



UNIVERSITAT<sup>DE</sup>  
BARCELONA

# Aplicació de la metodologia QSPR al càlcul de propietats de compostos inorgànics i de sistemes multicomponents

Jesús Jover Modrego



Aquesta tesi doctoral està subjecta a la llicència **Reconeixement 4.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento 4.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution 4.0. Spain License.**

TD Jov



**Departament de Química Inorgànica**

**Aplicació de la metodologia QSPR al càlcul  
de propietats de compostos inorgànics i de  
sistemes multicomponents**

**Jesús Jover Modrego**

**Tesi Doctoral**

BIBLIOTECA DE LA UNIVERSITAT DE BARCELONA



0700561294

A large, stylized handwritten signature in black ink is positioned at the bottom right of the page, overlapping the library information.

## **Capítol 4. Sistemes multicomponents**

Com ja s'ha indicat, la metodologia QSPR ha estat àmpliament utilitzada per calcular un elevat nombre de propietats químiques, sobretot de compostos orgànics. Normalment, els sistemes analitzats són constituïts per conjunts de molècules d'estructura semblant i el tipus de la propietat i/o les condicions en què es determina fa que el seu valor depengui únicament de la naturalesa dels compostos estudiats. En una revisió recent,<sup>1</sup> ja es comenta la necessitat d'aplicar la metodologia QSPR a situacions més complicades on la propietat depengui a la vegada de dos o més components del sistema, com, per exemple, la densitat d'una mescla de líquids o la solubilitat d'una substància en diferents solvents. Aquests sistemes complexos on la propietat depèn a la vegada de més d'un factor s'anomenen, de manera general, sistemes multicomponents.

L'aplicació de la metodologia QSPR a aquest tipus de sistemes ha estat, fins ara, molt escassa. Els pocs treballs trobats en la bibliografia referits a aquests sistemes multicomponents corresponen a estudis de propietats de compostos orgànics on l'únic component addicional sol ser la temperatura, d'aquesta manera s'han calculat: índexs de refracció, constants dielèctriques, moments dipolars i viscositats.<sup>2-4</sup> D'altra banda s'ha trobat un treball on s'analitza la cinètica d'hidròlisi àcida d'èsters carboxílics en diferents mescles de solvents incloent-hi, també, variacions en la temperatura.<sup>5</sup>

En aquesta Tesi s'ha utilitzat la metodologia QSPR de manera innovadora en el tractament de sistemes multicomponents. Per tal d'estudiar propietats d'aquest tipus de sistemes s'han construït models que contenen alhora descriptors dels dos components. D'aquesta manera s'obtenen models que calculen aquestes propietats i que, a més, permeten relacionar la propietat estudiada amb les característiques de cadascun dels elements que el formen i de les interaccions entre ells. Les propietats estudiades en aquest capítol són l'afinitat i la basicitat dels 20  $\alpha$ -aminoàcids habituals front alguns cations metàl·lics en fase gasosa i el  $pK_a$  de diversos àcids orgànics en diferents solvents.



Els resultats obtinguts de la basicitat front el catió liti de les molècules orgàniques estudiades van fer pensar en aplicar la metodologia QSPR a un conjunt de compostos semblant però alhora més complex. Donat que entre els compostos analitzats hi havia alguns aminoàcids es va considerar interessant estudiar les interaccions dels aminoàcids habituals amb cations monovalents. La cerca bibliogràfica va demostrar que no existeix cap conjunt de valors experimentals coherent de les afinitats i basicitats catióniques, i que sovint es troben diferències significatives en els valors proposats per diferents autors. En la bibliografia es troba, també, un gran nombre de valors calculats per diferents aproximacions, la major part d'ells amb mètodes de càlcul rigorosos com els DFT, d'un bon grapat d'entalpies i energies lliures de reacció entre els aminoàcids i alguns cations. Per tal de poder aplicar els mètodes QSPR a un conjunt coherent i prou gran de dades, s'ha realitzat un estudi teòric detallat i sistemàtic d'aquestes dues propietats termodinàmiques dels 20  $\alpha$ -aminoàcids habituals amb els cations  $\text{Li}^+$ ,  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Cu}^+$  i  $\text{Ag}^+$ . Aquests càlculs s'han fet amb càlculs a nivell DFT utilitzant el funcional B3LYP combinat amb una base DZVP. Els valors d'afinitat i basicitat així calculats s'han pres com a valors de referència i s'hi ha aplicat la metodologia QSPR per tal de trobar un model que en permeti la seva estimació. En el tractament per construir el model s'han utilitzat alguns descriptors externs característics dels cations metàl·lics com radis, relacions càrrega/radi, potencials d'ionització, afinitats electròniques, diferents escales d'electronegativitat, etc.

L'altra propietat estudiada de sistemes multicomponents és el  $\text{pK}_a$  d'algunes famílies de compostos orgànics en diferents solvents. Les famílies estudiades són fenols, àcids benzoics, àcids carboxílics alifàtics i anilines, compostos adients per ser estudiats amb aquesta metodologia ja que es disposa de suficients valors experimentals d'un elevat nombre de derivats. Aquests compostos presenten una diversitat química gran a causa dels diferents substituents que tenen i de les posicions de substitució disponibles. Els solvents considerats en aquest estudi són de naturalesa variada i es poden classificar en pròtics, com ara aigua, metanol, etanol, i apròtics: dimetilsulfòxid, acetonitril, nitrometà, acetona, entre d'altres. Els valors dels descriptors externs utilitzats per caracteritzar els solvents s'han extret de la bibliografia i corresponen, per un costat, a diferents propietats físiques, com la constant dielèctrica, la polaritzabilitat, el moment dipolar, l'entalpia de vaporització; i, per l'altra, als paràmetres de les escales

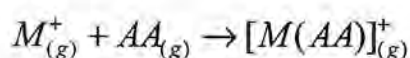
empíriques de polaritat dels solvents més populars com les de Kamlet i Taft, Reichardt, Koppel i Palm, Gutmann, Drago.

La naturalesa de les interaccions entre els components d'aquests sistemes i, en conseqüència, les veritables raons responsables de les propietats analitzades, són d'una gran complexitat, fet que obliga, pràcticament, a deduir els models corresponents només per mètodes no lineals. No sembla que una simple relació lineal pugui incloure, de manera general, les complicades relacions solut/solvent o catió/aminoàcid, estudiades.

Com a norma general, s'han emprat algorismes genètics implementats en ADAPT en la selecció dels descriptors i s'han utilitzat xarxes neuronals de tipus BPNN i *Support Vector Machines* en la deducció dels corresponents models.

### 4.1. Afinitat i basicitat catióniques dels aminoàcids habituals

La formació de complexos entre els 20  $\alpha$ -aminoàcids habituals i els cations metàl·lics és un procés bioinorgànic de gran importància; de fet, tots els metal·loenzims coneguts presenten, com a mínim, un dels seus aminoàcids coordinat al catió. Tal i com s'ha descrit anteriorment amb el catió liti, la basicitat i l'afinitat d'un aminoàcid davant d'un catió metàl·lic qualsevol, es defineixen com el valor negatiu de l'energia lliure de Gibbs i de l'entalpia de la reacció de la següent equació, a 25°C,<sup>6-8</sup> respectivament.



La Figura 4.1 conté l'estructura dels 20 aminoàcids habituals, els seus noms i les abreviatures més corrents que s'utilitzen per anomenar-los.

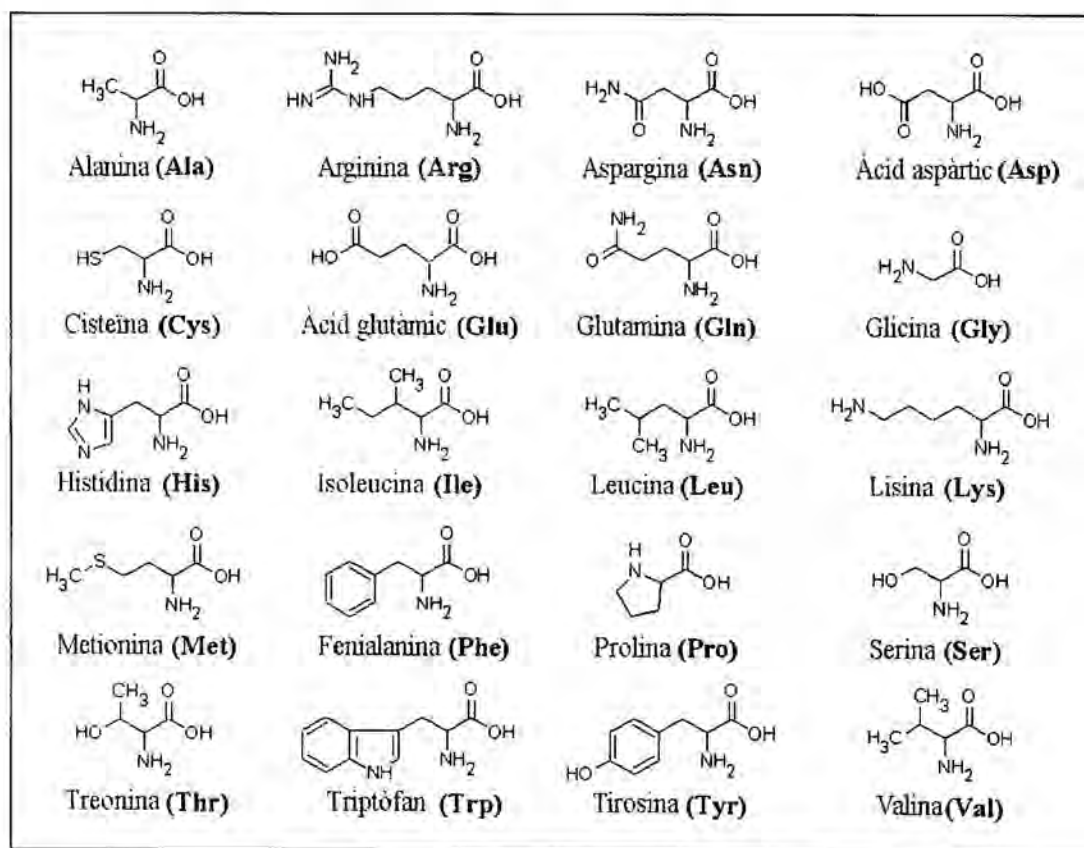


Figura 4.1. Estructura dels 20  $\alpha$ -aminoàcids habituals.

Donada la importància d'aquestes dues propietats, s'han desenvolupat diversos mètodes per determinar-les experimentalment.<sup>9-13</sup> El mètode més emprat és el mètode cinètic de Cooks,<sup>14,15</sup> que es basa en la dissociació unimolecular competitiva de dímers iònics de fórmula general  $[L_{ref}XL_x]^+$ . Aquest mètode només permet obtenir valors relatius entre els dímers estudiats, així doncs, per tal d'obtenir els valors absoluts cal emprar un valor "àncora". Una de les bases d'aquest mètode és que l'increment d'entropia de les dissociacions competitives es manté constant en tota la sèrie de compostos analitzats, cosa que, com es veurà més endavant, no sempre es compleix. D'altra banda, els grans avenços que s'han fet en el desenvolupament del maquinari i del programari permeten obtenir resultats cada cop més precisos per a aquests sistemes a partir de càlculs teòrics. De fet, alguns dels valors "àncora" han estat determinats a partir de càlculs teòrics, d'aquesta manera s'han pogut establir els valors absoluts de diverses sèries de compostos. La majoria dels treballs on es realitzen càlculs només tracten la glicina i en alguns casos l'alanina; tot i així, s'han trobat alguns articles on s'ha utilitzat un nombre més gran d'aminoàcids amb els cations sodi,<sup>16</sup> potassi<sup>17</sup> i plata,<sup>18</sup> encara que només en l'últim cas s'han fet els càlculs amb els vint aminoàcids habituals. Pel que fa als metalls estudiats, el major nombre de càlculs han estat realitzats amb els metalls alcalins: liti, sodi i potassi; i els metalls de transició coure i plata, tots ells en estat d'oxidació (I). Tot i així, es poden trobar altres treballs amb diversos elements en diferents estats d'oxidació, com per exemple cesi(I),<sup>19</sup> calci(II),<sup>20</sup> magnesi(II)<sup>20-24</sup> i coure(II)<sup>25</sup> i alumini(III).<sup>21-24</sup> És difícil establir una comparació clara entre els resultats que s'obtenen amb els diferents treballs per als mateixos sistemes degut a les diferències en el nivell de càlcul emprat: Hartree-Fock, DFT amb diferents funcionals i mètodes MP*n*; i del tractament donat a l'error de superposició de base en cadascun d'ells.

Els objectius d'aquest capítol són dos; el primer consistirà en construir, a partir de càlculs teòrics, una base de dades que contingui els valors de basicitat i afinitat en fase gas dels 20  $\alpha$ -aminoàcids habituals amb els cations monovalents de liti, sodi, potassi, coure i plata. La metodologia emprada en aquest procediment ha de permetre calcular els valors d'afinitat i basicitat per als sistemes estudiats però, a més, ha de servir com a base per estudiar, en un futur, sistemes semblants d'aminoàcids amb altres metalls en diferents estats d'oxidació. El segon objectiu consisteix en utilitzar la metodologia QSPR per trobar models que permetin el càlcul de les propietats estudiades

en funció de l'estructura molecular dels aminoàcids i de les característiques dels cations metàl·lics. En aquest cas es prendran com a valors experimentals els calculats aquí ja que conformen una base de dades uniforme i completa, cosa que no es troba a la bibliografia.

#### 4.1.1. Afinitat i basicitat catióniques amb càlculs teòrics

Per tal d'obtenir teòricament uns resultats acurats de basicitat i afinitat en fase gas cal emprar unes funcions de base de qualitat elevada. En la bibliografia, les funcions de base utilitzades es troben, normalment, entre 6-31G i 6-311G amb funcions difuses i de polarització. Alguns dels funcionals emprats en aquest tipus de càlculs són BP86, B3LYP i BHandHLYP. La tria de les funcions de base i del funcional és molt important ja que, en aquest cas, és necessari que l'optimització de geometries i el càlcul de freqüències tardin un temps raonable i que, a més, aquests paràmetres estiguin disponibles per a tots els elements presents en els sistemes metall/aminoàcid que es volen estudiar. Per aquestes raons s'ha triat el funcional B3LYP en combinació amb les funcions de base DZVP de Salahub.<sup>26</sup> Aquestes funcions de base semblen prou adients ja que han estat optimitzades per a càlculs DFT, tenen qualitat doble zeta amb funcions de polarització i estan disponibles per a tots els elements fins el xenó, incloent-hi les dues primeres sèries de transició. Aquesta combinació de paràmetres ha estat utilitzada anteriorment pel grup de Hopkinson,<sup>18</sup> amb bons resultats, en el càlcul de l'afinitat de la plata front els vint aminoàcids habituals. És important tenir en compte l'error de superposició de base (BSSE), en el cas aquí tractat s'ha trobat que els valors de BSSE són relativament petits (menys de 8.5 kJ/mol) per a tots els metalls excepte el coure, que presenta uns valors al voltant dels 25 kJ/mol. La correcció d'aquest error s'ha fet, en tots els casos, amb el mètode *counterpoise*.<sup>27,28</sup>

Un altre punt de gran importància en aquest estudi és la manera com es coordinen els àtoms metàl·lics als aminoàcids. La Figura 4.2 mostra tots els possibles modes de coordinació per al cas dels aminoàcids més senzills, com la glicina o l'alanina.



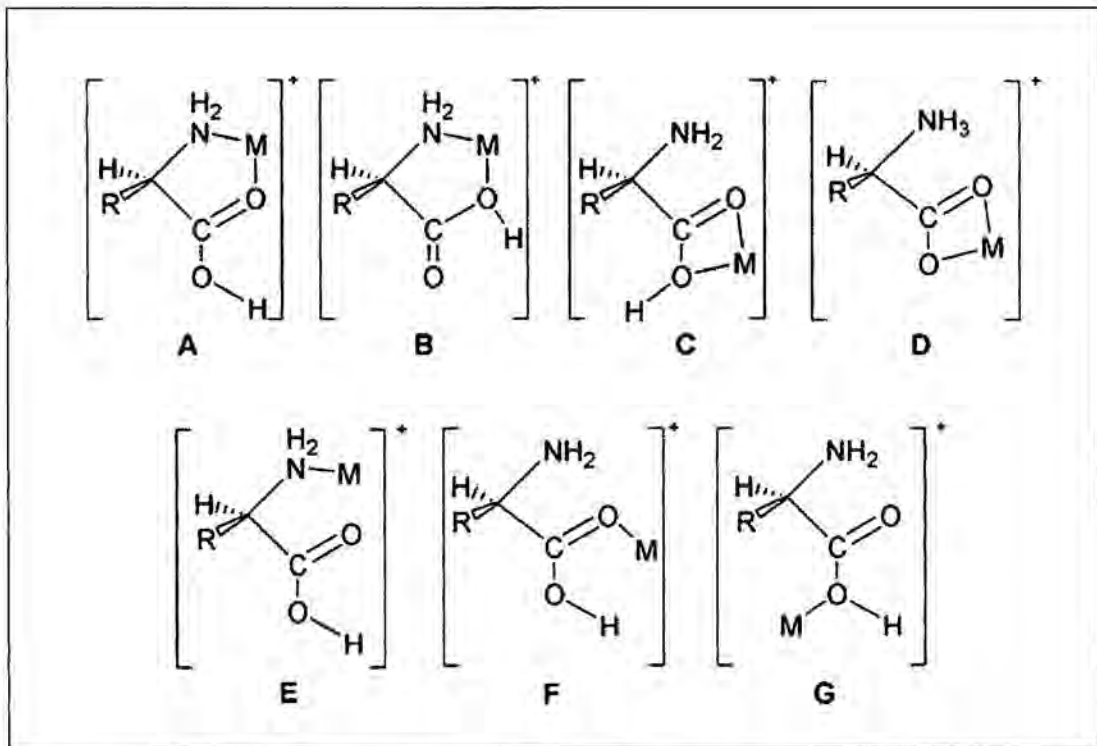


Figura 4.2. Possibles modes de coordinació  $[M(AA)]^+$  amb tres àtoms donadors.

En aquest cas hi ha tres àtoms donadors amb els quals es poden coordinar els cations i, per tant, es poden trobar quatre casos possibles en els quals l'aminoàcid actua com a lligand quelat, amb dos heteroàtoms diferents coordinant el metall. L'estructura tipus **A** utilitza l'oxigen carbonílic i el nitrogen com a àtoms donadors, en l'estructura **B** la coordinació es dóna a través de l'oxigen carboxílic i el nitrogen. En les estructures **C** i **D** la unió es produeix mitjançant els dos àtoms d'oxigen, en el cas de la primera l'aminoàcid es manté en la seva forma neutra mentre que en la segona es troba amb caràcter zwitteriònic. En principi, la forma zwitteriònica d'un aminoàcid no es forma en fase gas, almenys en els casos en què no hi ha un heteroàtom en la cadena lateral, però es pot estabilitzar mitjançant la coordinació a l'àtom metàl·lic.<sup>29</sup> A priori, és possible que la coordinació del catió es doni solament amb un dels heteroàtoms presents: el nitrogen (Estructura **E**) i els oxígens carbonílic (**F**) i carboxílic (**G**). Tot i així, és d'esperar que les formes monodentades siguin menys estables que les que comporten la formació d'un quelat. D'acord amb la bibliografia,<sup>30,31</sup> quan la grandària del catió és petita l'estructura **A** és la més afavorida mentre que un increment de la mida afavoreix

la formació de les estructures **C** i **D**; l'única excepció a aquesta regla és la prolina, per la qual es troba que la forma zwitteriònica sol ser la més estable.

En el cas en què l'aminoàcid tingui més de tres àtoms donadors, el nombre de modes de coordinació augmenta considerablement. La Figura 4.3 en mostra alguns, excloent-hi les formes monodentades.

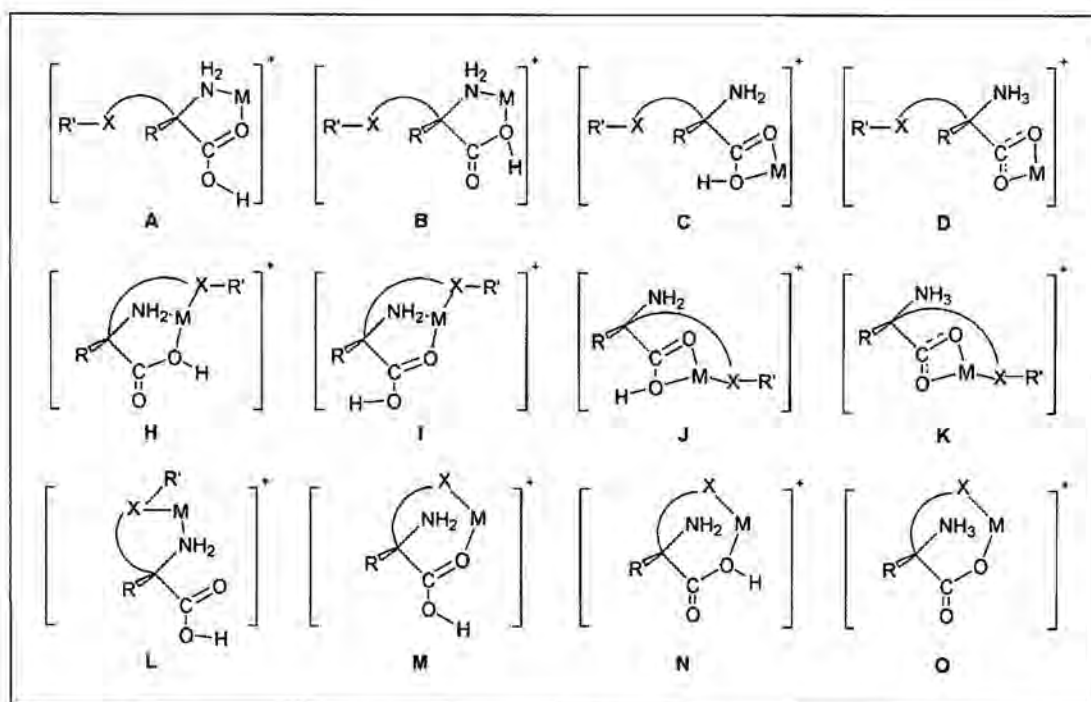


Figura 4.3. Possibles modes de coordinació  $[M(AA)]^+$  amb més de tres àtoms donadors.

A més dels modes de coordinació ja comentats, mono- (**E-G**) i didentats (**A-D**), hi ha quatre possibles modes tridentats (**H-K**) i quatre modes didentats (**L-O**) en els quals el segon enllaç es forma entre el catió i un dels àtoms donadors presents en la cadena lateral. Els aminoàcids que tenen grups aromàtics en la cadena lateral: fenilalanina, tirosina, histidina i triptòfan, poden interaccionar amb el catió a través del sistema  $\pi$ . Els aminoàcids que tenen un nombre d'heteroàtoms més gran en la cadena lateral: arginina, asparagina, àcid aspàrtic, glutamina i àcid glutàmic, poden presentar un nombre encara més gran de possibles modes de coordinació.

La flexibilitat de la cadena lateral comporta un gran nombre de confòrmers diferents dels aminoàcids, ja siguin lliures o coordinats. Per tal d'obtenir els valors

correctes de basicitat i afinitat és necessari conèixer, en cada cas, el confòrmer de mínima energia. Aquests mínims d'energia s'han trobat analitzant diferents confòrmeres amb càlculs de mecànica molecular amb el camp de forces MMFF<sup>32</sup> implementat en el programa Spartan;<sup>33</sup> en el cas dels aminoàcids lliures s'ha realitzat també un estudi a nivell HF/3-21G\*. Els confòrmeres de mínima energia trobats, a l'igual que els que s'han trobat en la bibliografia, s'han optimitzat al nivell B3LYP/DZVP amb el programa Gaussian03<sup>34</sup> i s'han seleccionat els més estables per a la posterior realització dels càlculs definitius. Els valors calculats d'afinitat ( $\Delta H^\circ$ ) i de basicitat ( $\Delta G^\circ$ ) catiónica s'han obtingut a partir de l'anàlisi termoquímica a 298 K utilitzant el càlcul de freqüències vibracionals a nivell B3LYP/DZVP. Per tal de mantenir el nombre de càlculs en un nivell raonable s'ha seguit el següent procediment per als compostos [catió/aminoàcid]<sup>+</sup>: en primer lloc s'han calculat els quatre modes de coordinació (A-D) de l'aminoàcid més simple, la glicina, amb tots els cations; el complexos més estables formats amb liti, sodi, coure i plata adopten l'estructura A. S'accepta que aquest mateix tipus de coordinació serà el que es trobarà amb alanina, isoleucina, leucina i valina per a aquests cations, ja que no presenten cap altre heteroàtom en la cadena lateral. S'ha trobat que el mode de coordinació C és el més estable per al sistema [KGly]<sup>+</sup>; així doncs, per a la resta d'aminoàcids sense heteroàtoms en la cadena lateral s'han calculat els modes de coordinació A i C. En tots els casos, els sistemes de tipus C són els més estables. En el cas dels sistemes formats per liti, sodi, coure i plata amb aminoàcids que presenten quatre heteroàtoms, s'han calculat, també, els modes de coordinació de tipus I, relacionats directament amb el mode A; mentre que per als sistemes formats pel potassi i els aminoàcids arginina, histidina, lisina, metionina i fenilalanina, s'han calculat també els modes de coordinació J. En tots els casos, aquests són menys estables que els de tipus I. Amb els aminoàcids que presenten grups carboxílics s'han estudiat preferentment els sistemes en els quals la interacció es dona a través de l'oxigen carbonílic, enfront de l'oxigen hidroxílic; aquesta decisió es basa en la major estabilitat que presenta el mode de coordinació A enfront del B. En últim lloc, per als aminoàcids asparagina i glutamina, s'ha trobat que els sistemes de tipus I en els quals l'oxigen amídic de la cadena lateral es coordina al metall són més estables que aquells en què la cadena lateral s'uneix a través del nitrogen. La geometria final seleccionada per a tots els sistemes [M(AA)]<sup>+</sup> es pot veure en la Figura 4.4. Per tal d'analitzar els resultats obtinguts es presentaran, en primer lloc, els que s'obtenen amb les diferents famílies de cations i després es realitzarà un estudi comparatiu de tots ells alhora.



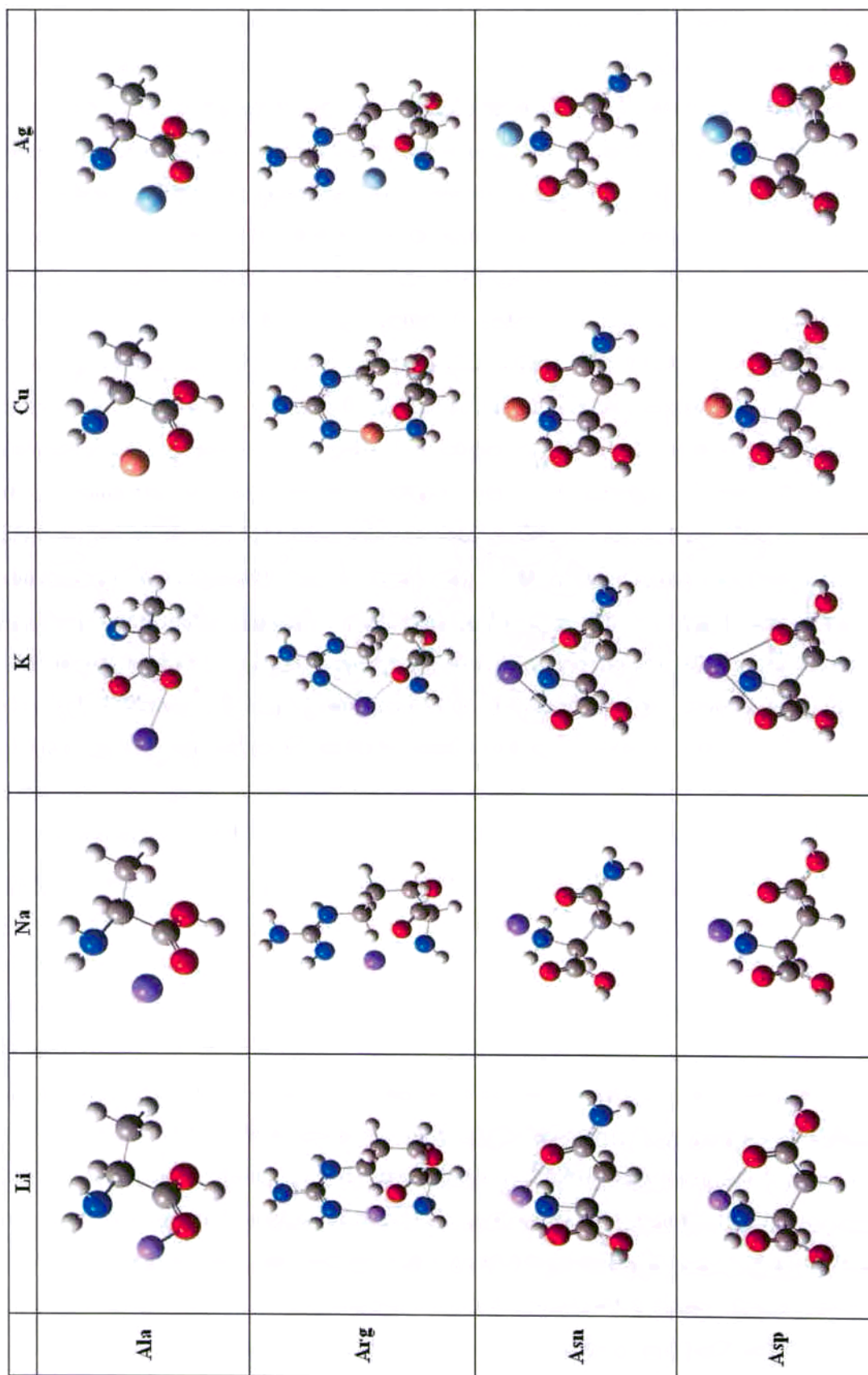


Figura 4.4. Geometria seleccionada per als sistemes  $[M(AA)]^+$ .

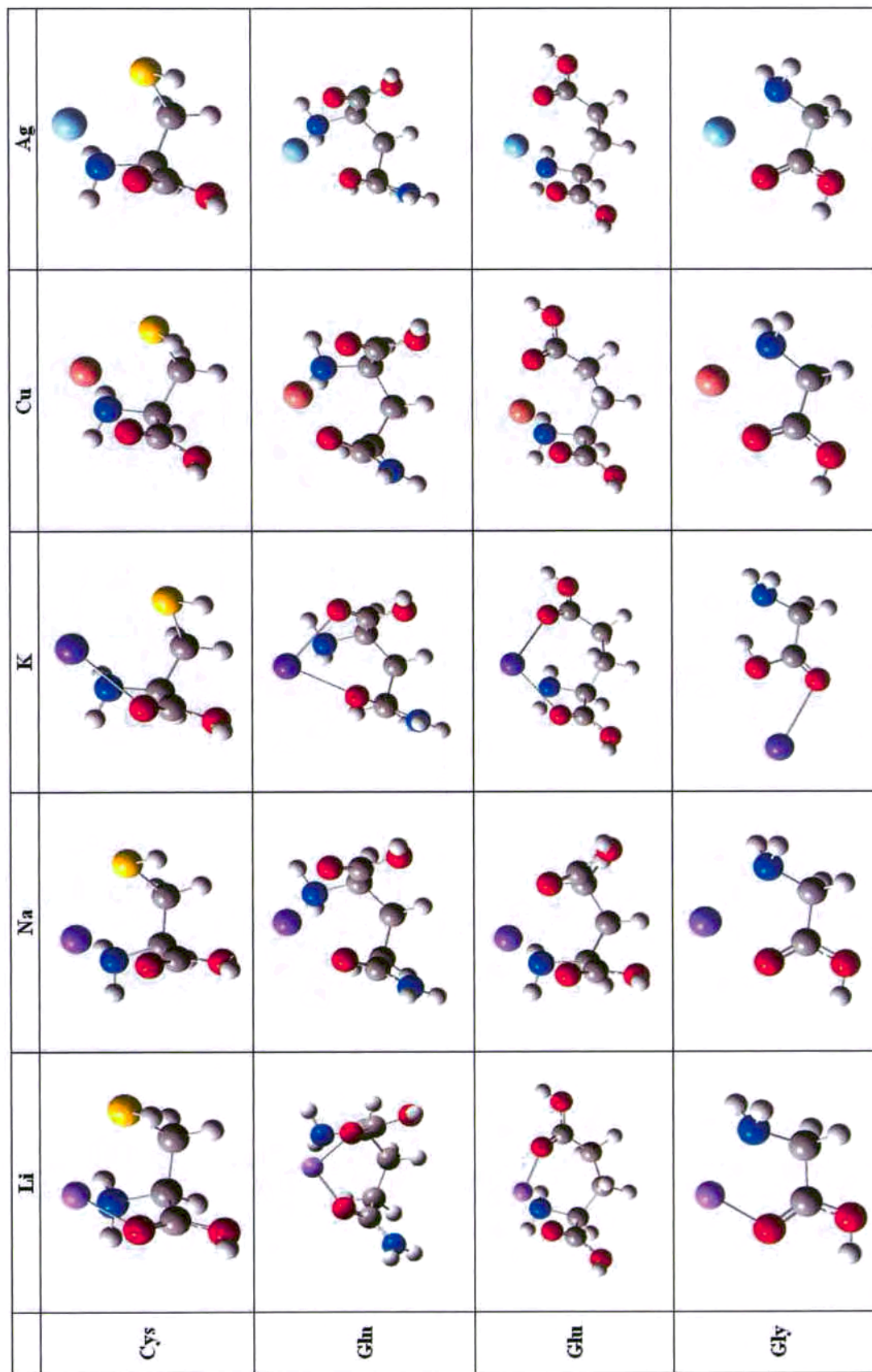


Figura 4.4. Geometria seleccionada per als sistemes  $[M(AA)]^+$ .

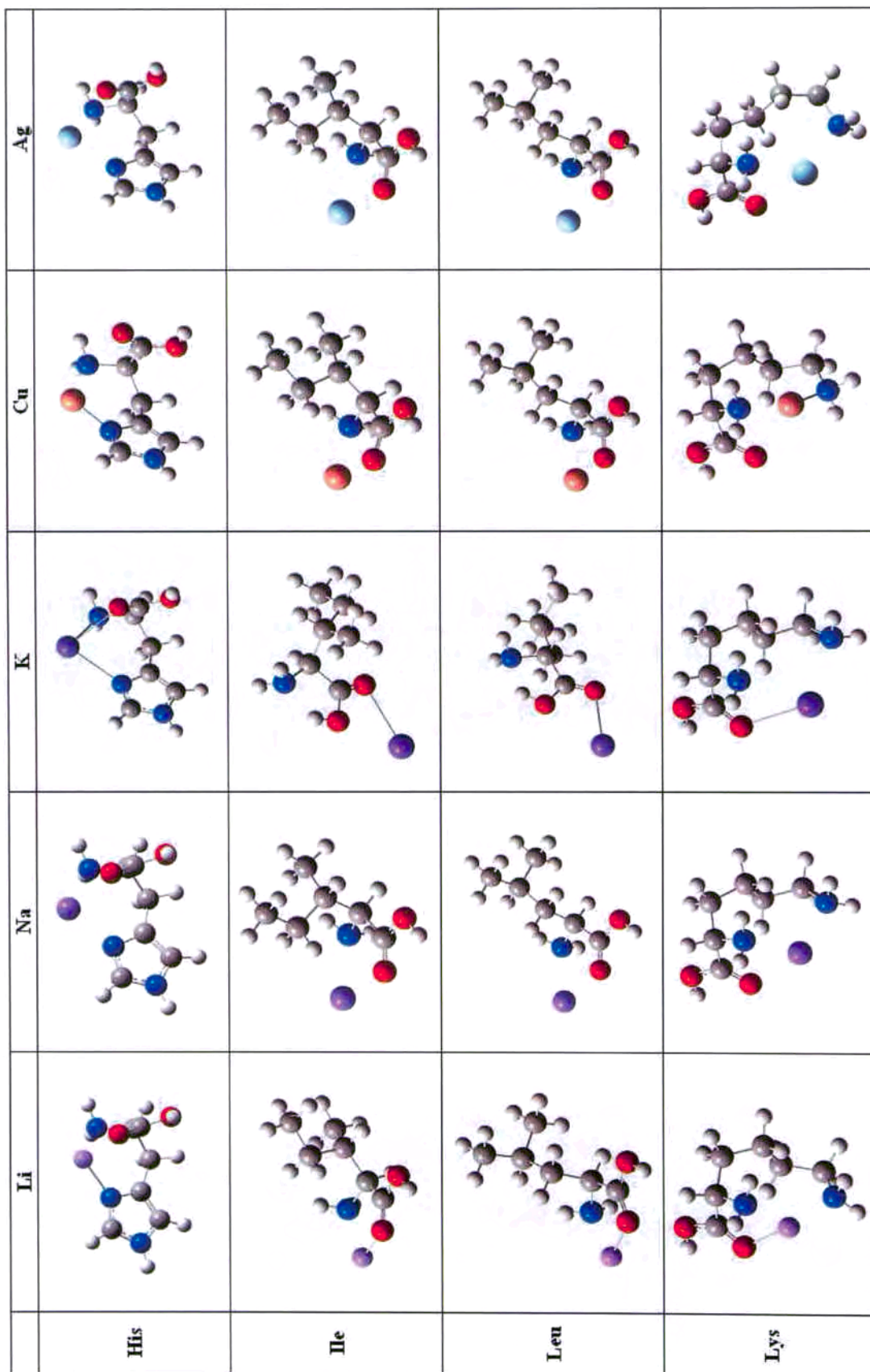


Figura 4.4. Geometria seleccionada per als sistemes  $[M(AA)]^+$ .

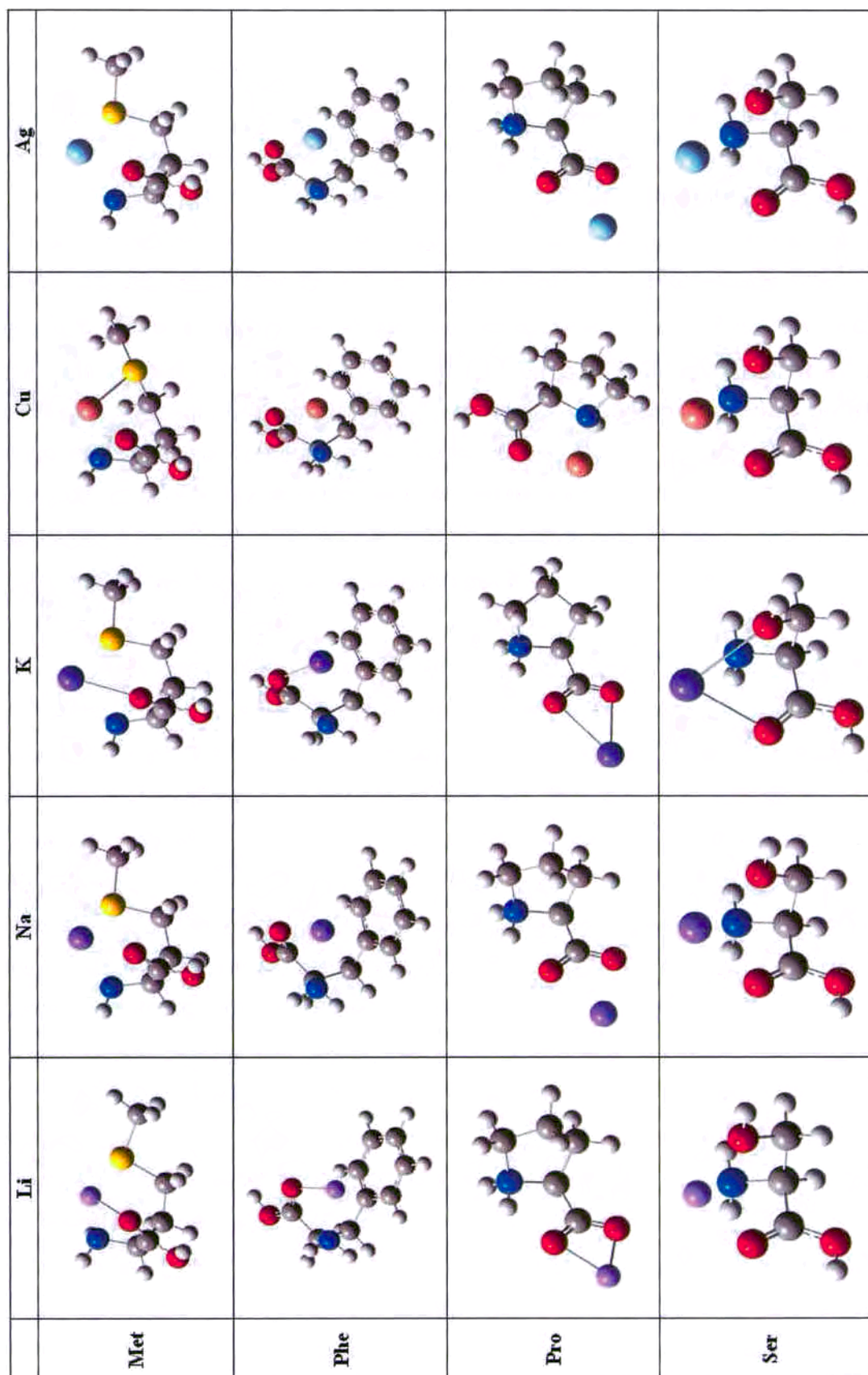


Figura 4.4. Geometria seleccionada per als sistemes  $[M(AA)]^+$ .



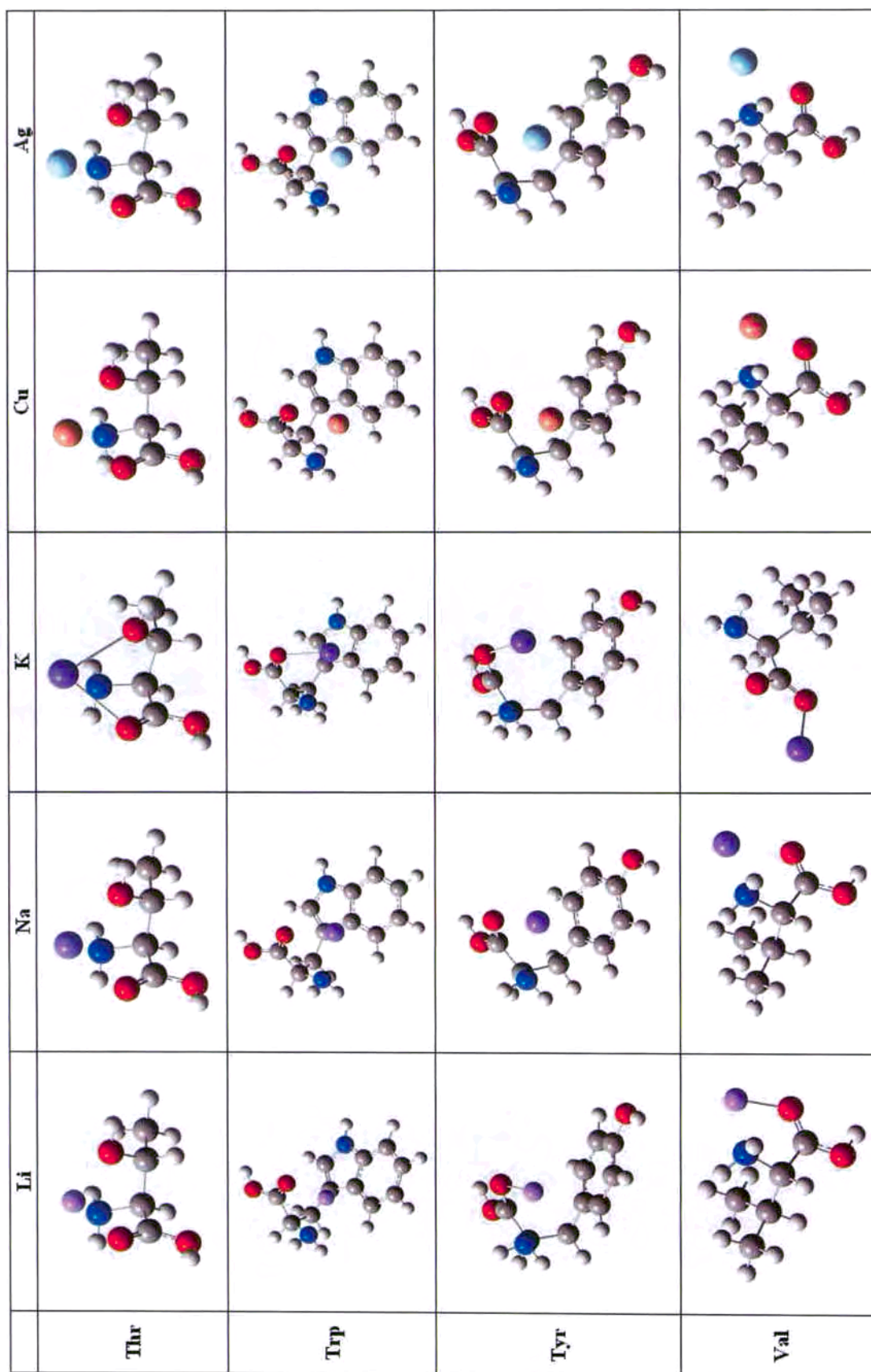


Figura 4.4. Geometria seleccionada per als sistemes  $[M(AA)]^+$ .

**Catió liti**

Com ja s'ha comentat, el primer pas a fer per poder calcular les afinitats i les basicitats consisteix en determinar el mode de coordinació de l'aminoàcid amb el catió metàl·lic. Alguns treballs teòrics previs<sup>31,35-38</sup> demostren que el sistema  $[\text{LiGly}]^+$  adopta el mode de coordinació de tipus **A**, on el catió liti s'enllaça a l'aminoàcid a través de l'oxigen carbonílic i el nitrogen; mentre que els isòmers on el catió es troba enllaçat a través dels dos àtoms d'oxigen en forma neutra (tipus **C**) i zwitteriònica (tipus **D**) presenten valors d'energia més elevats. En aquest treball s'han calculat els quatre possibles modes de coordinació (**A-D**) per al sistema  $[\text{LiGly}]^+$ ; el més estable de tots és el de tipus **A**, mentre que els de tipus **B** i **D** es troben a 41.2 kJ/mol i 20.9 kJ/mol per sobre, respectivament. En aquest cas, el sistema de tipus **C** no es correspon a un mínim d'energia, en la geometria final obtinguda el catió liti només es troba enllaçat a l'oxigen carbonílic, a una distància de més de 3 Å de l'oxigen hidroxílic; així, es pot dir que aquesta geometria correspon a un sistema de tipus **F**.

Alguns treballs previs demostren que els modes de coordinació més estables per a l'alanina<sup>6,35,39,40</sup> i la valina<sup>41</sup> són també els de tipus **A**. Per aquesta raó, per a aquests aminoàcids i per a la leucina i la isoleucina, que no presenten cap altre mode de coordinació alternatiu, aquesta és l'única forma que s'ha tingut en compte. El cas de la prolina és una excepció ja que aquest sistema adopta un mode de coordinació zwitteriònic (**D**), 14.6 kJ/mol més estable que el de tipus **A**, i 49.7kJ/mol més estable que el de tipus **C**. Per a la resta d'aminoàcids s'han calculat els modes de coordinació de tipus **I**, tenint en compte les diferents conformacions que es poden donar en la cadena lateral quan un dels seus heteroàtoms s'uneix al catió liti.

Els valors calculats d'afinitat ( $\Delta H^\circ$ ) i de basicitat ( $\Delta G^\circ$ ) de tots els sistemes  $[\text{Li(AA)}]^+$  es troben en la Taula 4.1, on també s'han recollit tots els valors experimentals i calculats trobats en la bibliografia. Bojesen i Andersen<sup>42</sup> han determinat l'ordre següent d'afinitat utilitzant el mètode cinètic: Arg > His > Gln > Asn > Lys > Trp > Glu > Asp > Tyr > Met > Phe > Thr > Pro > Ser > Ile ~ Leu > Val > Cys > Ala > Gly. Els resultats aquí obtinguts mostren una tendència pràcticament idèntica sent les úniques excepcions la lisina i l'asparagina que, d'acord amb els càlculs realitzats, tenen una afinitat major; i la fenilalanina, que té menor afinitat de l'esperada. El grup de Gronert<sup>43</sup> ha trobat experimentalment els valors de  $\Delta G^\circ$  corresponents als sistemes  $[\text{Li(AA)}]^+$

utilitzant escales de referència única a partir de determinacions amb el mètode cinètic. La comparació dels valors aquí obtinguts amb els determinats experimentalment permet veure que els càlculs proporcionen sistemàticament valors més grans que els experimentals; tot i així, s'ha reportat que el mètode cinètic tendeix a subestimar els valors d'energia d'enllaç.<sup>44</sup> La Figura 4.5 mostra la representació gràfica dels valors de  $\Delta G^\circ$  calculats vs. experimentals per als sistemes  $[\text{Li}(\text{AA})]^+$ .

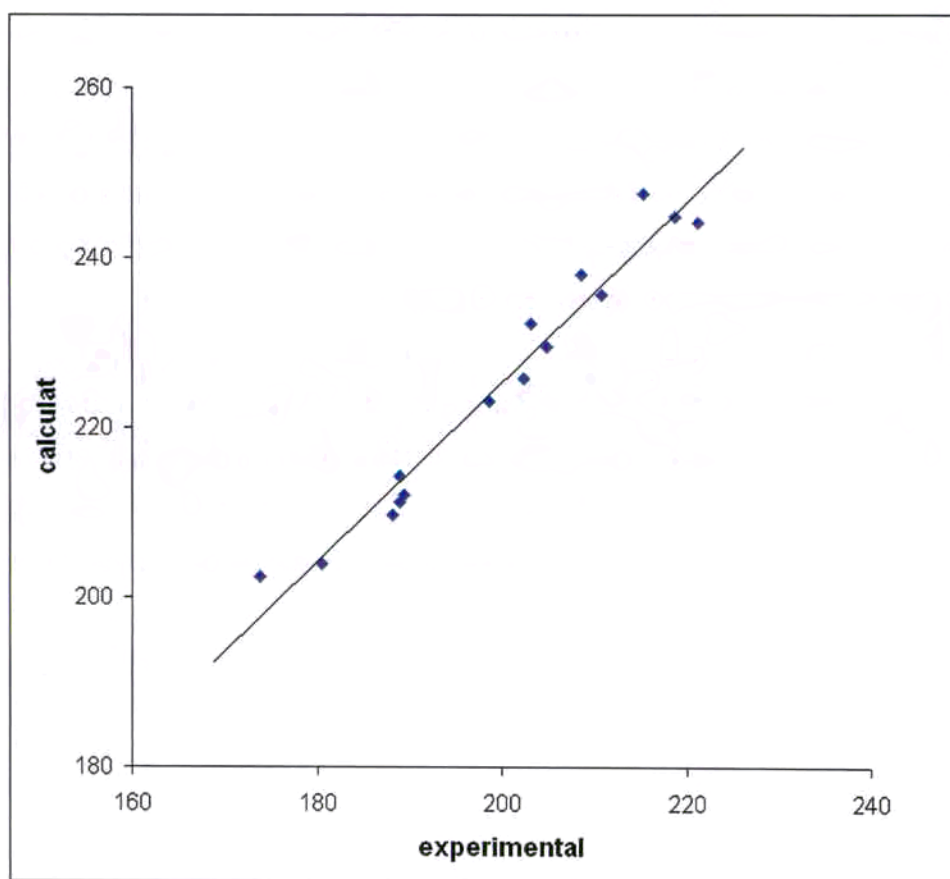


Figura 4.5.  $\Delta G^\circ$  calculat vs. experimental per als sistemes  $[\text{Li}(\text{AA})]^+$

Taula 4.1.  $\Delta H^\circ$  i  $\Delta G^\circ$  experimentals i calculats per als compostos  $[\text{Li}(\text{AA})]^+$  (kJ/mol).

Aminoàcid	$\Delta H^\circ$	$\Delta H^\circ(\text{calc})$	$\Delta H^\circ(\text{exp})$	$\Delta G^\circ$	$\Delta G^\circ(\text{calc})$	$\Delta G^\circ(\text{exp})$
Ala	234.1	236-248 <sup>a</sup>	220 <sup>b</sup>	204.0	209.4 <sup>c</sup>	180 <sup>d</sup>
Arg	347.4			318.5		
Asn	297.2			267.1		
Asp	279.2			247.5		215 <sup>d</sup>
Cys	241.6			211.1		189 <sup>d</sup>
Gln	307.2			271.7		
Glu	275.9			244.1		221 <sup>d</sup>
Gly	235.8	228-278 <sup>e</sup>	213 <sup>b</sup> ;226 <sup>f</sup>	202.3	202-223 <sup>g</sup>	174 <sup>d</sup> ;176 <sup>f</sup>
His	318.9			288.4		
Ile	242.4			211.9		189 <sup>d</sup>
Leu	242.9			214.0		189 <sup>d</sup>
Lys	307.6			278.8		
Met	272.1			231.6		211 <sup>d</sup>
Phe	257.5		275 <sup>h</sup>	225.7		
Pro	256.7	263-274 <sup>i</sup>		222.8	247.9 <sup>j</sup>	199 <sup>d</sup>
Ser	262.9			232.4		203 <sup>d</sup>
Thr	268.4			237.8		209 <sup>d</sup>
Trp	277.6			244.5		219 <sup>d</sup>
Tyr	258.7			229.5		205 <sup>d</sup>
Val	245.4			209.8		188 <sup>d</sup>

<sup>a</sup> Valors calculats: 236.2 (B3LYP, TZVP amb DZVP pels àtoms metàl·lics, no BSSE)<sup>35</sup>; 245.8 (B3LYP/6-311++G\*\*, correcció BSSE)<sup>6</sup>; 248.3 (B3LYP/6-311++G\*\*, no BSSE)<sup>35</sup>. <sup>b</sup> Mètode cinètic<sup>45</sup>. <sup>c</sup> B3LYP/6-311++G\*\*, correcció BSSE<sup>6</sup>. <sup>d</sup> Mètode cinètic<sup>43</sup>. <sup>e</sup> Valors calculats: 227.8 (MP2/6-311+G\*\*//MP2/6-31G\*)<sup>31</sup>; 231.6 (B3LYP, TZVP amb DZVP pels àtoms metàl·lics, no BSSE)<sup>35</sup>; 236.0 (Mètode CBS-QB3, no BSSE)<sup>36</sup>; 243.7 (B3LYP/6-311++G\*\*, no BSSE)<sup>35</sup>; 244.5 (MP2/6-311+G\*\*//MP2/6-31G\*, no BSSE)<sup>37</sup>; 250.0 (B3LYP/6-311+G\*\*, no BSSE)<sup>36</sup>; 257.1 (B3LYP/6-31+G\*, no BSSE)<sup>36</sup>; 277.6 (MP2/6-31G\*\*//MP2/6-31G\*, no BSSE)<sup>38</sup>. <sup>f</sup> Ref.<sup>31</sup>, recalculat de la Ref.<sup>45</sup>. <sup>g</sup> Valors calculats: 192.7 (MP2/6-311+G\*\*//MP2/6-31G\*)<sup>31</sup>; 201.7 (Mètode CBS-QB3, no BSSE)<sup>36</sup>; 215.7 (B3LYP/6-311+G\*\*, no BSSE)<sup>36</sup>; 222.7 (B3LYP/6-31+G\*, no BSSE)<sup>36</sup>. <sup>h</sup> Ref.<sup>31</sup>, recalculat de la Ref.<sup>45</sup>. <sup>i</sup> Valors calculats: 263.3 (MPW1PW91/6-311++G\*\*, correcció BSSE)<sup>46</sup>; 266.3 (B88LYP, 6-311++G\*\*, correcció BSSE)<sup>46</sup>; 267.1 (PW91PW91/6-311++G\*\*, correcció BSSE)<sup>46</sup>; 268.4 (B3LYP/6-311++G\*\*, correcció BSSE)<sup>46</sup>; 269.6 (B1LYP/6-311++G\*\*, correcció BSSE)<sup>46</sup>; 273.4 (BHandHLYP/6-311++G\*\*, correcció BSSE)<sup>46</sup>. <sup>j</sup> Dissociació induïda per col·lisió<sup>47</sup>.



### *Catió sodi*

Hi ha un gran nombre de publicacions que estudien l'afinitat dels sistemes  $[\text{Na}(\text{AA})]^+$ , tant des del punt de vista experimental com teòric. La Taula 4.2 mostra els valors de  $\Delta H^\circ$  i de  $\Delta G^\circ$  aquí calculats així com tots els experimentals i calculats trobats en la bibliografia.

Entre tots els exemples reportats en la bibliografia, hi ha algunes discrepàncies en les estructures dels sistemes  $[\text{Na}(\text{AA})]^+$ . Els càlculs teòrics apunten a que els modes de coordinació més estables són els corresponents a formes de tipus **A** i **I**, depenent de quants heteroàtoms tingui l'aminoàcid; mentre que algunes mesures experimentals apunten a què, almenys en els derivats de glicina i alanina, el tipus de coordinació **D** és el més estable. Com ja s'ha comentat, es pot estabilitzar la forma zwitteriònica d'un aminoàcid en fase gas mitjançant la seva coordinació al catió metàl·lic. Tot i així, en aquest cas sembla que les discrepàncies trobades es poden explicar mitjançant una transició de la forma **A** a la forma **D**, en les condicions experimentals de mesura.<sup>48</sup> Els càlculs realitzats en aquest treball mostren que el mode de coordinació més estable per al sistema  $[\text{NaGly}]^+$  és el de tipus **A**, mentre que la forma zwitteriònica (**D**) i les formes **B** i **C** presenten valors d'energia 8.8, 13.8 i 37.6 kJ/mol superiors, respectivament. Els càlculs amb l'alanina i la valina, a l'igual que en altres exemples trobats en la bibliografia,<sup>35,49</sup> demostren que els modes de coordinació més estables són els de tipus **A**. En el sistema  $[\text{NaPro}]^+$  el mode de coordinació **D** és el més estable, 32.7 kJ/mol més estable que la forma **A**, confirmant així el comportament trobat en la bibliografia.<sup>47</sup>

Els valors d'afinitat obtinguts per als sistemes de sodi segueixen la mateixa tendència, determinada experimentalment, trobada pel grup de Wesdemiotis<sup>50</sup> utilitzant el mètode cinètic, amb inversions en l'ordre entre diverses parelles d'aminoàcids que presenten valors d'afinitat molt semblants: Trp/Asn, Glu/Asp i Ile/Val. La Figura 4.6 mostra la representació gràfica dels valors d'afinitat aquí calculats enfront dels experimentals del grup de Wesdemiotis. Tal i com passava en el cas dels sistemes de liti, els valors calculats són sempre més alts que els experimentals; cal recordar, però, que el mètode cinètic tendeix a subestimar aquest tipus de propietats.<sup>44</sup>

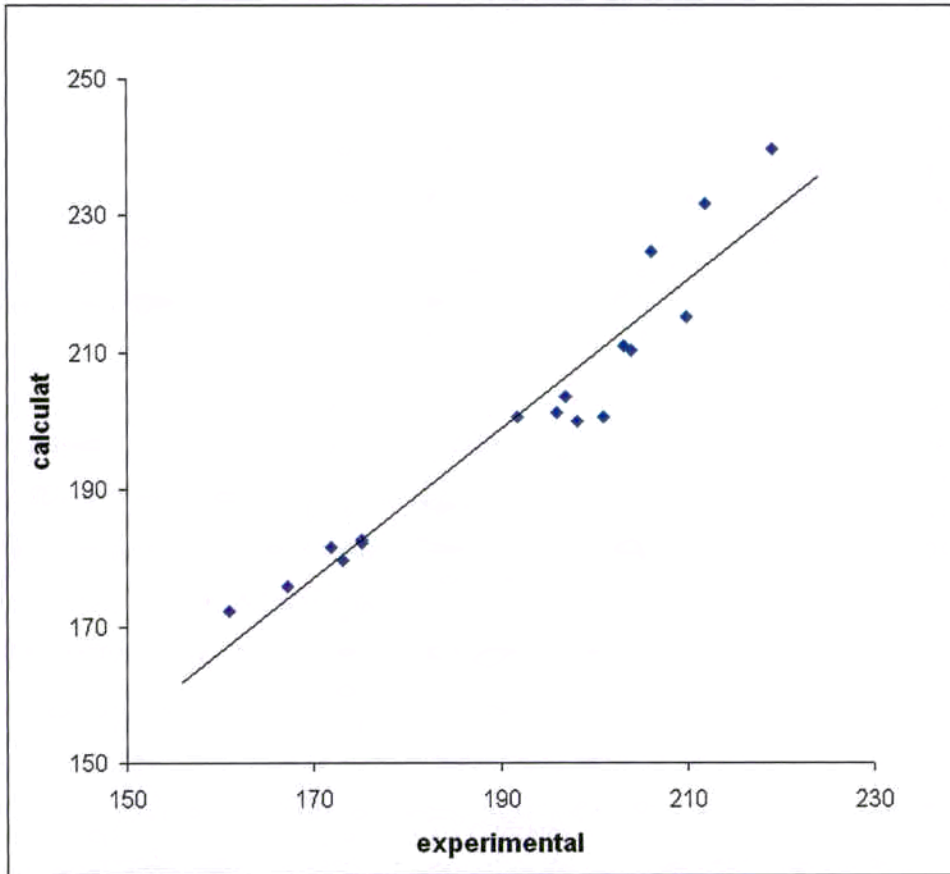


Figura 4.6  $\Delta H^\circ$  calculat vs. experimental per als sistemes  $[\text{Na}(\text{AA})]^+$ .

