valors pEC₅₀ per a cada una de les components principals. Resultats per al MPP dels furans.

Furans	pEC_C1	pEC_C2	pEC_C3	pECexp
23478df	7,43718	7,57252	7,61095	7,82
2347dfxy	7,51186	7,32409	7,48969	7,60
2378dfy	7,14118	7,44439	7,52484	7,39
234678df	6,93531	7,30488	7,3412	7,33
12378dfxy	6,77144	7,09652	7,07674	7,12
1237dfy	7,18033	6,94697	7,02527	6,96
13478dfx	6,37981	6,55964	6,64479	6,70
23479dfx	6,90088	6,72311	6,78254	6,70
2348dfxy	7,1444	6,77522	6,77542	6,70
123478dfxy	6,98371	6,82366	6,48517	6,64
123678dfy	6,77352	6,87614	6,64645	6,57
12379dfy	6,74696	6,61652	6,58478	6,40
12478df	5,96089	6,01656	5,97743	5,89
12468dfy	5,46253	5,24977	5,36474	5,51

Taula 5.5.7. Contribució als valors pEC₅₀ per a cada una de les components principals. Resultats per al MEI dels furans.

Furans	pEC_C1	pEC_C2	pEC_C3	pECexp
23478df	7,38627	7,63499	7,70008	7,82
2347dfxy	7,32439	7,53155	7,57549	7,60
2378dfy	7,44128	7,27228	7,33678	7,39
234678df	7,15377	7,35599	7,48399	7,33
12378dfxy	7,1959	7,02735	7,0769	7,12
1237dfy	6,94679	7,03519	6,94265	6,96
13478dfx	6,73154	6,51739	6,67674	6,70
23479dfx	6,58663	6,65298	6,74814	6,70
2348dfxy	6,82949	6,73934	6,62089	6,70
123478dfxy	6,87905	6,87102	6,60134	6,64
123678dfy	6,87613	6,77211	6,63157	6,57
12379dfy	6,65774	6,69051	6,59347	6,40
12478df	6,14077	5,83148	5,91637	5,89
12468dfy	5,18026	5,39782	5,4256	5,51

Taula 5.5.8. Contribució als valors pEC₅₀ per a cada una de les components principals. Resultats per al MEP del conjunt de dioxines i furans.

Dioxines i Furans	pEC_C1	pEC_C2	pEC_C3	pECexp
tcdd	7,55002	7,41819	7,6145	8
12378ddxy	7,05741	6,99709	7,21571	7,10
2367ddxy	7,15191	7,15537	6,77974	6,80
236ddy	6,20927	6,70787	6,25959	6,66
123478ddx	6,60423	6,26461	6,38447	6,55
1378ddxy	6,75606	6,83328	6,91667	6,10
12478dd	6,52273	6,14104	5,72589	5,96
1234dd	4,93552	5,65054	5,5664	5,89
237dd	7,15045	7,45883	7,2786	7,15
28dd	5,97158	5,53183	5,2437	5,49
12347ddy	6,253	6,28836	6,01391	5,19
124dd	4,42881	4,75013	4,71836	4,89
ocdd	5,98809	5,19545	5,12838	4,30
1mcddx	3,72634	3,49974	3,64083	3,30
23478df	7,10949	7,3627	7,90066	7,82
2347dfxy	6,85466	7,44302	7,58393	7,60
2378dfy	7,30764	7,14582	7,21723	7,39
234678df	6,87511	6,74715	7,08195	7,33
12378dfxy	6,9834	6,85582	6,79949	7,12
1237dfy	6,53298	6,83451	7,19897	6,96
13478dfx	6,93996	6,79742	6,57893	6,70
23479dfx	6,32916	6,7311	6,81925	6,70
2348dfxy	6,58872	6,97845	6,79468	6,70
123478dfxy	6,58067	6,12774	6,21026	6,64
123678dfy	6,90824	6,81968	6,7927	6,57
12379dfy	6,30676	6,51681	6,6208	6,40
12478df	6,1329	5,93618	5,66704	5,89
12468dfy	4,95488	4,52125	4,95737	5,51

Taula 5.5.9. Contribució als valors pEC₅₀ per a cada una de les components principals. Resultats per al MPP del conjunt de dioxines i furans.

Dioxines i Furans	pEC_C1	pEC_C2	pEC_C3	pECexp
tcdd	7,05451	7,72673	7,16697	8
12378ddxy	7,18439	7,65054	7,28373	7,10
2367ddxy	6,57807	7,05562	6,93488	6,80
236ddy	5,93305	6,30719	6,46944	6,66
123478ddx	7,09234	6,8238	5,90502	6,55
1378ddxy	6,60345	6,85534	6,48096	6,10
12478dd	6,53225	6,05353	5,70583	5,96
1234dd	5,65407	5,50149	5,47551	5,89
237dd	6,41177	7,39475	7,13524	7,15
28dd	5,79131	6,42007	5,70336	5,49
12347ddy	6,44602	6,51379	5,95771	5,19
124dd	5,02878	4,49759	3,98777	4,89
ocdd	7,27286	5,71509	4,82394	4,30
1mcddx	4,66949	4,72271	4,12088	3,30
23478df	6,80857	7,08197	7,4838	7,82
2347dfxy	6,3197	7,05325	7,60827	7,60
2378dfy	6,62862	7,06812	7,31588	7,39
234678df	7,02358	6,83812	7,50304	7,33
12378dfxy	6,75965	6,83717	7,24573	7,12
1237dfy	6,20978	6,72217	7,16033	6,96
13478dfx	6,32346	5,96164	6,47864	6,70
23479dfx	6,23862	6,04073	6,57249	6,70
2348dfxy	6,0735	6,27839	6,58379	6,70
123478dfxy	6,78422	6,42032	6,47844	6,64
123678dfy	6,89299	6,52844	6,87703	6,57
12379dfy	6,4208	6,48946	6,96311	6,40
12478df	6,17541	5,43585	5,88638	5,89
12468dfy	5,79873	4,71614	5,40183	5,51

Taula 5.5.10. Contribució als valors pEC50 per a cada una de les components principals. Resultats per al MEI del conjunt de dioxines i furans.

Dioxines i Furans	pEC_C1	pEC_C2	pEC_C3	pECexp
tcdd	6,86654	6,8891	7,13137	8
12378ddxy	5,06057	5,79007	5,63636	7,10
2367ddxy	7,1876	7,53056	7,28911	6,80
236ddy	6,24848	6,60276	6,38557	6,66
123478ddx	4,56425	4,69229	4,46806	6,55
1378ddxy	6,13126	5,57657	5,24082	6,10
12478dd	7,70766	7,6092	7,78722	5,96
1234dd	7,25285	7,25003	7,29179	5,89
237dd	7,13405	6,99895	6,67459	7,15
28dd	6,71327	6,39756	6,325	5,49
12347ddy	6,61881	5,87927	5,66015	5,19
124dd	6,23786	6,30812	5,86933	4,89
ocdd	6,2031	5,13531	5,05289	4,30
1mcddx	3,91003	3,74892	3,82861	3,30
23478df	7,03856	7,50097	7,84627	7,82
2347dfxy	6,69027	7,52219	7,47013	7,60
2378dfy	7,24519	7,11884	7,26131	7,39
234678df	6,82192	6,72367	7,05079	7,33
12378dfxy	6,94781	6,78197	6,80474	7,12
1237dfy	6,4239	6,97329	7,201	6,96
13478dfx	6,67812	6,43268	6,43779	6,70
23479dfx	6,24683	6,69538	6,70731	6,70
2348dfxy	6,4056	6,96015	6,73508	6,70
123478dfxy	6,40573	6,09987	6,21439	6,64
123678dfy	6,77939	6,75026	6,80486	6,57
12379dfy	6,23551	6,56332	6,67851	6,40
12478df	6,09024	5,62134	5,6595	5,89
12468dfy	4,86461	4,55737	5,19745	5,51

Els resultats obtinguts en aquesta anàlisi es poden comparar amb altres càlculs d'estudis de relació estructura-activitat fets per altres mètodes sobre aquestes molècules i publicades a la literatura. A la Taula 5.5.11 hi ha un resum d'aquest estudi comparatiu. Així per exemple, amb càlculs de regressió lineal sobre un conjunt de més d'una trentena de benzofurans utilitzant diversos descriptors monodimensionals derivats del potencial d'ionització i de l'afinitat electrònica s'obtenen regressions lineals amb coeficients al voltant del 0.75 (Arulmozhiraja, 2004; Hirokawa, 2005). En un treball del mateix grup utilitzant descriptors de moment quadrupolar, polaritzabilitats i de les energies del HOMO i LUMO s'obtenen resultats semblants amb un coeficient de regressió de 0,75 (Hirokawa, 2005). A un treball posterior Luan i col·laboradors (Luan, 2006), utilitzant la mateixa metodologia, van estudiar les relacions estructura-activitat de un total de 93 compostos incloent dioxines, bezofurans i bifenils i, utilitzant diversos descriptors topològics i diferents metodologies per triar els descriptors, els autors van obtenir correlacions amb un coeficient de regressió més alt, fins d'un 0.88. En un estudi realitzat amb el mètode CoMFA

(Comparative Molecular Field Analysis) Ashek i col·laboradors (Ashek 2006) van estudiar les relacions estructura-activitat d'un total de 15 dioxines i 35 furans utilitzant PLS amb els mapes electrostàtic i hidrofòbic. Els resultats obtinguts van ser d'un coeficient de regressió de 0.88 amb un coeficient de validació creuada de 0.74. Finalment, Gu i col·laboradors (Gu, 2007; Gu,2009) van dur a terme un estudi de correlació mitjançant el mètode PLS utilitzant setze descriptors relacionats amb els moments permanents i induïts del conjunt de molècules. En aquest cas es van aconseguir millors resultats amb un coeficient de regressió de 0.88.

Taula 5.5.11. Taula comparativa de la predicció de la toxicitat de dioxines i furans d'aquest treball i comparació amb altres models publicats a la literatura. R^2 és el coeficient de regressió; Q^2 és el coeficient de correlació de validació creuada; RMSEE és la desviació quadràtica mitjana; N és el nombre de molècules utilitzades en l' estudi; No és el nombre de descriptors per a les diferents aproximacions. Per el càlculs PLS indiquem el nombre de components significatives, obtingudes prèviament a l'anàlisi, amb el mètode de cross-validation (CV).

Model	R^2	Q^2	RMSEE	N	No.	Mètode de càlcul	Referència
1	0.776		0.657	33	3	MLR	Arulmozhiraja, 2004
2	0.747	0.572		34	7	MLR	Hirokawa, 2005
3	0.841	0.814		74	5	HM	Luan, 2006
4	0.883	0.742	0.568	60		CoMFA	Ashek,2006
5	0.881	0.827	0.542	40	16	PLS	Gu, 2007
Dioxines MEP	0.944	0.653	0.351	14	3	PLS	Aquest treball
Dioxines MPP	0.977	0.686	0.223	14	2	PLS	Aquest treball
Dioxines MEI	0.956	0.609	0.423	14	3	PLS	Aquest treball
Furans MEP	0.974	0.745	0.148	14	3	PLS	Aquest treball
Furans MPP	0.990	0.743	0.137	14	2	PLS	Aquest treball
Furans MINT	0.992	0.783	0.102	14	3	PLS	Aquest treball
Dioxines i furans							
MEP	0.950	0.672	0.587	28	3	PLS	Aquest treball
MPP	0.959	0.631	0.491	28	2	PLS	Aquest treball
MINT	0.971	0.689	0.463	28	3	PLS	Aquest treball

La comparativa amb els diferents models publicats a la literatura ens permet observar que a aquest treball s'obtenen valors del coeficient de regressió R^2 millors que els proporcionats en càlculs previs, així com valors de Q^2 (valor acumulat de la fracció de la variació per validació creuada) superiors a 0.5, el que demostra que el nostre model es predictiu. Respecte al RMSEE comentar la clara disminució dels nostres càlculs en comparació amb els models anteriors deguda a una dispersió de resultats més petita.

5.5.2 Estudi de les zones de màxima variabilitat

La utilització de la tècnica PLS en els estudis d'estructura-activitat té com a objectiu fer un ajust parcial de variables estructurals per a l'obtenció del millor ajust possible als valors experimentals d'una propietat, en aquest cas la toxicitat. En el present estudi prenem com a variables independents cada un dels punts dels mapes MEP, MPP o MEI i fem el millor ajust possible a través d'una anàlisi de components principals condicionada a obtenir el millor ajust possible. D'aquestes variables utilitzades per fer la correlació n'hi ha algunes que tenen associada una variabilitat més alta del sistema i per tant, que aporten més pes a la correlació. Com cada punt dels mapes està associat a unes coordenades espacials, es poden identificar gràficament. En aquest capítol hem fet un estudi de relacions estructura-activitat utilitzant un mínim de punts dels mapes MEP, MPP i MEI triats a través de la representació dels mapes de variabilitat.

A les Figures 5.5.1-5.5.9 es mostren els mapes de variabilitat pels mapes MEP, MPP i MEI per al conjunt de dioxines, furans i els dos junts.

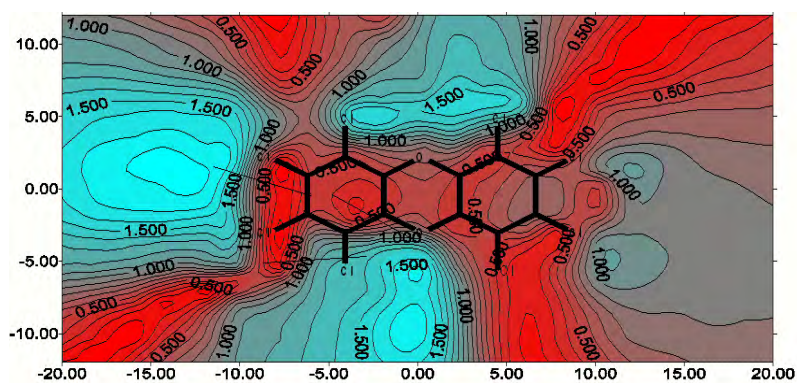


Figura 5.5.1. Mapa de la variabilitat corresponent a l'ajust PLS de les dades del MEP de dioxines.

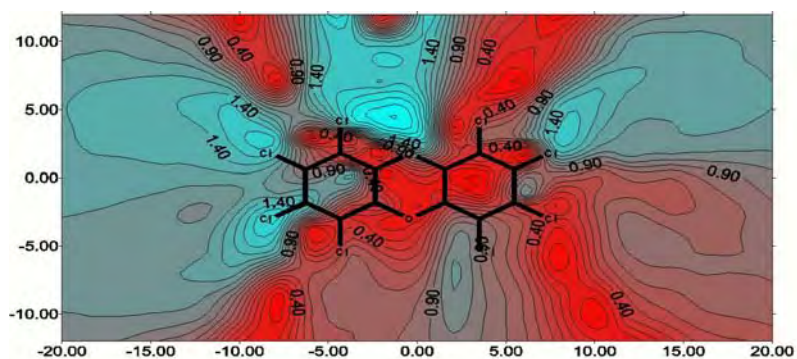


Figura 5.5.2. Mapa de la variabilitat corresponent a l'ajust PLS de les dades del MPP de dioxines.

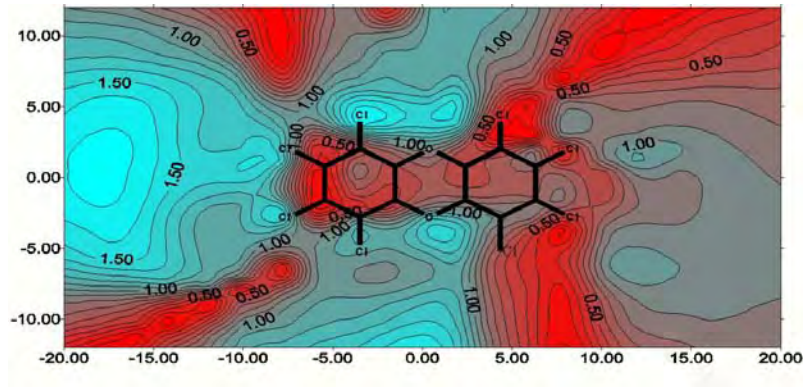


Figura 5.5.3. Mapa de la variabilitat corresponent a l'ajust PLS de les dades del MEI de dioxines.

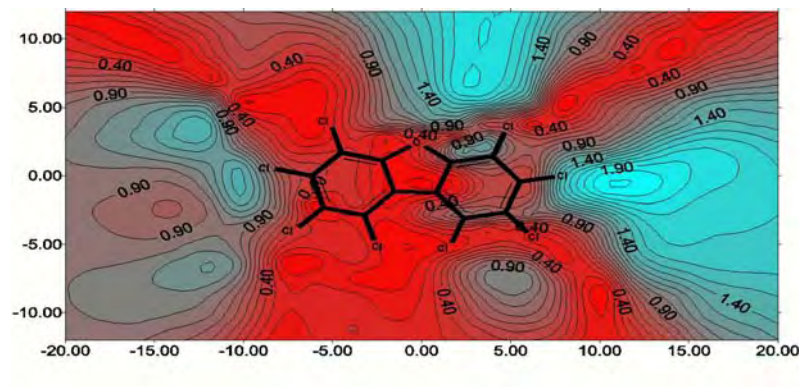


Figura 5.5.4. Mapa de la variabilitat corresponent a l'ajust PLS de les dades del MEP de furans.

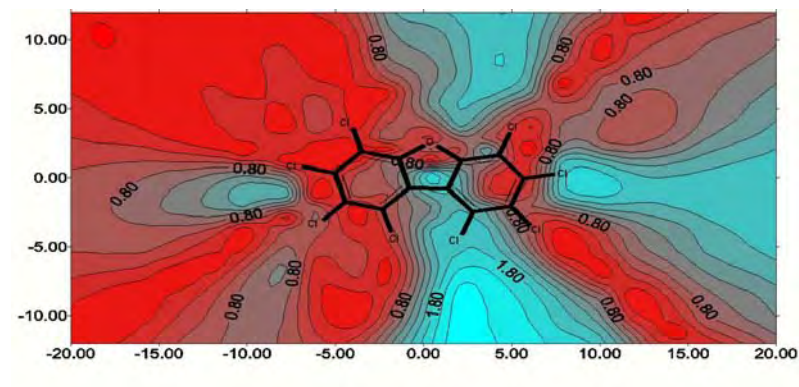


Figura 5.5.5. Mapa de la variabilitat corresponent a l'ajust PLS de les dades del MPP de furans.

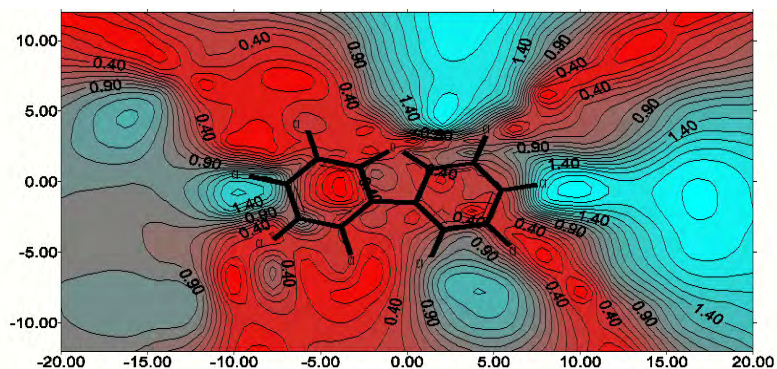


Figura 5.5.6. Mapa de la variabilitat corresponent a l'ajust PLS de les dades del MEI de furans.

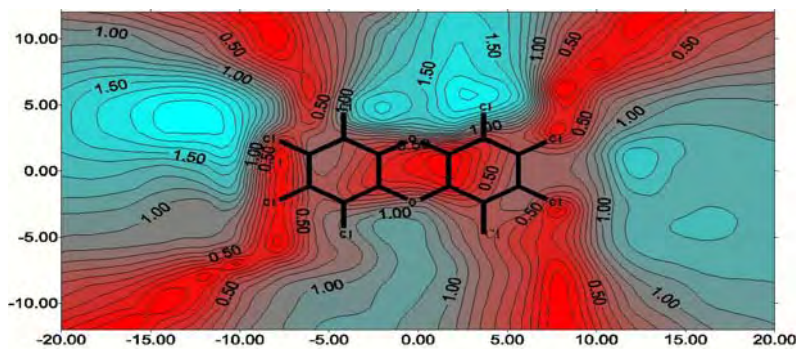


Figura 5.5.7. Mapa de la variabilitat corresponent a l'ajust PLS de les dades del MEP del conjunt de dioxines i furans.

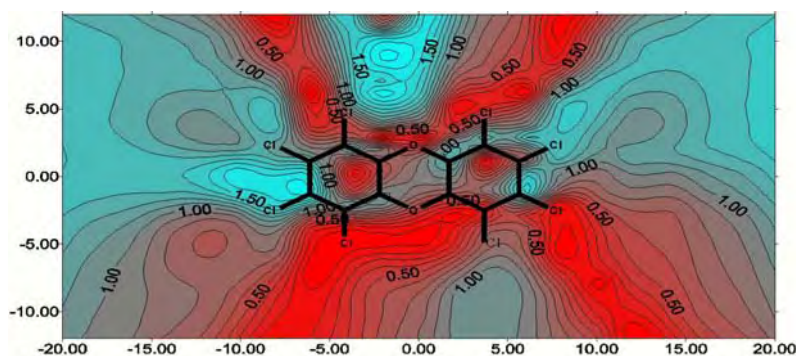


Figura 5.5.8. Mapa de la variabilitat corresponent a l'ajust PLS de les dades del MPP del conjunt de dioxines i furans.

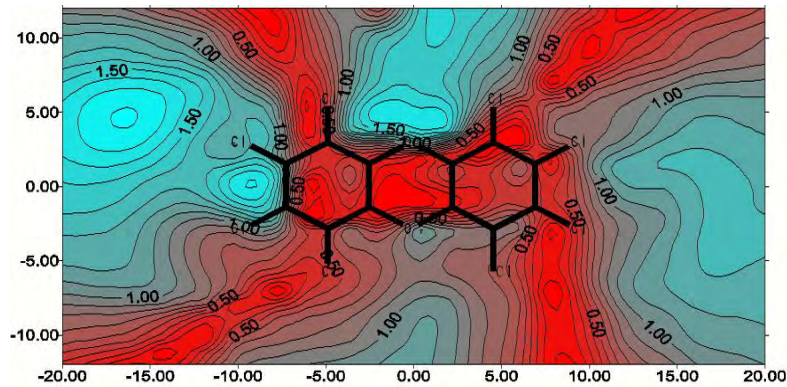


Figura 5.5.9. Mapa de la variabilitat corresponent a l'ajust PLS de les dades del MEI del conjunt de dioxines i furans.

De l'estudi d'aquests mapes es poden identificar les zones de màxima variabilitat i triar els punts més representatius, amb els que realitzarem un estudi de correlació. A continuació es presenten els resultats duts a terme en el conjunt de dioxines i furans junts.

De la inspecció de la Figura 5.5.7 varem triar un conjunt inicial de 10 punts que es troben llistats a la Taula 5.5.12.

Taula 5.5.12. Llistat dels 10 punts inicials triats del MEP del conjunt de dioxines i furans

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
	MEP	MEP	MEP	MEP	MEP	MEP	MEP	MEP	MEP	MEP
x	-14	2	12	2	-4	0	2	12	2	-4
y	4	6	1	-7	6	1	6	1	-7	6
variable	92	294	414	281	744	789	294	414	281	744

Si es fa una correlació del valor de toxicitat utilitzant aquestes variables s'obté un coeficient de regressió, $R^2 = 0.943$. Si reduïm el nombre de variables, el coeficient disminueix encara que no tots els punts tenen el mateix pes. De forma que amb un mínim de les tres millors variables s'obté un coeficient $R^2 = 0.930$.

Si ara utilitzem el potencial de polarització, a partir de la Figura 5.5.8 vam triar 10 variables de les àrees de màxima variabilitat, tal i com es mostren a la Taula 5.5.13.

Taula 5.5.13. Llistat dels 10 punts inicials triats del MPP del conjunt de dioxines i furans

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
	MPP	MPP	MPP	MPP	MPP	MPP	MPP	MPP	MPP	MPP
x	-4	0	-2	0	-8	2	6	-14	2	2
y	7	2	7	12	-1	2	0	4	6	-7
variable	220	265	245	275	162	290	338	92	294	281

Si es fa una correlació del valors de toxicitat utilitzant aquestes variables s'obté un coeficient de regressió, $R^2 = 0.927$. Si reduïm el nombre de variables triades el coeficient disminueix, de forma que amb un mínim de cinc variables s'obté un coeficient $R^2 = 0.912$.

Es pot millorar l'ajust si fem una anàlisi agafant els punts de màxima variabilitat combinats del MEP i del MPP. A la Taula 5.5.14 està llistat un conjunt de set punts triats del MEP i MPP obtenint un coeficient de correlació de $R^2 = 0.963$. Aquest conjunt es pot reduir utilitzant tres punts del MEP i un punt del MPP amb coeficient de $R^2 = 0.955$. Aquest resultat donen suport a la idea que el potencial de polarització millora la descripció del sistema.

Taula 5.5.14. Llistat d'un combinat de 6 punts triats del MEP i MPP del conjunt de dioxines i furans

	x1	x2	x3	x4	x5	x6
	MEP	MEP	MEP	MEP	MPP	MPP
x	-14	2	12	2	-4	0
y	4	6	1	-7	6	1
variable	92	294	414	281	744	789

Finalment, si s'utilitzen els mapes MEI obtenim pitjors coeficients de correlació. Així doncs triant onze punts directament del mapa de variabilitats, tal i com es mostra a la Taula 5.5.15, obtenim un coeficient de correlació $R^2 = 0.929$.

Taula 5.5.15. Llistat d'un combinat de 10 punts triats de les zones de màxima variabilitat del MEI del conjunt de dioxines i furans

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
	MEI	MEI	MEI	MEI	MEI	MEI	MEI	MEI	MEI	MEI
x	-16	-16	-14	-14	-10	-8	0	-18	18	0
y	4	6	4	6	0	0	4	0	-6	-8
variable	59	69	92	94	138	163	267	488	482	255

De tota manera, hem trobat que si utilitzem com a variables els set punts de l'anàlisi conjunt MEP/MPP que es troben a la Figura 5.5.14 s'obté una correlació de $R^2 = 0.957$. Aquest resultat és molt interessant encara que no és fàcil d'interpretar. Òbviament utilitzant els mapes MEP i MPP hi ha més informació que utilitzant únicament el MEI. Així doncs, si mirem els mapes de les Figures 5.5.7-5.5.9 veurem que el MEI conté màxims que apareixen tant al MEP com al MPP, encara que hi ha punts que no apareixen al MEI. Es pot concloure que la tria dels punts als MEI no ha d'estar basada únicament per les dades de variabilitat, si no que és força important la variabilitat dels MEP i MPP.

6 CONCLUSIONS

1.- Respecte al càlcul del valor esperat del moment dipolar s'observa que els valors calculats amb el mètode Hartree-Fock i MP2 i MP4 estan sistemàticament per sobre dels valors experimentals, essent més grans els valors Hartree Fock que els MP2 i MP4. Augmentant la base es troba que els valors calculats s'aproximen més als experimentals. En canvi, quan s'utilitza la teoria del funcional de la densitat, els valors estan per sota dels experimentals, tenint un millor comportament el funcional B3LYP.

2.- Respecte al moment quadrupolar s'observa que els resultats a nivell Hartree-Fock s'aproximen més als valors experimentals i milloren al augmentar la grandària de la base. La inclusió de la correlació electrònica empitjora els resultats sistemàticament. El funcional de la densitat dona resultats de menys qualitat que el mètode Hartree Fock

3.- Els valors de la polaritzabilitat milloren amb la inclusió de la correlació electrònica. En qualsevol cas la inclusió de funcions difuses és fonamental per obtenir resultats propers als experimentals. El funcional de la densitat sobreestima sistemàticament els valors d'aquestes propietats.

4.- Els valors de la hiperpolaritzabilitat milloren de forma desigual les diferents components del tensor i es fa palesa la necessitat d'utilitzar funcions difuses per a obtenir resultats acurats.

5.- Els mapes de potencial electrostàtic no presenten grans diferències quan s'utilitzen bases més esteses o quan es calculen a diferents nivells de teoria.

6.- La inclusió de la correlació electrònica al càlcul dels potencials de polarització millora de l'ordre del 2%, al igual que s'observa en el càlcul del tensor de polaritzabilitat. Per altre banda, la inclusió de funcions difuses millora substancialment els punts del mapa allunyats del nucli de la molècula , amb diferències fins a un 30%.

7.- Els càlculs fets amb el funcional de la densitat presenten desviaments en zones on la polaritzabilitat és petita. També s'observen sobreestimacions del potencial a les regions més llunyanes als nuclis. En aquests tipus de càlculs la dependència amb la base és petita.

8.- La comparació dels càlculs del potencial de polarització amb el mètode de la supermolècula i la teoria de pertorbacions de Rayleigh-Schrödinger permet estimar errors de fins a un 10% quan s'utilitza una base amb funcions difuses.

9.- L'anàlisi per components principals dels mapes de potencial electrostàtic i de polarització permet descriure trets comuns en grups de molècules amb una propietat biològica semblant. L'anàlisi de components principals permet deduir les característiques estereoelèctriques que descriuen el conjunt. La polarització permet una classificació més acurada.

10.- L'anàlisi pel mètode de ajust parcial per mínims quadrats utilitzant els mapes de potencial electrostàtic i de polarització permet obtenir valors alts de correlació estructura-activitat, i permet una predicció de les propietats biològiques per molècules que actuen a través del mateix receptor.

7 BIBLIOGRAFIA

- Alkorta I, Villar H, Perez J J, 1993**, *Effect of the Basis-Set on the Computation of Molecular-Polarization*. Journal of Physical Chemistry, **97**: 9113
- Alkorta I, Perez J. J, Villar H, 1994**, *Molecular polarization potential maps of the nucleic acid bases*. Journal of Molecular Graphics, **12**: 3
- Alkorta I, Perez J. J, 1995**, *Approximations in the computation of molecular polarization maps*. Electronic Journal of Theoretical Chemistry, **1**: 26
- Alkorta I, Perez, J J, 1996**, *Molecular polarization potential maps of the nucleic acid bases*. International Journal of Quantum Chemistry, **57**: 123
- Alparone A, Millefiori S, 2005**, *Gas and solution phase electronic and vibrational (hyper)polarizabilities in the series formaldehyde, formamide and urea: CCSD(T) and DFT theoretical study*. Chemical Physics Letters, **416**: 282
- Alvarado Y, 2000**, *Theoretical study of the molecular polarizability on triafulvene and analogue*. Boletín De La Sociedad Chilena De Química, **45**: 339
- Alvarado Y, 2002**, *Dipole polarizability of the pyrazabole molecule in solution*. Journal of Physical Organic Chemistry, **15**: 835
- Arulmozhiraja S, T. Fujii, 2000**, *Electron affinity for the most toxic 2, 3, 7, 8-tetrachlorodibenzo-p-dioxin (TCDD): A density functional theory study*. The Journal of Physical Chemistry A, **104**: 7068
- Arulmozhiraja S, Morita M., 2004.**, *Structure-activity relationships for the toxicity of polychlorinated dibenzofurans: approach through density functional theory-based descriptors*. Chemical Research in Toxicology, **17**: 348
- Applequist J, Carl J R, Fung K, 1972**, *Atom dipole interaction model for molecular polarizability. Application to polyatomic molecules and determination of atom polarizabilities*. Journal American Chemical Society, **94**: 2953
- Augspurger J D, Dykstra C E, 1992**, *Evolution of Polarizabilities and Hyperpolarizabilities with Molecular Aggregation, a Model Study of Acetylene Clusters*. International Journal of Quantum Chemistry, **43**: 135
- Asatryan R S , 2002** *Electronic elasticity-toxicity relationships for polychlorinated dibenzo-p-dioxin congeners*, Chemosphere **48**: 227
- Ashek A, Lee C, 2006**, *3D QSAR studies of dioxins and dioxin-like compounds using COMFA and COMSIA.*, Chemosphere **65** : 521
- Bandiera S, T. Sawyer, 1983**, *Polychlorinated dibenzofurans (PCDFs): effects of structure on binding to the 2,3,7,8-TCDD cytosolic receptor protein, AHH induction and toxicity*, Toxicology **32**:131
- Baranowska A, A Zawada, B Fernández, W Bartkowiak, D Kędziera, A Kaczmarek-Kędziera, 2010a**, *Interaction-induced electric properties and cooperative effects in model systems*, Physical Chemistry Chemical Physics **12**, 852
- Baranowska A, Sadlej A J, 2010b**, *Polarized basis sets for accurate calculations of static and dynamic electric properties of molecules*. Journal of Computational Chemistry, **31**: 552

- Barnett G**, Trsic M, Willette RE, 1978, *Quantitative Structure Activity Relationships of Analgesics, Narcotic Antagonists, and Hallucinoen*, eds QSAR, NIDA research Monograph 22. National Institute on Drug Abuse. Washington.
- Bartkowiak W**, 2002, *The influence of the solute/solvent interactions on the first-order hyperpolarizability in urea molecule. A quantum chemical study*. Chemical Physics Letters, **362**: 224
- Becke A D**, 1988, *Density-functional exchange-energy approximation with correct asymptotic behavior*. Physical Review, **A38**: 3098
- Beger, R. D.**, J.G. Wilkes, 2001, *Models of polychlorinated dibenzodioxins, dibenzofurans, and biphenyls binding affinity to the aryl hydrocarbon receptor developed using C-13 NMR data*. Journal of Chemical Information and Computer Sciences, **41**: 1322
- Benková, Z**, I. Ernušák, 2007, *Electric properties of formaldehyde, thioformaldehyde, urea, formamide, and thioformamide—Post HF and DFT study*. International Journal of Quantum Chemistry, **107**: 2133
- Bonaccorsi R**, Scrocco E, Tomasi J, 1976, *Group contributions to the electrostatic molecular potential*. Theoretical Chimica Acta , **43**: 63
- Bonati, L** 1991, *A Comparative-Analysis of Molecular Electrostatic Potentials of Chlorinated Dibenzo-P-Dioxins in Relation to Their Biological-Activity*. Journal De Chimie Physique Et De Physico-Chimie Biologique, **88**: 2631
- Bonati, L.** 1994, *Toward a Mechanistic Understanding of Pccd Biological-Activity Based on Molecular Electrostatic Potential Modeling*. Theochem-Journal of Molecular Structure, **109**: 43
- Bonati, L.** 1995, *A Hypothesis on the Mechanism of Pccd Biological-Activity Based on Molecular Electrostatic Potential Modeling .2*. Theochem-Journal of Molecular Structure, **340**: 83
- Bounds D G**, 1974, *Dipole-Moment of HCN in Crystal*. Chemical Physics Letters, **29**: 600.
- Bowen J. P**, N. L. Allinger, 1991, *Molecular mechanics: The art and science of parameterization*. Reviews in computational chemistry, Wiley & Sons. New York, NY: 81-97.
- Bruna**, 1997, *Benchmark ab initio calculations of formaldehyde, H₂CO₃*. Journal of Molecular Structure: THEOCHEM, **400**: 177.
- Buckingham A D**, 1959, *Direct Method of Measuring Molecular Quadrupole Moments* Journal of Chemical Physics, **30**: 1580.
- Buckingham A D**, 1967, *Permanent and Induced Molecular Moments and Long Range Intermolecular Forces*, Advanced Chemical Physics **12**: 107.
- Bündgen P**, Grein F, Thakkar A, 1995, *Dipole and quadrupole moments of small molecules. An ab initio study using perturbatively corrected, multi-reference configuration interaction wave functions*, Journal of Molecular Structure: THEOCHEM, **334**: 7
- Chablo A**, 1981, *Quadrupole-Moment Calculations for Some Aromatic-Hydrocarbons*. Chemical Physics Letters, **78**(3): 424
- Chen W**, Li Z R, Wu D, Li R Y, Sun C C, 2005, *Theoretical investigation of the large nonlinear optical properties of (HCN)(n) clusters with Li atom*. Journal of Physical Chemistry, **B109**: 601

- Chelkowski A, 1980, *Dielectric Physics*, Studies in physical and theoretical chemistry, Elsevier, Amsterdam, The Netherlands
- Clare B W, 2006, *QSAR of aromatic substances: toxicity of polychlorodibenzofurans*. Journal of Molecular Structure, **763**: 205
- Clark T, 1983, *Efficient Diffuse Function-Augmented Basis-Sets for Anion Calculations The 3-21+G Basis Set for 1st-Row Elements, Li-F*. Journal of Computational Chemistry, **4**: 294
- Colwell S. M, C. W. Murray, 1993,. *The determination of hyperpolarisabilities using density functional theory*. Chemical physics letters, **210**: 261
- Denomme M.A, K. Homonko, 1986, *Substituted polychlorinated dibenzofuran receptor binding affinities and aryl hydrocarbon hydroxylase induction potencies, a QSAR analysis*. Chemo-Biological Interactions, **57**: 175
- Diao J, Yang L, 2010, *QSAR Models for Predicting Toxicity of Polychlorinated Dibenzo-p-dioxins and Dibenzofurans Using Quantum Chemical Descriptors*. Bulletin of Environmental Contamination and Toxicology, **85**:109
- Dixon A, Matsuzawa N, 1994, *Density Functional Study of the Structures and Nonlinear Optical Properties of Urea*. Journal of Physical Chemistry, **98**: 3967
- Dovesi R, M. Causa, 1990, *Ab initio approach to molecular crystals: A periodic Hartree–Fock study of crystalline urea*. The Journal of Chemical Physics, **92**: 7402.
- Duffy E. M, D. L. Severance, 1993, *Urea: potential functions, log P, and free energy of hydration*. Israel journal of chemistry, **33**: 323
- Dunning T H, 1968, *Non empirical Calculations on Excited States : The Formaldehyde Molecule*. Journal of chemical Physics, **48**: 5263
- Dunning T H, 1970, *Gaussian Basis Functions for Use in Molecular Calculations. I. Contraction of (9s5p). Atomic Basis Sets for the First-Row Atoms*. Journal of Chemical Physics, **53**: 2823
- Dunning T H, 1976, *Gaussian basis sets for molecular calculations*. In: H.F. Schaefer III (Ed.), Modern Theoretical Chemistry, vol 3, Plenum, New York, pp. 1-28.
- Dunning T H, Schaefer H F, 1977, *Gaussian Basis Sets for Molecular Calculations. In Methods of Electronic Structure Theory*; Modern Theoretical Chemistry; Plenum Press: New York.
- Duijneveldt, Jeanne G. C. M. van Duijneveldt-van De Rijdt, Frans B, 1982, *Gaussian basis sets which yield accurate Hartree—Fock electric moments and polarizabilities* Original Research Article Journal of Molecular Structure: THEOCHEM, **89**: 185
- Dykstra C E, 1993, *Electrostatic Interaction Potentials in Molecular-Force Fields*. Chemical Reviews, **93**: 2339
- Engkvist O, Astrand P O, Karlstrom G, 2000, *Accurate intermolecular potentials obtained from molecular wave functions: Bridging the gap between quantum chemistry and molecular simulations*. Chemical Reviews, **100**: 4087.
- Francl M M, 1985, *Polarization Corrections to Electrostatic Potentials*. Journal of Physical Chemistry, **89**: 428

Fraschini E, Bonati L; **1995**. *A hypothesis on the mechanism of PCDD biological activity based on molecular electrostatic potential modeling. Part 2*. Journal of Molecular Structure (THEOCHEM), **340**: 83

Frisch M, **1977**, *Gaussian 94 Gaussian Inc., Pittsburgh, PA, 1995; MJS Dewar and W. Thiel*. Journal of American Chemical Society, **99**: 4899.