El grup de Dunbar<sup>16</sup> va determinar experimentalment algunes afinitats per a sistemes  $[\text{Na}(\text{AA})]^+$ ; en aquest cas, els valors obtinguts amb la prolina i la histidina, 163.9 i 226.1 kJ/mol respectivament, semblen massa baixos.

Taula 4.2.  $\Delta H^\circ$  i  $\Delta G^\circ$  experimentals i calculats per als compostos  $[\text{Na}(\text{AA})]^+$  (kJ/mol).

Aminoàcid	$\Delta H^\circ$	$\Delta H^\circ(\text{calc})$	$\Delta H^\circ(\text{exp})$	$\Delta G^\circ$	$\Delta G^\circ(\text{calc})$	$\Delta G^\circ(\text{exp})$
Ala	176.0	157-175 <sup>a</sup>	165-169 <sup>b</sup>	140.9	123.0 <sup>c</sup> ;134.6 <sup>d</sup>	
Arg	264.2		>225 <sup>e</sup>	231.6		
Asn	224.9	188.1 <sup>f</sup>	206 <sup>e</sup>	191.9		
Asp	211.1		203 <sup>e</sup>	176.4		
Cys	182.2	168-180 <sup>g</sup>	175 <sup>e</sup> ;179 <sup>f</sup>	148.4		
Gln	231.6		212 <sup>e</sup>	193.1		
Glu	210.7		204 <sup>e</sup>	173.9		
Gly	172.2	151-182 <sup>h</sup>	153-170 <sup>i</sup>	140.9	116-149 <sup>j</sup>	132 <sup>k</sup>
His	239.9	226.1 <sup>f</sup>	185 <sup>f</sup> ;219 <sup>e</sup>	205.7		
Ile	181.4		176 <sup>e</sup>	145.5		
Leu	182.7		175 <sup>e</sup>	149.2		
Lys	232.0		>213 <sup>e</sup>	198.1		
Met	206.1			165.5		
Phe	200.2	188-201 <sup>l</sup>	173-198 <sup>m</sup>	164.3		
Pro	201.5	164-200 <sup>n</sup>	175-196 <sup>o</sup>	168.9		156 <sup>p</sup>
Ser	200.6	188-200 <sup>q</sup>	188 <sup>f</sup> ;192 <sup>e</sup>	165.5		
Thr	203.6		197 <sup>e</sup>	169.3		
Trp	215.3	216.9 <sup>r</sup> ;217.4 <sup>r</sup>	180-210 <sup>s</sup>	178.5		
Tyr	200.6	201.9 <sup>r</sup>	175 <sup>r</sup> ;201 <sup>e</sup>	166.4		
Val	179.7		171-173 <sup>n</sup>	145.0		

<sup>a</sup> Valors calculats: 157.2 (MP2/6-31+G(d), correcció BSSE,  $D_0$ )<sup>51</sup>; 158.0 (MP2/6-311+G(2d,2p)/MP2/6-31G\*, correcció BSSE)<sup>8</sup>; 167.0 (MP2/6-311+G(2d,2p)/MP2/6-31G\*, no BSSE)<sup>50</sup>; 168.9 (DFT-B3P86/6-311+g(2df,2pd), correcció BSSE)<sup>16</sup>; 168.9 (B3LYP/6-311++G\*\*, no BSSE)<sup>35</sup>; 169.3 (B3LYP/6-311++G\*\*, correcció BSSE)<sup>6</sup>; 170.1 (B3LYP/6-31+G(d), correcció BSSE)<sup>44</sup>; 170.1 (B3LYP/6-31+G(d), correcció BSSE,  $D_0$ )<sup>51</sup>; 173.9 (B3-LYP/6-311+G(3df,2p)/B3-LYP/6-31+G(d), correcció BSSE, 0 K)<sup>40</sup>; 175.1 (B3LYP/TZVP, DZVP per Na, no BSSE)<sup>35</sup>. <sup>b</sup> Valors experimentals: 165 (Mètode cinètic)<sup>45</sup>; 167 (Mètode cinètic)<sup>50</sup>; 169 (Intercanvi de lligands en equilibri)<sup>16</sup>. <sup>c</sup> MP2/6-311+G(2d,2p)/MP2/6-31G\*, correcció BSSE <sup>8</sup>. <sup>d</sup> (B3LYP/6-311++G\*\*, correcció BSSE)<sup>6</sup>. <sup>e</sup> Mètode cinètic<sup>50</sup>. <sup>f</sup> (DFT-B3P86, energies finals calculades amb 6-311+g(2df,2pd), correcció BSSE)<sup>16</sup>. <sup>g</sup> Valors calculats: 167.6 (MP2/6-311+G(2d,2p)/MP2/6-31G\*, correcció BSSE)<sup>52</sup>; 173.1 (B3P86/6-311+g(2df,2pd), correcció BSSE)<sup>16</sup>; 180.2 (MP2/6-311+G(2d,2p)/MP2/6-31G\*, no BSSE)<sup>50</sup>. <sup>h</sup> Valors calculats: 150.9 (Mètode CBS-QB3)<sup>36</sup>; 151.7 (MP2/6-311+G(2d,2p)/MP2/6-31G\*, correcció BSSE)<sup>31</sup>; 154.0 (MP2/6-311+G(2d,2p)/MP2/6-31G\*, correcció BSSE)<sup>8</sup>; 157.2 (MP2/6-31G\*/HF/6-31G\*, no BSSE)<sup>38</sup>; 163.0 (DFT-B3P86, energies finals calculades amb 6-311+g(2df,2pd), correcció BSSE)<sup>16</sup>; 163.4 (B3LYP/6-311++G\*\*, no BSSE)<sup>35</sup>; 163.9 (energia calculada amb MP2/6-311+G(2d,2p), geometria optimitzada amb MP2/6-31G\*, no BSSE, 298 K)<sup>50</sup>; 165.9 (MP2/6-311+G(2d,2p)/MP2(full)/6-31G\*, correcció BSSE)<sup>53</sup>; 167.6 (B3LYP/6-31+G(d), correcció BSSE)<sup>44</sup>; 170.1 (B3LYP/TZVP, DZVP per Na, no BSSE)<sup>35</sup>; 172.5 (B3LYP/6-311+G(d,p), no BSSE)<sup>36</sup>; 182.4 (B3LYP/6-31+G\*, no BSSE)<sup>36</sup>. <sup>i</sup> Valors experimentals: 152 (Dissociació induïda per col·lisió)<sup>54</sup>; 158 (Mètode cinètic)<sup>45</sup>; 161 (Mètode cinètic)<sup>50</sup>; 161 (Intercanvi de lligands en equilibri)<sup>16</sup>; 167 (Dissociació induïda per col·lisió)<sup>55</sup>; 168<sup>31</sup>; revisat de la Ref.<sup>45</sup>. <sup>j</sup> Valors calculats: 116.2 (mètode CBS-QB3)<sup>36</sup>; 117.0 (MP2/6-311+G(2d,2p)/MP2/6-31G\*, correcció BSSE)<sup>31</sup>; 120.0 (MP2/6-311+G(2d,2p)/MP2/6-31G\*, correcció BSSE)<sup>8</sup>; 132.1 (MP2-(full)/6-311+G(2d,2p)/MP2(full)/6-31G\*, correcció BSSE)<sup>53</sup>; 137.9 (B3LYP/6-311+G(d,p), no BSSE)<sup>36</sup>; 148.6 (B3LYP/6-31+G(d), no BSSE)<sup>36</sup>. <sup>k</sup> Dissociació induïda per col·lisió<sup>56</sup>. <sup>l</sup> Valors calculats: 188.1 (DFT-B3P86/6-311+g(2df,2pd); correcció BSSE)<sup>16</sup>; 195.2 (B3LYP/6-31+G(d), correcció BSSE,  $D_0$ )<sup>51</sup>; 200.6 (B3LYP/6-31+G(d), correcció BSSE,  $D_0$ )<sup>44</sup>; 200.6 (B3LYP/6-31+G(d), correcció BSSE,  $D_0$ )<sup>51</sup>. <sup>m</sup> Valors experimentals: 174 (Mètode cinètic)<sup>44</sup>; 198 (Mètode cinètic)<sup>50</sup>; 198 (Intercanvi de lligands en equilibri)<sup>16</sup>. <sup>n</sup> Valors calculats: 163.9 (DFT-B3P86, energies finals amb 6-311+g(2df,2pd), correcció BSSE)<sup>16</sup>; 184.8 (MP2/6-

311+G(2d,2p)//MP2/6-31G\*, correcció BSSE)<sup>52</sup>; 192.3 (MPW1PW91/6-311++G\*\*, correcció BSSE)<sup>46</sup>; 195.2 (energia final amb MP2/6-311+G(2d,2p), geometria optimitzada amb MP2/6-31G\*, no BSSE, 298 K)<sup>50</sup>; 197.8 (B3LYP/6-311++G\*\*, correcció BSSE)<sup>46</sup>; 199.8 (PW91PW91/6-311++G\*\*, correcció BSSE)<sup>46</sup>.<sup>o</sup> Valors experimentals: 175 (Intercanvi de lligands en equilibri)<sup>16</sup>; 188 (Dissociació induïda per col·lisió)<sup>56</sup>; 188 (Dissociació induïda per col·lisió)<sup>47</sup>; 196 (Mètode cinètic)<sup>50</sup>.<sup>p</sup> Dissociació induïda per col·lisió<sup>56</sup>.<sup>q</sup> Valors calculats: 187.7 (MP2/6-311+G(2d,2p)//MP2/6-31G\*, correcció BSSE)<sup>52</sup>; 188.9 (DFT-B3P86, energies finals amb 6-311+g(2df,2pd), correcció BSSE)<sup>16</sup>; 199.8 (energia final amb MP2/6-311+G(2d,2p), geometria optimitzada amb MP2/6-31G\*, no BSSE, 298 K)<sup>50</sup>.<sup>r</sup> B3LYP/6-31+G(d), correcció BSSE<sup>44</sup>.<sup>s</sup> Valors experimentals: 180 (Mètode cinètic)<sup>44</sup>; 205 (Intercanvi de lligands en equilibri)<sup>16</sup>; 210 (Mètode cinètic)<sup>50</sup>.<sup>t</sup> Mètode cinètic<sup>44</sup>.<sup>u</sup> Valors experimentals: 171.4 (Mètode cinètic)<sup>45</sup>; 173.1 (Intercanvi de lligands en equilibri)<sup>16</sup>; 173.1 (Mètode cinètic)<sup>50</sup>.

### **Catió potassi**

Els sistemes  $[K(AA)]^+$  presenten un comportament diferent de la resta de sistemes estudiats. Amb els aminoàcids que no tenen cap heteroàtom ni sistema  $\pi$  en la cadena lateral, el mode de coordinació més estable es produeix amb els dos oxígens del grup carboxílic (tipus **C**). Per exemple, en el cas de la glicina la forma **C** és 2.0 kJ/mol més estable que la forma **A** i 10.5 kJ/mol més estable que la forma zwitteriònica (**D**). Aquests resultats estan d'acord amb els obtinguts per Tsang,<sup>57</sup> amb els quals es troba que la forma **C** és 2.7 kJ/mol més estable que la forma **A**. El tipus d'estructura **C** és també més estable que el tipus **A** en els casos de l'alanina (3.2 kJ/mol), la isoleucina (4.3 kJ/mol) i la leucina (4.1 kJ/mol). Aquests resultats coincideixen amb els obtinguts pel grup d'Ohanessian,<sup>31</sup> els quals indiquen que un augment en la grandària del catió metàl·lic implica l'estabilització del mode de coordinació **C**. Així doncs, en el cas del liti i el sodi, la forma **A** serà la més estable; mentre que el potassi, el rubidi i el cesi, afavoreixen la coordinació a través dels dos oxígens del grup carboxílic.

En canvi, els aminoàcids que tenen algun heteroàtom en la cadena lateral, els modes de coordinació més estables corresponen a la forma **I**, en què el catió potassi s'enllaça a l'oxigen carbonílic, al nitrogen del grup amino i a l'heteroàtom (o el sistema  $\pi$ ) de la cadena lateral. Per exemple, en el cas del  $[K\text{Phe}]^+$ , el complex de tipus **I** és 14.1, 20.1 i 37.8 kJ/mol més estable que les respectives formes **C**, **J** i **A**. Aquesta tendència és vàlida també per a la resta d'aminoàcids amb més de tres heteroàtoms com són la lisina, la metionina i l'arginina, que presenten modes de coordinació de tipus **I** 11.7, 8.8 i 24.2 kJ/mol més estables que les formes **J**, respectivament.

La Taula 4.3 mostra els valors calculats de  $\Delta H^\circ$  i  $\Delta G^\circ$  per als compostos  $[K(AA)]^+$ , en aquesta taula es recullen també tots els valors experimentals i calculats de la bibliografia. L'ordre d'afinitat obtingut és el següent: Arg > Gln > His > Asn ~ Lys > Trp > Asp ~ Glu > Pro > Phe > Tyr > Thr > Met > Ser > Leu > Val > Ile > Ala > Cys > Gly. La determinació d'afinitat experimental feta per Tsang i col·laboradors<sup>17</sup> per a la sèrie d'aminoàcids: Gly-Ala-Ile-Leu-Val no permet establir una correlació significativa amb els valors aquí calculats; tot i això, els valors experimentals i els calculats són bastant semblants, excepte en el cas de la isoleucina. Utilitzant el mètode cinètic, el grup de Ryzhov<sup>44</sup> va determinar l'afinitat de triptòfan, fenilalanina i tirosina. La comparació d'aquests valors experimentals amb els calculats en aquest treball permet comprovar que la tendència en els valors d'afinitat és la mateixa tot i que els valors experimentals semblen massa baixos, potser a causa del mètode utilitzat per determinar-los. La Figura 4.7 mostra la representació gràfica dels valors calculats enfront els valors experimentals obtinguts pels grups de Tsang i Ryzhov.

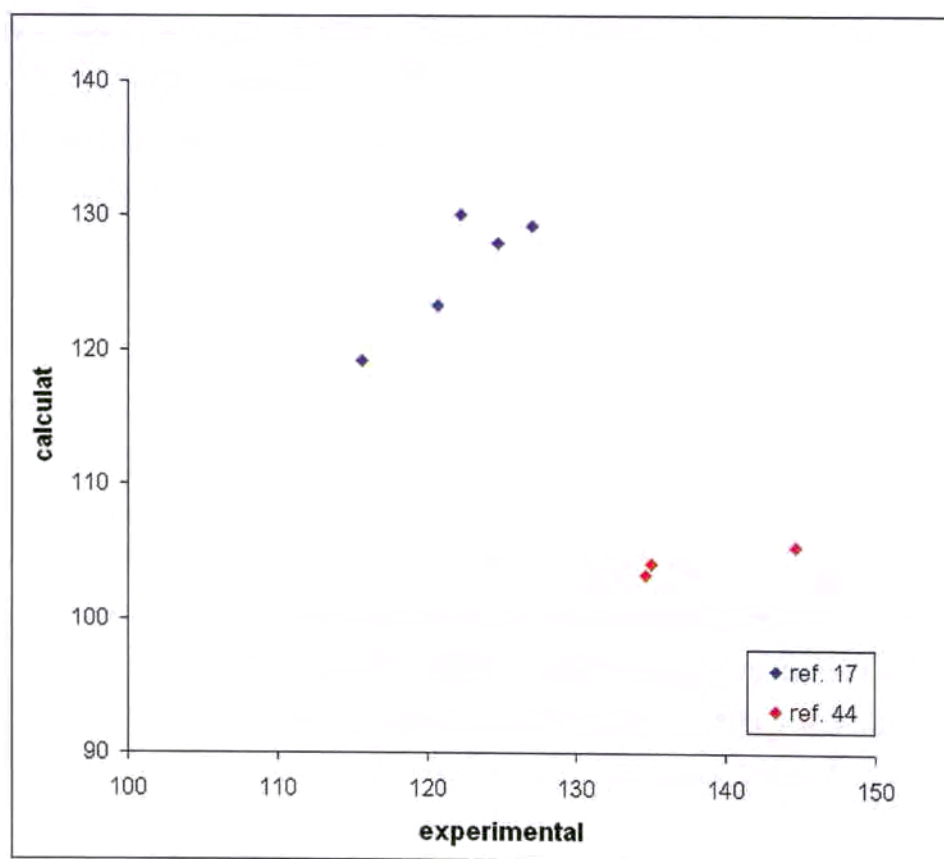


Figura 4.7.  $\Delta H^\circ$  calculat vs. experimental per als sistemes  $[K(AA)]^+$ .



Hi ha un gran nombre de treballs publicats en els que es calculen l'afinitat i la basicitat dels sistemes  $[K(AA)]^+$  amb diferents nivells de càlcul. La Figura 4.8 mostra la representació gràfica dels valors calculats en aquest treball enfront els obtinguts pel grup de Tsang.<sup>17</sup> En general, els valors aquí calculats són més baixos que els obtinguts en la resta de treballs publicats, excepte els obtinguts pel grup de Ryzhov amb nivell de càlcul B3LYP/6-31G\*.<sup>44</sup>

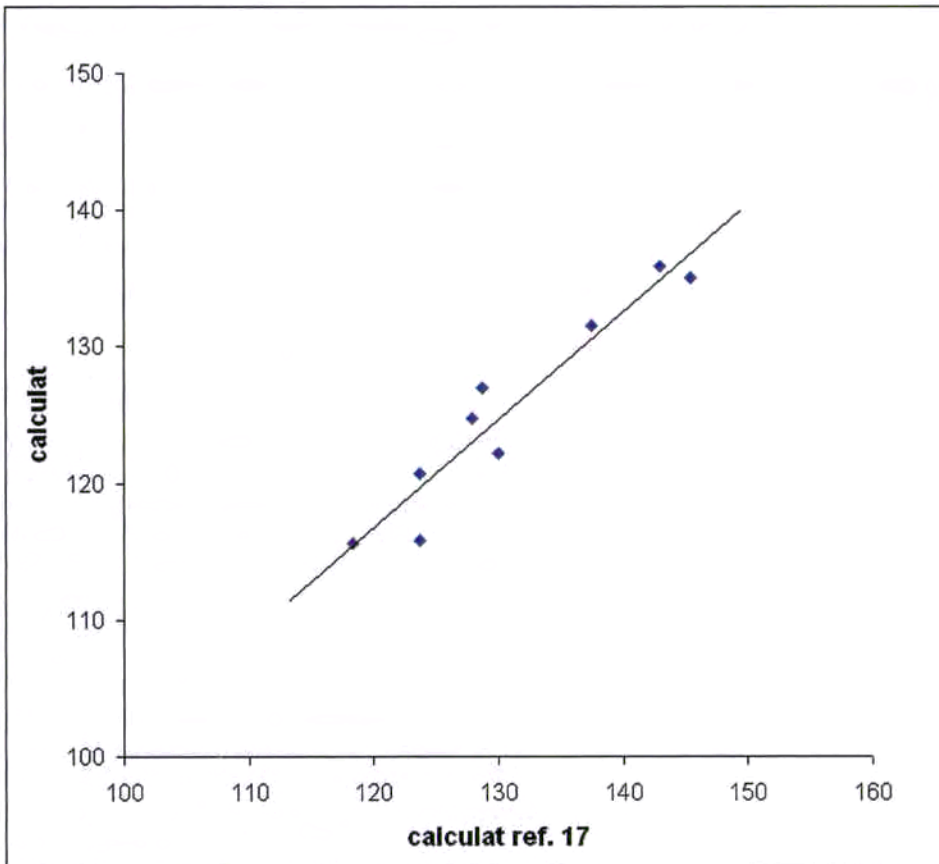


Figura 4.8.  $\Delta H^\circ$  calculats vs. referència 17 per als sistemes  $[K(AA)]^+$ .

Taula 4.3.  $\Delta H^\circ$  i  $\Delta G^\circ$  experimentals i calculats per als compostos  $[K(AA)]^+$  (kJ/mol).

Aminoàcid	$\Delta H^\circ$	$\Delta H^\circ(\text{calc})$	$\Delta H^\circ(\text{exp})$	$\Delta G^\circ$	$\Delta G^\circ(\text{calc})$	$\Delta G^\circ(\text{exp})$
Ala	120.8	115-124 <sup>a</sup>	123 <sup>b</sup>	91.1	88.6 <sup>c</sup>	
Arg	170.5			142.1		
Asn	150.5			119.5		
Asp	140.0			107.8		
Cys	115.8	123.7 <sup>d</sup>		85.7		
Gln	154.2			117.9		
Glu	137.1			103.7		
Gly	115.8	109-124 <sup>e</sup>	119-125 <sup>f</sup>	92.8	82-93 <sup>g</sup>	92 <sup>h</sup>
His	151.7			119.5		
Ile	122.1	130.0 <sup>d</sup>	130 <sup>b</sup>	89.9		
Leu	127.1	128.7 <sup>d</sup>	129 <sup>b</sup>	99.1		
Lys	150.1			120.0		
Met	131.7			95.7		
Phe	135.0	145.0 <sup>i</sup> ;145.5 <sup>d</sup>	104 <sup>j</sup>	99.9		
Pro	135.9	143-150 <sup>k</sup>	145 <sup>l</sup>	100.3		114 <sup>l</sup>
Ser	131.7	137.5 <sup>d</sup>		99.5		
Thr	133.3			101.6		
Trp	144.6	157.2 <sup>i</sup>	105 <sup>j</sup>	111.6		
Tyr	134.6	145.0 <sup>i</sup>	103 <sup>j</sup>	102.4		
Val	124.6	127.9 <sup>d</sup>	127 <sup>b</sup>	94.5		

<sup>a</sup> Valors calculats: 114.5 (B3LYP/6-31+G(d), correcció BSSE) <sup>44</sup>; 119.1 (MP2/6-311+G(2d,2p)//MP2/6-31G, amb funcions de base estàndard no equivalents per K, correcció BSSE) <sup>8</sup>; 123.7 (B3LYP/6-31+G(3df,29)//B3LYP/6-31G(d)) <sup>17</sup>. <sup>b</sup> Mètode cinètic <sup>17</sup>. <sup>c</sup> (MP2/6-311+G(2d,2p)//MP2/6-31G, amb funcions de base estàndard no equivalents per K, correcció BSSE) <sup>8</sup>. <sup>d</sup> (B3LYP/6-31+G(3df,29)//B3LYP/6-31G(d)) <sup>17</sup>. <sup>e</sup> Valors calculats: 109.1 (B3LYP/6-311+G(3df,2p)//B3LYP/6-31G(d), no BSSE) <sup>37</sup>; 110.8 (MP2/6-311+G(2d,2p)//MP2/6-31G\*, correcció BSSE) <sup>31</sup>; 112.4 (B3LYP/6-31+G(d), correcció BSSE) <sup>44</sup>; 112.9 (MP2/6-311+G(2d,2p)//MP2/6-31G, amb funcions de base estàndard no equivalents per K, correcció BSSE) <sup>8</sup>; 114.0 (MP2(full)/6-311pG(2d,2p)//MP2(full)/6-31pG(d,p), correcció BSSE, 0 K) <sup>58</sup>; 118.3 (B3LYP/6-31+G(3df,29)//B3LYP/6-31G(d)) <sup>17</sup>; 119.7 (B3LYP/6-311+G(d,p), no BSSE) <sup>36</sup>; 124.2 (B3LYP/6-31+G(d), no BSSE) <sup>36</sup>. <sup>f</sup> 119 (Mètode cinètic) <sup>17</sup>; 123 (Dissociació induïda per col·lisió) <sup>58</sup>; 125 (Dissociació induïda per col·lisió) <sup>54</sup>. <sup>g</sup> Valors calculats: 81.9 (MP2/6-311+G(2d,2p)//MP2/6-31G, amb funcions de base estàndard no equivalents per K, correcció BSSE) <sup>8</sup>; 87.4 (B3LYP/6-311+G(d,p), no BSSE) <sup>36</sup>; 92.8 (B3LYP/6-311+G(d,p), no BSSE) <sup>36</sup>. <sup>h</sup> Dissociació induïda per col·lisió <sup>58</sup>. <sup>i</sup> B3LYP/6-31+G(d), correcció BSSE. <sup>44</sup>. <sup>j</sup> Mètode cinètic <sup>44</sup>. <sup>k</sup> Valors calculats: 143.0 (B3LYP/6-31+G(3df,29)//B3LYP/6-31G(d)) <sup>17</sup>; 144.7 (B3LYP/6-311++G\*\*//B3LYP/6-311++G\*\*, correcció BSSE) <sup>46</sup>. <sup>l</sup> Mètode cinètic <sup>47</sup>.

*Catió coure*

El primer pas per tal de trobar el mode de coordinació més estable amb el  $\text{Cu}^+$  consisteix a calcular l'energia de les formes **A-D** per al complex de glicina. En aquest cas, el mode de coordinació de tipus **A** és el més estable mentre que les geometries de tipus **B**, **C** i **D** presenten valors d'energia 45.6, 46.3 i 30.5 kJ/mol superiors, respectivament, d'acord amb els resultats obtinguts per Russo.<sup>35</sup> El mode de coordinació **A** és també el més estable amb l'alanina, d'acord amb d'altres resultats trobats en la bibliografia.<sup>35</sup> El  $[\text{CuPro}]^+$  és l'únic en què aquest aminoàcid no adopta la forma zwitteriònica en coordinar-se al catió; així, en aquest cas el mode de coordinació **A** és 2.1 kJ/mol més estable que el **D**, coincidint amb els resultats trobats en la bibliografia.<sup>59</sup> D'acord amb els resultats obtinguts amb aquests aminoàcids, per a la resta de compostos  $[\text{Cu}(\text{AA})]^+$  s'ha calculat el mode de coordinació **A** si l'aminoàcid no presenta cap heteroàtom en la cadena lateral i el mode **I** quan sí en presenten, tenint en compte les possibles conformacions que es poden donar en la cadena lateral.

La Taula 4.4 mostra els valors de  $\Delta H^\circ$  i de  $\Delta G^\circ$  per als complexos  $[\text{Cu}(\text{AA})]^+$  calculats en aquest treball i trobats en la bibliografia. L'única determinació experimental per a aquests compostos ha estat realitzada pel grup de Wesdemiotis utilitzant el mètode cinètic,<sup>10,60</sup> assumint que les contribucions entròpiques de l'aminoàcid en la reacció són similars per a tots ells. Els valors obtinguts són relatius respecte al del sistema de la glicina, el valor del qual ha estat emprat com a àncora. En un primer pas els autors van mesurar totes les afinitats relatives,<sup>10</sup> excepte per a lisina i l'arginina ja que presenten diferències entròpiques molt grans respecte la resta d'aminoàcids. L'afinitat d'aquests dos aminoàcids es va mesurar, en un treball posterior, emprant espectrometria de masses de resolució angular.<sup>60</sup> Un cop s'havia establert l'ordre d'afinitat i les diferències d'energia entre els sistemes  $[\text{Cu}(\text{AA})]^+$ , Hoyau i Ohanessian van calcular l'afinitat del complex de glicina (271.3 kJ/mol)<sup>61</sup> i el van utilitzar com a àncora per calcular els valors d'afinitat absoluts d'aquesta família de compostos. La comparació dels valors obtinguts per Hoyau i Ohanessian amb els que s'obtenen en aquest treball permet comprovar que els aquí obtinguts són força més grans. Les diferències en el nivell de càlcul utilitzat poden ser una possible explicació per a aquest fet; Hoyau i Ohanessian van utilitzar el mètode MP2 realitzant, a més, un gran nombre d'aproximacions per estimar l'efecte de la grandària de les funcions de base i la correlació electrònica. També van fer càlculs DFT amb el funcional B3LYP i



van obtenir un valor per al  $[\text{CuGly}]^+$  (330.2 kJ/mol) superior a l'aquí trobat (302.2 kJ/mol). Sodupe i col·laboradors<sup>62</sup> van calcular l'afinitat per a aquest mateix sistema i van obtenir resultats de 279.6 i 313.5 kJ/mol utilitzant mètodes CCSD(T) i B3LYP/DFT, respectivament. Recentment, Hoppiliard i Ohanessian<sup>63</sup> proposen que les diferències obtingudes entre els mètodes MP2, CCSD(T) i B3LYP es deuen a què els mètodes  $\text{MP}_n$  i el CCSD(T) utilitzen orbitals moleculars poc apropiats i, per tant, condueixen a resultats erronis; tal i com es demostra en algunes publicacions.<sup>64,65</sup> Els càlculs d'afinitat realitzats pel grup de Sodupe<sup>66</sup> amb la fenilalanina, tirosina, triptòfan i histidina amb B3LYP són molt semblants als que s'obtenen en aquest treball.

És interessant comparar els resultats d'aquest treball amb els valors relatius determinats experimentalment, utilitzant el valor del sistema de glicina com a àncora, Figura 4.9.

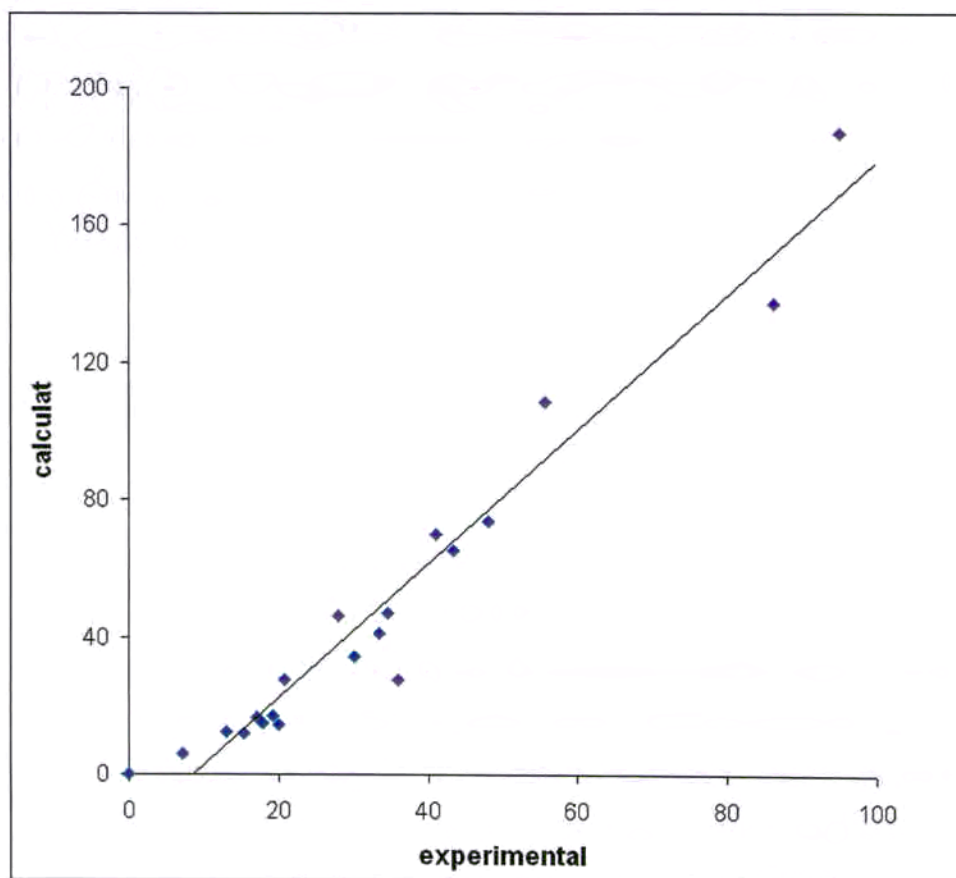


Figura 4.9.  $\Delta H^\circ$  relatiu calculat vs. experimental per als sistemes  $[\text{Cu}(\text{AA})]^+$ .

Com es pot comprovar, el pendent d'aquesta representació és molt més elevat que per a la resta de cations estudiats. Aquest fet, que també va ser constatat pel grup de Sodupe en el seu estudi dels sistemes de fenilalanina, tirosina, triptòfan i histidina,<sup>66</sup> és probablement degut a què la suposició que les contribucions entròpiques són similars per a tots els aminoàcids, quan s'utilitza el mètode cinètic és errònia i, per tant, els resultats que s'obtenen amb els aminoàcids didentats i tridentats no són comparables. Per tal de comprovar la certesa d'aquesta nova hipòtesi s'ha fet la representació gràfica dels valors d'afinitat relatius experimentals i calculats dels aminoàcids que no presenten cap substituent enllaçant en la cadena lateral: glicina, alanina, valina, isoleucina i leucina, Figura 4.10. En aquest cas, el pendent és molt més semblant a l'obtingut amb la resta de famílies estudiades i, com era d'esperar, l'ordenada a l'origen és zero.

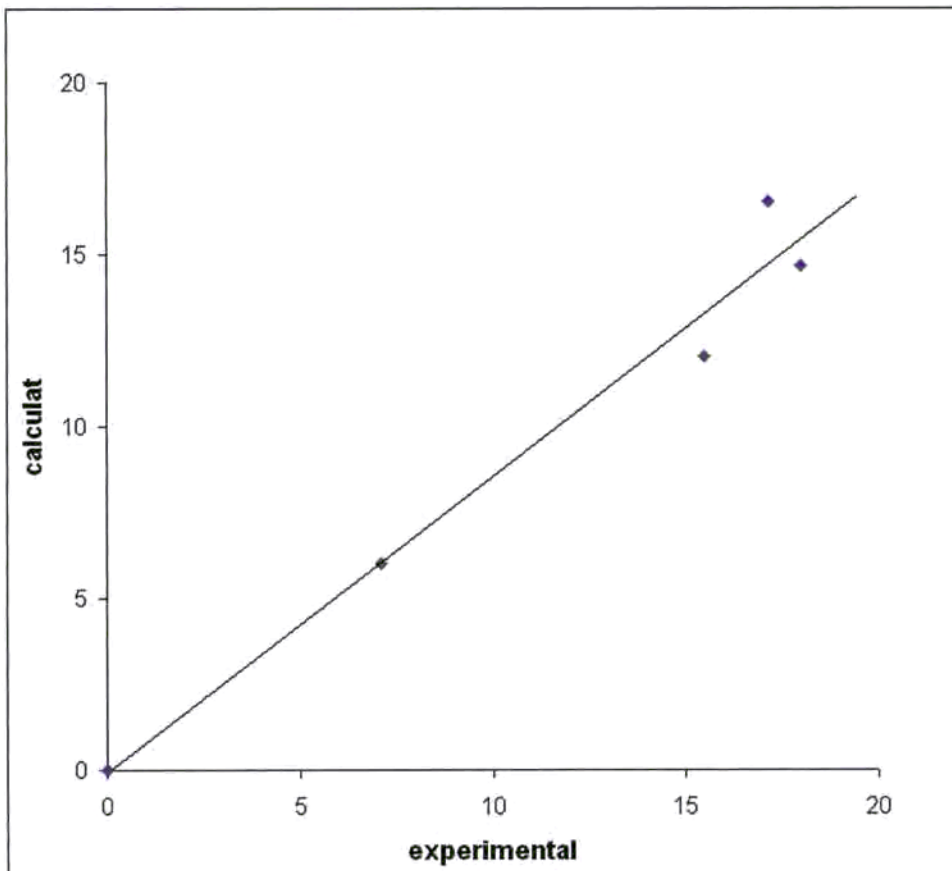


Figura 4.10.  $\Delta H^\circ$  relatiu calculat vs. experimental per als sistemes  $[\text{Cu}(\text{AA-didentat})]^+$ .

Taula 4.4.  $\Delta H^\circ$  i  $\Delta G^\circ$  experimentals i calculats per als compostos  $[\text{Cu}(\text{AA})]^\dagger$  (kJ/mol). Els valors absoluts s'obtenen dels relatius a partir del valor de  $\Delta H^\circ$  calculat per als compostos  $[\text{CuGly}]^\dagger$ .

Aminoàcid	$\Delta H^\circ$	$\Delta H^\circ(\text{calc})$	$\Delta H^\circ(\text{rel})^a$	$\Delta H^\circ(\text{abs})$	$\Delta G^\circ$	$\Delta G^\circ(\text{calc})$
Ala	308.5	285-306 <sup>b</sup>	7.1	309.3	271.3	
Arg	489.5		>55.6	>357.8	453.1	
Asn	348.6		28.0	330.2	314.8	
Asp	329.8		20.9	323.1	293.9	
Cys	329.8		35.9	338.2	293.4	
Gln	372.9		41.0	343.2	330.2	
Glu	336.5		30.1	332.3	298.0	
Gly	302.2	269-315 <sup>c</sup>	0.0	302.2	268.4	235.3 <sup>d</sup> ;258.3 <sup>e</sup>
His	410.9	408.0 <sup>f</sup>	>55.6	>357.8	374.9	372.4 <sup>f</sup>
Ile	317.3		18.0	320.2	280.1	
Leu	318.9		17.1	319.4	285.1	
Lys	439.7		55.6	357.8	401.7	
Met	367.4		43.5	345.7	323.5	
Phe	343.6	343.2 <sup>f</sup>	33.4	335.7	304.7	302.2 <sup>f</sup>
Pro	316.8	295.1 <sup>h</sup>	20.1	322.3	283.4	263.8 <sup>g</sup>
Ser	314.8		13.0	315.2	280.1	
Thr	319.4		19.2	321.4	284.7	
Trp	376.2	369.1 <sup>f</sup>	48.1	350.3	335.7	327.3 <sup>f</sup>
Tyr	349.4	346.1 <sup>f</sup>	34.7	336.9	310.6	305.6 <sup>f</sup>
Val	308.5		15.5	317.7	278.0	

<sup>a</sup> Mètode cinètic <sup>60</sup>. <sup>b</sup> Valors calculats: 285.1 (BHandHLYP/6-311++G\*\*, no BSSE) <sup>25</sup>; 295.9 (B3LYP, 6-311++G\*\*, LANL2DZ per Cu, no BSSE) <sup>35</sup>; 300.1 (B1LYP/6-311++G\*\*, no BSSE) <sup>25</sup>; 306.0 (B3LYP/6-311++G\*\*, no BSSE) <sup>25</sup>. <sup>c</sup> Valors calculats: 268.8 (MP2/6-311+G(2f,2d,2p)/MP2/6-31G\*) <sup>25</sup>; 271.3 (MP2/6-31G\*, correcció BSSE) <sup>61</sup>; 284.7 (CCSD(T)/DZP + funcions difoses per C, H, O i N, Watchers per Cu, geometria optimitzada amb B3LYP) <sup>62</sup>; 286.7 (MP2/6-311++G(2df,2p)/B3LYP/DZVP, correcció BSSE) <sup>52</sup>; 294.7 (B3LYP/DZVP, correcció BSSE) <sup>67</sup>; 295.9 (B3LYP/6-311+G(2d,2p)/B3LYP/6-31G\*) <sup>63</sup>; 300.1 (B3LYP, 6-311++G\*\*, LANL2DZ per Cu, no BSSE) <sup>35</sup>; 315.2 (B3LYP,D95++(d,p) + funció de polarització per Cu; no BSSE) <sup>62</sup>. <sup>d</sup> (MP2/6-31G\*, correcció BSSE) <sup>61</sup>. <sup>e</sup> (B3LYP/DZVP) <sup>68</sup>. <sup>f</sup> (B3LYP, 6-31++G(d,p), Watchers + funcions difoses per Cu, correcció BSSE) <sup>66</sup>. <sup>g</sup> (MP2, 6-31G\* for h, C, N, O; Watchers per Cu) <sup>59</sup>.

### *Catió plata*

L'únic exemple conegut de determinació experimental de la basicitat de sistemes  $[\text{Ag}(\text{AA})]^+$  ha estat realitzat pel grup de Siu.<sup>69</sup> Aquesta determinació s'ha fet emprant el mètode cinètic per a tots els aminoàcids excepte la cisteïna. D'altra banda, el grup de Hopkinson<sup>18</sup> ha deduït  $\Delta H^\circ$  i  $\Delta G^\circ$  amb el mateix mètode de càlcul que s'ha utilitzat en aquest treball: B3LYP/DZVP, per als vint aminoàcids habituals. Els modes de coordinació més estables són, el de tipus **A** per als aminoàcids didentats: glicina, alanina, valina, isoleucina i leucina, el de tipus **D** per a la prolina i el de tipus **I** per a la resta d'aminoàcids tridentats. Per tal de poder realitzar una anàlisi completa i comparable amb la resta dels cations monovalents estudiats, s'han refet els càlculs de Hopkinson; obtenint amb alguns aminoàcids valors lleugerament diferents, degut a què s'han trobat algunes conformacions dels aminoàcids lliures que són més estables que les suposades per aquell grup. La Taula 4.5 mostra els valors de  $\Delta H^\circ$  i  $\Delta G^\circ$  calculats, així com els valors determinats experimentalment per Siu i els calculats per Hopkinson. Donat que els valors determinats per Siu són relatius, s'ha utilitzat el valor del  $[\text{AgGly}]^+$  aquí trobat per convertir-los en valors absoluts. Com es pot comprovar, la tendència obtinguda amb els valors experimentals i els aquí calculats és la mateixa, amb diferències numèriques molt petites. La Figura 4.11 mostra la representació gràfica dels valors de basicitat calculats enfront dels experimentals per als compostos  $[\text{Ag}(\text{AA})]^+$ .

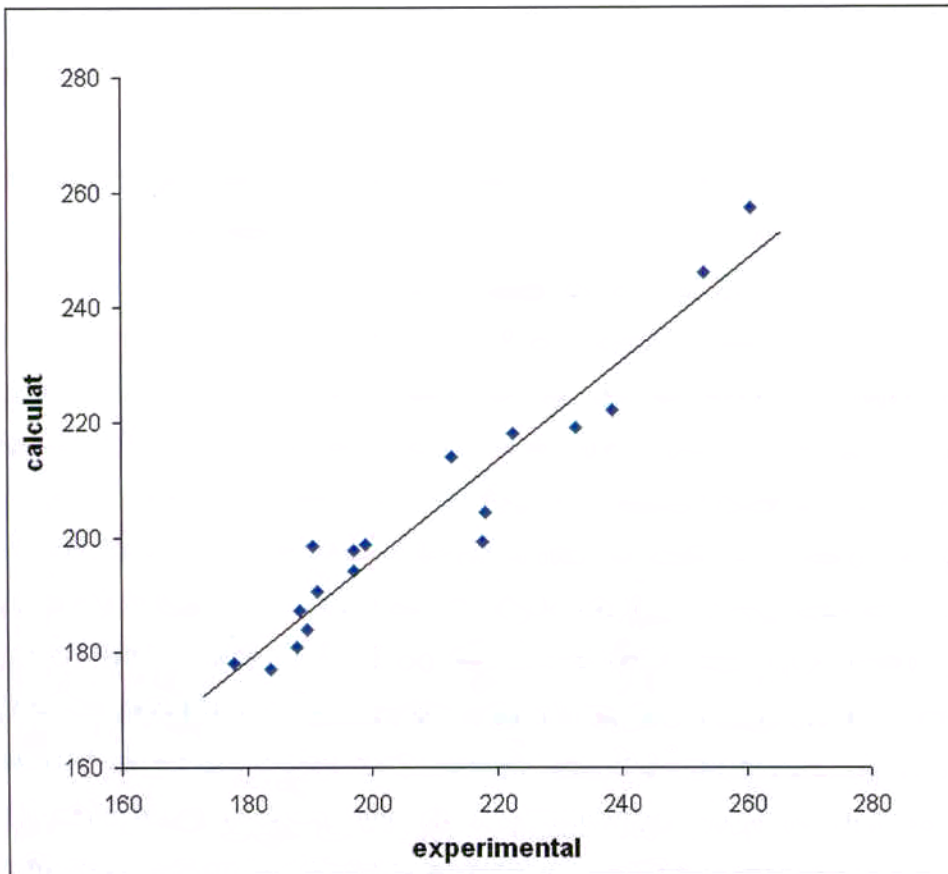


Figura 4.11.  $\Delta G^\circ$  calculat vs. experimental per als sistemes  $[\text{Ag}(\text{AA})]^+$ .

Taula 4.5.  $\Delta H^\circ$  i  $\Delta G^\circ$  experimentals i calculats per als compostos  $[\text{Ag}(\text{AA})]^+$  (kJ/mol).

Aminoàcid	$\Delta H^\circ$	$\Delta H^\circ(\text{calc})^a$	$\Delta G^\circ$	$\Delta G^\circ(\text{calc})^a$	$\Delta G^\circ(\text{rel})^b$	$\Delta G^\circ(\text{abs})^c$
Ala	212.8	212.8	177.2	176.8	5.9	183.9
Arg	325.6	336.5	291.3	293.9	>112.0	>290.1
Asn	245.8	250.8	214.0	217.4	34.7	212.8
Asp	232.4	232.4	198.6	199.0	12.5	190.6
Cys	228.2	230.3	193.5	194.4	-	-
Gln	256.7	264.2	218.2	225.7	44.7	222.8
Glu	234.1	239.9	197.7	203.1	19.2	197.3
Gly	206.1	206.1	178.1	170.1	0.0	178.1
His	281.7	284.2	245.8	249.1	75.2	253.3
Ile	219.0	221.1	183.5	188.9	11.7	189.8
Leu	221.1	219.5	187.3	185.6	10.5	188.5
Lys	290.9	296.8	257.1	260.8	82.8	260.8
Met	260.0	262.1	219.0	219.9	54.8	232.8
Phe	236.2	236.2	199.0	198.6	39.7	217.8
Pro	230.3	234.5	199.0	199.8	20.9	199.0
Ser	224.9	224.5	190.6	190.2	13.4	191.4
Thr	227.8	233.2	194.0	199.8	19.2	197.3
Trp	260.0	260.0	222.0	221.5	60.6	238.7
Tyr	240.4	239.9	204.4	202.3	40.1	218.2
Val	216.5	216.1	180.6	181.0	10.0	188.1

<sup>a</sup> B3LYP/DZVP, correcció BSSE <sup>18</sup>. <sup>b</sup> Valors relatius a Gly; mètode cinètic <sup>69</sup>. <sup>c</sup> Valors absoluts, recalculats de la Ref. <sup>69</sup> utilitzant el valor aquí calculat de glicina/Ag<sup>+</sup> com a àncora.

### Comparació dels valors obtinguts amb els diferents cations

Per tal de racionalitzar els resultats obtinguts s'ha seguit l'esquema desenvolupat per Álvarez i col·laboradors,<sup>70</sup> amb el qual es descompon l'energia de reacció en diferents contribucions. En el cas d'aquest treball, s'ha descompost l'energia de reacció  $E_{\text{tot}}$  corresponent a la reacció global en les dues contribucions:  $E_{\text{int}}$  i  $E_{\text{pre}}$ , que es mostren en la Figura 4.12.  $E_{\text{int}}$  és l'energia d'interacció entre el catió i l'aminoàcid, aquest paràmetre es calcula com la diferència d'energies entre el sistema  $[M(\text{AA})]^+$  i la suma d'energies del catió i de l'aminoàcid lliure en la mateixa conformació que quan està coordinat. El terme  $E_{\text{pre}}$  és l'energia de preorganització de l'aminoàcid, que es calcula com la diferència d'energies entre l'aminoàcid lliure i l'aminoàcid en la conformació en la que es troba coordinat al metall. La Taula 4.6 mostra els valors de les energies  $E_{\text{tot}}$ ,  $E_{\text{int}}$  i  $E_{\text{pre}}$  per a tots els sistemes  $[M(\text{AA})]^+$  estudiats; en tots els casos,  $E_{\text{pre}}$  és positiva ja que l'energia de l'aminoàcid coordinat és més alta que la de l'aminoàcid lliure; mentre que  $E_{\text{int}}$  és negativa donat que la interacció entre el catió metàl·lic i l'aminoàcid sempre és d'atracció.

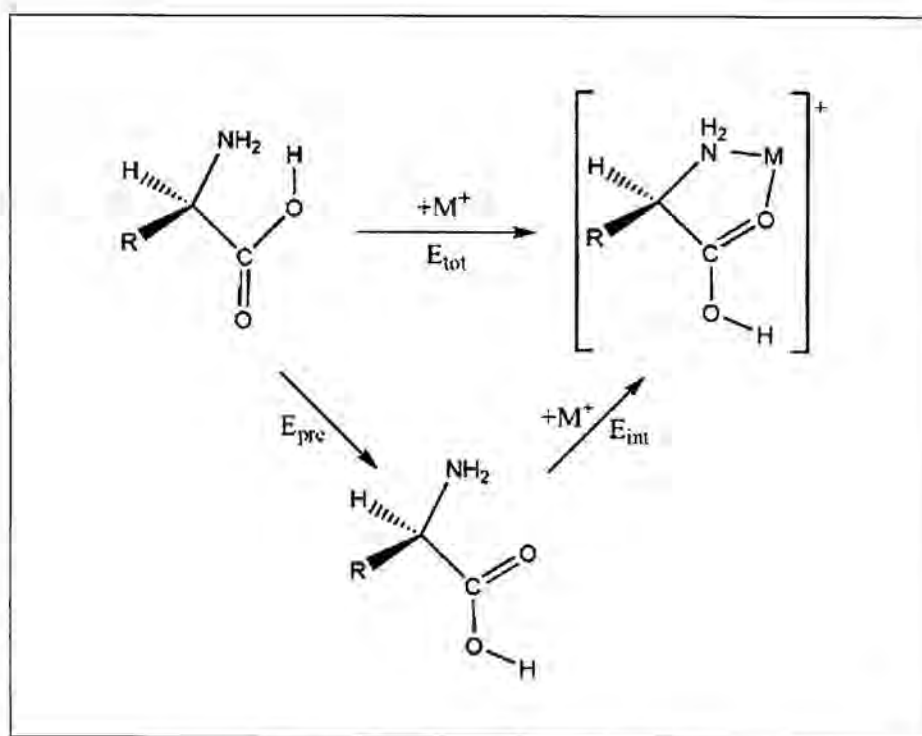


Figura 4.12. Descomposició de l'energia de reacció.



Taula 4.6. Descomposició de les energies de reacció per als compostos  $[M(AA)]^+$  (kJ/mol).

	Li <sup>+</sup>			Na <sup>+</sup>			K <sup>+</sup>			Cu <sup>+</sup>			Ag <sup>+</sup>		
	E <sub>tot</sub>	E <sub>pre</sub>	E <sub>int</sub>	E <sub>tot</sub>	E <sub>pre</sub>	E <sub>int</sub>	E <sub>tot</sub>	E <sub>pre</sub>	E <sub>int</sub>	E <sub>tot</sub>	E <sub>pre</sub>	E <sub>int</sub>	E <sub>tot</sub>	E <sub>pre</sub>	E <sub>int</sub>
Ala	-245.9	32.5	-278.4	-179.7	25.3	-205.0	-122.0	21.7	-143.6	-312.9	30.1	-343.0	-216.9	24.7	-241.6
Arg	-359.9	85.0	-444.8	-266.3	70.0	-336.2	-171.7	61.6	-233.2	-428.0	79.8	-507.9	-326.6	81.2	-407.8
Asn	-307.4	84.1	-391.5	-226.5	69.5	-296.0	-151.0	55.7	-206.7	-351.2	78.9	-430.1	-247.3	68.4	-315.7
Asp	-293.5	63.0	-356.4	-216.8	50.0	-266.7	-144.8	38.2	-183.0	-336.3	57.1	-393.4	-238.0	48.8	-286.7
Cys	-252.8	52.6	-305.4	-185.3	45.1	-230.4	-117.8	39.9	-157.8	-332.2	49.6	-381.8	-231.4	42.2	-273.6
Gln	-320.4	86.7	-407.1	-231.1	76.7	-307.8	-154.8	62.7	-217.6	-377.1	81.6	-458.6	-256.2	70.7	-326.9
Glu	-288.0	85.7	-373.6	-214.3	60.8	-275.1	-139.0	53.1	-192.2	-339.9	78.2	-418.0	-236.8	59.0	-296.7
Gly	-239.0	32.9	-271.8	-173.7	28.6	-202.3	-114.5	22.0	-136.6	-304.5	30.9	-335.4	-210.3	25.2	-235.5
His	-329.3	79.3	-409.0	-241.3	67.8	-309.1	-154.3	62.1	-216.4	-415.8	56.7	-472.3	-285.5	64.7	-350.1
Ile	-254.1	35.3	-289.4	-180.9	26.8	-207.7	-123.5	8.2	-131.7	-321.7	33.3	-354.9	-223.0	30.2	-253.2
Leu	-254.4	32.6	-287.0	-186.3	26.0	-212.3	-122.4	22.5	-144.9	-322.9	30.5	-353.4	-224.8	25.5	-250.3
Lys	-320.7	80.6	-401.2	-235.8	61.0	-296.9	-152.5	48.9	-201.4	-446.4	71.2	-517.5	-293.9	61.4	-355.3
Met	-276.0	58.1	-334.1	-204.0	46.8	-250.8	-130.5	40.3	-170.8	-374.8	57.1	-431.8	-267.0	43.6	-310.6
Phe	-269.7	44.3	-314.0	-204.4	37.9	-242.3	-138.4	47.4	-185.9	-348.0	48.8	-396.8	-240.3	39.0	-279.3
Pro	-264.1	72.1	-336.2	-207.2	64.9	-272.1	-141.0	65.7	-206.7	-321.2	33.4	-354.6	-236.2	67.6	-303.8
Ser	-273.8	68.2	-342.0	-203.4	56.2	-259.5	-133.5	49.0	-182.4	-317.3	60.7	-378.0	-227.0	53.5	-280.5
Thr	-278.9	70.4	-349.3	-206.3	58.0	-264.3	-135.0	50.1	-185.1	-322.0	63.5	-385.6	-230.2	55.6	-285.7
Trp	-289.2	52.6	-341.8	-219.0	44.2	-263.1	-147.4	39.0	-186.3	-379.5	60.6	-440.2	-256.9	46.6	-303.4
Tyr	-269.5	47.3	-316.9	-203.8	41.0	-244.8	-136.7	36.5	-173.3	-352.2	50.7	-402.9	-242.7	38.5	-281.2
Val	-251.1	34.7	-285.8	-183.4	26.9	-210.4	-121.2	4.8	-126.0	-318.8	33.4	-352.2	-220.6	27.6	-248.2

\* Els valors d'energia representats en la taula per glicina, alanina, isoleucina, leucina i valina amb potassi corresponen a estructures de tipus C. Els valors de  $E_{\text{tot}}$ ,  $E_{\text{pre}}$  i  $E_{\text{int}}$  per les geometries de tipus A d'aquests sistemes són: Gly: -114.49, 22.11, -135.34 kJ/mol; Ala: -117.92, 21.65, -139.57 kJ/mol; Ile: -117.29, 27.67, -143.71 kJ/mol; Leu: -122.43, 22.49, -144.92 kJ/mol; Val: -119.76, 24.54, -144.29 kJ/mol, respectivament.



L'energia d'interacció per als aminoàcids que no presenten cap heteroàtom ni sistema  $\pi$  en la cadena lateral: glicina, alanina, valina, isoleucina i leucina, es troba restringida en intervals petits: -271.7/-289.2 kJ/mol per al  $\text{Li}^+$ , -202.3/-212.3 kJ/mol per al  $\text{Na}^+$ , -124.1/-132.5 kJ/mol per al  $\text{K}^+$ , -335.2/-359.4 kJ/mol per al  $\text{Cu}^+$  i -235.3/-252.9 kJ/mol per al  $\text{Ag}^+$ . L'ordre de les energies d'interacció, en valor absolut, és:  $\text{K}^+ < \text{Na}^+ < \text{Ag}^+ < \text{Li}^+ < \text{Cu}^+$ , equivalent a l'ordre creixent de caràcter polaritzant d'aquests cations. Abans de poder fer aquesta afirmació, cal tenir en compte que la naturalesa dels compostos de potassi amb aquests aminoàcids és diferent de la resta: el complexos de potassi amb aquests aminoàcids presenten estructures de tipus **C** mentre que la resta de cations presenten estructures de tipus **A**. De tota manera, si s'analitzen les estructures de tipus **A** per als sistemes de potassi amb aquests aminoàcids (des de -136.7 a -145.0 kJ/mol), l'ordre d'energia d'interacció no varia. D'altra banda, l'increment en el nombre de carbonis en la cadena lateral implica un augment en l'energia d'interacció, degut a l'efecte inductiu.<sup>10</sup> L'energia de preorganització d'aquests mateixos aminoàcids es troba, també, restringida en intervals relativament petits: 32.5/35.3 kJ/mol per al  $\text{Li}^+$ , 25.3/28.6 kJ/mol per al  $\text{Na}^+$ , 30.1/33.4 kJ/mol per al  $\text{Cu}^+$  i 24.7/30.2 kJ/mol per al  $\text{Ag}^+$ . El cas del potassi és més complex de racionalitzar degut al diferent mode de coordinació que presenten els aminoàcids amb aquest catió: de tipus **C**, on lògicament l'energia de preorganització és menor (al voltant dels 6 kJ/mol). Per tal d'obtenir comparacions significatives s'ha actuat de la mateixa manera que s'ha fet abans, a partir dels càlculs d'aquests compostos  $[\text{K}(\text{AA})]^+$  amb mode de coordinació **A**. D'aquesta manera, el rang de l'energia de preorganització dels sistemes de potassi és 21.7/22.5 kJ/mol; i l'ordre resultant d'aquesta energia, per als sistemes de tipus **A**, és:  $\text{K}^+ < \text{Ag}^+ < \text{Na}^+ < \text{Cu}^+ < \text{Li}^+$ , suggerint una relació inversa amb el radi catiónic. Tsang i col·laboradors<sup>40</sup> van arribar a una conclusió equivalent en un estudi dels sistemes  $[\text{MAla}]^+$  amb liti, sodi i potassi, obtenint unes energies de preorganització de 35, 26 i 5 kJ/mol, respectivament.

La presència d'un heteroàtom o d'un sistema  $\pi$  en la cadena lateral dels aminoàcids té, principalment, dos efectes. El primer consisteix a fer augmentar significativament el valor de l'energia de preorganització; lògicament, els requeriments geomètrics són molt més importants en aquesta ocasió ja que l'aminoàcid s'ha de col·locar en la conformació idònia per fer interactuar el grup extra amb el catió. Per exemple, en el cas del  $\text{Li}^+$  l'energia de preorganització passa de 32.5/35.3 kJ/mol en els sistemes didentats a 52.6/86.7 kJ/mol en els sistemes tridentats. Quan la cadena lateral

conté un sistema  $\pi$ , l'energia de preorganització també augmenta, però no de manera tan marcada. El segon efecte que produeix l'heteroàtom addicional és l'augment, en valor absolut, de l'energia d'interacció. El canvi en aquesta magnitud, que es recull en la Taula 4.7, ve determinat alhora pel tipus d'heteroàtom i el tipus de catió implicats. Els valors d'aquesta taula han estat calculats com la diferència entre el complex tridentat de tipus **I** i el compost de geometria **A** de la valina amb un mateix catió. La raó per triar aquest aminoàcid és que el nombre d'àtoms de carboni en la cadena lateral, i per tant l'efecte inductiu d'aquesta, és la mitjana de la sèrie Gly-Ala-Val-Leu-Ile; qualsevol cadena lateral interacciona amb el catió metàl·lic en funció de les característiques electrostàtiques d'aquesta, sent les més importants les que es produeixen per fenòmens de polaritat en les cadenes laterals que presenten moment dipolar important.<sup>44</sup> De tota manera, la influència que exerceixen els carbonis de la cadena lateral en el valor d'energia d'interacció és sempre molt més petita que aquella que produeixen els heteroàtoms. Els aminoàcids de referència triats per estimar l'augment de les energies d'interacció de cada catió són: lisina per al nitrogen, serina per a l'oxigen  $sp^3$ , àcid glutàmic per a l'oxigen  $sp^2$ , metionina per al sofre  $sp^3$  i fenilalanina per al sistema  $\pi$ .

Taula 4.7. Contribució dels àtoms donadors de la cadena lateral a  $E_{int}$  (kJ/mol).

	$Li^+$	$Na^+$	$K^+$	$Cu^+$	$Ag^+$
N	115.4	86.5	57.1	165.3	107.1
O ( $sp^3$ )	56.2	49.1	38.1	25.8	32.3
O ( $sp^2$ )	87.8	64.7	47.9	65.8	48.5
S ( $sp^3$ )	48.3	40.4	26.5	79.6	62.4
Sistema $\pi$	28.2	31.9	41.6	44.6	31.1

Per a tots els cations estudiats en aquest treball, el nitrogen és l'heteroàtom que, quan és present en la cadena lateral produeix el major augment en l'energia d'interacció. En tots els casos aquest heteroàtom contribueix aproximadament en un 30% del valor total de l'energia d'interacció.

L'oxigen carbonílic interactua més fortament amb els cations que l'oxigen hidroxílic. És interessant comparar els aminoàcids que presenten el mateix tipus d'heteroàtom en la cadena lateral; per exemple, la serina i la treonina presenten un oxigen hidroxílic en la cadena lateral, que té la mateixa llargada per a tots dos; tot i així, l'energia d'interacció de la treonina amb tots els cations és un 2% més gran que la de la serina. Aquest fet és degut, probablement, al major efecte inductiu que presenta la

cadena lateral de la treonina degut al seu grup metílic.<sup>10</sup> Una altra comparació d'aquest tipus es pot establir entre l'àcid aspàrtic i l'àcid glutàmic; en aquest cas l'energia d'interacció per a l'àcid glutàmic és sempre un 3-5% més gran que la de l'àcid aspàrtic. Una possible raó d'aquest comportament és la major llargada de cadena de l'àcid glutàmic, que permet una millor interacció entre els tres heteroàtoms i el catió.

L'increment en l'energia d'interacció produït pels àtoms de sofre en la cadena lateral és sempre major per a la metionina que per a la cisteïna, aquest comportament és anàleg al que ja s'ha comentat entre la treonina i la serina. Aquest tipus d'àtom produeix, a més, una gran diferència entre els metalls alcalins i el coure i la plata, sent l'increment en l'energia d'interacció molt més gran per a aquests últims; aquest fenomen no és gens estrany ja que el sofre és un element tou a l'igual que els cations coure i plata monovalents. Així doncs, l'ordre en l'increment d'energia d'interacció dels cations alcalins en funció de l'heteroàtom present en la cadena lateral queda de la següent manera:  $N(sp^3) > O(sp^2) > O(sp^3) > S(sp^3)$ , mentre que per a coure i plata és:  $N(sp^3) > S(sp^3) > O(sp^2) > O(sp^3)$ .

La coordinació del catió metàl·lic amb el sistema  $\pi$  de la cadena lateral és l'últim factor que s'ha considerat. Els estudis de Ryzhov<sup>44</sup> de l'efecte del sistema  $\pi$  en la coordinació de fenilalanina, tirosina i triptòfan amb liti, sodi i potassi quantifiquen l'increment en l'energia d'interacció en uns 29/33 kJ/mol per als dos primers aminoàcids i en uns 40/50 kJ/mol per al triptòfan, coincidint amb els resultats aquí obtinguts. Altres treballs en els quals es calcula l'afinitat dels metalls alcalins amb el toluè<sup>71</sup> proposen que el valor de l'energia d'interacció decreix en augmentar el radi iònic del catió i han relacionat aquest comportament amb el tipus d'interaccions, principalment ió-dipol, ió-dipol induït i ió-quadrupol, que es produeixen en els sistemes estudiats. En el cas dels aminoàcids és més difícil establir aquest tipus de relacions ja que, en aquest cas, s'ha de tenir en compte l'efecte dels heteroàtoms presents i els possibles enllaços d'hidrogen intramoleculars que es formen. Així, l'ordre en l'increment d'energia d'interacció per coordinació amb el sistema  $\pi$  dels aminoàcids és el següent:  $Li^+ < Ag^+ < Na^+ < K^+ < Cu^+$ ; en el cas del potassi, aquest tipus d'interacció és més forta que amb els heteroàtoms  $O(sp^3)$  i  $S(sp^3)$ .

Aquesta metodologia es pot utilitzar també per explorar les estabilitats relatives dels diferents modes de coordinació. Aquest estudi però, s'hauria de fer per a cadascun dels sistemes  $[M(AA)]^+$ . Donat que aquest procediment seria massa llarg s'ha optat per comparar només l'estabilitat relativa entre els modes de coordinació **A-D** dels sistemes més senzills:  $[M(\text{Gly})]^+$ . La Taula 4.8 mostra els valors d'energia de preorganització i d'interacció calculats per als modes de coordinació d'aquest sistema.

Taula 4.8. Valors de  $E_{\text{pre}}$  i  $E_{\text{int}}$  dels compostos  $[M\text{Gly}]^+$  (kJ/mol).

	Tipus A		Tipus B		Tipus C		Tipus D	
	$E_{\text{pre}}$	$E_{\text{int}}$	$E_{\text{pre}}$	$E_{\text{int}}$	$E_{\text{pre}}$	$E_{\text{int}}$	$E_{\text{pre}}$	$E_{\text{int}}$
<b>Li<sup>+</sup></b>	32.9	-271.9	29.1	-227.0	11.2	-218.9	98.8	-316.6
<b>Na<sup>+</sup></b>	28.6	-202.3	24.1	-160.2	9.4	-169.4	91.6	-256.5
<b>K<sup>+</sup></b>	22.0	-136.6	22.0	-102.2	6.7	-124.1	85.6	-191.0
<b>Cu<sup>+</sup></b>	30.9	-335.4	27.4	-286.1	14.7	-275.8	85.6	-359.4
<b>Ag<sup>+</sup></b>	25.2	-235.5	23.2	-193.4	10.4	-194.8	94.2	-285.7

En tots els casos l'ordre en el valor de l'energia de preorganització és: **C** < **B** < **A** << **D**; lògicament, el valor més baix en aquest paràmetre correspon al mode de coordinació **C** ja que d'aquesta manera es manté l'enllaç d'hidrogen intramolecular entre l'hidrogen hidroxílic i el nitrogen; mentre que el mode de coordinació **D** presenta el valor més elevat ja que conté l'energia necessària per passar de la forma neutra a la zwitteriònica en fase gas. Es pot comprovar que amb el potassi els modes de coordinació **A** i **B** tenen idèntica energia de preorganització mentre que el mode **C** presenta un valor més baix que la resta de cations; Tsang<sup>57</sup> va arribar a la mateixa conclusió: d'acord amb els seus càlculs per al sistema  $[\text{KGly}]^+$ , les energies de preorganització en els modes de coordinació **A**, **B**, **C** i **D** són 24.1, 24.0, 7.7 i 86.2 kJ/mol, respectivament.

El mode de coordinació de tipus **D** presenta els valors més grans, en valor absolut, d'energia d'interacció, seguit pel de tipus **A**. En el cas dels compostos de tipus **D**, la presència d'una càrrega negativa sobre el grup carboxilat augmenta de manera considerable la capacitat de l'aminoàcid per interaccionar amb el catió. La resta de modes de coordinació presenten valors de  $E_{\text{int}}$  menors i similars entre sí, encara que els diferents cations presenten diferents ordres: mentre que amb liti i coure la geometria de tipus **C** té un valor més gran que la **B**; amb sodi, potassi i plata aquesta tendència és la inversa.

La preferència d'un mode de coordinació sobre un altre és el resultat de l'equilibri entre les energies de preorganització i d'interacció. Per exemple, encara que en el tipus de coordinació **D** el valor de  $E_{\text{int}}$  és el més gran, no és suficient per compensar l'energia de preorganització necessària per formar el zwitterió. Així, el mode de coordinació **A**, que presenta el segon valor més alt d'energia d'interacció i un valor d'energia de preorganització relativament petit, es converteix en l'opció preferida per a tots els aminoàcids excepte el potassi. En el cas d'aquest catió, la combinació de les energies de preorganització i d'interacció determinen que la geometria preferida serà la de tipus **C**. Cal remarcar que, en el cas dels cations alcalins, quan augmenta el radi iònic la diferència entre les energies d'interacció entre **A** i **C** disminueix molt més de pressa del que augmenta l'energia de preorganització, fent així que el mode de coordinació **C** sigui el preferit quan augmenta la grandària del catió metàl·lic.



### 4.1.2 Afinitat catiònica amb la metodologia QSPR

Com ja s'ha comentat anteriorment, en aquest apartat s'utilitzarà la metodologia QSPR per establir models que permetin el càlcul de l'afinitat dels vint aminoàcids habituals front els cinc cations metàl·lics monovalents de liti, sodi, potassi, coure i plata. Només s'han desenvolupat models per calcular l'afinitat catiònica ja que, aquesta propietat i la basicitat són molt semblants i estan molt lligades entre sí; així, seria d'esperar que els models obtinguts per a totes dues serien anàlegs. Una possible manera de realitzar aquest pas seria treballar a partir de les geometries optimitzades en l'apartat anterior de tots els sistemes  $[M(AA)]^+$ , calculant els descriptors amb CODESSA i establint, posteriorment, els models lineals i no lineals corresponents. Tot i ser adequat, aquest procediment és equivalent al que s'ha desenvolupat en l'apartat anterior, amb els confòrmers més estables per a tots els sistemes estudiats no caldria establir models per calcular l'afinitat i la basicitat ja que aquestes propietats es poden extreure fàcilment un cop es coneix l'estructura molecular dels compostos. El que es pretén fer ara és desenvolupar models que permetin el càlcul de l'afinitat dels sistemes  $[M(AA)]^+$  a partir de l'estructura molecular dels aminoàcids lliures combinada amb alguna de les característiques pròpies dels metalls amb què es coordinen. Això vol dir que els models han de contenir descriptors moleculars de l'aminoàcid, en la seva conformació més estable i descriptors externs que caracteritzin d'alguna manera les propietats dels cations metàl·lics. La combinació dels descriptors dels dos elements estudiats ha de conduir a l'establiment de models amb els quals sigui possible el càlcul de les propietats d'interès. Aquest procediment presenta com a principal avantatge, en relació a l'anterior, la seva rapidesa i senzillesa ja que només cal optimitzar l'estructura dels aminoàcids lliures. El fet de treballar seguint aquesta metodologia converteix aquest conjunt de compostos, fruit de la combinació dels aminoàcids amb els metalls, en un sistema multicomponent. Així doncs, tots els models desenvolupats per calcular l'afinitat seran de tipus no lineal, ja que és d'esperar que els models derivats per regressió multilineal no siguin capaços d'assumir les complexes interaccions que es produeixen, com els diferents tipus de coordinació dels cations amb un determinat aminoàcid

El mètode utilitzat per establir aquests models no lineals serà el proporcionat pel programa ADAPT: tria de descriptors mitjançant algorismes genètics i ús de les xarxes neuronals de tipus BPNN com a funcions avaluadores dels models construïts. Els



models no lineals han de contenir descriptors dels dos components del sistema estudiat, així, els aminoàcids seran tractats de la següent manera: es prendran les geometries optimitzades trobades en l'apartat anterior i, mitjançant el mètode AM1 es farà un càlcul *single point* que proporciona el fitxer necessari per tal de calcular els descriptors amb el programa CODESSA. Per la seva part, la caracterització dels cations metàl·lics es fa amb els mateixos descriptors externs ja utilitzats en el càlcul de les entalpies de formació dels compostos organometàl·lics, Taula 3.42.<sup>72</sup>

### **Conjunt de dades**

Els valors triats de d'afinitat ( $\Delta H^\circ$ ) per establir els models són els que s'han determinat mitjançant els càlculs teòrics en aquest treball.

Donat que es volen establir models no lineals amb les xarxes neuronals BPNN incorporades en el programa ADAPT, s'ha de dividir el conjunt de dades en *tset*, *pset* i *cvset*. Aquests nous subconjunts de dades contenen 70, 20 i 10 valors, respectivament. Aquestes particions del conjunt de dades s'han fet de manera que en cadascun dels subconjunts es mantingui la proporció inicial del conjunt de dades, per exemple, el *pset* conté 20 valors corresponent als vint aminoàcids estudiats, quatre de cada metall. El *cvset* conté 10 valors, dos de cada metall, també sense repetició de cap aminoàcid. La Taula 4.9 conté els valors calculats mitjançant els càlculs teòrics i els calculats amb els models desenvolupats amb la metodologia QSPR.

### **Model no lineal: Xarxes neuronals BPNN**

El nombre total de descriptors calculats amb CODESSA per al conjunt dels 20 aminoàcids és de 490. Afegint els 31 descriptors externs propis dels metalls, comentats en l'apartat 3.3, s'obtenen els 521 descriptors definitius inicials. La reducció de descriptors realitzada amb el mètode heurístic implementat en CODESSA és capaç de reduir-lo a 177 descriptors, 158 dels aminoàcids i 19 dels metalls. Com que aquest nombre és massa elevat per permetre utilitzar les xarxes neuronals BPNN d'ADAPT, es fa una segona reducció en el nombre de descriptors amb les rutines de selecció objectiva que implementa aquest programa. D'aquesta manera, el nombre de descriptors es veu reduït a 73, dels quals 66 són dels aminoàcids i 7 dels metalls.

L'arquitectura seleccionada per a les xarxes neuronals construïdes per a aquest sistema és 7-3-1. D'aquesta manera, el nombre de paràmetres ajustables és de 28 i el *ratio* de valors del *tset* sobre aquest nombre és 2.5, per sobre del 2.0 mínim recomanable. Els descriptors continguts en el model es troben en la Taula 4.10.

Taula 4.10 Descriptors continguts en el model BPNN.

Descriptor	Component	Classe
Nombre d'heteròtoms en la cadena lateral	Aminoàcid	Constitucional
Àrea superficial relativa carregada positivament, RPCS	Aminoàcid	CPSA
Atracció electró-nucli mínima d'un àtom d'hidrogen	Aminoàcid	Quàntic
Índex màxim de reacció nucleòfila d'un àtom de nitrogen	Aminoàcid	Quàntic
Índex màxim de reacció nucleòfila d'un àtom d'oxigen	Aminoàcid	Quàntic
Relació càrrega/radi iònic de Shannon	Metall	-
Electronegativitat de Pauling	Metall	-

Aquest model conté cinc descriptors dels aminoàcids, un constitucional: el nombre d'heteroàtoms en la cadena lateral, un electrostàtic de tipus CPSA: RPCS, i tres quàntics: els índex màxims de reacció nucleòfila d'un àtom de nitrogen i d'un àtom d'oxigen i l'atracció electró-nucli mínima d'un àtom d'hidrogen; i dos dels cations metàl·lics: la relació càrrega/radi iònic de Shannon i l'electronegativitat de Pauling.

El nombre d'heteroàtoms en la cadena lateral es correspon simplement amb la suma d'àtoms d'oxigen, nitrogen i sofre presents en aquesta la cadena. Aquest descriptor es pot relacionar amb la capacitat de l'aminoàcid per formar compostos tridentats amb el catió metàl·lic

El descriptor RPCS, es calcula de la manera següent:

$$RPCS = S_{max} \frac{\delta_{max}}{\sum_A \delta_A}$$

on  $S_{max}$  és la superfície atòmica amb major càrrega positiva,  $\delta_{max}$  és la càrrega atòmica positiva parcial màxima en la molècula, i  $\delta_A$  és la càrrega atòmica positiva parcial de cada àtom. Les càrregues parcials utilitzades per calcular aquest descriptor han estat derivades amb el mètode de Zefirov.<sup>73</sup> Aquest descriptor es relaciona clarament amb les interaccions de tipus polar que poden establir els aminoàcids, per exemple de tipus ió-dipol, amb el catió metàl·lic.

Els índex màxims de reactivitat nucleòfila dels àtoms de nitrogen i oxigen, derivats de la teoria de Fukui,<sup>74</sup> es calculen:

$$N_A = \sum_{i \in A} c_{iHOMO}^2 / (1 - \epsilon_{HOMO})$$

on  $c_{iHOMO}$  són els coeficients atòmics d'aquests elements en l'orbital HOMO i  $\epsilon_{HOMO}$  és l'energia d'aquest mateix orbital molecular. Aquests dos descriptors, que es poden relacionar amb la capacitat que presenta l'aminoàcid per enllaçar-se al metall, es troben situats majoritàriament sobre el nitrogen amínic i l'oxigen carbonílic de l'aminoàcid. Aquests dos àtoms, tal i com s'ha comprovat amb el model teòric desenvolupat, s'uneixen en tots els casos, excepte amb el potassi, als metalls per formar compostos quelats. Cal recordar que els compostos de potassi que no presenten cap heteroàtom en la cadena lateral se solen coordinar amb els cations a través dels dos oxígens carboxílics (Tipus C), tot i així, cal dir que la diferència d'energia entre aquesta conformació i la que implica un enllaç amb els grups carbonil i amino de l'aminoàcid (Tipus A) es troba a menys de 4 kJ de diferència. En el cas on aquests descriptors no es troben sobre aquests dos àtoms, s'ha comprovat que els valors seleccionats són molt pròxims als dels àtoms habituals i el seu intercanvi no produeix variacions significatives en el model.

El descriptor atracció electró-nucli mínima d'un àtom d'hidrogen se situa sempre sobre l'hidrogen del grup carboxílic de l'aminoàcid, mai sobre els grups àcids continguts en les cadenes laterals de l'àcid glutàmic i l'àcid aspàrtic. Aquest descriptor es pot relacionar amb els enllaços d'hidrogen intramoleculars que es donen en els aminoàcids, que es poden veure modificats quan es produeix la interacció amb els cations metàl·lics.

Els dos descriptors dels metalls inclosos en el model són l'electronegativitat de Pauling i la relació càrrega/radi iònic de Shannon.<sup>75</sup> Tots dos es poden relacionar amb la naturalesa de l'enllaç que es forma en els compostos  $[M(AA)]^+$ . La diferència d'electronegativitat entre el catió metàl·lic i els àtoms donadors és indicativa de la força de l'enllaç que es pot establir i de les interaccions ió-dipol que es poden donar. La relació càrrega/radi és una mesura del caràcter polaritzant i de la duresa del catió metàl·lic. Aquest descriptor es pot relacionar amb la intensitat de les interaccions dels metalls amb els diferents heteroàtoms de l'aminoàcid.

Els resultats que s'obtenen per als tres conjunts de dades: *tset*, *cvset* i *pset*, amb aquest model són molt bons, Taula 4.11. Els valors d'afinitat calculats amb aquest model es troben en la Taula 4.9 i la seva representació gràfica enfront dels valors experimentals es pot veure en la Figura 4.13.

Taula 4.11. Paràmetres estadístics obtinguts amb el model BPNN.

Conjunt	n	R <sup>2</sup>	F	rmse
<i>tset</i>	70	0.998	42780.8	3.16
<i>cvset</i>	10	0.999	5563.5	3.13
<i>pset</i>	20	0.999	25350.2	2.08

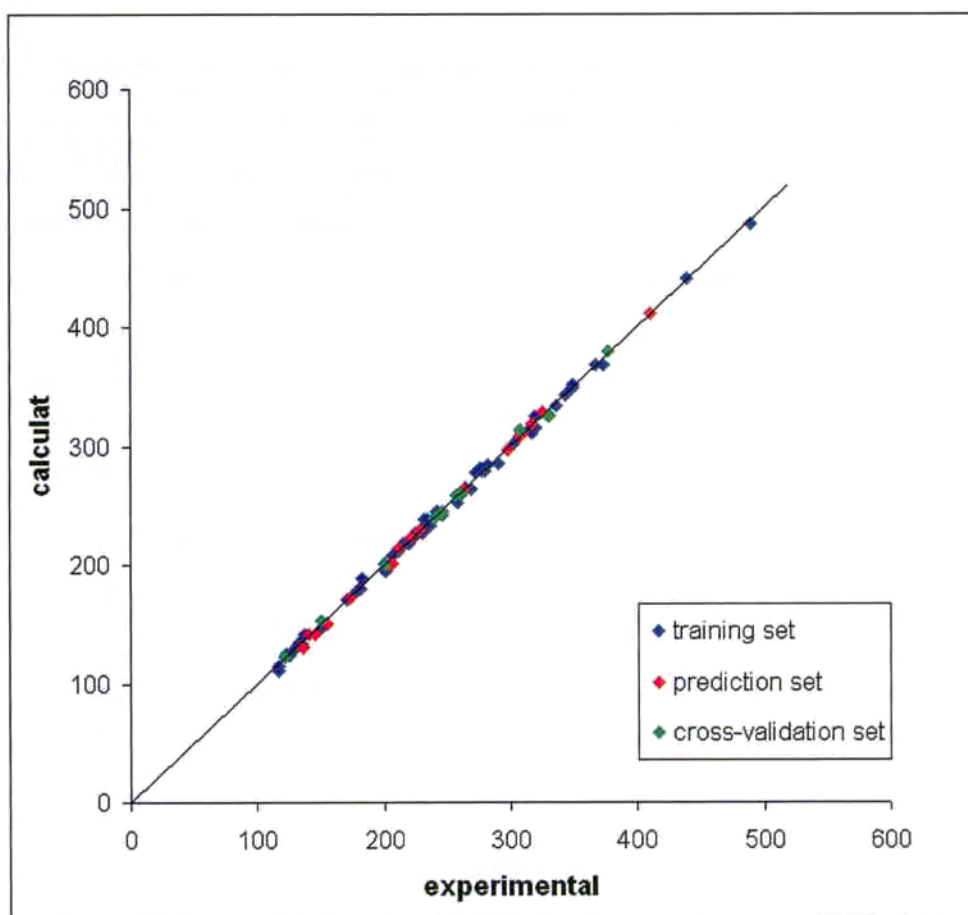


Figura 4.13. Representació gràfica de  $\Delta H^\circ$  calculat vs. experimental amb BPNN.

Donat que el nombre d'observacions és relativament baix, només es tenen 100 valors de la propietat, s'ha utilitzat un mètode de validació extra per tal d'assegurar que el model trobat funciona correctament. Aquest mètode consisteix en fer múltiples *training*, *cross-validation* i *prediction sets*, establir el model corresponent amb els descriptors de la Taula 4.10 i avaluar els seus resultats, comprovant que els valors mitjans de coeficient de determinació  $R^2$  i d'error *rmse* que es troben amb els nous subconjunts no varien de forma significativa respecte els que es troben en la Taula 4.11. S'han fet cinc subdivisions diferents del conjunt de dades, construint *psets* que contenen, en grups de 20, tots els compostos presents en l'estudi. Els valors mitjans de  $R^2$  i *rmse* obtinguts per a *tset*, *cvset* i *pset* són: 0.998 i 3.25, 0.998 i 3.35, 0.998 i 3.81, respectivament. Aquests resultats confirmen l'estabilitat del model obtingut.

El valor de *rmse* obtingut pel model amb tots els compostos estudiats és 3.0. No hi ha cap compost que presenti un residual tres vegades més gran que aquest valor i només n'hi ha quatre que tenen residuals dues vegades més grans. La Taula 4.12 mostra aquests quatre compostos i els seus residuals. Els compostos que apareixen pertanyen a quatre aminoàcids de naturalesa prou diferent: cisteïna, prolina, glutamina i isoleucina; combinats dos amb el sodi, un amb liti i un amb coure, respectivament. No sembla haver relació entre els errors obtinguts per a aquests compostos i la seva estructura.

Taula 4.12. Prediccions d'afinitat amb valors més grans que dues vegades el *rmse*.

Aminoàcid	Metall	Exp.	Calc.	Residual
Cys	Na	182.2	188.8	-6.6
Gln	Li	307.2	313.4	-6.2
Pro	Na	201.5	195.0	6.5
Ile	Cu	317.3	310.7	6.6

L'anàlisi dels resultats per als diferents subconjunts de metalls es troben en la Taula 4.13. Com es pot comprovar, els millors s'obtenen amb la plata, mentre que la resta de metalls presenten resultats semblants. Sobta potser que els pitjors resultats siguin per als compostos de sodi, sobretot tenint en compte que aquesta família no presenta grans diferències estructurals respecte a la resta. En canvi, la família del potassi, on hi ha variacions estructurals respecte a les altres, dona uns resultats bastant



bons que no semblen veure's afectats per aquest fet. La representació gràfica dels valors experimentals i calculats per a aquestes famílies es troba en la Figura 4.14, on es pot constatar la bona correlació obtinguda en cada cas.

Taula 4.13. Estadística dels subconjunts de metalls.

Metall	n	R <sup>2</sup>	F	rmse
Liti	20	0.989	1560.5	3.41
Sodi	20	0.979	848.9	3.62
Potassi	20	0.962	451.4	2.86
Coure	20	0.996	4073.1	3.32
Plata	20	0.994	3195.9	2.27

En resum, es pot concloure que ha estat possible establir un mètode genèric de càlcul per a l'afinitat i la basicitat dels  $\alpha$ -aminoàcids habituals front els cations metàl·lics seleccionats. Aquest mètode es podria utilitzar, en principi, per calcular aquestes propietats per a altres cations monovalents i fins i tot, per a d'altres cations amb càrregues més elevades, sent l'únic requisit la disponibilitat de funcions de base del metall.

El model desenvolupat amb la metodologia QSPR proporciona resultats molt bons i, a més, està format per descriptors que contenen informació que es pot relacionar amb les interaccions que es donen en els sistemes  $[M(AA)]^+$ , que són determinants per al valor de les dues propietats estudiades.

Cal dir, com a punt final, que els errors que s'obtenen aquí, tant amb els càlculs teòrics com amb el model QSPR, són molt baixos i que són del mateix ordre, o fins i tot menors, que els que es produeixen amb altres mètodes de càlcul i amb les determinacions experimentals.



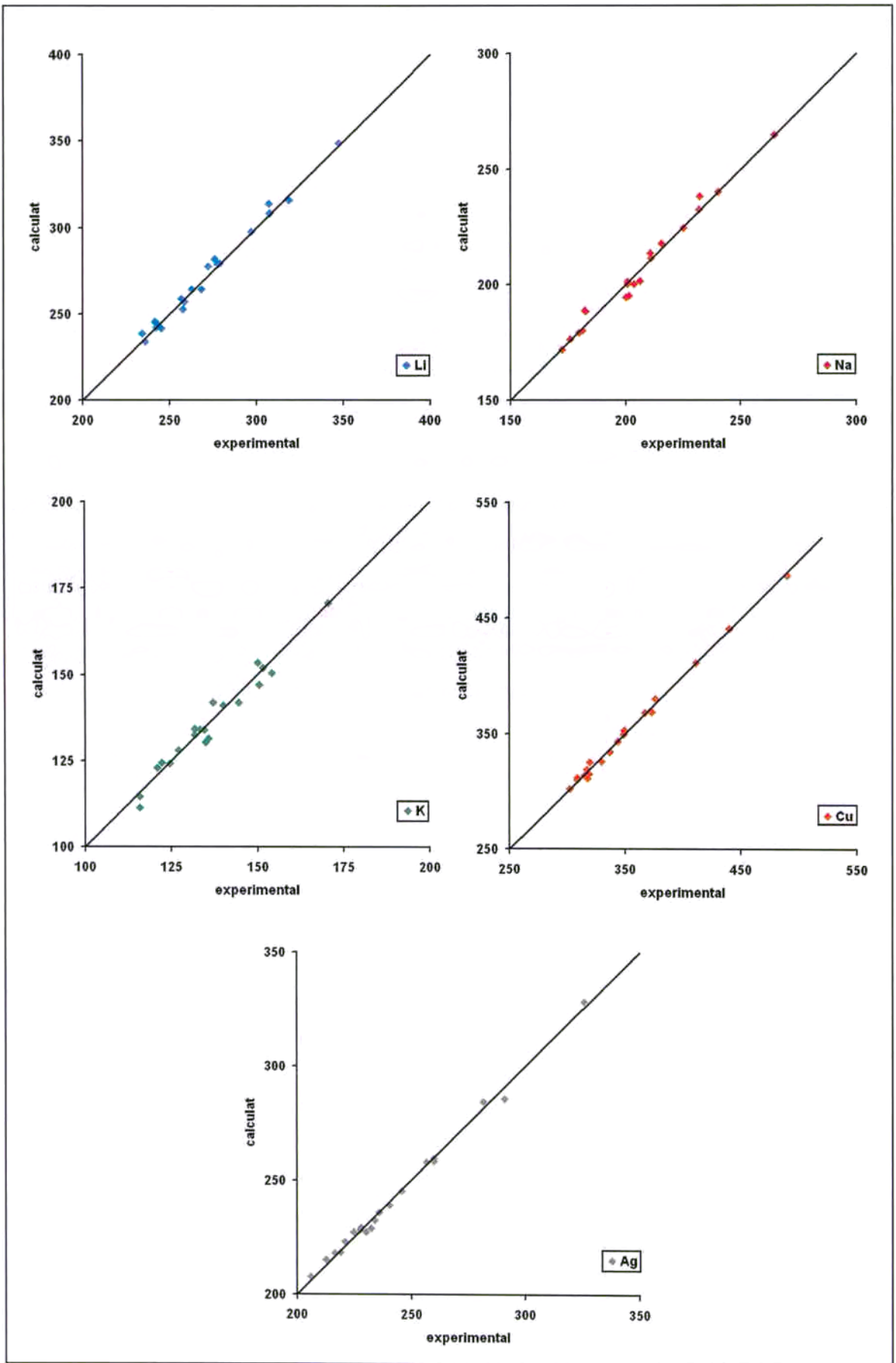


Figura 4.14.  $\Delta H^\circ$  calculat vs. experimental per a les famílies dels metalls.

Taula 4.9 Valors *experimentals* i calculats d'afinitat catiònica (kJ/mol).

#	Aminoàcid	Metall	Exp.	BPNN
1	Ala <sup>a</sup>	Ag	212.8	215.0
2	Ala <sup>c</sup>	Cu	308.5	309.5
3	Ala <sup>b</sup>	K	120.8	122.8
4	Ala <sup>a</sup>	Li	234.1	238.5
5	Ala <sup>a</sup>	Na	176.0	176.1
6	Arg <sup>c</sup>	Ag	325.6	328.1
7	Arg <sup>a</sup>	Cu	489.5	486.6
8	Arg <sup>a</sup>	K	170.5	170.7
9	Arg <sup>a</sup>	Li	347.4	348.2
10	Arg <sup>a</sup>	Na	264.2	264.7
11	Asn <sup>b</sup>	Ag	245.8	245.2
12	Asn <sup>a</sup>	Cu	348.6	349.0
13	Asn <sup>a</sup>	K	150.5	146.8
14	Asn <sup>c</sup>	Li	297.2	297.0
15	Asn <sup>a</sup>	Na	224.9	224.5
16	Asp <sup>a</sup>	Ag	232.4	228.9
17	Asp <sup>b</sup>	Cu	329.8	325.2
18	Asp <sup>c</sup>	K	140.0	140.9
19	Asp <sup>a</sup>	Li	279.2	278.8
20	Asp <sup>a</sup>	Na	211.1	211.4
21	Cys <sup>c</sup>	Ag	228.2	228.9
22	Cys <sup>a</sup>	Cu	329.8	325.6
23	Cys <sup>a</sup>	K	115.8	114.5
24	Cys <sup>a</sup>	Li	241.6	245.2
25	Cys <sup>a</sup>	Na	182.2	188.8
26	Gln <sup>a</sup>	Ag	256.7	258.0
27	Gln <sup>a</sup>	Cu	372.9	368.1
28	Gln <sup>c</sup>	K	154.2	150.3
29	Gln <sup>b</sup>	Li	307.2	313.4
30	Gln <sup>a</sup>	Na	231.6	232.5
31	Glu <sup>a</sup>	Ag	234.1	232.4
32	Glu <sup>a</sup>	Cu	336.5	333.8
33	Glu <sup>a</sup>	K	137.1	141.8
34	Glu <sup>a</sup>	Li	275.9	281.4
35	Glu <sup>c</sup>	Na	210.7	213.4
36	Gly <sup>a</sup>	Ag	206.1	207.6
37	Gly <sup>a</sup>	Cu	302.2	301.8
38	Gly <sup>a</sup>	K	115.8	111.3
39	Gly <sup>a</sup>	Li	235.8	233.5
40	Gly <sup>c</sup>	Na	172.2	171.6
41	His <sup>a</sup>	Ag	281.7	284.4
42	His <sup>c</sup>	Cu	410.9	411.4
43	His <sup>a</sup>	K	151.7	151.9
44	His <sup>a</sup>	Li	318.9	315.7
45	His <sup>b</sup>	Na	239.9	240.0
46	Ile <sup>a</sup>	Ag	219.0	218.3
47	Ile <sup>a</sup>	Cu	317.3	310.7
48	Ile <sup>a</sup>	K	122.1	124.4
49	Ile <sup>c</sup>	Li	242.4	241.9
50	Ile <sup>a</sup>	Na	181.4	180.1

#	Aminoàcid	Metall	Exp.	BPNN
51	Leu <sup>c</sup>	Ag	221.1	223.0
52	Leu <sup>a</sup>	Cu	318.9	314.5
53	Leu <sup>a</sup>	K	127.1	128.0
54	Leu <sup>c</sup>	Li	242.9	244.4
55	Leu <sup>a</sup>	Na	182.7	188.3
56	Lys <sup>a</sup>	Ag	290.9	285.8
57	Lys <sup>a</sup>	Cu	439.7	440.9
58	Lys <sup>b</sup>	K	150.1	153.3
59	Lys <sup>a</sup>	Li	307.6	308.0
60	Lys <sup>a</sup>	Na	232.0	238.0
61	Met <sup>b</sup>	Ag	260.0	259.3
62	Met <sup>a</sup>	Cu	367.4	367.8
63	Met <sup>a</sup>	K	131.7	132.3
64	Met <sup>a</sup>	Li	272.1	277.2
65	Met <sup>c</sup>	Na	206.1	201.4
66	Phe <sup>a</sup>	Ag	236.2	235.7
67	Phe <sup>a</sup>	Cu	343.6	342.5
68	Phe <sup>c</sup>	K	135.0	130.3
69	Phe <sup>a</sup>	Li	257.5	252.3
70	Phe <sup>a</sup>	Na	200.2	194.5
71	Pro <sup>a</sup>	Ag	230.3	227.1
72	Pro <sup>c</sup>	Cu	316.8	318.6
73	Pro <sup>a</sup>	K	135.9	131.4
74	Pro <sup>b</sup>	Li	256.7	258.4
75	Pro <sup>a</sup>	Na	201.5	195.0
76	Ser <sup>a</sup>	Ag	224.9	227.3
77	Ser <sup>a</sup>	Cu	314.8	313.1
78	Ser <sup>a</sup>	K	131.7	134.1
79	Ser <sup>c</sup>	Li	262.9	264.2
80	Ser <sup>b</sup>	Na	200.6	200.2
81	Thr <sup>a</sup>	Ag	227.8	228.1
82	Thr <sup>a</sup>	Cu	319.4	324.7
83	Thr <sup>a</sup>	K	133.3	133.8
84	Thr <sup>a</sup>	Li	268.4	264.0
85	Thr <sup>c</sup>	Na	203.6	200.0
86	Trp <sup>a</sup>	Ag	260.0	258.3
87	Trp <sup>b</sup>	Cu	376.2	379.7
88	Trp <sup>c</sup>	K	144.6	141.7
89	Trp <sup>a</sup>	Li	277.6	278.4
90	Trp <sup>a</sup>	Na	215.3	217.6
91	Tyr <sup>c</sup>	Ag	240.4	238.9
92	Tyr <sup>a</sup>	Cu	349.4	352.1
93	Tyr <sup>a</sup>	K	134.6	133.9
94	Tyr <sup>a</sup>	Li	258.7	256.8
95	Tyr <sup>a</sup>	Na	200.6	201.1
96	Val <sup>a</sup>	Ag	216.5	217.9
97	Val <sup>c</sup>	Cu	308.5	311.4
98	Val <sup>a</sup>	K	124.6	124.0
99	Val <sup>a</sup>	Li	245.4	241.5
100	Val <sup>a</sup>	Na	179.7	179.2

a Training set. b Cross-validation set. c Prediction set.

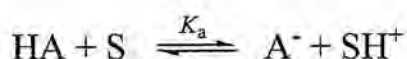
## 4.2. Constants d'acidesa de compostos orgànics en diferents solvents

Els processos àcid-base són un dels conjunts de reaccions més importants en química i bioquímica. El caràcter àcid o bàsic d'un compost químic ve donat per les seves constants de dissociació,  $K_a$ . Aquestes constants descriuen l'extensió amb què un compost es dissocia en fase gasosa o en solució. El  $pK_a$  és una propietat clau que té importància capital en la reactivitat general de les substàncies químiques amb altres compostos i també les interaccions amb els solvents. És igualment important en multitud de processos fisicoquímics complexos, com poden ser els processos cromatogràfics, de partició i de transferència de fase, etc.

Normalment, aquesta constant es mesura en aigua com a solvent però, en els últims temps, la seva determinació en solvents no aquosos està guanyant força interès i ja s'han fet determinacions d'aquest tipus per a un nombre significatiu de compostos orgànics i inorgànics, tant en solvents pròtics com en apròtics.

D'altra banda, és ben conegut que les propietats farmacocinètiques dels compostos químics: biodisponibilitat, capacitat de difusió a través de membranes, etc., estan fortament lligades al seu  $pK_a$ . De fet, el  $pK_a$  és considerat, juntament amb la integritat, la lipofília, la solubilitat i la permeabilitat, un dels cinc factors fisicoquímics clau que permeten la comprensió de propietats ADME (*Absorption Distribution Metabolism Excretion*),<sup>76</sup> de gran importància.

D'acord amb la definició de Brønsted-Lowry, un àcid és qualsevol compost donador de protons i una base és un acceptor de protons. Donat que en solució no poden existir protons aïllats, és lògic pensar que la reacció de transferència protònica tindrà lloc només en presència d'una base. La majoria de solvents que es coneixen presenten propietats àcid-base, així, l'acidesa d'un compost depèn marcadament en el medi en què es dissol. El procés de dissociació és una reacció complexa que es pot representar com un equilibri entre les molècules del compost no dissociat i el seu anió conjugat i el protó, ambdós solvatats pel solvent. L'equació següent mostra l'esquema clàssic d'aquest tipus de reaccions:



El procés de solvatació de les espècies es compon de diversos sub processos independents complexos, que es presenten sota el nom genèric d'interaccions solut/solvent. Aquestes interaccions es poden dividir, de manera simplificada, en dos grups. El primer, que es refereix als efectes macroscòpics o no específics, inclou les interaccions de llarg abast i no direccionals. Alguns exemples d'aquest tipus d'interaccions poden ser la polarització dielèctrica, la dispersió isotròpica i la formació de cavitats en el solvent. El segon tipus d'interaccions, les anomenades interaccions específiques, són de tipus direccional, curt abast i tenen molt en compte les interaccions individuals que es donen entre molècules de solut i solvent. En general, aquest tipus d'interaccions es relacionen amb formació d'enllaços d'hidrogen i interaccions donador/acceptor (àcid/base de Lewis).<sup>77</sup>

Es disposa d'uns quants mètodes de determinació experimental<sup>78,79</sup> de  $K_a$ , entre els quals poden es troben les valoracions potenciomètriques,<sup>80,81</sup> les determinacions espectrofotomètriques,<sup>82-84</sup> la ressonància magnètica nuclear,<sup>85,86</sup> la cromatografia de líquids,<sup>87,88</sup> l'electroforesi capil·lar,<sup>89,90</sup> etc. Donada la gran importància d'aquesta propietat s'han desenvolupat un bon nombre de mètodes teòrics que pretenen calcular les constants de dissociació dels compostos químics, entre d'ells caldria destacar el model clàssic de Hammett i Taft, basat en constants dels substituents.

Les propietats àcid-base de molècules petites en fase gasosa es poden calcular amb una exactitud molt elevada mitjançant mètodes de càlcul post-HF i DFT.<sup>91-104</sup> Aquests mètodes permeten el càlcul de l'acidesa, la basicitat i l'afinitat protònica de molècules petites orgàniques i inorgàniques.

Dissortadament, aquests mètodes computacionals no funcionen igual de bé quan es treballa en solució. El problema ve donat majoritàriament per la impossibilitat de calcular quantitativament de manera satisfactòria les energies de solvatació de les espècies implicades en la reacció de dissociació. Per tal de resoldre aquest problema s'han introduït diferents aproximacions per tractar els efectes de la solvatació. Alguns exemples són: simulacions moleculars,<sup>105-112</sup> els models dipolars de Langevin,<sup>113</sup> tècniques d'equacions integrals,<sup>114-118</sup> mètodes semiempírics<sup>119-122</sup> i mètodes de dielèctric continu.<sup>123-129</sup> A partir d'aquest últim mètode s'han implementat els mètodes del model continu polaritzat<sup>130-140</sup> i el model de clúster continu,<sup>141,142</sup> que permeten el càlcul de constants de dissociació d'un gran nombre d'àcids en aigua.

Tot i l'èxit que han tingut aquests últims mètodes per calcular les constants de dissociació en aigua, hi ha molt poca informació disponible sobre el càlcul d'aquesta propietat en solvents orgànics. Recentment, s'han publicat alguns treballs en què es calcula el  $pK_a$  d'alguns àcids orgànics en dimetilsulfòxid i acetonitril.<sup>143-146</sup>

La metodologia QSPR ha estat emprada per calcular constants de dissociació de diferents famílies de compostos orgànics en aigua.<sup>147-154</sup> Molt recentment s'ha publicat un estudi sobre el càlcul de  $pK_a$  d'amines orgàniques i fosfines en acetonitril.<sup>155</sup> S'ha publicat, també, un estudi comparatiu dels models de Hammet-Taft i de Drago on s'estimen constants de dissociació d'àcids carboxílics, àcids benzoics, fenols i anilines protonades en metanol.<sup>156</sup>

L'objectiu que es persegueix en aquest capítol és establir, mitjançant la metodologia QSPR, models que permetin calcular el  $pK_a$  de compostos orgànics en diferents solvents. La tria de les famílies de compostos orgànics i solvents ve donada per la disponibilitat de valors en la bibliografia, així, com a soluts es van triar quatre famílies: fenols, àcids benzoics, àcids carboxílics alifàtics i anilines, ja que de totes elles es disposa d'un gran nombre de valors experimentals en diferents solvents. Aquests solvents es poden dividir en pròtics: aigua, metanol (MeOH), etanol (EtOH), isopropanol (<sup>i</sup>PrOH) i *tert*-butanol (<sup>t</sup>BuOH); i apròtics: dimetilsulfòxid (DMSO), *N,N*-dimetilformamida (DMF), *N,N*-dimetilacetamida (DMA), acetonitril (AN), nitrometà (NM) i acetona (Ac). Per a les tres primeres famílies de soluts es volen deduir models que permetin el càlcul de  $pK_a$ , mentre que per a les anilines els models proposats serviran per calcular  $pK_b$ . Els models que s'establiran per aquestes quatre famílies de compostos seran del tipus no lineal ja que els sistemes multicomponents estudiats són massa complexos com per ser tractats amb mètodes de regressió multilineal. Dins dels mètodes no lineals, s'establiran correlacions amb xarxes neuronals: BPNN implementades en el programa ADAPT i amb SVM.



Abans de començar, però, cal decidir la manera en què es volen establir els models, és a dir, cal triar quins seran els descriptors implicats en els models. Per als soluts es generaran els descriptors moleculars de la manera habitual amb CODESSA. Per contra, les característiques del solvent no vindran donades pels seus descriptors moleculars ja que aquests descriptors defineixen la molècula com a tal i poden no servir per interpretar el seu comportament com a solvent. Així doncs, els descriptors triats per recollir les característiques dels solvents han estat propietats físiques: índex de refracció, densitat, permitivitat relativa, etc.; i les anomenades escales de solvent. El següent apartat conté un recull de les diferents propietats físiques i escales de solvent seleccionades per convertir-se en els descriptors externs dels solvents.

### *Descriptors externs dels solvents*

En aquest apartat s'enunciaran els paràmetres que caracteritzen els solvents com a tals i no com a molècules individuals. El fet de fer servir aquests paràmetres experimentals implica construir una nova base de dades que contingui les propietats físiques i les escales de solvents trobades en la bibliografia per a tots els solvents enumerats.

Les propietats físiques<sup>157</sup> incloses com a descriptors dels solvents són les següents: moment dipolar ( $\mu$ ), alfa polaritzabilitat (*alfa*), pes molecular ( $M$ ), índex de refracció ( $n_D$ ), volum molar a 25°C ( $V_m$ ), refractivitat molar ( $R$ ), densitat a 25°C ( $d$ ), entalpia ( $\Delta_{vap}H^\circ$ ) i energia ( $\Delta_{vap}U^\circ$ ) de vaporització molars, constant dielèctrica ( $\epsilon_r$ ) i solubilitat de Hildebrand ( $\delta$ ).

Les escales multiparamètriques de solvents han estat desenvolupades per facilitar la comprensió de les propietats dels solvents que no corresponen amb exactitud a cap de les seves característiques físiques.<sup>158,159</sup> Aquestes escales, a més, proporcionen les bases necessàries per triar el solvent adequat en cada situació. En general, aquestes escales es basen en diversos fenòmens fisicoquímics com poden ser entalpies de reacció, efectes solvatocròmics i velocitats de reacció, entre d'altres. La Taula 4.14 conté les escales utilitzades en aquest estudi com a descriptors dels solvents, tot seguit s'explicarà breument quin significat físic tenen aquestes escales i com es mesuren els seus valors.

Taula 4.14. Escales de solvents emprades en aquest estudi.

Descriptor	Definició (Autor)	Ref.
$\alpha$	Capacitat de donació d'enllaç d'hidrogen (Kamlet, Taft)	160
$\beta$	Capacitat de donació de parells electrònics (Kamlet, Taft)	161
$\pi^*$	Polaritat/Polaritzabilitat (Kamlet, Taft)	162
$E_T^{30}$	Índex de polaritat (Dimroth, Reichardt)	163
<b>Y</b>	Polaritat (Koppel, Palm)	164
<b>P</b>	Polaritzabilitat (Koppel, Palm)	164
<b>E</b>	Acidesa (Koppel, Palm)	164
<b>B</b>	Basicitat (Koppel, Palm)	164
<b>B'</b>	Formació acid-base d'enllaços d'hidrogen (Koppel, Paju)	165
<b>DN</b>	Nombre donador (Gutmann)	166
<b>AN</b>	Nombre acceptor (Mayer, Gutmann, Gerger)	167
<b>Z</b>	Capacitat de formació d'enllaços d'hidrogen (Kosower)	168
<b>SA</b>	Escala d'acidesa (Catalán)	169
<b>SB</b>	Escala de basicitat (Catalán)	170
<b>SPP</b>	Polaritat (Catalán)	171
<b>B<sub>j</sub></b>	Tendència a solvatar cations (Swain)	172
<b>A<sub>j</sub></b>	Tendència a solvatar anions (Swain)	172
<b>S'</b>	Escala de polaritat no específica (Drago)	173
<b>E<sub>B</sub></b>	Entalpia electrostàtica bàsica (Drago)	174
<b>C<sub>B</sub></b>	Entalpia covalent bàsica (Drago)	174
<b>S</b>	Capacitat de formació d'enllaços d'hidrogen (Brownstein)	175
<b>E<sub>b</sub></b>	Acidesa de Lewis (Janowski)	176
<b>M</b>	Expressió de l'índex de refracció (McRae)	177
<b>J</b>	Expressió de la constant dielèctrica (Kirkwood, David)	177
<b>N</b>	Combinació de M i P (McRae)	177

- Capacitat de donació d'enllaç d'hidrogen ( $\alpha$ ),<sup>160</sup> capacitat de donació de parells electrònics ( $\beta$ )<sup>161</sup> i polaritat/polaritzabilitat ( $\pi^*$ )<sup>162</sup> de Kamlet i Taft. Els dos primers es poden associar a les característiques àcid-base de Lewis. Les escales  $\alpha$  i  $\beta$  s'obtenen mitjançant un mètode de comparació amb parelles homomòrfiques que permet establir els desplaçaments solvatocròmics de diferents colorants. L'escala  $\pi^*$  es basa en els efectes del solvent en les transicions  $\pi \rightarrow \pi^*$  per a diferents compostos nitroaromàtics.
- Índex de polaritat  $E_T^{30}$  de Dimroth i Reichardt.<sup>163</sup> Aquest paràmetre, que s'obté a partir de l'energia de transició de la banda d'absorció solvatocròmica de major longitud d'ona del colorant *N*-fenolat betaina de piridini, és una mesura tant de la polaritat del solvent com de la seva capacitat per formar enllaços d'hidrogen per donació amb altres molècules.

- Acidesa (E), basicitat (B), polaritat (Y) i polaritzabilitat (P) de Koppel i Palm.<sup>164</sup> Aquestes quatre escales es combinen en una equació que permet relacionar la variació d'una propietat qualsevol A amb dues característiques específiques: E i B; i dues característiques no específiques: P i Y, del solvent.

$$A = A_0 + yY + pP + bB + eE$$

Els paràmetres E i B es poden relacionar directament amb l'acidesa i basicitat de Lewis, respectivament. Els paràmetres P i Y s'obtenen, respectivament, a partir de diferents expressions de l'índex de refracció i de la constant dielèctrica.

- B' de Koppel i Paju.<sup>165</sup> Aquest paràmetre, que és anàleg a la basicitat de Lewis, s'obté a partir de mesures de la freqüència de la vibració de l'enllaç O-H del fenol en tetraclorometà en absència i en presència del solvent per al qual es vol determinar.
- Nombres donador (DN)<sup>166</sup> i acceptor (AN)<sup>167</sup> de Guttman. El nombre acceptor (AN), que és una mesura de l'acidesa de Lewis del solvent, es deriva de desplaçaments de RMN de fòsfor de l'adducte 1:1 Et<sub>3</sub>PO-SbCl<sub>5</sub>. El nombre donador, que és una mesura de la basicitat de Lewis del solvent, es troba mitjançant calorimetria i és l'entalpia de coordinació del SbCl<sub>5</sub> amb un solvent qualsevol a temperatura ambient.
- El paràmetre Z de Kosower,<sup>168</sup> que és una mesura de la capacitat del solvent per establir enllaços d'hidrogen i interaccions polars. Aquesta escala s'obté experimentalment a partir de mesures de l'energia de la transició molar d'absorció de transferència de càrrega per al compost, iodur d'l-etil- 4(metoxicarbonil)piridini en el solvent triat.
- Acidesa (SA),<sup>169</sup> basicitat (SB)<sup>170</sup> i polaritat (SPP)<sup>171</sup> de Catalán. Aquestes tres escales, que es basen també en mesures UV/Vis de parelles solvatocròmiques i homomòrfiques de colorants, contenen valors per a més de dos-cents solvents diferents.

- Tendències a solvatar anions ( $A_j$ ) i cations ( $B_j$ ) de Swain.<sup>172</sup> Aquests dos paràmetres, anàlegs a l'acidesa i basicitat de Lewis, respectivament, es calculen a partir de reaccions sensibles a canvis de solvent i propietats fisicoquímiques: velocitats de reacció, constants d'equilibri, IR, UV/Vis., RSE, RMN, etc.
- Escala de polaritat no específica del solvent ( $S'$ ),<sup>173</sup> entalpies de basicitat electrostàtica ( $E_B$ ) i covalent ( $C_B$ ) de Drago.<sup>174</sup> L'escala  $S'$  conté les forces electrostàtiques no específiques que permeten al solvent interaccionar amb els soluts iònics, apolars i polars. Els paràmetres  $E_B$  i  $C_B$  formen part de l'equació de quatre membres desenvolupada per correlacionar l'entalpia de reacció entre un compost acceptor i un de donador en un solvent inert (tetraclorometà o hexà):

$$-\Delta H_{AB}^{\circ} = E_A E_B + C_A C_B$$

on  $E_A$  i  $C_A$  són els paràmetres acceptors i  $E_B$  i  $C_B$  són els paràmetres donadors. Els dos paràmetres  $E$  mesuren la tendència d'un àcid o una base per participar en interaccions electrostàtiques mentre que els paràmetres  $C$  mesuren la tendència a participar en interaccions covalents. En aquest estudi només s'usen els paràmetres  $E_B$  i  $C_B$  perquè de  $E_A$  i  $C_A$  no es tenen els valors per a tots els solvents.

- $S$  de Brownstein.<sup>175</sup> Aquesta escala, que deriva de l'escala  $Z$  de Kosower, utilitza la refractivitat molar del solvent per tal de millorar la sensibilitat d'aquesta.
- Acidesa de Lewis ( $E_b$ ) de Janowski.<sup>176</sup> Aquesta escala es mesura experimentalment a partir de l'espectre d'absorció de la *N*-oxil-2,2,6,6-tetrametilpiperidina (TMPNO), és a dir, els valors de  $E_b$  són l'energia de la transició  $n \rightarrow \pi^*$  de la TMPNO en el solvent triat.
- Per últim, els descriptors de McRae, Kirkwood i David:<sup>177</sup>  $M$ ,  $J$  i  $N$  són, respectivament, diferents expressions de l'índex de refracció, de la constant dielèctrica i una combinació d'ambdós.

### 4.2.1. $pK_a$ de fenols

La primera família de compostos estudiada és la dels fenols. Aquests compostos semblen molt aptes per fer el primer pas d'utilitzar la metodologia QSPR al càlcul de propietats de sistemes multicomponents per diverses raons: són de fàcil accés i es disposa de gran nombre de valors de  $pK_a$  en diferents solvents, presenten gran analogia entre ells degut a l'anell fenòlic però alhora les diferències químiques entre ells són notables dependent de quins substituents tinguin i en quines posicions es trobin.

#### *Conjunt de dades*

El conjunt de dades triat conté 276 valors de  $pK_a$ ; fruit de la combinació de 94 fenols diferents en deu solvents. Aquests deu solvents es poden dividir entre pròtics:  $H_2O$ , MeOH,  $i$ -PrOH i  $t$ -BuOH; i apròtics: Ac, AN, DMSO, DMF, DMA i NM. Els valors de  $pK_a$  en aigua i alcohols s'han obtingut de publicacions de Bosch i col·laboradors,<sup>178-182</sup> els valors de  $pK_a$  en solvents apròtics s'han obtingut preferentment de la compilació d'Izutsku<sup>183</sup> i de les publicacions de Vianello i col·laboradors.<sup>184</sup>

Els 94 fenols presenten diferents substituents sobre l'anell aromàtic. Apart del fenol, hi ha 39 fenols monosubstituïts, 30 disubstituïts, 18 trisubstituïts, 4 tetrasubstituïts i 2 pentasubstituïts. Aquests substituents, 17 en total, són: metil, etil, *tert*-butil, fenil, metoxi, fluorur, clorur, bromur, iodur, hidroxil, nitro, trifluorometil, cianur, acetil, amino, metilsulfonyl i trifluorometilsulfonyl. Donat la seva diferent naturalesa química presenten efectes electrònics i estèrics molt diferenciats. En els compostos polihidroxilats el valor de  $pK_a$  triat es correspon amb el de la primera constant de dissociació del fenol. Del conjunt de 94 fenols, n'hi ha 51 que presenten substituents en *orto*-, aquest tipus de substitució és força important ja que els substituents en aquesta posició poden interactuar amb el grup OH que es dissocia formant enllaços d'hidrogen o afectant a la dissociació d'aquest enllaç per efectes estèrics. En el conjunt de dades es troben 141 valors de  $pK_a$  de fenols *orto*-substituïts i, a més del mateix fenol, 135 valors de fenols substituïts en *meta*- o en *para*-. El nombre de fenols diferents en cada solvent és molt variable: 83 en aigua, 39 en metanol, 16 en isopropanol, 16 en *tert*-butanol, 44 en DMSO, 29 en DMF, 25 en acetonitril, 10 en nitrometà i acetona i 4 en DMA. Aquesta distribució dels valors experimentals no és gaire homogènia, i dels 276 valors



de  $pK_a$  experimentals, 154 corresponen a mesures en solvents pròtics i 122 corresponen a solvents apròtics. El rang de  $pK_a$  que cobreixen aquests compostos en els deu solvents esmentats és molt ampli i va des de -0.3 per al 2,4,6-trinitrofenol (àcid pícric) en DMSO a 27.5 per al 2-metilfenol en acetonitril. La Taula 4.15 conté tots els valors de  $pK_a$  experimentals i calculats per a aquesta família de compostos.

La divisió del conjunt de dades en *training* i *prediction sets* proporciona subconjunts de 221 i 55 valors, respectivament. Aquests dos subconjunts de dades seran els que s'usaran en el model construït amb SVM. Per tal d'establir els models no lineals amb ADAPT és necessari l'anomenat *cross-validation set*, que conté un 10% dels valors del *tset*. D'aquesta manera el *tset* d'ADAPT conté 199 valors de  $pK_a$  mentre que el *cvset* en té 22. Totes aquestes particions del conjunt de dades inicials es fan de manera que s'asseguri la presència de tots els solvents i tots els substituents en els tres nous subconjunts de dades creats, intentant també mantenir-ne les mateixes proporcions.

#### **Model no lineal: Xarxes neuronals BPNN**

El nombre inicial de descriptors calculats per als fenols amb CODESSA és de 677, a aquest nombre cal sumar-hi els 37 descriptors dels solvents. La reducció de descriptors implementada en el mètode heurístic de CODESSA redueix aquest nombre inicial a 240 descriptors, dels quals 215 pertanyen als fenols i 25 als solvents. Com que per establir els models BPNN s'utilitza el programa ADAPT, aquest nombre de descriptors es redueix encara més, fins a un nombre definitiu de 114, 91 dels fenols i 23 dels solvents. L'arquitectura seleccionada per a les xarxes neuronals és 7-5-1, amb la qual cosa el nombre de paràmetres ajustables es troba fixat en 46. El *ratio* del nombre de compostos del *tset*, 199, sobre el nombre de paràmetres ajustables és de 4.3, bastant per sobre del 2.0 recomanat.

El model trobat amb les xarxes neuronals, amb una arquitectura 7-5-1, conté els descriptors que es mostren en la Taula 4.16. Aquest model conté cinc descriptors dels fenols i dos descriptors dels solvents.

Taula 4.16. Descriptors continguts en el model dels fenols amb BPNN.