Frisch M J, Pople J A, Binkley J S, **1984**, *Self-consistent molecular orbital methods: XXV. Supplementary functions for Gaussian basis sets*. Journal of Chemical Physics, **80**: 3265

Frisch, M. J, Trucks, G. W, Schlegel, H. B, Gill, P. M. W, Johnson, B. G, Robb, M. A, Cheeseman, J. R, Keith, T, Petersson, G. A.; Montgomery, J. A, Raghavachari, K, Al-Laham, M. A, Zakrzewski, V. G, Ortiz, J. V, Foresman, J. B, Cioslowski, J.; Stefanov, B. B. Nanayakkara, A, Challacombe, M, Peng, C. Y, Ayala, P. Y, Chen, W, Wong, M. W, Andres, J. L, Replogle, E. S, Gomperts, R, Martin, R. L, Fox, D. J, Binkley, J. S, Defrees, D. J, Baker, J, Stewart, J. J. P, Head-Gordon, M, Gonzalez, C, Pople, J. A, **1995**, *GAUSSIAN 94*; Gaussian: Pittsburgh, PA

Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery JA, Stratman RE, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Baboul AG, Stefanov BB, Liu C, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Gonzalez C, Challacombe M, Gill PMW, Johnson BG, Chen W, Wong MW, Andres JL, Gonzalez C, Head-Gordon M, Replogle ES, Pople JA, **1998**, *Gaussian 98*. Gaussian, Pittsburgh PA.

Galvez Gonzalez O, Murray JS, Peralta-Inga Z, **2001**, *Molecular surface electrostatic potentials in relation to noncovalent interactions in biological systems*. International Journal of Quantum Chemistry, **85**: 676

Gaumann, T, **1958**, *Dielektrische Messungen an polaren Gemischen. 6. Mitteilung Harnstoffe* Helvetical Chimica Acta, **41 B**: 1956

Goldstein S, Kuik R, Lebowitz J, **1989**, *From PCA's to equilibrium systems and back*. Communications in Mathematical Physics, **125**: 71

Gready, J. E, G. B. Bacskay, N. S. Hush, **1978**, *finite-field method calculations IV. Higher order moments, dipole moment gradients, polarisability gradients and field-induced shifts in molecular properties: Application to N₂, CO, CN, HCN and HNC*. Chemical Physics, **31**: 467

Gu CG, Jiang X, **2007**, *DFT study on the structure-toxicity relationship of dioxin compounds using PLS analysis*. SAR and QSAR in Environmental Research, **18**: 2007

Gu CG, Jiang X, **2009**, *QSARs for congener-specific toxicity of polyhalogenated dibenzo-p-dioxins with DFT and WHIM theory*. Ecotoxicology and Environmental Safety, **72**: 60

Haenlein M, Kaplan A, **2004**, *A Beginner's Guide to Partial Least Squares Analysis, Understanding Statistics*, **3**: 283

Halgren T A, Damm W, **2001**, *Polarizable force fields*. Current Opinion in Structural Biology, **11**: 236

Hansch C, Steinmetz W E, **2003**, *On the Role of Polarizability in Chemical Biological Interactions*. Journal of Chemical Informatic Computational Science, **43**: 120

- Harihara P, Pople J A, 1973, *Influence of Polarization Functions on Molecular-Orbital Hydrogenation Energies*. *Theoretica Chimica Acta*, **28**: 213
- Hartford S L, Allen W C, 1973, , *The molecular Zeeman effect in HCP, HCN, and HCCBr and a comparison with similar molecules*. *Chemical Physics Letters*, **18**: 153
- Hasegawa, K, 2002, *New molecular surface-based 3D-QSAR method using Kohonen neural network and 3-way PLS*. *Computers & Chemistry*, **26**: 583
- Haskopoulos A, Maroulis G, 2006, *Interaction dipole moment in Rg-Xe (Rg =He, Ne, Ar, and Kr) heterodiatoms from conventional ab initio and density functional theory calculations*. *Journal of Mathematical Chemistry*, **40**: 233
- Hehre W J, Ditchfie R, Pople J A, 1972, *Self-Consistent Molecular-Orbital Methods .12. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular-Orbital Studies of Organic-Molecules*. *Journal of Chemical Physics*, **56**: 2257
- Hinchliffe H, Bounds D G, 1978 *Theoretical Chemistry, A Specialists Periodical Report*, R.N. Dixon and C. Thomsón. The Chemical Society. Bartholomeus Press, Dorking, U K
- Hinchliffe, A, Munn, R. W, 1985, *Molecular Electromagnetism*, John Wiley and Sons Ltd: Chichester, UK
- Hinchliffe A, 1996, *Modelling Molecular Structures*, John Wiley and Sons Ltd, Chichester, U K
- Hinchliffe A, Perez J J, Soscun Machado H J, 1997, *Density functional studies of molecular polarizabilities Part 4: the C₁₀H₈ molecules azulene, fulvalene and naphthalene*. *Electronic Journal of Theoretical Chemistry*, **2**: 325
- Hinchliffe A, Machado H J, 2000a, *Density functional studies of molecular polarizabilities Part 9; quinoline and isoquinoline*. *Asian Journal of Spectroscopy*, **4**: 21
- Hinchliffe A, Machado H J, 2000b *Density functional studies of molecular polarizabilities. Part 8. Acenaphthene and acenaphthylene*. *Chemical Physics Letters*, **321**: 151
- Hirokawa, S., Imasaka, 2005, *Chlorine substitution pattern, molecular electronic properties, and the nature of the ligand-receptor interaction: quantitative property-activity relationships of polychlorinated dibenzofurans*. *Chemical Research in Toxicology*, **18** : 232
- Hohenberg P, Kohn W, 1964, *Inhomogenous Electron Gas*. *Physical Review* **B864**:136,
- Hobza P, Zahradník R, 1981, *Formation of hydrogen-bonded van der waals molecules:applicability of 3-21G intermolecular interaction energies*. *Chemical Physics Letters*, **82**:473
- Hurdis E C 1943, *The Structural Effects of Unsaturation and Hyperconjugation in Aldehydes, Nitriles and Chlorides as Shown by their Dipole Moments in the Vapor State*. *Journal of American Chemical Society*, **65**: 89.
- Hussein W, Walker CG, Peralta-Inga Z, Murray JS, 2001, *Computed electrostatic potentials and average local ionization energies of the molecular surfaces of some tetracyclines*. *International Journal of Quantum Chemistry*, **82**:160
- Huzinaga S, Y. Sakai, 1969, *Gaussian Type Functions for Polyatomic Systems II*. *Journal of Chemical Physics*, **50**: 1371

- Jacquemin D, Andre J M, Perpete E A, **2004**, *Geometry, dipole moment, polarizability and first hyperpolarizability of polymethineimine: An assessment of electron correlation contributions*. Journal of Chemical Physics, **121**: 4389.
- Jaszunski M, Jorgensen P, Koch H, Agren H, Helgaker T, **1993**, *Frequency-Dependent Hyperpolarizabilities of Polynes*. Journal of Chemical Physics, **98**: 7229.
- Jensen L, **2000**, *Frequency-dependent molecular polarizability calculated within an interaction model*. Journal of Physical Chemistry A, **104**: 1563.
- Kafafi S.A., H.Y. Afeefy, **1993**, *Relationship between aryl hydrocarbon receptor binding, induction of aryl hydroxylase and 7-ethoxyresorufin O-deethylase enzymes, and toxic activities of aromatic xenobiotics in animals. A new model*. Chemical Research in Toxicology, **6**: 328
- Keith L, Choudhary G, Rappe C, **1983**, *Chlorinated Dioxins and Dibenzofurans in the total environment*. Butterworth Publ, 80 Montavele, Stoneham, USA.
- Kohn W, L. J. Sham, **1965**, *Self-Consistent Equations Including Exchange and Correlation Effects*. Physical Review, **140**: A1133
- Kramer, R **1998**, *Chemometric Techniques for Quantitative Analysis*, Marcel Dekker: New York.
- Kukolich S. G, **1971**, *Molecular Beam Measurement of the Magnetic Susceptibility Anisotropies and Molecular Quadrupole Moment in HCO*. The Journal of Chemical Physics, **54**: 8.
- Landolt-Bornstein, **1951**, *Atom und Molecularphysik*, Springer Verlag, Berlin, Vol 1 part 3
- Langhoff P W, Hurst R. P, **1965**, *Multipole Polarizabilities and Shielding Factors from Hartree-Fock Wave Functions*. Physical Review, **A139**: 1415
- Ledoux I, Zyss J, **1982**, *Influence of the Molecular Environment in Solution Measurements of the 2Nd-Order Optical Susceptibility for Urea and Derivatives*. Chemical Physics, **73**: 203
- Luan, F. **2006**, *QSAR study of polychlorinated dibenzodioxins, dibenzofurans, and Biphenyls using the heuristic method and support vector machine*. Qsar & Combinatorial Science, **5**: 46
- Luque F J, Orozco M, Illas F, Rubio J, **1991**, *Effect of electron correlation on the electrostatic potential distribution of molecules*. Journal of American Chemical Society, **113**: 5203
- Ma Y G, Politzer P, **2004**, *Calculation of electrostatic and polarization energies from electron densities*. Journal of Chemical Physics, **120**: 3152
- Maroulis G, **1998a**, *A systematic study of basis set, electron correlation, and geometry effects on the electric multipole moments, polarizability and hyperpolarizability of HCl*. Journal of Chemical Physics, **108**: 5432
- Maroulis G, Pouchan C, **1998b**, *Molecules in static electric fields: Linear and nonlinear polarizability of HC N and HC P*. Physical Review, **A57**: 2440
- Masia M, Probst M, Rey R, **2005**, *On the performance of molecular polarization methods close to a point charge*. Computer Physics Communications, **169**: 331
- Masón G, Farrell K, **1986**. *Polychlorinated dibenzo-p-dioxins: quantitative in vitro and in vivo structure-activity relationships*. S. Toxicology, **41**:21.

- McKinney J. D, Darden T, Lyerly J. A, Pedersen L. G, 1984, *Dioxin and related compound binding to the Ah receptor(s). Theoretical model based on molecular parameters and molecular mechanics*. Quantitative Structure Activity Relationships 1985, **4**:166
- McLean, A, M. Yoshimine, 1967, *Computed Ground State Properties of FH and CH*. The Journal of Chemical Physics, **47**: 3256.
- Merchan M, Roos B, 1995, *A theoretical determination of the electronic spectrum of formaldehyde*. Theoretical Chemistry Accounts: Theory, Computation and Modeling (Theoretica Chimica Acta), **92**: 227
- Mhin, B.J., J. Choi, W. Choi, 2001, *A simple rule for classification of polychlorinated dibenzo-p-dioxin congeners on the basis of IR frequency patterns*. Journal of the American Chemical Society, **123**: 3584
- Mhin, B.J., J.E. Lee, W.Y. Choi, 2002, *Understanding the congener-specific toxicity in polychlorinated dibenzo-p-dioxins: Chlorination pattern and molecular quadrupole moment*. Journal of the American Chemical Society, **124**: 144
- Miller K J, 1990, *Calculation of the Molecular Polarizability Tensor*. Journal of the American Chemical Society, **112**: 8543
- Miller Francl M, QCPE No. 490 Quantum Chemistry Program Exchange, Department of Chemistry, Indiana University, Bloomington, IN 47405, USA
- McClellan A L, 1963, *Tables of Experimental Dipole Moments*, Freeman. San Francisco, USA
- Møller C, Plesset M.S, 1934, *Note on an Approximation Treatment for Many Electron Systems*. Physical Review, **46**: 618
- Murray J.S, Evans P, Politzer P. 1987, *A comparative analysis of the electrostatic potentials of some structural analogues of 2,3,7,8-Tetrachlorodibenzo-p-dioxin and of related aromatic systems*. International Journal Quantum Chemistry, **37**: 271
- Murray J S, Seminario J M, Politzer P, Sjoberg P, 1990, *Average Local Ionization Energies Computed on the Surfaces of Some Strained Molecules*. International Journal of Quantum Chemistry, **38**: 645
- Naray-Szabo G, Ferenczy GGN,1995, *Molecular Electrostatics*. Chemical Review, **95**: 829.
- Neuman, Moskowitz W, 1969, *One Electron Properties of Near Hartree-Fock Wavefunctions. II. HCHO, CO*. Journal of Chemical Physics, **50**: 2216
- Oberg, T, J. Bergstrom 1989, *Indicator Parameters for Pcdd/Pcdf*. Chemosphere, **19**: 337
- Oprea, T.I, 2001, *MTD-PLS: A PLS-based variant of the MTD method. A 3D-QSAR analysis of receptor affinities for a series of halogenated dibenzoxin and biphenyl derivatives*. Sar and Qsar in Environmental Research, **12**: 75
- Parr R G, Yang W, 1989, *Density-functional theory of atoms and molecules*. Oxford Univ. Press, Oxford, UK.
- Patra S. K, 1999, *State of aggregation of bilirubin in aqueous solution: principal component analysis approach*. Journal of Photochemistry & Photobiology A:Chemistry, **122**:23

- Pérez J J, Dupuis, 1991, *Hydrogen-Bonds and (Hyper)Polarizabilities in Molecular-Crystals an Abinitio Scf Study of Urea*. Journal of Physical Chemistry, **95**(17): 6525
- Pluta T, A. J. Sadlej, 1998, *HyPol basis sets for high-level-correlated calculations of electric dipole hyperpolarizabilities*. Chemical Physics Letters, **297**: 391
- Pluta T, Sadlej A J, 2001, *Elèctric properties of urea and thiourea*. Journal of Chemical Physics, **114**: 136
- Poland A. Knutsón JC. 1982, 2,3,7,8-Tetrachlorodibenzo-p-dioxin and related halogenated aromatic hydrocarbur : examination of the mechanism of toxicity. Annual Review of Pharmacology and toxicology, **22**: 517
- Politzer, P, Truhlar, D. G, 1981, *Eds. Chemical Applications of Atomic and Molecular Electrostatic Potentials*, Plenum: New York, NY USA
- Politzer P, 1988, *Computational approaches to the identification of suspect toxic molecules*. Toxicological Letters , **43**: 257
- Politzer P, Murray J S, 1991, In: (Lipskowitz K B , Boyd D B Eds) *Reviews of Computational Chemistry Vol 2*, VCH Publishers, New York
- Poso A, K Tuppurainen, 1993. *Binding of some dioxins and dibenzofurans to the Ah receptor. A QSAR model based on comparative molecular field analysis (CoMFA)*. Journal of Molecular Structure (Theochem), **282** : 259
- Procopio, M, 2002, *A model for recognition of polychlorinated dibenzo-p-dioxins by the aryl hydrocarbon receptor*. European Journal of Biochemistry, **269**: 13
- Putzrath, R.M, 1997 , *Estimating relative potency for receptor-mediated toxicity: Reevaluating the toxicity equivalence factor (TEF) model*. Regulatory Toxicology and Pharmacology, **25**: 68
- Rabinowitz, J.R., S.B. Little, E.M. Gifford, 1998, *Interactions between chlorinated dioxins and a positively charged molecular probe: New molecular interaction potential*. Journal of Computational Chemistry, **19**: 673
- Reimers J, Zeng J, Hush N, 1996, *Vibrational Stark Spectroscopy. 2. Application to the CN Stretch in HCN and Acetonitrile*. Journal of Physical Chemistry, **100**: 1498
- Reis V, M. G. Papadopoulos, R. W. Munn, 1998, *Calculation of macroscopic first-, second-, and third-order optical susceptibilities for the urea crystal*. Chemical Physics, **109**: 6828
- Rice J E, Amos R D, Colwell S M, Handy N C, Sanz J, 1990, *Frequency dependent hyperpolarizabilities with application to formaldehyde and methyl fluoride*. Journal of Chemical Physics, **93**: 8828
- Rick S W, Stuart S J, 2002, *Potentials and algorithms for incorporating polarizability in computer simulations*. Reviews in Computational Chemistry, **18**: 89
- Risikrishna P, 2006, *Density functional theoretical (DFT) study for the prediction of spectroscopic parameters of CICC*. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, **65**: 874
- Rizzo A, Helgaker T, Ruud K, Barszczewicz A, Jaszunski M, Jorgensen P, 1995, *Elèctric-Field Dependence of Magnetic-Properties - Multiconfigurational Self-Consistent-Field Calculations of*

Hypermagnetizabilities and Nuclear Shielding Polarizabilities of N₂, C₂H₂, HCN and H₂O. Journal of Chemical Physics, **102**: 8953

Roset L, 2007, *Actividad Biológica y Toxicidad de Dioxinas y Furanos. relación estructura Química Actividad Biológica*. A: "Revista de Toxicología". Dra. M. del Prado Mínguez Santiyán Universidad de Extremadura p. **103**.

Roset L, Rubio-Martinez J, Pérez J J, 2009; *Analysis of the basis set and correlation effects on the computation of molecular polarizabilities using molecular polarization maps*. Theoretical Chemistry Accounts, **123**: 113

Roset L, Perez J J, 2010, *Assessment of the performace of the functional density theory*. Journal of Computational Methods in Sciences and Engineering, **10**: 521

Rosipal R, Kramer N, 2006, *Overview and Recent Advances in Partial Least Squares, in Subspace, Latent Structure and Feature Selection Techniques*. C. Saunders pp. 34–51. Berlin, Germany.

Sachse T I, Kleinekathofer U, 2002, *Generalized Heitler-London theory for H₃: a comparisón of the surface integral method with perturbation theory*. European Physical Journal D, **18**: 61

Sadlej A. J, 1988, *Medium-size polarized basis sets for high-level correlated calculations of molecular electric properties*. Collection of Czechoslovak Chemical Communications, **53**: 1995

Sadlej A J, 1991, *Medium-size polarized basis sets for high-level-correlated calculations of molecular elèctric properties*. Theoretical Chimica Acta, **81**: 339

Safe S,1986, *Comparative toxicology and mechanism of action of polychlorinated dibenzo-p-dioxins and dibenzofurans*, Annual Review of Pharmacology, **26**: 371

Safe S, 1990, *Polychlorinated biphenyls (PCBs), dibenzo-p-dioxins(PCDDs), dibenzofurans (PCDFs), and related compounds: environmental and mechanistic consideration which support the development of toxic equivalency factors (TEFs)*. Critical Reviews in Toxicology, **21**: 51

Sarkar, U, 2006,, *Toxicity analysis of polychlorinated dibenzofurans through global and local electrophilicities*. Journal of Molecular Structure-Theochem, **758**: 119

Scrocco E, Tomasi J, 1973, In: Topics in Current Chemistry, **Vol 42**, Springer-Verlag, Berlin, Germany

Shaw P, 2003, *Multivariate Statistics for the Environmental Sciences*. Arnold Publishers, Malta.

Sim F, Salahub D R, Chin S, 1992, *The accurate calculation of dipole moments and dipole polarizabilities using Gaussian-based density functional methods*. International Journal of Quantum Chemistry, **43**: 463

Sjoberg P, Murray, J. S., Brinck, T., Evans, P., and Politzer, P. 1990, *The Use of the Electrostatic Potential at the Molecular-Surface in Recognition Interactions - Dibenzo-P-Dioxins and Related Systems*. Journal of Molecular Graphics, **8**: 81

Soliva, 1997, *Suitability of density functional methods for calculation of electrostatic properties*. Journal of Computational Chemistry, **18**: 980

Soscun H, 2000, *Ab initio and DFT calculations of the dipole polarizability and the second dipole hyperpolarizability of benzene*. Abstracts of Papers of the American Chemical Society, **220**: U211

- Soscun H, 2002, *Ab initio and density functional theory calculations of the dipole polarizability and the second dipole hyperpolarizability of benzene*. International Journal of Quantum Chemistry, **90**: 497.
- Spackman M. A, P. G. Byrom, 1999. *Influence of intermolecular interactions on multipole-refined electron densities*. Acta Crystallographica Section A: Foundations of Crystallography, **55**: 30.
- Surfer 6.0 Golden Software, 2002. SURFER for Windows, Version 6. , Golden Software Inc, Golden, CO. *A Powerful Contouring, Gridding, and Surface Mapping Package for Scientists and Engineers* . Colorado.
- Szabo A, 1977, *The correlation energy in the random phase approximation: Intermolecular forces between closed shell systems* .The Journal of Chemical Physics, **67**: 4351
- Szabo A, Ostlund N S, 1996 *Modern Quantum Chemistry: Introduction to advanced electronic structure theory*. Courier Dover Publications. N Y
- Tyrrell J, 1979, *Electronic structure and one-electron properties of the isoelectronic molecules HCN, HNC, HBO, HOB, HBF+, and HFB+*. Journal of Physical Chemistry, **83**: 2906.
- Tuppurainen, K, Ruuskanen, J, 2000. *Electronic eigenvalue (EEVA): a new QSAR/QSPR descriptor for electronic substituent effects based on molecular orbital energies. A QSAR approach to the Ah receptor binding affinity of polychlorinated biphenyls (PCBs), dibenzo-p-dioxins (PCDDs) and dibenzofurans (PCDFs)*. Chemosphere, **41**: 843
- Umetrics AB, 2002, *SIMCA-P version 10.0*. Umetrics AB, Umea, Sweden.
- Vandenberg M, 1994, *The Toxicokinetics and Metabolism of Polychlorinated Dibenzo-P-Dioxins (Pcdds) and Dibenzofurans (Pcdfs) and Their Relevance for Toxicity*. Critical Reviews in Toxicology, **24**: 1
- Wagener M, Sadowski J, 1995, *Autocorrelation of molecular surface properties for modeling corticosteroid binding globulin and cytosolic Ah receptor activity by neural networks*. Journal of American Chemical Society, **117**: 7769
- Waller, C.L, McKinney, J.D, 1992. *Comparative molecular field analysis of polyhalogenated dibenzo-p-dioxins, dibenzofurans, and biphenyls*. Journal of Medicinal Chemistry, **35**: 3660
- Waller, C.L, McKinney, J.D, 1995. *Three-dimensional quantitative-activity relationships of dioxins and dioxin-like compounds: model validation and Ah receptor characterization*. Chemical Research in Toxicology, **8**: 847
- Weast R C, 1980, Handbook of Chemistry and Physics, 61 ed.; CRC Press: Boca Raton, FL
- Weerasingue, 2003, *Cavity formation and preferential interactions in urea solutions: Dependence on urea aggregation*. Journal of Chemical Physics, **118**: 5901
- Weinstein H, 1975, *Some new quantum chemical procedures for the analysis of drug-receptor interactions*. International Journal of Quantum Chemistry, **9**: 59
- Whitten A E, 2006, *Dipole and quadrupole moments of molecules in crystals: A novel approach based on integration over Hirshfeld surfaces*. Journal of Chemical Physics, **124**: 74106
- Wold, H., 1966. *Estimation of principal components and related models by iterative least squares*. In: Krishnaiah, P.R, Multivariate Analysis, Academic Press, New York, pp. 391-420.

Wold, H., 1973. *Non-linear iterative partial least squares (NIPALS) modelling. Some current developments.* In: Krishnaiah, P.R. (Ed.), *Multivariate Analysis*, Vol. III. Academic Press, New York, pp. 383–407.

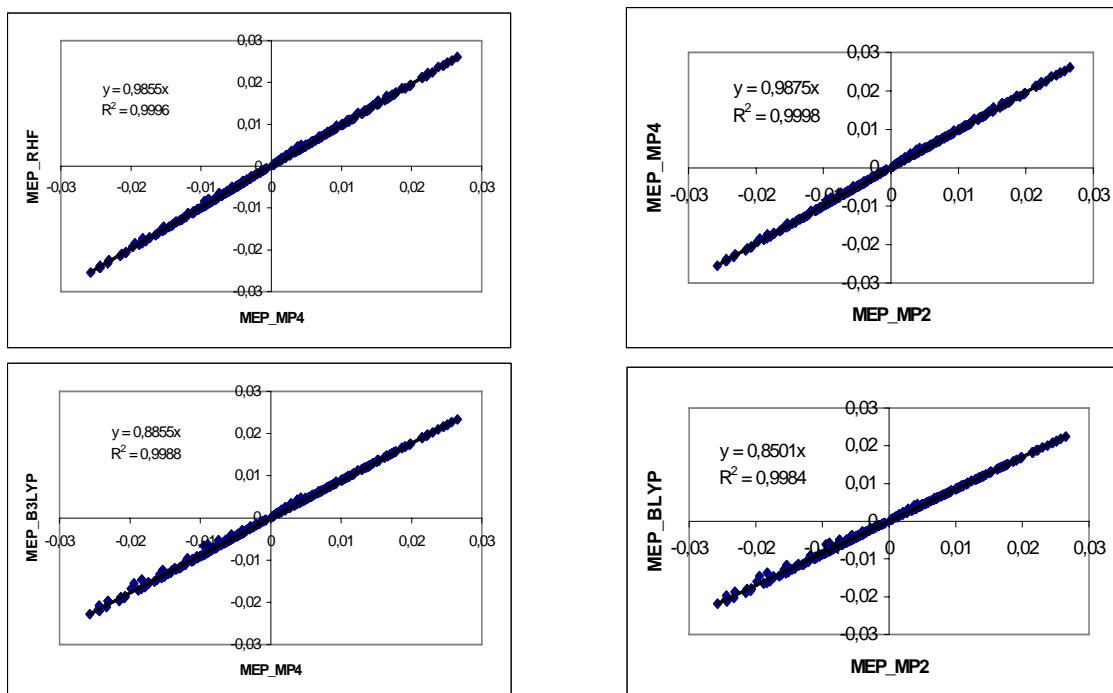
Wold, S, 1978. *Cross-validation estimation of the number of components infactor and principal components analysis.* *Technometrics*, **24**: 397

Wold, S, Martens, H., Wold, H, 1983, *The multivariate calibration problem in chemistry solved by the PLS method.* *Lecture Notes in Mathematics*, **973**: 286

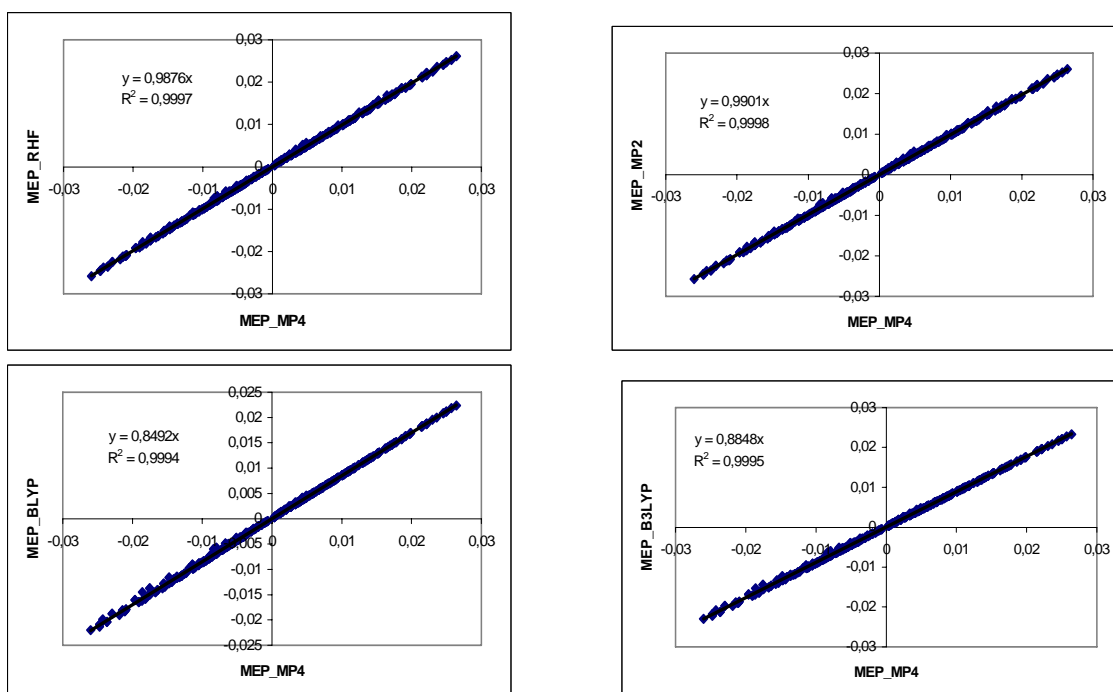
Yang G, Yu J 2007, *QSPR Study on the Aqueous Solubility ($-\lg S_w$) and n-Octanol/Water Partition Coefficients ($\lg K_{ow}$) of Polychlorinated Dibenzo-p-dioxins (PCDDs).* *QSAR & Combinatorial Science*, **26**: 352

8 ANNEXOS

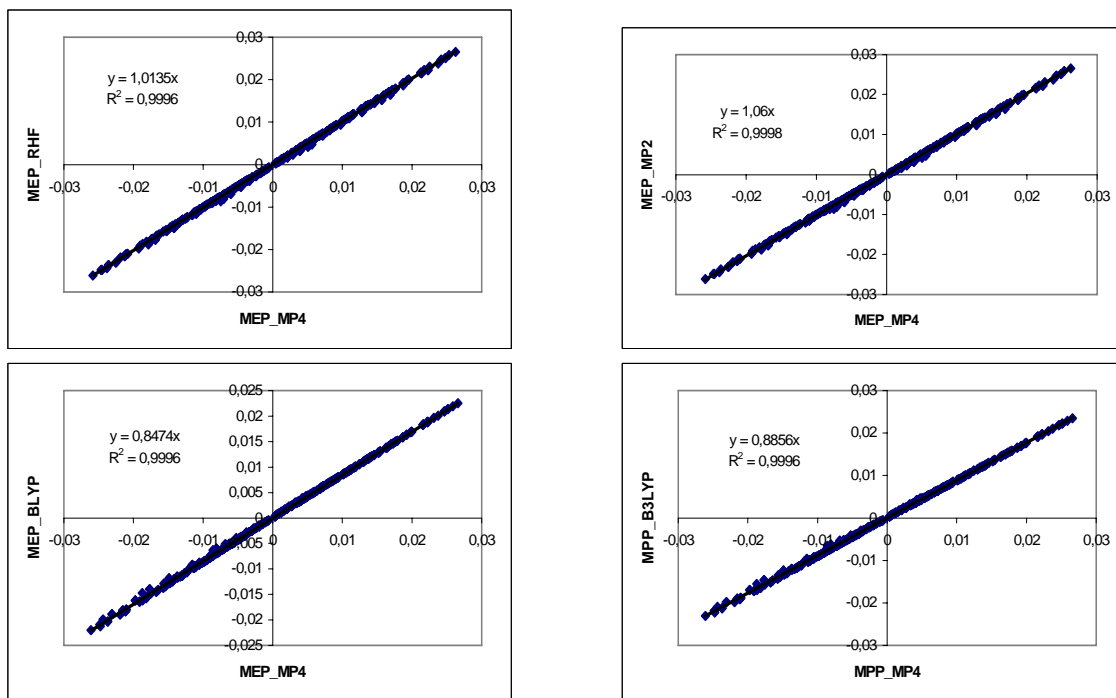
8.1 Annex 1. Anàlisi de correlació. Càlculs MEP i MPP



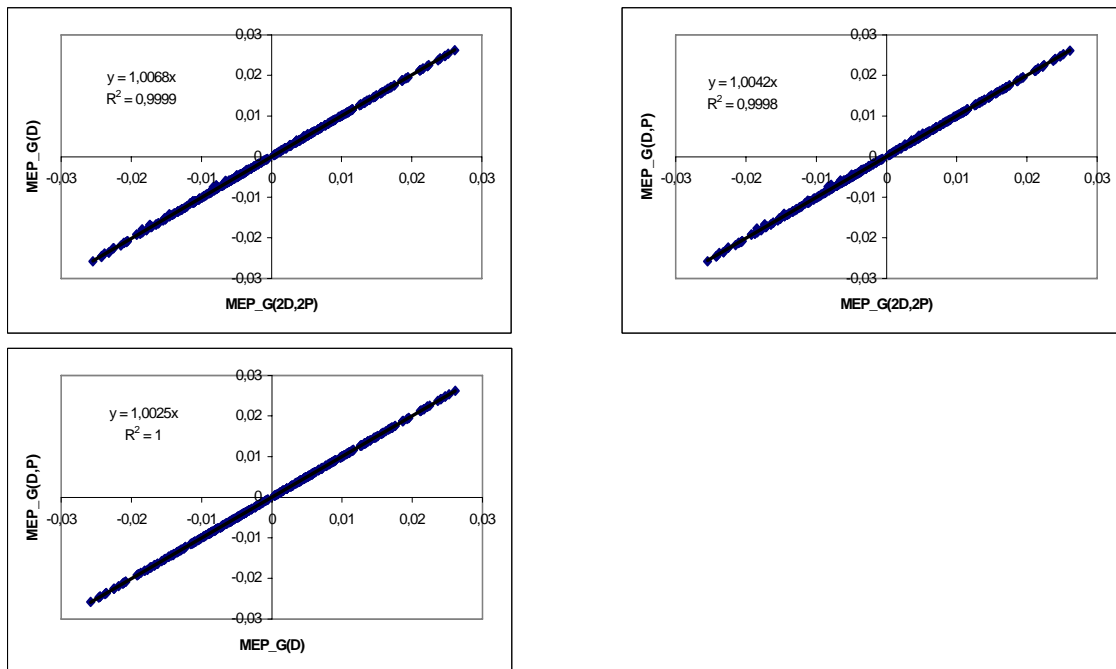
Gràfica 8.1.1 Molècula HCN. Gaussian. Anàlisi correlació. MEP. Base 6-31G(2d,2p).