Descriptor	Component	Classe
Càrrega parcial màxima d'un àtom d'hidrogen	Solut	Electroestàtic
Alfa polaritzabilitat	Solut	Quàntic
Energia LUMO+1	Solut	Quàntic
Atracció electró-nucli màxima d'un enllaç C-O	Solut	Quàntic
Àrea superficial relativa carregada positivament, RPCS	Solut	Quàntic
Capacitat de donació d'enllaç d'hidrogen de Kamlet i Taft, $\alpha$	Solvent	-
Moment dipolar, $\mu$	Solvent	-

Com es pot comprovar, quatre descriptors dels fenols pertanyen a la classe dels quàntics; el cinquè descriptor, la càrrega parcial màxima sobre un àtom d'hidrogen, és electroestàtic. Aquest descriptor se situa, en tots els casos, sobre l'hidrogen hidroxílic dels fenols, és a dir, sobre el protó que es dissocia de la molècula. Aquesta càrrega parcial ha estat calculada amb el mètode proposat per Zefirov,<sup>73</sup> a partir de les electronegativitats de Sanderson. Els descriptors quàntics s'obtenen directament amb el càlcul semiempíric AM1 implementat en el programa MOPAC; aquests descriptors són l'alfa polaritzabilitat, l'energia de l'orbital LUMO+1, l'energia d'atracció electró-nucli màxima d'un enllaç C-O i l'àrea superficial relativa carregada positivament, RPCS. El descriptor  $\alpha$  del solvent pertany a l'escala desenvolupada per Kamlet i Taft,<sup>159</sup> que és una de les més utilitzades per definir les interaccions a través d'enllaços d'hidrogen dels solvents. Aquest descriptor mesura la capacitat de les molècules de solvent per establir enllaços donadors d'hidrogen amb els soluts. L'últim descriptor del model és el moment dipolar del solvent,  $\mu$ .

Els resultats estadístics que s'obtenen amb aquest model són molt bons, la Taula 4.17 conté els valors del coeficient de determinació  $R^2$ , el test- $F$  i el  $rmse$  per als tres subconjunts de dades: *tset*, *cvset* i *pset*. Els valors de  $pK_a$  calculats amb aquest model es troben en la Taula 4.15 i la seva representació gràfica enfront dels valors experimentals es pot veure en la Figura 4.15.

Taula 4.17. Paràmetres estadístics obtinguts amb el model BPNN.

Conjunt	n	$R^2$	F	$rmse$
<i>tset</i>	199	0.982	10916.74	0.75
<i>cvset</i>	22	0.975	777.08	0.95
<i>pset</i>	55	0.977	2267.66	0.83

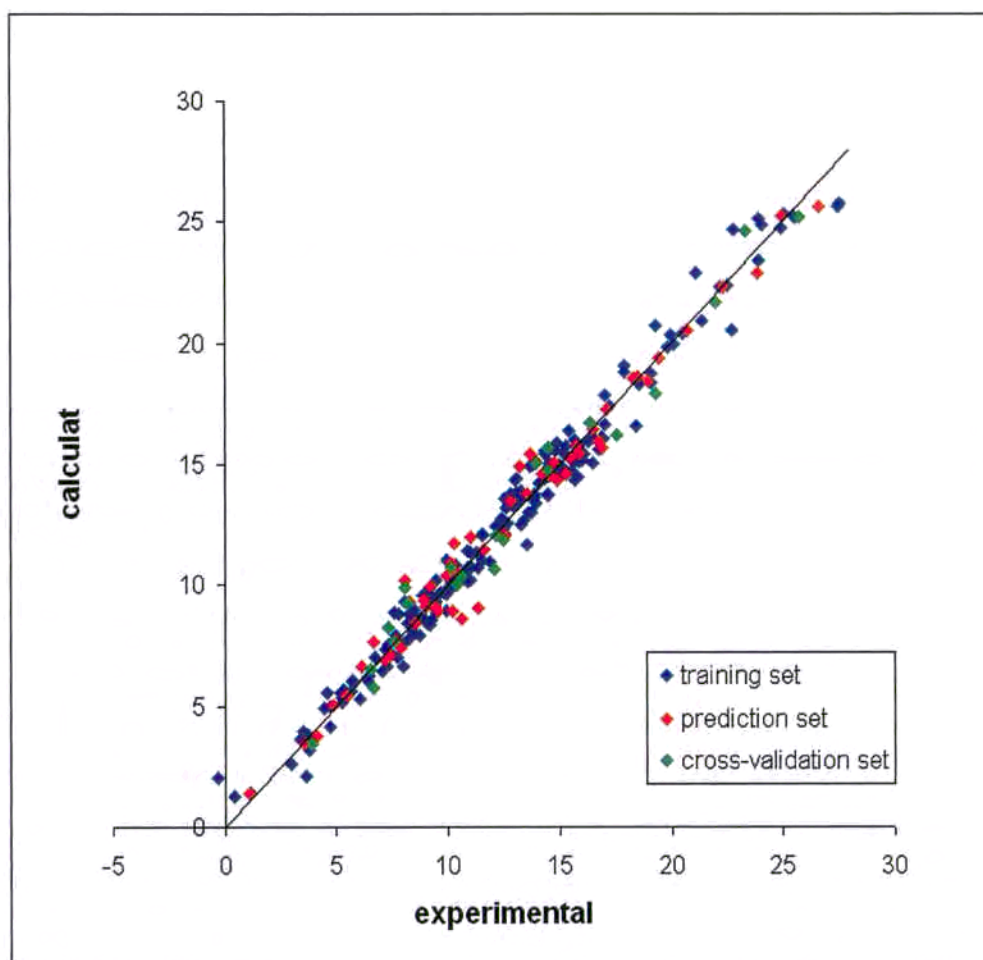


Figura 4.15. Representació gràfica de  $pK_a$  calculat vs. experimental per als fenols amb BPNN.

Els descriptors que inclou el model permeten relacionar el procés de dissociació que es dona amb les interaccions solut/solvent que es produeixen en aquest sistema. Cal recordar que les interaccions solut/solvent es divideixen en dos grups: les interaccions no específiques, que es refereixen al comportament global del sistema i que se solen relacionar amb els fenòmens de polarització dielèctrica, dispersió isotròpica i formació de cavitats en el solvent; i les interaccions específiques, que es relacionen formació d'enllaços d'hidrogen i interaccions donador/acceptor (àcid/base de Lewis).<sup>77</sup> Analitzant els descriptors que conté el model, es pot dir que només hi ha un descriptor referent a les interaccions específiques: l'energia de l'orbital LUMO+1, els altres quatre descriptors dels fenols estan relacionats amb les interaccions no específiques. Els descriptors del solvent es poden classificar de la mateixa manera: el moment dipolar fa

referència a les interaccions no específiques mentre que el descriptor  $\alpha$  es refereix a les específiques.

Com ja s'ha dit, el descriptor càrrega parcial màxima sobre un àtom d'hidrogen se situa sempre sobre el protó hidroxílic, aquest descriptor reflecteix la polaritat de l'enllaç O-H que es dissocia i mesura la facilitat amb la qual aquest enllaç es trenca per donar el fenolat conjugat i un protó. D'acord amb els efectes electrònics produïts pels substituents en els anells aromàtics, els valors més elevats d'aquest descriptor es troben en fenols que tenen un gran nombre de substituents electroatracients: policlorats i polinitrats; per contra, els fenols amb valors més baixos per aquest descriptor solen ser els que presenten substituents alquilics. Anàlogament, el descriptor energia d'atracció electró-nucli màxima d'un enllaç C-O es relaciona amb l'energia dels enllaços O-H ja que, en tots els casos, aquest descriptor se situa en l'enllaç C-O del grup fenòlic. El descriptor RPCS pertany als anomenats descriptors quàntics de tipus CPSA, aquest tipus de descriptors contenen alhora informació de la forma de la molècula i de la distribució de càrrega, així doncs, aquest descriptor codifica informació sobre les possibles interaccions polars que es poden donar entre les molècules de solut i solvent. El descriptor alfa polaritzabilitat del solut i el descriptor moment dipolar del solvent són mesures de la polaritat, dins de les interaccions no específiques, aquests descriptors recullen la informació referent a les interaccions electrostàtiques que es donen entre solut i solvent. Les interaccions específiques, de tipus acceptor/donador, es poden relacionar amb el descriptor dels soluts energia de l'orbital LUMO+1. Aquest descriptor correspon a la segona afinitat electrònica dels fenols que, a la seva vegada es pot relacionar amb el caràcter àcid de Lewis d'aquestes molècules. Per últim, el descriptor  $\alpha$  es està relacionat, clarament, amb aquest mateix tipus d'interaccions ja que descriu la formació dels enllaços d'hidrogen entre el fenol i el solvent en el procés de dissociació.

La importància relativa de cada descriptor dintre del model s'ha avaluat amb el mètode de Jurs. Així, la permuta d'un descriptor condueix a un increment en el *rmse* del model, que serà més gran com més important sigui el descriptor. La taula 4.18 conté els increments de *rmse* que pateix el model quan es permuten individualment els descriptors.



Taula 4.18. Increments de *rmse* per la permuta de cada descriptor.

Descriptor	Increment <i>rmse</i>	Importància
Alfa polaritzabilitat	0.34	7
Càrrega parcial màxima d'un àtom d'hidrogen	0.97	4
Energia LUMO+1	2.79	3
Atracció electró-nucli màxima d'un enllaç C-O	0.61	5
Àrea superficial relativa carregada positivament, RPCS	0.53	6
Capacitat de donació d'enllaç d'hidrogen de Kamlet i Taft, $\alpha$	5.61	1
Moment dipolar, $\mu$	4.10	2

Com es pot comprovar, els descriptors més importants són els dos que pertanyen als solvents:  $\alpha$  i  $\mu$ , el tercer descriptor en importància, i primer dels soluts, és l'energia de l'orbital LUMO+1.

El valor de *rmse* per als 276 valors de  $pK_a$  és 0.76. En la Taula 4.19 es troben aquells compostos pels qual s'obté un valor de residual més gran que dues vegades el *rmse* del model. D'entre tots aquests, només el 2,4,6-trinitrofenol en DMSO presenta un residual major que tres vegades el *rmse* del model i, com a conseqüència, es podria considerar com a *outlier*. Dels altres catorze valors, fruit de la combinació de deu fenols en diferents solvents, només dos corresponen a mesures en solvent pròtics: 2-cloro-4-fenilfenol i 4-acetilfenol en aigua. Els altres dotze valors corresponen a mesures en solvents apròtics on, en general, la precisió de les mesures sol ser més baixa.

Taula 4.19. Prediccions de  $pK_a$  amb valors més grans que dues vegades el *rmse*\*

Solut	Solvent	Exper.	Calc.	residual
2,4,6-tribromofenol	Ac	21.10	22.89	1.79
2,4-dinitrofenol	AN	18.40	16.57	-1.83
2-cloro-4-fenilfenol	H <sub>2</sub> O	8.07	10.23	2.16
2-metilfenol	AN	27.50	25.76	-1.74
3-(metilsulfonil)fenol	DMSO	13.56	11.70	-1.86
3,4-diclorofenol	DMF	13.22	14.96	1.74
3,5-dinitrofenol	Ac	22.70	20.57	-2.13
3,5-dinitrofenol	DMF	11.30	9.07	-2.23
3,5-dinitrofenol	DMSO	10.60	8.65	-1.95
4-acetilfenol	DMSO	13.68	15.44	1.76
4-acetilfenol	H <sub>2</sub> O	8.05	9.90	1.85
4-cianofenol	AN	22.77	24.70	1.93
4-metilfenol	AN	27.45	25.66	-1.79
4- <i>tert</i> -butilfenol	AN	27.48	25.64	-1.84
2,4,6-trinitrofenol	DMSO	-0.30	2.04	2.34

\* Els compostos ombrejats tenen residuals més grans que tres vegades el *rmse*



Per comprovar la robustesa del model, es van separar els compostos en diferents subconjunts. La Taula 4.20 mostra tots aquests subconjunts i els paràmetres estadístics que s'obtenen per a cadascun amb els valors de  $pK_a$  predits pel model.

Taula 4.20. Estadística dels subconjunts de fenols.

Subconjunt	n	R <sup>2</sup>	F	rmse
<i>orto</i>	141	0.983	8261.7	0.69
No <i>orto</i>	135	0.974	4941.4	0.81
Pròtics	154	0.975	6020.2	0.58
Apròtics	122	0.978	5414.7	0.93
H <sub>2</sub> O	83	0.936	1194.6	0.56
MeOH	39	0.963	962.9	0.58
<sup>t</sup> PrOH	16	0.928	179.6	0.77
<sup>t</sup> BuOH	16	0.979	663.4	0.51
DMSO	44	0.965	1166.2	0.81
DMF	29	0.955	573.0	0.94
AN	25	0.937	340.0	1.05
Ac	10	0.949	148.9	1.13
NM	10	0.976	319.8	0.73
DMA	4	0.918	22.7	0.75

Com es pot comprovar, els resultats que s'obtenen amb els fenols mono- o disubstituïts en *orto* i els que no tenen substituïda aquesta posició (no *orto*) són molt semblants. Això fa indicar que l'anomenat efecte *orto* és poc important en aquesta família de compostos. Aquest fet coincideix amb altres estudis on també s'ha trobat aquest comportament per a la família dels fenols.<sup>156,185</sup> Els altres dos grans subconjunts que es poden fer es basen en la classificació en solvents pròtics i apròtics. En aquest cas, la diferència entre els dos conjunts és del mateix ordre que la vista entre fenols *orto* i no *orto*. Pel que fa als paràmetres estadístics individuals de cada solvent, tots els resultats que s'obtenen són bons, en especial en els solvents pròtics.

**Model no lineal: Support Vector Machines**

El model desenvolupat amb *Support Vector Machines* conté els mateixos set descriptors triats pel model BPNN. Així doncs, l'optimització dels paràmetres  $\gamma$ ,  $C$  i  $\epsilon$ , Figura 4.16, condueix a l'establiment del model definitiu. L'optimització del paràmetre  $\gamma$  (amb  $\epsilon=0.1$  i  $C=100$ ) troba que el valor òptim d'aquest paràmetre és 1.9. Un cop optimitzat aquest paràmetre, es passa a optimitzar  $\epsilon$  (amb  $\gamma=1.9$  i  $C=100$ ), l'optimització d'aquest paràmetre troba que el millor valor per a  $\epsilon$  és 0.09. Per últim, el paràmetre  $C$  s'optimitza amb  $\gamma=1.9$  i  $\epsilon=0.09$ ; el valor òptim per aquest paràmetre és 105.

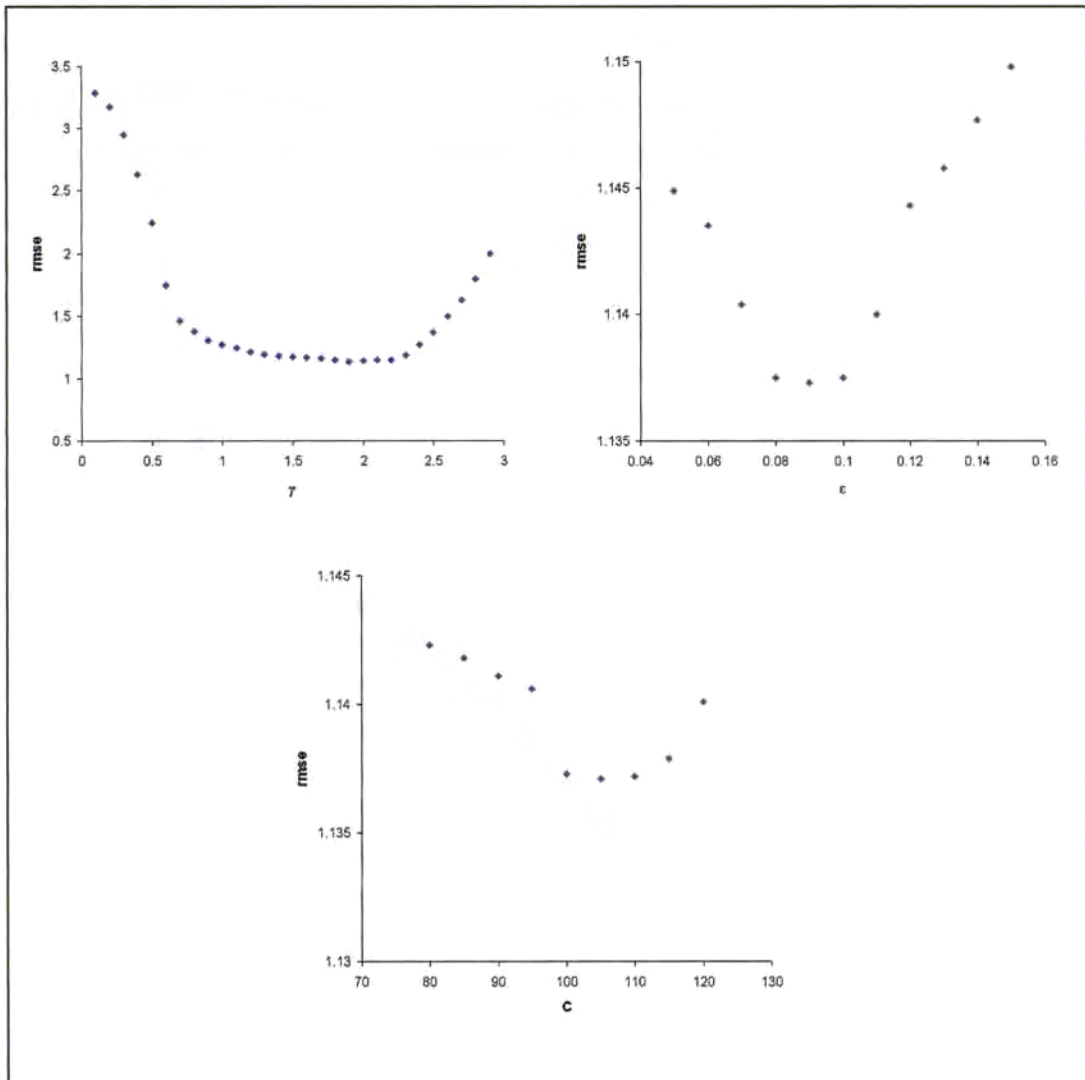


Figura 4.16. Optimització dels paràmetres  $\gamma$ ,  $\epsilon$  i  $C$ .

El model calculat per LibSVM consta de 203 vectors suport, és a dir, s'utilitza aquest nombre de compostos d'un total de 221 del *tset* per establir el model. Els resultats estadístics que s'obtenen amb aquest model són semblants als que s'obtenien amb el model BPNN, la Taula 4.21 conté els valors del coeficient de determinació  $R^2$ , el test- $F$  i el *rmse* per als dos subconjunts de dades: *tset* i *pset*.

Taula 4.21. Paràmetres estadístics obtinguts amb el model SVM.

Conjunt	n	$R^2$	F	<i>rmse</i>
<i>tset</i>	221	0.987	17166.6	0.60
<i>pset</i>	55	0.965	1458.4	0.96

Els valors de  $pK_a$  calculats amb aquest model es troben en la Taula 4.15 i la seva representació gràfica enfront dels valors experimentals es pot veure en la Figura 4.17.

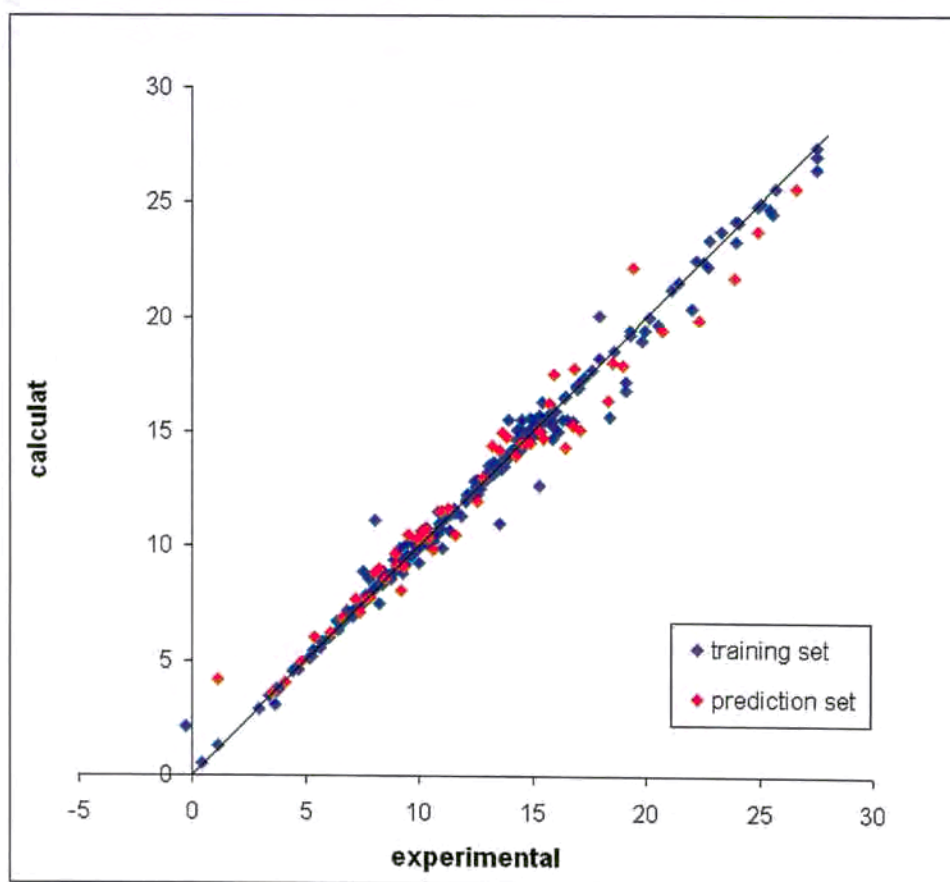


Figura 4.17. Representació gràfica de  $pK_a$  calculat vs. experimental per als fenols amb SVM.

De nou, s'examinen els paràmetres estadístics que s'obtenen en separar el conjunt de compostos en subconjunts de naturalesa anàloga. La Taula 4.22 mostra aquests subconjunts de compostos i els seus paràmetres estadístics. En aquest cas es pot comprovar que l'establiment del model amb el mètode SVM comporta una millora dels valors en solvents pròtics i un empitjorament en els apròtics. En general, sembla que en aplicar aquest mètode, es milloren els resultats per als solvents majoritaris: aigua, metanol, DMSO, DMF i acetonitril; i s'empitjoren els dels solvents minoritaris: acetona, nitrometà i DMA.

Taula 4.22. Estadística dels subconjunts de fenols.

<b>Subconjunt</b>	<b>n</b>	<b>R<sup>2</sup></b>	<b>F</b>	<b>rmse</b>
<i>orto</i>	141	0.980	6455.9	0.75
No <i>orto</i>	135	0.983	8132.3	0.65
Pròtics	154	0.988	12543.2	0.40
Apròtics	122	0.976	4904.3	0.94
H <sub>2</sub> O	83	0.973	2964.5	0.35
MeOH	39	0.982	2061.9	0.39
<sup>t</sup> PrOH	16	0.993	1936.8	0.25
<sup>t</sup> BuOH	16	0.962	357.8	0.67
DMSO	44	0.975	1629.8	0.71
DMF	29	0.973	960.4	0.67
AN	25	0.962	584.9	0.93
Ac	10	0.964	216.0	0.90
NM	10	0.939	123.9	1.05
DMA	4	0.754	6.1	1.20

Taula 4.15. Valors experimentals i calculats de  $pK_a$  per a la família dels fenols.

#	Fenol	Solvent	Exp.	BPNN	SVM
1	2,3,4,6-tetraclorofenol <sup>a,c</sup>	DMSO	7.55	8.88	8.89
2	2,3,4,6-tetraclorofenol <sup>a,c</sup>	H <sub>2</sub> O	5.63	5.63	5.54
3	2,3,4,6-tetraclorofenol <sup>d</sup>	DMF	9.50	8.96	10.46
4	2,3,4,6-tetranitrofenol <sup>d</sup>	DMF	1.11	1.40	4.13
5	2,3-dihidroxifenol <sup>a,c</sup>	H <sub>2</sub> O	9.01	8.54	8.92
6	2,3-dimetilfenol <sup>a,c</sup>	H <sub>2</sub> O	10.54	10.41	10.30
7	2,3-dimetilfenol <sup>a,c</sup>	MeOH	15.08	14.93	14.90
8	2,3-dinitrofenol <sup>a,c</sup>	H <sub>2</sub> O	5.24	5.37	5.15
9	2,3-dinitrofenol <sup>a,c</sup>	MeOH	9.43	9.34	9.52
10	2,4,5-triclorofenol <sup>a,c</sup>	DMF	12.46	12.31	12.83
11	2,4,5-triclorofenol <sup>d</sup>	DMSO	10.97	12.01	11.52
12	2,4,6-tribromofenol <sup>a,c</sup>	Ac	21.10	22.89	21.19
13	2,4,6-tribromofenol <sup>d</sup>	H <sub>2</sub> O	6.10	6.63	6.18
14	2,4,6-tribromofenol <sup>d</sup>	MeOH	10.10	10.94	10.57
15	2,4,6-triclorofenol <sup>a,c</sup>	Ac	22.50	22.37	22.41
16	2,4,6-triclorofenol <sup>a,c</sup>	DMSO	10.19	10.64	10.28
17	2,4,6-triclorofenol <sup>a,c</sup>	H <sub>2</sub> O	6.42	6.29	6.34
18	2,4,6-triclorofenol <sup>a,c</sup>	<sup>t</sup> BuOH	14.82	15.57	14.73
19	2,4,6-triclorofenol <sup>b,c</sup>	DMF	12.05	10.64	11.96
20	2,4,6-triclorofenol <sup>d</sup>	<sup>i</sup> PrOH	12.55	12.05	11.92
21	2,4,6-trimetil-3-nitrofenol <sup>a,c</sup>	H <sub>2</sub> O	8.98	9.51	8.89
22	2,4,6-trimetilfenol <sup>a,c</sup>	H <sub>2</sub> O	10.86	10.59	10.77
23	2,4,6-trimetilfenol <sup>a,c</sup>	MeOH	15.53	14.78	15.44
24	2,4,6-trinitrofenol <sup>a,c</sup>	<sup>i</sup> PrOH	3.70	3.81	3.61
25	2,4,6-trinitrofenol <sup>a,c</sup>	Ac	9.20	8.34	9.11
26	2,4,6-trinitrofenol <sup>a,c</sup>	AN	11.00	10.80	9.86
27	2,4,6-trinitrofenol <sup>a,c</sup>	DMF	3.65	2.13	3.08
28	2,4,6-trinitrofenol <sup>a,c</sup>	DMSO	-0.30	2.04	2.09
29	2,4,6-trinitrofenol <sup>a,c</sup>	H <sub>2</sub> O	0.43	1.30	0.52
30	2,4,6-trinitrofenol <sup>a,c</sup>	NM	8.02	9.33	11.11
31	2,4,6-trinitrofenol <sup>a,c</sup>	<sup>t</sup> BuOH	4.70	4.97	4.79
32	2,4,6-trinitrofenol <sup>b,c</sup>	MeOH	3.90	3.47	3.81
33	2,4-diclorofenol <sup>a,c</sup>	<sup>i</sup> PrOH	14.48	13.70	14.21
34	2,4-diclorofenol <sup>a,c</sup>	DMSO	13.25	13.55	13.16
35	2,4-diclorofenol <sup>a,c</sup>	H <sub>2</sub> O	7.65	7.79	7.85
36	2,4-diclorofenol <sup>a,c</sup>	<sup>t</sup> BuOH	17.25	17.40	17.34
37	2,4-diclorofenol <sup>d</sup>	DMF	13.56	13.78	14.17
38	2,4-dimetilfenol <sup>a,c</sup>	MeOH	15.04	14.84	14.83
39	2,4-dimetilfenol <sup>b,c</sup>	H <sub>2</sub> O	10.60	10.42	10.19
40	2,4-dinitrofenol <sup>a,c</sup>	Ac	15.70	16.03	15.79
41	2,4-dinitrofenol <sup>a,c</sup>	AN	18.40	16.57	15.62
42	2,4-dinitrofenol <sup>a,c</sup>	DMF	6.36	6.06	6.69
43	2,4-dinitrofenol <sup>a,c</sup>	DMSO	5.32	5.67	5.41
44	2,4-dinitrofenol <sup>d</sup>	H <sub>2</sub> O	4.10	3.75	4.05
45	2,4-dinitrofenol <sup>d</sup>	MeOH	7.82	7.43	7.70
46	2,4-dinitrofenol <sup>d</sup>	NM	15.90	15.45	17.46
47	2,4-di-t-butilfenol <sup>a,c</sup>	MeOH	16.77	15.63	15.45
48	2,4-di-t-butilfenol <sup>d</sup>	H <sub>2</sub> O	11.57	11.52	10.50
49	2,5-diclorofenol <sup>a,c</sup>	DMF	13.85	13.64	13.94
50	2,5-dimetilfenol <sup>a,c</sup>	MeOH	14.91	15.02	14.96



#	Fenol	Solvent	Exp.	BPNN	SVM
51	2,5-dimetilfenol <sup>d</sup>	H <sub>2</sub> O	10.41	10.51	10.32
52	2,5-dinitrofenol <sup>a,c</sup>	Ac	19.80	19.82	18.94
53	2,5-dinitrofenol <sup>a,c</sup>	DMF	8.78	8.68	8.76
54	2,5-dinitrofenol <sup>a,c</sup>	H <sub>2</sub> O	5.22	5.19	5.13
55	2,5-dinitrofenol <sup>a,c</sup>	NM	17.90	18.86	20.05
56	2,5-dinitrofenol <sup>b,c</sup>	DMSO	7.32	8.23	7.41
57	2,5-dinitrofenol <sup>d</sup>	MeOH	8.93	9.41	9.63
58	2,6-dibromo-4-nitrofenol <sup>a,c</sup>	DMF	5.71	6.08	5.72
59	2,6-dibromo-4-nitrofenol <sup>a,c</sup>	DMSO	5.17	5.56	5.08
60	2,6-dibromo-4-nitrofenol <sup>a,c</sup>	H <sub>2</sub> O	3.38	3.65	3.47
61	2,6-dibromo-4-nitrofenol <sup>a,c</sup>	MeOH	7.31	7.50	7.40
62	2,6-dicloro-4-hidroxifenol <sup>d</sup>	H <sub>2</sub> O	7.38	7.09	7.09
63	2,6-dicloro-4-nitrofenol <sup>a,c</sup>	DMF	5.72	6.01	5.81
64	2,6-dicloro-4-nitrofenol <sup>a,c</sup>	MeOH	7.40	7.24	7.31
65	2,6-dicloro-4-nitrofenol <sup>d</sup>	H <sub>2</sub> O	3.55	3.42	3.58
66	2,6-diclorofenol <sup>a,c</sup>	<sup>i</sup> PrOH	13.58	13.00	13.49
67	2,6-diclorofenol <sup>a,c</sup>	Ac	23.90	23.44	23.31
68	2,6-diclorofenol <sup>a,c</sup>	DMF	12.55	12.13	12.91
69	2,6-diclorofenol <sup>a,c</sup>	DMSO	11.54	12.11	11.40
70	2,6-diclorofenol <sup>a,c</sup>	H <sub>2</sub> O	6.79	6.95	7.15
71	2,6-diclorofenol <sup>b,c</sup>	<sup>t</sup> BuOH	16.38	16.73	16.47
72	2,6-dimetil-4-cianofenol <sup>d</sup>	H <sub>2</sub> O	8.27	9.35	9.02
73	2,6-dimetil-4-nitrofenol <sup>a,c</sup>	DMSO	9.00	9.12	8.91
74	2,6-dimetil-4-nitrofenol <sup>a,c</sup>	H <sub>2</sub> O	7.19	7.31	7.25
75	2,6-dimetilfenol <sup>a,c</sup>	H <sub>2</sub> O	10.59	10.32	10.49
76	2,6-dimetilfenol <sup>d</sup>	MeOH	15.26	14.62	15.01
77	2,6-dinitro-4-hidroxifenol <sup>a,c</sup>	H <sub>2</sub> O	4.42	4.92	4.51
78	2,6-dinitrofenol <sup>a,c</sup>	Ac	13.78	14.97	13.87
79	2,6-dinitrofenol <sup>a,c</sup>	DMF	6.07	5.29	5.98
80	2,6-dinitrofenol <sup>a,c</sup>	H <sub>2</sub> O	3.74	3.18	3.76
81	2,6-dinitrofenol <sup>a,c</sup>	MeOH	7.70	7.00	7.79
82	2,6-dinitrofenol <sup>a,c</sup>	NM	16.00	15.21	15.91
83	2,6-dinitrofenol <sup>d</sup>	AN	16.45	16.48	14.31
84	2,6-dinitrofenol <sup>d</sup>	DMSO	4.82	5.03	4.90
85	2,6-di-t-butil-4-nitrofenol <sup>a,c</sup>	AN	19.10	18.74	16.80
86	2,6-di-t-butil-4-nitrofenol <sup>a,c</sup>	DMF	8.27	8.79	8.36
87	2,6-di-t-butil-4-nitrofenol <sup>a,c</sup>	MeOH	10.89	10.73	10.98
88	2,6-di-t-butil-4-nitrofenol <sup>b,c</sup>	DMSO	7.60	7.68	7.69
89	2,6-di-t-butil-4-nitrofenol <sup>d</sup>	H <sub>2</sub> O	6.62	7.68	6.80
90	2,6-di-t-butilfenol <sup>d</sup>	DMSO	16.85	15.69	17.77
91	2-aminofenol <sup>a,c</sup>	H <sub>2</sub> O	9.44	10.19	10.09
92	2-bromofenol <sup>a,c</sup>	AN	23.92	25.12	24.17
93	2-bromofenol <sup>a,c</sup>	H <sub>2</sub> O	8.39	8.60	8.50
94	2-bromofenol <sup>d</sup>	DMF	13.85	15.20	14.76
95	2-cloro-4-bromofenol <sup>a,c</sup>	MeOH	12.70	12.59	12.52
96	2-cloro-4-bromofenol <sup>d</sup>	H <sub>2</sub> O	7.64	7.73	7.66
97	2-cloro-4-fenilfenol <sup>a,c</sup>	<sup>i</sup> PrOH	13.70	14.95	13.79
98	2-cloro-4-fenilfenol <sup>a,c</sup>	AN	24.90	24.80	24.81
99	2-cloro-4-fenilfenol <sup>a,c</sup>	DMF	15.70	15.13	15.79
100	2-cloro-4-fenilfenol <sup>a,c</sup>	DMSO	14.90	14.45	14.81
101	2-cloro-4-fenilfenol <sup>a,c</sup>	<sup>t</sup> BuOH	18.60	18.37	18.51

#	Fenol	Solvent	Exp.	BPNN	SVM
102	2-cloro-4-fenilfenol <sup>d</sup>	H <sub>2</sub> O	8.07	10.23	8.83
103	2-clorofenol <sup>a,c</sup>	<sup>t</sup> PrOH	15.83	14.48	15.42
104	2-clorofenol <sup>a,c</sup>	<sup>t</sup> BuOH	18.54	18.31	18.50
105	2-clorofenol <sup>b,c</sup>	DMA	13.91	15.06	15.49
106	2-clorofenol <sup>d</sup>	H <sub>2</sub> O	8.51	8.45	8.58
107	2-clorofenol <sup>d</sup>	MeOH	12.83	13.48	12.94
108	2-etilfenol <sup>a,c</sup>	H <sub>2</sub> O	10.20	10.29	10.55
109	2-fluorofenol <sup>a,c</sup>	H <sub>2</sub> O	8.73	7.92	8.56
110	2-fluorofenol <sup>a,c</sup>	MeOH	12.14	12.48	12.23
111	2-hidroxí-3-nitrofenol <sup>a,c</sup>	H <sub>2</sub> O	6.68	7.03	6.68
112	2-hidroxifenol <sup>a,c</sup>	H <sub>2</sub> O	9.12	9.32	9.52
113	2-iodofenol <sup>a,c</sup>	H <sub>2</sub> O	8.46	9.01	8.87
114	2-metil-4,6-dinitrofenol <sup>a,c</sup>	DMSO	4.59	5.53	4.68
115	2-metilfenol <sup>a,c</sup>	AN	27.50	25.76	27.00
116	2-metilfenol <sup>a,c</sup>	H <sub>2</sub> O	10.31	9.99	10.22
117	2-metilfenol <sup>d</sup>	MeOH	14.90	14.36	14.51
118	2-metoxifenol <sup>a,c</sup>	H <sub>2</sub> O	9.90	9.62	9.81
119	2-metoxifenol <sup>a,c</sup>	MeOH	14.48	13.79	14.39
120	2-nitrofenol <sup>a,c</sup>	<sup>t</sup> PrOH	13.30	12.50	13.21
121	2-nitrofenol <sup>a,c</sup>	DMSO	11.00	11.37	11.51
122	2-nitrofenol <sup>a,c</sup>	H <sub>2</sub> O	7.23	6.65	7.14
123	2-nitrofenol <sup>a,c</sup>	MeOH	11.52	12.09	11.61
124	2-nitrofenol <sup>a,c</sup>	NM	21.40	20.96	21.49
125	2-nitrofenol <sup>a,c</sup>	<sup>t</sup> BuOH	15.88	15.81	15.79
126	2-nitrofenol <sup>b,c</sup>	AN	22.00	21.67	20.38
127	2-nitrofenol <sup>b,c</sup>	DMF	12.14	12.05	12.04
128	2-nitrofenol <sup>d</sup>	Ac	22.30	22.32	19.85
129	2-t-butilfenol <sup>a,c</sup>	H <sub>2</sub> O	11.34	10.73	10.67
130	2-t-butilfenol <sup>a,c</sup>	MeOH	16.50	15.07	15.48
131	3-(metilsulfonil)fenol <sup>a,c</sup>	DMSO	13.56	11.70	10.95
132	3-(trifluorometilsulfonil)fenol <sup>b,c</sup>	NM	19.30	17.95	19.21
133	3,4,5-triclorofenol <sup>a,c</sup>	DMSO	12.58	13.21	12.27
134	3,4,5-triclorofenol <sup>a,c</sup>	H <sub>2</sub> O	7.68	7.81	7.59
135	3,4-diclorofenol <sup>a,c</sup>	AN	24.06	24.86	24.15
136	3,4-diclorofenol <sup>a,c</sup>	H <sub>2</sub> O	8.51	8.60	8.42
137	3,4-diclorofenol <sup>d</sup>	DMF	13.22	14.96	14.39
138	3,4-diclorofenol <sup>d</sup>	DMSO	14.22	14.65	13.90
139	3,4-dimetilfenol <sup>a,c</sup>	H <sub>2</sub> O	10.36	10.41	10.34
140	3,4-dimetilfenol <sup>a,c</sup>	MeOH	14.63	14.80	14.96
141	3,4-dinitrofenol <sup>a,c</sup>	AN	17.90	19.10	18.19
142	3,4-dinitrofenol <sup>a,c</sup>	DMSO	7.97	7.62	8.06
143	3,4-dinitrofenol <sup>a,c</sup>	MeOH	9.46	8.87	9.55
144	3,4-dinitrofenol <sup>d</sup>	H <sub>2</sub> O	5.42	5.44	5.98
145	3,5-diclorofenol <sup>a,c</sup>	<sup>t</sup> PrOH	14.05	14.22	14.14
146	3,5-diclorofenol <sup>a,c</sup>	DMSO	13.09	14.44	13.54
147	3,5-diclorofenol <sup>a,c</sup>	H <sub>2</sub> O	8.18	8.42	8.27
148	3,5-diclorofenol <sup>a,c</sup>	MeOH	12.94	13.29	12.92
149	3,5-diclorofenol <sup>a,c</sup>	<sup>t</sup> BuOH	17.04	17.90	17.13
150	3,5-diclorofenol <sup>b,c</sup>	AN	23.31	24.66	23.73
151	3,5-dihidroxifenol <sup>a,c</sup>	H <sub>2</sub> O	8.45	8.71	8.54
152	3,5-dimetil-4-cianofenol <sup>b,c</sup>	H <sub>2</sub> O	8.21	9.23	8.75

#	Fenol	Solvent	Exp.	BPNN	SVM
153	3,5-dimetil-4-nitrofenol <sup>a,c</sup>	H <sub>2</sub> O	8.25	7.79	7.49
154	3,5-dimetilfenol <sup>a,c</sup>	H <sub>2</sub> O	10.20	10.44	10.29
155	3,5-dimetilfenol <sup>a,c</sup>	MeOH	14.62	14.96	14.91
156	3,5-dinitrofenol <sup>a,c</sup>	<sup>i</sup> PrOH	10.84	10.16	10.55
157	3,5-dinitrofenol <sup>a,c</sup>	Ac	22.70	20.57	22.18
158	3,5-dinitrofenol <sup>a,c</sup>	AN	20.50	20.45	19.65
159	3,5-dinitrofenol <sup>a,c</sup>	<sup>t</sup> BuOH	13.40	12.72	13.49
160	3,5-dinitrofenol <sup>b,c</sup>	H <sub>2</sub> O	6.66	5.72	6.70
161	3,5-dinitrofenol <sup>d</sup>	DMF	11.30	9.07	11.60
162	3,5-dinitrofenol <sup>d</sup>	DMSO	10.60	8.65	9.83
163	3,5-dinitrofenol <sup>d</sup>	MeOH	10.20	8.92	10.34
164	3,5-dinitrofenol <sup>d</sup>	NM	19.40	19.42	22.13
165	3,5-di-t-butilfenol <sup>a,c</sup>	MeOH	14.89	15.92	15.55
166	3,5-di-t-butilfenol <sup>d</sup>	H <sub>2</sub> O	10.29	11.76	10.70
167	3-acetilfenol <sup>a,c</sup>	DMSO	15.14	15.52	15.23
168	3-acetilfenol <sup>d</sup>	H <sub>2</sub> O	9.19	9.98	8.06
169	3-aminofenol <sup>a,c</sup>	H <sub>2</sub> O	9.99	9.91	9.93
170	3-bromofenol <sup>a,c</sup>	<sup>i</sup> PrOH	14.83	14.87	15.03
171	3-bromofenol <sup>a,c</sup>	H <sub>2</sub> O	9.01	9.33	9.23
172	3-bromofenol <sup>a,c</sup>	MeOH	13.30	13.92	13.68
173	3-bromofenol <sup>d</sup>	<sup>t</sup> BuOH	18.52	18.61	18.00
174	3-cianofenol <sup>d</sup>	DMSO	14.76	15.09	14.48
175	3-cloro-2,4,6-trinitrofenol <sup>a,c</sup>	DMSO	1.16	1.43	1.25
176	3-cloro-4-nitrofenol <sup>a,c</sup>	AN	19.95	20.39	19.40
177	3-cloro-4-nitrofenol <sup>a,c</sup>	DMSO	9.80	9.73	10.03
178	3-cloro-4-nitrofenol <sup>b,c</sup>	H <sub>2</sub> O	6.49	6.49	6.60
179	3-clorofenol <sup>a,c</sup>	AN	25.04	25.34	24.95
180	3-clorofenol <sup>a,c</sup>	DMA	16.29	15.96	15.51
181	3-clorofenol <sup>a,c</sup>	DMSO	15.83	15.50	15.24
182	3-clorofenol <sup>a,c</sup>	MeOH	13.10	13.88	13.49
183	3-clorofenol <sup>d</sup>	H <sub>2</sub> O	9.02	9.13	9.13
184	3-etilfenol <sup>d</sup>	H <sub>2</sub> O	9.90	10.38	10.25
185	3-fluorofenol <sup>a,c</sup>	DMSO	15.88	15.16	14.70
186	3-fluorofenol <sup>a,c</sup>	H <sub>2</sub> O	9.28	8.53	8.78
187	3-hidroxifenol <sup>a,c</sup>	DMSO	15.30	15.74	15.71
188	3-hidroxifenol <sup>a,c</sup>	H <sub>2</sub> O	9.15	9.27	9.33
189	3-iodofenol <sup>a,c</sup>	H <sub>2</sub> O	8.88	9.55	9.35
190	3-metil-5-etilfenol <sup>b,c</sup>	H <sub>2</sub> O	10.10	10.71	10.64
191	3-metilfenol <sup>a,c</sup>	DMSO	16.86	16.18	16.95
192	3-metilfenol <sup>a,c</sup>	H <sub>2</sub> O	10.10	10.16	10.15
193	3-metilfenol <sup>b,c</sup>	MeOH	14.48	14.74	14.57
194	3-metoxifenol <sup>a,c</sup>	H <sub>2</sub> O	9.65	9.72	9.55
195	3-metoxifenol <sup>d</sup>	DMSO	15.72	15.79	16.20
196	3-nitrofenol <sup>a,c</sup>	<sup>i</sup> PrOH	13.92	13.42	14.01
197	3-nitrofenol <sup>a,c</sup>	DMF	13.85	13.79	13.76
198	3-nitrofenol <sup>a,c</sup>	DMSO	13.75	13.03	13.44
199	3-nitrofenol <sup>a,c</sup>	H <sub>2</sub> O	8.36	8.10	8.27
200	3-nitrofenol <sup>a,c</sup>	MeOH	12.40	12.74	12.31
201	3-nitrofenol <sup>a,c</sup>	NM	22.20	22.33	22.49
202	3-nitrofenol <sup>a,c</sup>	<sup>t</sup> BuOH	16.99	16.67	16.90
203	3-nitrofenol <sup>d</sup>	AN	23.85	22.91	21.73

#	Fenol	Solvent	Exp.	BPNN	SVM
204	3-trifluorometil-4-nitrofenol <sup>a, c</sup>	<sup>t</sup> PrOH	9.90	11.02	10.16
205	3-trifluorometil-4-nitrofenol <sup>a, c</sup>	AN	19.30	20.74	19.39
206	3-trifluorometil-4-nitrofenol <sup>a, c</sup>	DMSO	9.30	9.42	9.16
207	3-trifluorometil-4-nitrofenol <sup>a, c</sup>	H <sub>2</sub> O	6.41	6.18	6.32
208	3-trifluorometil-4-nitrofenol <sup>a, c</sup>	<sup>t</sup> BuOH	12.77	13.82	12.83
209	3-trifluorometil-4-nitrofenol <sup>b, c</sup>	DMF	10.40	10.00	10.31
210	3-trifluorometilfenol <sup>a, c</sup>	<sup>t</sup> PrOH	12.50	13.61	12.59
211	3-trifluorometilfenol <sup>a, c</sup>	DMF	15.70	14.39	15.61
212	3-trifluorometilfenol <sup>a, c</sup>	DMSO	14.30	14.56	15.07
213	3-trifluorometilfenol <sup>a, c</sup>	H <sub>2</sub> O	9.04	8.40	8.95
214	3-trifluorometilfenol <sup>a, c</sup>	MeOH	12.10	12.09	12.19
215	3-trifluorometilfenol <sup>d</sup>	AN	24.90	25.29	23.71
216	3-trifluorometilfenol <sup>d</sup>	<sup>t</sup> BuOH	17.10	17.32	15.08
217	4-acetilfenol <sup>b, c</sup>	H <sub>2</sub> O	8.05	9.90	8.14
218	4-acetilfenol <sup>d</sup>	DMSO	13.68	15.44	14.95
219	4-aminofenol <sup>a, c</sup>	H <sub>2</sub> O	10.43	10.39	10.34
220	4-bromofenol <sup>a, c</sup>	<sup>t</sup> PrOH	14.30	14.66	14.39
221	4-bromofenol <sup>a, c</sup>	AN	25.53	25.19	24.53
222	4-bromofenol <sup>a, c</sup>	DMF	14.34	15.56	14.70
223	4-bromofenol <sup>a, c</sup>	MeOH	13.63	13.70	13.37
224	4-bromofenol <sup>a, c</sup>	<sup>t</sup> BuOH	19.10	18.36	17.19
225	4-bromofenol <sup>d</sup>	DMSO	15.50	15.25	14.66
226	4-bromofenol <sup>d</sup>	H <sub>2</sub> O	9.36	9.16	9.05
227	4-cianofenol <sup>a, c</sup>	AN	22.77	24.70	23.37
228	4-cianofenol <sup>a, c</sup>	DMSO	13.01	14.39	13.24
229	4-cianofenol <sup>a, c</sup>	H <sub>2</sub> O	7.80	8.83	8.55
230	4-cloro-2,6-dinitrofenol <sup>a, c</sup>	AN	15.30	14.67	12.64
231	4-cloro-2,6-dinitrofenol <sup>a, c</sup>	DMF	4.68	4.15	4.61
232	4-cloro-2,6-dinitrofenol <sup>a, c</sup>	DMSO	3.51	3.93	3.60
233	4-cloro-2,6-dinitrofenol <sup>a, c</sup>	H <sub>2</sub> O	2.97	2.63	2.88
234	4-clorofenol <sup>a, c</sup>	<sup>t</sup> PrOH	15.31	14.71	14.99
235	4-clorofenol <sup>a, c</sup>	AN	25.44	25.28	24.73
236	4-clorofenol <sup>a, c</sup>	DMSO	16.10	15.43	15.03
237	4-clorofenol <sup>a, c</sup>	H <sub>2</sub> O	9.38	9.17	9.19
238	4-clorofenol <sup>a, c</sup>	MeOH	13.59	13.74	13.50
239	4-clorofenol <sup>b, c</sup>	DMF	14.50	15.72	15.09
240	4-clorofenol <sup>d</sup>	DMA	16.78	15.95	15.23
241	4-clorofenol <sup>d</sup>	<sup>t</sup> BuOH	18.96	18.45	17.86
242	4-etilfenol <sup>a, c</sup>	H <sub>2</sub> O	10.00	10.41	10.28
243	4-fluorofenol <sup>a, c</sup>	H <sub>2</sub> O	9.95	8.96	9.27
244	4-hidroxi-2,3,5,6-tetrametilfenol <sup>a, c</sup>	H <sub>2</sub> O	11.51	11.02	11.42
245	4-hidroxi-2-metilfenol <sup>a, c</sup>	H <sub>2</sub> O	10.20	10.22	10.10
246	4-hidroxi-2-nitrofenol <sup>a, c</sup>	H <sub>2</sub> O	7.63	7.92	7.72
247	4-hidroxifenol <sup>a, c</sup>	H <sub>2</sub> O	9.14	9.84	9.91
248	4-iodofenol <sup>a, c</sup>	DMF	14.31	15.47	14.62
249	4-iodofenol <sup>a, c</sup>	H <sub>2</sub> O	9.20	9.28	9.07
250	4-metilfenol <sup>a, c</sup>	AN	27.45	25.66	26.41
251	4-metilfenol <sup>a, c</sup>	DMSO	16.96	16.11	16.85
252	4-metilfenol <sup>a, c</sup>	H <sub>2</sub> O	10.28	10.10	10.07
253	4-metilfenol <sup>d</sup>	MeOH	14.52	14.58	14.51
254	4-metoxifenol <sup>a, c</sup>	H <sub>2</sub> O	10.27	10.33	10.25

#	Fenol	Solvent	Exp.	BPNN	SVM
255	4-metoxifenol <sup>b, c</sup>	DMSO	17.58	16.19	17.67
256	4-nitrofenol <sup>a, c</sup>	DMA	10.85	11.42	11.48
257	4-nitrofenol <sup>a, c</sup>	DMF	11.84	11.01	11.27
258	4-nitrofenol <sup>a, c</sup>	DMSO	11.00	10.20	11.09
259	4-nitrofenol <sup>a, c</sup>	MeOH	11.24	11.34	11.26
260	4-nitrofenol <sup>a, c</sup>	NM	20.10	19.95	20.01
261	4-nitrofenol <sup>a, c</sup>	<sup>t</sup> BuOH	14.60	14.71	14.69
262	4-nitrofenol <sup>b, c</sup>	<sup>i</sup> PrOH	12.45	11.85	12.36
263	4-nitrofenol <sup>d</sup>	AN	20.70	20.55	19.42
264	4-nitrofenol <sup>d</sup>	H <sub>2</sub> O	7.18	6.91	7.67
265	4-t-butilfenol <sup>a, c</sup>	AN	27.48	25.64	27.39
266	4-t-butilfenol <sup>a, c</sup>	H <sub>2</sub> O	10.31	10.86	10.80
267	4-t-butilfenol <sup>a, c</sup>	MeOH	14.52	15.06	15.53
268	fenol <sup>a, c</sup>	DMF	15.40	16.41	16.29
269	fenol <sup>a, c</sup>	DMSO	16.47	16.15	16.56
270	fenol <sup>a, c</sup>	H <sub>2</sub> O	9.99	9.78	10.08
271	fenol <sup>a, c</sup>	MeOH	14.32	14.45	14.23
272	fenol <sup>b, c</sup>	NM	25.70	25.20	25.61
273	fenol <sup>d</sup>	AN	26.60	25.67	25.57
274	pentaclorofenol <sup>a, c</sup>	DMF	7.97	6.66	8.06
275	pentaclorofenol <sup>a, c</sup>	DMSO	7.05	6.45	6.89
276	pentaclorofenol <sup>d</sup>	Ac	18.30	18.57	16.33

<sup>a</sup> Training set de BPNN. <sup>b</sup> Cross-validation set de BPNN.

<sup>c</sup> Training set de SVM. <sup>d</sup> Prediction set de BPNN i SVM.



### 4.2.2. $pK_a$ d'àcids benzoics

Aquesta família de compostos és, estructuralment parlant, molt semblant a la família dels fenols, així doncs, és d'esperar que la metodologia QSPR sigui capaç d'analitzar els valors de  $pK_a$  per als compostos que la formen. En aquest cas també existeix una gran analogia entre els àcids benzoics, ja que tots presenten un anell aromàtic; però alhora, degut a la gran quantitat de substituents diferents i a les diferents posicions de substitució, presenten diferències de  $pK_a$  molt grans, abastant un rang de més de vint unitats. Per totes aquestes raons, i degut a l'èxit aconseguit amb la família dels fenols, és possible pensar que la metodologia QSPR pot assumir l'establiment de models que permetin el càlcul de  $pK_a$  per aquesta família de compostos.

#### *Conjunt de dades*

El conjunt de dades construït a partir de la recerca bibliogràfica conté 519 valors de  $pK_a$  provinents de la combinació de 136 àcids benzoics diferents en 9 solvents. De nou, els solvents es poden dividir en pròtics:  $H_2O$ , MeOH, EtOH i  $iPrOH$ ; i apròtics: Ac, AN, DMSO, DMF i DMA. En aquest cas, les fonts bibliogràfiques d'on s'han obtingut aquests valors són molt més variades que en el cas dels fenols; així, els valors de  $pK_a$  en solvents pròtics s'han obtingut preferentment dels treballs de Pytela,<sup>186-191</sup> Bosch,<sup>181,192</sup> Tehan,<sup>147</sup> Kolthoff<sup>193,194</sup> i alguns altres;<sup>195-212</sup> mentre que els valors de  $pK_a$  en solvents apròtics s'han trobat preferentment en treballs de Pytela,<sup>186-191</sup> Exner<sup>195,213</sup> i en la compilació d'Izutsku.<sup>183</sup>

Els 136 àcids benzoics presenten diferents tipus de substitució, així com una gran varietat de substituents disponibles. Apart de l'àcid benzoic, hi ha 49 àcids benzoics monosubstituïts, 68 disubstituïts, 14 trisubstituïts, 3 tetrasubstituïts i 1 pentasubstituït. Els diferents substituents, 24 en total, són els següents: metil, etil, *tert*-butil, metoxi, etoxi, propoxi, isopropoxi, butoxi, acetoxi, acetamido, èster metil·lic, fluorur, clorur, bromur, iodur, hidroxil, nitro, trifluorometil, cianur, formil, acetil, amino, dimetilamino i *N,N*-dimetilaminocarbonil. Degut a la seva diferent naturalesa química, aquests substituents presenten efectes electrònics i estèrics molt diferenciats. De nou, en els compostos polihidroxilats el valor de  $pK_a$  triat es correspon amb el de la primera constant de dissociació de l'àcid benzoic. Del conjunt de 136 àcids benzoics, hi ha 66 que presenten substituents en *orto*-, aquest tipus de substitució és força important ja que

els substituents en aquesta posició poden interactuar amb el grup OH que es dissocia formant enllaços d'hidrogen o afectant a la dissociació d'aquest enllaç per efectes estèrics. En el conjunt de dades es troben 238 valors de  $pK_a$  d'àcids benzoics *orto*-substituïts i 281 valors per àcids benzoics *meta*- o *para*-substituïts, més el propi àcid benzoic. El nombre de valors de  $pK_a$  en cada solvent és molt diferent: 105 en aigua, 107 en metanol, 23 en etanol, 18 en isopropanol, 84 en DMSO, 58 en DMF, 63 en acetonitril, 53 en acetona i 6 en DMA. Així, dels 519 valors de  $pK_a$  experimentals, 253 corresponen a mesures en solvents pròtics i 266 corresponen a solvents apròtics. El rang de  $pK_a$  que cobreixen aquests compostos en els nou solvents esmentats és molt ampli: des de 0.65 per a l'àcid 2,4,6-trinitrobenzoic en aigua a 21.29 per a l'àcid 3-metil-4-metoxibenzoic en acetonitril. La Taula 4.23 conté tots els valors de  $pK_a$  experimentals i calculats per a aquesta família de compostos.

Abans de començar a treballar s'ha de dividir el conjunt de dades en *training*, *cross-validation* i *prediction sets*. El primer pas consisteix en crear *tset* i *pset* amb un 80% i un 20% dels valors totals, respectivament, aquests subconjunts de dades són els que s'empraran per establir el model amb el mètode SVM. Així, el *tset* consta de 421 valors de  $pK_a$  mentre que el *pset* en té 98. Un cop feta aquesta primera divisió, s'agafarà un 10% de valors del *tset* per construir el *cvset*, necessari per establir els models amb les xarxes neuronals que incorpora el programa ADAPT. D'aquesta manera, els tres subconjunts de dades d'ADAPT: *tset*, *cvset* i *pset*, contenen 379, 42 i 98 valors de  $pK_a$ , respectivament.

### **Model no lineal: Xarxes neuronals BPNN**

A partir de la geometria optimitzada per MOPAC, el programa CODESSA calcula 604 descriptors dels àcids benzoics. A aquests descriptors se'ls sumen els 37 descriptors dels solvents, els qual s'importen per donar un total de 641 descriptors inicials. El mètode heurístic de CODESSA redueix aquest nombre inicial de descriptors a 240, dels quals 19 són dels solvents. Per establir els models no lineals amb BPNN s'importen aquests descriptors al programa ADAPT i es procedeix a realitzar la segona reducció en el nombre de descriptors. En aquesta ocasió es passa dels 240 descriptors inicials a un total de 82, 73 dels àcids benzoics i 9 dels solvents. L'arquitectura optimitzada de les xarxes neuronals és 7-17-1, el nombre de paràmetres ajustables de la

xarxa neuronal amb aquesta arquitectura és de 154 i el seu *ratio* envers el nombre de compostos en el *training set*, 379, és de 2.43, per sobre del 2.0 mínim acceptable.

El model triat per la xarxa neuronal d'arquitectura 7-17-1 conté els descriptors que es troben en la Taula 4.24. Com en el cas dels fenols, aquest model té cinc descriptors dels soluts i dos dels solvents.

Taula 4.24. Descriptors continguts en el model dels àcids benzoics amb BPNN.

Descriptor	Component	Classe
Càrrega parcial màxima d'un àtom d'hidrogen	Solut	Electroestàtic
Àrea superficial positiva fraccionada, FPSA-2	Solut	Electroestàtic
Energia de ressonància mínima d'un enllaç H-O	Solut	Quàntic
València mínima d'un àtom de carboni	Solut	Quàntic
Repulsió electró-electró màxima d'un àtom d'oxigen	Solut	Quàntic
Capacitat de donació d'enllaç d'hidrogen de Kamlet i Taft, $\alpha$	Solvent	-
Energia de vaporització molar, $\Delta_{\text{vap}}U^\circ$	Solvent	-

Els descriptors dels àcids benzoics pertanyen a dues classes diferents, electroestàtics i quàntics. Els descriptors electroestàtics, que es deriven de la distribució de càrregues parcials calculades pel mètode de Zefirov<sup>73</sup> a partir de les electronegativitats de Sanderson, són la càrrega parcial màxima sobre un àtom d'hidrogen i l'àrea superficial positiva fraccionada, FPSA-2. Com en el cas dels fenols, la càrrega parcial màxima sobre un àtom d'hidrogen se situa sempre sobre el protó del grup àcid, és a dir, aquell que es dissocia de la molècula. El descriptor FPSA-2, de tipus CPSA, es calcula de la següent manera:

$$FPSA-2 = \frac{\sum_A q_A \sum_A S_A}{TMSA}$$

on  $q_A$  són les càrregues atòmiques parcials positives,  $S_A$  són les superfícies atòmiques accessibles al solvent carregades positivament i  $TMSA$  és la superfície total de la molècula. Els descriptors quàntics s'obtenen directament amb el càlcul semiempíric AM1 implementat en el programa MOPAC; aquests descriptors són l'energia de ressonància mínima d'un enllaç H-O, la valència mínima d'un àtom de carboni i l'energia de repulsió electró-electró màxima d'un àtom d'oxigen. Aquests descriptors semblen força adients per a la reacció que s'està estudiant ja que la seva disposició en la

La molècula permetrà la interpretació posterior del model: el descriptor energia de ressonància mínima d'un enllaç H-O es troba sempre localitzat en l'enllaç que es dissocia, la valència mínima d'un àtom de carboni es troba sempre sobre el carboni del grup àcid i, per últim, l'energia de repulsió electró-electró màxima d'un àtom d'oxigen se situa sobre l'oxigen carbonílic del grup àcid. La Figura 4.18 permet visualitzar més fàcilment on es troben localitzats aquests descriptors.

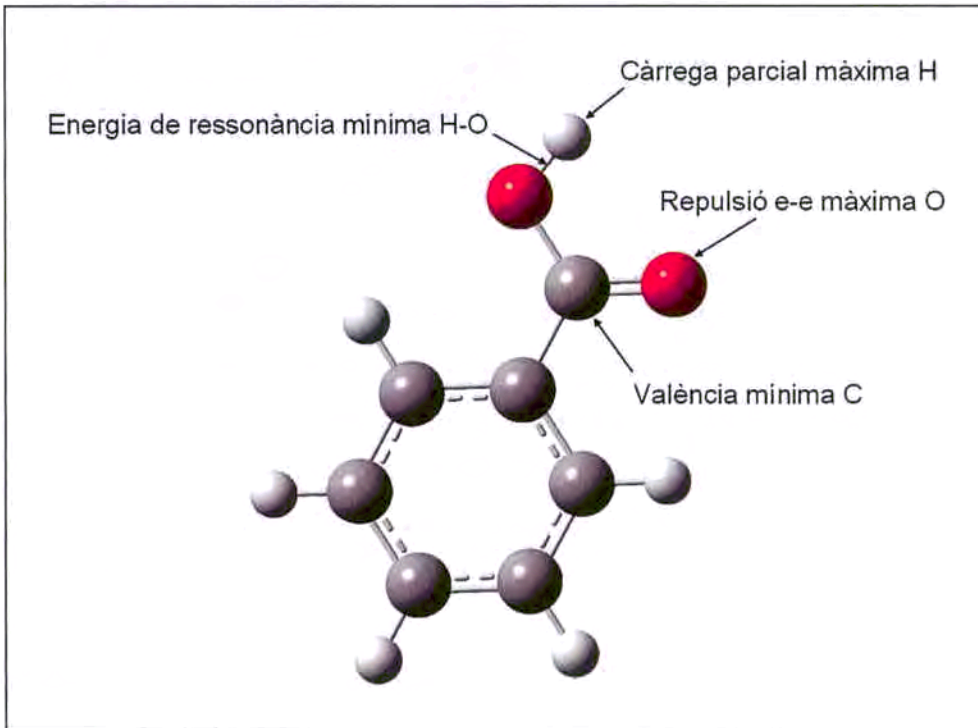


Figura 4.18. Localització dels descriptors del model sobre els àcids benzoïcs.

Els descriptors del solvent són  $\alpha$  de Kamlet i Taft,<sup>160</sup> que ja s'havia trobat en el model dels fenols i l'energia de vaporització molar,  $\Delta_{\text{vap}}U^\circ$ . El descriptor  $\alpha$  mesura la capacitat de les molècules de solvent per establir enllaços donadors d'hidrogen amb els soluts. L'energia de vaporització molar del solvent,  $\Delta_{\text{vap}}U^\circ$ , es defineix com la diferència d'energia molar entre el gas ideal a pressió normal i el líquid saturat a 25°C. Aquest descriptor és una mesura de les forces d'interacció entre les molècules del solvent (energia de cohesió), és a dir, mesura la capacitat del solvent per formar les cavitats en les quals s'han d'encabir els soluts.

Els resultats estadístics que s'obtenen amb aquest model són molt bons, la Taula 4.25 conté els valors del coeficient de determinació  $R^2$ , el test- $F$  i el  $rmse$  per als tres subconjunts de dades:  $tset$ ,  $cvset$  i  $pset$ . Els valors de  $pK_a$  calculats amb aquest model es troben en la Taula 4.23 i la seva representació gràfica enfront dels valors experimentals es pot veure en la Figura 4.19.

Taula 4.25. Paràmetres estadístics obtinguts amb el model BPNN.

Conjunt	n	$R^2$	F	$rmse$
<i>tset</i>	379	0.998	222858.0	0.21
<i>cvset</i>	42	0.998	24060.1	0.21
<i>pset</i>	98	0.998	56520.1	0.21

Els descriptors d'aquest model són coherents amb el procés de dissociació que es dona en funció de les interaccions solut/solvent que es produeixen en aquest sistema. Així, els descriptors càrrega parcial màxima sobre un àtom d'hidrogen, FPSA-2, energia de ressonància mínima d'un enllaç H-O i valència mínima per a un àtom de carboni de l'àcid benzoic fan referència a les interaccions solut/solvent no específiques. El descriptor de solvent energia de vaporització molar ( $\Delta_{vap}U^\circ$ ) conté informació sobre aquest mateix tipus d'interaccions. Els altres dos descriptors del model contenen informació sobre les interaccions específiques, o acceptor/donador, que s'estableixen en el sistema. Aquests descriptors són l'energia de repulsió electró-electró màxima per a un àtom d'oxigen i la  $\alpha$  de Kamlet i Taft, per a soluts i solvents respectivament.



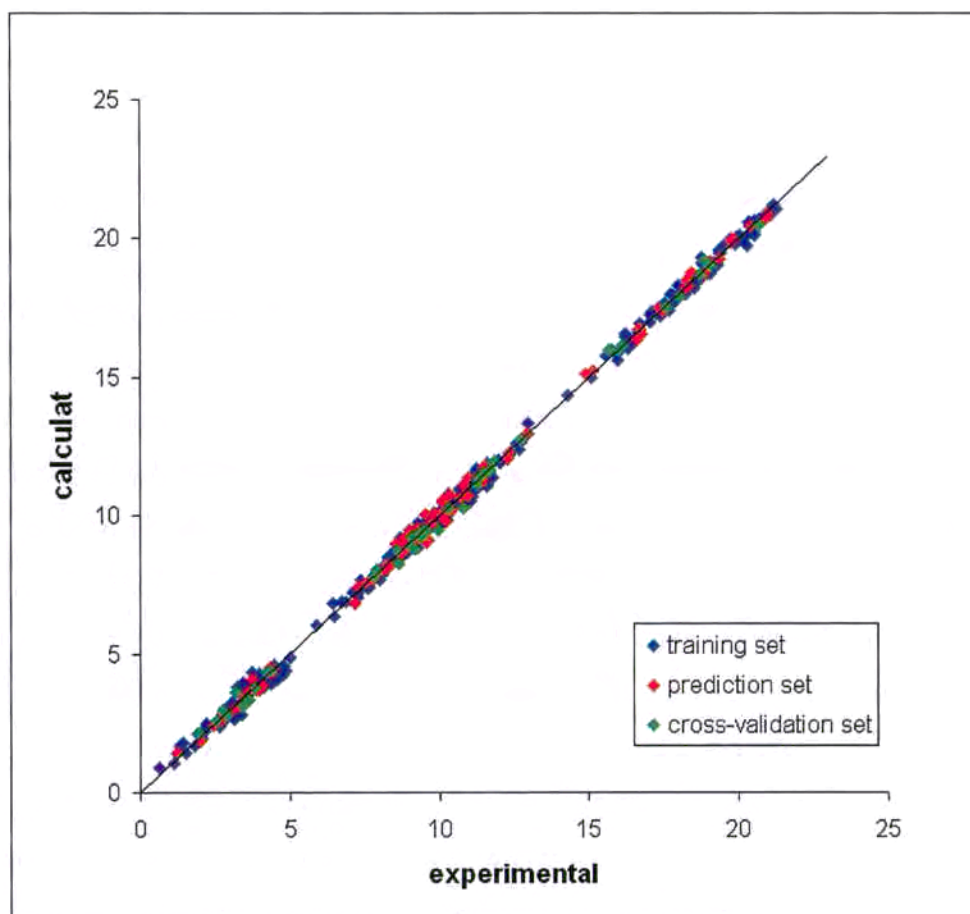


Figura 4.19. Representació gràfica de  $pK_a$  calculat vs. experimental per als àcids benzoics amb BPNN.

El descriptor càrrega parcial màxima sobre un àtom d'hidrogen se situa sempre sobre el protó carboxílic dels àcids benzoics, aquest descriptor reflecteix la polaritat de l'enllaç O-H que es trenca i mesura la facilitat amb la qual es trenca per donar el benzoat conjugat i un protó. A l'igual que passava en els fenols, els valors més elevats d'aquest descriptor es troben en àcids benzoics que tenen un gran nombre de substituents electroatracients; per contra, els àcids benzoics amb valors més baixos per aquest descriptor solen ser els que presenten substituents alquílics. De la mateixa manera, el descriptor energia de ressonància mínima d'un a enllaç H-O es relaciona amb l'enllaç que es trenca en la reacció de dissociació ja que, en tots els casos, aquest descriptor se situa en l'enllaç O-H del grup carboxílic. Tal i com passava amb el descriptor anterior, hi ha una relació directa entre el valor d'aquest descriptor i els substituents de l'anell aromàtic, sent més petit el valor d'aquest descriptor per aquells

àcids benzoics substituïts pels grups més electroatracients i en posicions *orto*. Aquest mateix comportament es troba pel tercer descriptor no específics dels àcids benzoics, la valència mínima d'un àtom de carboni presenta valors baixos per als àcids benzoics polisubstituïts amb grups electroatracients i valors més elevats per als grups alquilics, reflectint l'efecte electrònic dels substituents en l'anell aromàtic. Aquests tres descriptors fan referència a la polarització de l'enllaç que es trenca en la reacció de dissociació i recullen alhora els efectes dels substituents, tant estèrics com electrònics, i de les posicions de substitució sobre l'anell. El descriptor FPSA-2 pertany a la classe CPSA, conté informació sobre la distribució de càrrega en la molècula i codifica informació sobre les possibles interaccions electrostàtiques i polars que es poden donar entre les molècules de solut i solvent. El descriptor dels solvents energia de vaporització molar és, com ja s'ha comentat, una mesura de les forces de cohesió que s'estableixen entre les molècules de solvent. Del valor d'aquest descriptor depèn la capacitat del solvent per formar cavitats en el seu interior i, així, possibilitar la dissolució de les molècules de solut. Així doncs, aquest descriptor conté, a més de l'energia involucrada en la formació de cavitats, la relacionada amb les cavitats que contenen els ions generats en la reacció de dissociació. El descriptor quàntic energia de repulsió electró-electró màxima d'un àtom d'oxigen indica la capacitat d'interacció dels àcids benzoics a través d'enllaços d'hidrogen. Aquest descriptor, que habitualment se situa sobre l'oxigen carbonílic dels àcids benzoics, pren valors més elevats quan la densitat electrònica sobre aquest àtom és més gran, fent que la capacitat de l'àcid benzoic per formar enllaços d'hidrogen a través d'aquest àtom augmenti. La combinació d'aquest descriptor amb el descriptor  $\alpha$  del solvent descriu les interaccions específiques acceptor/donador, ja que aquest descriptor correspon a la capacitat dels solvents per formar enllaços d'hidrogen per donació.

Donat que el model que s'ha construït és no lineal, l'única manera d'avaluar la importància relativa de cada descriptor dintre del model consisteix en calcular els increments de *rmse* amb el mètode de Jurs. Aquest mètode proporciona increments en el *rmse* del model segons la importància de cadascun dels descriptors que conté quan es fan permutes internes. Els resultats que s'obtenen amb aquest mètode es troben en la Taula 4.26.

Taula 4.26. Increments de *rmse* per a la permuta de cada descriptor.

Descriptor	Increment <i>rmse</i>	Importància
Càrrega parcial màxima d'un àtom d'hidrogen	0.23	7
Àrea superficial positiva fraccionada, FPSA-2	0.26	6
Energia de ressonància mínima d'un enllaç H-O	0.44	4
València mínima d'un àtom de carboni	0.46	3
Repulsió e-e màxima d'un àtom d'oxigen	0.39	5
Capacitat de donació d'enllaç d'hidrogen de Kamlet i Taft, $\alpha$	6.30	1
Energia de vaporització molar, $\Delta_{\text{vap}}U^\circ$	3.70	2

Aquests resultats són molt anàlegs als obtinguts en el model dels fenols ja que, en tots dos casos, els descriptors dels solvents són molt més importants que els dels soluts. En aquest cas, els descriptors més importants, apart de  $\alpha$  i  $\Delta_{\text{vap}}U^\circ$ , són: valència mínima d'un àtom de carboni, energia de ressonància mínima de l'enllaç H-O i energia de repulsió electró-electró màxima d'un àtom d'oxigen.

El valor de *rmse* per al conjunt de 519 compostos és 0.21. Es pot comprovar que no hi ha cap compost que presenti un valor de residual superior a tres cops el valor de *rmse*. Hi ha 31 compostos per als quals s'obtenen residuals més grans que dues vegades l'error del model, aquests compostos es troben en la Taula 4.27. Tot i així, dintre d'aquests compostos no s'aprecia cap tipus de tendència en l'error comès deguda a la naturalesa de l'àcid benzoic ni del solvent. Aquests 31 valors corresponen a la combinació de 24 àcids benzoics diferents en sis solvents, hi ha 15 residuals positius i 16 negatius, 17 d'ells corresponen a mesures en solvents pròtics i 14 en solvents apròtics. D'altra banda, no s'aprecia cap diferenciació entre els àcids benzoics que presenten substituents en *orto*- (12) o en *meta*- o *para*- (19).

Taula 4.27. Prediccions amb valors més grans que dues vegades el *rmse*.

Solut	Solvent	$pK_a$ exp.	$pK_a$ calc.	residual
Àcid 2,3-dimetilbenzoic	H <sub>2</sub> O	3.74	4.33	-0.59
Àcid 3-metoxi-4-nitrobenzoic	AN	18.77	19.35	-0.58
Àcid 2,6-dimetilbenzoic	H <sub>2</sub> O	3.25	3.80	-0.55
Àcid 3-clorobenzoic	DMSO	9.51	10.05	-0.54
Àcid 2-metil-6-metoxibenzoic	H <sub>2</sub> O	3.46	3.98	-0.52
Àcid 2-metil-6-metoxibenzoic	MeOH	8.70	9.21	-0.51
Àcid 2,3-dimetilbenzoic	MeOH	8.98	9.48	-0.50
Àcid 4-clorobenzoic	DMSO	10.10	10.58	-0.48
Àcid 3-metoxi-4-nitrobenzoic	DMF	10.33	10.80	-0.47
Àcid 4-cianobenzoic	DMSO	9.27	9.73	-0.46
Àcid 2,3,5-trimetilbenzoic	MeOH	9.05	9.50	-0.45
Àcid 2,3,5-trimetilbenzoic	DMSO	11.24	11.69	-0.45
Àcid 3-clorobenzoic	DMF	10.95	11.39	-0.44
Àcid 2-metilèsterbenzoic	H <sub>2</sub> O	3.18	3.62	-0.44
Àcid 3,5-dibromobenzoic	DMF	10.11	10.54	-0.43
Àcid 4- <i>tert</i> -butilbenzoic	H <sub>2</sub> O	4.36	3.93	0.43
Àcid 4-metoxibenzoic	DMSO	11.80	11.37	0.43
Àcid 2-acetilbenzoic	MeOH	9.28	8.84	0.44
Àcid 2-aminobenzoic	H <sub>2</sub> O	4.87	4.41	0.46
Àcid 3-cloro-5-nitrobenzoic	H <sub>2</sub> O	3.13	2.66	0.47
Àcid 3-metil-4-nitrobenzoic	DMF	11.03	10.55	0.48
Àcid 3-aminobenzoic	H <sub>2</sub> O	4.74	4.25	0.49
Àcid 3,4-dimetilbenzoic	<i>i</i> PrOH	11.60	11.11	0.49
Àcid 2-formilbenzoic	H <sub>2</sub> O	4.55	4.05	0.50
Àcid 4- <i>tert</i> -butilbenzoic	MeOH	9.61	9.11	0.50
Àcid 2-nitrobenzoic	DMF	9.96	9.46	0.50
Àcid 3-aminobenzoic	DMSO	11.60	11.08	0.52
Àcid 4-nitrobenzoic	DMF	10.80	10.28	0.52
Àcid 3-hidroxibenzoic	MeOH	9.58	9.04	0.54
Àcid 4-bromobenzoic	AN	20.30	19.75	0.55
Àcid 3-amino-2,5-diclorobenzoic	H <sub>2</sub> O	3.40	2.82	0.58

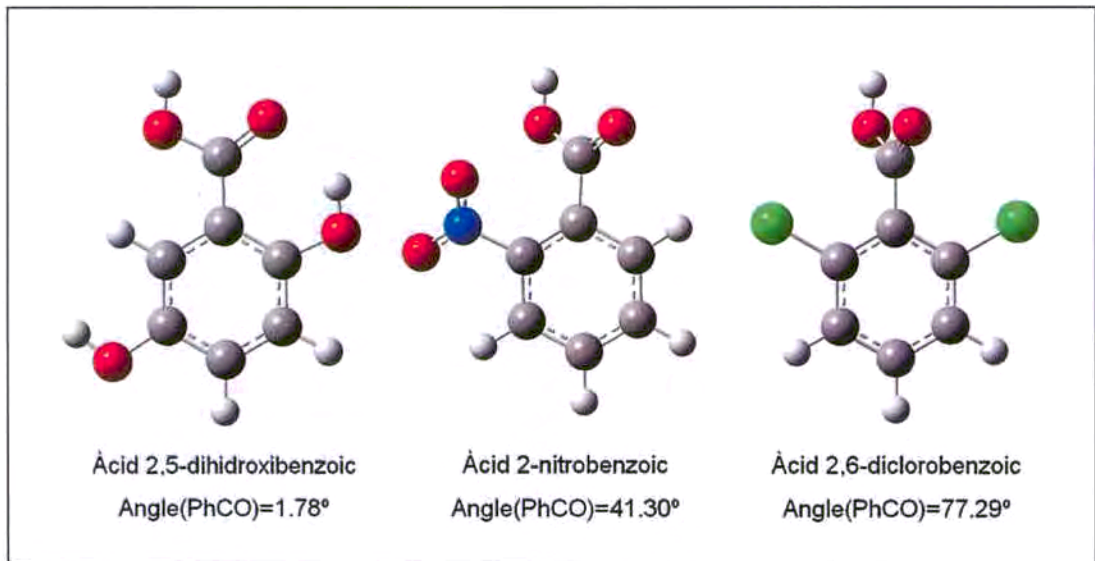
Es pot comprovar la robustesa del model analitzant per separat tots els subconjunts de compostos que es poden formar a partir del conjunt inicial de 519 valors de  $pK_a$ . La Taula 4.28 mostra tots aquests subconjunts i els paràmetres estadístics que s'obtenen per a cadascun d'ells.

Taula 4.28. Estadística dels subconjunts d'àcids benzoics.

Subconjunt	n	R <sup>2</sup>	F	rmse
<i>orto</i>	238	0.998	146062.4	0.21
No <i>orto</i>	281	0.998	156412.6	0.22
Pròtics	253	0.994	43692.0	0.22
Apròtics	266	0.997	105055.2	0.21
H <sub>2</sub> O	105	0.937	1537.2	0.22
MeOH	107	0.939	1605.6	0.20
EtOH	23	0.963	540.3	0.19
<sup>i</sup> PrOH	18	0.967	465.2	0.23
DMSO	84	0.966	2360.1	0.23
DMF	58	0.962	1417.6	0.23
AN	63	0.963	1589.8	0.21
Ac	55	0.980	2533.0	0.16
DMA	6	0.925	49.1	0.17

En aquest cas tampoc no es pot apreciar una diferència significativa entre els subconjunts pròtic/apròtic i *orto*/no *orto*. El comportament d'aquests últims subconjunts no s'adiu amb el trobat en altres estudis<sup>156</sup> on l'efecte *orto* és molt marcat per als àcids benzoics. En aquest cas s'ha comprovat que l'angle díedre de l'enllaç que formen el grup carboxílic i l'anell aromàtic pot variar entre 0 i 90 graus, alguns exemples d'aquest comportament es poden trobar en la Figura 4.20. Tot i així, el model construït és capaç de predir correctament el  $pK_a$  dels àcids benzoics *orto*-substituïts independentment d'aquest angle, per la qual cosa, es pot dir que l'efecte *orto* es troba incorporat en el model. Pel que fa als resultats que s'obtenen individualment per a cada solvent, es pot comprovar com els paràmetres estadístics de tots ells són molt bons, incloent valors molt regulars de desviació estàndard en el càlcul d'unes 0.2 unitats de  $pK_a$  per a tots els ells.




 Figura 4.20. Variació de l'angle diedre en àcids benzoics *ortho*-substituïts.

És possible avaluar també com funciona el model comprovant els resultats que s'obtenen per a un mateix àcid benzoic en diferents solvents. Així, la Taula 4.29 mostra els paràmetres estadístics que s'obtenen analitzant 38 àcids benzoics específics en diferents solvents, amb un mínim de sis.

Taula 4.29. Paràmetres estadístics obtinguts per a àcids benzoics en diferents solvents.

Solut	Solvents	R <sup>2</sup>	rmse
Àcid 2,6-dibromobenzoic	7	0.999	0.13
Àcid 2,6-dibutoxibenzoic	8	1.000	0.09
Àcid 2,6-diclorobenzoic	8	0.999	0.14
Àcid 2,6-dietoxibenzoic	7	1.000	0.09
Àcid 2,6-difluorobenzoic	7	1.000	0.03
Àcid 2,6-diiodobenzoic	7	0.997	0.25
Àcid 2,6-diisopropoxibenzoic	7	0.998	0.22
Àcid 2,6-dimetoxibenzoic	8	0.999	0.16
Àcid 2,6-dimetilbenzoic	8	0.999	0.14
Àcid 2,6-dinitrobenzoic	8	0.999	0.17
Àcid 2,6-dipropoxibenzoic	7	0.999	0.12
Àcid 2-acetilbenzoic	6	1.000	0.10
Àcid 2-cloro-6-nitrobenzoic	6	0.999	0.18
Àcid 2-clorobenzoic	6	0.999	0.18
Àcid 2-metil-6-clorobenzoic	6	1.000	0.13
Àcid 2-metil-6-metoxibenzoic	6	0.998	0.13
Àcid 2-metil-6-nitrobenzoic	6	0.999	0.18
Àcid 2-nitrobenzoic	7	0.999	0.17

Àcid 3,4-diclorobenzoic	7	0.999	0.16
Àcid 3,4-dimetoxibenzoic	6	1.000	0.14
Àcid 3,4-dimetilbenzoic	7	0.999	0.17
Àcid 3,4-dinitrobenzoic	6	0.999	0.17
Àcid 3,5-dimetoxibenzoic	6	1.000	0.14
Àcid 3,5-dimetilbenzoic	6	1.000	0.08
Àcid 3,5-dinitrobenzoic	7	0.998	0.26
Àcid 3-bromo-4-metilbenzoic	6	1.000	0.12
Àcid 3-bromobenzoic	7	1.000	0.11
Àcid 3-cloro-5-nitrobenzoic	6	1.000	0.13
Àcid 3-metoxi-4-metilbenzoic	6	1.000	0.09
Àcid 3-metil-4-clorobenzoic	6	1.000	0.06
Àcid 3-metil-4-metoxibenzoic	6	1.000	0.08
Àcid 3-metil-4-nitrobenzoic	6	1.000	0.14
Àcid 3-nitro-4clorobenzoic	7	1.000	0.12
Àcid 3-nitro-4-metoxienzoic	6	0.999	0.21
Àcid 3-nitro-4-metilbenzoic	6	1.000	0.11
Àcid 4-cianobenzoic	6	0.999	0.17
Àcid 4-nitrobenzoic	8	0.998	0.21
Àcid benzoic	9	1.000	0.06

La capacitat predictiva que mostra el model, tan globalment com individual, és clarament superior a la que s'ha obtingut amb altres intents previs.<sup>147,149</sup> Cal recordar que en aquests intents s'han trobat coeficients de determinació baixos ( $R^2 \approx 0.8$ ) per diferents conjunts d'àcids benzoics en aigua. Només, en un treball on s'utilitzen els models de Hammet-Taft i Drago<sup>156</sup> per predir el  $pK_a$  de 48 àcids benzoics en metanol s'obtenen resultats semblants als aquí trobats. La diferència trobada entre el model aquí establert i càlculs teòrics amb models de polarització contínua no es pot fer de manera directa ja que, en aquests treballs, el conjunt de compostos conté substàncies de naturalesa química molt variada. Tot i així, els estudis en aigua,<sup>143</sup> DMSO<sup>144-146</sup> i acetonitril<sup>144</sup> troben precisions d'entre una i dues unitats de  $pK_a$ , bastant per sota de les que s'aconsegueix amb el model aquí descrit.

**Model no lineal: Support Vector Machines**

El model desenvolupat per *Support Vector Machines* conté els mateixos descriptors que el model construït amb BPNN. L'optimització dels paràmetres  $\gamma$ ,  $C$  i  $\epsilon$  condueix a l'establiment definitiu del model no lineal, Figura 4.21. L'optimització del paràmetre  $\gamma$  (amb  $\epsilon=0.1$  i  $C=100$ ) presenta un valor de *rmse* mínim quan  $\gamma=0.9$ . L'optimització de  $\epsilon$  (amb  $\gamma=0.9$  i  $C=100$ ) condueix a un valor de 0.09 per a aquest paràmetre. L'optimització del paràmetre  $C$  (amb  $\gamma=0.9$  i  $\epsilon=0.09$ ) demostra que el seu valor òptim és 100.

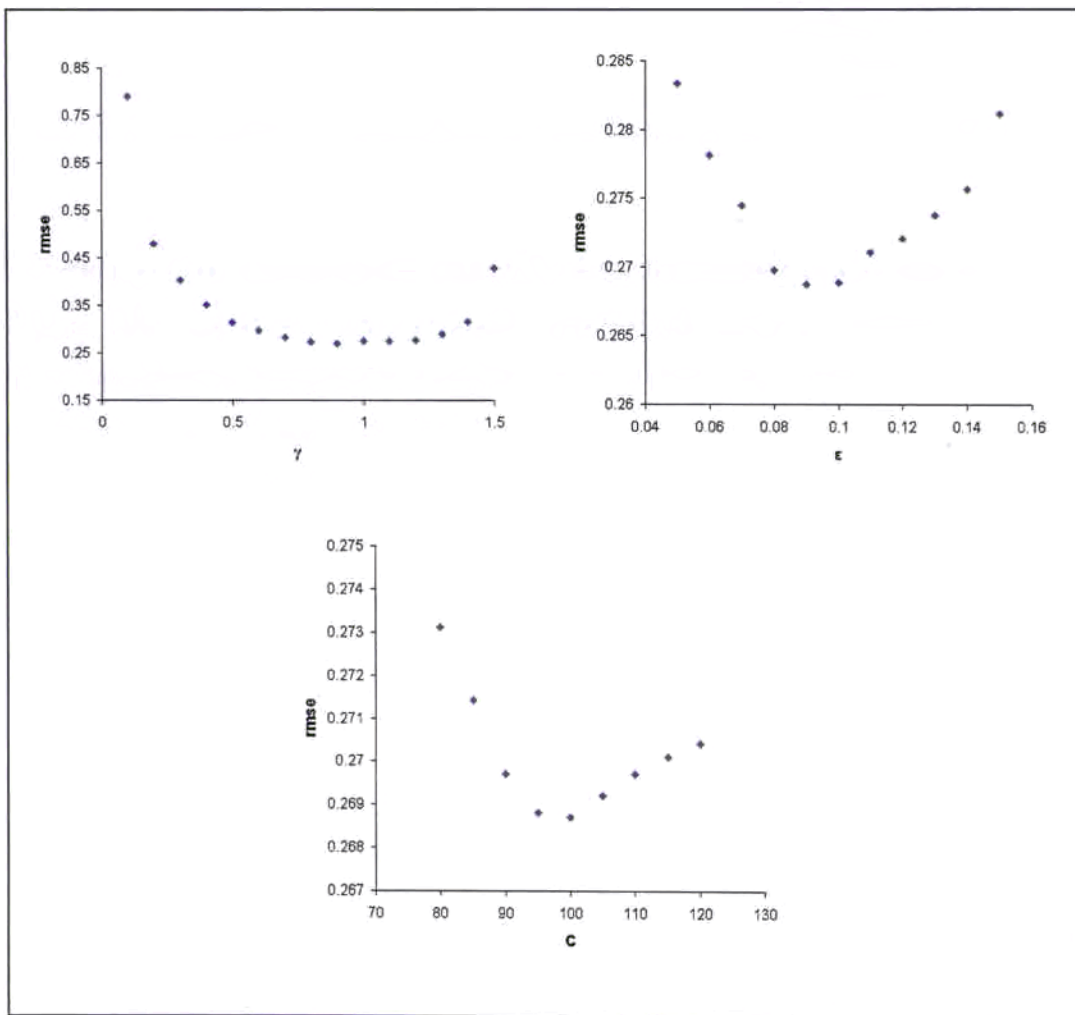


Figura 4.21. Optimització dels paràmetres  $\gamma$ ,  $\epsilon$  i  $C$ .

El model desenvolupat amb LibSVM amb aquests descriptors consta de 308 vectors suport escollits d'entre els 421 compostos que formen el *tset*. Els resultats que s'obtenen amb aquest model són pràcticament idèntics als que s'obtenen amb el model BPNN, la Taula 4.30 conté els valors del coeficient de determinació  $R^2$ , el test- $F$  i el *rmse* per als dos subconjunts de dades: *tset* i *pset*.

Taula 4.30. Paràmetres estadístics obtinguts amb el model SVM.

Conjunt	n	$R^2$	F	<i>rmse</i>
<i>tset</i>	421	0.998	237085.6	0.22
<i>pset</i>	98	0.998	56556.5	0.21

La representació gràfica dels valors de  $pK_a$  calculat enfront dels experimentals es mostra en la Figura 4.22. Tots els valors de  $pK_a$  calculat es troben a la Taula 4.23.

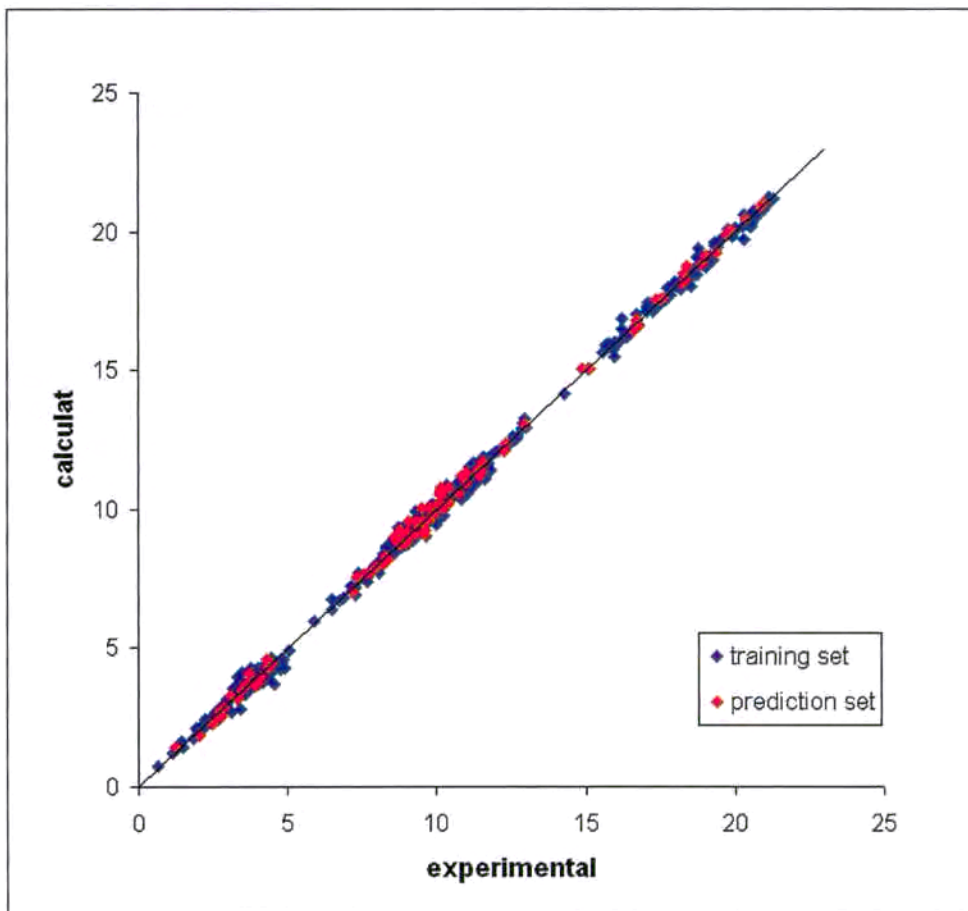


Figura 4.22. Representació gràfica de  $pK_a$  calculat vs. experimental per als àcids benzoics amb SVM.

Els resultats que s'obtenen quan s'analitzen els paràmetres estadístics en tots els subconjunts són també molt bons, Taula 4.31. Aquests resultats són, també, gairebé idèntics als que s'obtenen amb el model BPNN.

Taula 4.31. Estadística dels subconjunts d'àcids benzoïcs.

<b>Subconjunt</b>	<b>n</b>	<b>R<sup>2</sup></b>	<b>F</b>	<b>rmse</b>
<i>orto</i>	238	0.998	136997.4	0.21
No <i>orto</i>	281	0.998	152582.0	0.22
Pròtics	253	0.994	43058.9	0.22
Apròtics	266	0.997	99249.6	0.21
H <sub>2</sub> O	105	0.934	1450.9	0.23
MeOH	107	0.937	1553.7	0.20
EtOH	23	0.971	705.0	0.17
iPrOH	18	0.971	535.8	0.21
DMSO	84	0.966	2351.4	0.23
DMF	58	0.956	1217.3	0.24
AN	63	0.964	1646.9	0.21
Ac	55	0.975	2044.0	0.18
DMA	6	0.935	57.4	0.17



Taula 4.23. Valors experimentals i calculats de  $pK_a$  per a la família dels àcids benzoics.

#	Àcid benzoic	Solvent	Exp.	BPNN	SVM
1	Àcid 2,3,4,5,6-pentametilbenzoic <sup>a,c</sup>	DMSO	11.31	11.08	10.94
2	Àcid 2,3,4,5,6-pentametilbenzoic <sup>b,c</sup>	MeOH	8.66	8,25	8.57
3	Àcid 2,3,4,5-tetrametilbenzoic <sup>a,c</sup>	MeOH	9.24	9.38	9.51
4	Àcid 2,3,4,5-tetrametilbenzoic <sup>d</sup>	DMSO	11.58	11.71	11.70
5	Àcid 2,3,4,6-tetrametilbenzoic <sup>a,c</sup>	DMSO	11.17	11.12	11.11
6	Àcid 2,3,4,6-tetrametilbenzoic <sup>a,c</sup>	MeOH	8.61	8.71	8.70
7	Àcid 2,3,4-trimetilbenzoic <sup>b,c</sup>	DMSO	11.26	11.15	11.16
8	Àcid 2,3,4-trimetilbenzoic <sup>a,c</sup>	MeOH	9.19	8.97	9.08
9	Àcid 2,3,5,6-tetrametilbenzoic <sup>a,c</sup>	DMSO	10.89	11.22	11.12
10	Àcid 2,3,5,6-tetrametilbenzoic <sup>a,c</sup>	MeOH	8.51	8.57	8.80
11	Àcid 2,3,5-trimetilbenzoic <sup>a,c</sup>	DMSO	11.24	11.69	11.70
12	Àcid 2,3,5-trimetilbenzoic <sup>d</sup>	MeOH	9.05	9.50	9.58
13	Àcid 2,3,6-triclorobenzoic <sup>a,c</sup>	H <sub>2</sub> O	1.50	1.44	1.45
14	Àcid 2,3,6-trimetilbenzoic <sup>a,c</sup>	DMSO	11.07	11.03	10.98
15	Àcid 2,3,6-trimetilbenzoic <sup>a,c</sup>	MeOH	8.42	8.55	8.75
16	Àcid 2,3-dihidroxibenzoic <sup>b,c</sup>	H <sub>2</sub> O	2.91	2.81	2.88
17	Àcid 2,3-dimetoxibenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.98	4.00	4.08
18	Àcid 2,3-dimetilbenzoic <sup>a,c</sup>	DMSO	11.09	11.51	11.52
19	Àcid 2,3-dimetilbenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.74	4.33	4.30
20	Àcid 2,3-dimetilbenzoic <sup>a,c</sup>	MeOH	8.98	9.48	9.54
21	Àcid 2,4,5-trimetilbenzoic <sup>a,c</sup>	DMSO	11.58	11.89	11.90
22	Àcid 2,4,5-trimetilbenzoic <sup>a,c</sup>	MeOH	9.60	9.75	9.80
23	Àcid 2,4,6-trimetilbenzoic <sup>a,c</sup>	DMSO	10.88	10.77	10.76
24	Àcid 2,4,6-trimetilbenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.45	3.52	3.36
25	Àcid 2,4,6-trimetilbenzoic <sup>a,c</sup>	MeOH	8.57	8.57	8.60
26	Àcid 2,4,6-trimetilbenzoic <sup>a,c</sup>	AN	20.55	20.33	20.31
27	Àcid 2,4,6-trinitrobenzoic <sup>a,c</sup>	H <sub>2</sub> O	0.65	0.88	0.74
28	Àcid 2,4-diclorobenzoic <sup>a,c</sup>	DMSO	9.12	9.06	9.03
29	Àcid 2,4-diclorobenzoic <sup>d</sup>	H <sub>2</sub> O	2.76	2.63	2.53
30	Àcid 2,4-diclorobenzoic <sup>a,c</sup>	MeOH	7.80	7.81	7.79
31	Àcid 2,4-diclorobenzoic <sup>d</sup>	AN	18.40	18.82	18.75
32	Àcid 2,4-dimetoxibenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.36	4.22	4.27
33	Àcid 2,4-dimetilbenzoic <sup>a,c</sup>	DMSO	11.42	11.49	11.50
34	Àcid 2,4-dimetilbenzoic <sup>b,c</sup>	H <sub>2</sub> O	4.22	4.35	4.33
35	Àcid 2,4-dimetilbenzoic <sup>d</sup>	MeOH	9.50	9.53	9.55
36	Àcid 2,4-dinitrobenzoic <sup>a,c</sup>	H <sub>2</sub> O	1.43	1.81	1.63
37	Àcid 2,4-dinitrobenzoic <sup>a,c</sup>	MeOH	6.45	6.81	6.75
38	Àcid 2,5-diclorobenzoic <sup>d</sup>	H <sub>2</sub> O	2.47	2.42	2.29
39	Àcid 2,5-dihidroxibenzoic <sup>a,c</sup>	H <sub>2</sub> O	2.95	2.77	2.86
40	Àcid 2,5-dihidroxibenzoic <sup>a,c</sup>	MeOH	8.04	7.95	8.13
41	Àcid 2,5-dimetilbenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.98	4.29	4.27
42	Àcid 2,5-dimetilbenzoic <sup>d</sup>	DMSO	11.36	11.43	11.44
43	Àcid 2,5-dimetilbenzoic <sup>d</sup>	MeOH	9.29	9.45	9.49
44	Àcid 2,6-dibromobenzoic <sup>a,c</sup>	<sup>t</sup> PrOH	8.61	8.68	8.60
45	Àcid 2,6-dibromobenzoic <sup>a,c</sup>	Ac	15.89	15.83	15.77
46	Àcid 2,6-dibromobenzoic <sup>b,c</sup>	AN	18.04	17.90	18.03
47	Àcid 2,6-dibromobenzoic <sup>a,c</sup>	DMSO	8.61	8.37	8.52
48	Àcid 2,6-dibromobenzoic <sup>b,c</sup>	EtOH	7.93	8.00	7.96
49	Àcid 2,6-dibromobenzoic <sup>a,c</sup>	MeOH	7.14	7.27	7.10
50	Àcid 2,6-dibromobenzoic <sup>d</sup>	DMF	9.72	9.75	10.04

#	Àcid benzoic	Solvent	Exp.	BPNN	SVM
51	Àcid 2,6-dibutoxibenzoic <sup>a,c</sup>	<sup>t</sup> PrOH	11.48	11.28	11.20
52	Àcid 2,6-dibutoxibenzoic <sup>a,c</sup>	Ac	19.00	18.89	18.91
53	Àcid 2,6-dibutoxibenzoic <sup>a,c</sup>	AN	21.13	21.10	21.04
54	Àcid 2,6-dibutoxibenzoic <sup>a,c</sup>	DMF	12.90	12.90	13.10
55	Àcid 2,6-dibutoxibenzoic <sup>a,c</sup>	DMSO	11.45	11.38	11.36
56	Àcid 2,6-dibutoxibenzoic <sup>a,c</sup>	EtOH	10.19	10.28	10.28
57	Àcid 2,6-dibutoxibenzoic <sup>a,c</sup>	MeOH	9.13	9.12	9.12
58	Àcid 2,6-dibutoxibenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.60	3.69	3.69
59	Àcid 2,6-diclorobenzoic <sup>b,c</sup>	Ac	15.68	15.99	15.93
60	Àcid 2,6-diclorobenzoic <sup>a,c</sup>	AN	18.16	18.16	18.25
61	Àcid 2,6-diclorobenzoic <sup>a,c</sup>	DMF	9.71	9.81	10.11
62	Àcid 2,6-diclorobenzoic <sup>a,c</sup>	DMSO	8.31	8.45	8.63
63	Àcid 2,6-diclorobenzoic <sup>a,c</sup>	EtOH	8.03	7.90	8.09
64	Àcid 2,6-diclorobenzoic <sup>a,c</sup>	H <sub>2</sub> O	1.82	1.67	1.73
65	Àcid 2,6-diclorobenzoic <sup>a,c</sup>	MeOH	7.26	7.18	7.17
66	Àcid 2,6-diclorobenzoic <sup>d</sup>	<sup>t</sup> PrOH	8.71	8.60	8.78
67	Àcid 2,6-dietoxibenzoic <sup>a,c</sup>	<sup>t</sup> PrOH	11.09	11.10	11.18
68	Àcid 2,6-dietoxibenzoic <sup>a,c</sup>	AN	20.97	20.93	21.06
69	Àcid 2,6-dietoxibenzoic <sup>a,c</sup>	DMF	12.97	13.00	12.93
70	Àcid 2,6-dietoxibenzoic <sup>b,c</sup>	DMSO	11.38	11.46	11.44
71	Àcid 2,6-dietoxibenzoic <sup>b,c</sup>	MeOH	9.05	8.91	8.96
72	Àcid 2,6-dietoxibenzoic <sup>d</sup>	Ac	18.88	18.73	18.80
73	Àcid 2,6-dietoxibenzoic <sup>d</sup>	EtOH	10.03	10.05	10.19
74	Àcid 2,6-difluorobenzoic <sup>a,c</sup>	<sup>t</sup> PrOH	8.93	8.98	8.88
75	Àcid 2,6-difluorobenzoic <sup>a,c</sup>	Ac	16.03	15.96	15.94
76	Àcid 2,6-difluorobenzoic <sup>a,c</sup>	AN	18.30	18.24	18.21
77	Àcid 2,6-difluorobenzoic <sup>a,c</sup>	DMF	9.91	9.93	10.17
78	Àcid 2,6-difluorobenzoic <sup>a,c</sup>	DMSO	8.59	8.57	8.50
79	Àcid 2,6-difluorobenzoic <sup>a,c</sup>	EtOH	8.22	8.24	8.28
80	Àcid 2,6-difluorobenzoic <sup>a,c</sup>	MeOH	7.39	7.45	7.48
81	Àcid 2,6-dihidroxibenzoic <sup>d</sup>	H <sub>2</sub> O	1.22	1.43	1.43
82	Àcid 2,6-diiodobenzoic <sup>b,c</sup>	<sup>t</sup> PrOH	9.18	8.79	9.01
83	Àcid 2,6-diiodobenzoic <sup>a,c</sup>	Ac	16.34	16.44	16.25
84	Àcid 2,6-diiodobenzoic <sup>a,c</sup>	AN	18.27	18.27	18.36
85	Àcid 2,6-diiodobenzoic <sup>a,c</sup>	EtOH	8.22	8.12	8.41
86	Àcid 2,6-diiodobenzoic <sup>d</sup>	DMF	10.13	10.47	10.80
87	Àcid 2,6-diiodobenzoic <sup>d</sup>	DMSO	8.75	8.85	9.27
88	Àcid 2,6-diiodobenzoic <sup>d</sup>	MeOH	7.29	7.42	7.58
89	Àcid 2,6-diisopropoxibenzoic <sup>a,c</sup>	<sup>t</sup> PrOH	11.61	11.46	11.44
90	Àcid 2,6-diisopropoxibenzoic <sup>a,c</sup>	Ac	19.06	19.20	19.10
91	Àcid 2,6-diisopropoxibenzoic <sup>a,c</sup>	AN	21.17	21.23	21.26
92	Àcid 2,6-diisopropoxibenzoic <sup>a,c</sup>	DMF	12.95	13.37	13.28
93	Àcid 2,6-diisopropoxibenzoic <sup>a,c</sup>	EtOH	10.34	10.34	10.43
94	Àcid 2,6-diisopropoxibenzoic <sup>a,c</sup>	MeOH	9.13	9.08	9.15
95	Àcid 2,6-diisopropoxibenzoic <sup>d</sup>	DMSO	11.50	11.80	11.69
96	Àcid 2,6-dimetoxibenzoic <sup>a,c</sup>	Ac	18.65	18.49	18.56
97	Àcid 2,6-dimetoxibenzoic <sup>b,c</sup>	DMF	12.72	12.79	12.63
98	Àcid 2,6-dimetoxibenzoic <sup>a,c</sup>	DMSO	11.29	11.23	11.20
99	Àcid 2,6-dimetoxibenzoic <sup>a,c</sup>	EtOH	9.84	10.09	10.21
100	Àcid 2,6-dimetoxibenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.44	3.56	3.49
101	Àcid 2,6-dimetoxibenzoic <sup>d</sup>	<sup>t</sup> PrOH	10.81	11.08	11.18

#	Àcid benzoic	Solvent	Exp.	BPNN	SVM
102	Àcid 2,6-dimetoxibenzoic <sup>d</sup>	AN	20.87	20.73	20.87
103	Àcid 2,6-dimetoxibenzoic <sup>d</sup>	MeOH	8.57	8.99	8.98
104	Àcid 2,6-dimetilbenzoic <sup>a,c</sup>	<sup>t</sup> PrOH	10.32	10.50	10.41
105	Àcid 2,6-dimetilbenzoic <sup>a,c</sup>	Ac	17.74	17.99	17.94
106	Àcid 2,6-dimetilbenzoic <sup>a,c</sup>	AN	20.31	20.59	20.64
107	Àcid 2,6-dimetilbenzoic <sup>a,c</sup>	DMF	11.96	12.01	12.07
108	Àcid 2,6-dimetilbenzoic <sup>a,c</sup>	DMSO	10.68	10.94	10.94
109	Àcid 2,6-dimetilbenzoic <sup>a,c</sup>	EtOH	9.42	9.66	9.71
110	Àcid 2,6-dimetilbenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.25	3.80	3.65
111	Àcid 2,6-dimetilbenzoic <sup>a,c</sup>	MeOH	8.57	8.78	8.90
112	Àcid 2,6-dinitrobenzoic <sup>a,c</sup>	<sup>t</sup> PrOH	7.59	7.34	7.50
113	Àcid 2,6-dinitrobenzoic <sup>a,c</sup>	Ac	14.28	14.37	14.19
114	Àcid 2,6-dinitrobenzoic <sup>a,c</sup>	AN	16.67	16.63	16.59
115	Àcid 2,6-dinitrobenzoic <sup>a,c</sup>	DMF	8.26	8.33	8.35
116	Àcid 2,6-dinitrobenzoic <sup>a,c</sup>	DMSO	6.72	6.89	6.74
117	Àcid 2,6-dinitrobenzoic <sup>a,c</sup>	H <sub>2</sub> O	1.14	1.07	1.20
118	Àcid 2,6-dinitrobenzoic <sup>a,c</sup>	MeOH	6.49	6.37	6.40
119	Àcid 2,6-dinitrobenzoic <sup>d</sup>	EtOH	7.15	6.84	7.03
120	Àcid 2,6-dipropoxibenzoic <sup>a,c</sup>	<sup>t</sup> PrOH	11.43	11.47	11.34
121	Àcid 2,6-dipropoxibenzoic <sup>a,c</sup>	Ac	18.98	18.77	18.89
122	Àcid 2,6-dipropoxibenzoic <sup>a,c</sup>	AN	21.08	20.91	21.06
123	Àcid 2,6-dipropoxibenzoic <sup>a,c</sup>	DMSO	11.52	11.50	11.43
124	Àcid 2,6-dipropoxibenzoic <sup>a,c</sup>	EtOH	10.19	10.34	10.36
125	Àcid 2,6-dipropoxibenzoic <sup>a,c</sup>	MeOH	9.19	9.02	9.10
126	Àcid 2,6-dipropoxibenzoic <sup>d</sup>	DMF	12.92	12.99	13.04
127	Àcid 2-acetamidobenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.67	3.75	3.76
128	Àcid 2-acetilbenzoic <sup>a,c</sup>	Ac	17.67	17.41	17.54
129	Àcid 2-acetilbenzoic <sup>a,c</sup>	AN	20.51	20.10	20.21
130	Àcid 2-acetilbenzoic <sup>a,c</sup>	DMF	11.68	11.49	11.53
131	Àcid 2-acetilbenzoic <sup>a,c</sup>	EtOH	10.07	9.73	9.84
132	Àcid 2-acetilbenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.13	3.83	3.74
133	Àcid 2-acetilbenzoic <sup>a,c</sup>	MeOH	9.28	8.84	8.96
134	Àcid 2-aminobenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.87	4.41	4.29
135	Àcid 2-bromobenzoic <sup>a,c</sup>	EtOH	8.98	9.03	8.89
136	Àcid 2-bromobenzoic <sup>d</sup>	H <sub>2</sub> O	2.85	2.85	2.86
137	Àcid 2-bromobenzoic <sup>d</sup>	MeOH	8.19	8.12	8.03
138	Àcid 2-bromobenzoic <sup>d</sup>	DMSO	9.57	9.24	9.24
139	Àcid 2-cloro-3-nitrobenzoic <sup>d</sup>	H <sub>2</sub> O	2.02	1.84	1.85
140	Àcid 2-cloro-4-nitrobenzoic <sup>a,c</sup>	DMSO	7.96	7.85	7.95
141	Àcid 2-cloro-4-nitrobenzoic <sup>a,c</sup>	H <sub>2</sub> O	1.92	2.14	2.09
142	Àcid 2-cloro-4-nitrobenzoic <sup>a,c</sup>	MeOH	7.12	7.21	7.27
143	Àcid 2-cloro-5-nitrobenzoic <sup>a,c</sup>	H <sub>2</sub> O	2.17	2.24	2.12
144	Àcid 2-cloro-5-nitrobenzoic <sup>d</sup>	DMSO	7.95	7.92	7.94
145	Àcid 2-cloro-6-metilbenzoic <sup>b,c</sup>	H <sub>2</sub> O	2.75	2.98	2.91
146	Àcid 2-cloro-6-nitrobenzoic <sup>a,c</sup>	AN	17.53	17.67	17.51
147	Àcid 2-cloro-6-nitrobenzoic <sup>a,c</sup>	DMF	9.14	9.03	9.11
148	Àcid 2-cloro-6-nitrobenzoic <sup>a,c</sup>	DMSO	7.66	7.66	7.57
149	Àcid 2-cloro-6-nitrobenzoic <sup>a,c</sup>	MeOH	6.86	6.90	6.80
150	Àcid 2-cloro-6-nitrobenzoic <sup>d</sup>	Ac	15.12	15.27	15.08
151	Àcid 2-cloro-6-nitrobenzoic <sup>a,c</sup>	H <sub>2</sub> O	1.34	1.70	1.43
152	Àcid 2-clorobenzoic <sup>a,c</sup>	DMA	9.60	9.51	9.43

#	Àcid benzoic	Solvent	Exp.	BPNN	SVM
153	Àcid 2-clorobenzoic <sup>b,c</sup>	EtOH	9.08	9.21	9.14
154	Àcid 2-clorobenzoic <sup>a,c</sup>	H <sub>2</sub> O	2.92	3.06	2.98
155	Àcid 2-clorobenzoic <sup>a,c</sup>	MeOH	8.31	8.23	8.19
156	Àcid 2-clorobenzoic <sup>a,c</sup>	DMSO	9.16	9.51	9.43
157	Àcid 2-clorobenzoic <sup>d</sup>	AN	19.00	19.17	19.14
158	Àcid 2-cianobenzoic <sup>b,c</sup>	AN	19.07	19.02	18.98
159	Àcid 2-cianobenzoic <sup>a,c</sup>	MeOH	8.15	8.09	8.24
160	Àcid 2-cianobenzoic <sup>d</sup>	Ac	16.65	16.72	16.58
161	Àcid 2-cianobenzoic <sup>d</sup>	DMF	10.37	10.75	10.78
162	Àcid 2-cianobenzoic <sup>d</sup>	H <sub>2</sub> O	3.08	2.99	3.25
163	Àcid 2-fluorobenzoic <sup>a,c</sup>	EtOH	9.52	9.35	9.43
164	Àcid 2-fluorobenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.57	3.36	3.34
165	Àcid 2-fluorobenzoic <sup>a,c</sup>	MeOH	8.41	8.51	8.51
166	Àcid 2-formilbenzoic <sup>a,c</sup>	Ac	17.39	17.41	17.30
167	Àcid 2-formilbenzoic <sup>a,c</sup>	AN	20.38	20.48	20.29
168	Àcid 2-formilbenzoic <sup>a,c</sup>	DMF	11.21	11.11	11.30
169	Àcid 2-formilbenzoic <sup>a,c</sup>	MeOH	9.19	8.91	8.89
170	Àcid 2-formilbenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.55	4.05	3.72
171	Àcid 2-hidroxi-3,5-dinitrobenzoic <sup>a,c</sup>	MeOH	5.88	6.03	5.97
172	Àcid 2-hidroxi-5-bromobenzoic <sup>b,c</sup>	H <sub>2</sub> O	2.54	2.58	2.60
173	Àcid 2-hidroxi-5-bromobenzoic <sup>d</sup>	MeOH	7.76	7.70	7.70
174	Àcid 2-hidroxi-5-clorobenzoic <sup>a,c</sup>	H <sub>2</sub> O	2.59	2.64	2.68
175	Àcid 2-hidroxi-5-clorobenzoic <sup>b,c</sup>	MeOH	7.81	7.80	7.84
176	Àcid 2-hidroxi-5-fluorobenzoic <sup>a,c</sup>	MeOH	7.91	7.89	7.91
177	Àcid 2-hidroxi-5-fluorobenzoic <sup>d</sup>	H <sub>2</sub> O	2.70	2.62	2.72
178	Àcid 2-hidroxi-5-iodobenzoic <sup>d</sup>	H <sub>2</sub> O	2.62	2.64	2.56
179	Àcid 2-hidroxi-5-nitrobenzoic <sup>a,c</sup>	H <sub>2</sub> O	2.05	1.94	2.14
180	Àcid 2-hidroxi-5-nitrobenzoic <sup>a,c</sup>	MeOH	7.25	7.04	6.94
181	Àcid 2-hidroxibenzoic <sup>a,c</sup>	EtOH	8.45	8.50	8.36
182	Àcid 2-hidroxibenzoic <sup>a,c</sup>	H <sub>2</sub> O	2.77	2.88	2.77
183	Àcid 2-hidroxibenzoic <sup>a,c</sup>	MeOH	7.92	8.03	7.98
184	Àcid 2-iodobenzoic <sup>a,c</sup>	EtOH	9.09	8.99	9.00
185	Àcid 2-iodobenzoic <sup>a,c</sup>	H <sub>2</sub> O	2.96	2.86	3.15
186	Àcid 2-iodobenzoic <sup>d</sup>	MeOH	8.24	8.10	8.16
187	Àcid 2-iodobenzoic <sup>d</sup>	DMSO	9.90	9.54	9.80
188	Àcid 2-metoxi-6-clorobenzoic <sup>a,c</sup>	Ac	17.24	17.30	17.15
189	Àcid 2-metoxi-6-clorobenzoic <sup>a,c</sup>	AN	19.33	19.50	19.51
190	Àcid 2-metoxi-6-clorobenzoic <sup>a,c</sup>	DMF	11.24	11.35	11.22
191	Àcid 2-metoxi-6-clorobenzoic <sup>a,c</sup>	DMSO	9.76	9.96	9.85
192	Àcid 2-metoxi-6-clorobenzoic <sup>a,c</sup>	MeOH	8.06	8.11	8.02
193	Àcid 2-metoxi-6-nitrobenzoic <sup>a,c</sup>	Ac	16.65	16.72	16.56
194	Àcid 2-metoxi-6-nitrobenzoic <sup>a,c</sup>	AN	18.98	18.87	18.89
195	Àcid 2-metoxi-6-nitrobenzoic <sup>a,c</sup>	DMF	10.65	10.49	10.58
196	Àcid 2-metoxi-6-nitrobenzoic <sup>a,c</sup>	DMSO	9.16	8.85	9.07
197	Àcid 2-metoxi-6-nitrobenzoic <sup>a,c</sup>	MeOH	7.65	7.60	7.56
198	Àcid 2-metoxibenzoic <sup>a,c</sup>	MeOH	9.26	9.13	9.13
199	Àcid 2-metoxibenzoic <sup>d</sup>	EtOH	10.21	9.79	10.02
200	Àcid 2-metoxibenzoic <sup>d</sup>	H <sub>2</sub> O	4.09	3.86	3.98
201	Àcid 2-metil-6-clorobenzoic <sup>a,c</sup>	AN	19.17	19.11	19.16
202	Àcid 2-metil-6-clorobenzoic <sup>a,c</sup>	DMF	10.89	10.63	10.66
203	Àcid 2-metil-6-clorobenzoic <sup>b,c</sup>	DMSO	9.48	9.49	9.42