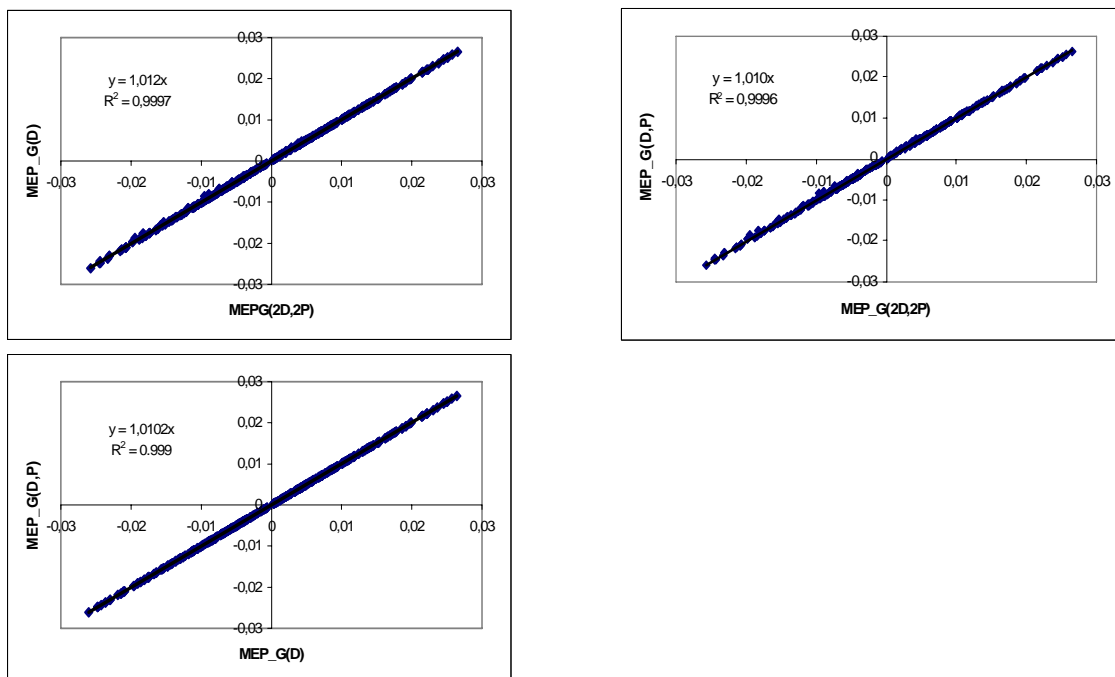
Gràfica 8.1.2 Molècula HCN. Gaussian. Anàlisi correlació. MEP. Base 6-31G(d,p)



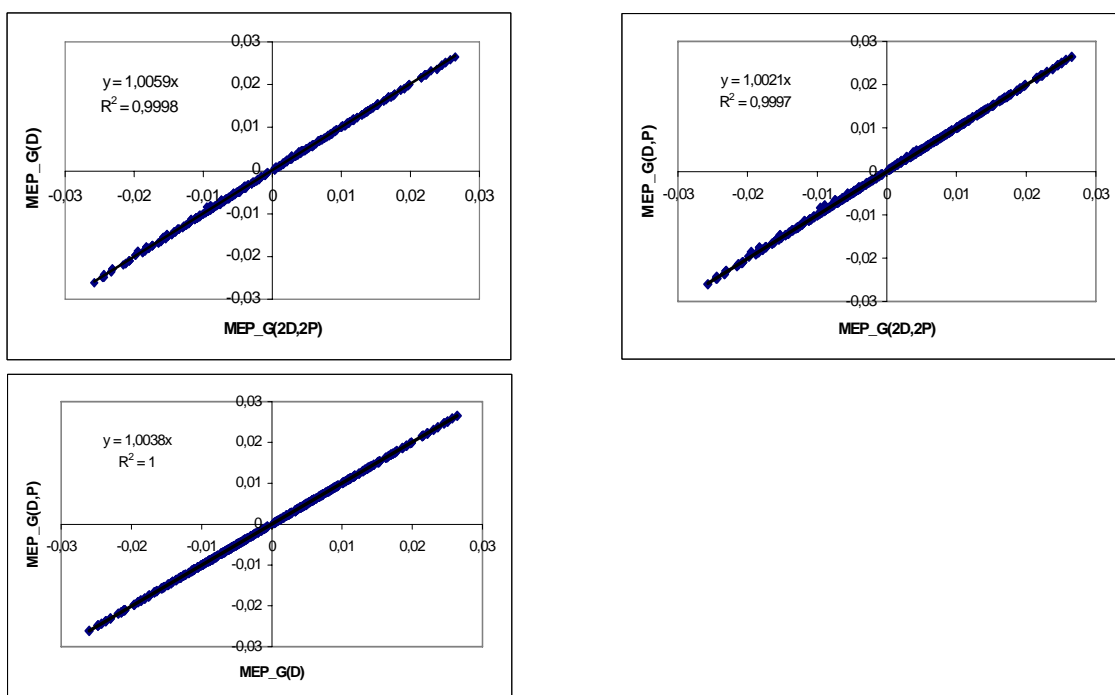
Gràfica.8.1.3 Molècula HCN. Gaussian. Anàlisi correlació. MEP. Base 6-31G(d).



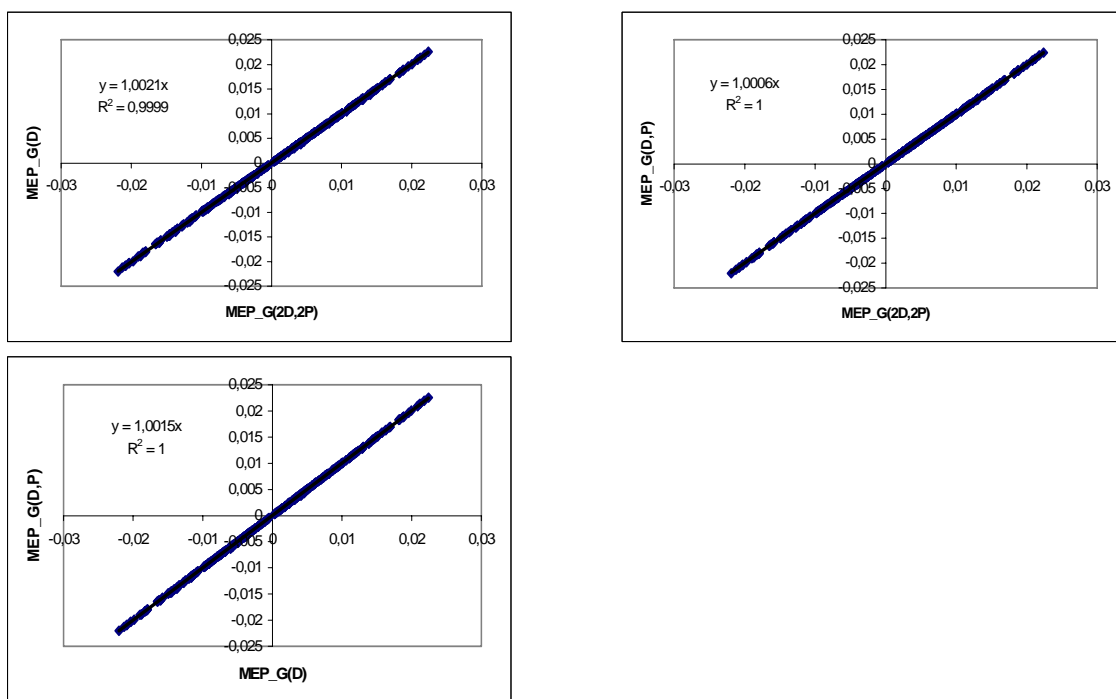
Gràfica 8.1.4 Molècula HCN. Gaussian. Anàlisi correlació. MEP. Càlculs RHF.



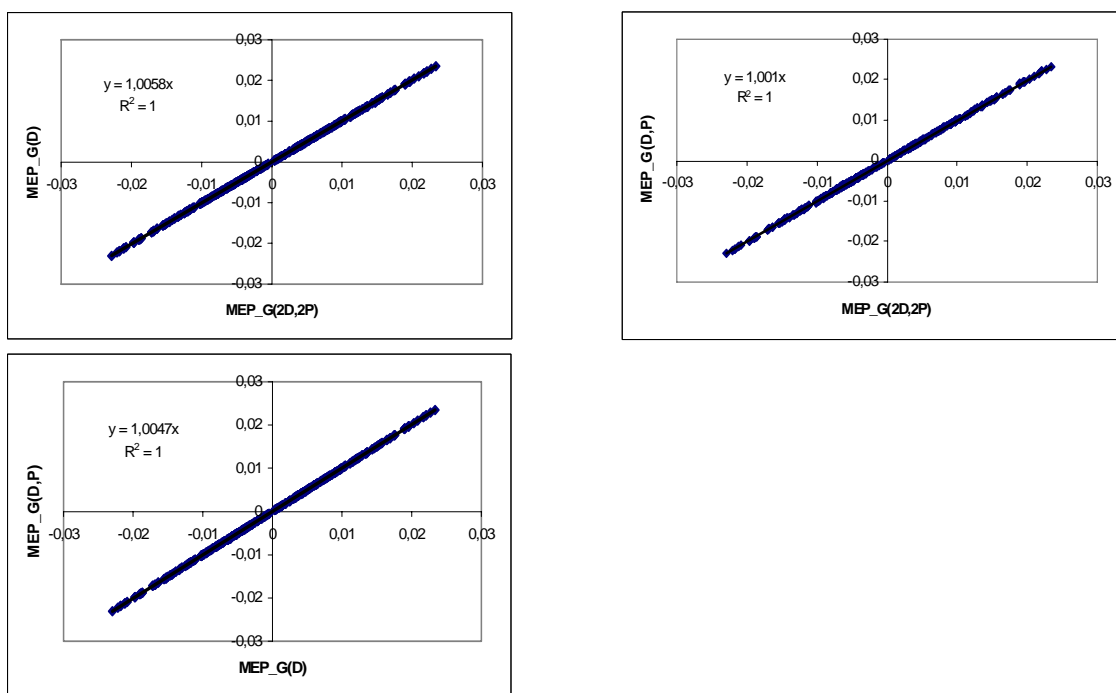
Gràfica 8.1.5 Molècula HCN. Gaussian. Anàlisi correlació. MEP. Càlculs MP2.



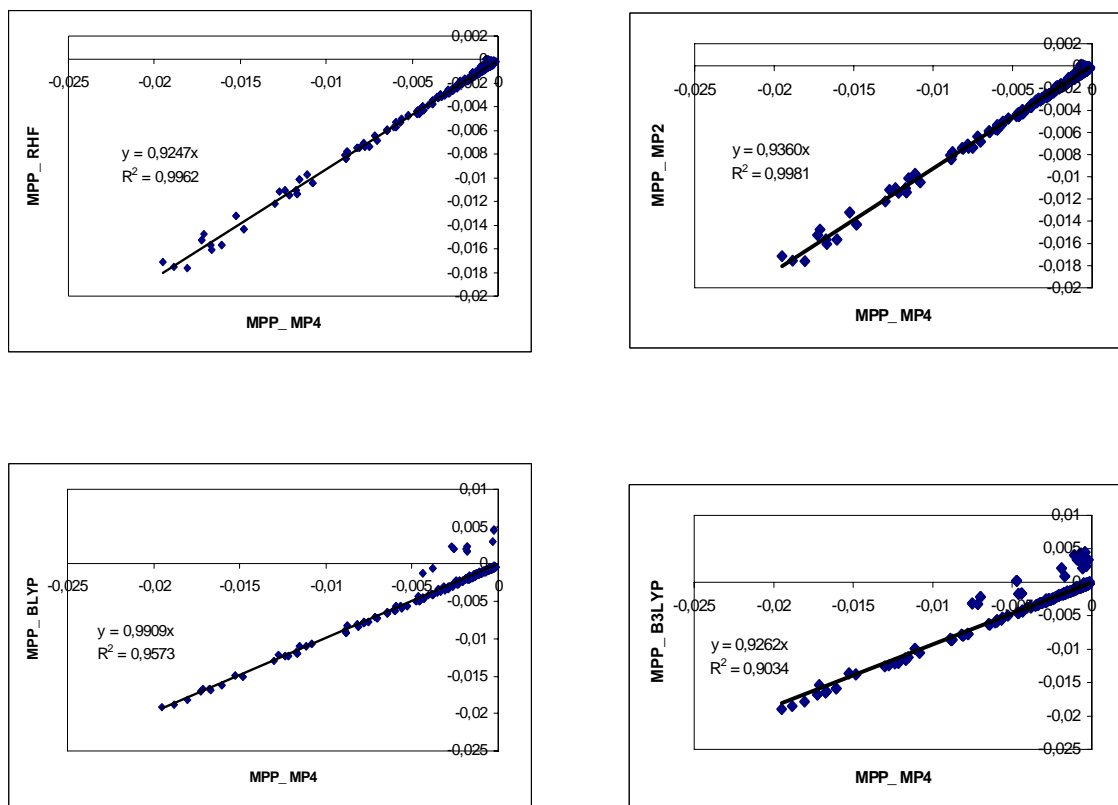
Gràfica 8.1.6 Molècula HCN. Gaussian. Anàlisi correlació. MEP. Càlculs MP4.



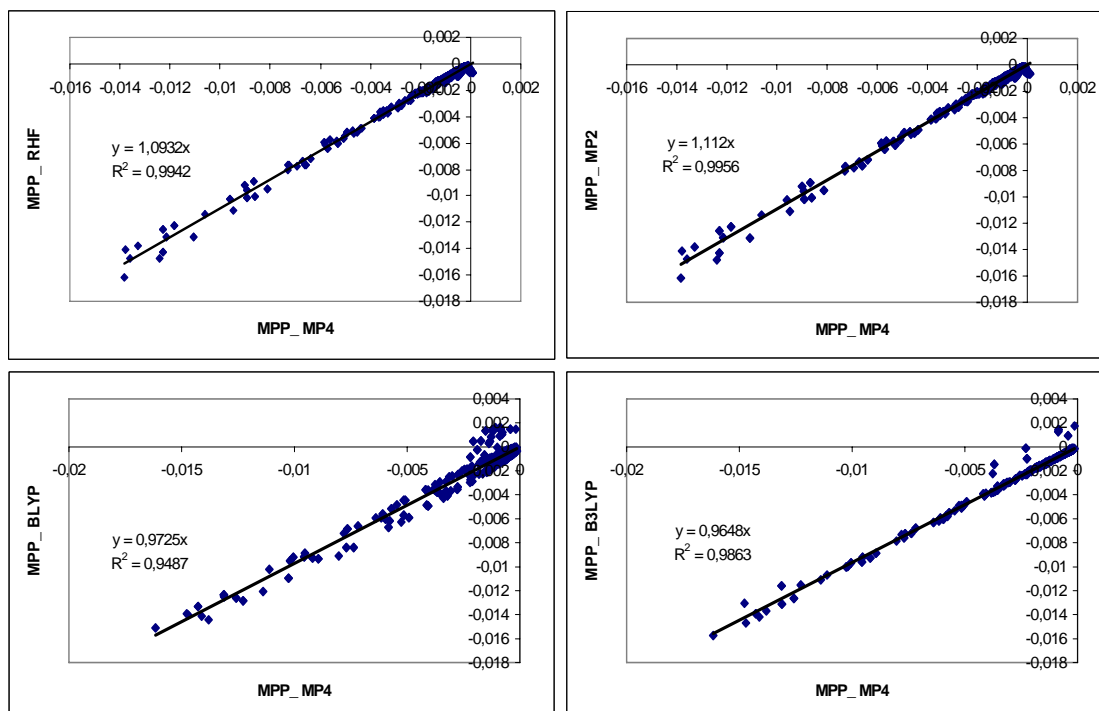
Gràfica 8.1.7. Molècula HCN. Gaussian. Anàlisi correlació. MEP. Càlculs BLYP.



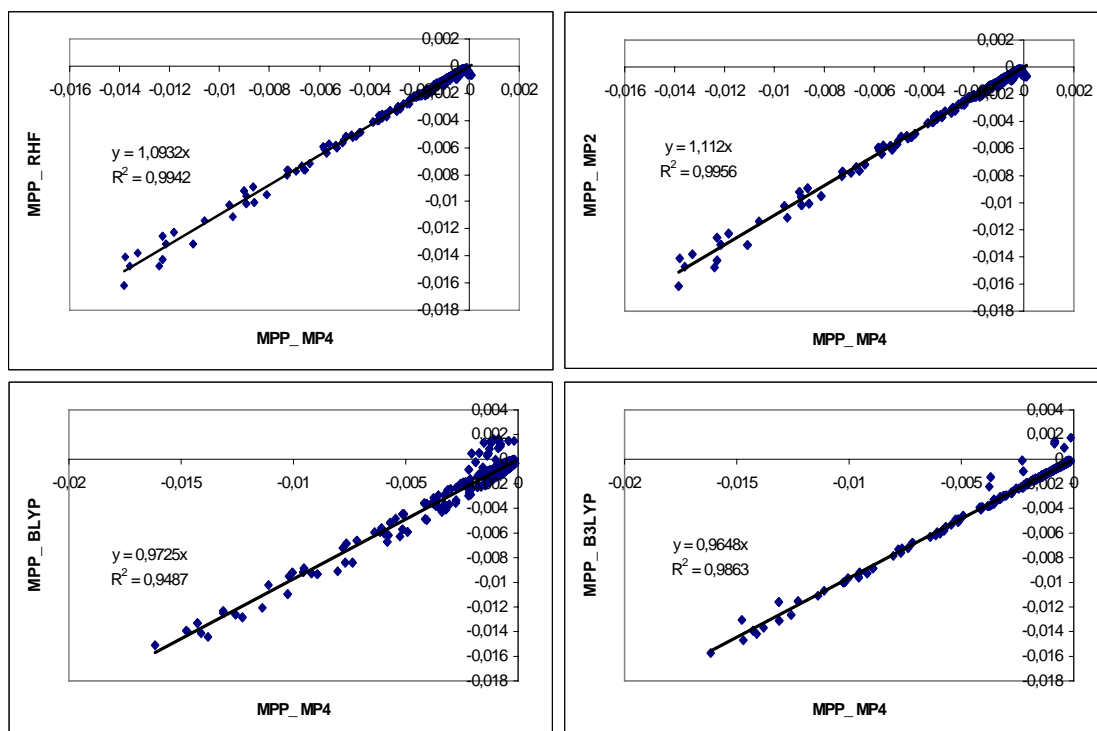
Gràfica 8.1.8 Molècula HCN. Gaussian. Anàlisi correlació. MEP. Càlculs B3LYP



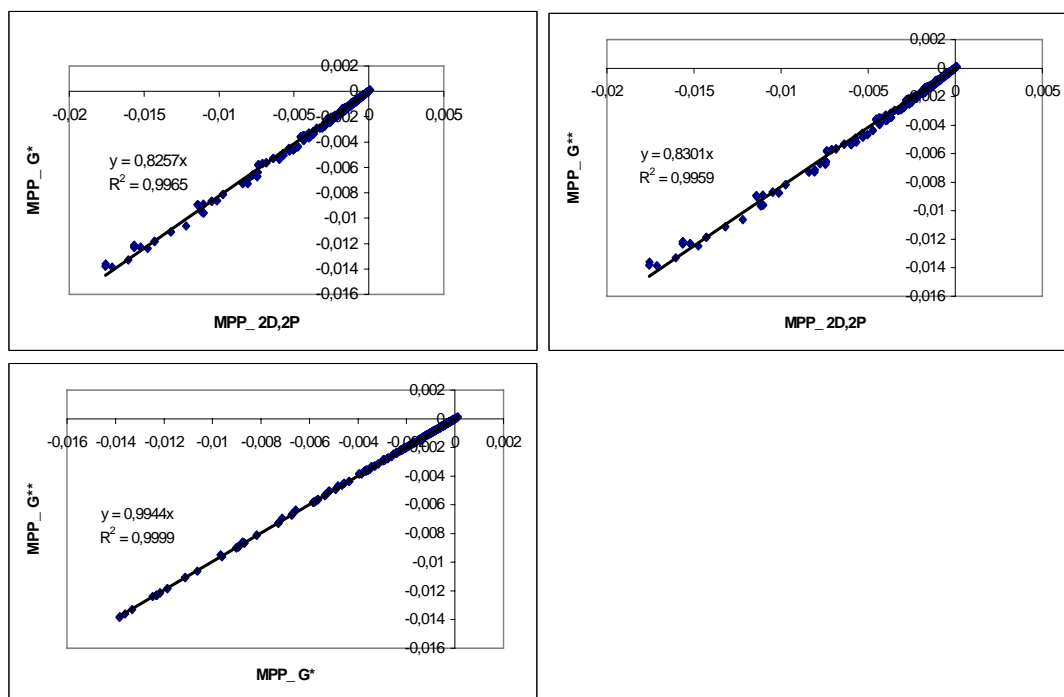
Gràfica 8.1.9 Molècula HCN. MPP. Anàlisi de correlació. Base 6-31G(2d,2p)



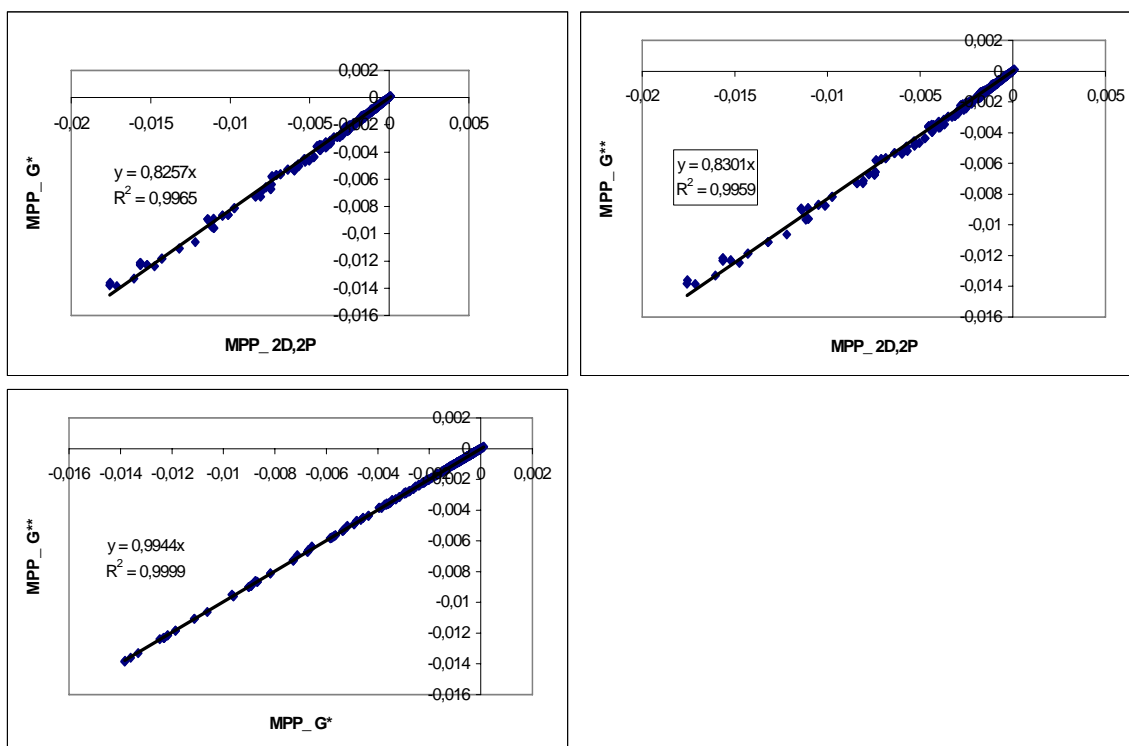
Gràfica.8.1.10 Molècula HCN. MPP. Anàlisi de correlació. Base 6-31G(d,p)



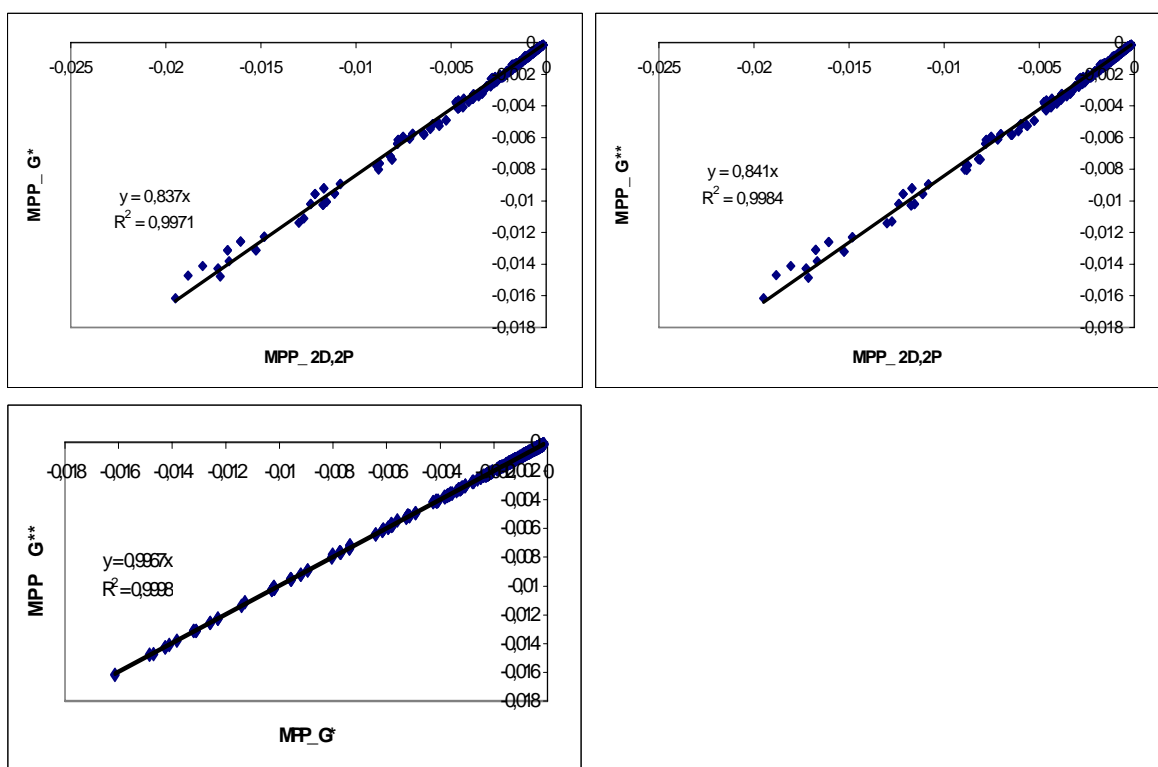
Gràfica 8.1.11 Molècula HCN. MPP. Anàlisi de correlació. Base 6-31G(d).



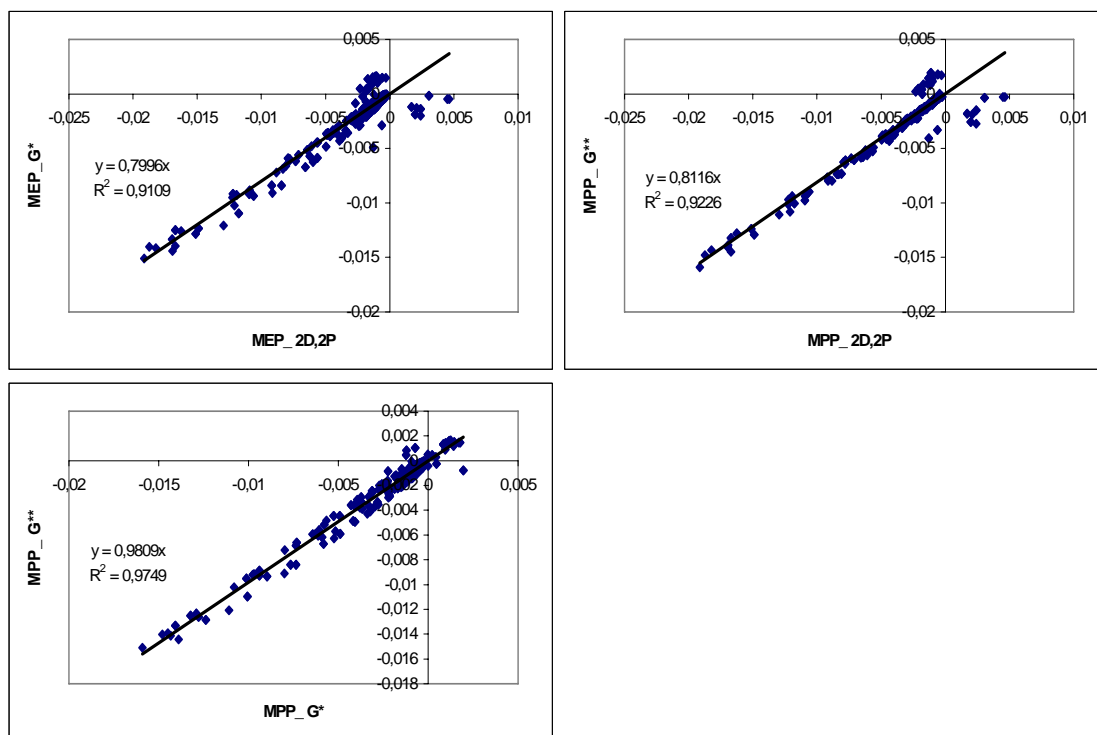
Gràfica 8.1.12 Molècula HCN. MPP. Anàlisi de correlació. Càlculs RHF.



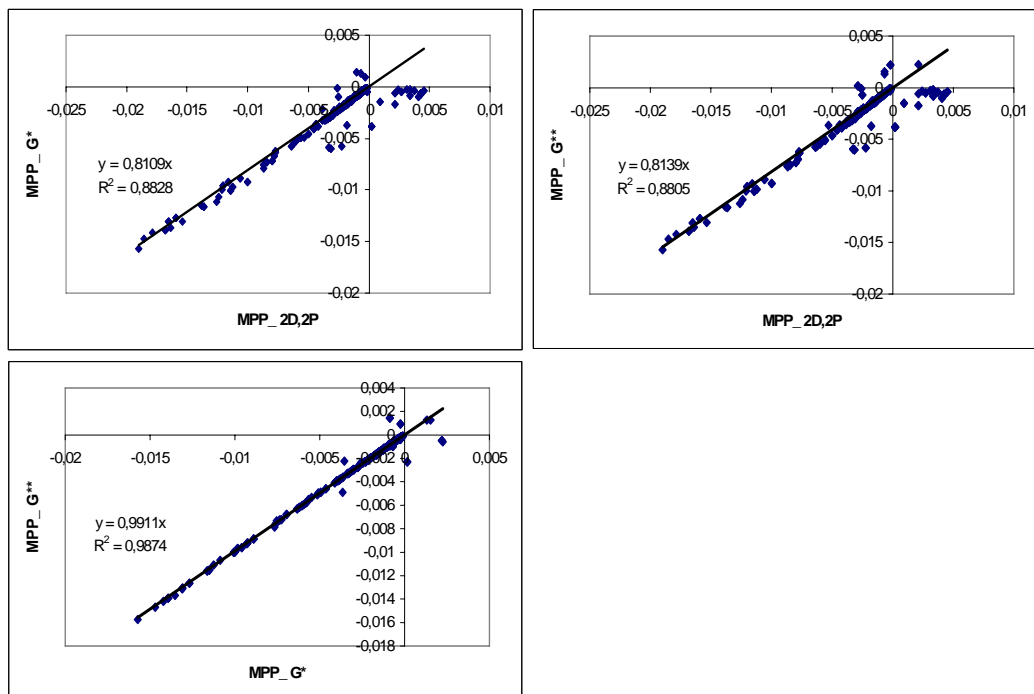
Gràfica.8.1.13 Molècula HCN. MPP. Anàlisi de correlació. Càlculs RHF.



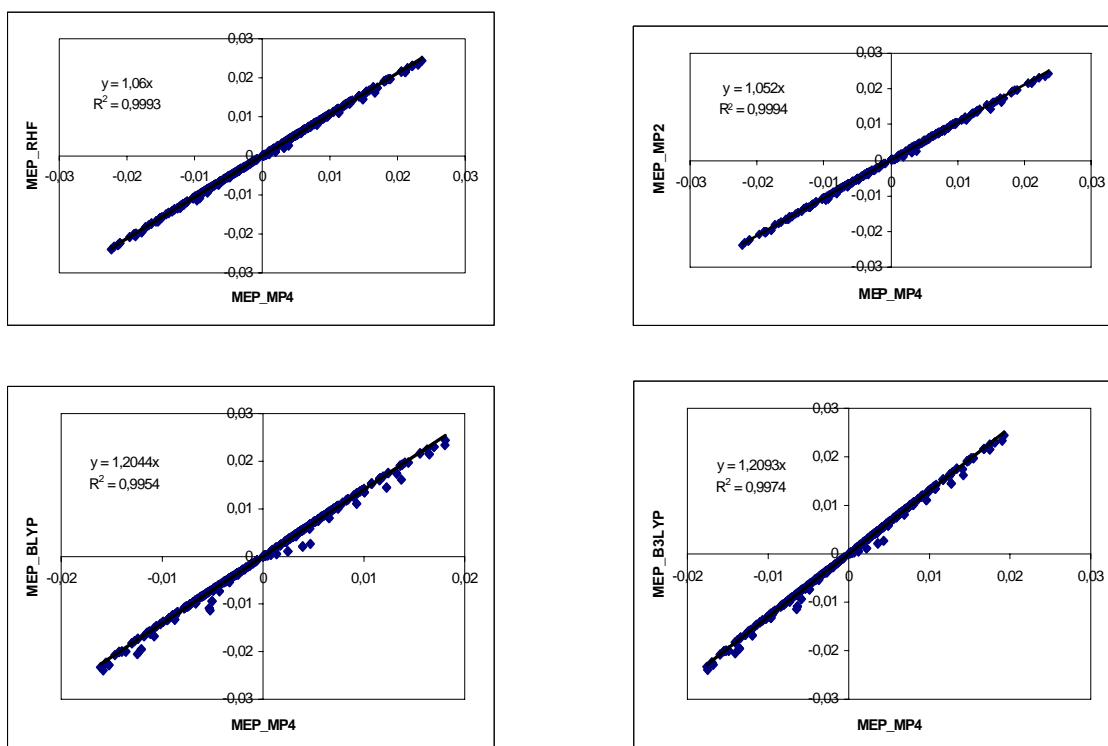
Gràfica 8.1.14 Molècula HCN. MPP. Anàlisi de correlació. Càlculs MP4.



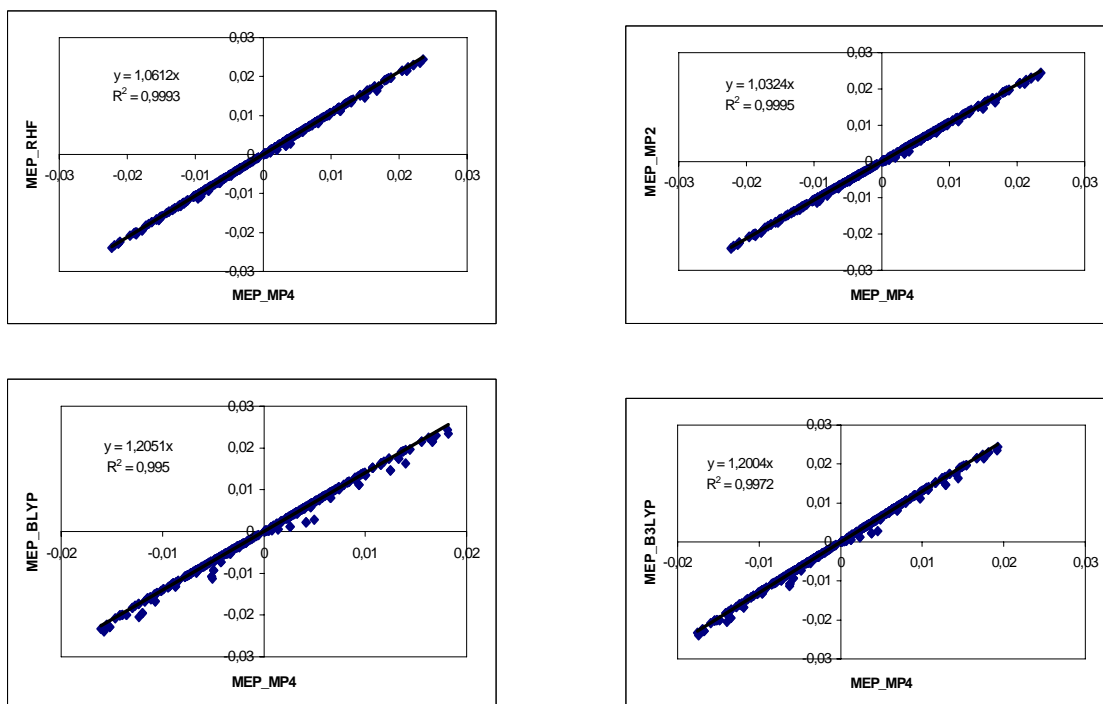
Gràfica 8.1.15 Molècula HCN. MPP. Anàlisi de correlació. Càlculs BLYP.



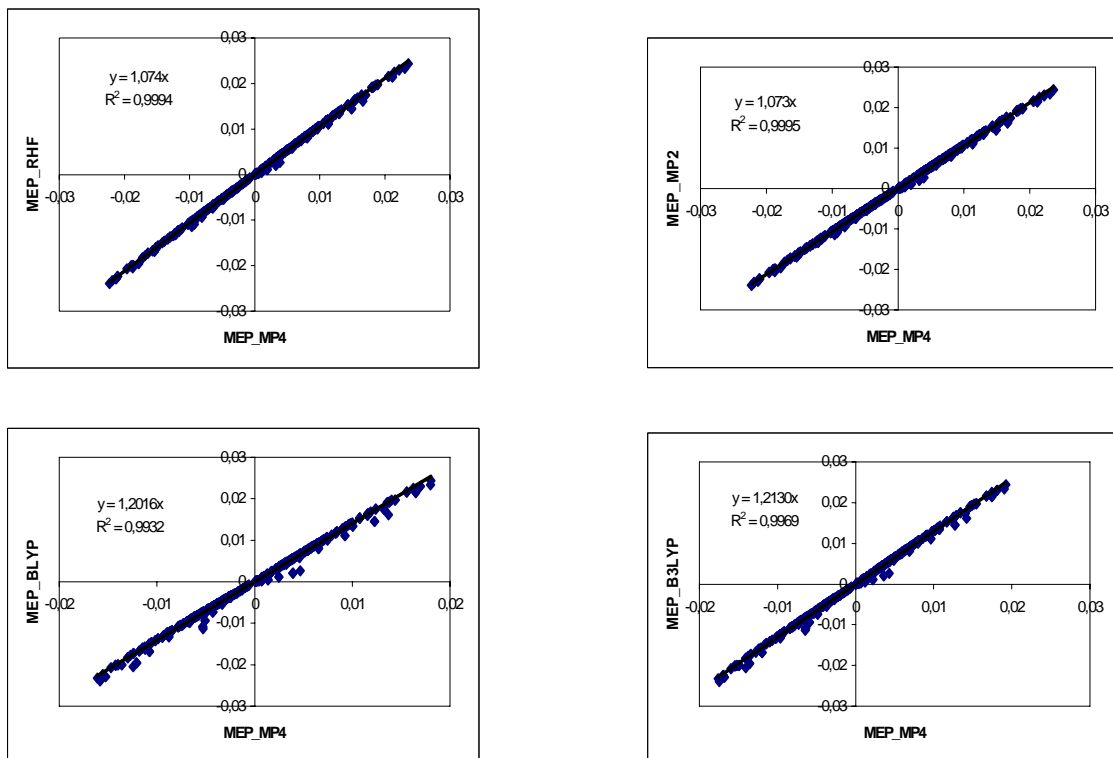
Gràfica.8.1.16 Molècula HCN. MPP. Anàlisi de correlació. Càlculs B3LYP.



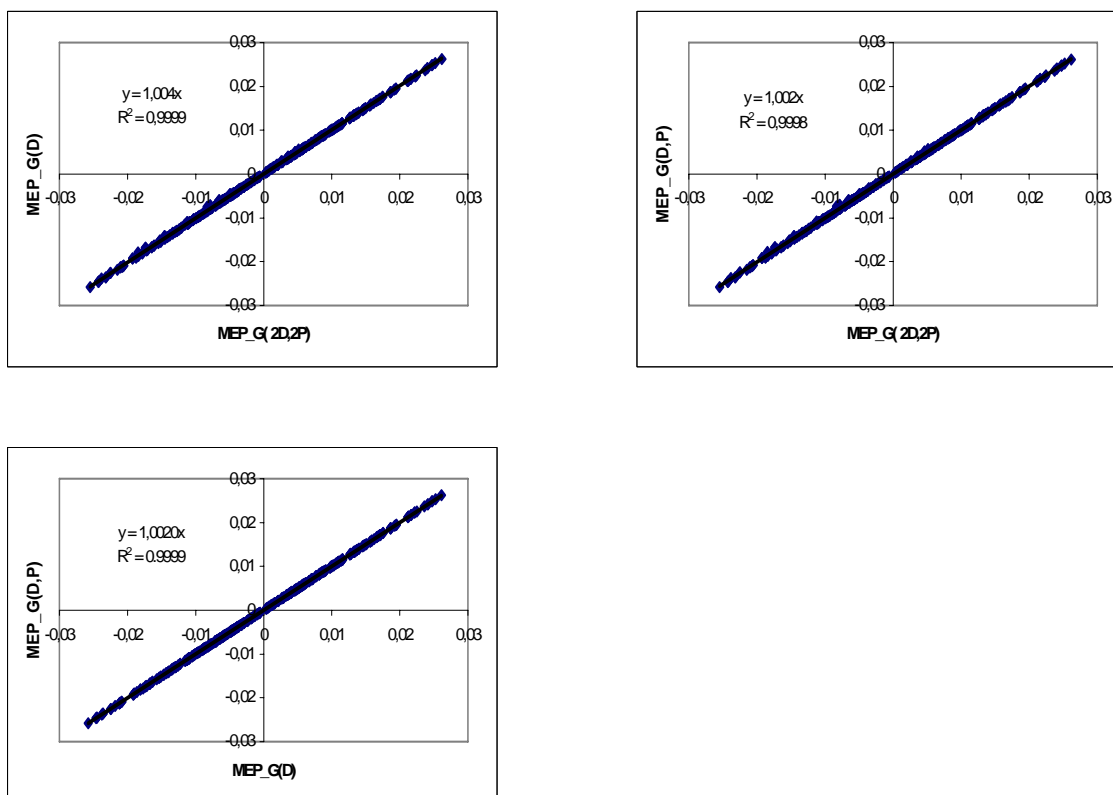
Gràfica 8.1.17 Molècula formaldehid. Gaussian. Anàlisi de correlació. MEP. Base 6-31G(d).



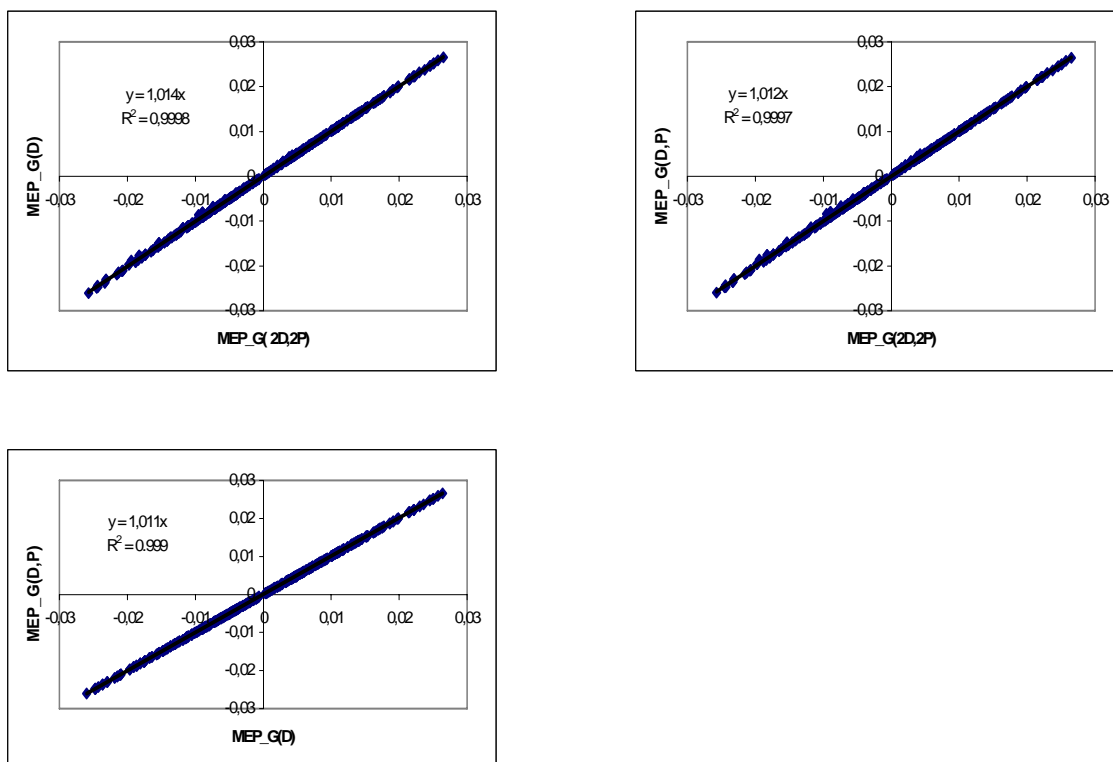
Gràfica 8.1.18 Molècula formaldehid. Gaussian. Anàlisi de correlació. MEP. Base 6-31G(d,p).



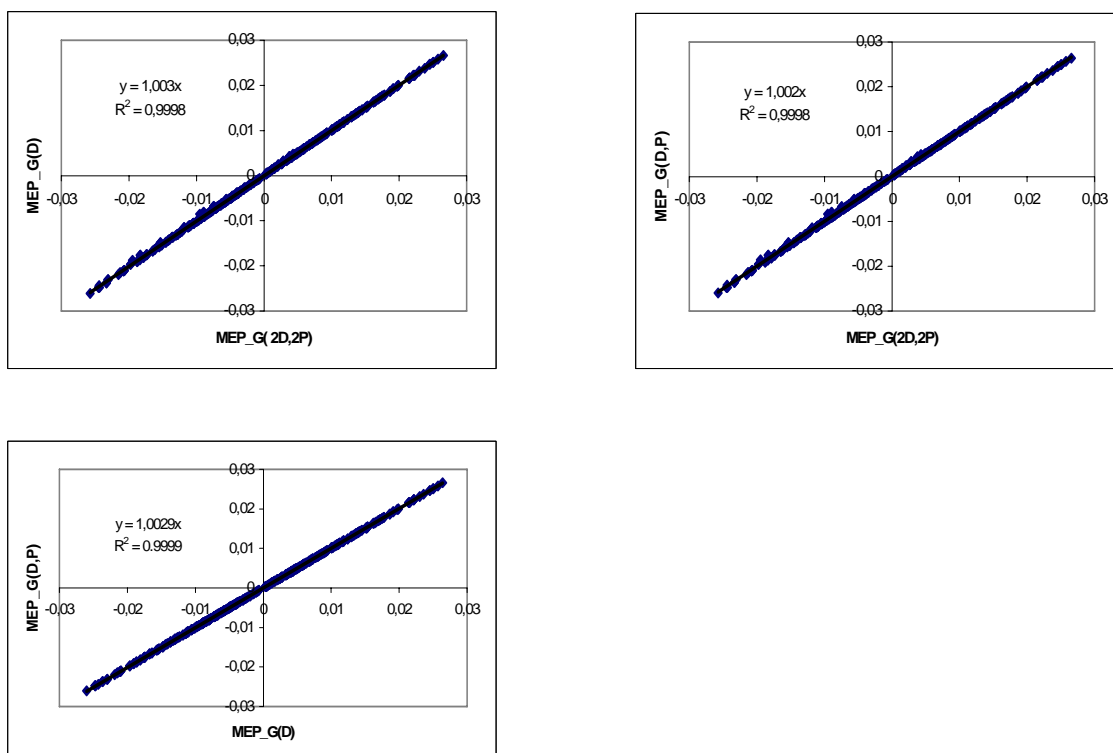
Gràfica 8.1.19 Molècula formaldehid. Gaussian. Anàlisi correlació. MEP. Base 6-31G(2d,2p).



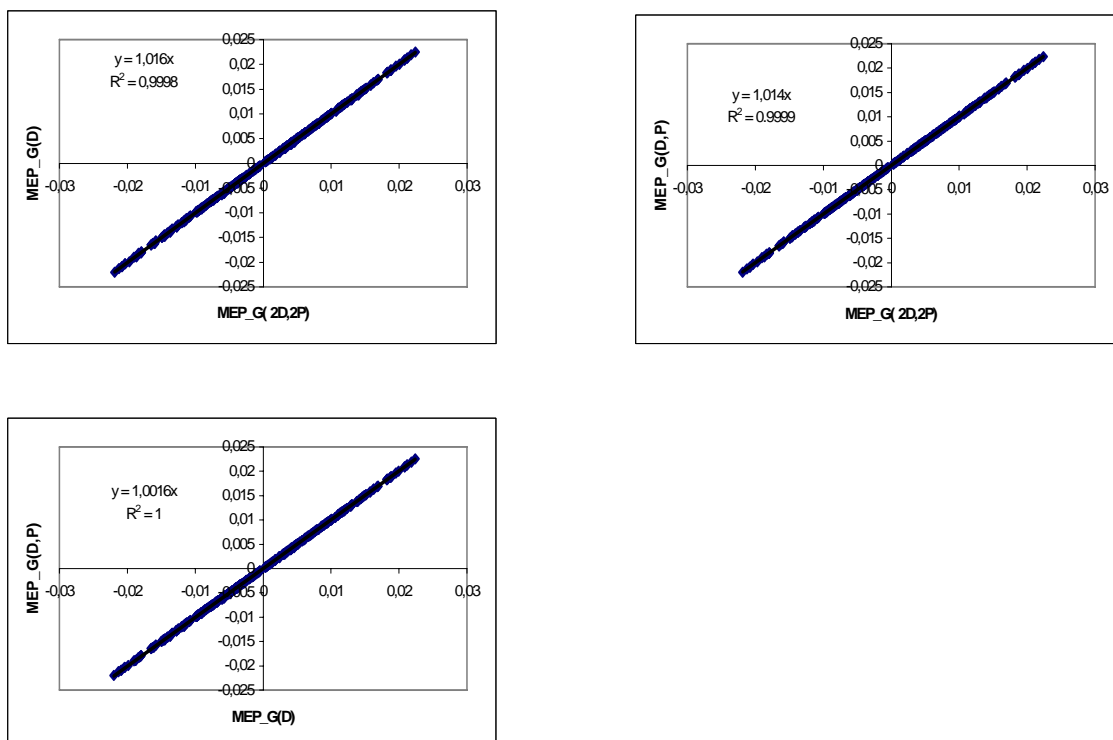
Gràfica 8.1.20 Molècula formaldehid. Gaussian. Anàlisi correlació. MEP. Càlcul RHF.



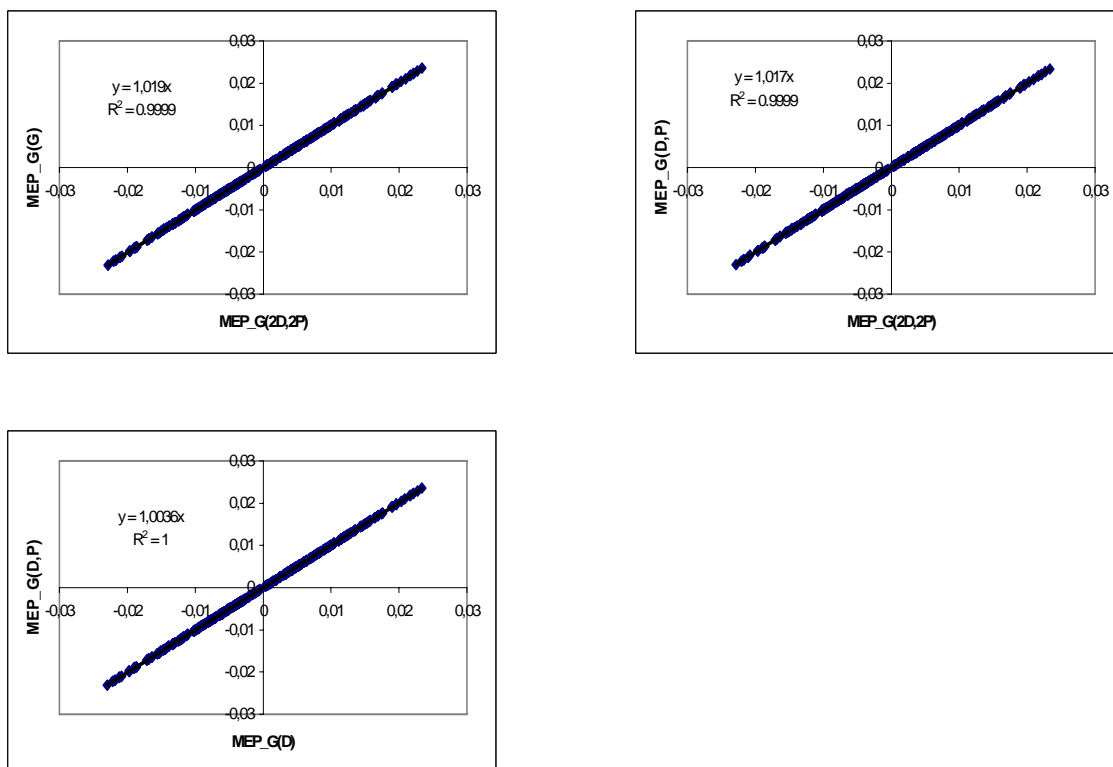
Gràfica 8.1.21 Molècula formaldehid. Gausssian. Anàlisi correlació. MEP. MP2.



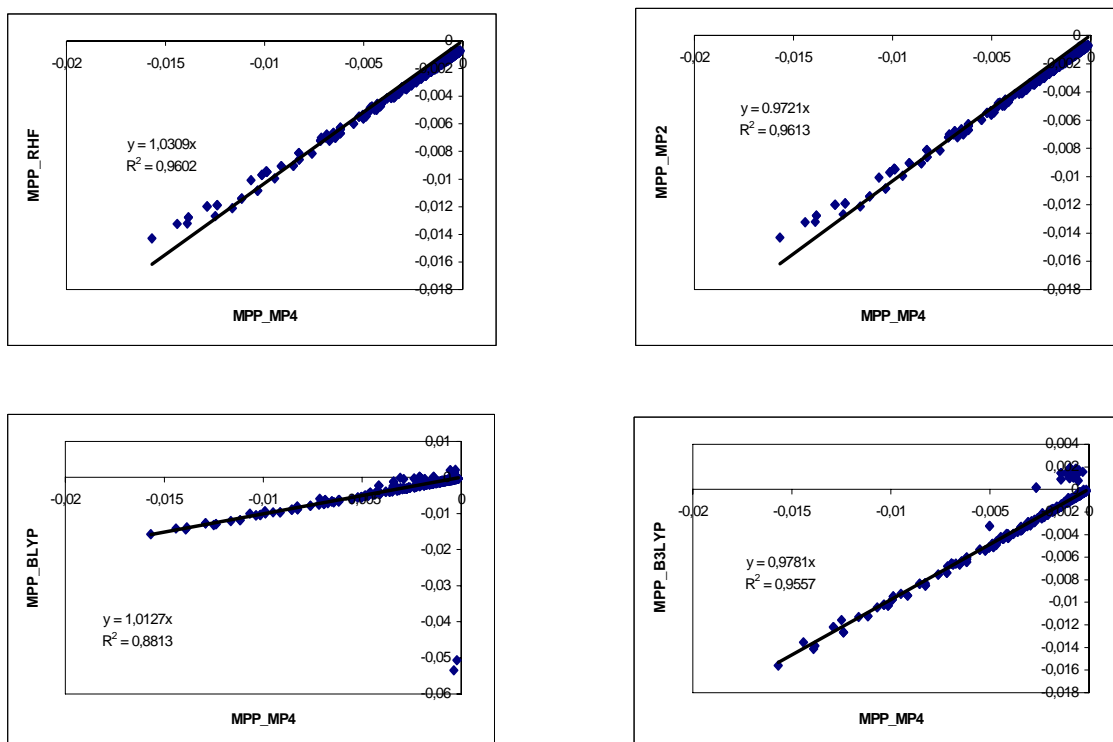
Gràfica 8.1.22 Molècula formaldehid. Gausssian. Anàlisi correlació. MEP. MP4.



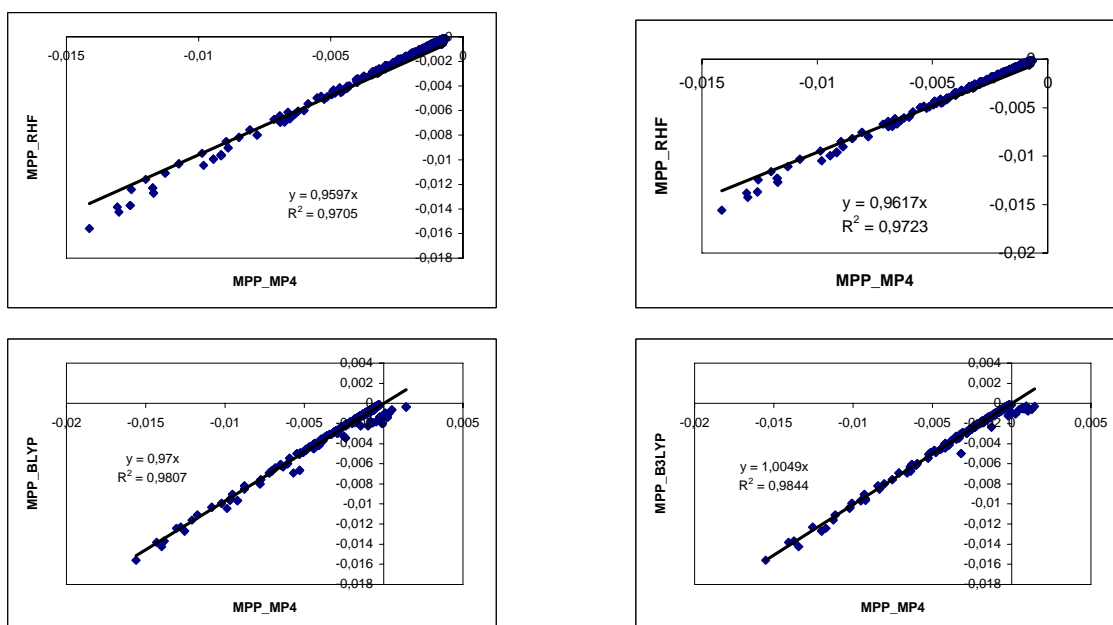
Gràfica 8.1.23 Molècula formaldehid. Gaussian. Anàlisi correlació. MEP. BLYP



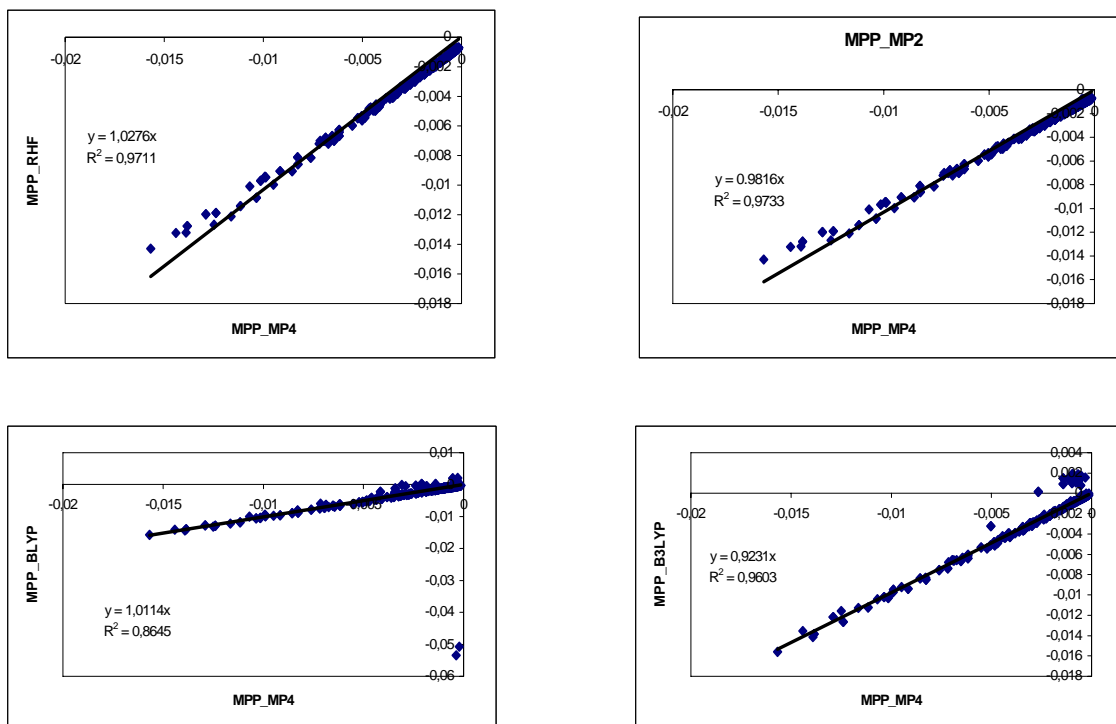
Gràfica 8.1.24 Molècula formaldehid. Gaussian. Anàlisi correlació. MEP. B3LYP.



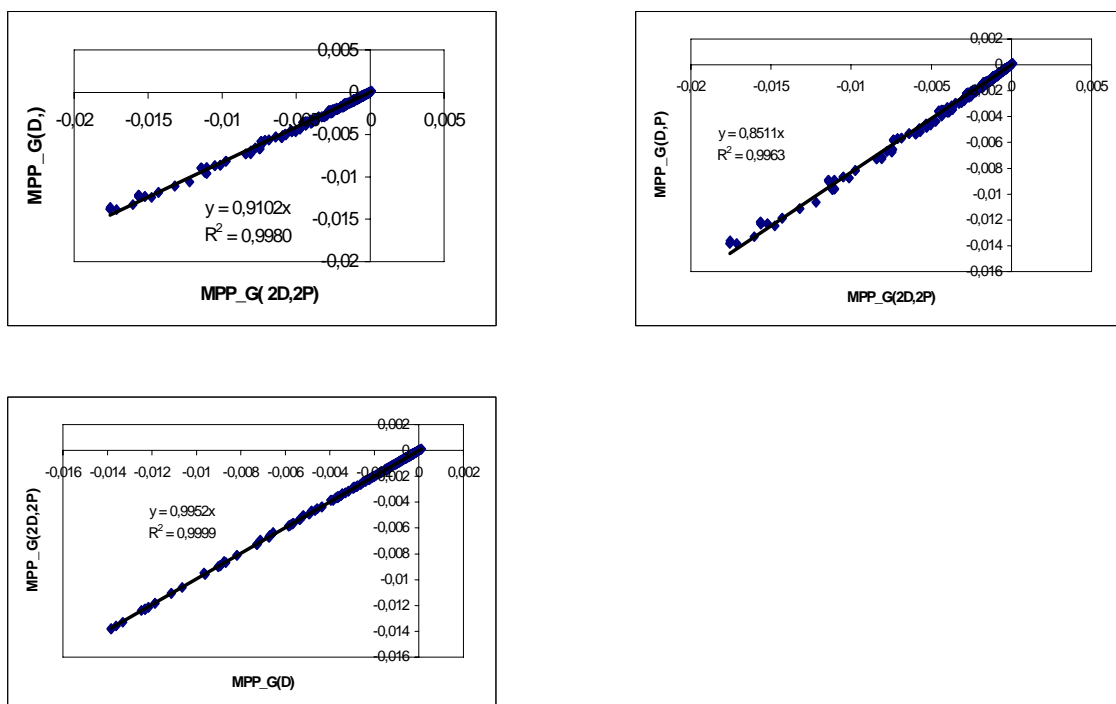
Gràfica 8.1.25 Molècula formaldehid. Gaussian. Anàlisi correlació. MPP. Base 6-31G(d,p)



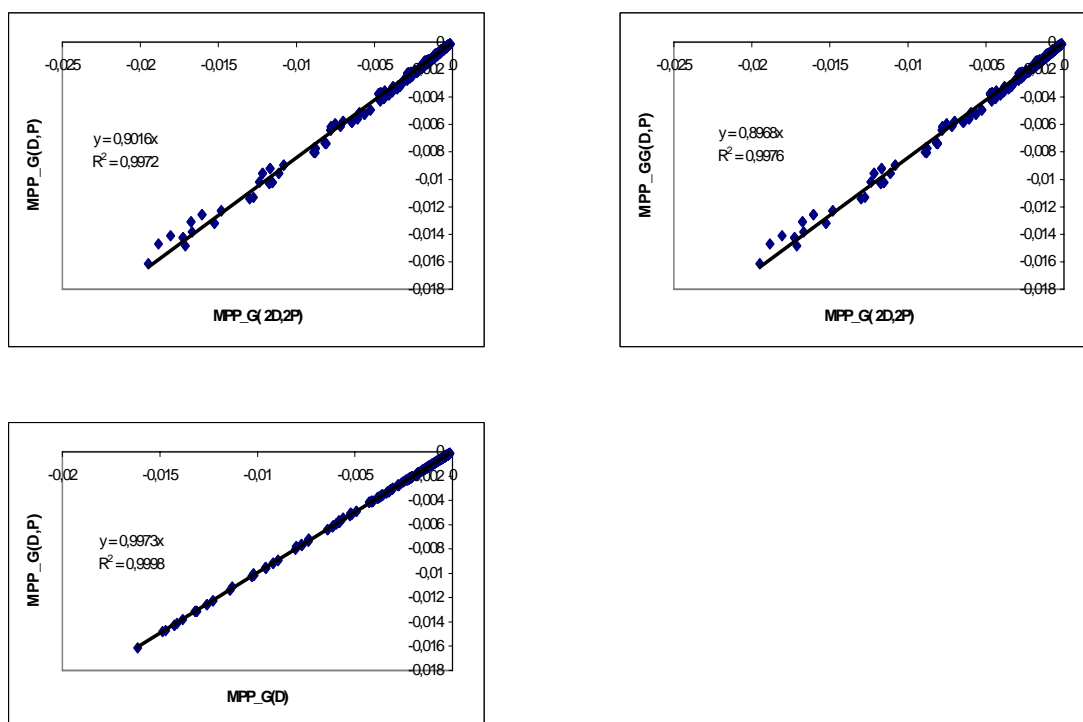
Gràfica 8.1.26 Molècula formaldehid. Gaussian. Anàlisi correlació. MPP. Base 6-31G(d).



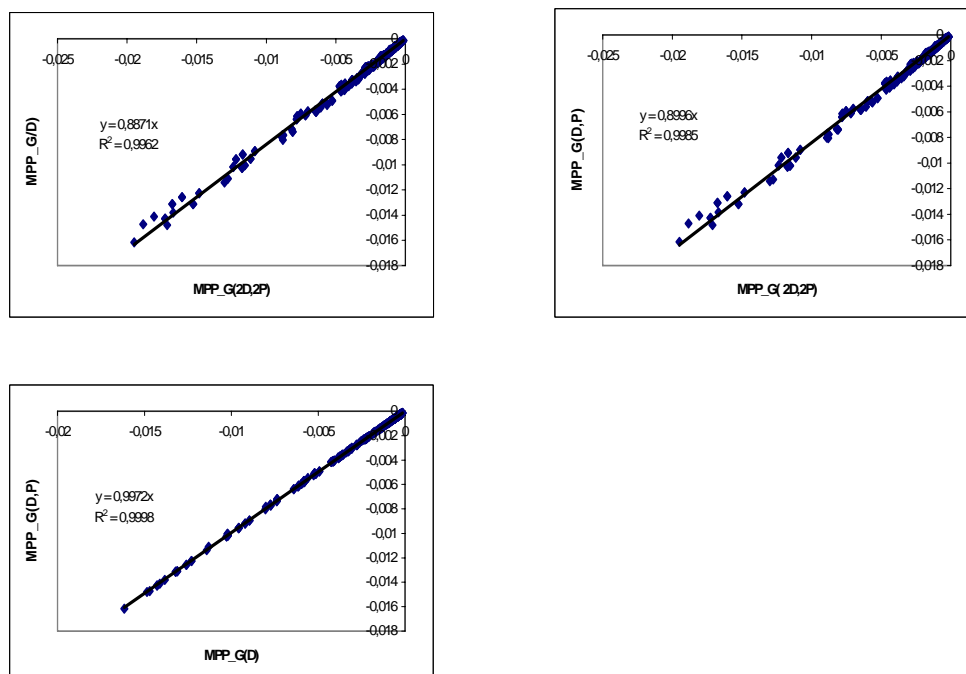
Gràfica 8.1.27 Molècula formaldehid . Gaussian. Anàlisi correlació. MPP. Base 6-31G(2d,2p).



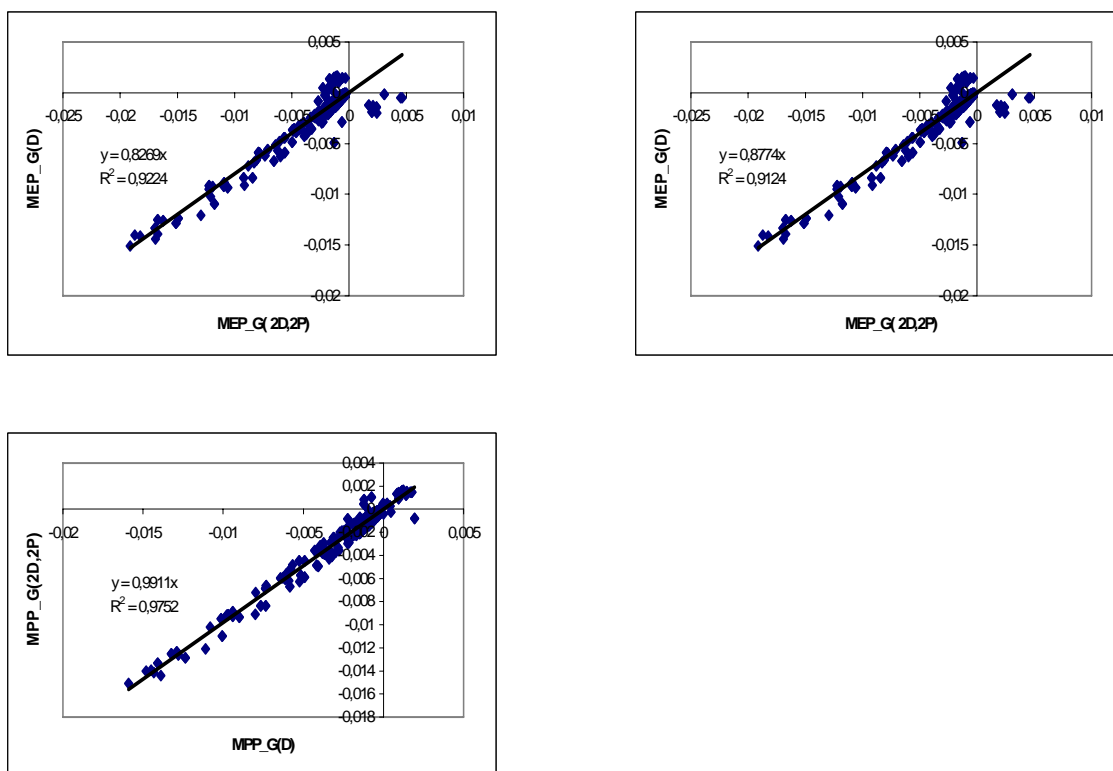
Gràfica 8.1.28 Molècula formaldehid. Gaussian. Anàlisi correlació. MPP. Càlculs RHF.