#	Àcid benzoic	Solvent	Exp.	BPNN	SVM
204	Àcid 2-metil-6-clorobenzoic <sup>a,c</sup>	H <sub>2</sub> O	2.75	2.71	2.71
205	Àcid 2-metil-6-clorobenzoic <sup>a,c</sup>	MeOH	7.91	7.99	8.00
206	Àcid 2-metil-6-clorobenzoic <sup>d</sup>	Ac	16.74	16.60	16.61
207	Àcid 2-metil-6-metoxibenzoic <sup>a,c</sup>	AN	20.52	20.65	20.54
208	Àcid 2-metil-6-metoxibenzoic <sup>a,c</sup>	DMF	12.34	12.20	12.25
209	Àcid 2-metil-6-metoxibenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.46	3.98	4.12
210	Àcid 2-metil-6-metoxibenzoic <sup>a,c</sup>	MeOH	8.70	9.21	9.34
211	Àcid 2-metil-6-metoxibenzoic <sup>d</sup>	Ac	18.24	18.24	18.10
212	Àcid 2-metil-6-metoxibenzoic <sup>d</sup>	DMSO	10.95	10.71	10.97
213	Àcid 2-metil-6-nitrobenzoic <sup>b,c</sup>	Ac	16.13	16.16	16.04
214	Àcid 2-metil-6-nitrobenzoic <sup>a,c</sup>	AN	18.63	18.64	18.55
215	Àcid 2-metil-6-nitrobenzoic <sup>a,c</sup>	DMSO	8.85	8.63	8.68
216	Àcid 2-metil-6-nitrobenzoic <sup>d</sup>	DMF	10.13	10.01	10.02
217	Àcid 2-metil-6-nitrobenzoic <sup>d</sup>	MeOH	7.56	7.59	7.68
218	Àcid 2-metil-6-nitrobenzoic <sup>a,c</sup>	H <sub>2</sub> O	2.21	2.50	2.42
219	Àcid 2-metilbenzoic <sup>a,c</sup>	DMSO	11.07	11.15	11.16
220	Àcid 2-metilbenzoic <sup>a,c</sup>	EtOH	10.23	10.18	10.21
221	Àcid 2-metilbenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.91	4.24	4.19
222	Àcid 2-metilbenzoic <sup>b,c</sup>	MeOH	9.31	9.40	9.40
223	Àcid 2-metilèsterbenzoic <sup>a,c</sup>	Ac	17.34	17.53	17.48
224	Àcid 2-metilèsterbenzoic <sup>a,c</sup>	DMF	11.33	11.60	11.54
225	Àcid 2-metilèsterbenzoic <sup>b,c</sup>	H <sub>2</sub> O	3.18	3.62	3.52
226	Àcid 2-metilèsterbenzoic <sup>a,c</sup>	MeOH	8.47	8.55	8.72
227	Àcid 2-metilèsterbenzoic <sup>d</sup>	AN	19.78	19.99	19.96
228	Àcid 2-nitrobenzoic <sup>a,c</sup>	Ac	15.98	15.64	15.48
229	Àcid 2-nitrobenzoic <sup>a,c</sup>	AN	18.54	18.21	18.02
230	Àcid 2-nitrobenzoic <sup>b,c</sup>	DMF	9.96	9.46	9.46
231	Àcid 2-nitrobenzoic <sup>a,c</sup>	H <sub>2</sub> O	2.19	2.38	2.28
232	Àcid 2-nitrobenzoic <sup>a,c</sup>	MeOH	7.59	7.44	7.45
233	Àcid 2-nitrobenzoic <sup>d</sup>	DMSO	8.18	8.09	8.08
234	Àcid 2-nitrobenzoic <sup>d</sup>	EtOH	8.27	8.23	8.28
235	Àcid 3,4,5-trihidroxibenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.21	4.00	4.01
236	Àcid 3,4,5-trimetoxibenzoic <sup>a,c</sup>	DMSO	10.30	10.37	10.39
237	Àcid 3,4,5-trimetilbenzoic <sup>a,c</sup>	DMSO	11.54	11.49	11.50
238	Àcid 3,4,5-trimetilbenzoic <sup>a,c</sup>	MeOH	9.66	9.57	9.58
239	Àcid 3,4-diclorobenzoic <sup>a,c</sup>	<sup>i</sup> PrOH	9.80	10.14	10.08
240	Àcid 3,4-diclorobenzoic <sup>a,c</sup>	AN	19.42	19.47	19.42
241	Àcid 3,4-diclorobenzoic <sup>a,c</sup>	DMF	10.79	10.84	10.77
242	Àcid 3,4-diclorobenzoic <sup>a,c</sup>	DMSO	9.60	9.51	9.51
243	Àcid 3,4-diclorobenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.64	3.56	3.51
244	Àcid 3,4-diclorobenzoic <sup>a,c</sup>	MeOH	8.64	8.60	8.64
245	Àcid 3,4-diclorobenzoic <sup>d</sup>	Ac	16.70	16.76	16.79
246	Àcid 3,4-dihidroxibenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.48	4.18	4.19
247	Àcid 3,4-dimetoxibenzoic <sup>a,c</sup>	Ac	18.59	18.61	18.40
248	Àcid 3,4-dimetoxibenzoic <sup>a,c</sup>	AN	20.95	20.95	20.77
249	Àcid 3,4-dimetoxibenzoic <sup>a,c</sup>	DMF	12.65	12.41	12.51
250	Àcid 3,4-dimetoxibenzoic <sup>a,c</sup>	DMSO	11.40	11.18	11.14
251	Àcid 3,4-dimetoxibenzoic <sup>a,c</sup>	MeOH	9.54	9.60	9.59
252	Àcid 3,4-dimetoxibenzoic <sup>d</sup>	H <sub>2</sub> O	4.44	4.43	4.46
253	Àcid 3,4-dimetilbenzoic <sup>a,c</sup>	<sup>i</sup> PrOH	11.60	11.11	11.14
254	Àcid 3,4-dimetilbenzoic <sup>a,c</sup>	Ac	18.71	18.61	18.46



#	Àcid benzoic	Solvent	Exp.	BPNN	SVM
255	Àcid 3,4-dimetilbenzoic <sup>a,c</sup>	AN	21.05	20.89	20.99
256	Àcid 3,4-dimetilbenzoic <sup>a,c</sup>	DMF	12.70	12.64	12.61
257	Àcid 3,4-dimetilbenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.40	4.37	4.39
258	Àcid 3,4-dimetilbenzoic <sup>a,c</sup>	MeOH	9.63	9.55	9.57
259	Àcid 3,4-dimetilbenzoic <sup>d</sup>	DMSO	11.46	11.33	11.34
260	Àcid 3,4-dinitrobenzoic <sup>a,c</sup>	Ac	15.07	15.00	15.04
261	Àcid 3,4-dinitrobenzoic <sup>a,c</sup>	DMF	9.07	9.09	9.16
262	Àcid 3,4-dinitrobenzoic <sup>a,c</sup>	DMSO	7.88	7.77	7.82
263	Àcid 3,4-dinitrobenzoic <sup>a,c</sup>	H <sub>2</sub> O	2.82	2.57	2.72
264	Àcid 3,4-dinitrobenzoic <sup>a,c</sup>	MeOH	7.44	7.64	7.56
265	Àcid 3,4-dinitrobenzoic <sup>d</sup>	AN	17.50	17.37	17.55
266	Àcid 3,5-dibromobenzoic <sup>a,c</sup>	Ac	16.23	16.47	16.50
267	Àcid 3,5-dibromobenzoic <sup>b,c</sup>	AN	18.86	19.22	19.13
268	Àcid 3,5-dibromobenzoic <sup>a,c</sup>	DMSO	8.88	9.21	9.23
269	Àcid 3,5-dibromobenzoic <sup>a,c</sup>	MeOH	8.29	8.47	8.45
270	Àcid 3,5-dibromobenzoic <sup>d</sup>	DMF	10.11	10.54	10.53
271	Àcid 3,5-diclorobenzoic <sup>a,c</sup>	DMSO	8.81	9.10	9.10
272	Àcid 3,5-diclorobenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.10	3.29	3.22
273	Àcid 3,5-diclorobenzoic <sup>a,c</sup>	MeOH	8.26	8.32	8.35
274	Àcid 3,5-diclorobenzoic <sup>a,c</sup>	AN	18.75	19.12	19.06
275	Àcid 3,5-diclorobenzoic <sup>a,c</sup>	DMF	10.43	10.39	10.37
276	Àcid 3,5-dihidroxibenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.04	3.98	3.95
277	Àcid 3,5-dimetoxibenzoic <sup>a,c</sup>	Ac	18.02	18.05	18.06
278	Àcid 3,5-dimetoxibenzoic <sup>a,c</sup>	AN	20.39	20.59	20.48
279	Àcid 3,5-dimetoxibenzoic <sup>a,c</sup>	DMF	12.04	11.92	12.13
280	Àcid 3,5-dimetoxibenzoic <sup>a,c</sup>	DMSO	10.71	10.73	10.77
281	Àcid 3,5-dimetoxibenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.97	4.17	4.19
282	Àcid 3,5-dimetoxibenzoic <sup>a,c</sup>	MeOH	9.27	9.29	9.34
283	Àcid 3,5-dimetilbenzoic <sup>a,c</sup>	Ac	18.54	18.60	18.46
284	Àcid 3,5-dimetilbenzoic <sup>a,c</sup>	DMF	12.57	12.62	12.61
285	Àcid 3,5-dimetilbenzoic <sup>b,c</sup>	DMSO	11.29	11.34	11.36
286	Àcid 3,5-dimetilbenzoic <sup>b,c</sup>	MeOH	9.59	9.46	9.50
287	Àcid 3,5-dimetilbenzoic <sup>d</sup>	AN	20.95	20.90	21.02
288	Àcid 3,5-dimetilbenzoic <sup>d</sup>	H <sub>2</sub> O	4.30	4.32	4.31
289	Àcid 3,5-dinitrobenzoic <sup>a,c</sup>	<sup>t</sup> PrOH	8.30	8.50	8.64
290	Àcid 3,5-dinitrobenzoic <sup>a,c</sup>	AN	17.36	17.42	17.52
291	Àcid 3,5-dinitrobenzoic <sup>a,c</sup>	DMF	8.76	9.10	9.13
292	Àcid 3,5-dinitrobenzoic <sup>a,c</sup>	DMSO	7.38	7.67	7.73
293	Àcid 3,5-dinitrobenzoic <sup>a,c</sup>	H <sub>2</sub> O	2.67	2.37	2.45
294	Àcid 3,5-dinitrobenzoic <sup>a,c</sup>	MeOH	7.64	7.39	7.38
295	Àcid 3,5-dinitrobenzoic <sup>d</sup>	Ac	14.87	15.14	15.04
296	Àcid 3,6-dicloro-2-metoxibenzoic <sup>b,c</sup>	H <sub>2</sub> O	1.97	2.09	2.02
297	Àcid 3-acetamidobenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.06	4.17	4.15
298	Àcid 3-acetamidobenzoic <sup>a,c</sup>	MeOH	9.25	9.28	9.29
299	Àcid 3-acetilbenzoic <sup>a,c</sup>	DMSO	10.22	10.50	10.53
300	Àcid 3-acetilbenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.83	4.07	4.10
301	Àcid 3-acetilbenzoic <sup>a,c</sup>	MeOH	8.87	9.25	9.20
302	Àcid 3-amino-2,5-diclorobenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.40	2.82	2.78
303	Àcid 3-aminobenzoic <sup>a,c</sup>	DMSO	11.60	11.08	11.08
304	Àcid 3-aminobenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.74	4.25	4.21
305	Àcid 3-bromo-4-metoxibenzoic <sup>a,c</sup>	Ac	17.82	17.93	17.91

#	Àcid benzoic	Solvent	Exp.	BPNN	SVM
306	Àcid 3-bromo-4-metoxibenzoic <sup>b,c</sup>	DMF	11.83	11.98	11.95
307	Àcid 3-bromo-4-metoxibenzoic <sup>a,c</sup>	DMSO	10.61	10.63	10.68
308	Àcid 3-bromo-4-metoxibenzoic <sup>a,c</sup>	MeOH	9.27	9.34	9.39
309	Àcid 3-bromo-4-metoxibenzoic <sup>d</sup>	AN	20.38	20.42	20.45
310	Àcid 3-bromo-4-metilbenzoic <sup>b,c</sup>	Ac	17.59	17.51	17.50
311	Àcid 3-bromo-4-metilbenzoic <sup>a,c</sup>	AN	20.03	20.11	20.12
312	Àcid 3-bromo-4-metilbenzoic <sup>a,c</sup>	DMF	11.49	11.62	11.56
313	Àcid 3-bromo-4-metilbenzoic <sup>b,c</sup>	DMSO	10.30	10.29	10.30
314	Àcid 3-bromo-4-metilbenzoic <sup>d</sup>	MeOH	9.06	9.10	9.10
315	Àcid 3-bromo-4-metilbenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.29	3.64	3.97
316	Àcid 3-bromobenzoic <sup>a,c</sup>	<sup>i</sup> PrOH	10.11	10.26	10.14
317	Àcid 3-bromobenzoic <sup>a,c</sup>	DMSO	9.68	9.84	9.88
318	Àcid 3-bromobenzoic <sup>a,c</sup>	EtOH	9.40	9.65	9.58
319	Àcid 3-bromobenzoic <sup>b,c</sup>	H <sub>2</sub> O	3.81	3.78	3.74
320	Àcid 3-bromobenzoic <sup>a,c</sup>	MeOH	8.80	8.89	8.84
321	Àcid 3-bromobenzoic <sup>a,c</sup>	AN	19.50	19.76	19.74
322	Àcid 3-bromobenzoic <sup>a,c</sup>	DMF	11.20	11.16	11.15
323	Àcid 3-cloro-4-nitrobenzoic <sup>a,c</sup>	Ac	15.78	15.98	15.95
324	Àcid 3-cloro-4-nitrobenzoic <sup>a,c</sup>	DMF	10.07	9.99	10.01
325	Àcid 3-cloro-4-nitrobenzoic <sup>a,c</sup>	DMSO	8.61	8.65	8.71
326	Àcid 3-cloro-4-nitrobenzoic <sup>a,c</sup>	MeOH	8.09	7.91	7.96
327	Àcid 3-cloro-4-nitrobenzoic <sup>d</sup>	AN	18.28	18.52	18.49
328	Àcid 3-cloro-5-nitrobenzoic <sup>a,c</sup>	Ac	15.57	15.77	15.67
329	Àcid 3-cloro-5-nitrobenzoic <sup>a,c</sup>	AN	17.98	18.30	18.21
330	Àcid 3-cloro-5-nitrobenzoic <sup>a,c</sup>	DMF	9.60	9.68	9.69
331	Àcid 3-cloro-5-nitrobenzoic <sup>a,c</sup>	DMSO	8.37	8.30	8.35
332	Àcid 3-cloro-5-nitrobenzoic <sup>a,c</sup>	MeOH	8.03	7.67	7.72
333	Àcid 3-cloro-5-nitrobenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.13	2.66	2.67
334	Àcid 3-clorobenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.80	3.90	3.86
335	Àcid 3-clorobenzoic <sup>d</sup>	DMA	9.80	10.05	10.03
336	Àcid 3-clorobenzoic <sup>d</sup>	DMSO	9.51	10.05	10.03
337	Àcid 3-clorobenzoic <sup>d</sup>	MeOH	8.83	8.99	8.98
338	Àcid 3-clorobenzoic <sup>d</sup>	DMF	10.95	11.39	11.31
339	Àcid 3-cianobenzoic <sup>a,c</sup>	DMSO	9.44	9.69	9.81
340	Àcid 3-cianobenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.60	3.64	3.69
341	Àcid 3-cianobenzoic <sup>a,c</sup>	MeOH	8.53	8.75	8.74
342	Àcid 3-fluorobenzoic <sup>a,c</sup>	MeOH	8.87	8.74	8.79
343	Àcid 3-fluorobenzoic <sup>d</sup>	H <sub>2</sub> O	3.88	3.71	3.66
344	Àcid 3-formilbenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.84	3.97	3.84
345	Àcid 3-hidroxibenzoic <sup>a,c</sup>	DMA	10.75	10.45	10.47
346	Àcid 3-hidroxibenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.03	3.97	3.93
347	Àcid 3-hidroxibenzoic <sup>d</sup>	MeOH	9.58	9.04	9.08
348	Àcid 3-iodobenzoic <sup>a,c</sup>	DMSO	10.16	9.96	10.13
349	Àcid 3-iodobenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.82	3.86	3.91
350	Àcid 3-iodobenzoic <sup>a,c</sup>	MeOH	8.89	9.02	8.96
351	Àcid 3-metoxi-4-clorobenzoic <sup>a,c</sup>	Ac	17.39	17.32	17.46
352	Àcid 3-metoxi-4-clorobenzoic <sup>a,c</sup>	AN	20.02	19.97	20.03
353	Àcid 3-metoxi-4-clorobenzoic <sup>a,c</sup>	DMF	11.58	11.42	11.49
354	Àcid 3-metoxi-4-clorobenzoic <sup>d</sup>	DMSO	10.39	10.22	10.26
355	Àcid 3-metoxi-4-clorobenzoic <sup>d</sup>	MeOH	8.98	8.88	8.94
356	Àcid 3-metoxi-4-metilbenzoic <sup>a,c</sup>	AN	20.76	20.69	20.74

#	Àcid benzoic	Solvent	Exp.	BPNN	SVM
357	Àcid 3-metoxi-4-metilbenzoic <sup>a,c</sup>	DMF	12.39	12.30	12.33
358	Àcid 3-metoxi-4-metilbenzoic <sup>a,c</sup>	DMSO	11.06	11.10	11.09
359	Àcid 3-metoxi-4-metilbenzoic <sup>b,c</sup>	MeOH	9.49	9.34	9.38
360	Àcid 3-metoxi-4-metilbenzoic <sup>d</sup>	Ac	18.32	18.27	18.23
361	Àcid 3-metoxi-4-metilbenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.13	4.23	4.22
362	Àcid 3-metoxi-4-nitrobenzoic <sup>a,c</sup>	Ac	16.21	16.57	16.84
363	Àcid 3-metoxi-4-nitrobenzoic <sup>a,c</sup>	AN	18.77	19.35	19.39
364	Àcid 3-metoxi-4-nitrobenzoic <sup>a,c</sup>	DMF	10.33	10.80	10.87
365	Àcid 3-metoxi-4-nitrobenzoic <sup>a,c</sup>	DMSO	9.22	9.58	9.60
366	Àcid 3-metoxi-4-nitrobenzoic <sup>a,c</sup>	MeOH	9.28	9.33	9.06
367	Àcid 3-metoxi-5-bromobenzoic <sup>a,c</sup>	Ac	17.03	17.25	17.30
368	Àcid 3-metoxi-5-bromobenzoic <sup>a,c</sup>	DMF	11.28	11.32	11.31
369	Àcid 3-metoxi-5-bromobenzoic <sup>a,c</sup>	DMSO	9.79	10.03	10.05
370	Àcid 3-metoxi-5-bromobenzoic <sup>a,c</sup>	MeOH	8.79	8.80	8.89
371	Àcid 3-metoxi-5-bromobenzoic <sup>d</sup>	AN	19.73	19.85	19.86
372	Àcid 3-metoxi-5-nitrobenzoic <sup>a,c</sup>	Ac	16.31	16.03	16.22
373	Àcid 3-metoxi-5-nitrobenzoic <sup>b,c</sup>	AN	18.72	18.71	18.77
374	Àcid 3-metoxi-5-nitrobenzoic <sup>a,c</sup>	MeOH	8.40	8.57	8.45
375	Àcid 3-metoxi-5-nitrobenzoic <sup>d</sup>	DMF	10.37	10.17	10.20
376	Àcid 3-metoxi-5-nitrobenzoic <sup>d</sup>	DMSO	9.01	8.90	8.91
377	Àcid 3-metoxibenzoic <sup>a,c</sup>	DMSO	11.10	10.72	10.74
378	Àcid 3-metoxibenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.12	4.06	4.03
379	Àcid 3-metoxibenzoic <sup>a,c</sup>	MeOH	9.30	9.16	9.19
380	Àcid 3-metil-4-clorobenzoic <sup>a,c</sup>	AN	20.27	20.14	20.18
381	Àcid 3-metil-4-clorobenzoic <sup>b,c</sup>	DMF	11.70	11.67	11.68
382	Àcid 3-metil-4-clorobenzoic <sup>a,c</sup>	DMSO	10.54	10.39	10.43
383	Àcid 3-metil-4-clorobenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.07	3.95	3.96
384	Àcid 3-metil-4-clorobenzoic <sup>d</sup>	Ac	17.59	17.55	17.58
385	Àcid 3-metil-4-clorobenzoic <sup>d</sup>	MeOH	9.12	9.07	9.07
386	Àcid 3-metil-4-metoxibenzoic <sup>a,c</sup>	Ac	19.01	18.90	18.73
387	Àcid 3-metil-4-metoxibenzoic <sup>a,c</sup>	AN	21.29	21.09	21.20
388	Àcid 3-metil-4-metoxibenzoic <sup>a,c</sup>	DMF	12.84	12.87	12.89
389	Àcid 3-metil-4-metoxibenzoic <sup>a,c</sup>	DMSO	11.71	11.59	11.62
390	Àcid 3-metil-4-metoxibenzoic <sup>a,c</sup>	MeOH	9.81	9.69	9.70
391	Àcid 3-metil-4-metoxibenzoic <sup>d</sup>	H <sub>2</sub> O	4.35	4.47	4.50
392	Àcid 3-metil-4-nitrobenzoic <sup>a,c</sup>	Ac	16.55	16.38	16.44
393	Àcid 3-metil-4-nitrobenzoic <sup>a,c</sup>	AN	19.21	18.97	18.99
394	Àcid 3-metil-4-nitrobenzoic <sup>a,c</sup>	DMF	11.03	10.55	10.56
395	Àcid 3-metil-4-nitrobenzoic <sup>a,c</sup>	DMSO	9.48	9.29	9.29
396	Àcid 3-metil-4-nitrobenzoic <sup>b,c</sup>	H <sub>2</sub> O	3.65	3.31	3.42
397	Àcid 3-metil-4-nitrobenzoic <sup>a,c</sup>	MeOH	8.54	8.40	8.38
398	Àcid 3-metil-5-clorobenzoic <sup>a,c</sup>	Ac	17.40	17.40	17.40
399	Àcid 3-metil-5-clorobenzoic <sup>a,c</sup>	DMF	11.35	11.50	11.44
400	Àcid 3-metil-5-clorobenzoic <sup>a,c</sup>	DMSO	10.04	10.18	10.17
401	Àcid 3-metil-5-clorobenzoic <sup>a,c</sup>	MeOH	8.94	8.95	8.99
402	Àcid 3-metil-5-clorobenzoic <sup>d</sup>	AN	19.85	19.99	20.01
403	Àcid 3-metil-5-metoxibenzoic <sup>a,c</sup>	Ac	18.26	18.26	18.26
404	Àcid 3-metil-5-metoxibenzoic <sup>a,c</sup>	AN	20.66	20.71	20.78
405	Àcid 3-metil-5-metoxibenzoic <sup>a,c</sup>	DMSO	10.93	11.08	11.12
406	Àcid 3-metil-5-metoxibenzoic <sup>a,c</sup>	MeOH	9.44	9.27	9.35
407	Àcid 3-metil-5-metoxibenzoic <sup>d</sup>	DMF	12.32	12.26	12.35

#	Àcid benzoic	Solvent	Exp.	BPNN	SVM
408	Àcid 3-metil-5-nitrobenzoic <sup>a,c</sup>	AN	19.18	18.98	18.97
409	Àcid 3-metil-5-nitrobenzoic <sup>a,c</sup>	DMF	10.92	10.51	10.49
410	Àcid 3-metil-5-nitrobenzoic <sup>b,c</sup>	DMSO	9.36	9.22	9.21
411	Àcid 3-metil-5-nitrobenzoic <sup>a,c</sup>	MeOH	8.55	8.43	8.41
412	Àcid 3-metil-5-nitrobenzoic <sup>d</sup>	Ac	16.62	16.37	16.42
413	Àcid 3-metilbenzoic <sup>a,c</sup>	DMSO	11.26	10.98	11.01
414	Àcid 3-metilbenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.24	4.21	4.21
415	Àcid 3-metilbenzoic <sup>a,c</sup>	MeOH	9.47	9.36	9.37
416	Àcid 3-nitro-4-clorobenzoic <sup>b,c</sup>	Ac	15.98	15.99	16.01
417	Àcid 3-nitro-4-clorobenzoic <sup>a,c</sup>	AN	18.57	18.67	18.60
418	Àcid 3-nitro-4-clorobenzoic <sup>a,c</sup>	DMF	9.98	10.03	10.02
419	Àcid 3-nitro-4-clorobenzoic <sup>a,c</sup>	MeOH	8.24	8.16	8.17
420	Àcid 3-nitro-4-clorobenzoic <sup>d</sup>	iPrOH	9.30	9.52	9.54
421	Àcid 3-nitro-4-clorobenzoic <sup>d</sup>	DMSO	8.74	8.68	8.73
422	Àcid 3-nitro-4-clorobenzoic <sup>d</sup>	H <sub>2</sub> O	3.29	3.14	3.17
423	Àcid 3-nitro-4-metoxibenzoic <sup>a,c</sup>	Ac	17.04	17.01	17.14
424	Àcid 3-nitro-4-metoxibenzoic <sup>a,c</sup>	AN	19.51	19.70	19.66
425	Àcid 3-nitro-4-metoxibenzoic <sup>a,c</sup>	DMF	11.18	11.12	11.15
426	Àcid 3-nitro-4-metoxibenzoic <sup>a,c</sup>	DMSO	9.85	9.79	9.86
427	Àcid 3-nitro-4-metoxibenzoic <sup>d</sup>	MeOH	8.83	9.22	9.11
428	Àcid 3-nitro-4-metoxibenzoic <sup>d</sup>	H <sub>2</sub> O	3.72	4.07	4.07
429	Àcid 3-nitro-4-metilbenzoic <sup>a,c</sup>	Ac	16.75	16.59	16.66
430	Àcid 3-nitro-4-metilbenzoic <sup>a,c</sup>	DMF	10.93	10.73	10.70
431	Àcid 3-nitro-4-metilbenzoic <sup>a,c</sup>	DMSO	9.45	9.41	9.44
432	Àcid 3-nitro-4-metilbenzoic <sup>b,c</sup>	MeOH	8.61	8.74	8.69
433	Àcid 3-nitro-4-metilbenzoic <sup>d</sup>	AN	19.36	19.28	19.25
434	Àcid 3-nitro-4-metilbenzoic <sup>d</sup>	H <sub>2</sub> O	3.62	3.64	3.69
435	Àcid 3-nitrobenzoic <sup>a,c</sup>	DMSO	9.17	9.12	9.18
436	Àcid 3-nitrobenzoic <sup>a,c</sup>	MeOH	8.32	8.50	8.49
437	Àcid 3-nitrobenzoic <sup>d</sup>	H <sub>2</sub> O	3.46	3.43	3.50
438	Àcid 3-nitrobenzoic <sup>a,c</sup>	AN	19.29	19.06	19.00
439	Àcid 3-trifluorometilbenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.75	4.07	3.73
440	Àcid 3-trifluorometilbenzoic <sup>a,c</sup>	MeOH	8.69	8.94	8.78
441	Àcid 4-(dimetilamino)benzoic <sup>a,c</sup>	H <sub>2</sub> O	5.03	4.86	4.94
442	Àcid 4-( <i>N,N</i> -dimetilaminocarbonil)benzoic <sup>a,c</sup>	Ac	17.83	17.70	17.68
443	Àcid 4-( <i>N,N</i> -dimetilaminocarbonil)benzoic <sup>a,c</sup>	AN	20.55	20.21	20.20
444	Àcid 4-( <i>N,N</i> -dimetilaminocarbonil)benzoic <sup>d</sup>	MeOH	9.00	9.24	9.14
445	Àcid 4-acetamidobenzoic <sup>a,c</sup>	MeOH	9.57	9.78	9.72
446	Àcid 4-acetamidobenzoic <sup>d</sup>	H <sub>2</sub> O	4.30	4.52	4.57
447	Àcid 4-acetilbenzoic <sup>a,c</sup>	AN	19.95	20.07	20.10
448	Àcid 4-acetilbenzoic <sup>a,c</sup>	DMF	11.38	11.65	11.68
449	Àcid 4-acetilbenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.74	3.89	3.94
450	Àcid 4-acetilbenzoic <sup>a,c</sup>	MeOH	8.84	9.02	9.01
451	Àcid 4-acetilbenzoic <sup>d</sup>	Ac	17.32	17.49	17.53
452	Àcid 4-aminobenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.82	4.56	4.58
453	Àcid 4-aminobenzoic <sup>a,c</sup>	MeOH	10.25	9.84	9.80
454	Àcid 4-bromobenzoic <sup>a,c</sup>	DMSO	10.20	9.79	9.91
455	Àcid 4-bromobenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.99	3.72	3.72
456	Àcid 4-bromobenzoic <sup>d</sup>	MeOH	8.93	8.84	8.80
457	Àcid 4-bromobenzoic <sup>a,c</sup>	AN	20.30	19.75	19.72
458	Àcid 4-clorobenzoic <sup>a,c</sup>	DMA	10.30	10.21	10.24



#	Àcid benzoic	Solvent	Exp.	BPNN	SVM
459	Àcid 4-clorobenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.00	4.13	4.18
460	Àcid 4-clorobenzoic <sup>a,c</sup>	MeOH	9.09	9.26	9.28
461	Àcid 4-clorobenzoic <sup>d</sup>	DMSO	10.10	10.58	10.61
462	Àcid 4-cianobenzoic <sup>a,c</sup>	Ac	16.69	16.95	17.03
463	Àcid 4-cianobenzoic <sup>a,c</sup>	AN	19.34	19.62	19.61
464	Àcid 4-cianobenzoic <sup>a,c</sup>	DMF	11.02	11.05	11.19
465	Àcid 4-cianobenzoic <sup>a,c</sup>	DMSO	9.27	9.73	9.92
466	Àcid 4-cianobenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.53	3.56	3.68
467	Àcid 4-cianobenzoic <sup>a,c</sup>	MeOH	8.49	8.66	8.70
468	Àcid 4-etoxibenzoic <sup>b,c</sup>	DMSO	11.50	11.52	11.55
469	Àcid 4-etoxibenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.80	4.44	4.47
470	Àcid 4-etilbenzoic <sup>d</sup>	H <sub>2</sub> O	4.35	4.27	4.28
471	Àcid 4-fluorobenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.15	3.99	3.98
472	Àcid 4-fluorobenzoic <sup>d</sup>	MeOH	9.23	9.13	9.10
473	Àcid 4-formilbenzoic <sup>a,c</sup>	Ac	17.10	17.38	17.44
474	Àcid 4-formilbenzoic <sup>a,c</sup>	AN	19.74	20.04	20.08
475	Àcid 4-formilbenzoic <sup>a,c</sup>	DMF	11.19	11.49	11.59
476	Àcid 4-formilbenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.69	3.68	3.60
477	Àcid 4-formilbenzoic <sup>a,c</sup>	MeOH	8.94	8.68	8.72
478	Àcid 4-hidroxi-3,5-dimetoxibenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.34	4.38	4.42
479	Àcid 4-hidroxi-3-metoxibenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.51	4.31	4.33
480	Àcid 4-hidroxiibenzoic <sup>a,c</sup>	AN	20.80	20.77	20.84
481	Àcid 4-hidroxiibenzoic <sup>a,c</sup>	DMA	11.25	11.15	11.13
482	Àcid 4-hidroxiibenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.55	4.38	4.40
483	Àcid 4-hidroxiibenzoic <sup>a,c</sup>	MeOH	9.99	9.59	9.57
484	Àcid 4-iodo-3-nitrobenzoic <sup>a,c</sup>	DMSO	8.65	8.43	8.53
485	Àcid 4-iodobenzoic <sup>a,c</sup>	MeOH	9.04	8.84	8.79
486	Àcid 4-iodobenzoic <sup>d</sup>	H <sub>2</sub> O	3.98	3.70	3.75
487	Àcid 4-metoxibenzoic <sup>a,c</sup>	DMSO	11.80	11.37	11.39
488	Àcid 4-metoxibenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.25	4.40	4.42
489	Àcid 4-metoxibenzoic <sup>a,c</sup>	MeOH	9.79	9.62	9.61
490	Àcid 4-metilbenzoic <sup>b,c</sup>	H <sub>2</sub> O	4.38	4.37	4.39
491	Àcid 4-metilbenzoic <sup>d</sup>	DMSO	11.42	11.20	11.22
492	Àcid 4-metilbenzoic <sup>d</sup>	MeOH	9.61	9.55	9.55
493	Àcid 4-metilèsterbenzoic <sup>a,c</sup>	Ac	17.37	17.23	17.28
494	Àcid 4-metilèsterbenzoic <sup>a,c</sup>	AN	19.91	19.82	19.82
495	Àcid 4-metilèsterbenzoic <sup>a,c</sup>	H <sub>2</sub> O	3.74	3.90	3.91
496	Àcid 4-metilèsterbenzoic <sup>a,c</sup>	MeOH	8.84	9.07	8.93
497	Àcid 4-metilèsterbenzoic <sup>d</sup>	DMF	11.45	11.50	11.41
498	Àcid 4-nitrobenzoic <sup>a,c</sup>	<sup>t</sup> PrOH	9.60	9.60	9.56
499	Àcid 4-nitrobenzoic <sup>a,c</sup>	Ac	16.42	16.16	16.23
500	Àcid 4-nitrobenzoic <sup>a,c</sup>	AN	19.08	18.78	18.78
501	Àcid 4-nitrobenzoic <sup>b,c</sup>	DMF	10.80	10.28	10.36
502	Àcid 4-nitrobenzoic <sup>a,c</sup>	DMSO	9.04	8.98	9.09
503	Àcid 4-nitrobenzoic <sup>a,c</sup>	EtOH	8.90	8.99	8.99
504	Àcid 4-nitrobenzoic <sup>b,c</sup>	H <sub>2</sub> O	3.43	3.13	3.27
505	Àcid 4-nitrobenzoic <sup>d</sup>	MeOH	8.36	8.20	8.22
506	Àcid 4-propoxibenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.46	4.59	4.65
507	Àcid 4- <i>t</i> -butilbenzoic <sup>a,c</sup>	H <sub>2</sub> O	4.36	3.93	3.89
508	Àcid 4- <i>t</i> -butilbenzoic <sup>a,c</sup>	MeOH	9.61	9.11	9.03
509	Àcid acetilsalicílic <sup>a,c</sup>	DMF	11.30	11.26	11.39



#	Àcid benzoic	Solvent	Exp.	BPNN	SVM
510	Àcid acetilsalicílic <sup>a,c</sup>	H <sub>2</sub> O	3.69	3.73	3.78
511	Àcid benzoic <sup>a,c</sup>	Ac	18.20	18.00	17.91
512	Àcid benzoic <sup>b,c</sup>	AN	20.70	20.49	20.51
513	Àcid benzoic <sup>a,c</sup>	DMA	11.00	10.74	10.82
514	Àcid benzoic <sup>a,c</sup>	DMSO	11.00	10.73	10.82
515	Àcid benzoic <sup>a,c</sup>	EtOH	10.25	9.99	9.98
516	Àcid benzoic <sup>a,c</sup>	H <sub>2</sub> O	4.21	4.07	4.09
517	Àcid benzoic <sup>b,c</sup>	MeOH	9.41	9.25	9.22
518	Àcid benzoic <sup>d</sup>	<sup>t</sup> PrOH	10.71	10.62	10.57
519	Àcid benzoic <sup>d</sup>	DMF	12.27	12.05	12.09

<sup>a</sup> Training set de BPNN. <sup>b</sup> Cross-validation set de BPNN.

<sup>c</sup> Training set de SVM. <sup>d</sup> Prediction set de BPNN i SVM.

### 4.2.3. Anàlisi conjunta de fenols i àcids benzoics

Degut als bons resultats obtinguts amb les famílies dels fenols i els àcids benzoics es va pensar en fer una anàlisi conjunta d'aquestes dues famílies. La intenció d'aquest punt és complicar, encara una mica més, el sistema multicomponent d'estudi per veure la resposta que proporciona la metodologia QSPR en aquestes condicions. Un sistema que conté alhora fenols i àcids benzoics és, per suposat, més complex que qualsevol dels dos sistemes individuals però, tot i així, les semblances estructurals entre els compostos de les dues famílies són força elevades: totes dues tenen l'anell benzènic, els substituents i els solvents en els quals es disposa de valors experimentals són gairebé els mateixos, el rang de  $pK_a$  en el qual s'estenen els seus valors és el mateix, etc. A més, els descriptors que s'obtenen en els dos models són del mateix tipus i, en algun cas, fins i tot coincideixen: càrrega parcial màxima sobre un àtom d'hidrogen dels soluts i  $\alpha$  de Kamlet i Taft dels solvents.

#### *Conjunt de dades*

En aquesta ocasió, el conjunt de dades escollit és la suma dels 276 valors de  $pK_a$  dels fenols i dels 519 valors dels àcids benzoics. Cal dir, però, que s'han eliminat dos compostos de la família dels fenols ja que presenten substituents que no coincideixen, i que a més són molt diferents, amb els que es troben en els àcids benzoics: el 3-(metilsulfonyl)fenol en DMSO i el 3-(trifluorometilsulfonyl)fenol en nitrometà. Així, el conjunt de dades definitiu conté 793 valors experimentals de  $pK_a$ , 519 d'àcids benzoics i 274 de fenols. Els solvents inclosos en aquest estudi són onze: H<sub>2</sub>O, MeOH, EtOH, <sup>i</sup>PrOH, <sup>t</sup>BuOH, Ac, AN, DMSO, DMF, DMA i NM. Val a dir que en alguns d'aquests solvents només es tenen valors experimentals per a una de les famílies, en etanol només es tenen mesures d'àcids benzoics i en nitrometà i *tert*-butanol només n'hi ha de fenols. De nou, el nombre de valors experimentals en cada solvent és molt diferent, així, es troben 188 en aigua, 146 en metanol, 23 en etanol, 34 en isopropanol, 16 en *tert*-butanol, 65 en acetona, 88 en acetonitril, 127 en DMSO, 87 en DMF, 10 en DMA i 9 en nitrometà.

La divisió del conjunt de dades en *tset* i *pset* proporciona subconjunts de 640 i 153 valors experimentals, respectivament. En aquesta ocasió no s'emprarà el 10% dels

compostos del *tset* per crear el *cvset* ja que, només es construirà el model no lineal amb *Support Vector Machines*, per raons que s'explicaran més endavant. Els subconjunts generats, *tset* i *pset*, coincideixen completament amb els que s'han utilitzat anteriorment en l'estudi individual de cada família de compostos. La Taula 4.32 conté els valors experimentals i calculats per a tots els compostos estudiats.

### ***Model no lineal: Support Vector Machines***

El gran nombre de valors experimentals ha fet impossible treballar amb el programa ADAPT ja que aquest només permet treballar amb conjunts de dades no superiors a 500 valors. Donat que el *tset* d'aquest estudi conté 640 valors experimentals, ha estat impossible la tria de descriptors amb algorismes genètics i l'establiment dels models no lineals amb BPNN. D'altra banda, el mètode SVM permet treballar amb conjunts de molècules molt grans, sent la seva limitació més important la potència de l'ordinador amb el qual es vulguin fer aquest tipus de càlculs.

Donat que no ha estat possible fer una tria de descriptors específica per a aquest conjunt de dades utilitzant els algorismes genètics implementats per ADAPT, es va decidir que s'emprarien els descriptors obtinguts pel model dels àcids benzoics per al nou conjunt de compostos. Aquesta elecció sembla a priori encertada, ja que els descriptors d'aquest model són molt semblants als que s'obtenen en el model dels fenols. A més, els descriptors que conté aquest model semblen adients per descriure també la dissociació dels fenols, la seva situació en la molècula coincideix amb la dels àcids benzoics, tal i com es mostra en la Figura 4.23.

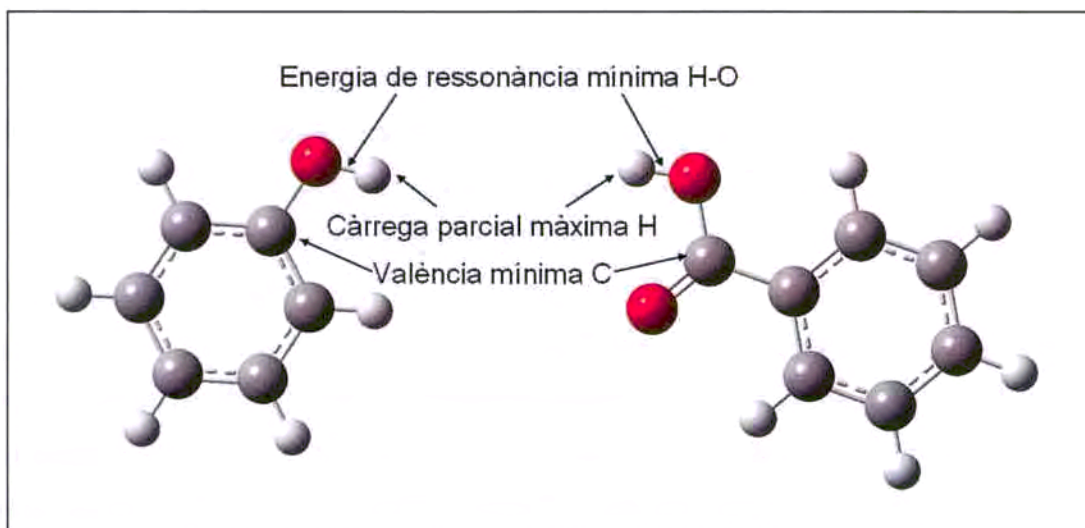
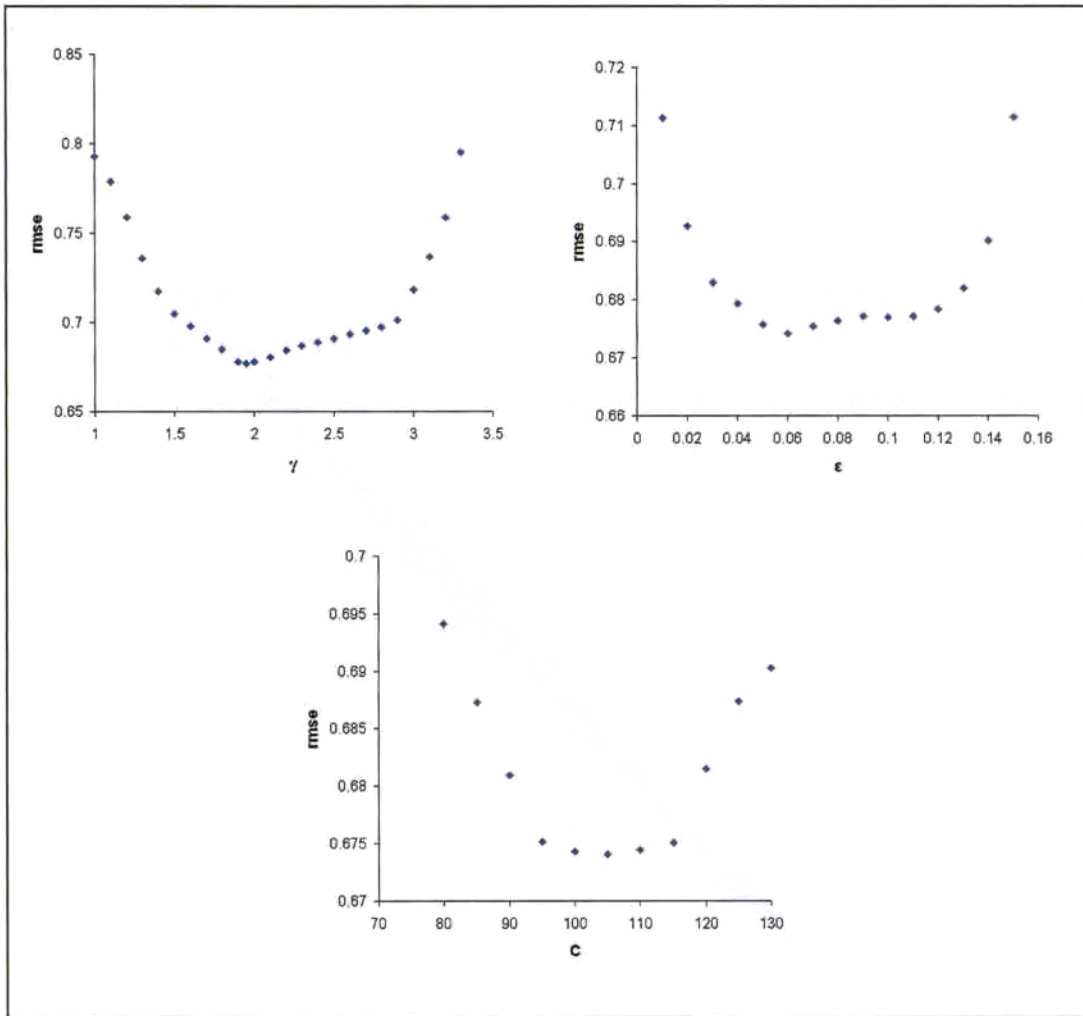


Figura 4.23. Localització dels descriptors del model sobre fenols i àcids benzoics.

Així doncs, els descriptors triats per establir el model no lineal per a les famílies dels fenols i dels àcids benzoics són: la càrrega parcial màxima sobre un àtom d'hidrogen, l'àrea superficial positiva fraccionada (FPSA-2), l'energia de ressonància mínima d'un enllaç H-O, la valència mínima d'un àtom de carboni, la repulsió electró-electró màxima d'un àtom d'oxigen dels soluts; i  $\alpha$  de Kamlet i Taft i l'energia de vaporització molar ( $\Delta_{\text{vap}}U^\circ$ ) dels solvents.

Com s'ha pogut comprovar amb els models obtinguts amb les famílies dels fenols i dels àcids benzoics, els resultats que s'obtenen utilitzant els mètodes BPNN i SVM són pràcticament idèntics. Així doncs, s'espera que en desenvolupar el model conjunt amb fenols i àcids benzoics amb SVM, els resultats que s'obtinguin no haurien de ser molt diferents a aquells que s'obtidrien amb BPNN. Per tal de construir el model SVM cal optimitzar els seus paràmetres característics:  $\gamma$ ,  $\epsilon$  i  $C$ , amb el mètode de minimització de l'error *leave-one-out* per als compostos del *tset*, Figura 4.24. El valor òptim pel paràmetre  $\gamma$  (amb  $\epsilon=0.1$  i  $C=100$ ) és de 1.95, en aquest punt el *rmse* del model és mínim. El valor del paràmetre  $\epsilon$  que condueix a un *rmse* mínim (amb  $\gamma=1.95$  i  $C=100$ ) és de 0.06. Per últim, s'optimitza el paràmetre  $C$  (amb  $\gamma=1.95$  i  $\epsilon=0.06$ ) i es troba que el valor òptim per aquest paràmetre és 105.


 Figura 4.24. Optimització dels paràmetres  $\gamma$ ,  $\epsilon$  i  $C$ .

Un cop optimitzats els tres paràmetres característics, es construeix el model no lineal definitiu. Aquest model consta de 562 vectors suport, triats d'entre els 640 compostos de què consta el *tset*. Els resultats que s'obtenen amb aquest model són molt bons, la Taula 4.33 conté els valors del coeficient de determinació  $R^2$ , el test- $F$  i el *rmse* per als dos subconjunts de dades: *tset* i *pset*.

Taula 4.33. Paràmetres estadístics obtinguts amb el model SVM.

Conjunt	n	$R^2$	F	rmse
<i>tset</i>	640	0.993	92678.9	0.44
<i>pset</i>	153	0.988	12445.0	0.59



La representació gràfica dels valors de  $pK_a$  calculat enfront dels experimentals es troba en la Figura 4.25. Els valors de  $pK_a$  calculat i experimental es troben a la Taula 4.32.

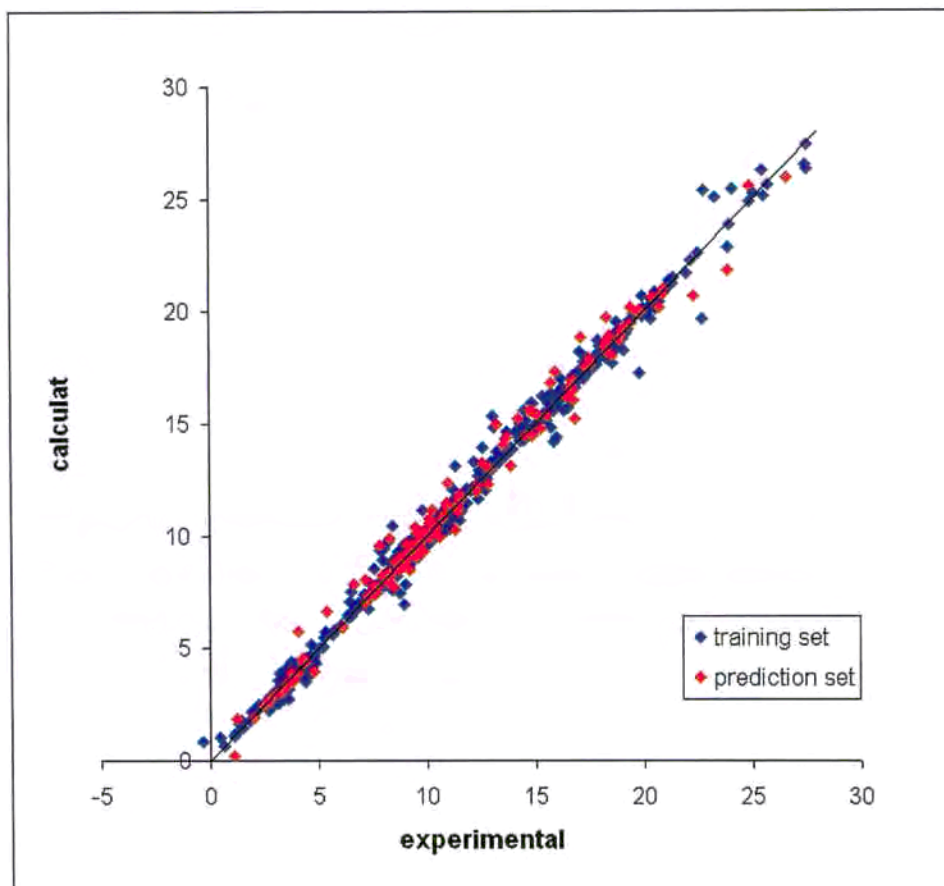


Figura 4.25. Representació gràfica de  $pK_a$  calculat vs. experimental amb SVM.

La interpretació fisicoquímica que es pot fer dels descriptors inclosos en el model és exactament la mateixa que s'ha fet per la família dels àcids benzoics. Els descriptors càrrega parcial màxima sobre un àtom d'hidrogen, energia de ressonància mínima de l'enllaç H-O i valència mínima d'un àtom de carboni fan referència a la polarització de l'enllaç que es trenca en la reacció de dissociació i recullen, alhora, els efectes dels substituents, tan estèrics com electrònics, i de les posicions de substitució sobre l'anell. El descriptor electrostàtic àrea superficial positiva fraccionada, FPSA-2, conté la informació relacionada amb les interaccions electrostàtiques i polars que es donen entre les molècules de solut i solvent. El descriptor de solvent energia de

vaporització molar ( $\Delta_{\text{vap}}U^\circ$ ) mesura l'energia involucrada en la formació de cavitats en el procés de dissolució de les espècies implicades en la reacció de dissociació. Els últims dos descriptors: repulsió electró-electró màxima d'un àtom d'oxigen i  $\alpha$  de Kamlet i Taft, fan referència a les interaccions donador/acceptor d'enllaços d'hidrogen que s'estableixen entre soluts i solvents.

Per comprovar la robustesa del model obtingut s'analitzen per separat tots els subconjunts de compostos que es poden formar a partir del conjunt inicial de 793 valors de  $pK_a$ . La Taula 4.34 mostra els paràmetres estadístics que s'obtenen per cadascun d'ells.

Taula 4.34. Estadística dels subconjunts d'àcids benzoïcs i fenols.

Subconjunt	n	R <sup>2</sup>	F	rmse
Fenols	274	0.982	15253.0	0.72
Àcids benzoïcs	519	0.997	205152.1	0.26
<i>orto</i>	378	0.992	47245.5	0.46
No <i>orto</i>	415	0.992	52192.9	0.47
Pròtics	407	0.988	34706.5	0.40
Apròtics	386	0.988	32678.7	0.54
H <sub>2</sub> O	188	0.976	7690.8	0.45
MeOH	146	0.982	7800.3	0.32
EtOH	23	0.970	686.3	0.16
PrOH	34	0.975	1261.5	0.41
BuOH	16	0.962	350.8	0.72
DMSO	127	0.971	4258.3	0.50
DMF	87	0.965	2386.6	0.48
AN	88	0.960	2052.4	0.53
Ac	65	0.922	748.2	0.54
NM	9	0.968	211.8	0.91
DMA	10	0.990	788.4	0.25

Com es pot comprovar, els resultats per a tots els subconjunts de compostos són molt bons. Un cop més, els resultats que s'obtenen per als conjunts de compostos *orto*/no *orto*-substituïts són molt semblants, indicant que el model establert és capaç d'encabir l'efecte *orto*. Els resultats que s'obtenen per als compostos pròtics i apròtics són, també, molt semblants entre ells. Cal destacar els resultats que s'obtenen en els solvents no coincidents en les dues famílies. En el cas dels fenols, els resultats en són molt semblants (*tert*-butanol) o millors (nitrometà) als que s'obtenien amb el seu model SVM individual. També s'aprecia una petita millora en els resultats dels àcids benzoïcs en etanol respecte als resultats que s'obtenien en el model SVM individual. Així doncs,

es pot dir que el model establert dóna bons resultats per a tots els conjunts de compostos i que, a més, permet fer una generalització molt bona en la predicció del  $pK_a$  d'aquestes dues famílies de compostos.

Taula 4.32. Valors experimentals i calculats de  $pK_a$  per a àcids benzoics i fenols.