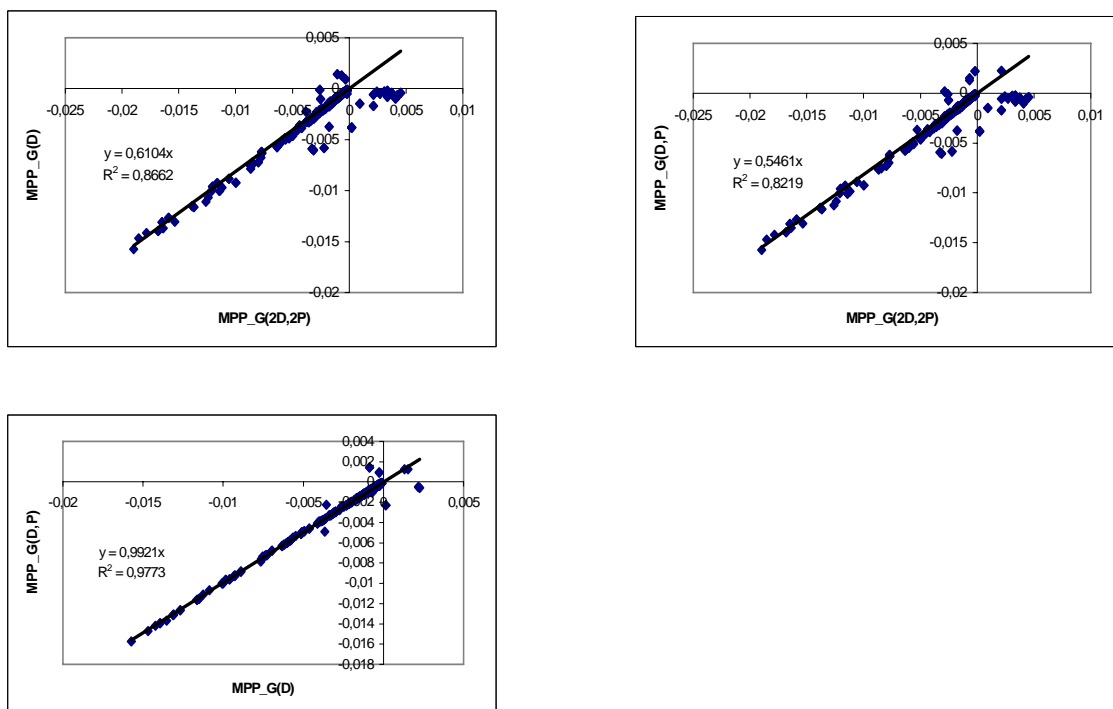
Gràfica 8.1.29 Molècula formaldehid . Gausisian. Anàlisi correlació. MPP.Càlculs MP2.



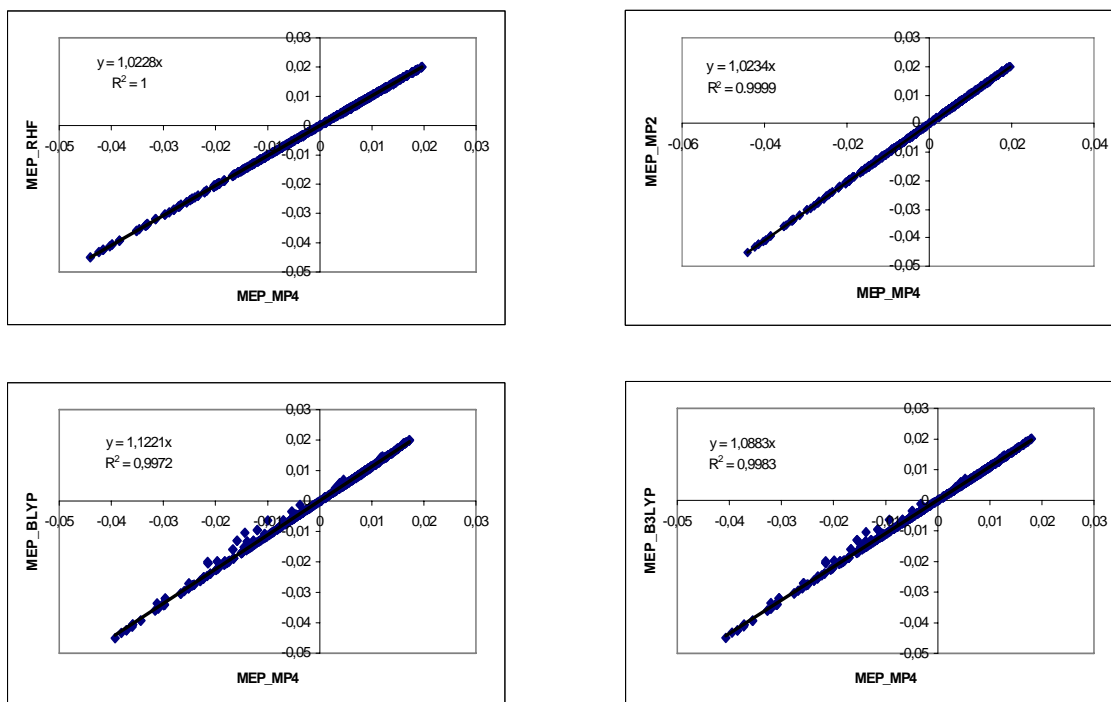
Gràfica 8.1.30 Molècula formaldehid. Gausisian. Anàlisi correlació. MPP. Càlculs MP4.



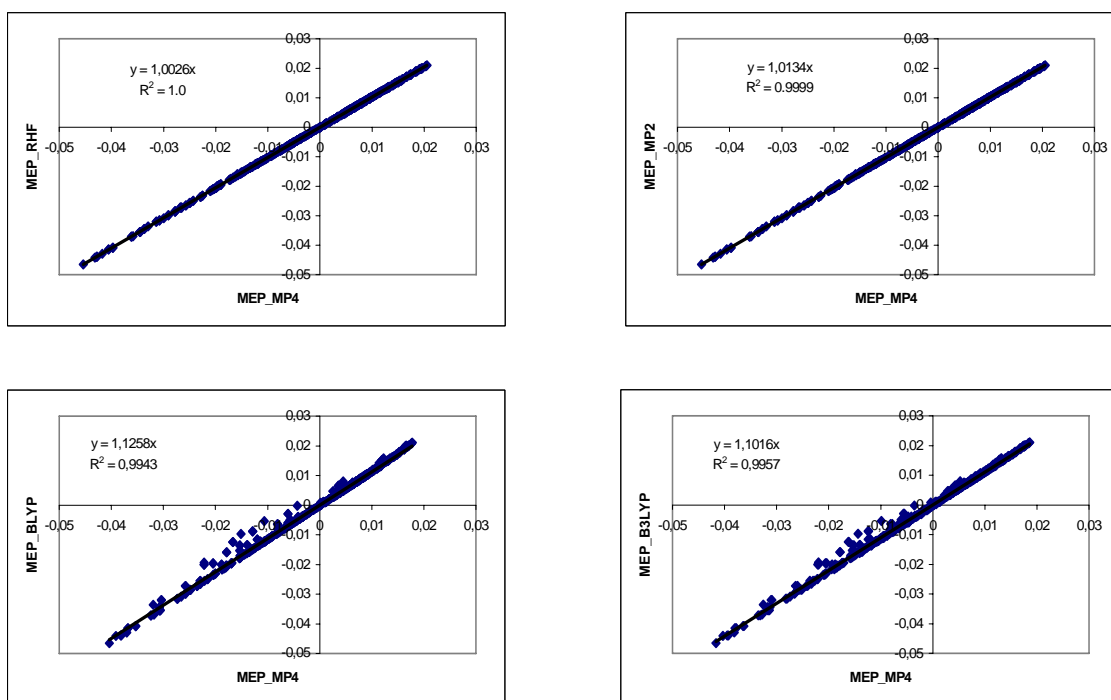
Gràfica 8.1.31 Molècula formaldehid. Gaussian. Anàlisi correlació. MPP.Càlculs BLYP.



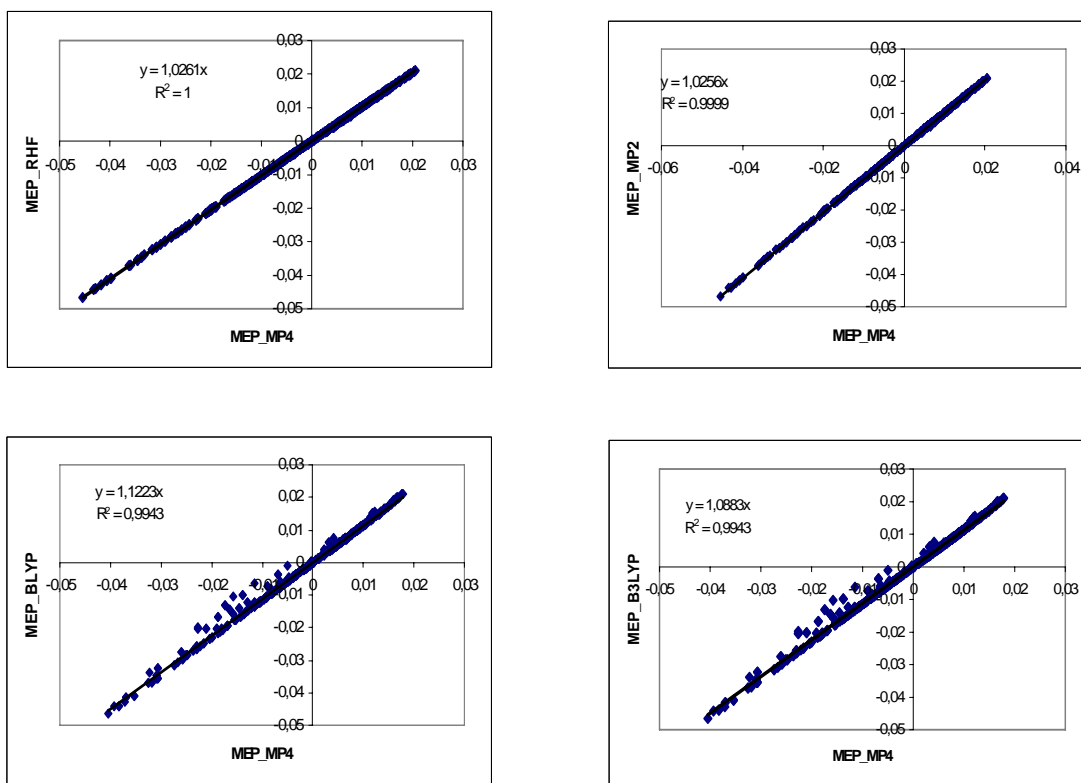
Gràfica 8.1.32 Molècula formaldehid. Gaussian. Anàlisi correlació. MPP Càlculs B3LYP.



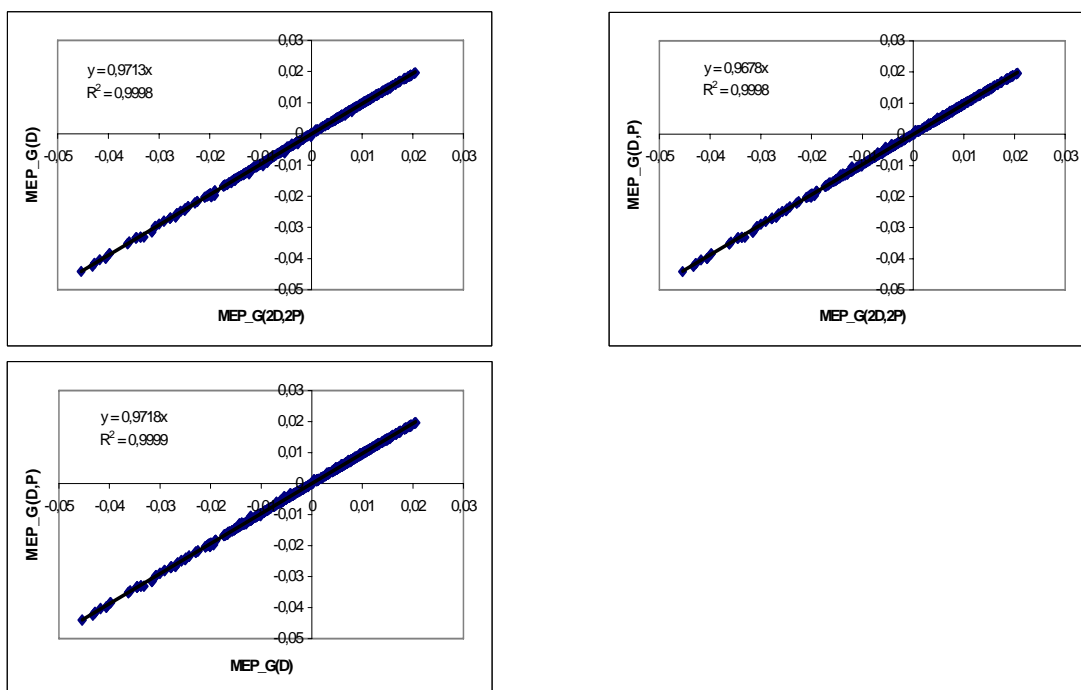
Gràfica 8.1.33 Molècula urea. Gaussian. Anàlisi correlació. MEP. Base 6-31G(2d,2p).



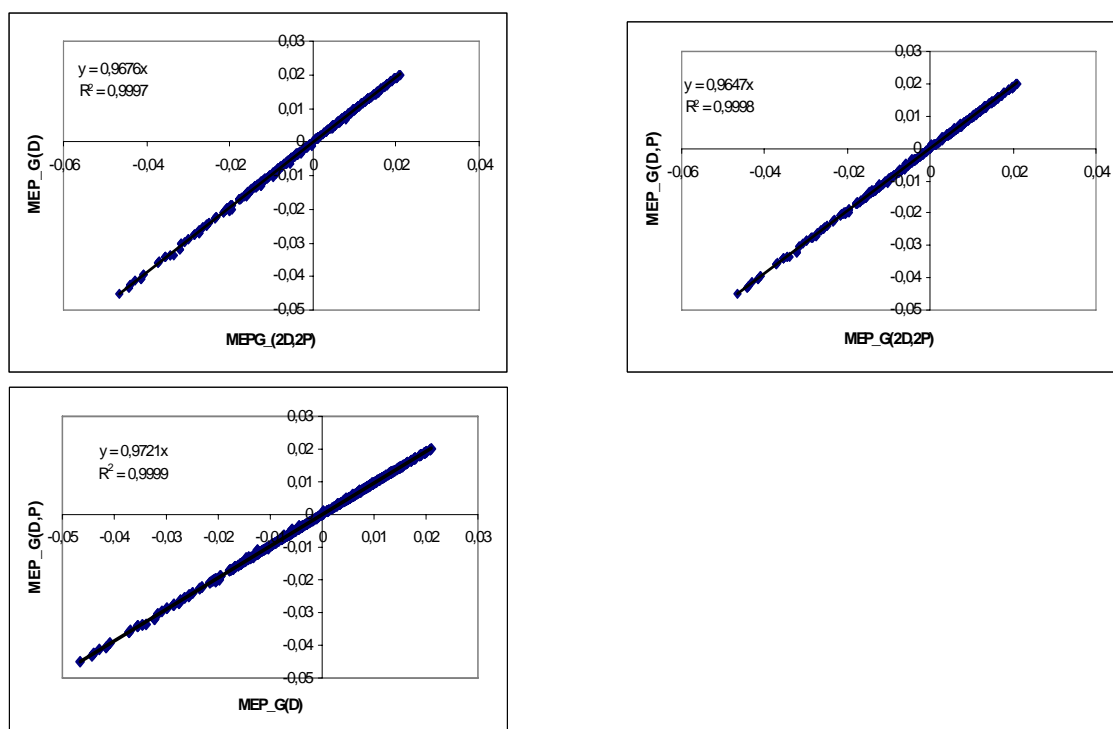
Gràfica 8.1.34 Molècula urea. Gaussian. Anàlisi correlació. MEP. Base 6-31G(d,p).



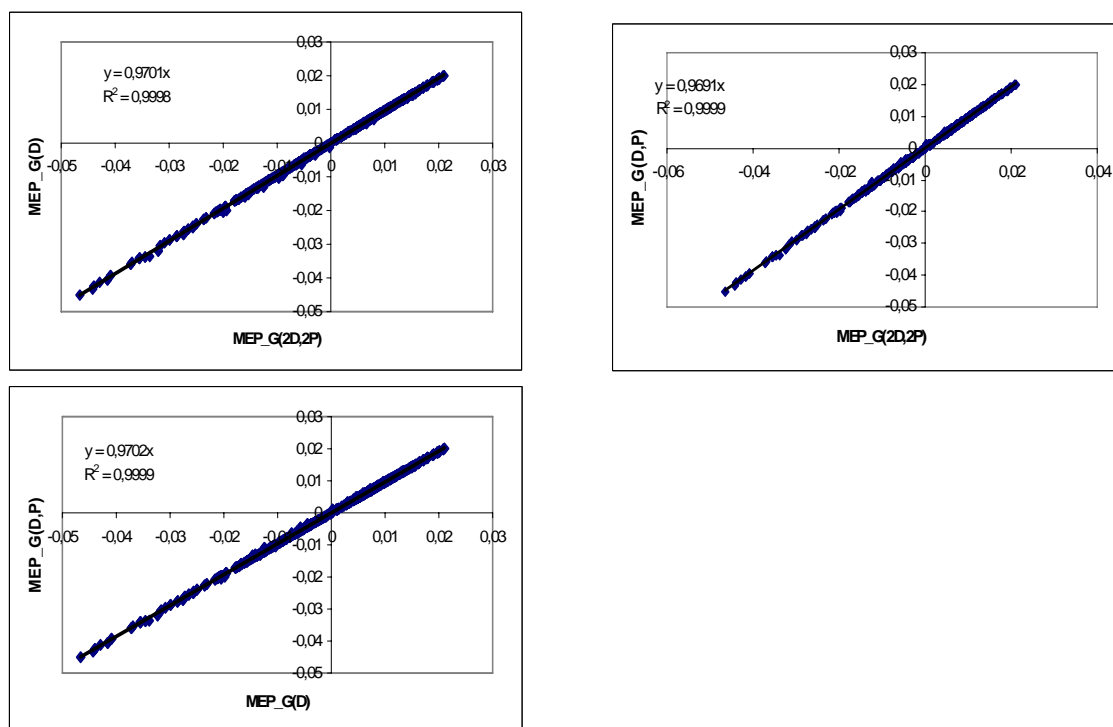
Gràfica 8.1.35 Molècula urea. Gaussian. Anàlisi correlació. MEP. Base 6-31G(d).



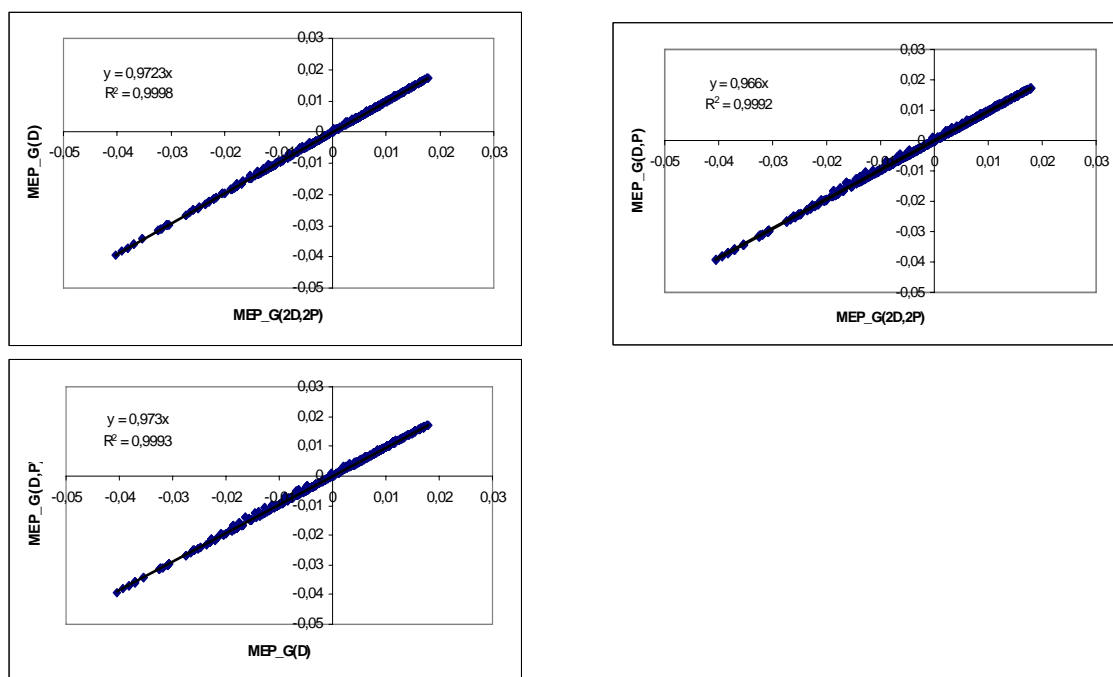
Gràfica 8.1.36 Molècula urea. Gaussian. Anàlisi correlació. MEP. Càlculs RHF.



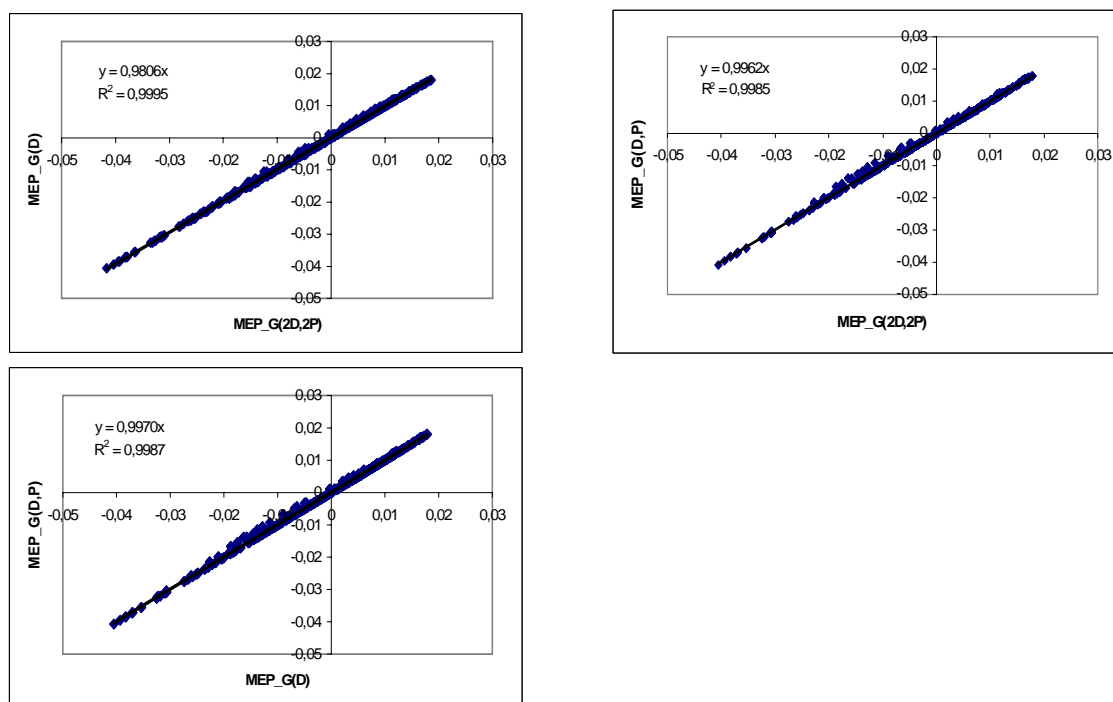
Gràfica 8.1.37 Molècula urea. Gaussian. Anàlisi correlació. MEP. Càlculs MP2.



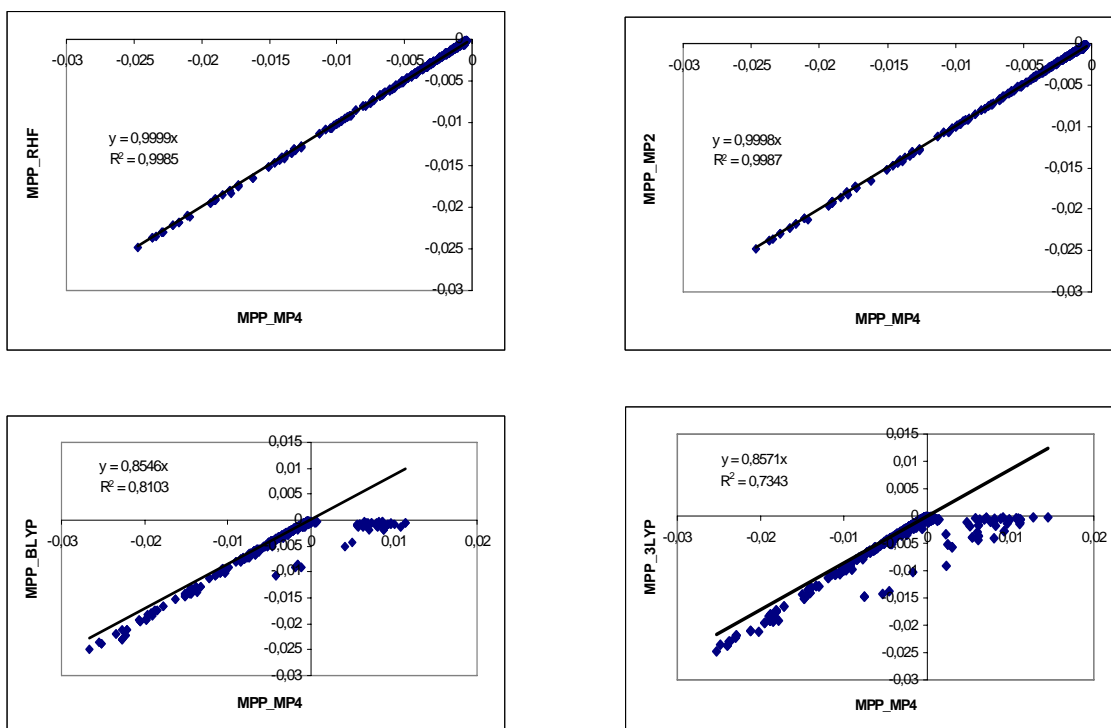
Gràfica 8.1.38 Molècula urea. Gaussian. Anàlisi correlació. MEP. Càlculs MP4.



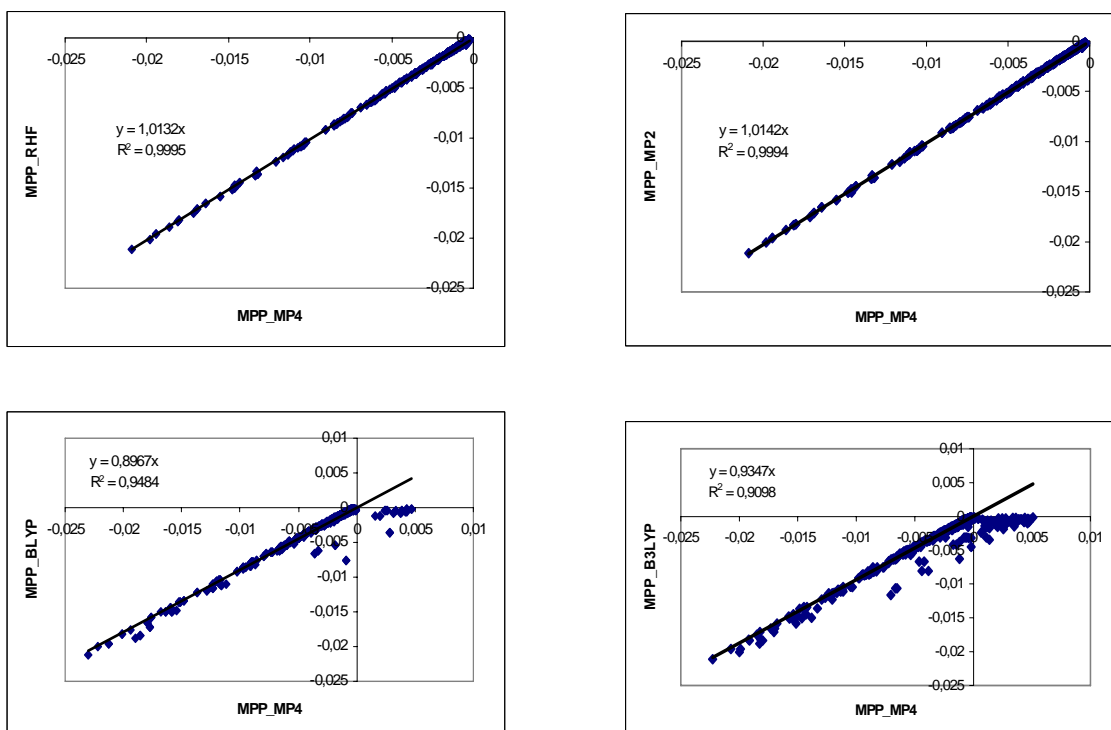
Gràfica 8.1.39 Molècula urea. Gaussian. Anàlisi correlació. MEP. Càlculs BLYP.



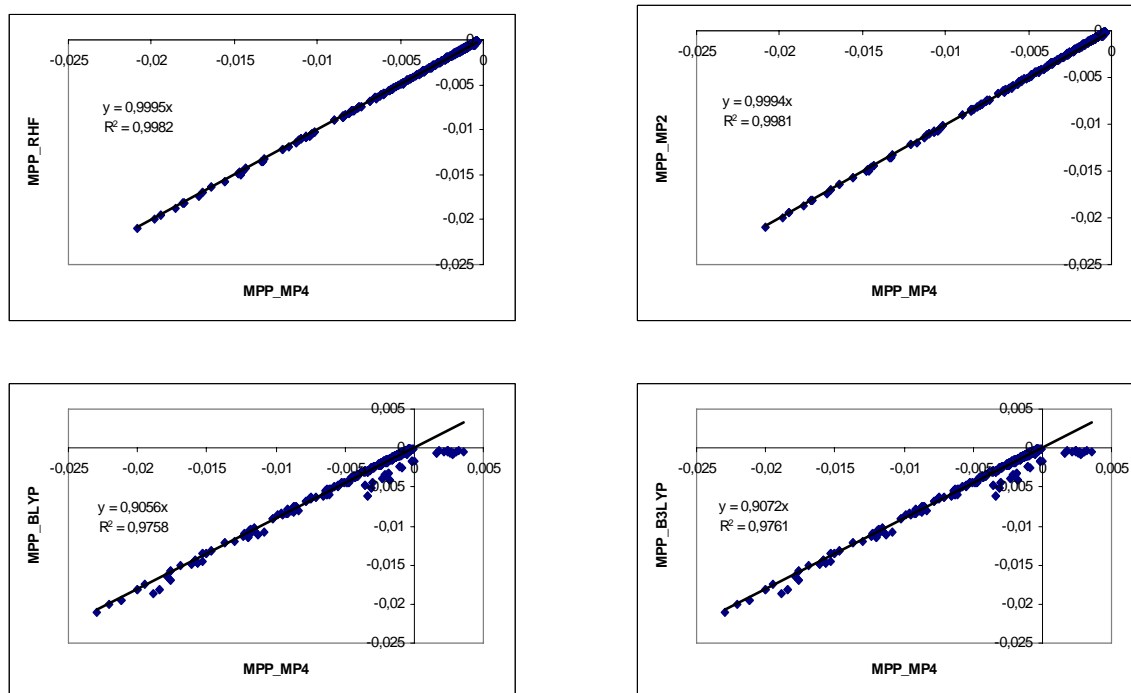
Gràfica 8.1.40 Molècula urea. Gaussian. Anàlisi correlació. MEP. Càlculs B3LYP.



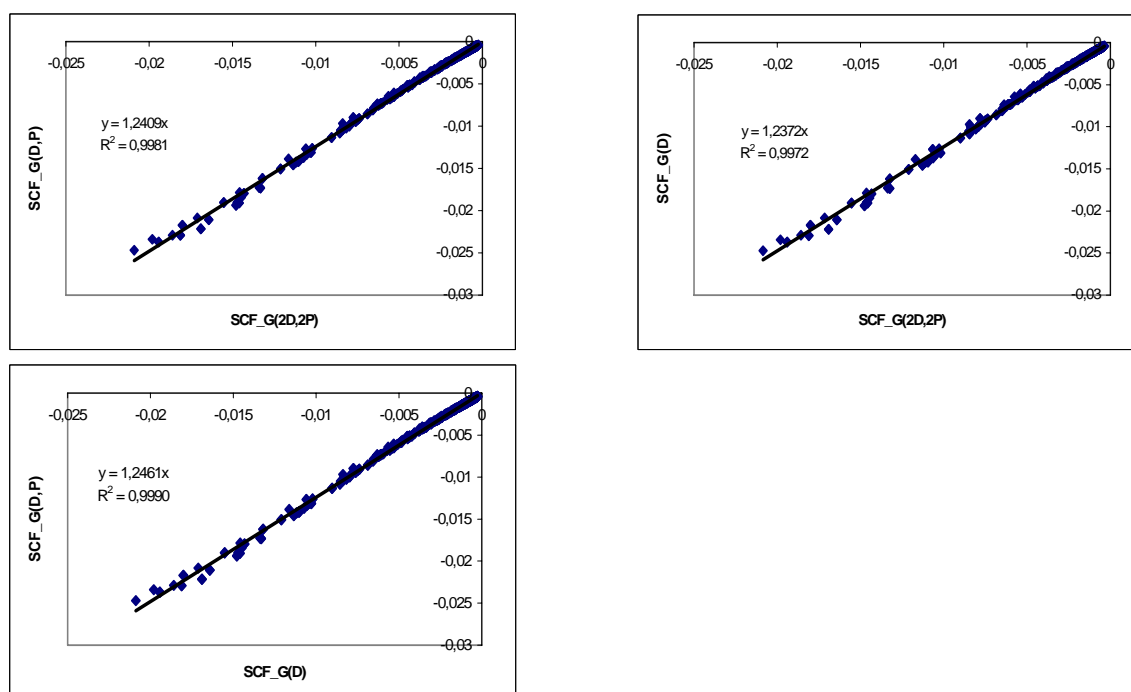
Gràfica 8.1.41 Molècula urea. Gaussian. Anàlisi correlació. MPP. Base 6-31G(2d,2p).



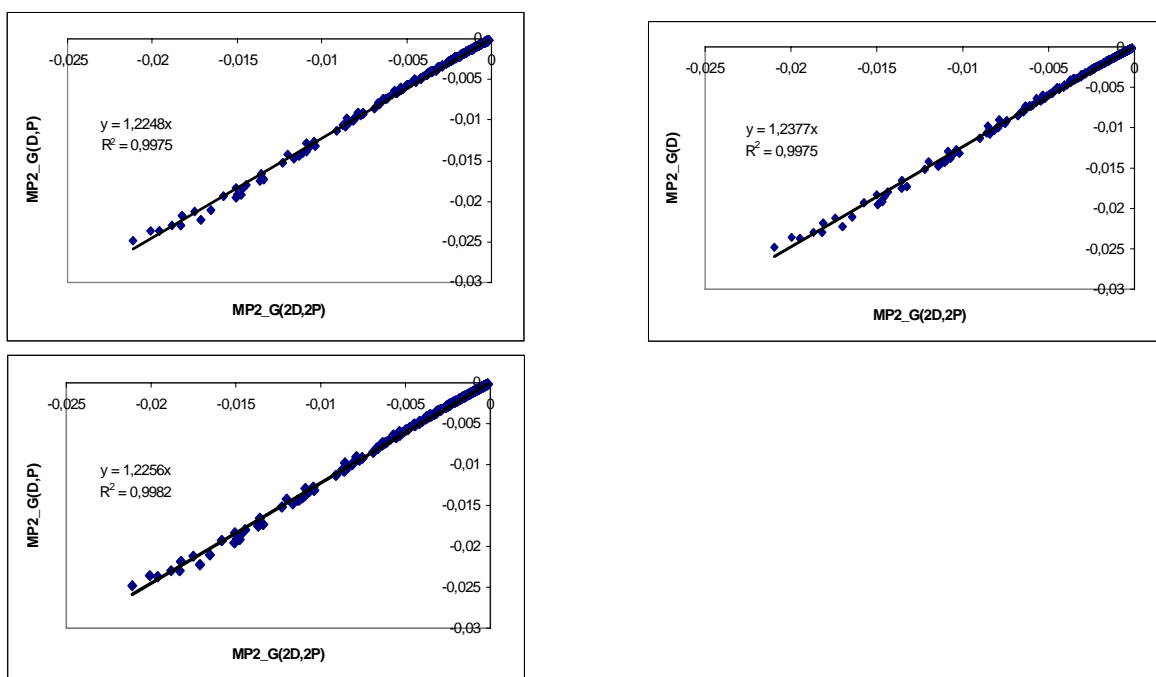
Gràfica 8.1.42 Molècula urea. Gaussian. Anàlisi correlació. MPP. Base 6-31G(d,p).



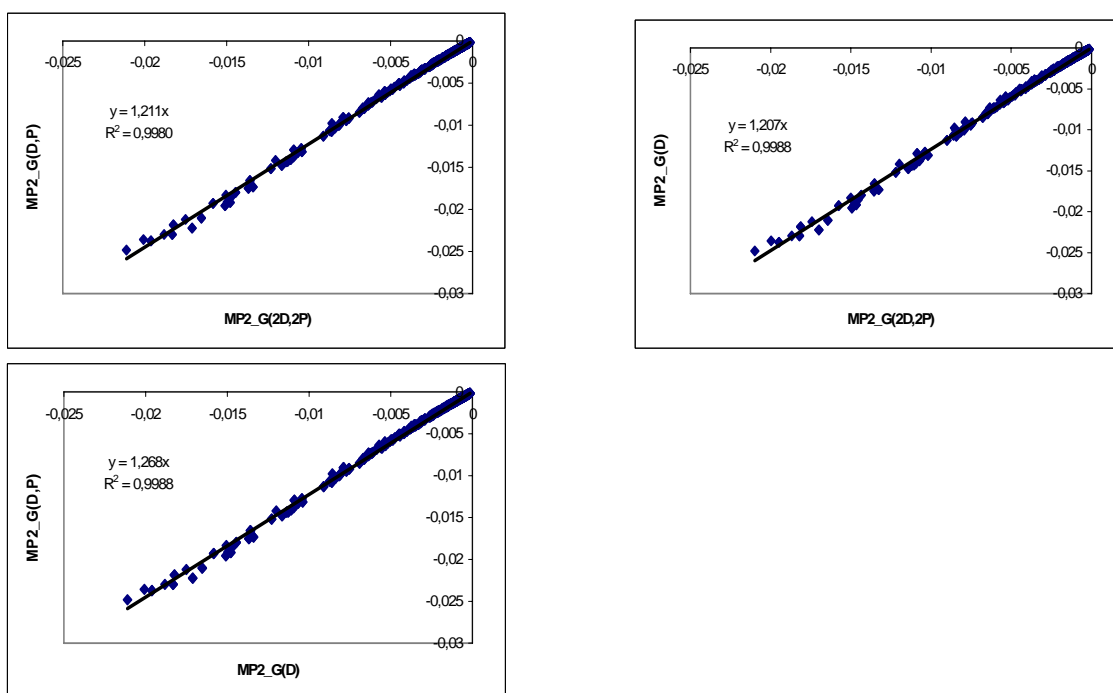
Gràfica 8.1.43 Molècula urea. Gaussian. Anàlisi correlació. MPP. Base 6-31G(D).



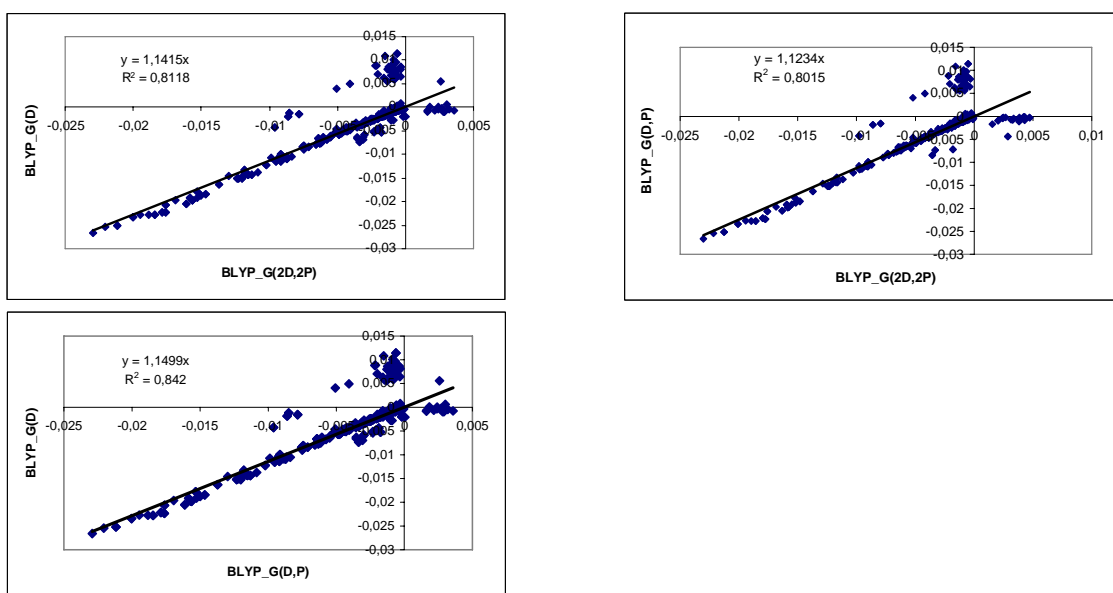
Gràfica 8.1.44 Molècula urea. Gaussian. Anàlisi correlació. MPP. Càlculs RHF.



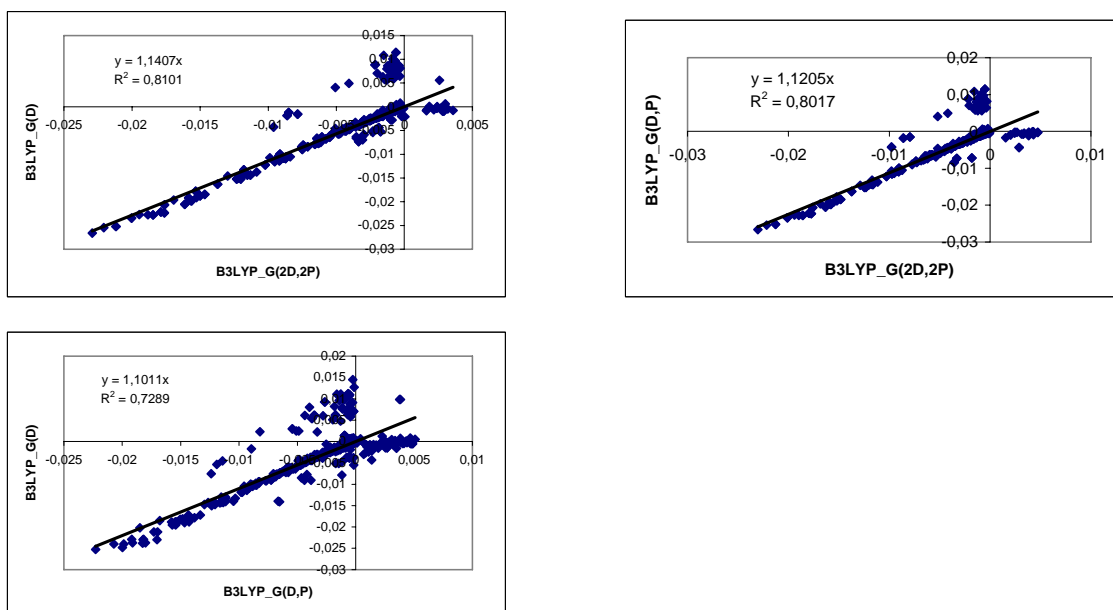
Gràfica 8.1.45 Molècula urea. Gaussian. Anàlisi correlació. MPP. Càlculs MP2.



Gràfica 8.1.46 Molècula urea. Gaussian. Anàlisi correlació. MPP. Càlculs MP4



Gràfica 8.1.47 Molècula urea. Gaussian. Anàlisi correlació. MPP. Càlculs BLYP.



Gràfica 8.1.48 Molècula urea. Gaussian. Anàlisi correlació. MPP. Càlculs B3LYP

8.2 Annex 2. Geometries optimitzades de dioxines i furans

Taula 8.2.1. Geometries optimitzades de les molècules 1mcdd, 28dd, 234dd, 236dd.

1mcdd	X	Y	Z
Cl	2.618053	-2.087265	0.000025
O	-0.097757	-1.007445	-0.000972
O	-0.775268	1.710007	-0.001003
O	-4.055139	0.161605	0.000688
C	-3.719131	-1.182381	0.000696
C	-2.387906	-1.564198	0.000142
C	-1.400379	-.600437	-0.000389
C	-1.734990	0.740798	-0.000381
C	-3.060497	1.125298	0.000136
C	0.523172	1.304869	-0.000396
C	1.506092	2.271623	0.000187
C	2.838209	1.894798	0.000726
C	3.183930	0.556692	0.000621
C	2.188560	-0.405940	0.000123
C	0.853152	-0.042968	-0.000389
H	-3.291969	2.174133	0.000117
H	-5.086580	0.462706	0.001094
H	-4.487338	-1.933580	0.001125
H	-2.096780	-2.597962	0.000144
H	4.211942	0.249844	0.000969
H	3.606552	2.645390	0.001162
H	1.211036	3.303932	0.000189

28dd	X	Y	Z
H	4.48932	1.62194	-0.00645
H	-2.33461	2.85282	-0.0034
H	-4.48932	1.62193	-0.00634
H	-2.32747	-2.06422	0.00373
Cl	5.04301	-1.17263	-0.00386
Cl	-5.04301	-1.17263	-0.00357
O	0	-1.00334	0.00858
O	0	1.80421	0.00448
C	3.55761	1.09049	-0.00332
C	3.54101	-0.29135	-0.00127
C	2.34932	-0.99198	0.00264
C	1.16039	-0.29186	0.00457
C	1.1608	1.09219	0.00252
C	2.35665	1.7793	0.00144
C	-1.16079	-1.09219	0.00219
C	-2.35665	1.7793	-0.00188
C	-3.5576	1.09049	-0.00348
C	-3.54102	-0.29136	-0.00136
C	-2.34932	-0.99199	0.00216
C	-1.16038	-0.29186	0.00417
H	2.33462	2.85283	-0.00306
H	2.32747	-2.06422	0.00416

234dd	X	Y	Z
Cl	-0.6782	3.45399	-0.00006
Cl	-4.16587	-0.61108	-0.00374
Cl	-1.72176	-2.68276	0.00132
O	0.84766	-1.39316	0.00569
O	1.34795	1.3527	0.00519
C	4.72226	0.02402	-0.00476
C	4.47363	-1.33934	-0.00457
C	3.17068	-1.80826	-0.00131
C	2.12577	-0.9078	0.00198
C	2.37322	0.44945	0.00175
C	3.66886	0.92286	-0.00168
C	0.08788	0.86654	0.00283
C	-0.97447	1.7498	0.00039
C	-2.2752	1.28407	-0.00153
C	-2.51995	-0.07407	-0.00123
C	-1.46475	-0.9777	0.00098
C	-0.16351	-0.49853	0.0031
H	3.8308	1.98447	-0.00178
H	5.73229	0.39016	-0.00732
H	5.28934	-2.03851	-0.00701
H	2.94722	-2.85865	-0.00111
H	-3.09193	1.97741	-0.00329

236dd	X	Y	Z
Cl	3.91972	1.88748	-0.00444
Cl	3.61585	2.23003	-0.00165
H	5.02965	-2.39361	-0.01571
H	5.41571	0.04609	-0.01464
H	2.70425	-3.26797	-0.00375
Cl	-4.415	-1.29524	-0.00546
O	0.57994	-1.85923	0.0114
O	1.00899	0.90897	0.01271
C	-2.60912	0.75487	-0.00008
C	-2.82212	-0.61421	-0.00053
C	-1.73935	-1.47955	0.00289
C	-0.45644	-0.98206	0.00722
C	-0.2443	0.38503	0.00777
C	-1.31437	1.24972	0.00389
C	2.04661	0.03285	0.00483
C	3.34242	0.51803	-0.00216
C	4.41946	-0.35179	-0.00946
C	4.19629	-1.71617	-0.00999
C	2.90454	-2.21327	-0.00338
C	1.84005	-1.33783	0.00414
H	-1.13273	2.30655	0.00426
H	-1.88947	-2.54131	0.0025

Taula 8.2.2. Geometries optimitzades de les molècules 237dd, 1234dd, 1378dd, 2367dd.

237dd	X	Y	Z
H	3.08995	-2.00454	0.00196
Cl	5.70827	-0.85957	-0.00357
Cl	-4.30753	-1.84787	-0.00162
H	4.89348	1.86977	-0.00428
H	2.63233	2.89208	-0.00066
Cl	-4.61175	1.35932	-0.00284
O	0.40645	1.62772	0.00427
O	0.67173	-1.16797	0.00537
C	-2.93194	-0.79507	-0.0002
C	-3.06266	0.58402	-0.0008
C	-1.92966	1.38248	0.00041
C	-0.67814	0.81039	0.00245
C	-0.54777	-0.56828	0.00302
C	-1.66885	-1.36595	0.00153
C	1.76099	-0.34983	0.00302
C	3.01036	-0.93504	0.00143
C	4.13031	-0.12471	-0.00107
C	4.0161	1.2528	-0.0021
C	2.7558	1.82548	-0.00005
C	1.63114	1.02759	0.00244
H	-1.55225	-2.43193	0.00203
H	-2.01593	2.45133	0.00007

1234dd	X	Y	Z
Cl	-3.57264	-1.56826	0.00571
Cl	-0.82123	-3.10927	-0.00258
Cl	-3.57264	1.56826	0.00575
Cl	-0.82121	3.10928	-0.0026
O	1.47121	1.39418	-0.01107
O	1.47121	-1.39419	-0.01107
C	5.02952	-0.69286	0.01054
C	5.02952	0.69285	0.01053
C	3.83197	1.38855	0.00366
C	2.64331	0.6895	-0.0035
C	2.64331	-0.68951	-0.00352
C	3.83197	-1.38856	0.00365
C	0.31901	-0.69266	-0.00658
C	-0.87723	-1.38637	-0.00255
C	-2.08532	-0.69477	0.00108
C	-2.08532	0.69478	0.00108
C	-0.87722	1.38637	-0.00257
C	0.31901	0.69266	-0.00658
H	3.80092	-2.46187	0.00354
H	5.95736	-1.23418	0.01596
H	5.95736	1.23417	0.01593
H	3.80093	2.46186	0.00356

1378dd	X	Y	Z
H	1.95988	-2.84422	0.00101
Cl	4.91218	1.004	-0.00207
Cl	-5.32874	-1.60075	-0.00166
Cl	-5.15299	1.61421	-0.00197
Cl	2.24826	2.71096	0.00053
Cl	4.68108	-2.16521	-0.00172
O	-0.30146	-1.66345	0.00356
O	-0.15097	1.13232	0.00332
C	-3.73708	0.61869	-0.00042
C	-3.81272	-0.76526	-0.00029
C	-2.6492	-1.51899	0.00094
C	-1.424	-0.89439	0.00214
C	-1.34878	0.48573	0.00202
C	-2.49814	1.24071	0.0007
C	0.96556	0.3659	0.002
C	2.2024	0.98726	0.00064
C	3.3727	0.22792	-0.0005
C	3.2748	-1.15636	-0.00035
C	2.03805	-1.7752	0.0009
C	0.89287	-1.01625	0.00212
H	-2.42295	2.31032	0.00061
H	-2.69278	-2.59035	0.00104

2367dd	X	Y	Z
Cl	-2.98532	2.18124	0.00175
Cl	-5.44753	0.1331	-0.00664
Cl	4.49561	1.91905	-0.00351
H	-4.3978	-2.46368	-0.00603
H	-2.06954	-3.31732	-0.00042
Cl	5.03462	-1.25571	-0.00457
O	0.04699	-1.88721	0.00747
O	-0.4194	0.87286	0.00848
C	3.20096	0.76916	-0.00042
C	3.43289	-0.59735	-0.00089
C	2.36255	-1.47781	0.00153
C	1.07344	-0.99699	0.00465
C	0.84262	0.36611	0.00512
C	1.89973	1.24636	0.00248
C	-1.44417	-0.01508	0.0044
C	-2.74343	0.47314	0.00101
C	-3.80627	-0.42141	-0.00261
C	-3.56818	-1.78411	-0.00315
C	-2.27343	-2.26352	-0.00004
C	-1.2166	-1.37981	0.00386
H	1.70337	2.3006	0.00283
H	2.52724	-2.53745	0.00115

Taula 8.2.3. Geometries optimitzades de les molècules. 12378dd, 123478dd, 12347dd, 12478dd.

12378dd	X	Y	Z
H	1.95988	-2.84422	0.00101
Cl	4.91218	1.004	-0.00207
Cl	-5.32874	-1.60075	-0.00166
Cl	-5.15299	1.61421	-0.00197
Cl	2.24826	2.71096	0.00053
Cl	4.68108	-2.16521	-0.00172
O	-0.30146	-1.66345	0.00356
O	-0.15097	1.13232	0.00332
C	-3.73708	0.61869	-0.00042
C	-3.81272	-0.76526	-0.00029
C	-2.6492	-1.51899	0.00094
C	-1.424	-0.89439	0.00214
C	-1.34878	0.48573	0.00202
C	-2.49814	1.24071	0.0007
C	0.96556	0.3659	0.002
C	2.2024	0.98726	0.00064
C	3.3727	0.22792	-0.0005
C	3.2748	-1.15636	-0.00035
C	2.03805	-1.7752	0.0009
C	0.89287	-1.01625	0.00212
H	-2.42295	2.31032	0.00061
H	-2.69278	-2.59035	0.00104

123478dd	X	Y	Z
Cl	4.65908	1.56798	-0.01352
Cl	-5.40777	-1.60928	-0.01211
Cl	-5.40779	1.60926	-0.01154
Cl	1.90793	3.10946	0.00461
Cl	4.6591	-1.56796	-0.01324
Cl	1.90799	-3.10945	0.00488
O	-0.38495	-1.39476	0.02075
O	-0.38496	1.39476	0.02088
C	-3.93974	0.69324	-0.0022
C	-3.93972	-0.69325	-0.00241
C	-2.73706	-1.38305	0.00498
C	-1.54975	-0.69006	0.01284
C	-1.54976	0.69006	0.01296
C	-2.73708	1.38304	0.00532
C	0.77066	0.6918	0.01256
C	1.96534	1.3878	0.00443
C	3.17265	0.69531	-0.00343
C	3.17266	-0.6953	-0.00335
C	1.96536	-1.38779	0.00452
C	0.77067	-0.6918	0.01255
H	-2.72074	2.45515	0.00532

12347dd	X	Y	Z
Cl	-4.26864	1.9968	-0.00009
Cl	-5.02686	-1.10939	-0.00016
Cl	5.02685	-1.1094	-0.00023
Cl	4.26863	1.99681	-0.00008
O	0.0	-1.94874	0.00014
C	-3.37464	-0.58923	-0.00006
C	-3.04244	0.76862	-0.00001
C	-1.7191	1.17005	0.00015
C	-0.72817	0.20362	0.00019
C	-1.0851	-1.13819	0.00015
C	-2.39241	-1.56515	0.00008
C	0.72818	0.20361	0.00021
C	1.7191	1.17005	0.00023
C	3.04244	0.76863	0.00009
C	3.37465	-0.58924	0.00002
C	2.39242	-1.56514	0.00009
C	1.08511	-1.13819	0.00014
H	-1.48435	2.21706	0.0002
H	-2.64666	-2.60644	0.0001
H	1.48435	2.21706	0.00026
H	2.64665	-2.60644	0.00006

12478dd	X	Y	Z
Cl	-4.84428	-1.89547	-0.00048
Cl	-5.12625	1.31088	-0.00067
Cl	2.10821	3.44207	-0.00003
Cl	5.20468	-0.92819	-0.001
Cl	2.57965	-2.76537	0.00024
O	0.14015	-1.24392	0.0012
O	-0.10397	1.53753	0.00106
C	-3.58285	0.52709	-0.00014
C	-3.4614	-0.8543	-0.00005
C	-2.20329	-1.43565	0.00034
C	-1.08027	-0.64162	0.00071
C	-1.20113	0.73358	0.00064
C	-2.44564	1.31925	0.00017
C	1.10923	0.93575	0.00061
C	2.24718	1.71922	0.0001
C	3.49806	1.13367	-0.00038
C	3.61659	-0.24211	-0.00034
C	2.48309	-1.04489	0.0002
C	1.23308	-0.44498	0.00066
H	-2.52297	2.38879	0.00011
H	-2.09256	-2.50225	0.00038
H	4.37585	1.74794	-0.00081

Taula 8.2.4. Geometries optimitzades de les molècules 124dd, ocdd.

124dd	X	Y	Z
Cl	-0.6782	3.45399	-0.00006
Cl	-4.16587	-0.61108	-0.00374
Cl	-1.72176	-2.68276	0.00132
O	0.84766	-1.39316	0.00569
O	1.34795	1.3527	0.00519
C	4.72226	0.02402	-0.00476
C	4.47363	-1.33934	-0.00457
C	3.17068	-1.80826	-0.00131
C	2.12577	-0.9078	0.00198
C	2.37322	0.44945	0.00175
C	3.66886	0.92286	-0.00168
C	0.08788	0.86654	0.00283
C	-0.97447	1.7498	0.00039
C	-2.2752	1.28407	-0.00153
C	-2.51995	-0.07407	-0.00123
C	-1.46475	-0.9777	0.00098
C	-0.16351	-0.49853	0.0031
H	3.8308	1.98447	-0.00178
H	5.73229	0.39016	-0.00732
H	5.28934	-2.03851	-0.00701
H	2.94722	-2.85865	-0.00111
H	-3.09193	1.97741	-0.00329

ocdd	X	Y	Z
Cl	4.65908	1.56798	-0.01352
Cl	-5.40777	-1.60928	-0.01211
Cl	-5.40779	1.60926	-0.01154
Cl	1.90793	3.10946	0.00461
Cl	4.6591	-1.56796	-0.01324
Cl	1.90799	-3.10945	0.00488
Cl	-1.9	3.1	0.0
Cl	-1.9	-3.1	0.0
O	-0.38495	-1.39476	0.02075
O	-0.38496	1.39476	0.02088
C	-3.93974	0.69324	-0.0022
C	-3.93972	-0.69325	-0.00241
C	-2.73706	-1.38305	0.00498
C	-1.54975	-0.69006	0.01284
C	-1.54976	0.69006	0.01296
C	-2.73708	1.38304	0.00532
C	0.77066	0.6918	0.01256
C	1.96534	1.3878	0.00443
C	3.17265	0.69531	-0.00343
C	3.17266	-0.6953	-0.00335
C	1.96536	-1.38779	0.00452
C	0.77067	-0.6918	0.01255
H	-2.72074	2.45515	0.00532

Taula 8.2.5. Geometries optimitzades de les molècules 2347df, 2348df, 12379df, 12478df,

2347df	X	Y	Z
Cl	1.87348	2.7897	0.00015
Cl	-5.7934	0.25149	-0.00015
Cl	4.278	0.70388	0.00004
Cl	3.62487	-2.39115	-0.00025
O	-0.8011	1.36368	-0.00019
C	-4.09838	-0.14206	-0.00007
C	-3.72233	-1.48522	0.00044
C	-2.38569	-1.82956	0.00065
C	-1.43441	-0.81748	0.00027
C	-1.85168	0.50557	-0.00027
C	-3.17733	0.8864	-0.00042
C	0.01999	-0.75101	-0.0001
C	1.05059	-1.6733	-0.0002
C	2.35179	-1.21175	-0.00009
C	2.64195	0.15845	0.00017
C	1.606	1.08752	0.00027
C	0.3137	0.60484	0.00011
H	-2.09677	-2.8645	0.0009
H	-3.47132	1.91721	-0.00069
H	0.86234	-2.72922	-0.00025
H	-4.48056	-2.24396	0.00066

2348df	X	Y	Z
Cl	-5.43346	1.20229	-0.00015
Cl	2.44374	-2.62036	-0.00013
Cl	4.30719	-0.04002	-0.00008
Cl	2.9644	2.82385	0.00011
O	-0.48547	-1.84325	0.00012
C	-4.0556	-1.12751	0.00016
C	-3.96394	0.26411	-0.00007
C	-2.74967	0.91631	-0.0002
C	-1.60097	0.13524	-0.00005
C	-1.70578	-1.24756	0.0002
C	-2.91593	-1.90897	0.0003
C	-0.16956	0.40274	0.00008
C	0.62258	1.5367	0.00014
C	1.99443	1.38514	0.00008
C	2.58993	0.11673	-0.00005
C	1.79431	-1.02443	-0.00011
C	0.42526	-0.85035	-0.00003
H	-2.70479	1.98853	-0.0003
H	-2.96718	-2.98082	0.00044
H	0.19745	2.52146	0.00016
H	-5.02386	-1.58983	0.00025

12379df	X	Y	Z
Cl	1.86437	2.74161	-0.00122
Cl	-1.35122	2.45626	0.00107
Cl	5.54853	-1.06893	0.00038
Cl	-4.44173	-1.97533	-0.00075
Cl	-4.14891	-1.15577	0.00014
O	0.59586	-2.15599	0.00013
C	3.86063	-0.66784	0.00012
C	3.48679	0.66878	-0.00037
C	2.15346	1.03369	-0.00066
C	1.15102	0.05351	-0.00008
C	1.60697	-1.26806	0.00022
C	2.92353	-1.67275	0.0003
C	-0.32608	-0.07618	0.00035
C	-1.46624	0.73784	0.00084
C	-2.73495	0.16049	0.00041
C	-2.88043	-1.23087	-0.00006
C	-1.78088	-2.05565	-0.00011
C	-0.54759	-1.451	0.00005
H	4.23866	1.43151	-0.00059
H	3.18504	-2.71195	0.00051
H	-1.8755	-3.12285	-0.00038

12478df	X	Y	Z
Cl	2.86749	3.30467	0.00014
Cl	-4.19095	-1.93566	-0.00019
Cl	1.39526	-2.78962	0.00012
Cl	-4.94231	1.17083	0.00008
Cl	4.44551	-1.8305	-0.00007
O	0.07898	1.99829	0.00011
C	-3.29261	0.64769	0.00002
C	-2.96143	-0.71174	-0.00011
C	-1.64106	-1.11906	-0.00013
C	-0.64318	-0.15486	0.0
C	-1.003	1.18515	0.00013
C	-2.30836	1.61883	0.00015
C	0.81624	-0.14886	0.00013
C	1.81644	-1.11475	0.00023
C	3.13623	-0.69553	0.0001
C	3.458	0.65993	-0.00004
C	2.47318	1.62371	-0.00005
C	1.16147	1.19756	0.00001
H	-1.41246	-2.16474	-0.00018
H	-2.55507	2.66182	0.00023
H	4.48899	0.95298	-0.00016

Taula 8.2.6. Geometries optimitzades de les molècules 13478df, 23478df, 2378df, 23479df.

13478df	X	Y	Z
Cl	-3.22017	-2.52558	-0.00019
Cl	4.58069	1.45028	-0.00034
Cl	-0.84352	3.27437	0.00021
Cl	4.77967	-1.73941	0.00021
Cl	-5.13106	0.07298	-0.00008
O	-0.31117	-1.67893	0.00003
C	3.24606	0.937	0.00013
C	3.15686	0.45919	-0.00017
C	1.92726	1.09028	-0.0003
C	0.77842	0.31335	-0.00008
C	0.89806	-1.06858	0.0002
C	2.10762	-1.72278	0.0003
C	-0.65739	0.5613	0.00011
C	-1.49321	1.66764	0.00023
C	-2.8591	1.49143	0.00016
C	-3.40734	0.21046	0.00002
C	-2.60351	-0.91855	-0.00009
C	-1.23803	-0.70205	-0.00003
H	1.88218	2.16021	-0.00043
H	2.16987	-2.79274	0.00042
H	-3.5121	2.3409	0.00017

23478df	X	Y	Z
Cl	-2.7097	-2.72875	-0.0002
Cl	4.70404	1.92136	-0.00026
Cl	5.20291	1.23549	0.00024
Cl	-4.82833	-0.35346	-0.00005
Cl	-3.78795	2.6325	0.00019
O	0.12438	-1.65565	-0.00001
C	3.59974	-0.58225	0.00015
C	3.38055	0.79996	-0.00015
C	2.09589	1.30941	-0.0003
C	1.02885	0.42654	-0.00013
C	1.27397	-0.93823	0.00014
C	2.54112	-1.47359	0.00026
C	-0.4222	0.54724	0.00008
C	-1.3262	1.59414	0.00018
C	-2.67511	1.3019	0.00015
C	-3.13687	-0.0212	0.00001
C	-2.2284	1.07488	-0.00011
C	-0.88495	-0.7601	-0.00006
H	1.94773	2.37206	-0.00042
H	2.70697	-2.53247	0.00037
H	-1.00465	2.61736	0.00019

2378df	X	Y	Z
Cl	-4.26864	1.9968	-0.00009
Cl	-5.02686	-1.10939	-0.00016
Cl	5.02685	-1.1094	-0.00023
Cl	4.26863	1.99681	-0.00008
O	0.0	-1.94874	0.00014
C	-3.37464	-0.58923	-0.00006
C	-3.04244	0.76862	-0.00001
C	-1.7191	1.17005	0.00015
C	-0.72817	0.20362	0.00019
C	-1.0851	-1.13819	0.00015
C	-2.39241	-1.56515	0.00008
C	0.72818	0.20361	0.00021
C	1.7191	1.17005	0.00023
C	3.04244	0.76863	0.00009
C	3.37465	-0.58924	0.00002
C	2.39242	-1.56514	0.00009
C	1.08511	-1.13819	0.00014
H	-1.48435	2.21706	0.0002
H	-2.64666	-2.60644	0.0001
H	1.48435	2.21706	0.00026
H	2.64665	-2.60644	0.00006

23479df	X	Y	Z
O	0.002	-1.898	0.219
C	-1.117	-1.022	0.219
C	-0.701	0.335	0.213
C	0.715	0.434	0.209
C	1.122	-1.023	0.219
C	-1.669	1.375	0.226
C	-3.053	1.056	0.244
C	-3.469	-0.302	0.197
C	-2.501	-1.341	0.211
C	2.498	-1.376	0.215
C	3.466	-0.272	0.235
C	3.084	1.095	0.226
C	1.709	1.448	0.212
Cl	-5.142	-0.686	0.108
Cl	-1.166	3.018	0.203
Cl	4.287	2.324	0.215
Cl	5.142	-0.653	0.259
Cl	3.006	-3.018	0.170
H	-3.802	1.860	0.285
H	-2.823	-2.393	0.206
H	1.413	2.508	0.195

Taula 8.2.7. Geometries optimitzades de les molècules .. [.234678df](#),
[12468df](#), [123678df](#), [123478df](#).

234678df	X	Y	Z
O	0.002	-1.643	0.011
C	-1.117	-0.767	0.011
C	-0.701	0.591	0.005
C	0.715	0.689	0.001
C	1.122	-0.768	0.011
C	-1.669	1.630	0.018
C	-3.053	1.311	0.036
C	-3.469	-0.047	-0.011
C	-2.501	-1.086	0.003
C	2.498	-1.121	0.007
C	3.466	-0.017	0.028
C	3.084	1.351	0.018
C	1.709	1.704	0.005
Cl	-3.005	-2.729	0.025
Cl	-5.142	-0.431	-0.100
Cl	-4.224	2.568	0.100
Cl	4.287	2.579	0.007
Cl	5.142	-0.398	0.051
Cl	3.006	-2.763	-0.038
H	-1.347	2.681	0.023
H	1.413	2.763	-0.012

12468df	X	Y	Z
O	0.003	-1.941	-0.005
C	-1.117	-1.065	-0.005
C	-0.701	0.293	-0.011
C	0.716	0.391	-0.015
C	1.123	-1.066	-0.005
C	-1.669	1.332	0.002
C	-3.052	1.013	0.020
C	-3.468	-0.345	-0.027
C	-2.500	-1.384	-0.013
C	2.498	-1.419	-0.009
C	3.466	-0.315	0.012
C	3.084	1.052	0.002
C	1.709	1.405	-0.011
Cl	-3.004	-3.027	0.009
Cl	-4.223	2.270	0.084
Cl	1.247	3.061	-0.008
Cl	4.287	2.281	-0.009
Cl	3.006	-3.061	-0.054
H	-1.346	2.383	0.007
H	-4.539	-0.591	-0.084

123678df	X	Y	Z
O	-0.525	-1.970	0.011
C	-1.278	-0.764	0.011
C	-0.422	0.369	0.005
C	0.942	-0.023	0.001
C	0.826	-1.531	0.011
C	-0.977	1.677	0.018
C	-2.386	1.851	0.036
C	-3.241	0.717	-0.011
C	-2.687	-0.590	0.003
C	1.998	-2.333	0.007
C	3.286	-1.627	0.028
C	3.394	-0.211	0.018
C	2.223	0.591	0.005
Cl	-3.722	-1.962	0.025
Cl	-4.945	0.928	-0.100
Cl	-3.056	3.432	0.100
Cl	2.354	2.305	0.008
Cl	4.945	0.532	0.007
Cl	4.730	-2.558	0.051
H	-0.314	2.555	0.023
H	1.944	-3.432	-0.002

123478df	X	Y	Z
O	0.002	-1.941	0.011
C	-1.117	-1.065	0.011
C	-0.701	0.293	0.005
C	0.715	0.391	0.001
C	1.122	-1.066	0.011
C	-1.669	1.332	0.018
C	-3.053	1.013	0.036
C	-3.469	-0.345	-0.011
C	-2.501	-1.384	0.003
C	2.498	-1.419	0.007
C	3.466	-0.315	0.028
C	3.084	1.052	0.018
C	1.709	1.405	0.005
Cl	-5.142	-0.729	-0.100
Cl	-4.224	2.270	0.100
Cl	1.246	3.061	0.008
Cl	4.287	2.281	0.007
Cl	5.142	-0.696	0.051
Cl	3.006	-3.061	-0.038
H	-1.347	2.383	0.023
H	-2.823	-2.436	-0.002

Taula 8.2.8. Geometries optimitzades de les molècules .. 12378df, 1237df.

12378df	X	Y	Z
Cl	4.3506	1.87088	0.00023
Cl	-1.2679	2.53699	-0.00015
Cl	5.19686	-1.21172	-0.00004
Cl	-4.83756	-1.5741	0.00014
Cl	-4.25842	1.53137	-0.00009
O	0.2021	-2.19358	-0.00034
C	3.53181	-0.73834	-0.00007
C	3.15946	0.60933	0.00003
C	1.82646	0.97594	0.0001
C	0.85866	-0.0175	0.00008
C	1.25914	-1.3471	-0.00008
C	2.57705	-1.73944	-0.00019
C	-0.59934	-0.07133	-0.0001
C	-1.63787	0.84896	-0.00021
C	-2.95333	0.40025	0.00002
C	-3.21292	-0.97659	0.00032
C	-2.19413	-1.90901	0.00022
C	-0.90809	-1.42506	0.00011
H	1.56594	2.01416	0.00023
H	2.8567	-2.77411	-0.00032
H	-2.40171	-2.96022	0.00016

1237df	X	Y	Z
Cl	-1.1194	2.62181	0.00002
Cl	5.73978	-0.30911	-0.00026
Cl	-4.19032	-1.87489	-0.00026
Cl	-3.97307	1.27707	-0.00007
O	0.88742	-1.90772	0.00024
C	4.01452	-0.0819	-0.00001
C	3.50722	1.21554	-0.00009
C	2.14292	1.42951	0.00002
C	1.29051	0.33061	0.00015
C	1.84152	-0.94513	0.00023
C	3.19739	-1.19439	0.00019
C	-0.15157	0.10809	0.0002
C	-1.28978	0.90072	0.00016
C	-2.54547	0.30333	0.00014
C	-2.64478	-1.09284	0.00001
C	-1.52463	-1.90153	0.0001
C	-0.30319	-1.27306	0.00021
H	4.1854	2.04667	-0.00026
H	1.75936	2.43013	-0.00004
H	3.58916	-2.19217	0.00024
H	-1.61001	-2.96968	0.00001

8.3 Annex 3. Programes de càlcul emprats

GAUSSIAN 94 (FRISCH 1995)

Gaussian és un paquet per el càlcul d'estructura electrònica desenvolupat per Gaussian Inc. disponible per les arquitectures IBM-SP2, SGI-ORIGIN2000 i Digital ALPHA-400 (multiprocessador), tant per a córrer en sèrie como en paral·lel. Gaussian pot realitzar càlculs a diferents nivells de teoria: ab-initio, semiempírics, teoria del funcional de la densitat i interacció de configuració; permetent modelar sistemes químics que van des de àtoms fins a molècules d'un elevat nombre de electrons, de molts electrons, en estats base i excitats, en fase gasosa o en solució.

Exemples de propietats que podem calcular:

- Energies i estructures moleculars.
- Energies i estructures d'estats de transició.
- Freqüències vibracionals.
- Espectres infrarroig i Raman.
- Propietats termodinàmiques.
- Energies de reacció i de enllaç.
- Trajectòries de reacció.
- Orbitals moleculars
- Càrregues atòmiques.
- Moments multipolars.
- Afinitats electròniques i potencials de ionització.
- Polarització i hiperpolarització. Potencials electrostàtics i densitats electròniques.

En el present treball s'han utilitzat els programes GAUSSIAN 94¹ i GAUSSIAN 98².

¹ www.gaussian.com (The official Gaussian Website)

² www.gaussian.com (The official Gaussian Website)

MEPHISTO (FRANCL 1985)

El programa MEPHISTO, basat en la teoria de pertorbacions de Rayleigh-Schrödinger, ha estat descrit a l'apartat de 'Mètodes de càlcul de la funció d'ona multieletrònica', per a calcular el potencial de polarització utilitzant la teoria de pertorbació de Rayleigh-Schrödinger a partir de la funció d'ona HF, avaluant les correccions de primer i segon ordre a l'energia d'una molècula interaccionant amb una càrrega puntual.