#	Àcid	Solvent	Exp.	SVM
1	Àcid 2,3,4,5,6-pentametilbenzoic <sup>a</sup>	DMSO	11.31	11.10
2	Àcid 2,3,4,5,6-pentametilbenzoic <sup>a</sup>	MeOH	8.66	8.77
3	Àcid 2,3,4,5-tetrametilbenzoic <sup>b</sup>	DMSO	11.58	11.60
4	Àcid 2,3,4,5-tetrametilbenzoic <sup>a</sup>	MeOH	9.24	9.55
5	Àcid 2,3,4,6-tetrametilbenzoic <sup>a</sup>	DMSO	11.17	11.11
6	Àcid 2,3,4,6-tetrametilbenzoic <sup>a</sup>	MeOH	8.61	8.76
7	Àcid 2,3,4-trimetilbenzoic <sup>a</sup>	DMSO	11.26	11.10
8	Àcid 2,3,4-trimetilbenzoic <sup>a</sup>	MeOH	9.19	9.16
9	Àcid 2,3,5,6-tetrametilbenzoic <sup>a</sup>	DMSO	10.89	11.11
10	Àcid 2,3,5,6-tetrametilbenzoic <sup>a</sup>	MeOH	8.51	8.97
11	Àcid 2,3,5-trimetilbenzoic <sup>b</sup>	MeOH	9.05	9.60
12	Àcid 2,3,5-trimetilbenzoic <sup>a</sup>	DMSO	11.24	11.64
13	Àcid 2,3,6-triclorobenzoic <sup>a</sup>	H <sub>2</sub> O	1.50	1.45
14	Àcid 2,3,6-trimetilbenzoic <sup>a</sup>	DMSO	11.07	11.01
15	Àcid 2,3,6-trimetilbenzoic <sup>a</sup>	MeOH	8.42	8.92
16	Àcid 2,3-dihidroxibenzoic <sup>a</sup>	H <sub>2</sub> O	2.91	2.92
17	Àcid 2,3-dimetoxibenzoic <sup>a</sup>	H <sub>2</sub> O	3.98	4.09
18	Àcid 2,3-dimetilbenzoic <sup>a</sup>	DMSO	11.09	11.49
19	Àcid 2,3-dimetilbenzoic <sup>a</sup>	H <sub>2</sub> O	3.74	4.40
20	Àcid 2,3-dimetilbenzoic <sup>a</sup>	MeOH	8.98	9.57
21	Àcid 2,4,5-trimetilbenzoic <sup>a</sup>	DMSO	11.58	11.84
22	Àcid 2,4,5-trimetilbenzoic <sup>a</sup>	MeOH	9.60	9.81
23	Àcid 2,4,6-trimetilbenzoic <sup>a</sup>	DMSO	10.88	10.78
24	Àcid 2,4,6-trimetilbenzoic <sup>a</sup>	H <sub>2</sub> O	3.45	3.63
25	Àcid 2,4,6-trimetilbenzoic <sup>a</sup>	MeOH	8.57	8.72
26	Àcid 2,4,6-trimetilbenzoic <sup>a</sup>	AN	20.55	20.20
27	Àcid 2,4,6-trinitrobenzoic <sup>a</sup>	H <sub>2</sub> O	0.65	0.66
28	Àcid 2,4-diclorobenzoic <sup>b</sup>	H <sub>2</sub> O	2.76	2.62
29	Àcid 2,4-diclorobenzoic <sup>a</sup>	DMSO	9.12	9.05
30	Àcid 2,4-diclorobenzoic <sup>b</sup>	AN	18.40	18.86
31	Àcid 2,4-diclorobenzoic <sup>a</sup>	MeOH	7.80	7.91
32	Àcid 2,4-dimetoxibenzoic <sup>a</sup>	H <sub>2</sub> O	4.36	4.18
33	Àcid 2,4-dimetilbenzoic <sup>a</sup>	DMSO	11.42	11.48
34	Àcid 2,4-dimetilbenzoic <sup>a</sup>	H <sub>2</sub> O	4.22	4.40
35	Àcid 2,4-dimetilbenzoic <sup>b</sup>	MeOH	9.50	9.57
36	Àcid 2,4-dinitrobenzoic <sup>a</sup>	H <sub>2</sub> O	1.43	1.79
37	Àcid 2,4-dinitrobenzoic <sup>a</sup>	MeOH	6.45	7.07
38	Àcid 2,5-diclorobenzoic <sup>b</sup>	H <sub>2</sub> O	2.47	2.42
39	Àcid 2,5-dihidroxibenzoic <sup>a</sup>	H <sub>2</sub> O	2.95	2.95
40	Àcid 2,5-dihidroxibenzoic <sup>a</sup>	MeOH	8.04	8.10
41	Àcid 2,5-dimetilbenzoic <sup>a</sup>	H <sub>2</sub> O	3.98	4.35
42	Àcid 2,5-dimetilbenzoic <sup>b</sup>	DMSO	11.36	11.39
43	Àcid 2,5-dimetilbenzoic <sup>b</sup>	MeOH	9.29	9.52
44	Àcid 2,6-dibromobenzoic <sup>a</sup>	PrOH	8.61	8.67
45	Àcid 2,6-dibromobenzoic <sup>a</sup>	Ac	15.89	15.83
46	Àcid 2,6-dibromobenzoic <sup>a</sup>	AN	18.04	18.03
47	Àcid 2,6-dibromobenzoic <sup>a</sup>	DMSO	8.61	8.20
48	Àcid 2,6-dibromobenzoic <sup>a</sup>	EtOH	7.93	7.99
49	Àcid 2,6-dibromobenzoic <sup>a</sup>	MeOH	7.14	7.08
50	Àcid 2,6-dibromobenzoic <sup>b</sup>	DMF	9.72	9.70

#	Àcid	Solvent	Exp.	SVM
51	Àcid 2,6-dibutoxibenzoic <sup>a</sup>	<sup>t</sup> PrOH	11.48	11.42
52	Àcid 2,6-dibutoxibenzoic <sup>a</sup>	Ac	19.00	18.94
53	Àcid 2,6-dibutoxibenzoic <sup>a</sup>	AN	21.13	21.07
54	Àcid 2,6-dibutoxibenzoic <sup>a</sup>	DMF	12.90	12.96
55	Àcid 2,6-dibutoxibenzoic <sup>a</sup>	DMSO	11.45	11.39
56	Àcid 2,6-dibutoxibenzoic <sup>a</sup>	EtOH	10.19	10.25
57	Àcid 2,6-dibutoxibenzoic <sup>a</sup>	MeOH	9.13	9.19
58	Àcid 2,6-dibutoxibenzoic <sup>a</sup>	H <sub>2</sub> O	3.60	3.66
59	Àcid 2,6-diclorobenzoic <sup>a</sup>	Ac	15.68	15.88
60	Àcid 2,6-diclorobenzoic <sup>a</sup>	AN	18.16	18.21
61	Àcid 2,6-diclorobenzoic <sup>a</sup>	DMF	9.71	9.89
62	Àcid 2,6-diclorobenzoic <sup>a</sup>	DMSO	8.31	8.43
63	Àcid 2,6-diclorobenzoic <sup>a</sup>	EtOH	8.03	7.97
64	Àcid 2,6-diclorobenzoic <sup>a</sup>	H <sub>2</sub> O	1.82	1.76
65	Àcid 2,6-diclorobenzoic <sup>a</sup>	MeOH	7.26	7.15
66	Àcid 2,6-diclorobenzoic <sup>b</sup>	<sup>t</sup> PrOH	8.71	8.63
67	Àcid 2,6-dietoxibenzoic <sup>a</sup>	<sup>t</sup> PrOH	11.09	11.09
68	Àcid 2,6-dietoxibenzoic <sup>a</sup>	AN	20.97	20.91
69	Àcid 2,6-dietoxibenzoic <sup>a</sup>	DMF	12.97	12.91
70	Àcid 2,6-dietoxibenzoic <sup>a</sup>	DMSO	11.38	11.53
71	Àcid 2,6-dietoxibenzoic <sup>a</sup>	MeOH	9.05	8.86
72	Àcid 2,6-dietoxibenzoic <sup>b</sup>	Ac	18.88	18.66
73	Àcid 2,6-dietoxibenzoic <sup>b</sup>	EtOH	10.03	9.96
74	Àcid 2,6-difluorobenzoic <sup>a</sup>	<sup>t</sup> PrOH	8.93	8.87
75	Àcid 2,6-difluorobenzoic <sup>a</sup>	Ac	16.03	15.97
76	Àcid 2,6-difluorobenzoic <sup>a</sup>	AN	18.30	18.24
77	Àcid 2,6-difluorobenzoic <sup>a</sup>	DMF	9.91	9.97
78	Àcid 2,6-difluorobenzoic <sup>a</sup>	DMSO	8.59	8.38
79	Àcid 2,6-difluorobenzoic <sup>a</sup>	EtOH	8.22	8.23
80	Àcid 2,6-difluorobenzoic <sup>a</sup>	MeOH	7.39	7.40
81	Àcid 2,6-dihidroxibenzoic <sup>b</sup>	H <sub>2</sub> O	1.22	1.81
82	Àcid 2,6-diiodobenzoic <sup>a</sup>	<sup>t</sup> PrOH	9.18	9.07
83	Àcid 2,6-diiodobenzoic <sup>a</sup>	Ac	16.34	16.26
84	Àcid 2,6-diiodobenzoic <sup>a</sup>	AN	18.27	18.33
85	Àcid 2,6-diiodobenzoic <sup>a</sup>	EtOH	8.22	8.32
86	Àcid 2,6-diiodobenzoic <sup>b</sup>	DMF	10.13	9.94
87	Àcid 2,6-diiodobenzoic <sup>b</sup>	DMSO	8.75	8.44
88	Àcid 2,6-diiodobenzoic <sup>b</sup>	MeOH	7.29	7.29
89	Àcid 2,6-diisopropoxibenzoic <sup>a</sup>	<sup>t</sup> PrOH	11.61	11.55
90	Àcid 2,6-diisopropoxibenzoic <sup>a</sup>	Ac	19.06	19.17
91	Àcid 2,6-diisopropoxibenzoic <sup>a</sup>	AN	21.17	21.33
92	Àcid 2,6-diisopropoxibenzoic <sup>a</sup>	DMF	12.95	13.26
93	Àcid 2,6-diisopropoxibenzoic <sup>a</sup>	EtOH	10.34	10.36
94	Àcid 2,6-diisopropoxibenzoic <sup>a</sup>	MeOH	9.13	9.19
95	Àcid 2,6-diisopropoxibenzoic <sup>b</sup>	DMSO	11.50	11.75
96	Àcid 2,6-dimetoxibenzoic <sup>a</sup>	Ac	18.65	18.40
97	Àcid 2,6-dimetoxibenzoic <sup>a</sup>	DMF	12.72	12.66
98	Àcid 2,6-dimetoxibenzoic <sup>a</sup>	DMSO	11.29	11.32
99	Àcid 2,6-dimetoxibenzoic <sup>a</sup>	EtOH	9.84	9.90
100	Àcid 2,6-dimetoxibenzoic <sup>a</sup>	H <sub>2</sub> O	3.44	3.50
101	Àcid 2,6-dimetoxibenzoic <sup>b</sup>	<sup>t</sup> PrOH	10.81	10.94



#	Àcid	Solvent	Exp.	SVM
102	Àcid 2,6-dimetoxibenzoic <sup>b</sup>	AN	20.87	20.71
103	Àcid 2,6-dimetoxibenzoic <sup>b</sup>	MeOH	8.57	8.85
104	Àcid 2,6-dimetilbenzoic <sup>a</sup>	<sup>t</sup> PrOH	10.32	10.43
105	Àcid 2,6-dimetilbenzoic <sup>a</sup>	Ac	17.74	17.80
106	Àcid 2,6-dimetilbenzoic <sup>a</sup>	AN	20.31	20.50
107	Àcid 2,6-dimetilbenzoic <sup>a</sup>	DMF	11.96	11.98
108	Àcid 2,6-dimetilbenzoic <sup>a</sup>	DMSO	10.68	10.94
109	Àcid 2,6-dimetilbenzoic <sup>a</sup>	EtOH	9.42	9.78
110	Àcid 2,6-dimetilbenzoic <sup>a</sup>	H <sub>2</sub> O	3.25	3.90
111	Àcid 2,6-dimetilbenzoic <sup>a</sup>	MeOH	8.57	9.02
112	Àcid 2,6-dinitrobenzoic <sup>a</sup>	<sup>t</sup> PrOH	7.59	7.53
113	Àcid 2,6-dinitrobenzoic <sup>a</sup>	Ac	14.28	14.26
114	Àcid 2,6-dinitrobenzoic <sup>a</sup>	AN	16.67	16.61
115	Àcid 2,6-dinitrobenzoic <sup>a</sup>	DMF	8.26	8.20
116	Àcid 2,6-dinitrobenzoic <sup>a</sup>	DMSO	6.72	6.66
117	Àcid 2,6-dinitrobenzoic <sup>a</sup>	H <sub>2</sub> O	1.14	1.11
118	Àcid 2,6-dinitrobenzoic <sup>a</sup>	MeOH	6.49	6.43
119	Àcid 2,6-dinitrobenzoic <sup>b</sup>	EtOH	7.15	7.01
120	Àcid 2,6-dipropoxibenzoic <sup>a</sup>	<sup>t</sup> PrOH	11.43	11.37
121	Àcid 2,6-dipropoxibenzoic <sup>a</sup>	Ac	18.98	18.92
122	Àcid 2,6-dipropoxibenzoic <sup>a</sup>	AN	21.08	21.07
123	Àcid 2,6-dipropoxibenzoic <sup>a</sup>	DMSO	11.52	11.46
124	Àcid 2,6-dipropoxibenzoic <sup>a</sup>	EtOH	10.19	10.21
125	Àcid 2,6-dipropoxibenzoic <sup>a</sup>	MeOH	9.19	9.12
126	Àcid 2,6-dipropoxibenzoic <sup>b</sup>	DMF	12.92	12.99
127	Àcid 2-acetamidobenzoic <sup>a</sup>	H <sub>2</sub> O	3.67	3.99
128	Àcid 2-acetilbenzoic <sup>a</sup>	Ac	17.67	17.48
129	Àcid 2-acetilbenzoic <sup>a</sup>	AN	20.51	20.11
130	Àcid 2-acetilbenzoic <sup>a</sup>	DMF	11.68	11.59
131	Àcid 2-acetilbenzoic <sup>a</sup>	EtOH	10.07	9.70
132	Àcid 2-acetilbenzoic <sup>a</sup>	H <sub>2</sub> O	4.13	3.75
133	Àcid 2-acetilbenzoic <sup>a</sup>	MeOH	9.28	8.91
134	Àcid 2-aminobenzoic <sup>a</sup>	H <sub>2</sub> O	4.87	4.34
135	Àcid 2-bromobenzoic <sup>a</sup>	EtOH	8.98	8.76
136	Àcid 2-bromobenzoic <sup>b</sup>	H <sub>2</sub> O	2.85	2.79
137	Àcid 2-bromobenzoic <sup>b</sup>	MeOH	8.19	7.95
138	Àcid 2-bromobenzoic <sup>b</sup>	DMSO	9.57	9.05
139	Àcid 2-cloro-3-nitrobenzoic <sup>b</sup>	H <sub>2</sub> O	2.02	1.89
140	Àcid 2-cloro-4-nitrobenzoic <sup>a</sup>	DMSO	7.96	8.08
141	Àcid 2-cloro-4-nitrobenzoic <sup>a</sup>	H <sub>2</sub> O	1.92	2.14
142	Àcid 2-cloro-4-nitrobenzoic <sup>a</sup>	MeOH	7.12	7.39
143	Àcid 2-cloro-5-nitrobenzoic <sup>a</sup>	H <sub>2</sub> O	2.17	2.23
144	Àcid 2-cloro-5-nitrobenzoic <sup>b</sup>	DMSO	7.95	8.19
145	Àcid 2-cloro-6-metilbenzoic <sup>a</sup>	H <sub>2</sub> O	2.75	2.86
146	Àcid 2-cloro-6-nitrobenzoic <sup>a</sup>	AN	17.53	17.71
147	Àcid 2-cloro-6-nitrobenzoic <sup>a</sup>	DMF	9.14	9.15
148	Àcid 2-cloro-6-nitrobenzoic <sup>a</sup>	DMSO	7.66	7.72
149	Àcid 2-cloro-6-nitrobenzoic <sup>a</sup>	MeOH	6.86	7.00
150	Àcid 2-cloro-6-nitrobenzoic <sup>b</sup>	Ac	15.12	15.30
151	Àcid 2-cloro-6-nitrobenzoic <sup>a</sup>	H <sub>2</sub> O	1.34	1.63
152	Àcid 2-clorobenzoic <sup>a</sup>	DMA	9.60	9.39

#	Àcid	Solvent	Exp.	SVM
153	Àcid 2-clorobenzoic <sup>a</sup>	EtOH	9.08	9.12
154	Àcid 2-clorobenzoic <sup>a</sup>	H <sub>2</sub> O	2.92	2.98
155	Àcid 2-clorobenzoic <sup>a</sup>	MeOH	8.31	8.22
156	Àcid 2-clorobenzoic <sup>a</sup>	DMSO	9.16	9.39
157	Àcid 2-clorobenzoic <sup>b</sup>	AN	19.00	19.12
158	Àcid 2-cianobenzoic <sup>a</sup>	AN	19.07	19.01
159	Àcid 2-cianobenzoic <sup>a</sup>	MeOH	8.15	8.18
160	Àcid 2-cianobenzoic <sup>b</sup>	Ac	16.65	16.67
161	Àcid 2-cianobenzoic <sup>b</sup>	DMF	10.37	10.41
162	Àcid 2-cianobenzoic <sup>b</sup>	H <sub>2</sub> O	3.08	2.99
163	Àcid 2-fluorobenzoic <sup>a</sup>	EtOH	9.52	9.31
164	Àcid 2-fluorobenzoic <sup>a</sup>	H <sub>2</sub> O	3.57	3.30
165	Àcid 2-fluorobenzoic <sup>a</sup>	MeOH	8.41	8.51
166	Àcid 2-formilbenzoic <sup>a</sup>	Ac	17.39	17.45
167	Àcid 2-formilbenzoic <sup>a</sup>	AN	20.38	20.32
168	Àcid 2-formilbenzoic <sup>a</sup>	DMF	11.21	12.05
169	Àcid 2-formilbenzoic <sup>a</sup>	MeOH	9.19	8.72
170	Àcid 2-formilbenzoic <sup>a</sup>	H <sub>2</sub> O	4.55	3.76
171	Àcid 2-hidroxi-3,5-dinitrobenzoic <sup>a</sup>	MeOH	5.88	5.82
172	Àcid 2-hidroxi-5-bromobenzoic <sup>a</sup>	H <sub>2</sub> O	2.54	2.54
173	Àcid 2-hidroxi-5-bromobenzoic <sup>b</sup>	MeOH	7.76	7.68
174	Àcid 2-hidroxi-5-clorobenzoic <sup>a</sup>	H <sub>2</sub> O	2.59	2.63
175	Àcid 2-hidroxi-5-clorobenzoic <sup>a</sup>	MeOH	7.81	7.81
176	Àcid 2-hidroxi-5-fluorobenzoic <sup>a</sup>	MeOH	7.91	7.83
177	Àcid 2-hidroxi-5-fluorobenzoic <sup>b</sup>	H <sub>2</sub> O	2.70	2.64
178	Àcid 2-hidroxi-5-iodobenzoic <sup>b</sup>	H <sub>2</sub> O	2.62	2.61
179	Àcid 2-hidroxi-5-nitrobenzoic <sup>a</sup>	H <sub>2</sub> O	2.05	2.11
180	Àcid 2-hidroxi-5-nitrobenzoic <sup>a</sup>	MeOH	7.25	7.19
181	Àcid 2-hidroxibenzoic <sup>a</sup>	EtOH	8.45	8.39
182	Àcid 2-hidroxibenzoic <sup>a</sup>	H <sub>2</sub> O	2.77	2.83
183	Àcid 2-hidroxibenzoic <sup>a</sup>	MeOH	7.92	7.98
184	Àcid 2-iodobenzoic <sup>a</sup>	EtOH	9.09	9.03
185	Àcid 2-iodobenzoic <sup>a</sup>	H <sub>2</sub> O	2.96	3.03
186	Àcid 2-iodobenzoic <sup>b</sup>	MeOH	8.24	8.13
187	Àcid 2-iodobenzoic <sup>b</sup>	DMSO	9.90	9.30
188	Àcid 2-metoxi-6-clorobenzoic <sup>a</sup>	Ac	17.24	16.99
189	Àcid 2-metoxi-6-clorobenzoic <sup>a</sup>	AN	19.33	19.39
190	Àcid 2-metoxi-6-clorobenzoic <sup>a</sup>	DMF	11.24	11.18
191	Àcid 2-metoxi-6-clorobenzoic <sup>a</sup>	DMSO	9.76	9.84
192	Àcid 2-metoxi-6-clorobenzoic <sup>a</sup>	MeOH	8.06	8.00
193	Àcid 2-metoxi-6-nitrobenzoic <sup>a</sup>	Ac	16.65	16.55
194	Àcid 2-metoxi-6-nitrobenzoic <sup>a</sup>	AN	18.98	18.84
195	Àcid 2-metoxi-6-nitrobenzoic <sup>a</sup>	DMF	10.65	10.61
196	Àcid 2-metoxi-6-nitrobenzoic <sup>a</sup>	DMSO	9.16	9.14
197	Àcid 2-metoxi-6-nitrobenzoic <sup>a</sup>	MeOH	7.65	7.75
198	Àcid 2-metoxibenzoic <sup>a</sup>	MeOH	9.26	9.06
199	Àcid 2-metoxibenzoic <sup>b</sup>	EtOH	10.21	9.90
200	Àcid 2-metoxibenzoic <sup>b</sup>	H <sub>2</sub> O	4.09	3.89
201	Àcid 2-metil-6-clorobenzoic <sup>a</sup>	AN	19.17	18.93
202	Àcid 2-metil-6-clorobenzoic <sup>a</sup>	DMF	10.89	10.52
203	Àcid 2-metil-6-clorobenzoic <sup>a</sup>	DMSO	9.48	9.32

#	Àcid	Solvent	Exp.	SVM
204	Àcid 2-metil-6-clorobenzoic <sup>a</sup>	H <sub>2</sub> O	2.75	2.69
205	Àcid 2-metil-6-clorobenzoic <sup>a</sup>	MeOH	7.91	7.93
206	Àcid 2-metil-6-clorobenzoic <sup>b</sup>	Ac	16.74	16.38
207	Àcid 2-metil-6-metoxibenzoic <sup>a</sup>	AN	20.52	20.82
208	Àcid 2-metil-6-metoxibenzoic <sup>a</sup>	DMF	12.34	12.68
209	Àcid 2-metil-6-metoxibenzoic <sup>a</sup>	H <sub>2</sub> O	3.46	4.09
210	Àcid 2-metil-6-metoxibenzoic <sup>a</sup>	MeOH	8.70	9.25
211	Àcid 2-metil-6-metoxibenzoic <sup>b</sup>	Ac	18.24	18.41
212	Àcid 2-metil-6-metoxibenzoic <sup>b</sup>	DMSO	10.95	11.44
213	Àcid 2-metil-6-nitrobenzoic <sup>a</sup>	Ac	16.13	16.12
214	Àcid 2-metil-6-nitrobenzoic <sup>a</sup>	AN	18.63	18.61
215	Àcid 2-metil-6-nitrobenzoic <sup>a</sup>	DMSO	8.85	8.78
216	Àcid 2-metil-6-nitrobenzoic <sup>b</sup>	DMF	10.13	10.05
217	Àcid 2-metil-6-nitrobenzoic <sup>b</sup>	MeOH	7.56	7.78
218	Àcid 2-metil-6-nitrobenzoic <sup>a</sup>	H <sub>2</sub> O	2.21	2.51
219	Àcid 2-metilbenzoic <sup>a</sup>	DMSO	11.07	11.16
220	Àcid 2-metilbenzoic <sup>a</sup>	EtOH	10.23	10.27
221	Àcid 2-metilbenzoic <sup>a</sup>	H <sub>2</sub> O	3.91	4.27
222	Àcid 2-metilbenzoic <sup>a</sup>	MeOH	9.31	9.43
223	Àcid 2-metilèsterbenzoic <sup>a</sup>	Ac	17.34	17.54
224	Àcid 2-metilèsterbenzoic <sup>a</sup>	DMF	11.33	11.58
225	Àcid 2-metilèsterbenzoic <sup>a</sup>	H <sub>2</sub> O	3.18	3.57
226	Àcid 2-metilèsterbenzoic <sup>a</sup>	MeOH	8.47	8.79
227	Àcid 2-metilèsterbenzoic <sup>b</sup>	AN	19.78	20.01
228	Àcid 2-nitrobenzoic <sup>a</sup>	Ac	15.98	15.73
229	Àcid 2-nitrobenzoic <sup>a</sup>	AN	18.54	18.23
230	Àcid 2-nitrobenzoic <sup>a</sup>	DMF	9.96	9.56
231	Àcid 2-nitrobenzoic <sup>a</sup>	H <sub>2</sub> O	2.19	2.33
232	Àcid 2-nitrobenzoic <sup>a</sup>	MeOH	7.59	7.56
233	Àcid 2-nitrobenzoic <sup>b</sup>	DMSO	8.18	8.28
234	Àcid 2-nitrobenzoic <sup>b</sup>	EtOH	8.27	8.34
235	Àcid 3,4,5-trihidroxibenzoic <sup>a</sup>	H <sub>2</sub> O	4.21	3.98
236	Àcid 3,4,5-trimetoxibenzoic <sup>a</sup>	DMSO	10.30	10.44
237	Àcid 3,4,5-trimetilbenzoic <sup>a</sup>	DMSO	11.54	11.41
238	Àcid 3,4,5-trimetilbenzoic <sup>a</sup>	MeOH	9.66	9.59
239	Àcid 3,4-diclorobenzoic <sup>a</sup>	<sup>t</sup> PrOH	9.80	9.97
240	Àcid 3,4-diclorobenzoic <sup>a</sup>	AN	19.42	19.45
241	Àcid 3,4-diclorobenzoic <sup>a</sup>	DMF	10.79	10.84
242	Àcid 3,4-diclorobenzoic <sup>a</sup>	DMSO	9.60	9.66
243	Àcid 3,4-diclorobenzoic <sup>a</sup>	H <sub>2</sub> O	3.64	3.42
244	Àcid 3,4-diclorobenzoic <sup>a</sup>	MeOH	8.64	8.60
245	Àcid 3,4-diclorobenzoic <sup>b</sup>	Ac	16.70	16.88
246	Àcid 3,4-dihidroxibenzoic <sup>a</sup>	H <sub>2</sub> O	4.48	4.12
247	Àcid 3,4-dimetoxibenzoic <sup>a</sup>	Ac	18.59	18.44
248	Àcid 3,4-dimetoxibenzoic <sup>a</sup>	AN	20.95	20.81
249	Àcid 3,4-dimetoxibenzoic <sup>a</sup>	DMF	12.65	12.38
250	Àcid 3,4-dimetoxibenzoic <sup>a</sup>	DMSO	11.40	11.05
251	Àcid 3,4-dimetoxibenzoic <sup>a</sup>	MeOH	9.54	9.50
252	Àcid 3,4-dimetoxibenzoic <sup>b</sup>	H <sub>2</sub> O	4.44	4.36
253	Àcid 3,4-dimetilbenzoic <sup>a</sup>	<sup>t</sup> PrOH	11.60	11.12
254	Àcid 3,4-dimetilbenzoic <sup>a</sup>	Ac	18.71	18.46

#	Àcid	Solvent	Exp.	SVM
255	Àcid 3,4-dimetilbenzoic <sup>a</sup>	AN	21.05	20.96
256	Àcid 3,4-dimetilbenzoic <sup>a</sup>	DMF	12.70	12.48
257	Àcid 3,4-dimetilbenzoic <sup>a</sup>	H <sub>2</sub> O	4.40	4.40
258	Àcid 3,4-dimetilbenzoic <sup>a</sup>	MeOH	9.63	9.58
259	Àcid 3,4-dimetilbenzoic <sup>b</sup>	DMSO	11.46	11.26
260	Àcid 3,4-dinitrobenzoic <sup>a</sup>	Ac	15.07	15.37
261	Àcid 3,4-dinitrobenzoic <sup>a</sup>	DMF	9.07	9.04
262	Àcid 3,4-dinitrobenzoic <sup>a</sup>	DMSO	7.88	7.82
263	Àcid 3,4-dinitrobenzoic <sup>a</sup>	H <sub>2</sub> O	2.82	2.37
264	Àcid 3,4-dinitrobenzoic <sup>a</sup>	MeOH	7.44	7.50
265	Àcid 3,4-dinitrobenzoic <sup>b</sup>	AN	17.50	17.89
266	Àcid 3,5-dibromobenzoic <sup>a</sup>	Ac	16.23	16.61
267	Àcid 3,5-dibromobenzoic <sup>a</sup>	AN	18.86	19.13
268	Àcid 3,5-dibromobenzoic <sup>a</sup>	DMSO	8.88	9.34
269	Àcid 3,5-dibromobenzoic <sup>a</sup>	MeOH	8.29	8.40
270	Àcid 3,5-dibromobenzoic <sup>b</sup>	DMF	10.11	10.54
271	Àcid 3,5-diclorobenzoic <sup>a</sup>	DMSO	8.81	9.27
272	Àcid 3,5-diclorobenzoic <sup>a</sup>	H <sub>2</sub> O	3.10	3.15
273	Àcid 3,5-diclorobenzoic <sup>a</sup>	MeOH	8.26	8.32
274	Àcid 3,5-diclorobenzoic <sup>a</sup>	AN	18.75	19.08
275	Àcid 3,5-diclorobenzoic <sup>a</sup>	DMF	10.43	10.44
276	Àcid 3,5-dihidroxi benzoic <sup>a</sup>	H <sub>2</sub> O	4.04	3.92
277	Àcid 3,5-dimetoxibenzoic <sup>a</sup>	Ac	18.02	18.08
278	Àcid 3,5-dimetoxibenzoic <sup>a</sup>	AN	20.39	20.48
279	Àcid 3,5-dimetoxibenzoic <sup>a</sup>	DMF	12.04	12.05
280	Àcid 3,5-dimetoxibenzoic <sup>a</sup>	DMSO	10.71	10.75
281	Àcid 3,5-dimetoxibenzoic <sup>a</sup>	H <sub>2</sub> O	3.97	4.12
282	Àcid 3,5-dimetoxibenzoic <sup>a</sup>	MeOH	9.27	9.27
283	Àcid 3,5-dimetilbenzoic <sup>a</sup>	Ac	18.54	18.42
284	Àcid 3,5-dimetilbenzoic <sup>a</sup>	DMF	12.57	12.47
285	Àcid 3,5-dimetilbenzoic <sup>a</sup>	DMSO	11.29	11.27
286	Àcid 3,5-dimetilbenzoic <sup>a</sup>	MeOH	9.59	9.53
287	Àcid 3,5-dimetilbenzoic <sup>b</sup>	AN	20.95	20.95
288	Àcid 3,5-dimetilbenzoic <sup>b</sup>	H <sub>2</sub> O	4.30	4.36
289	Àcid 3,5-dinitrobenzoic <sup>a</sup>	PrOH	8.30	8.78
290	Àcid 3,5-dinitrobenzoic <sup>a</sup>	AN	17.36	17.96
291	Àcid 3,5-dinitrobenzoic <sup>a</sup>	DMF	8.76	9.14
292	Àcid 3,5-dinitrobenzoic <sup>a</sup>	DMSO	7.38	7.86
293	Àcid 3,5-dinitrobenzoic <sup>a</sup>	H <sub>2</sub> O	2.67	2.24
294	Àcid 3,5-dinitrobenzoic <sup>a</sup>	MeOH	7.64	7.42
295	Àcid 3,5-dinitrobenzoic <sup>b</sup>	Ac	14.87	15.50
296	Àcid 3,6-dicloro-2-metoxibenzoic <sup>a</sup>	H <sub>2</sub> O	1.97	2.17
297	Àcid 3-acetamidobenzoic <sup>a</sup>	H <sub>2</sub> O	4.06	4.14
298	Àcid 3-acetamidobenzoic <sup>a</sup>	MeOH	9.25	9.28
299	Àcid 3-acetilbenzoic <sup>a</sup>	DMSO	10.22	10.48
300	Àcid 3-acetilbenzoic <sup>a</sup>	H <sub>2</sub> O	3.83	4.01
301	Àcid 3-acetilbenzoic <sup>a</sup>	MeOH	8.87	9.18
302	Àcid 3-amino-2,5-diclorobenzoic <sup>a</sup>	H <sub>2</sub> O	3.40	2.75
303	Àcid 3-aminobenzoic <sup>a</sup>	DMSO	11.60	11.04
304	Àcid 3-aminobenzoic <sup>a</sup>	H <sub>2</sub> O	4.74	4.23
305	Àcid 3-bromo-4-metoxibenzoic <sup>a</sup>	Ac	17.82	18.06

#	Àcid	Solvent	Exp.	SVM
306	Àcid 3-bromo-4-metoxibenzoic <sup>a</sup>	DMF	11.83	12.08
307	Àcid 3-bromo-4-metoxibenzoic <sup>a</sup>	DMSO	10.61	10.86
308	Àcid 3-bromo-4-metoxibenzoic <sup>a</sup>	MeOH	9.27	9.32
309	Àcid 3-bromo-4-metoxibenzoic <sup>b</sup>	AN	20.38	20.57
310	Àcid 3-bromo-4-metilbenzoic <sup>a</sup>	Ac	17.59	17.55
311	Àcid 3-bromo-4-metilbenzoic <sup>a</sup>	AN	20.03	20.09
312	Àcid 3-bromo-4-metilbenzoic <sup>a</sup>	DMF	11.49	11.56
313	Àcid 3-bromo-4-metilbenzoic <sup>a</sup>	DMSO	10.30	10.36
314	Àcid 3-bromo-4-metilbenzoic <sup>b</sup>	MeOH	9.06	9.07
315	Àcid 3-bromo-4-metilbenzoic <sup>a</sup>	H <sub>2</sub> O	3.29	3.90
316	Àcid 3-bromobenzoic <sup>a</sup>	PrOH	10.11	10.17
317	Àcid 3-bromobenzoic <sup>a</sup>	DMSO	9.68	9.96
318	Àcid 3-bromobenzoic <sup>a</sup>	EtOH	9.40	9.61
319	Àcid 3-bromobenzoic <sup>a</sup>	H <sub>2</sub> O	3.81	3.68
320	Àcid 3-bromobenzoic <sup>a</sup>	MeOH	8.80	8.81
321	Àcid 3-bromobenzoic <sup>a</sup>	AN	19.50	19.67
322	Àcid 3-bromobenzoic <sup>a</sup>	DMF	11.20	11.15
323	Àcid 3-cloro-4-nitrobenzoic <sup>a</sup>	Ac	15.78	16.10
324	Àcid 3-cloro-4-nitrobenzoic <sup>a</sup>	DMF	10.07	9.89
325	Àcid 3-cloro-4-nitrobenzoic <sup>a</sup>	DMSO	8.61	8.68
326	Àcid 3-cloro-4-nitrobenzoic <sup>a</sup>	MeOH	8.09	7.96
327	Àcid 3-cloro-4-nitrobenzoic <sup>b</sup>	AN	18.28	18.64
328	Àcid 3-cloro-5-nitrobenzoic <sup>a</sup>	Ac	15.57	15.91
329	Àcid 3-cloro-5-nitrobenzoic <sup>a</sup>	AN	17.98	18.43
330	Àcid 3-cloro-5-nitrobenzoic <sup>a</sup>	DMF	9.60	9.68
331	Àcid 3-cloro-5-nitrobenzoic <sup>a</sup>	DMSO	8.37	8.44
332	Àcid 3-cloro-5-nitrobenzoic <sup>a</sup>	MeOH	8.03	7.77
333	Àcid 3-cloro-5-nitrobenzoic <sup>a</sup>	H <sub>2</sub> O	3.13	2.56
334	Àcid 3-clorobenzoic <sup>a</sup>	H <sub>2</sub> O	3.80	3.77
335	Àcid 3-clorobenzoic <sup>b</sup>	DMA	9.80	10.13
336	Àcid 3-clorobenzoic <sup>b</sup>	DMSO	9.51	10.12
337	Àcid 3-clorobenzoic <sup>b</sup>	MeOH	8.83	8.91
338	Àcid 3-clorobenzoic <sup>b</sup>	DMF	10.95	11.33
339	Àcid 3-cianobenzoic <sup>a</sup>	DMSO	9.44	9.79
340	Àcid 3-cianobenzoic <sup>a</sup>	H <sub>2</sub> O	3.60	3.61
341	Àcid 3-cianobenzoic <sup>a</sup>	MeOH	8.53	8.75
342	Àcid 3-fluorobenzoic <sup>a</sup>	MeOH	8.87	8.76
343	Àcid 3-fluorobenzoic <sup>b</sup>	H <sub>2</sub> O	3.88	3.59
344	Àcid 3-formilbenzoic <sup>a</sup>	H <sub>2</sub> O	3.84	3.93
345	Àcid 3-hidroxibenzoic <sup>a</sup>	DMA	10.75	10.45
346	Àcid 3-hidroxibenzoic <sup>a</sup>	H <sub>2</sub> O	4.03	3.89
347	Àcid 3-hidroxibenzoic <sup>b</sup>	MeOH	9.58	9.06
348	Àcid 3-iodobenzoic <sup>a</sup>	DMSO	10.16	10.15
349	Àcid 3-iodobenzoic <sup>a</sup>	H <sub>2</sub> O	3.82	3.88
350	Àcid 3-iodobenzoic <sup>a</sup>	MeOH	8.89	8.95
351	Àcid 3-metoxi-4-clorobenzoic <sup>a</sup>	Ac	17.39	17.42
352	Àcid 3-metoxi-4-clorobenzoic <sup>a</sup>	AN	20.02	19.98
353	Àcid 3-metoxi-4-clorobenzoic <sup>a</sup>	DMF	11.58	11.43
354	Àcid 3-metoxi-4-clorobenzoic <sup>b</sup>	DMSO	10.39	10.27
355	Àcid 3-metoxi-4-clorobenzoic <sup>b</sup>	MeOH	8.98	8.90
356	Àcid 3-metoxi-4-metilbenzoic <sup>a</sup>	AN	20.76	20.70



#	Àcid	Solvent	Exp.	SVM
357	Àcid 3-metoxi-4-metilbenzoic <sup>a</sup>	DMF	12.39	12.19
358	Àcid 3-metoxi-4-metilbenzoic <sup>a</sup>	DMSO	11.06	11.00
359	Àcid 3-metoxi-4-metilbenzoic <sup>a</sup>	MeOH	9.49	9.33
360	Àcid 3-metoxi-4-metilbenzoic <sup>b</sup>	Ac	18.32	18.18
361	Àcid 3-metoxi-4-metilbenzoic <sup>a</sup>	H <sub>2</sub> O	4.13	4.19
362	Àcid 3-metoxi-4-nitrobenzoic <sup>a</sup>	Ac	16.21	16.98
363	Àcid 3-metoxi-4-nitrobenzoic <sup>a</sup>	AN	18.77	19.51
364	Àcid 3-metoxi-4-nitrobenzoic <sup>a</sup>	DMF	10.33	10.88
365	Àcid 3-metoxi-4-nitrobenzoic <sup>a</sup>	DMSO	9.22	9.67
366	Àcid 3-metoxi-4-nitrobenzoic <sup>a</sup>	MeOH	9.28	8.92
367	Àcid 3-metoxi-5-bromobenzoic <sup>a</sup>	Ac	17.03	17.38
368	Àcid 3-metoxi-5-bromobenzoic <sup>a</sup>	DMF	11.28	11.35
369	Àcid 3-metoxi-5-bromobenzoic <sup>a</sup>	DMSO	9.79	10.15
370	Àcid 3-metoxi-5-bromobenzoic <sup>a</sup>	MeOH	8.79	8.88
371	Àcid 3-metoxi-5-bromobenzoic <sup>b</sup>	AN	19.73	19.91
372	Àcid 3-metoxi-5-nitrobenzoic <sup>a</sup>	Ac	16.31	16.44
373	Àcid 3-metoxi-5-nitrobenzoic <sup>a</sup>	AN	18.72	18.95
374	Àcid 3-metoxi-5-nitrobenzoic <sup>a</sup>	MeOH	8.40	8.46
375	Àcid 3-metoxi-5-nitrobenzoic <sup>b</sup>	DMF	10.37	10.33
376	Àcid 3-metoxi-5-nitrobenzoic <sup>b</sup>	DMSO	9.01	9.11
377	Àcid 3-metoxibenzoic <sup>a</sup>	DMSO	11.10	10.68
378	Àcid 3-metoxibenzoic <sup>a</sup>	H <sub>2</sub> O	4.12	4.00
379	Àcid 3-metoxibenzoic <sup>a</sup>	MeOH	9.30	9.15
380	Àcid 3-metil-4-clorobenzoic <sup>a</sup>	AN	20.27	20.15
381	Àcid 3-metil-4-clorobenzoic <sup>a</sup>	DMF	11.70	11.54
382	Àcid 3-metil-4-clorobenzoic <sup>a</sup>	DMSO	10.54	10.37
383	Àcid 3-metil-4-clorobenzoic <sup>a</sup>	H <sub>2</sub> O	4.07	3.91
384	Àcid 3-metil-4-clorobenzoic <sup>b</sup>	Ac	17.59	17.58
385	Àcid 3-metil-4-clorobenzoic <sup>b</sup>	MeOH	9.12	9.07
386	Àcid 3-metil-4-metoxibenzoic <sup>a</sup>	Ac	19.01	18.74
387	Àcid 3-metil-4-metoxibenzoic <sup>a</sup>	AN	21.29	21.22
388	Àcid 3-metil-4-metoxibenzoic <sup>a</sup>	DMF	12.84	12.78
389	Àcid 3-metil-4-metoxibenzoic <sup>a</sup>	DMSO	11.71	11.55
390	Àcid 3-metil-4-metoxibenzoic <sup>a</sup>	MeOH	9.81	9.63
391	Àcid 3-metil-4-metoxibenzoic <sup>b</sup>	H <sub>2</sub> O	4.35	4.48
392	Àcid 3-metil-4-nitrobenzoic <sup>a</sup>	Ac	16.55	16.49
393	Àcid 3-metil-4-nitrobenzoic <sup>a</sup>	AN	19.21	19.06
394	Àcid 3-metil-4-nitrobenzoic <sup>a</sup>	DMF	11.03	10.28
395	Àcid 3-metil-4-nitrobenzoic <sup>a</sup>	DMSO	9.48	9.11
396	Àcid 3-metil-4-nitrobenzoic <sup>a</sup>	H <sub>2</sub> O	3.65	3.18
397	Àcid 3-metil-4-nitrobenzoic <sup>a</sup>	MeOH	8.54	8.34
398	Àcid 3-metil-5-clorobenzoic <sup>a</sup>	Ac	17.40	17.42
399	Àcid 3-metil-5-clorobenzoic <sup>a</sup>	DMF	11.35	11.41
400	Àcid 3-metil-5-clorobenzoic <sup>a</sup>	DMSO	10.04	10.21
401	Àcid 3-metil-5-clorobenzoic <sup>a</sup>	MeOH	8.94	8.95
402	Àcid 3-metil-5-clorobenzoic <sup>b</sup>	AN	19.85	19.95
403	Àcid 3-metil-5-metoxibenzoic <sup>a</sup>	Ac	18.26	18.17
404	Àcid 3-metil-5-metoxibenzoic <sup>a</sup>	AN	20.66	20.72
405	Àcid 3-metil-5-metoxibenzoic <sup>a</sup>	DMSO	10.93	11.06
406	Àcid 3-metil-5-metoxibenzoic <sup>a</sup>	MeOH	9.44	9.29
407	Àcid 3-metil-5-metoxibenzoic <sup>b</sup>	DMF	12.32	12.22

#	Àcid	Solvent	Exp.	SVM
408	Àcid 3-metil-5-nitrobenzoic <sup>a</sup>	AN	19.18	19.09
409	Àcid 3-metil-5-nitrobenzoic <sup>a</sup>	DMF	10.92	10.33
410	Àcid 3-metil-5-nitrobenzoic <sup>a</sup>	DMSO	9.36	9.14
411	Àcid 3-metil-5-nitrobenzoic <sup>a</sup>	MeOH	8.55	8.38
412	Àcid 3-metil-5-nitrobenzoic <sup>b</sup>	Ac	16.62	16.53
413	Àcid 3-metilbenzoic <sup>a</sup>	DMSO	11.26	10.95
414	Àcid 3-metilbenzoic <sup>a</sup>	H <sub>2</sub> O	4.24	4.23
415	Àcid 3-metilbenzoic <sup>a</sup>	MeOH	9.47	9.40
416	Àcid 3-nitro-4-clorobenzoic <sup>a</sup>	Ac	15.98	16.25
417	Àcid 3-nitro-4-clorobenzoic <sup>a</sup>	AN	18.57	18.82
418	Àcid 3-nitro-4-clorobenzoic <sup>a</sup>	DMF	9.98	10.04
419	Àcid 3-nitro-4-clorobenzoic <sup>a</sup>	MeOH	8.24	8.18
420	Àcid 3-nitro-4-clorobenzoic <sup>b</sup>	PrOH	9.30	9.52
421	Àcid 3-nitro-4-clorobenzoic <sup>b</sup>	DMSO	8.74	8.85
422	Àcid 3-nitro-4-clorobenzoic <sup>b</sup>	H <sub>2</sub> O	3.29	3.00
423	Àcid 3-nitro-4-metoxibenzoic <sup>a</sup>	Ac	17.04	17.36
424	Àcid 3-nitro-4-metoxibenzoic <sup>a</sup>	AN	19.51	19.86
425	Àcid 3-nitro-4-metoxibenzoic <sup>a</sup>	DMF	11.18	11.28
426	Àcid 3-nitro-4-metoxibenzoic <sup>a</sup>	DMSO	9.85	10.04
427	Àcid 3-nitro-4-metoxibenzoic <sup>b</sup>	MeOH	8.83	9.04
428	Àcid 3-nitro-4-metoxibenzoic <sup>b</sup>	H <sub>2</sub> O	3.72	3.90
429	Àcid 3-nitro-4-metilbenzoic <sup>a</sup>	Ac	16.75	16.84
430	Àcid 3-nitro-4-metilbenzoic <sup>a</sup>	DMF	10.93	10.67
431	Àcid 3-nitro-4-metilbenzoic <sup>a</sup>	DMSO	9.45	9.49
432	Àcid 3-nitro-4-metilbenzoic <sup>a</sup>	MeOH	8.61	8.65
433	Àcid 3-nitro-4-metilbenzoic <sup>b</sup>	AN	19.36	19.42
434	Àcid 3-nitro-4-metilbenzoic <sup>b</sup>	H <sub>2</sub> O	3.62	3.49
435	Àcid 3-nitrobenzoic <sup>a</sup>	DMSO	9.17	9.23
436	Àcid 3-nitrobenzoic <sup>a</sup>	MeOH	8.32	8.46
437	Àcid 3-nitrobenzoic <sup>b</sup>	H <sub>2</sub> O	3.46	3.30
438	Àcid 3-nitrobenzoic <sup>a</sup>	AN	19.29	19.19
439	Àcid 3-trifluorometilbenzoic <sup>a</sup>	H <sub>2</sub> O	3.75	3.81
440	Àcid 3-trifluorometilbenzoic <sup>a</sup>	MeOH	8.69	8.75
441	Àcid 4-(dimetilamino)benzoic <sup>a</sup>	H <sub>2</sub> O	5.03	4.97
442	Àcid 4-( <i>N,N</i> -dimetilaminocarbonil)benzoic <sup>a</sup>	Ac	17.83	17.59
443	Àcid 4-( <i>N,N</i> -dimetilaminocarbonil)benzoic <sup>a</sup>	AN	20.55	20.15
444	Àcid 4-( <i>N,N</i> -dimetilaminocarbonil)benzoic <sup>b</sup>	MeOH	9.00	9.08
445	Àcid 4-acetamidobenzoic <sup>a</sup>	MeOH	9.57	9.68
446	Àcid 4-acetamidobenzoic <sup>b</sup>	H <sub>2</sub> O	4.30	4.52
447	Àcid 4-acetilbenzoic <sup>a</sup>	AN	19.95	20.07
448	Àcid 4-acetilbenzoic <sup>a</sup>	DMF	11.38	11.40
449	Àcid 4-acetilbenzoic <sup>a</sup>	H <sub>2</sub> O	3.74	3.86
450	Àcid 4-acetilbenzoic <sup>a</sup>	MeOH	8.84	9.02
451	Àcid 4-acetilbenzoic <sup>b</sup>	Ac	17.32	17.49
452	Àcid 4-aminobenzoic <sup>a</sup>	H <sub>2</sub> O	4.82	4.61
453	Àcid 4-aminobenzoic <sup>a</sup>	MeOH	10.25	9.79
454	Àcid 4-bromobenzoic <sup>a</sup>	DMSO	10.20	9.92
455	Àcid 4-bromobenzoic <sup>a</sup>	H <sub>2</sub> O	3.99	3.68
456	Àcid 4-bromobenzoic <sup>b</sup>	MeOH	8.93	8.79
457	Àcid 4-bromobenzoic <sup>a</sup>	AN	20.30	19.61
458	Àcid 4-clorobenzoic <sup>a</sup>	DMA	10.30	10.57

#	Àcid	Solvent	Exp.	SVM
459	Àcid 4-clorobenzoic <sup>a</sup>	H <sub>2</sub> O	4.00	4.09
460	Àcid 4-clorobenzoic <sup>a</sup>	MeOH	9.09	9.26
461	Àcid 4-clorobenzoic <sup>b</sup>	DMSO	10.10	10.57
462	Àcid 4-cianobenzoic <sup>a</sup>	Ac	16.69	17.07
463	Àcid 4-cianobenzoic <sup>a</sup>	AN	19.34	19.59
464	Àcid 4-cianobenzoic <sup>a</sup>	DMF	11.02	10.96
465	Àcid 4-cianobenzoic <sup>a</sup>	DMSO	9.27	9.76
466	Àcid 4-cianobenzoic <sup>a</sup>	H <sub>2</sub> O	3.53	3.59
467	Àcid 4-cianobenzoic <sup>a</sup>	MeOH	8.49	8.73
468	Àcid 4-etoxibenzoic <sup>a</sup>	DMSO	11.50	11.49
469	Àcid 4-etoxibenzoic <sup>a</sup>	H <sub>2</sub> O	4.80	4.44
470	Àcid 4-etilbenzoic <sup>b</sup>	H <sub>2</sub> O	4.35	4.30
471	Àcid 4-fluorobenzoic <sup>a</sup>	H <sub>2</sub> O	4.15	3.90
472	Àcid 4-fluorobenzoic <sup>b</sup>	MeOH	9.23	9.05
473	Àcid 4-formilbenzoic <sup>a</sup>	Ac	17.10	17.30
474	Àcid 4-formilbenzoic <sup>a</sup>	AN	19.74	19.92
475	Àcid 4-formilbenzoic <sup>a</sup>	DMF	11.19	11.32
476	Àcid 4-formilbenzoic <sup>a</sup>	H <sub>2</sub> O	3.69	3.70
477	Àcid 4-formilbenzoic <sup>a</sup>	MeOH	8.94	8.84
478	Àcid 4-hidroxi-3,5-dimetoxibenzoic <sup>a</sup>	H <sub>2</sub> O	4.34	4.28
479	Àcid 4-hidroxi-3-metoxibenzoic <sup>a</sup>	H <sub>2</sub> O	4.51	4.27
480	Àcid 4-hidroxibenzoic <sup>a</sup>	AN	20.80	20.86
481	Àcid 4-hidroxibenzoic <sup>a</sup>	DMA	11.25	11.18
482	Àcid 4-hidroxibenzoic <sup>a</sup>	H <sub>2</sub> O	4.55	4.36
483	Àcid 4-hidroxibenzoic <sup>a</sup>	MeOH	9.99	9.54
484	Àcid 4-iodo-3-nitrobenzoic <sup>a</sup>	DMSO	8.65	8.52
485	Àcid 4-iodobenzoic <sup>a</sup>	MeOH	9.04	8.80
486	Àcid 4-iodobenzoic <sup>b</sup>	H <sub>2</sub> O	3.98	3.74
487	Àcid 4-metoxibenzoic <sup>a</sup>	DMSO	11.80	11.38
488	Àcid 4-metoxibenzoic <sup>a</sup>	H <sub>2</sub> O	4.25	4.40
489	Àcid 4-metoxibenzoic <sup>a</sup>	MeOH	9.79	9.56
490	Àcid 4-metilbenzoic <sup>a</sup>	H <sub>2</sub> O	4.38	4.39
491	Àcid 4-metilbenzoic <sup>b</sup>	DMSO	11.42	11.18
492	Àcid 4-metilbenzoic <sup>b</sup>	MeOH	9.61	9.56
493	Àcid 4-metilèsterbenzoic <sup>a</sup>	Ac	17.37	17.18
494	Àcid 4-metilèsterbenzoic <sup>a</sup>	AN	19.91	19.77
495	Àcid 4-metilèsterbenzoic <sup>a</sup>	H <sub>2</sub> O	3.74	3.74
496	Àcid 4-metilèsterbenzoic <sup>a</sup>	MeOH	8.84	8.86
497	Àcid 4-metilèsterbenzoic <sup>b</sup>	DMF	11.45	11.09
498	Àcid 4-nitrobenzoic <sup>a</sup>	PrOH	9.60	9.53
499	Àcid 4-nitrobenzoic <sup>a</sup>	Ac	16.42	16.32
500	Àcid 4-nitrobenzoic <sup>a</sup>	AN	19.08	18.88
501	Àcid 4-nitrobenzoic <sup>a</sup>	DMF	10.80	10.08
502	Àcid 4-nitrobenzoic <sup>a</sup>	DMSO	9.04	8.89
503	Àcid 4-nitrobenzoic <sup>a</sup>	EtOH	8.90	8.96
504	Àcid 4-nitrobenzoic <sup>a</sup>	H <sub>2</sub> O	3.43	3.03
505	Àcid 4-nitrobenzoic <sup>b</sup>	MeOH	8.36	8.18
506	Àcid 4-propoxibenzoic <sup>a</sup>	H <sub>2</sub> O	4.46	4.59
507	Àcid 4- <i>t</i> -butilbenzoic <sup>a</sup>	H <sub>2</sub> O	4.36	3.92
508	Àcid 4- <i>t</i> -butilbenzoic <sup>a</sup>	MeOH	9.61	9.03
509	Àcid acetilsalicílic <sup>a</sup>	DMF	11.30	13.12

#	Àcid	Solvent	Exp.	SVM
510	Àcid acetilsalicilic <sup>a</sup>	H <sub>2</sub> O	3.69	3.54
511	Àcid benzoic <sup>a</sup>	Ac	18.2	17.9
512	Àcid benzoic <sup>a</sup>	AN	20.7	20.41
513	Àcid benzoic <sup>a</sup>	DMA	11	10.74
514	Àcid benzoic <sup>a</sup>	DMSO	11	10.74
515	Àcid benzoic <sup>a</sup>	EtOH	10.25	10.13
516	Àcid benzoic <sup>a</sup>	H <sub>2</sub> O	4.21	4.15
517	Àcid benzoic <sup>a</sup>	MeOH	9.41	9.29
518	Àcid benzoic <sup>b</sup>	<sup>t</sup> PrOH	10.71	10.71
519	Àcid benzoic <sup>b</sup>	DMF	12.27	11.94
520	2,3,4,6-tetraclorofenol <sup>a</sup>	DMSO	7.55	8.49
521	2,3,4,6-tetraclorofenol <sup>a</sup>	H <sub>2</sub> O	5.63	5.57
522	2,3,4,6-tetraclorofenol <sup>b</sup>	DMF	9.50	10.35
523	2,3,4,6-tetranitrofenol <sup>b</sup>	DMF	1.11	0.17
524	2,3-dihidroxifenol <sup>a</sup>	H <sub>2</sub> O	9.01	8.95
525	2,3-dimetilfenol <sup>a</sup>	H <sub>2</sub> O	10.54	10.35
526	2,3-dimetilfenol <sup>a</sup>	MeOH	15.08	14.95
527	2,3-dinitrofenol <sup>a</sup>	H <sub>2</sub> O	5.24	5.46
528	2,3-dinitrofenol <sup>a</sup>	MeOH	9.43	9.49
529	2,4,5-triclorofenol <sup>a</sup>	DMF	12.46	12.92
530	2,4,5-triclorofenol <sup>b</sup>	DMSO	10.97	12.30
531	2,4,6-tribromofenol <sup>a</sup>	Ac	21.10	21.04
532	2,4,6-tribromofenol <sup>b</sup>	H <sub>2</sub> O	6.10	5.92
533	2,4,6-tribromofenol <sup>b</sup>	MeOH	10.10	10.76
534	2,4,6-triclorofenol <sup>a</sup>	Ac	22.50	22.56
535	2,4,6-triclorofenol <sup>a</sup>	DMSO	10.19	10.13
536	2,4,6-triclorofenol <sup>a</sup>	H <sub>2</sub> O	6.42	6.56
537	2,4,6-triclorofenol <sup>a</sup>	<sup>t</sup> BuOH	14.82	15.87
538	2,4,6-triclorofenol <sup>a</sup>	DMF	12.05	11.99
539	2,4,6-triclorofenol <sup>b</sup>	<sup>t</sup> PrOH	12.55	13.19
540	2,4,6-trimetil-3-nitrofenol <sup>a</sup>	H <sub>2</sub> O	8.98	6.96
541	2,4,6-trimetilfenol <sup>a</sup>	H <sub>2</sub> O	10.86	10.44
542	2,4,6-trimetilfenol <sup>a</sup>	MeOH	15.53	15.12
543	2,4,6-trinitrofenol <sup>a</sup>	<sup>t</sup> PrOH	3.70	3.73
544	2,4,6-trinitrofenol <sup>a</sup>	Ac	9.20	9.14
545	2,4,6-trinitrofenol <sup>a</sup>	AN	11.00	11.12
546	2,4,6-trinitrofenol <sup>a</sup>	DMF	3.65	2.70
547	2,4,6-trinitrofenol <sup>a</sup>	DMSO	-0.30	0.84
548	2,4,6-trinitrofenol <sup>a</sup>	H <sub>2</sub> O	0.43	1.00
549	2,4,6-trinitrofenol <sup>a</sup>	NM	8.02	9.44
550	2,4,6-trinitrofenol <sup>a</sup>	<sup>t</sup> BuOH	4.70	4.64
551	2,4,6-trinitrofenol <sup>a</sup>	MeOH	3.90	3.96
552	2,4-diclorofenol <sup>a</sup>	<sup>t</sup> PrOH	14.48	14.21
553	2,4-diclorofenol <sup>a</sup>	DMSO	13.25	13.72
554	2,4-diclorofenol <sup>a</sup>	H <sub>2</sub> O	7.65	7.59
555	2,4-diclorofenol <sup>a</sup>	<sup>t</sup> BuOH	17.25	17.70
556	2,4-diclorofenol <sup>b</sup>	DMF	13.56	14.05
557	2,4-dimetilfenol <sup>a</sup>	MeOH	15.04	14.78
558	2,4-dimetilfenol <sup>a</sup>	H <sub>2</sub> O	10.60	10.19
559	2,4-dinitrofenol <sup>a</sup>	Ac	15.70	16.33
560	2,4-dinitrofenol <sup>a</sup>	AN	18.40	18.46

#	Àcid	Solvent	Exp.	SVM
561	2,4-dinitrofenol <sup>a</sup>	DMF	6.36	6.48
562	2,4-dinitrofenol <sup>a</sup>	DMSO	5.32	5.71
563	2,4-dinitrofenol <sup>b</sup>	H <sub>2</sub> O	4.10	5.72
564	2,4-dinitrofenol <sup>b</sup>	MeOH	7.82	9.55
565	2,4-dinitrofenol <sup>b</sup>	NM	15.90	17.26
566	2,4-di-t-butilfenol <sup>a</sup>	MeOH	16.77	16.71
567	2,4-di-t-butilfenol <sup>b</sup>	H <sub>2</sub> O	11.57	11.79
568	2,5-diclorofenol <sup>a</sup>	DMF	13.85	13.91
569	2,5-dimetilfenol <sup>a</sup>	MeOH	14.91	14.85
570	2,5-dimetilfenol <sup>b</sup>	H <sub>2</sub> O	10.41	10.24
571	2,5-dinitrofenol <sup>a</sup>	Ac	19.80	17.21
572	2,5-dinitrofenol <sup>a</sup>	DMF	8.78	7.46
573	2,5-dinitrofenol <sup>a</sup>	H <sub>2</sub> O	5.22	5.02
574	2,5-dinitrofenol <sup>a</sup>	NM	17.90	17.84
575	2,5-dinitrofenol <sup>a</sup>	DMSO	7.32	6.71
576	2,5-dinitrofenol <sup>b</sup>	MeOH	8.93	8.75
577	2,6-dibromo-4-nitrofenol <sup>a</sup>	DMF	5.71	5.77
578	2,6-dibromo-4-nitrofenol <sup>a</sup>	DMSO	5.17	5.11
579	2,6-dibromo-4-nitrofenol <sup>a</sup>	H <sub>2</sub> O	3.38	3.44
580	2,6-dibromo-4-nitrofenol <sup>a</sup>	MeOH	7.31	7.37
581	2,6-dicloro-4-hidroxifenol <sup>b</sup>	H <sub>2</sub> O	7.38	7.33
582	2,6-dicloro-4-nitrofenol <sup>a</sup>	DMF	5.72	5.80
583	2,6-dicloro-4-nitrofenol <sup>a</sup>	MeOH	7.40	7.34
584	2,6-dicloro-4-nitrofenol <sup>b</sup>	H <sub>2</sub> O	3.55	3.28
585	2,6-diclorofenol <sup>a</sup>	<sup>i</sup> PrOH	13.58	13.52
586	2,6-diclorofenol <sup>a</sup>	Ac	23.90	22.82
587	2,6-diclorofenol <sup>a</sup>	DMF	12.55	12.30
588	2,6-diclorofenol <sup>a</sup>	DMSO	11.54	10.70
589	2,6-diclorofenol <sup>a</sup>	H <sub>2</sub> O	6.79	6.85
590	2,6-diclorofenol <sup>a</sup>	<sup>i</sup> BuOH	16.38	16.32
591	2,6-dimetil-4-cianofenol <sup>b</sup>	H <sub>2</sub> O	8.27	9.86
592	2,6-dimetil-4-nitrofenol <sup>a</sup>	DMSO	9.00	7.84
593	2,6-dimetil-4-nitrofenol <sup>a</sup>	H <sub>2</sub> O	7.19	7.35
594	2,6-dimetilfenol <sup>a</sup>	H <sub>2</sub> O	10.59	10.17
595	2,6-dimetilfenol <sup>b</sup>	MeOH	15.26	14.77
596	2,6-dinitro-4-hidroxifenol <sup>a</sup>	H <sub>2</sub> O	4.42	3.48
597	2,6-dinitrofenol <sup>a</sup>	Ac	13.78	13.72
598	2,6-dinitrofenol <sup>a</sup>	DMF	6.07	6.01
599	2,6-dinitrofenol <sup>a</sup>	H <sub>2</sub> O	3.74	3.68
600	2,6-dinitrofenol <sup>a</sup>	MeOH	7.70	7.47
601	2,6-dinitrofenol <sup>a</sup>	NM	16.00	14.37
602	2,6-dinitrofenol <sup>b</sup>	AN	16.45	16.11
603	2,6-dinitrofenol <sup>b</sup>	DMSO	4.82	3.96
604	2,6-di-t-butil-4-nitrofenol <sup>a</sup>	AN	19.10	19.04
605	2,6-di-t-butil-4-nitrofenol <sup>a</sup>	DMF	8.27	8.33
606	2,6-di-t-butil-4-nitrofenol <sup>a</sup>	MeOH	10.89	10.83
607	2,6-di-t-butil-4-nitrofenol <sup>a</sup>	DMSO	7.60	7.54
608	2,6-di-t-butil-4-nitrofenol <sup>b</sup>	H <sub>2</sub> O	6.62	7.80
609	2,6-di-t-butilfenol <sup>b</sup>	DMSO	16.85	15.22
610	2-aminofenol <sup>a</sup>	H <sub>2</sub> O	9.44	9.50
611	2-bromofenol <sup>a</sup>	AN	23.92	23.86