Tant la descripció del programa com el programa Fortran es troben al QCPE (Quantum Chemistry Program Exchange)³

SURFER 6.0

El programa Surfer 6.0⁴ és un software de contorn i visualització 3D que funciona en Microsoft Windows. Permet la transformació de dades de contorn, superfície, vectors, imatges i mapes, així com el tractament numèric de les dades representades:

- Contorn de Mapes
- Superfície de mapes en 3D
- Mapes de vectors
- Imatge de Mapes
- Grids
- Variogrames

Es presenta un breu resum dels continguts del programa :

Contour Maps	Map Layers
3D Surface Maps	Stacking Maps
Image Maps	Map Projections
Shaded Relief Maps	Customize Your Map
Post Maps	Superior Gridding
3D Wireframe Maps	Variograms
Vector Maps	Faults and Breaklines
Base Maps	Grid Functions

³ M. Miller Francl, QCPE No. 490

⁴ www.goldensoftware.com

SIMCA 10.0

El programa Simca 10.0⁵ ens permet treballar amb tecnologia multivariant, en base al disseny d'experiments i els diferents mètodes multivariants amb els mètodes apropiats de l'anàlisi multivariant, com l'anàlisi de components principals (PCA) i de components principals parcials (PLS)

Es presenta un breu resum:

Analysis of variance	<i>A statistical technique to separate and estimate different causes of variation.</i>
ANOVA	<i>See Analysis of Variance.</i>
BSPC	<i>Batch Statistical Process Control.</i>
Chemometrics	<i>The application of mathematical and statistical methods to chemical data.</i>
Collinearity	<i>A high level of correlation between variables.</i>
Correlation	<i>Also correlation coefficient, the strength of the relationship between variables.</i>
Cross-validation	<i>A procedure of calculations to simulate the predictive power of a model, in order to determine its significance.</i>
D-optimal design	<i>A computer-generated design for non-standard conditions or when the experimental domains is distorted. The D in D-optimal stands for determinant.</i>
Design of Experiments	<i>A strategy for setting up a set of experiments in which all variables are varied in a systematic manner, for the purpose of determining the correlation between variables and to predict results.</i>
DOE	<i>See Design of Experiments.</i>
Interaction	<i>Also interaction coefficient, the strength of the relation between an independent variable and dependent variables, as a function of another independent variable.</i>
Metabonomics	<i>The study of excreted metabolites of a species or an individual organism, involving measurements of the response to an influence.</i>
MLR	<i>Multiple Linear Regression.</i>
MODDE	<i>See MODDE in the product menu.</i>
Model	<i>The mathematical description of the behaviour of a system.</i>
MSPC	<i>Multivariate Statistical Process Control.</i>
Multivariate Data Analysis	<i>Regression analysis by projection methods such as PCA and PLS.</i>
MVA	<i>See Multivariate Data Analysis.</i>
NIPALS	<i>See Nonlinear Iterative Partial Least Squares.</i>
Nonlinear Iterative Partial Least Squares	<i>Algorithm for calculating principal components.</i>
O-PLS	<i>Also Orthogonal PLS, a modification of PLS in which systematic variation in independent factors is divided into two parts: either related or non-related to the dependent responses.</i>

⁵ www.umetrics.com/simca

<i>OLS</i>	<i>Ordinary Least Squares, equivalent to MLR.</i>
<i>Omics</i>	<i>The study of a group or system of biomolecules.</i>
<i>Partial Least Squares Projections to Latent Structures</i>	<i>See Projections to Latent Structures.</i>
<i>PAT</i>	<i>See Process Analytical Technology.</i>
<i>PCA</i>	<i>See Principal Component Analysis</i>
<i>PCR</i>	<i>See Principal Component Regression.</i>
<i>PLS</i>	<i>See Projections to Latent Structures.</i>
<i>PLS-DA</i>	<i>Also PLS Discriminant Analysis, a PLS analysis involving a dummy variable for classification.</i>
<i>Prediction</i>	<i>A statement (usually quantitative) about what will happen under specific conditions, as a logical consequence of scientific theories.</i>
<i>Principal Component Analysis</i>	<i>A transformation where the data set receives a new coordinate system, in which new axes follow the direction of greatest variance in the data set.</i>
<i>Principal Component Regression</i>	<i>A regression technique that combines principal component calculations with MLR.</i>
<i>Process Analytical Technology</i>	<i>Systems for analysis and control of manufacturing processes based on timely measurements, during processing, of critical quality parameters and performance attributes of raw and in-process materials and processes to assure acceptable end product quality at the completion of the process.</i>
<i>Projections to Latent Structures</i>	<i>A regression technique for modelling the relationship between projections of dependent factors and independent responses.</i>
<i>QSAR</i>	<i>See Quantitative Structure-Activity Relationship.</i>
<i>Quantitative Structure-Activity Relationship</i>	<i>Estimation of the strength of a mathematical relation between chemical structure and pharmacological activity for a series of compounds.</i>
<i>Rectangular Experimental Designs for Multi-Unit Platforms</i>	<i>A set of designs, for experiments in 96-well plates using multi-pipettes.</i>
<i>REDMUP</i>	<i>See Rectangular Experimental Design for Multi-Unit Platforms.</i>
<i>Regression</i>	<i>The fitting of a curve to data points, expresses the mathematical relationship between variables.</i>
<i>SBOL</i>	<i>See SIMCA-Batch-On-Line in the product menu.</i>
<i>Semiconductor</i>	<i>A material with low conductivity, and hence an intermediate between a conductor and an insulator.</i>
<i>SIMCA</i>	<i>Soft Independent Modeling of Class Analogy. See also the SIMCA software family in the product menu.</i>
<i>Validity</i>	<i>Term stemming from logical argument, stating that an argument is valid if, for every model, all premises in the model are true, then the conclusion in the model is true.</i>
<i>Variability</i>	<i>The variation between samples in the same condition, without systematic error.</i>
<i>Variance</i>	<i>Measurement of variability, equal to the square of standard deviation.</i>

8.4 Annex 4. Publicacions

A continuació es presenten els articles i pòsters publicats amb resultats obtinguts amb el treball presentat :

Roset L, Rubio-Martinez J, Pérez J J, 2009; *Analysis of the basis set and correlation effects on the computation of molecular polarizabilities using molecular polarization maps.* Theoretical Chemistry Accounts, **123(1-2):** 113

Roset L, Perez J J, 2010, *Assessment of the performace of the functional density theory,* Journal of Computational Methods in Sciences and Engineering **10:** 521

Roset L, 2007, *Actividad Biológica y Toxicidad de Dioxinas y Furanos. relación estructura Química Actividad Biológica.* A: "Revista de Toxicología". Dra. M. del Prado Mínguez Santiyán Universidad de Extremadura **p. 103.**

Analysis of the basis set and correlation effects on the computation of molecular polarizabilities using molecular polarization maps

Lurdes Roset · Jaime Rubio-Martinez ·
Juan J. Perez

Received: 25 February 2009 / Accepted: 26 February 2009 / Published online: 17 March 2009
© Springer-Verlag 2009

Abstract In the present report, we analyze the effect of augmenting the basis set and the level of the theory in the computation of the molecular polarization potential and it is used as a tool to understand their spatial effects in the calculation of molecular polarizabilities. This is analyzed in the present report through the results obtained for three reference molecules: hydrogen cyanide, formaldehyde and urea. Accordingly, different basis sets were used for the calculations, including: 6-31G(D), 6-31G(D,P), 6-31G(2D,2P) and 6-31++G(2D,2P). On the other hand, calculations at different levels of the theory were considered, including: Hartree–Fock, Moller–Plesset second order (MP2) and fourth order (MP4) as well as the functionals BLYP and B3LYP at the density functional theory. The results of the polarization maps reveal similar trends observed with the calculation of the different components of the polarizability tensor. However, the use of the maps permits to understand spatial effects specific for each level of calculation and each basis set.

Keywords Polarization maps · Polarizability · Polarization · Basis set effect

Dedicated to Professor Santiago Olivella on the occasion of his 65th birthday and published as part of the Olivella Festschrift Issue.

L. Roset · J. J. Perez (✉)
Department of Chemical Engineering,
Technical University of Catalonia, ETS d'Enginyeria Industrial,
Av. Diagonal, 647, 08028 Barcelona, Spain
e-mail: juan.jesus.perez@upc.edu

J. Rubio-Martinez
Department of Physical Chemistry,
The Institut de Recerca en Química Teòrica i Computacional
(IQTCUB), University of Barcelona,
Martí i Franqués 1, 08028 Barcelona, Spain

1 Introduction

Molecular polarization potential maps (MPPs) [1–6] account for the spatial distribution of deformability of the electronic cloud of a molecule and can be considered as a correction to the molecular electrostatic potential maps (MEPs) [1, 7–9]. However, whereas the latter are graphical representations of the first order (coulombic) contribution of the interaction energy between a point charge and a molecule, the former represent the second order (induced) contribution to the interaction energy of a molecule in the same conditions. The information provided by MPPs at the van der Waals surface of the molecule and longer distances complements the picture provided by MEPs [1, 6] and may reveal new features in the description of the interaction energy between two molecules.

Molecular polarization potential maps were introduced as a tool for analyzing the early stages of the drug–receptor interaction [1, 8] and since then a large number of papers have been written devoted to describe its use in studies concerning different aspects of molecular recognition [9]. The success of these maps as a predictive tool is based in two factors: on the one hand, the importance of the electrostatic contribution to the interaction energy between two molecular systems and on the other, to the fact that they provide a picture of the interactions not attainable from a multipole expansion analysis. Despite their use underscoring the most important features involving molecular recognition between two molecules at long distances, polarization contributions at short distances may become equally important [9] and these effects can be properly accounted through MPPs [1–6]. Accordingly, MPPs have been used in the past to explain specific features of molecular interactions, besides being used to generate new force field parameters [10].

Alternatively, MPPs can also be used as a tool for understanding the effects of the basis set and method used in the evaluation of molecular polarizabilities. In fact, despite being known that in order to compute reasonable estimates of molecular polarizabilities a set of well tempered basis set including polarization and diffuse functions is necessary [11], the spatial distribution due to the increase of the basis set or due to the use of a different evaluation procedure has never been analyzed. Accordingly, we report in this work the analysis carried out on three small molecules: hydrogen cyanide, formaldehyde, and urea used as benchmark in a previous study [4].

2 Methods

Molecular polarizabilities were estimated by computing the analytical derivatives of the energy as described elsewhere [12]. For the calculation of the MPPs, molecules were placed in the *YZ* plane and polarization maps were computed at $x = 4$ bohr from the molecular plane. Grid points were computed every bohr and extended approximately 6 bohr beyond the van der Waals surface of the molecules. Since the molecules selected have a symmetry plane σ_s , only one half of the points of the grid needed to be computed. The total number of points calculated was 231 points for hydrogen cyanide, 242 for formaldehyde and 312 for urea. The value of the polarization at every point of the grid was computed using the finite field method with a probe charge of 0.1 a.u. For this purpose at each point, the scaled value of the electrostatic potential (V) was subtracted from the interaction energy of the molecule with the probe charge (ΔE). From these values the polarization potential (P) was easily obtained, from the following expression:

$$\Delta E(r) = eV(r) + e^2 P(r)$$

Molecular polarization potential maps were computed with three different basis sets and using different approximations. The molecular geometry was optimized for each basis set and method used. Basis sets selected for the present study include the standard 6-31G(d) [13], 6-31G(d,p) [13], 6-31 (2d,2p) [14] basis sets with polarization functions and the 6-311++G(2d,2p) [15] that includes diffuse functions. Calculation were done at the Hartree–Fock, second order Moller–Plesset (MP2), fourth order Moller–Plesset (MP4) levels [16] and using the functional density theory, with the BLYP [17] and B3LYP [18] functional. All the calculations were performed with the Gaussian94 package [19].

3 Results and discussion

In Table 1 are listed the estimates of the components of the polarizability tensor for the three molecules studied in the present work, together with the experimental results available. As can be seen in Table 1a–c, a second set of polarization functions improves all the components of the tensor around 10%, but only for the out-of-the *YZ* plane components. In contrast, the use of diffuse functions is indispensable to obtain reliable estimates with increases of around 30% in regard to the calculations carried out with the basis sets augmented with two sets of polarization functions, although this improvement is smaller for the *YZ* plane components for formaldehyde and urea. Present results confirm also that DFT reproduces well the results obtained at the ab initio level at the MP4 level, although they are about 5% overestimated in the case of the basis set augmented with diffuse functions (Fig. 1).

Molecular polarization potential maps provide the spatial distribution of these trends. MPPs of these molecules were described in a previous report [4]. Figure 2a–c show the MPP maps of the three molecules studied in the present study at the MP4 level with the 6-311++(2d,2p) basis set. In contrast to MEPs, MPPs do not present any stationary point, but the separation between level curves inform of the regions where the molecule is more polarizable. Thus, looking at the level curves, we can see that in formaldehyde the oxygen is more polarizable in the directions of its non-bonding pair and in the path that intersects the two hydrogens. In the case of hydrogen cyanide, the map shows that the molecule exhibits a similar polarizable pattern in any direction. Finally, in the case of urea it exhibits as more polarizability paths in the regions between the carbon and the oxygen as well as the hydrogen bisecting directions. The different basis sets and methods used in the present work reproduce well the main features of these maps.

Whereas qualitative differences found among different computing schemes carried out in this work are small, quantitative differences are noticeable. We have analyzed the effect of correlation in the description of the maps and the effect of the augmentation of the basis set. Improvement on the calculations follows the same trends as found on the estimates of the polarizability tensor shown in Table 1. Figure 2 shows the correlation between the MPPs of urea computed with the basis set at the MP4 level with other levels of calculation. Compared with the Hartree–Fock and MP2 all the points of the map behave similarly, independent of their value. Moreover, it can be seen that differences with MP2 are not noticeable and with the calculations carried out at the Hartree–Fock level, differences are around 2%. Interestingly, it had been previously

Table 1 Estimates of the polarizability tensor for the different molecules reported in the present work

	6-31G(d)	6-31G(d,p)	6-31G(2d,2p)	6-311 ++ G(2d,2p)
<i>(a) Results for hydrogen cyanide, with an average experimental value of 17.5 a.u. [21]</i>				
RHF	xx 7.47	7.53	8.70	11.36
	yy 7.47	7.53	8.70	11.36
	zz 19.52	19.64	19.68	26.12
	<α> 10.48	11.56	12.36	16.28
MP2	xx 7.49	7.55	9.35	11.51
	yy 7.49	7.55	9.35	11.51
	zz 18.28	18.41	18.61	26.34
	<α> 11.09	11.17	12.44	16.45
MP4	xx 7.62	7.70	9.42	11.44
	yy 7.62	7.70	9.42	11.44
	zz 18.81	18.89	18.85	26.99
	<α> 11.34	11.43	12.63	16.62
BLYP	xx 7.73	7.79	9.52	12.22
	yy 7.73	7.79	9.52	12.22
	zz 19.22	19.36	19.41	27.01
	<α> 11.56	11.64	12.82	17.15
B3LYP	xx 7.66	7.71	9.47	11.88
	yy 7.66	7.71	9.47	11.88
	zz 19.15	19.28	19.30	26.49
	<α> 11.49	11.57	12.75	16.75
<i>(b) Results for formaldehyde with an average experimental value of 16.5 a.u. [22]</i>				
RHF	xx 6.48	6.77	8.73	10.48
	yy 12.92	13.04	14.04	15.23
	zz 17.64	17.81	17.93	21.51
	<α> 12.35	12.54	13.57	15.74
MP2	xx 6.45	6.65	8.71	10.71
	yy 13.56	13.66	14.77	15.24
	zz 17.40	17.52	17.84	21.56
	<α> 12.48	12.61	13.78	15.89
MP4	xx 7.50	6.92	8.80	10.86
	yy 13.71	14.02	14.82	16.54
	zz 17.42	17.62	17.98	22.09
	<α> 12.88	12.85	13.87	16.50
BLYP	xx 6.76	6.96	8.84	11.17
	yy 14.16	14.31	15.43	17.90
	zz 17.95	18.09	18.32	23.58
	<α> 12.96	13.12	14.20	17.55
B3LYP	xx 6.69	6.89	8.80	10.99
	yy 13.74	13.84	14.97	16.96
	zz 17.74	17.69	18.10	22.72
	<α> 12.72	12.89	13.96	16.87
<i>(c) Results for urea, with an experimental value of 31.9 a.u. [23]</i>				
RHF	xx 11.30	12.16	13.01	22.61
	yy 28.12	28.66	28.92	32.47
	zz 30.90	31.42	32.60	35.37
	<α> 23.44	24.08	24.84	30.15

Table 1 continued

	6-31G(d)	6-31G(d,p)	6-31G(2d,2p)	6-311 ++ G(2d,2p)
MP2	xx 12.06	12.31	13.69	25.14
	yy 29.11	29.52	29.62	33.45
	zz 30.84	31.24	31.82	36.72
	<α> 24.01	24.36	25.04	31.77
MP4	xx 12.16	12.42	13.70	25.17
	yy 30.24	30.67	30.91	34.50
	zz 30.82	31.82	31.85	36.98
	<α> 24.41	24.97	25.48	32.22
BLYP	xx 12.25	12.51	14.71	26.45
	yy 31.89	32.30	30.73	40.77
	zz 32.15	32.49	32.28	42.10
	<α> 25.43	25.77	25.91	36.44
B3LYP	xx 12.16	12.41	14.70	25.13
	yy 30.26	30.67	28.99	38.01
	zz 31.05	31.40	32.19	39.91
	<α> 24.49	24.83	25.29	34.35

Values in every cell correspond to the different tensor components (xx , yy , zz) followed by the average value (in bold), computed as $\langle\alpha\rangle = (\alpha_{xx} + \alpha_{yy} + \alpha_{zz})/3$. All values are in atomic units ($1e^{2a_2}E_h^{-1} = 1.648778 \times 10^{-41}C^2 m^2 J^{-1}$)

reported that the effect of correlation does not affect the MEPs considerably [20]. In contrast, in the case of the DFT calculations, it can be seen that the correlation with MP4 calculations is poor. Specifically, in the case of using the B3LYP functional, there are a series of points that are not properly computed and are distributed at different values. In contrast, the BLYP functional exhibits a better behavior with points that differ in regard to the reference calculation found at values close to zero.

Spatial distributions of differences provides information about the areas that are described differentially. For this purpose, relative difference maps of the polarization computed with two different methods or basis sets were computed. Each point of the grid was computed using the formula:

$$\varepsilon_i = \frac{\{P(\text{method A})i - P(\text{method B})i\}}{P(\text{method A})i} \times 100$$

where P is the polarization at point i of the grid computed with one of the basis sets and with a specific method. The extent of the relative deviation between the two different methods is measured by the average values of ε for all points of the grid. Figure 3a–c shows the effect of augmenting the basis set in a series of calculations carried out on urea at the MP4 level and using as reference the 6-311++(2d,2p) basis set. As can be seen in Fig. 3a, differences with the basis that includes two sets of polarization functions are around 10% in the areas close to the nuclei and increase up to a 20% in the outermost

regions of the molecule. Interestingly, when we compare the differences in the computation of the polarizability as shown in Table 1, it can be clearly concluded that there is the need of a proper description of this region for a proper calculation of some of the components of the polarizability specially the out-of-the YZ plane component. However, when we compare the reference results with the basis set that includes one set of polarization functions (Fig. 3b, c), differences are larger, but more interestingly, the area close to the nuclei is not properly computed. Indeed, the map shows positive deviations in the region close to the nuclei, like if they were overestimated and negative differences in the outermost regions of the molecule with nearly zero deviations on top of the middle of the different bonds. The same trends are observed in the map obtained with the smaller basis set.

4 Conclusions

The aim of the present report is to assess the spatial distribution of the effects observed when using different basis sets and at different levels of the theory for the computation of molecular polarizabilities. For this purpose, we computed the molecular polarization potential of three small molecules: hydrogen cyanide, formaldehyde, and urea. The results show that the inclusion of a second set of polarization functions in the basis set increases about 10% the values of the polarizabilities and this improvement is

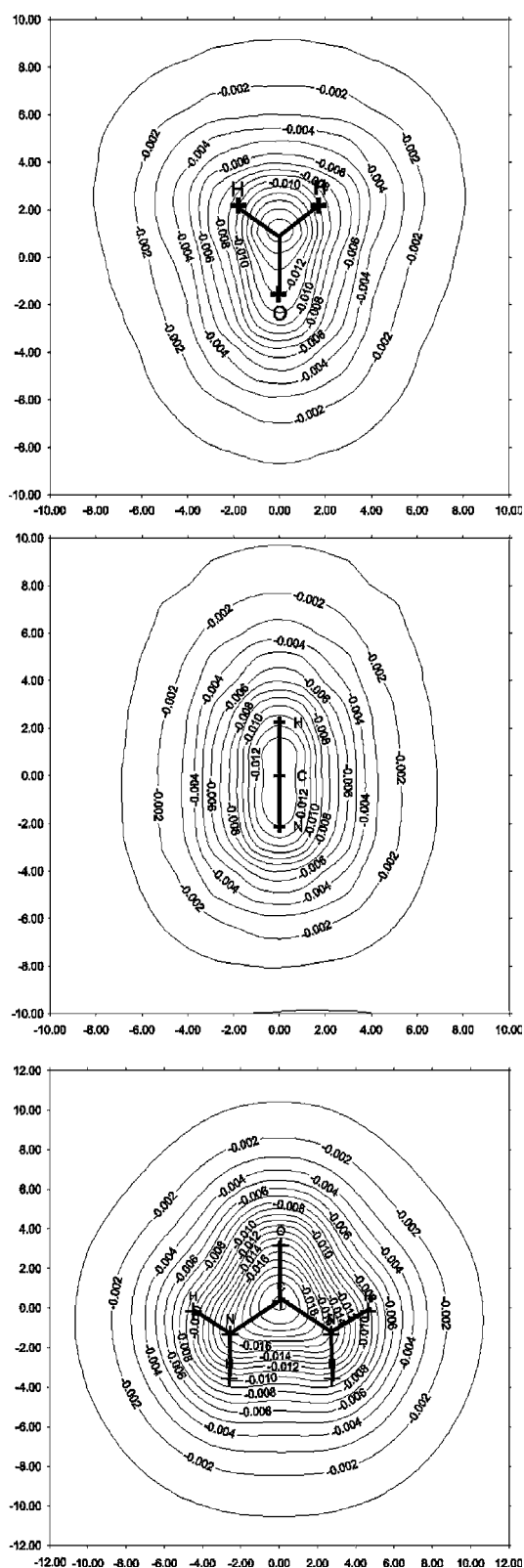


Fig. 1 Molecular polarization maps computed with a 6-311++G(2D,2P) basis set at the MP4 level of calculation. All the results are in a.u.: **a** formaldehyde, **b** hydrogen cyanide, **c** urea

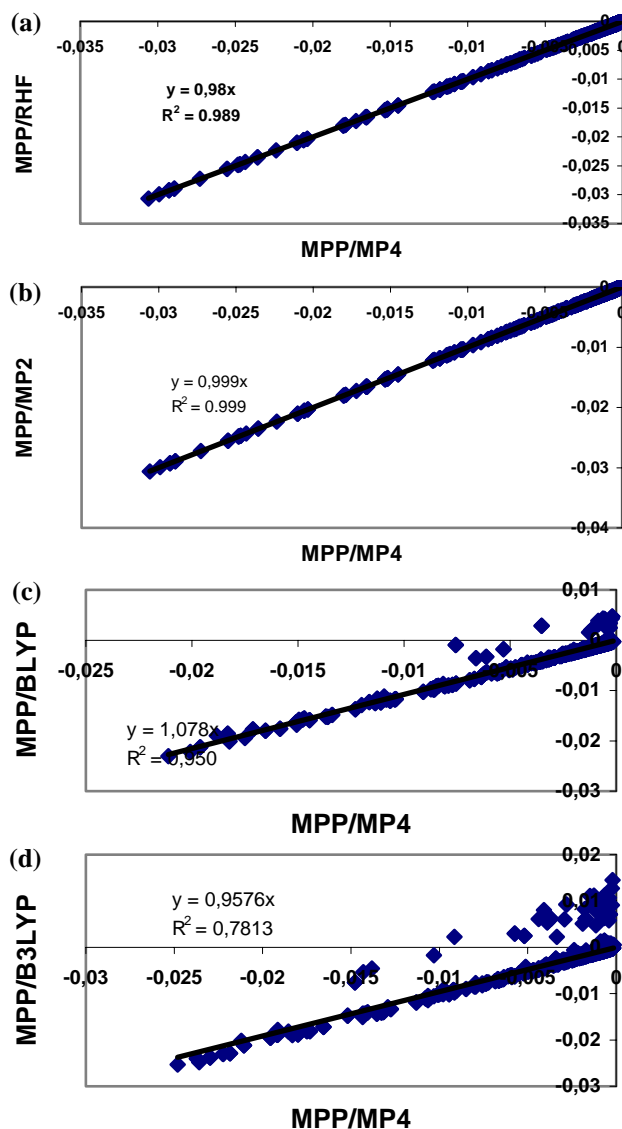


Fig. 2 Comparison of the values of the molecular polarization maps computed at different levels of theory using a 6-311++G(2D,2P) basis set and using the MP4 as reference. All the results are in a.u.: **a** Hartree–Fock, **b** MP2, **c** BLYP, **d** B3LYP

basically located in the most polarizable regions like lone pairs and bonds. This has a parallel in the evaluation of the different components of the tensor. On the other hand, basis augmentation with diffuse functions is critical for an accurate description of the less polarizable directions of the molecule. Increments of about 30% are obtained and mostly located in the outermost regions of the molecule. Inclusion of the electronic correlation does not alter much the spatial distribution of the molecular polarization maps, although there is a difference of around 2% in the values. In contrast, calculations carried out at the DFT level exhibit a poor correlation compared to MP4 calculations, with

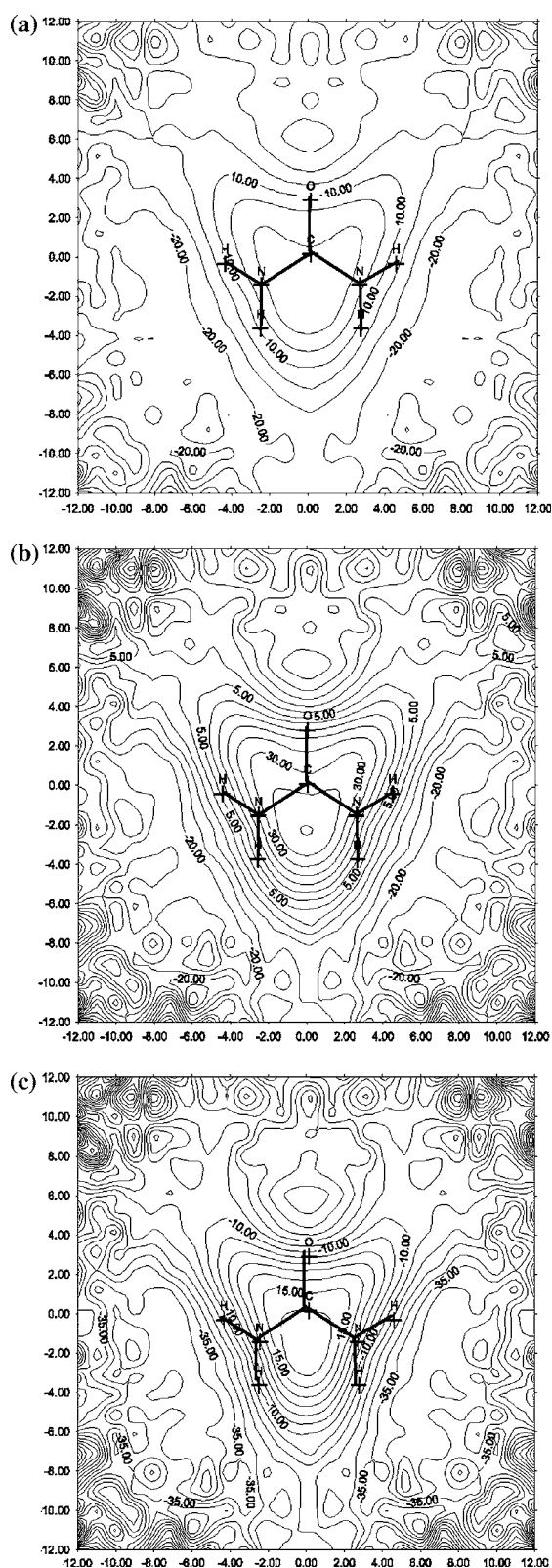


Fig. 3 Maps of the difference (in percentage) of molecular polarization maps for urea, using the 6-311++G(2D,2P) basis at the MP4 level of the theory as reference calculation: **a** comparison with 6-31G(2D,2P), **b** comparison with 6-31G(D,P), **c** comparison with 6-31G(D)

the BLYP functional exhibiting a better behavior. Interestingly, deviations are preferentially located in regions where the polarization values gets close to zero.

Acknowledgments The Barcelona Supercomputer Center is acknowledged for a generous allocation of computer time. The Spanish Ministry of Science and Technology supported this work through the projects CTQ2006-06588/BQU to J.R.M. and SAF2005-08148-C04-01 to J.J.P.

References

- Bonaccorsi R, Scrocco E, Tomasi J (1976) *Theor Chim Acta* 43:63. doi:10.1007/BF00548957
- Francl MM (1985) *J Phys Chem* 89:428. doi:10.1021/j100249a012
- Alkorta I, Perez JJ, Villar HO (1994) *J Mol Graph* 12:3. doi:10.1016/0263-7855(94)80002-2
- Alkorta I, Villar HO, Perez JJ (1993) *J Phys Chem* 97:9113. doi:10.1021/j100138a008
- Alkorta I, Perez JJ (1996) *Int J Quantum Chem* 57:123. doi:10.1002/(SICI)1097-461X(1996)57:1<123::AID-QUA14>3.0.CO;2-9
- Alkorta I, Perez JJ (1995) *Electron J Theor Chem* 1:26. doi:10.1002/ejtc.6
- Scrocco E, Tomasi J (1973) In: *Topics in current chemistry*, vol 42, Springer, Berlin
- Weinstein H (1975) *Int J Quantum Chem QBS2*:5969
- Politzer P, Murray JS (1991) In: Lipskowitz KB, Boyd DB (eds) *Reviews of computational chemistry vol 2*. VCH Publishers, New York
- Alkorta I, Bachs M, Perez JJ (1994) *Chem Phys Lett* 224:160. doi:10.1016/0009-2614(94)00512-5
- Sadlej AJ (1991) *Theor Chim Acta* 81:339. doi:10.1007/BF01118573
- Rice JE, Amos RD, Colwell SM, Handy NC, Sanz J (1990) *J Chem Phys* 93:8828. doi:10.1063/1.459221
- Hariharan PC, Pople JA (1973) *Theor Chim Acta* 28:213. doi:10.1007/BF00533485
- Frisch MJ, Pople JA, Binkley JS (1984) *J Chem Phys* 80:3265. doi:10.1063/1.447079
- Clark T, Chandrasekhar J, Spitznagel GW, Schleyer PR (1983) *J Comput Chem* 4:294. doi:10.1002/jcc.540040303
- Hinchliffe A (1996) *Modelling Molecular Structures*. Wiley and Sons, Chichester
- Becke AD (1988) *Phys Rev A* 38:3098. doi:10.1103/PhysRevA.38.3098
- Becke AD (1993) *J Chem Phys* 98:5648. doi:10.1063/1.464913
- Frisch MJ, Trucks G W, Schlegel HB, Gill P MW, Johnson BG, Robb MA, Cheeseman JR, Keith T, Peterson GA, Montgomery JA, Raghavachari K, Al-Laham MA, Zakrzewski VG, Ortiz JV, Foresman JB, Peng CY, Ayala PY, Chen W, Wong MW, Andres JL, Replogle ES, Gomperts R, Martin RL, Fox J, Binkley JS, Defrees DJ, Baker J, Stewart JJP, Head-Gordon M, Gonzales C, Pople JA (1995) *Gaussian94, Revision B 2*. Gaussian Inc, Pittsburgh PA
- Luque FJ, Orozco M, Illas F, Rubio J (1991) *J Am Chem Soc* 113:5203. doi:10.1021/ja00014a010
- McClellan AL (1963) *Tables of experimental dipole moments*. Freeman, San Francisco
- Landolt-Bornstein (1951) *Atom und Molecularphysik*, vol 1 part 3, Springer Berlin
- Appelquist J, Carl JR, Fung KK (1972) *J Am Chem Soc* 94:2953

Assessment of the performance of the functional density theory to compute molecular polarizabilities using molecular polarization maps

Lurdes Roset and Juan J. Perez*

Department of Chemical Engineering, Technical University of Catalonia ETS d'Enginyeria Industrial, Av. Diagonal, 647, 08028 Barcelona, Spain

Abstract. In the present report we analyze the performance of DFT to compute molecular polarizabilities through the analysis of molecular polarization potential that can be used as a tool to understand their spatial effects in the calculation of molecular polarizabilities. This is done in the present report through the results obtained for three reference molecules: hydrogen cyanide, formaldehyde and urea. Calculations were carried out at the density functional theory using the BLYP and B3LYP functionals. The results of the polarization maps reveal similar trends observed with the calculation of the different components of the polarizability tensor. However, the use of the maps permits to understand spatial effects specific for each level of calculation and each basis set.

1. Introduction

Dipole polarizability accounts for the linear contribution to the induced dipole of a charge distribution in the presence of an electric field [1]. The polarizability is an important magnitude involved in the description of molecular interactions with a determining role in some cases [2,3]. Since its experimental determination is not straightforward, computational methods provide a reasonable alternative to obtain estimates of this property [4]. Indeed, polarizability can routinely be calculated from the molecular electron density through either the computation of the analytical derivative of the dipole moment or computing the second derivative of the energy versus the electric field [5]. Alternatively, the Finite field theory (FFT) can also be used for this purpose; however it requires a careful choice of electric field to be in the position to disregard higher order contributions [6]. At the Hartree-Fock level this procedure is equivalent to use the perturbation theory known as the couple Hartree-Fock method (CPHF) [7].

The effect of the polarizability in molecular interactions and its distribution in space can be rationalized using the Molecular Polarization Potential (MPP) maps. They represent an interesting way to look at polarizabilities accounting for the spatial distribution of the deformability of the molecular electron density as a correction to the Molecular Electrostatic Potential (MEP) maps [8–11]. Indeed, the MEP represents the first order (coulombic) contribution of the interaction energy between a point charge and a

*Corresponding author. Tel.: +34 934016679; Fax: +34 934017150; E-mail: juan.jesus.perez@upc.edu.

molecule, whereas the MPP represents the second order (induced) contribution to the interaction energy of a molecule in the same conditions. However, MPP maps can also be used as a tool to provide a quantitative measure of how the spatial distribution of the electron density affects the calculation of the polarizability [12], providing new insights into the understanding of the effect of the extension of the basis sets and/or the deficiencies for describing the induced moments of alternative procedures to compute the electron density [12–18].

The purpose of this report is to analyze the performance of the DFT method to compute molecular polarizabilities. In this work we report the results of calculations using two different commonly used exchange functionals. DFT results reported in this work show an overestimation of polarizability estimates as had been previously published in the literature [19]. Furthermore, in order to understand the origin of such a behavior we carried out a comparative analysis of the MPP maps using the calculations computed at the MP4 level using a 6-311++(2d,2p) basis set as reference [20]. In the present work benchmark calculations have been performed on three small molecules: hydrogen cyanide, formaldehyde and urea.

2. Methods

The polarizability tensor was calculated from the second derivative of the energy on the optimized geometries within each computational method. For the computation of the MPP maps, molecules were placed on the XZ plane and the maps computed on the plane $y = 4$ bohr above the molecular plane. Grid points were computed every bohr and extended approximately 6 bohr beyond the van der Waals molecular surface. Since the molecules selected have a symmetry plane only one half of the points of the grid were computed. The total number of points calculated was 231 points for hydrogen cyanide, 242 for formaldehyde and 312 for urea.

The polarization was computed at each point of the grid by subtracting the coulomb contribution to the interaction energy (ΔE) computed using a probe charge of 0.1 a.u. The coulomb contribution (eV) was computed from the corresponding molecular electrostatic potential map. The polarization potential (P) can easily be obtained from the following expression:

$$\Delta E(r) = eV(r) + e^2P(r)$$

Calculations at the DFT level were carried out with the BLYP [21] and B3LYP [22] functionals, using a 6-311++(2d,2p) [20] basis set that includes two sets of polarization functions as well as diffuse functions. Calculations were also carried out at the SCF and MP2 and MP4 level with the same basis set to be used as reference. All the calculations reported in this work were performed with the Gaussian94 package [23].

3. Results and discussion

Estimates of the polarizability tensor for the three molecules studied in the present work are listed in Table 1 together with the available experimental results. Inspection of the Table shows that whereas the MP4 improves systematically both, SCF and MP2 results by less than a 5%, the DFT method tends to overestimate polarizability calculations beyond MP4 results. However, the two functionals used in the present work exhibit a differential behavior, with the B3LYP providing closer values to MP4 estimates.

Table 1

Estimates of the polarizability tensor molecules reported in the present work computed with the 6-311++(2d,2p) basis set. Values correspond to the different tensor components (α_{xx} , α_{yy} , α_{zz}) followed by the average value $\langle \alpha \rangle$ (in bold), computed as $\langle \alpha \rangle = (\alpha_{xx} + \alpha_{yy} + \alpha_{zz})/3$. All values are in atomic units ($1e^2 a_0^2 E_h^{-1} = 1.648778 \times 10^{-41} \text{C}^2 \cdot \text{m}^2 \cdot \text{J}^{-1}$). Experimental values available are also listed

		RHF	MP2	MP4	BLYP	B3LYP	Exper. value $\langle \alpha \rangle$
Hydrogen cyanide	xx	11.36	11.51	11.44	12.22	11.88	
	yy	11.36	11.51	11.44	12.22	11.88	
	zz	26.12	26.34	26.99	27.01	26.49	
	$\langle \alpha \rangle$	16.28	16.45	16.62	17.15	16.75	17.5
Formaldehyde	xx	15.23	15.24	16.54	17.90	16.96	
	yy	10.48	10.71	10.86	11.17	10.99	
	zz	21.51	21.56	22.09	23.58	22.72	
	$\langle \alpha \rangle$	15.74	15.89	16.50	17.55	16.87	16.5
Urea	xx	32.47	33.45	34.50	40.77	38.01	
	yy	22.61	25.14	25.17	26.45	25.13	
	zz	35.37	36.72	36.98	42.10	39.91	
	$\langle \alpha \rangle$	30.15	31.77	32.22	36.44	34.35	31.9

Specifically, for formaldehyde and hydrogen cyanide the results are very similar to MP4 estimates, whereas for the case of urea estimates are about 5% above. In contrast, the BLYP exhibits a poorer behavior with deviations of about 5% for formaldehyde and hydrogen cyanide and even a 10% for urea.

Some insight of the performance of the DFT method can be obtained from the analysis of the MPP maps, however, more interesting the MPP themselves are the maps of the deviations of the polarization from a reference that can provide information about systematic differences between the different methods of calculation. Figures 1–3 show the polarization difference maps (in percentage), with the MP4 calculation being taken as reference. As can be seen, difference maps of the Hartree-Fock and MP2 calculations are smooth; the improvement is continuous, whereas the maps of the DFT calculations are rough with values underestimated in the neighborhood of the atoms and values overestimated in the outer-most regions of the molecule.

In the case of the hydrogen cyanide the largest differences between the Hartree-Fock calculations and the MP4 are around the nitrogen lone pair (Fig. 1a), with underestimations around 15%, whereas for the rest of the molecule these difference goes down to a 4%. Clearly the MP2 calculation is closer to the reference MP4 calculations (Fig. 1b), with deviations of around a 2% in the neighborhood of the nitrogen lone pair. This behavior clearly reflects the continuous improvement in the evaluation of the polarizability when the correlation energy is included. In contrast, in the case of the DFT calculations difference maps are uneven with zones where the polarization is overestimated and regions where it is underestimated (Figs 1c and 1d). The two functionals exhibit a similar behavior with an underestimation of the polarization in the neighborhood of the nuclei with peaks that reach a 10% around the nitrogen lone pair and an overestimation of the polarization of around a 4% in the outer-most regions of the molecule in the case of the BLYP functional and around a 8% for the B3LYP functional. Thus, although both maps show an uneven topography the B3LYP surface exhibits larger differences between lows and highs.

Similar results are obtained after the analysis performed to the formaldehyde molecule. The Hartree-Fock procedure underestimates the polarization in regard to the MP4 calculations up to 25% in the areas close to the oxygen lone pairs (Fig. 2a). This effect is reduced with differences up to 4% when the correlation energy is included at the MP2 level (Fig. 2b). In regard to the DFT calculations, the

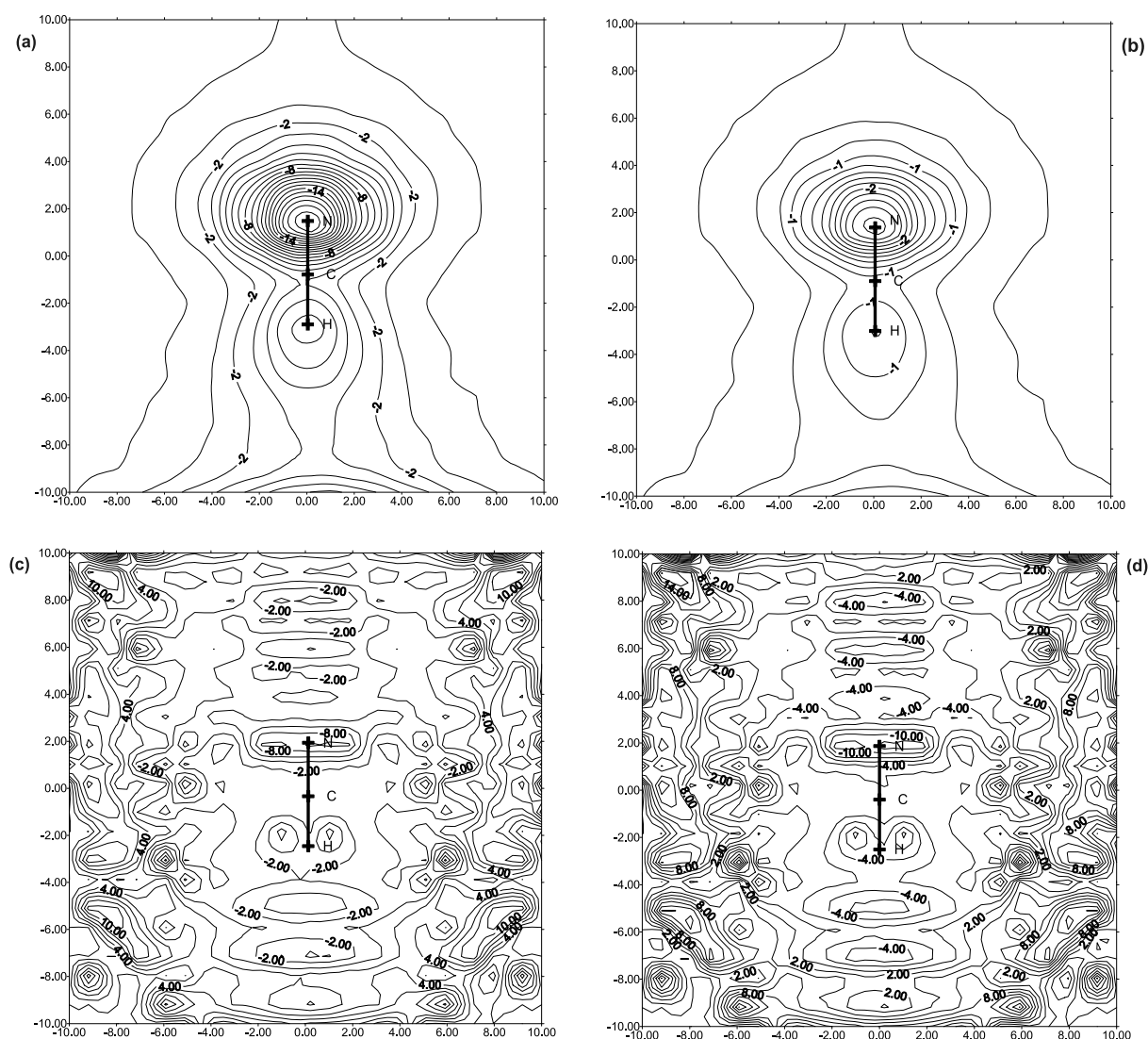


Fig. 1. Molecular polarization difference maps (in percentage) for HCN at different levels of the theory using the MP4 calculation as reference. (a) RHF/MP4 (b) MP2/MP4 (c) BLYP/MP4 (d) B3LYP/MP4.

two functionals overestimate the polarizability, although values computed with the B3LYP are closer to those obtained with the MP4 method. Difference maps reveal an uneven distribution of the polarization with underestimations in regions close to the nuclei and overestimations in the outer-most regions of the molecule (Figs 2c and 2d).

In regard to urea, Figs 3a–3d show the polarization difference maps taking the MP4 calculation as reference. The same trends as described above can be observed: whereas the SCF and MP2 maps are smooth, the DFT maps are uneven. Thus, the SCF calculation underestimates the polarization about a 25% (Fig. 3a) whereas the MP2 calculation only around 2% (Fig. 3b). On other hand, DFT calculations exhibit underestimations up to 50% close to the nuclei (Figs 3c and 3d) and overestimations of around 20% in the outer-most regions.

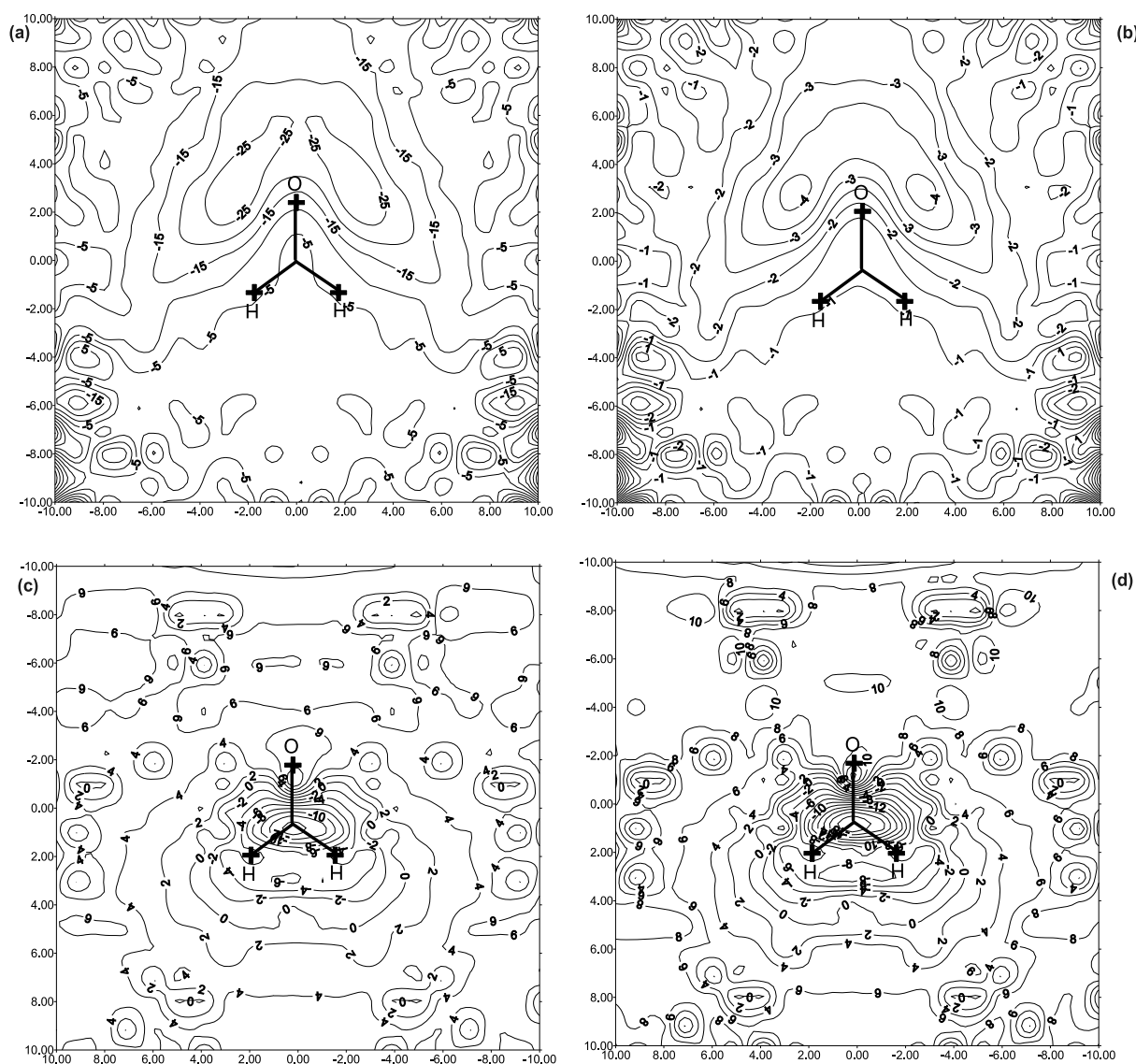


Fig. 2. Molecular polarization difference maps (in percentage) for formaldehyde at different levels of the theory, using the MP4 calculation as reference. (a) RHF/MP4 (b) MP2/MP4 (c) BLYP/MP4 (d) B3LYP/MP4.

A good description of the electronic density in the outer-most regions of the molecules is critical for good values of the polarizability. This is supported by the need to use extended basis sets including diffused functions to get good estimates [12]. Accordingly, comparison of the polarization maps may suggest that the overestimation of the polarizability obtained when using the DFT method in contrast to the Hartree-Fock and MP procedures must be due to the overestimation of the polarization in these regions. This effect is more remarkable in the perpendicular components.

Figures 4–6 show the correlations between the values of the polarization computed with different methods and the MP4. As expected, the correlation between points computed at the SCF, MP2 versus the MP4 method is good, being the fit better for points of lower polarization. In contrast, the correlation

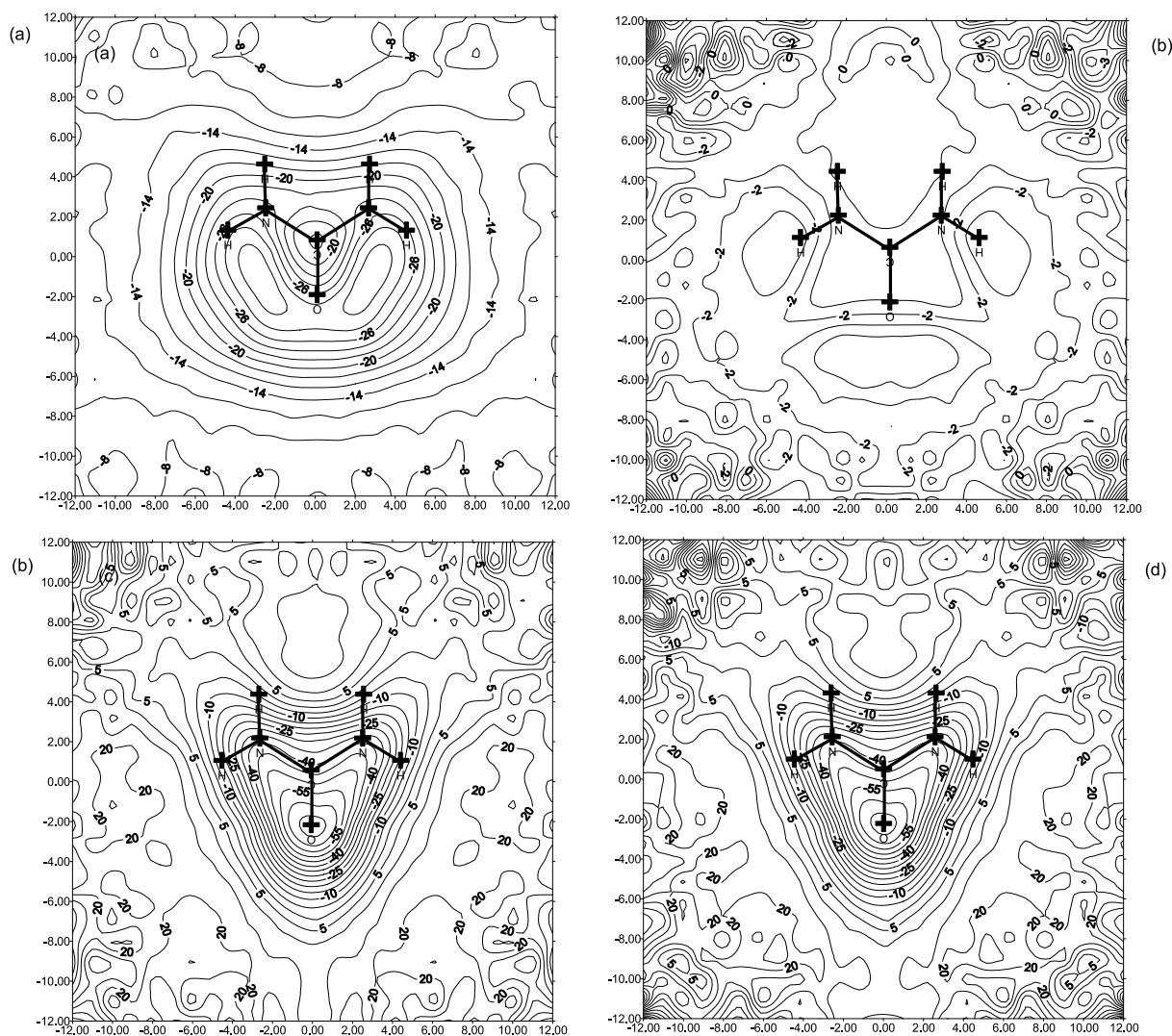


Fig. 3. Molecular polarization difference maps (in percentage) for urea at different levels of the theory using the MP4 calculation as reference. (a) RHF/MP4 (b) MP2/MP4 (c) BLYP/MP4 (d) B3LYP/MP4.

is worse in the DFT calculations. As can be seen there are points poorly computed, even some of them have the sign changed. However, most of these outsiders correspond to small values of polarization and are due to the lack of accuracy of the density function.

4. Conclusions

The aim of the present report is to assess the spatial distribution of the effects observed when using different basis sets and at different levels of the theory for the computation of molecular polarizabilities. For this purpose, we computed the molecular polarization potential of three small molecules: hydrogen cyanide, formaldehyde and urea. The results show that the DFT method overestimates the values of the polarization in the outer-most regions of the molecule, being this the reason for the overestimation of

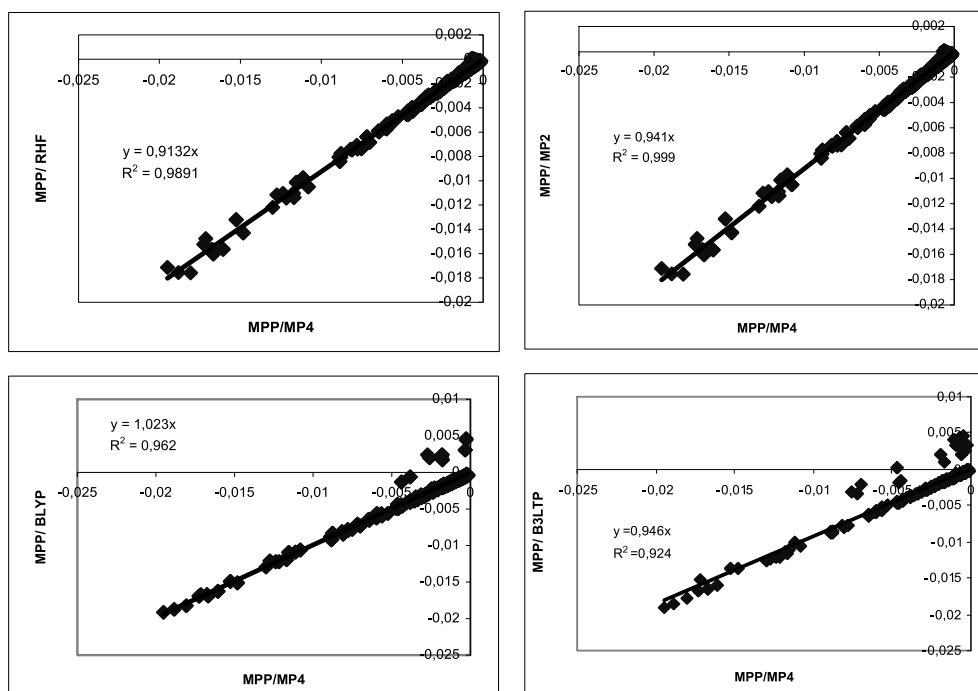


Fig. 4. Molecular polarization maps correlation plots of HCN computed at different levels using the MP4 level as reference. All the results are in a.u.

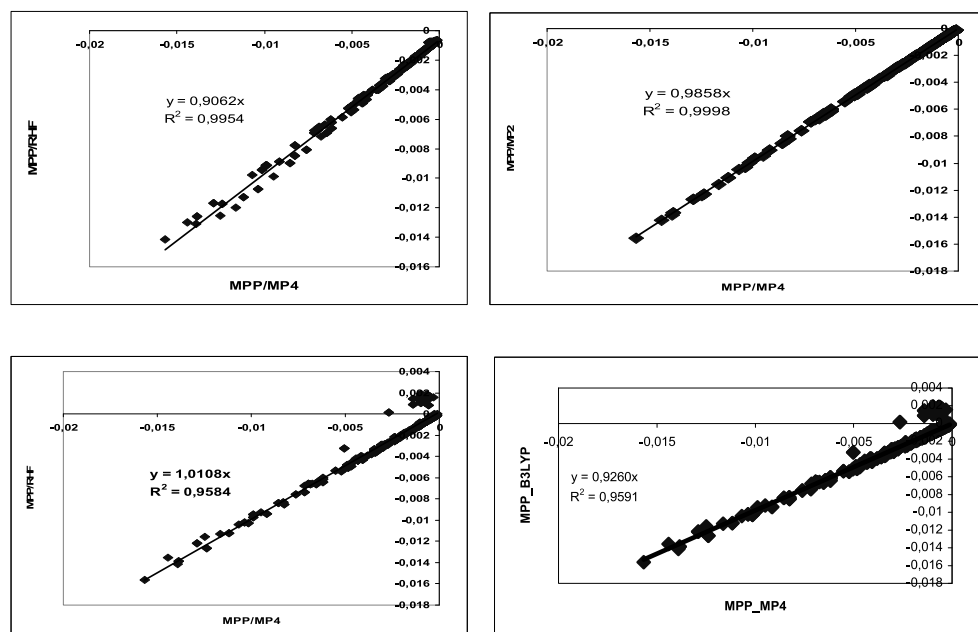


Fig. 5. Molecular polarization maps correlation plots of formaldehyde computed at different levels using the MP4 level as reference. All the results are in a.u.

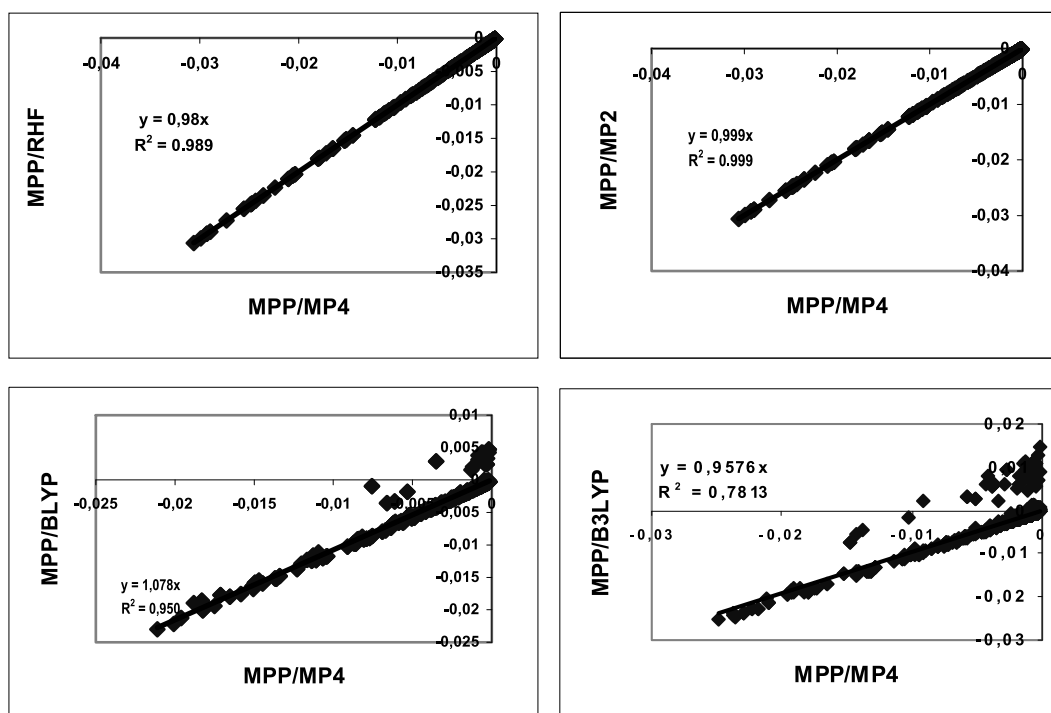


Fig. 6. Molecular polarization maps correlation plots of urea computed at different levels using the MP4 level as reference. All the results are in a.u.

polarizability estimates. Moreover, calculations carried out at the DFT level exhibit a poor correlation compared to MP4 calculations, with the BLYP functional exhibiting a better behavior. Interestingly, points outside of the correlation correspond to small values of the polarization and even there are some values with the sign changed suggesting that the accuracy of the density function is small for the electric field used in the calculations.

Acknowledgement

The Barcelona Supercomputer Center is acknowledged for a generous allocation of computer time. The Spanish Ministry of Science of Science and innovation supported this work through the project SAF2008-04943-C02-01.

References

- [1] A.D. Buckingham, *Quart Rev* **13** (1959), 183.
- [2] C. Hansch, W.E. Steinmetz, A.J. Leo, S.B. Mekapati, A. Kurup and D. Hoekman, *J Chem Inf Comput Sci* **43** (2003), 120.
- [3] A. Grossfield, P. Ren and J.W. Ponder, *J Am Chem Soc* **125** (2003), 15671.
- [4] A. Chelkowski, *Dielectric Physics*, Studies in physical and theoretical chemistry, Elsevier, Amsterdam, The Netherlands.
- [5] J.E. Rice, R.D. Amos, S.M. Colwell, N.C. Handy and J. Sanz, *J Chem Phys* **93** (1990), 8828.
- [6] H.D. Cohen and C.C.J. Roothaan, *J Chem Phys* **43** (1965), S34.
- [7] H. Hinchliffe and D.G. Bounds, in: *Theoretical Chemistry, A Specialists Periodical Report*, R.N. Dixon and C. Thomson, eds, The Chemical Society. Bartholomeus Press, Dorking, United Kingdom, 1978.

- [8] E. Scrocco and J. Tomasi, in: *Topics in Current Chemistry*, (Vol 42), Springer-Verlag, Berlin, 1973.
- [9] H. Weinstein, *Int J Quantum Chem QBS2* (1975), 5969.
- [10] P. Politzer and J.S. Murray, in: *Reviews of Computational Chemistry*, (Vol. 2), K.B. Lipskowitz, D.B. Boyd, eds, VCH Publishers, New York, 1991.
- [11] A. Hinchliffe, *Modelling Molecular Structures*, John Wiley and Sons Ltd, Chichester, 1996.
- [12] L. Roset, J. Rubio-Martinez and J.J. Perez, *Theor Chem Acc* (2009).
- [13] R. Bonaccorsi, E. Scrocco and J. Tomasi, *Theor Chim Acta* **43** (1976), 63.
- [14] M.M. Francl, *J Phys Chem* **89** (1985), 428.
- [15] I. Alkorta, J.J. Perez and H.O. Villar, *J Mol Graph* **12** (1994), 3.
- [16] I. Alkorta, H.O. Villar and J.J. Perez, *J Phys Chem* **97** (1993), 9113.
- [17] I. Alkorta and J.J. Perez, *Int J Quantum Chem* **57** (1996), 123.
- [18] I. Alkorta and J.J. Perez, *Elec J Theor Chem* **1** (1995), 26.
- [19] A. Hinchliffe, J.J. Perez and H.J. Soscun Machado, *Elec J Theor Chem* **2** (1997), 325.
- [20] T. Clark, J. Chandrasekhar, G.W. Spitznagel and P.R. Schleyer, *J Comp Chem* **4** (1983), 294.
- [21] A.D. Becke, *Phys Rev A* **38** (1988), 3098.
- [22] A.D. Becke, *J Chem Phys* **98** (1993), 5648.
- [23] M.J. Frisch, G.W. Trucks, H.B. Schlegel, P.M.W. Gill, B.G. Johnson, M.A. Robb, J.R. Cheeseman, T. Keith, G.A. Peterson, J.A. Montgomery, K. Raghavachari, M.A. Al-Laham, V.G. Zakrzewski, J.V. Ortiz, J.B. Foresman, C.Y. Peng, P.Y. Ayala, W. Chen, M.W. Wong, J.L. Andres, E.S. Replogle, R. Gomperts, R.L. Martin, J. Fox, J.S. Binkley, D.J. Defrees, J. Baker, J.J.P. Stewart, M. Head-Gordon, C. Gonzales and J.A. Pople, Gaussian94, Revision B 2; Gaussian Inc Pittsburgh PA, 1995.
- [24] A.L. McClellan, *Tables of Experimental Dipole Moments*, Freeman. San Francisco, 1963.
- [25] Landolt-Bornstein, *Atom und Molecularphysik*, Springer Verlag, Berlin, Vol1 part 3, 1951.
- [26] J. Applequist, J.R. Carl and K.K. Fung, *J Am Chem Soc* **94** (1972), 2953.

Actividad biológica y toxicidad de dioxinas y furanos. Relación Estructura Química-Actividad Biológica.

Lurdes Roset ⁽¹⁾, Montserrat Ruiz ⁽¹⁾, Patricia Gómez⁽²⁾ y Juan Jesús Pérez ⁽²⁾

⁽¹⁾Departament d'Enginyeria Química, UPC, EPSEVG, Avda. Víctor Balaguer s/n, 08800 Vilanova i la Geltrú (Barcelona).
⁽²⁾Departament d'Enginyeria Química, UPC, ETSEIB, Diagonal 647, 08028 Barcelona. E-mail: m.lurdes.rosat@upc.edu



Introducción

La importancia medioambiental del estudio de la actividad biológica y toxicidad de los compuestos aromáticos polihalogenados (PHAs), como las dibenzo-p-dioxinas (PCDDs o dioxinas) y los dibenzofuranos (PCDFs o furanos) ha sido descrita en numerosas ocasiones (Figura 1). La toxicidad de estos compuestos se mide a través del receptor AHR (receptor de hidrocarburos de arilo, Hestermann-2000) de forma que la toxicidad está relacionada directamente con la afinidad de los compuestos, siendo la toxicidad muy sensible al número y distribución de los halógenos en la molécula. En el presente trabajo se analizan las características de reconocimiento molecular de los compuestos bioactivos mencionados a través de cálculos de mapas de potencial electrostático (MEP), de potencial de polarización (MPP) y de potencial de interacción (MEI) (Pérez, J.J.-1994). Se realizan los cálculos para los conjuntos de dioxinas y furanos descritos, a fin de obtener una descripción teórica de la actividad biológica de los mismos (Mhin-2002). Por último se realiza un estudio de la predicción de la actividad biológica de las dioxinas y furanos mediante cálculos PLS (Partial Least Square) utilizando los cálculos de potencial electrostático, de polarización y de interacción. Estos cálculos permiten obtener un coeficiente de correlación que nos validará el estudio realizado.

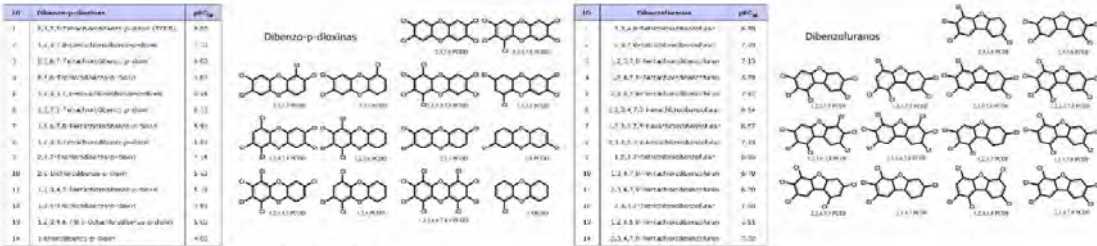


Figura 1. Conjunto de dioxinas y furanos. Toxicidad experimental (pEC₅₀).

Métodos

El procedimiento implica en primer lugar la optimización de las estructuras moleculares de todas las moléculas a estudio. Las estructuras se optimizan a nivel Hartree-Fock utilizando la base 6-31G de Dunning con inclusión de funciones polarización, utilizando el programa Gaussian 90. Los cálculos correspondientes al potencial electrostático se realizan a partir de las geometrías ya optimizadas sobre un plano a 4 Å sobre el plano molecular, y los cálculos de polarización utilizando el programa Mephisto. Los mapas de potencial electrostático, de polarización y de interacción se obtienen con el programa gráfico Surfer 6.0.

Resultados

- Se presentan los mapas de potencial electrostático, de polarización y de interacción para la dioxina 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD), ya que es la que presenta mayor actividad biológica (pEC₅₀) y se conocen numerosos estudios experimentales sobre la misma (Procopio-2002), (Figura 2).
- Se han realizado los cálculos para todo el conjunto de dioxinas y furanos descritos, y se realizan los cálculos PLS para determinar la validez de nuestro método, así como poder predecir la actividad biológica de otros compuestos de las familias mencionadas.
- Los mapas de potencial presentan mínimos cerca de los cloros y de los oxígenos, mucho mejor definidos en el mapa de interacción, debido a la importancia de la inclusión de la polarizabilidad en estos mapas.
- Se obtiene una excelente correlación entre la actividad biológica experimental (pEC₅₀ exp) y la actividad biológica calculada (pEC₅₀ calc) (Figura 3), de manera que podemos considerar el método de trabajo válido para determinar los puntos de interacción molecular de compuestos bioactivos, a partir de los mapas de potencial, y predecir actividades biológicas en función de la densidad electrónica y su distribución espacial.

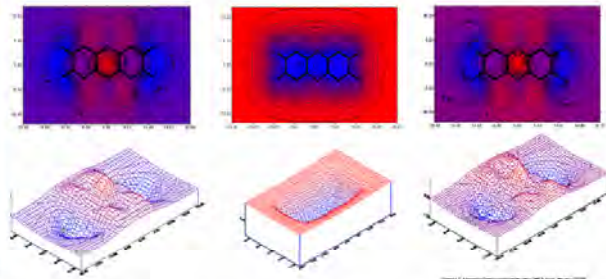


Figura 2. Mapas de potencial electrostático, de polarización y de interacción de la dioxina TCDD

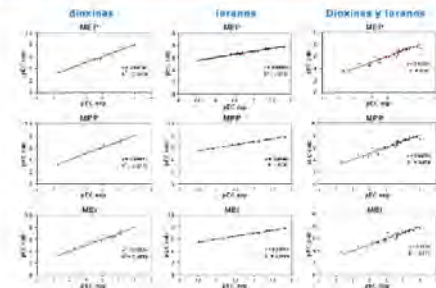


Figura 3. Cálculos PLS. Correlación actividad biológica experimental y calculada

Bibliografía

- Pérez, J. J.-1994 Pérez, J.J., Alkorta, I. *Molecular-polarization maps as tool for studies of intermolecular interactions and chemical-reactivity*. Journal of Molecular Graphics 12 (1), 3-13 (1994)
- Hestermann-2000 Hestermann, E., Stegeman, J., Hahn, M. *Relative contributions of Affinity and Intrinsic Efficacy to Aryl Hydrocarbon Receptor Ligand Potency*. Toxicology and Applied Pharmacology 168, 160-172 (2000).
- Mhin-2002: Mhin, B.J., J.E. Lee, and W.Y. Cho. *Understanding the congener-specific toxicity in polychlorinated dibenzo-p-dioxins: Chlorination pattern and molecular quadrupole moment*. Journal of the American Chemical Society, 2002, 124(1): p. 144-148.
- Procopio-2002: Procopio, M., et al., *A model for recognition of polychlorinated dibenzo-p-dioxins by the aryl hydrocarbon receptor*. European Journal of Biochemistry, 2002, 269(1), p. 13-18.

Conclusiones

- El estudio realizado con el conjunto de dioxinas y furanos descrito nos permite reconocer a nivel molecular la densidad electrónica y su distribución espacial, determinantes de las interacciones intermoleculares o de varios sistemas moleculares.
- Se ha llevado a cabo un estudio de la predicción de la actividad biológica de dioxinas y furanos mediante cálculos PLS en función del estudio de los mapas de potencial electrostático, de polarización y de interacción previamente calculados, obteniendo una buena correlación entre nuestros resultados y datos de toxicidad de que se dispone.
- Con la metodología descrita podremos realizar la predicción de toxicidad en compuestos bioactivos de interés farmacológico e industrial.