#	Àcid	Solvent	Exp.	SVM
612	2-bromofenol <sup>a</sup>	H <sub>2</sub> O	8.39	7.54
613	2-bromofenol <sup>b</sup>	DMF	13.85	13.11
614	2-cloro-4-bromofenol <sup>a</sup>	MeOH	12.70	12.02
615	2-cloro-4-bromofenol <sup>b</sup>	H <sub>2</sub> O	7.64	7.38
616	2-cloro-4-fenilfenol <sup>a</sup>	<sup>t</sup> PrOH	13.70	14.59
617	2-cloro-4-fenilfenol <sup>a</sup>	AN	24.90	24.84
618	2-cloro-4-fenilfenol <sup>a</sup>	DMF	15.70	14.82
619	2-cloro-4-fenilfenol <sup>a</sup>	DMSO	14.90	14.49
620	2-cloro-4-fenilfenol <sup>a</sup>	<sup>t</sup> BuOH	18.60	18.21
621	2-cloro-4-fenilfenol <sup>b</sup>	H <sub>2</sub> O	8.07	8.09
622	2-clorofenol <sup>a</sup>	<sup>t</sup> PrOH	15.83	14.19
623	2-clorofenol <sup>a</sup>	<sup>t</sup> BuOH	18.54	17.65
624	2-clorofenol <sup>a</sup>	DMA	13.91	13.85
625	2-clorofenol <sup>b</sup>	H <sub>2</sub> O	8.51	7.69
626	2-clorofenol <sup>b</sup>	MeOH	12.83	12.28
627	2-etilfenol <sup>a</sup>	H <sub>2</sub> O	10.20	10.26
628	2-fluorofenol <sup>a</sup>	H <sub>2</sub> O	8.73	8.59
629	2-fluorofenol <sup>a</sup>	MeOH	12.14	13.25
630	2-hidroxi-3-nitrofenol <sup>a</sup>	H <sub>2</sub> O	6.68	6.74
631	2-hidroxifenol <sup>a</sup>	H <sub>2</sub> O	9.12	8.57
632	2-iodofenol <sup>a</sup>	H <sub>2</sub> O	8.46	8.52
633	2-metil-4,6-dinitrofenol <sup>a</sup>	DMSO	4.59	4.65
634	2-metilfenol <sup>a</sup>	AN	27.50	26.30
635	2-metilfenol <sup>a</sup>	H <sub>2</sub> O	10.31	10.00
636	2-metilfenol <sup>b</sup>	MeOH	14.90	14.44
637	2-metoxifenol <sup>a</sup>	H <sub>2</sub> O	9.90	9.64
638	2-metoxifenol <sup>a</sup>	MeOH	14.48	14.42
639	2-nitrofenol <sup>a</sup>	<sup>t</sup> PrOH	13.30	13.24
640	2-nitrofenol <sup>a</sup>	DMSO	11.00	10.94
641	2-nitrofenol <sup>a</sup>	H <sub>2</sub> O	7.23	7.29
642	2-nitrofenol <sup>a</sup>	MeOH	11.52	11.46
643	2-nitrofenol <sup>a</sup>	NM	21.40	21.46
644	2-nitrofenol <sup>a</sup>	<sup>t</sup> BuOH	15.88	15.82
645	2-nitrofenol <sup>a</sup>	AN	22.00	21.65
646	2-nitrofenol <sup>a</sup>	DMF	12.14	12.08
647	2-nitrofenol <sup>b</sup>	Ac	22.30	20.67
648	2-t-butilfenol <sup>a</sup>	H <sub>2</sub> O	11.34	10.89
649	2-t-butilfenol <sup>a</sup>	MeOH	16.50	15.71
650	3,4,5-triclorofenol <sup>a</sup>	DMSO	12.58	13.92
651	3,4,5-triclorofenol <sup>a</sup>	H <sub>2</sub> O	7.68	7.85
652	3,4-diclorofenol <sup>a</sup>	AN	24.06	25.40
653	3,4-diclorofenol <sup>a</sup>	H <sub>2</sub> O	8.51	8.74
654	3,4-diclorofenol <sup>b</sup>	DMF	13.22	14.92
655	3,4-diclorofenol <sup>b</sup>	DMSO	14.22	15.21
656	3,4-dimetilfenol <sup>a</sup>	H <sub>2</sub> O	10.36	10.21
657	3,4-dimetilfenol <sup>a</sup>	MeOH	14.63	14.75
658	3,4-dinitrofenol <sup>a</sup>	AN	17.90	18.66
659	3,4-dinitrofenol <sup>a</sup>	DMSO	7.97	8.03
660	3,4-dinitrofenol <sup>a</sup>	MeOH	9.46	9.52
661	3,4-dinitrofenol <sup>b</sup>	H <sub>2</sub> O	5.42	6.58
662	3,5-diclorofenol <sup>a</sup>	<sup>t</sup> PrOH	14.05	14.62

#	Àcid	Solvent	Exp.	SVM
663	3,5-diclorofenol <sup>a</sup>	DMSO	13.09	14.82
664	3,5-diclorofenol <sup>a</sup>	H <sub>2</sub> O	8.18	8.39
665	3,5-diclorofenol <sup>a</sup>	MeOH	12.94	13.00
666	3,5-diclorofenol <sup>a</sup>	<sup>t</sup> BuOH	17.04	18.17
667	3,5-diclorofenol <sup>a</sup>	AN	23.31	25.02
668	3,5-dihidroxifenol <sup>a</sup>	H <sub>2</sub> O	8.45	10.40
669	3,5-dimetil-4-cianofenol <sup>a</sup>	H <sub>2</sub> O	8.21	9.76
670	3,5-dimetil-4-nitrofenol <sup>a</sup>	H <sub>2</sub> O	8.25	7.78
671	3,5-dimetilfenol <sup>a</sup>	H <sub>2</sub> O	10.20	10.20
672	3,5-dimetilfenol <sup>a</sup>	MeOH	14.62	14.73
673	3,5-dinitrofenol <sup>a</sup>	<sup>i</sup> PrOH	10.84	11.17
674	3,5-dinitrofenol <sup>a</sup>	Ac	22.70	19.66
675	3,5-dinitrofenol <sup>a</sup>	AN	20.50	20.44
676	3,5-dinitrofenol <sup>a</sup>	<sup>t</sup> BuOH	13.40	13.46
677	3,5-dinitrofenol <sup>a</sup>	H <sub>2</sub> O	6.66	6.72
678	3,5-dinitrofenol <sup>b</sup>	DMF	11.30	10.26
679	3,5-dinitrofenol <sup>b</sup>	DMSO	10.60	9.93
680	3,5-dinitrofenol <sup>b</sup>	MeOH	10.20	10.25
681	3,5-dinitrofenol <sup>b</sup>	NM	19.40	20.14
682	3,5-di-t-butilfenol <sup>a</sup>	MeOH	14.89	15.92
683	3,5-di-t-butilfenol <sup>b</sup>	H <sub>2</sub> O	10.29	11.13
684	3-acetilfenol <sup>a</sup>	DMSO	15.14	15.08
685	3-acetilfenol <sup>b</sup>	H <sub>2</sub> O	9.19	8.43
686	3-aminofenol <sup>a</sup>	H <sub>2</sub> O	9.99	9.97
687	3-bromofenol <sup>a</sup>	<sup>i</sup> PrOH	14.83	14.77
688	3-bromofenol <sup>a</sup>	H <sub>2</sub> O	9.01	8.95
689	3-bromofenol <sup>a</sup>	MeOH	13.30	13.50
690	3-bromofenol <sup>b</sup>	<sup>t</sup> BuOH	18.52	17.96
691	3-cianofenol <sup>b</sup>	DMSO	14.76	15.58
692	3-cloro-2,4,6-trinitrofenol <sup>a</sup>	DMSO	1.16	1.22
693	3-cloro-4-nitrofenol <sup>a</sup>	AN	19.95	20.63
694	3-cloro-4-nitrofenol <sup>a</sup>	DMSO	9.80	11.10
695	3-cloro-4-nitrofenol <sup>a</sup>	H <sub>2</sub> O	6.49	7.52
696	3-clorofenol <sup>a</sup>	AN	25.04	25.23
697	3-clorofenol <sup>a</sup>	DMA	16.29	15.56
698	3-clorofenol <sup>a</sup>	DMSO	15.83	15.58
699	3-clorofenol <sup>a</sup>	MeOH	13.10	13.23
700	3-clorofenol <sup>b</sup>	H <sub>2</sub> O	9.02	8.81
701	3-etilfenol <sup>b</sup>	H <sub>2</sub> O	9.90	10.17
702	3-fluorofenol <sup>a</sup>	DMSO	15.88	15.94
703	3-fluorofenol <sup>a</sup>	H <sub>2</sub> O	9.28	9.22
704	3-hidroxifenol <sup>a</sup>	DMSO	15.30	16.20
705	3-hidroxifenol <sup>a</sup>	H <sub>2</sub> O	9.15	9.68
706	3-iodofenol <sup>a</sup>	H <sub>2</sub> O	8.88	9.20
707	3-metil-5-etilfenol <sup>a</sup>	H <sub>2</sub> O	10.10	10.43
708	3-metilfenol <sup>a</sup>	DMSO	16.86	16.61
709	3-metilfenol <sup>a</sup>	H <sub>2</sub> O	10.10	10.10
710	3-metilfenol <sup>a</sup>	MeOH	14.48	14.57
711	3-metoxifenol <sup>a</sup>	H <sub>2</sub> O	9.65	10.11
712	3-metoxifenol <sup>b</sup>	DMSO	15.72	16.76
713	3-nitrofenol <sup>a</sup>	<sup>i</sup> PrOH	13.92	13.86

#	Àcid	Solvent	Exp.	SVM
714	3-nitrofenol <sup>a</sup>	DMF	13.85	13.91
715	3-nitrofenol <sup>a</sup>	DMSO	13.75	13.69
716	3-nitrofenol <sup>a</sup>	H <sub>2</sub> O	8.36	8.30
717	3-nitrofenol <sup>a</sup>	MeOH	12.40	11.64
718	3-nitrofenol <sup>a</sup>	NM	22.20	22.26
719	3-nitrofenol <sup>a</sup>	<sup>t</sup> BuOH	16.99	16.92
720	3-nitrofenol <sup>b</sup>	AN	23.85	21.80
721	3-trifluorometil-4-nitrofenol <sup>a</sup>	<sup>i</sup> PrOH	9.90	9.96
722	3-trifluorometil-4-nitrofenol <sup>a</sup>	AN	19.30	19.36
723	3-trifluorometil-4-nitrofenol <sup>a</sup>	DMSO	9.30	9.36
724	3-trifluorometil-4-nitrofenol <sup>a</sup>	H <sub>2</sub> O	6.41	6.35
725	3-trifluorometil-4-nitrofenol <sup>a</sup>	<sup>t</sup> BuOH	12.77	12.54
726	3-trifluorometil-4-nitrofenol <sup>a</sup>	DMF	10.40	10.46
727	3-trifluorometilfenol <sup>a</sup>	<sup>i</sup> PrOH	12.50	12.56
728	3-trifluorometilfenol <sup>a</sup>	DMF	15.70	15.64
729	3-trifluorometilfenol <sup>a</sup>	DMSO	14.30	14.36
730	3-trifluorometilfenol <sup>a</sup>	H <sub>2</sub> O	9.04	8.98
731	3-trifluorometilfenol <sup>a</sup>	MeOH	12.10	12.16
732	3-trifluorometilfenol <sup>b</sup>	AN	24.90	25.56
733	3-trifluorometilfenol <sup>b</sup>	<sup>t</sup> BuOH	17.10	18.79
734	4-acetilfenol <sup>a</sup>	H <sub>2</sub> O	8.05	8.11
735	4-acetilfenol <sup>b</sup>	DMSO	13.68	14.34
736	4-aminofenol <sup>a</sup>	H <sub>2</sub> O	10.43	10.01
737	4-bromofenol <sup>a</sup>	<sup>i</sup> PrOH	14.30	14.98
738	4-bromofenol <sup>a</sup>	AN	25.53	25.08
739	4-bromofenol <sup>a</sup>	DMF	14.34	14.40
740	4-bromofenol <sup>a</sup>	MeOH	13.63	13.57
741	4-bromofenol <sup>a</sup>	<sup>t</sup> BuOH	19.10	18.27
742	4-bromofenol <sup>b</sup>	DMSO	15.50	15.34
743	4-bromofenol <sup>b</sup>	H <sub>2</sub> O	9.36	9.12
744	4-cianofenol <sup>a</sup>	AN	22.77	25.36
745	4-cianofenol <sup>a</sup>	DMSO	13.01	15.30
746	4-cianofenol <sup>a</sup>	H <sub>2</sub> O	7.80	9.35
747	4-cloro-2,6-dinitrofenol <sup>a</sup>	AN	15.30	15.24
748	4-cloro-2,6-dinitrofenol <sup>a</sup>	DMF	4.68	5.17
749	4-cloro-2,6-dinitrofenol <sup>a</sup>	DMSO	3.51	3.25
750	4-cloro-2,6-dinitrofenol <sup>a</sup>	H <sub>2</sub> O	2.97	2.91
751	4-clorofenol <sup>a</sup>	<sup>i</sup> PrOH	15.31	15.35
752	4-clorofenol <sup>a</sup>	AN	25.44	26.25
753	4-clorofenol <sup>a</sup>	DMSO	16.10	16.04
754	4-clorofenol <sup>a</sup>	H <sub>2</sub> O	9.38	9.66
755	4-clorofenol <sup>a</sup>	MeOH	13.59	14.30
756	4-clorofenol <sup>a</sup>	DMF	14.50	15.57
757	4-clorofenol <sup>b</sup>	DMA	16.78	16.02
758	4-clorofenol <sup>b</sup>	<sup>t</sup> BuOH	18.96	18.85
759	4-etilfenol <sup>a</sup>	H <sub>2</sub> O	10.00	10.18
760	4-fluorofenol <sup>a</sup>	H <sub>2</sub> O	9.95	10.00
761	4-hidroxi-2,3,5,6-tetrametilfenol <sup>a</sup>	H <sub>2</sub> O	11.51	11.45
762	4-hidroxi-2-metilfenol <sup>a</sup>	H <sub>2</sub> O	10.20	10.51
763	4-hidroxi-2-nitrofenol <sup>a</sup>	H <sub>2</sub> O	7.63	7.57
764	4-hidroxifenol <sup>a</sup>	H <sub>2</sub> O	9.14	9.75

#	Àcid	Solvent	Exp.	SVM
765	4-iodofenol <sup>a</sup>	DMF	14.31	14.26
766	4-iodofenol <sup>a</sup>	H <sub>2</sub> O	9.20	9.21
767	4-metilfenol <sup>a</sup>	AN	27.45	26.49
768	4-metilfenol <sup>a</sup>	DMSO	16.96	17.01
769	4-metilfenol <sup>a</sup>	H <sub>2</sub> O	10.28	10.02
770	4-metilfenol <sup>b</sup>	MeOH	14.52	14.41
771	4-metoxifenol <sup>a</sup>	H <sub>2</sub> O	10.27	10.02
772	4-metoxifenol <sup>a</sup>	DMSO	17.58	17.52
773	4-nitrofenol <sup>a</sup>	DMA	10.85	11.05
774	4-nitrofenol <sup>a</sup>	DMF	11.84	11.47
775	4-nitrofenol <sup>a</sup>	DMSO	11.00	11.06
776	4-nitrofenol <sup>a</sup>	MeOH	11.24	11.30
777	4-nitrofenol <sup>a</sup>	NM	20.10	20.04
778	4-nitrofenol <sup>a</sup>	<sup>t</sup> BuOH	14.60	14.66
779	4-nitrofenol <sup>a</sup>	<sup>i</sup> PrOH	12.45	12.39
780	4-nitrofenol <sup>b</sup>	AN	20.70	20.15
781	4-nitrofenol <sup>b</sup>	H <sub>2</sub> O	7.18	8.01
782	4-t-butilfenol <sup>a</sup>	AN	27.48	27.42
783	4-t-butilfenol <sup>a</sup>	H <sub>2</sub> O	10.31	10.51
784	4-t-butilfenol <sup>a</sup>	MeOH	14.52	15.16
785	fenol <sup>a</sup>	DMF	15.40	15.46
786	fenol <sup>a</sup>	DMSO	16.47	16.41
787	fenol <sup>a</sup>	H <sub>2</sub> O	9.99	9.87
788	fenol <sup>a</sup>	MeOH	14.32	14.24
789	fenol <sup>a</sup>	NM	25.70	25.64
790	fenol <sup>b</sup>	AN	26.60	25.95
791	pentaclorofenol <sup>a</sup>	DMF	7.97	8.91
792	pentaclorofenol <sup>a</sup>	DMSO	7.05	6.99
793	pentaclorofenol <sup>b</sup>	Ac	18.30	19.71

<sup>a</sup> Training set. <sup>b</sup> Prediction set.

#### 4.4. $pK_a$ d'àcids carboxílics alifàtics

Els àcids carboxílics alifàtics constitueixen la tercera família de compostos en què s'ha aplicat la metodologia QSPR al càlcul de les constants de dissociació. Aquesta família és bastant diferent a les que s'han estudiat anteriorment ja que, en aquest cas, no contenen l'anell aromàtic present en els fenols i àcids benzoics. De tota manera, aquesta família d'àcids ha estat molt estudiada i es disposa de suficient informació experimental que permet l'aplicació de la metodologia QSPR.

##### *Conjunt de dades*

La base de dades creada per a l'anàlisi d'aquesta família de compostos conté 158 valors de  $pK_a$  experimental; que s'obtenen de la combinació de 43 àcids carboxílics en 11 solvents:  $H_2O$ , MeOH, EtOH,  $i$ PrOH,  $t$ BuOH, Ac, AN, DMSO, DMF, DMA i NM. Cal dir que, d'entre tots els àcids carboxílics que existeixen, només s'han escollit aquells que presenten una sola constant de dissociació; és a dir, els àcids carboxílics monoprotics. Els valors de  $pK_a$  en aigua i alcohols s'han obtingut de publicacions de Bosch,<sup>178-182</sup> Kolthoff,<sup>193,194</sup> Headley<sup>214</sup> i alguns altres;<sup>215-219</sup> mentre que els valors de  $pK_a$  en solvents apròtics s'han obtingut de treballs de Bartnicka<sup>220</sup> i de la compilació d'Izutsku.<sup>183</sup>

Els àcids carboxílics triats es caracteritzen per la llargada de la seva cadena, pels substituents que presenten i per la seva posició. Per tal de mantenir la màxima compacitat dintre del conjunt de dades s'han utilitzat els àcids carboxílics amb llargada de cadena d'entre dos i sis carbonis. Els substituents que presenten aquests compostos són força variats: halògens, hidroxils, cianurs, grups orgànics com el fenil o el vinil, etc. El nombre de valors de  $pK_a$  en cada solvent és variable, així es tenen 43 valors en aigua, 37 en metanol, 11 en etanol, 11 en isopropanol, 6 en *tert*-butanol, 12 en DMSO, 12 en acetonitril, 9 en nitrometà, 7 en DMF, 5 en acetona i 5 en DMA. És de sobra conegut que hi ha molts més valors disponibles de  $pK_a$  per aquest tipus de compostos en aigua però no s'han afegit a la base de dades per tal de no esbiaixar-la en favor d'aquest solvent ja que, en cas contrari, s'obtindrien models que funcionarien bé per a aquest solvent però fallarien per a la resta. D'aquesta manera, es tenen 108 valors de  $pK_a$  en solvents pròtics i només 50 en solvents apròtics. Tot i així, el rang de  $pK_a$  que cobreixen aquests compostos en els onze solvents és més gran de vint unitats: des de 0.7 per a



l'àcid tricloroacètic en aigua fins a 22.73 de l'àcid butanoic en acetonitril. La Taula 4.35 conté els valors experimentals i calculats de  $pK_a$  per aquesta família de compostos. De la mateixa manera que anteriorment, en cas que un àcid pugui presentar més d'una constant de dissociació, el valor de  $pK_a$  escollit correspon a la primera.

La divisió del conjunt de dades en *tset* i *pset* proporciona subconjunts de 130 i 28 compostos, respectivament. Aquests dos subconjunts s'empraran en l'establiment del model amb SVM. Per tal de realitzar l'estudi amb les xarxes neuronals implementades en el programa ADAPT es necessita construir el *cvset* amb aproximadament un 10% del *tset*. Així, en l'estudi BPNN, els tres sets *tset*, *cvset* i *pset* constaran de 114, 16 i 28 valors de  $pK_a$ , respectivament.

#### ***Model no lineal: Xarxes neuronals BPNN***

El nombre de descriptors inicial que calcula CODESSA per als àcids carboxílics és de 643. La importació dels descriptors dels solvents eleva aquest nombre inicial fins a 680. El mètode heurístic implementat per CODESSA fa la primera reducció en el nombre de descriptors, passant dels 680 inicials a 171, 153 dels àcids carboxílics i 18 dels solvents. El programa ADAPT realitza la segona reducció en el nombre de descriptors, així el nombre definitiu de descriptors és de 99, 88 dels àcids carboxílics i 11 dels solvents. L'arquitectura triada per les xarxes neuronals és 7-5-1. El nombre de paràmetres ajustables corresponents a aquesta arquitectura és 46, amb la qual cosa el *ratio* del nombre de compostos del *tset*, 114, sobre el nombre de paràmetres ajustables és de 2.48, per sobre del mínim 2.0 acceptat.

El model trobat amb les xarxes neuronals d'arquitectura 7-5-1 conté els descriptors que es mostren en la Taula 4.36. Com en les altres famílies de compostos estudiades, el model presenta cinc descriptors dels soluts i dos descriptors dels solvents. En aquest cas els descriptors dels àcids carboxílics pertanyen a tres classes diferents: un topològic, dos electrostàtics i dos quàntics.

Taula 4.36. Descriptors continguts en el model dels àcids carboxílics amb BPNN.

Descriptor	Component	Classe
Contingut d'informació mitjà d'ordre 0	Solut	Topològic
Càrrega parcial màxima d'un àtom d'hidrogen	Solut	Electrostàtic
Àrea superficial negativa fraccionada, FNSA-3	Solut	Electrostàtic
Moment Dipolar	Solut	Quàntic
Repulsió electró-electró mínima d'un enllaç H-O	Solut	Quàntic
Capacitat de donació d'enllaç d'hidrogen de Kamlet i Taft, $\alpha$	Solvent	-
Energia de vaporització molar, $\Delta_{\text{vap}}U^\circ$	Solvent	-

El descriptor topològic: contingut d'informació mitjà d'ordre 0, es deriva de la teoria d'informació de Shannon<sup>221,222</sup> i es calcula de la següent manera:

$${}^0\overline{IC} = -\sum_i \frac{n_i}{n} \log_2 \frac{n_i}{n}$$

on  $n_i$  és el nombre d'àtoms de cada element present en la molècula i  $n$  és el nombre total d'àtoms. Aquest descriptor, també anomenat Entropia de Shannon, es pot considerar com una mesura de la diversitat atòmica dintre de la molècula i es relaciona amb la simetria associada a l'estructura.

Els descriptors electrostàtics, que es calculen a partir de la distribució de càrregues parcials de Zefirov,<sup>73</sup> són la càrrega parcial màxima sobre un àtom d'hidrogen i l'àrea superficial negativa fraccionada, FNSA-3. Com en el cas dels fenols i els àcids benzoics, la càrrega parcial màxima sobre un àtom d'hidrogen se situa sempre sobre el protó del grup àcid, és a dir, aquell que es dissocia. El descriptor FNSA-3, de tipus CPSA, es calcula:

$$FNSA-3 = \frac{\sum_A q_A \cdot S_A}{TMSA}$$

on  $q_A$  són les càrregues atòmiques parcials negatives,  $S_A$  és la superfície total accessible al solvent carregada negativament i  $TMSA$  és la superfície total de la molècula. Els descriptors quàntics s'obtenen directament amb el càlcul semiempíric AM1 implementat en el programa MOPAC; aquests descriptors són el dipol total de la molècula i l'energia de repulsió electró-electró mínima d'un enllaç H-O. Aquests

descriptors semblen força adients per descriure la reacció de dissociació estudiada ja que el moment dipolar del solut es pot relacionar amb la capacitat d'interacció de l'àcid carboxílic amb el solvent i l'energia de repulsió electró-electró mínima de l'enllaç enllaç H-O se situa sempre sobre l'enllaç que es trenca en el procés de dissociació.

Els descriptors del solvent són els mateixos que s'han trobat abans per als àcids benzoics:  $\alpha$  de Kamlet i Taft<sup>160</sup> i l'energia de vaporització molar,  $\Delta_{\text{vap}}U^\circ$ . El descriptor  $\alpha$  mesura la capacitat de les molècules de solvent per establir enllaços donadors d'hidrogen amb els soluts. L'energia de vaporització molar del solvent,  $\Delta_{\text{vap}}U^\circ$ , s'anomena també energia de cohesió i mesura la capacitat del solvent per formar les cavitats en les quals s'han d'encabir els soluts.

Els resultats estadístics que s'aconsegueixen amb aquest model no lineal per als diferents subconjunts de valors: *tset*, *cvset* i *pset*, es troben en la Taula 4.37. Els valors de  $pK_a$  calculats es troben en la Taula 4.35 i la seva representació gràfica enfront dels valors experimentals es pot veure en la Figura 4.26.

Taula 4.37. Paràmetres estadístics obtinguts amb el model BPNN.

Conjunt	n	R <sup>2</sup>	F	rmse
<i>tset</i>	114	0.998	63087.9	0.19
<i>cvset</i>	16	0.999	11320.6	0.19
<i>pset</i>	28	0.998	10956.4	0.25

Com es pot comprovar, el model dóna també resultats bons i presenta, també, una robustesa molt elevada, amb valors de R<sup>2</sup> i *rmse* pràcticament iguals per als tres subconjunts de dades.

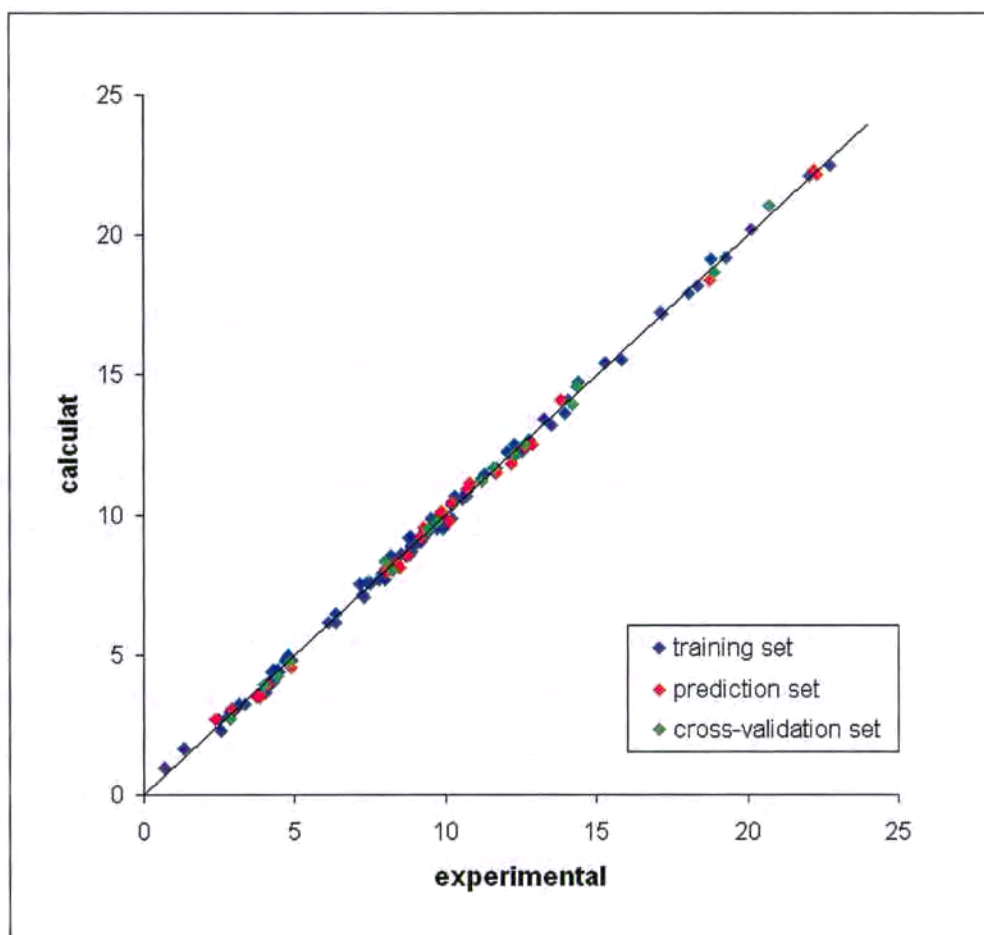


Figura 4.26. Representació gràfica de  $pK_a$  calculat vs. experimental per als àcids carboxílics amb BPNN.

Els descriptors inclosos en el model permeten, un cop més, interpretar la reacció de dissociació en funció de les complexes interaccions solut/solvent que es donen en el sistema. Així, els descriptors càrrega parcial màxima sobre un àtom d'hidrogen, moment dipolar total, FNSA-3 i energia de repulsió electró-electró mínima d'un enllaç H-O fan referència a les interaccions solut/solvent no específiques. El descriptor de solvent energia de vaporització molar ( $\Delta_{\text{vap}}U^\circ$ ) conté informació sobre aquest mateix tipus d'interaccions. Les interaccions de tipus específic (acceptor/donador) venen donades pel descriptor dels solvents  $\alpha$  de Kamlet i Taft. En canvi, l'últim descriptor del model: contingut d'informació mitjà d'ordre 0, és difícil d'interpretar en aquest sentit.

El descriptor càrrega parcial màxima sobre un àtom d'hidrogen se situa sempre sobre el protó carboxílic dels àcids, aquest descriptor reflecteix la polaritat de l'enllaç O-H que es trenca i mesura la facilitat amb la qual es trenca per donar el carboxilat

conjugat i un protó. Els valors més baixos d'aquest descriptor es troben en els àcids carboxílics de cadena més llarga i que presenten substituents alquilics; per contra, els àcids carboxílics que tenen substituents electroatracients presenten valors més alts per aquest descriptor. Un altre descriptor que reflecteix la polaritat de l'enllaç que es trenca és l'energia de repulsió electró-electró mínima d'un enllaç H-O, com ja s'ha comentat anteriorment, aquest descriptor se situa sempre sobre l'enllaç hidroxílic del grup àcid. Aquest descriptor també es pot relacionar amb els substituents, la posició de substitució i la llargada de la cadena, presentant els valor més baixos per compostos polihalogenats amb les cadenes carbonades més curtes. Els descriptors dipol total i FNSA-3 fan referència directa a la capacitat de l'àcid carboxílic per establir interaccions electrostàtiques i polars amb les molècules del solvent. El descriptor de solvent  $\Delta_{\text{vap}}U^\circ$  és una mesura de la capacitat del solvent per dissoldre les molècules de solut en les cavitats que es formen en el seu interior. Cal recordar que aquestes cavitats contenen tant les molècules neutres com els contraions que es formen en la reacció de dissociació. Les interaccions específiques, o acceptor/donador, es poden relacionar amb el descriptor dels solvents  $\alpha$  de Kamlet i Taft. Aquest descriptor mesura la capacitat del solvent per establir enllaços d'hidrogen per donació a les molècules de solut. Aquest descriptor és important en l'estabilització dels anions carboxilat que apareixen en la reacció de dissociació i, per tant, serà més important en solvents pròtics que en apròtics.

La importància relativa de cada descriptor dins del model s'ha avaluat amb el mètode de Jurs. Aquest mètode proposa que la permuta d'un descriptor condueix a un increment en el *rmse* proporcional a la importància del descriptor en el model. La taula 4.38 conté els increments de *rmse* que es donen quan es permuten individualment els descriptors.

 Taula 4.38. Increments de *rmse* per la permuta de cada descriptor.

Descriptor	Increment <i>rmse</i>	Importància
Càrrega parcial màxima d'un àtom d'hidrogen	1.24	3
Àrea superficial negativa fraccionada, FNSA-3	0.65	7
Contingut d'informació mitjà d'ordre 0	0.85	4
Moment Dipolar	0.78	6
Repulsió electró-electró mínima d'un enllaç H-O	0.85	5
Capacitat de donació d'enllaç d'hidrogen de Kamlet i Taft, $\alpha$	6.49	1
Energia de vaporització molar, $\Delta_{\text{vap}}U^\circ$	3.79	2



D'acord amb aquests increments de *rmse*, els més importants són, per aquest ordre: la  $\alpha$  de Kamlet i Taft, l'energia de vaporització molar ( $\Delta_{\text{vap}}U^\circ$ ) i la càrrega parcial màxima sobre un àtom d'hidrogen.

El valor de *rmse* del model per a tots els compostos és de 0.20. No hi ha cap compost que presenti un valor de residual per sobre de dues vegades aquest valor de *rmse*, provant la bondat del model establert. També es pot comprovar la robustesa del model no lineal analitzant amb detall els resultats estadístics que s'obtenen amb els subconjunts de compostos que es poden formar a partir del conjunt inicial. La Taula 4.39 mostra tots aquests subconjunts i els paràmetres estadístics que s'obtenen per a cadascun d'ells.

Taula 4.39. Estadística dels subconjunts d'àcids carboxílics.

Subconjunt	n	R <sup>2</sup>	F	rmse
Pròtics	108	0.996	28852.7	0.19
Apròtics	50	0.998	20351.8	0.23
H <sub>2</sub> O	43	0.965	1133.2	0.17
MeOH	37	0.976	1446.5	0.17
EtOH	11	0.982	494.2	0.19
<sup>t</sup> PrOH	11	0.981	463.1	0.23
<sup>t</sup> BuOH	6	0.991	434.0	0.19
DMSO	12	0.985	672.2	0.28
DMF	7	0.992	588.6	0.24
AN	12	0.997	3119.3	0.21
Ac	5	0.997	1068.4	0.08
NM	9	0.992	902.4	0.24
DMA	5	0.997	968.5	0.17

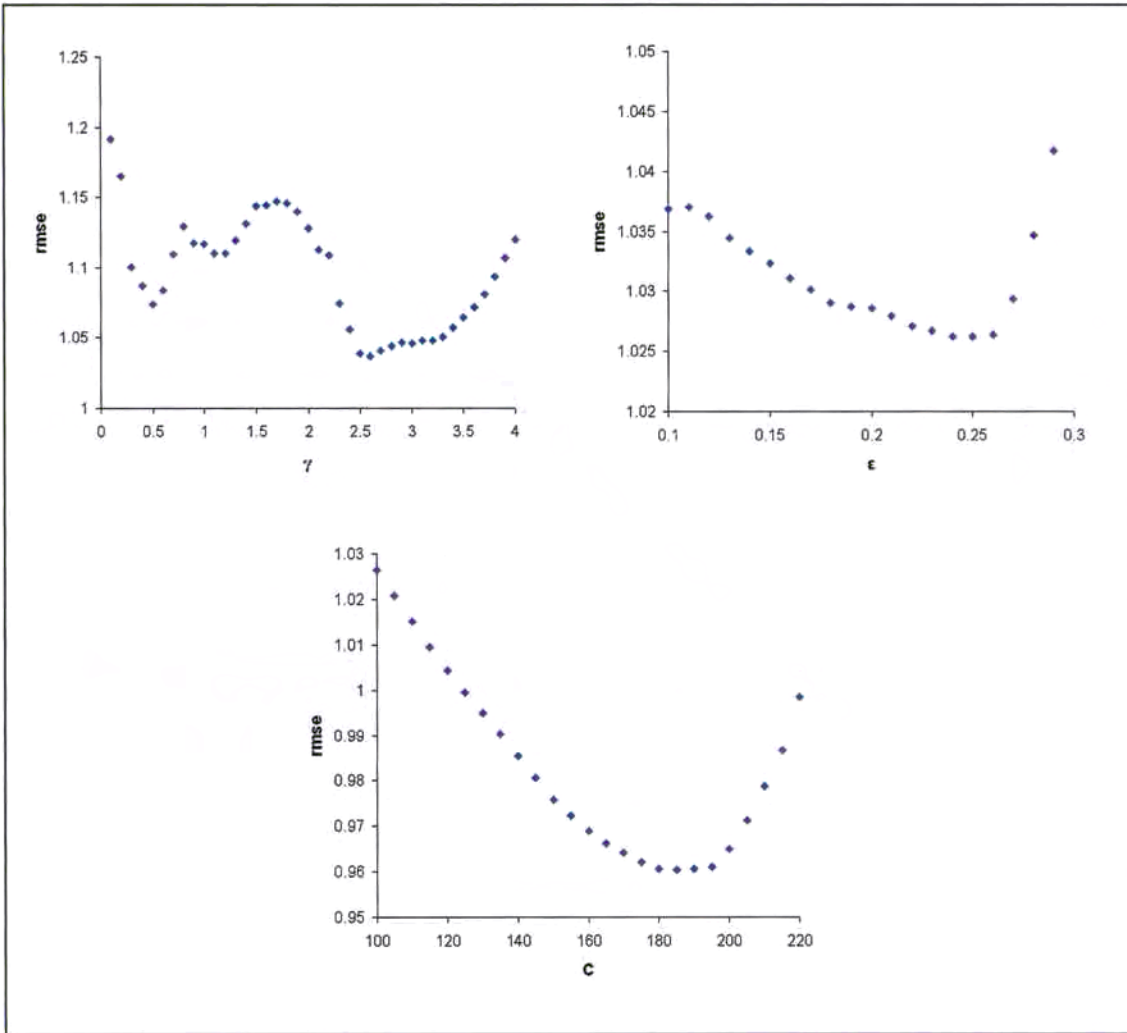
Com es pot observar, els resultats que s'obtenen amb aquests subconjunts són bons i molt semblants entre sí, donant proves de la robustesa adquirida pel model no lineal aquí exposat. Els resultats obtinguts amb el model no lineal són millors que els que s'obtenen en treballs anàlegs, si bé només en aigua com a solvent, ja publicats. Un dels primers intents va ser realitzat per Citra<sup>149</sup> amb un conjunt de 56 àcids carboxílics en aigua. Amb un model lineal de tres descriptors: càrregues parcials sobre els àtoms d'oxigen i hidrogen del grup carboxílic i ordre d'enllaç entre ells, s'obté un coeficient de determinació  $R^2=0.84$ . En un dels treballs de Tehan,<sup>148</sup> on s'ha construït un model lineal de dos descriptors: una autopolaritzabilitat atòmica i una superdeslocalització electrònica, per a 141 àcids carboxílics també en aigua, s'ha trobat un coeficient de

determinació  $R^2=0.80$  amb un  $rmse=0.39$ . Més recentment, Gasteiger i col·laboradors<sup>151</sup> van construir un model lineal amb cinc descriptors per a 1122 àcids carboxílics, també només en aigua. Els resultats que s'obtenen són:  $R^2=0.81$  i  $rmse=0.42$ , per sota dels que s'han obtingut amb el model no lineal aquí desenvolupat. Només en un treball s'obtenen resultats semblants als aquí obtinguts, en què s'utilitzen els models de Hammet-Taft i Drago<sup>156</sup> per predir el  $pK_a$  de 38 àcids carboxílics en metanol.

### **Model no lineal: Support Vector Machines**

El model desenvolupat amb *Support Vector Machines* conté els mateixos descriptors que el model de les xarxes neuronals. L'establiment del model es fa per optimització dels paràmetres  $\gamma$ ,  $C$  i  $\varepsilon$  amb el mètode de minimització de l'error de validació creuada per als compostos del *tset*, Figura 4.27. L'optimització del paràmetre  $\gamma$  (amb  $\varepsilon=0.1$  i  $C=100$ ) fixa el valor òptim d'aquest paràmetre en 2.6. L'optimització del paràmetre  $\varepsilon$  (amb  $\gamma=2.6$  i  $C=100$ ) troba que el seu millor valor és 0.24. El paràmetre  $C$  s'optimitza amb els valors dels altres paràmetres ja optimitzats:  $\gamma=2.6$  i  $\varepsilon=0.24$ . El valor òptim per aquest paràmetre és 185.

El model construït per LibSVM conté 82 vectors suport, triats d'entre els 130 que componen el *tset*. Els resultats que s'obtenen amb aquest mètode de càlcul són molt semblants als que s'obtenen amb el model BPNN, potser en aquest cas empitjoren una mica però no deixen de ser realment prou bons. La Taula 4.40 conté els valors del coeficient de determinació  $R^2$ , el test- $F$  i el  $rmse$  per als dos subconjunts de dades: *tset* i *pset*.


 Figura 4.27. Optimització dels paràmetres  $\gamma$ ,  $\epsilon$  i  $C$ .

Taula 4.40. Paràmetres estadístics obtinguts amb el model SVM.

Conjunt	n	$R^2$	F	rmse
<i>tset</i>	130	0.997	48787.7	0.24
<i>pset</i>	28	0.995	5524.7	0.35

La representació gràfica dels valors de  $pK_a$  calculat enfront dels experimentals es troba en la Figura 4.28. Tots els valors de  $pK_a$  calculat es troben a la Taula 4.35.

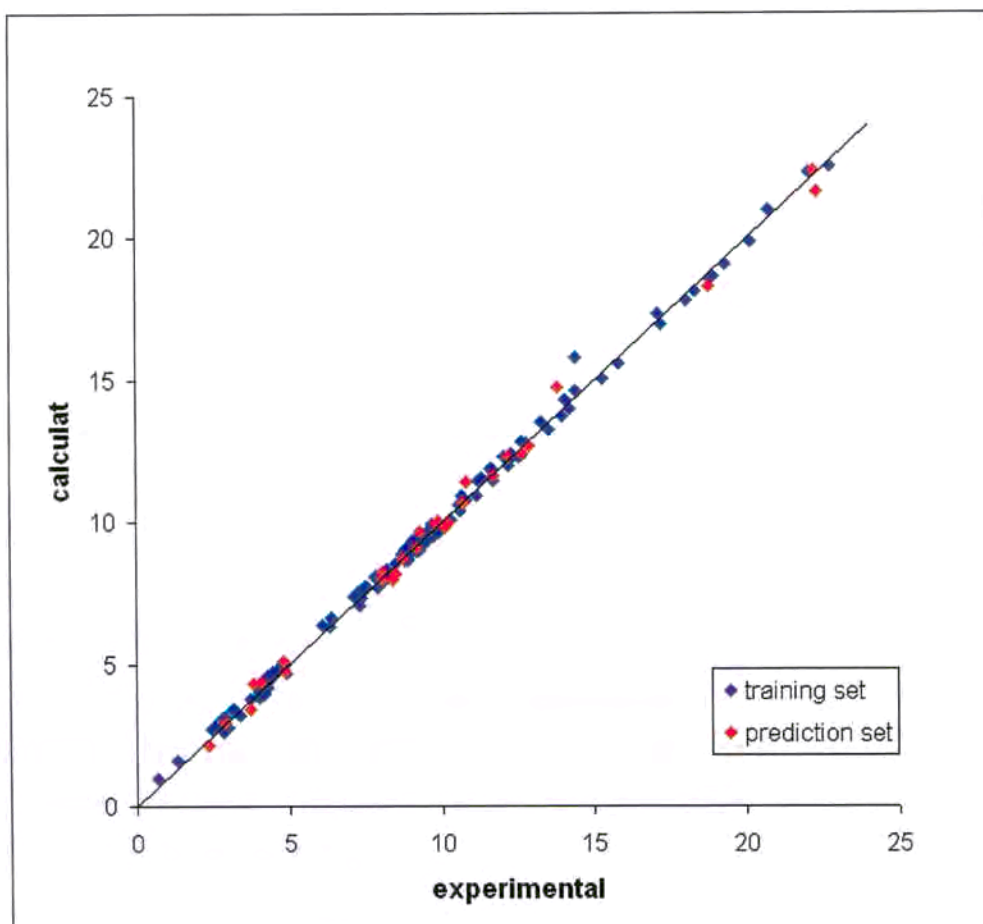


Figura 4.28. Representació gràfica de  $pK_a$  calculat vs. experimental per als àcids carboxílics amb SVM.

De nou, s'examinen els paràmetres estadístics que s'obtenen analitzant els subconjunts de solvents. La Taula 4.41 mostra aquests subconjunts de compostos i els paràmetres estadístics que s'obtenen per a tots ells amb aquest model.

Taula 4.41. Estadística dels subconjunts d'àcids carboxílics.

Subconjunt	n	$R^2$	F	rmse
Pròtics	108	0.996	23480.6	0.21
Apròtics	50	0.994	8568.8	0.34
H <sub>2</sub> O	43	0.954	855.2	0.21
MeOH	37	0.970	1127.2	0.20
EtOH	11	0.991	1025.4	0.13
iPrOH	11	0.977	390.2	0.25
iBuOH	6	1.000	71103.0	0.02
DMSO	12	0.991	1075.2	0.23
DMF	7	0.993	685.0	0.23
AN	12	0.993	1470.7	0.30
Ac	5	0.996	818.6	0.10
NM	9	0.979	334.2	0.42
DMA	5	0.994	533.9	0.22

Com es pot veure, encara que en la majoria dels solvents s'obtenen resultats lleugerament inferiors o bé iguals, en altres es donen millors resultats: isopropanol, DMSO, DMF. Menció especial requereix el cas del *tert*-butanol, pel qual s'obtenen uns resultats gairebé perfectes.

En aquesta ocasió, el valor de *rmse* del model amb tots els compostos és 0.26, una mica superior a l'obtingut amb BPNN. La representació gràfica dels valors calculats amb aquest model enfront dels valors experimentals mostra l'existència d'alguns *outliers*. La Taula 4.42 mostra els compostos que presenten un valor de residual superior a dues vegades el *rmse* del model.

Taula 4.42. Prediccions de  $pK_a$  amb valors més grans que dues vegades el *rmse*\*.

Solut	Solvent	Exper.	Calc.	residual
Àcid acètic	NM	14.41	15.76	-1.35
Àcid 2-fenilacètic	NM	13.80	14.75	-0.95
Àcid 2-fenilacètic	<sup>t</sup> PrOH	10.80	11.39	-0.59
Àcid acètic	AN	22.30	21.59	0.71

\* Els compostos ombrejats tenen residuals més grans que tres vegades el *rmse*

Com es pot comprovar, aquests quatre valors corresponen a dos àcids carboxílics: l'àcid acètic i l'àcid 2-fenilacètic en tres solvents diferents: nitrometà, isopropanol i acetonitril. Tot i així, no es pot dir que el model no funcioni per a aquests àcids carboxílics en concret ja que aquests dos compostos presenten 11 i 9 valors en diferents solvents, respectivament, i només amb dos de cadascun s'han detectat aquests problemes.



Taula 4.35. Valors experimentals i calculats de  $pK_a$  per a la família d'àcids carboxílics.

#	Àcid carboxílic	Solvent	Exp.	BPNN	SVM
1	Àcid 2-(3-clorofenil)acrílic <sup>a,c</sup>	H <sub>2</sub> O	4.29	4.06	4.13
2	Àcid 2-(3-clorofenil)acrílic <sup>a,c</sup>	MeOH	9.06	9.14	9.30
3	Àcid 2-(3-nitrofenil)acrílic <sup>a,c</sup>	H <sub>2</sub> O	4.12	3.90	4.36
4	Àcid 2-(3-nitrofenil)acrílic <sup>a,c</sup>	MeOH	9.89	9.58	9.65
5	Àcid 2-(4-nitrofenil)acrílic <sup>a,c</sup>	H <sub>2</sub> O	4.05	3.77	3.81
6	Àcid 2-(4-nitrofenil)acrílic <sup>a,c</sup>	MeOH	8.83	9.22	9.07
7	Àcid 2,2-dicloroacètic <sup>a,c</sup>	<sup>t</sup> PrOH	7.80	7.68	8.04
8	Àcid 2,2-dicloroacètic <sup>a,c</sup>	AN	15.81	15.57	15.57
9	Àcid 2,2-dicloroacètic <sup>a,c</sup>	DMA	6.10	6.16	6.34
10	Àcid 2,2-dicloroacètic <sup>a,c</sup>	DMF	7.90	7.81	7.66
11	Àcid 2,2-dicloroacètic <sup>a,c</sup>	DMSO	6.36	6.16	6.32
12	Àcid 2,2-dicloroacètic <sup>a,c</sup>	EtOH	7.30	7.06	7.06
13	Àcid 2,2-dicloroacètic <sup>a,c</sup>	H <sub>2</sub> O	1.34	1.67	1.58
14	Àcid 2,2-dicloroacètic <sup>a,c</sup>	MeOH	6.38	6.47	6.62
15	Àcid 2,2-dicloroacètic <sup>a,c</sup>	NM	8.88	9.21	9.12
16	Àcid 2,2-dicloroacètic <sup>a,c</sup>	<sup>t</sup> BuOH	10.20	10.44	9.96
17	Àcid 2,3-dibromopropanoic <sup>a,c</sup>	<sup>t</sup> PrOH	8.88	8.93	9.12
18	Àcid 2,3-dibromopropanoic <sup>a,c</sup>	AN	17.10	17.22	17.34
19	Àcid 2,3-dibromopropanoic <sup>a,c</sup>	DMSO	7.13	7.51	7.37
20	Àcid 2,3-dibromopropanoic <sup>a,c</sup>	MeOH	7.38	7.57	7.31
21	Àcid 2,3-dibromopropanoic <sup>a,c</sup>	<sup>t</sup> BuOH	11.70	11.70	11.46
22	Àcid 2,3-dibromopropanoic <sup>d</sup>	H <sub>2</sub> O	2.36	2.70	2.13
23	Àcid 2,3-dicloropropanoic <sup>a,c</sup>	MeOH	7.50	7.54	7.61
24	Àcid 2,3-dicloropropanoic <sup>b,c</sup>	H <sub>2</sub> O	2.85	2.69	2.61
25	Àcid 2-bromopropanoic <sup>a,c</sup>	H <sub>2</sub> O	3.00	3.01	2.76
26	Àcid 2-bromopropanoic <sup>b,c</sup>	MeOH	8.22	8.02	7.98
27	Àcid 2-clorobutanoic <sup>a,c</sup>	MeOH	8.11	8.03	7.90
28	Àcid 2-clorobutanoic <sup>d</sup>	H <sub>2</sub> O	2.92	3.03	2.89
29	Àcid 2-cloropropanoic <sup>a,c</sup>	H <sub>2</sub> O	2.90	3.08	3.14
30	Àcid 2-cloropropanoic <sup>d</sup>	MeOH	8.06	8.10	8.19
31	Àcid 2-fenilpropanoic <sup>a,c</sup>	MeOH	9.31	9.43	9.18
32	Àcid 2-fenilpropanoic <sup>b,c</sup>	H <sub>2</sub> O	4.44	4.26	4.68
33	Àcid 2-fluoroacrílic <sup>a,c</sup>	H <sub>2</sub> O	2.57	2.29	2.81
34	Àcid 2-propenoic <sup>a,c</sup>	H <sub>2</sub> O	4.25	4.40	4.14
35	Àcid 2-propenoic <sup>d</sup>	MeOH	9.27	9.57	9.62
36	Àcid 3-(2-furil)acrílic <sup>a,c</sup>	H <sub>2</sub> O	4.29	4.46	4.53
37	Àcid 3-bromobutanoic <sup>a,c</sup>	MeOH	9.12	9.16	9.03
38	Àcid 3-bromobutanoic <sup>b,c</sup>	H <sub>2</sub> O	4.01	3.92	3.81
39	Àcid 3-bromopropanoic <sup>a,c</sup>	H <sub>2</sub> O	4.04	3.88	3.98
40	Àcid 3-bromopropanoic <sup>a,c</sup>	MeOH	9.00	8.96	8.87
41	Àcid 3-clorobutanoic <sup>a,c</sup>	H <sub>2</sub> O	4.17	3.94	3.93
42	Àcid 3-clorobutanoic <sup>d</sup>	MeOH	9.18	9.16	9.04
43	Àcid 3-cloropropanoic <sup>a,c</sup>	MeOH	9.18	9.00	8.95
44	Àcid 3-cloropropanoic <sup>d</sup>	H <sub>2</sub> O	4.09	3.95	4.36
45	Àcid 3-fenilpropanoic <sup>a,c</sup>	H <sub>2</sub> O	4.45	4.46	4.69
46	Àcid 3-fenilpropanoic <sup>b,c</sup>	NM	14.37	14.60	14.61
47	Àcid 3-hidroxiopropanoic <sup>a,c</sup>	H <sub>2</sub> O	4.51	4.41	4.62
48	Àcid 3-hidroxiopropanoic <sup>b,c</sup>	MeOH	9.42	9.53	9.44
49	Àcid 3-iodopropanoic <sup>a,c</sup>	H <sub>2</sub> O	4.05	3.68	4.05
50	Àcid 3-iodopropanoic <sup>a,c</sup>	MeOH	8.89	8.83	8.67

#	Àcid carboxílic	Solvent	Exp.	BPNN	SVM
51	Àcid acètic <sup>a,c</sup>	<sup>t</sup> PrOH	11.30	11.49	11.54
52	Àcid acètic <sup>a,c</sup>	Ac	18.33	18.23	18.09
53	Àcid acètic <sup>a,c</sup>	DMF	13.25	13.45	13.49
54	Àcid acètic <sup>a,c</sup>	DMSO	12.30	12.51	12.35
55	Àcid acètic <sup>a,c</sup>	EtOH	10.30	10.64	10.06
56	Àcid acètic <sup>a,c</sup>	H <sub>2</sub> O	4.75	4.74	4.99
57	Àcid acètic <sup>a,c</sup>	MeOH	9.70	9.80	9.80
58	Àcid acètic <sup>a,c</sup>	NM	14.41	14.76	15.76
59	Àcid acètic <sup>b,c</sup>	DMA	12.60	12.51	12.36
60	Àcid acètic <sup>b,c</sup>	<sup>t</sup> BuOH	14.20	13.95	13.96
61	Àcid acètic <sup>d</sup>	AN	22.30	22.16	21.59
62	Àcid bromoacètic <sup>a,c</sup>	H <sub>2</sub> O	2.90	2.98	2.66
63	Àcid bromoacètic <sup>d</sup>	MeOH	8.06	7.96	7.95
64	Àcid bromoacètic <sup>d</sup>	NM	11.70	11.54	11.58
65	Àcid butanoic <sup>a,c</sup>	AN	22.73	22.53	22.49
66	Àcid butanoic <sup>a,c</sup>	EtOH	10.70	10.83	10.63
67	Àcid butanoic <sup>a,c</sup>	MeOH	9.68	9.89	9.92
68	Àcid butanoic <sup>b,c</sup>	H <sub>2</sub> O	4.82	4.73	4.80
69	Àcid butanoic <sup>d</sup>	<sup>t</sup> PrOH	12.20	11.86	12.27
70	Àcid butanoic <sup>d</sup>	DMSO	12.86	12.55	12.65
71	Àcid cianoacètic <sup>a,c</sup>	AN	18.04	17.95	17.80
72	Àcid cianoacètic <sup>a,c</sup>	H <sub>2</sub> O	2.46	2.69	2.70
73	Àcid cianoacètic <sup>a,c</sup>	MeOH	7.50	7.57	7.74
74	Àcid cianoacètic <sup>a,c</sup>	NM	10.68	10.64	10.92
75	Àcid cianoacètic <sup>b,c</sup>	EtOH	8.00	8.32	7.76
76	Àcid cianoacètic <sup>d</sup>	DMSO	8.50	8.11	8.15
77	Àcid cianoacètic <sup>d</sup>	<sup>t</sup> BuOH	12.60	12.41	12.40
78	Àcid cloroacètic <sup>a,c</sup>	<sup>t</sup> PrOH	9.20	9.20	9.44
79	Àcid cloroacètic <sup>a,c</sup>	AN	18.80	19.17	18.56
80	Àcid cloroacètic <sup>a,c</sup>	DMSO	8.90	8.70	8.66
81	Àcid cloroacètic <sup>a,c</sup>	EtOH	8.20	8.54	8.30
82	Àcid cloroacètic <sup>a,c</sup>	H <sub>2</sub> O	2.85	2.95	3.04
83	Àcid cloroacètic <sup>a,c</sup>	MeOH	7.88	7.91	8.12
84	Àcid cloroacètic <sup>a,c</sup>	NM	11.62	11.54	11.86
85	Àcid cloroacètic <sup>a,c</sup>	<sup>t</sup> BuOH	12.20	11.83	11.96
86	Àcid cloroacètic <sup>d</sup>	DMA	8.75	8.55	8.67
87	Àcid cloroacètic <sup>d</sup>	DMF	10.13	9.75	9.80
88	Àcid fenilacètic <sup>a,c</sup>	Ac	17.17	17.22	16.93
89	Àcid fenilacètic <sup>a,c</sup>	DMF	13.50	13.22	13.26
90	Àcid fenilacètic <sup>a,c</sup>	EtOH	10.20	10.34	9.96
91	Àcid fenilacètic <sup>a,c</sup>	H <sub>2</sub> O	4.31	4.20	4.55
92	Àcid fenilacètic <sup>a,c</sup>	MeOH	9.43	9.41	9.49
93	Àcid fenilacètic <sup>b,c</sup>	AN	20.73	21.10	20.95
94	Àcid fenilacètic <sup>b,c</sup>	DMSO	11.60	11.67	11.84
95	Àcid fenilacètic <sup>d</sup>	<sup>t</sup> PrOH	10.80	11.17	11.39
96	Àcid fenilacètic <sup>d</sup>	NM	13.80	14.13	14.75
97	Àcid fenoxiacètic <sup>a,c</sup>	<sup>t</sup> PrOH	9.50	9.85	9.26
98	Àcid fenoxiacètic <sup>a,c</sup>	Ac	15.28	15.42	15.04
99	Àcid fenoxiacètic <sup>a,c</sup>	AN	20.12	20.22	19.88
100	Àcid fenoxiacètic <sup>a,c</sup>	DMSO	9.71	9.52	9.47
101	Àcid fenoxiacètic <sup>a,c</sup>	H <sub>2</sub> O	3.17	3.24	3.41

#	Àcid carboxílic	Solvent	Exp.	BPNN	SVM
102	Àcid fenoxiacètic <sup>a,c</sup>	NM	12.60	12.35	12.84
103	Àcid fenoxiacètic <sup>b,c</sup>	DMF	11.17	11.21	10.93
104	Àcid fluoroacètic <sup>a,c</sup>	H <sub>2</sub> O	2.82	2.81	3.06
105	Àcid fluoroacètic <sup>a,c</sup>	MeOH	7.99	7.69	7.84
106	Àcid hexanoic <sup>a,c</sup>	MeOH	9.90	9.52	9.66
107	Àcid hexanoic <sup>d</sup>	H <sub>2</sub> O	4.83	4.60	5.10
108	Àcid hidroxiacètic <sup>a,c</sup>	AN	19.30	19.22	19.06
109	Àcid hidroxiacètic <sup>a,c</sup>	DMA	9.85	9.86	9.97
110	Àcid hidroxiacètic <sup>a,c</sup>	DMSO	10.20	9.86	9.96
111	Àcid hidroxiacètic <sup>a,c</sup>	EtOH	9.27	9.17	9.03
112	Àcid hidroxiacètic <sup>a,c</sup>	MeOH	8.68	8.55	8.83
113	Àcid hidroxiacètic <sup>d</sup>	H <sub>2</sub> O	3.85	3.52	4.28
114	Àcid iodoacètic <sup>a,c</sup>	H <sub>2</sub> O	3.13	3.15	3.37
115	Àcid iodoacètic <sup>a,c</sup>	MeOH	8.38	8.24	8.16
116	Àcid isobutanoic <sup>a,c</sup>	DMF	14.05	14.11	14.29
117	Àcid isobutanoic <sup>a,c</sup>	DMSO	12.79	12.68	12.76
118	Àcid isobutanoic <sup>a,c</sup>	H <sub>2</sub> O	4.88	4.80	4.95
119	Àcid isobutanoic <sup>a,c</sup>	MeOH	9.90	9.94	9.95
120	Àcid isobutanoic <sup>b,c</sup>	Ac	18.90	18.69	18.66
121	Àcid isobutanoic <sup>d</sup>	AN	22.20	22.32	22.37
122	Àcid isohexanoic <sup>a,c</sup>	H <sub>2</sub> O	4.80	5.01	5.01
123	Àcid isohexanoic <sup>d</sup>	MeOH	9.88	10.13	10.01
124	Àcid isopentanoic <sup>a,c</sup>	<sup>t</sup> PrOH	12.30	12.08	12.38
125	Àcid isopentanoic <sup>a,c</sup>	EtOH	10.80	10.96	10.73
126	Àcid isopentanoic <sup>a,c</sup>	H <sub>2</sub> O	4.74	4.90	4.98
127	Àcid isopentanoic <sup>d</sup>	MeOH	9.81	10.03	9.89
128	Àcid làctic <sup>a,c</sup>	H <sub>2</sub> O	3.73	3.55	3.76
129	Àcid làctic <sup>a,c</sup>	MeOH	8.81	8.62	8.64
130	Àcid mandèlic <sup>a,c</sup>	DMSO	8.79	9.17	9.03
131	Àcid mandèlic <sup>a,c</sup>	H <sub>2</sub> O	3.37	3.21	3.18
132	Àcid mandèlic <sup>d</sup>	MeOH	8.40	8.27	7.94
133	Àcid mercaptoacètic <sup>a,c</sup>	MeOH	8.52	8.60	8.53
134	Àcid mercaptoacètic <sup>d</sup>	H <sub>2</sub> O	3.73	3.51	3.38
135	Àcid metacrílic <sup>a,c</sup>	H <sub>2</sub> O	4.16	4.02	3.95
136	Àcid pentanoic <sup>a,c</sup>	H <sub>2</sub> O	4.25	4.42	4.17
137	Àcid pentanoic <sup>a,c</sup>	H <sub>2</sub> O	4.70	4.79	4.86
138	Àcid pentanoic <sup>a,c</sup>	MeOH	9.86	9.96	9.86
139	Àcid pentanoic <sup>b,c</sup>	<sup>t</sup> PrOH	12.30	12.17	12.25
140	Àcid pentanoic <sup>d</sup>	EtOH	10.70	10.92	10.64
141	Àcid propanoic <sup>a,c</sup>	<sup>t</sup> PrOH	12.20	12.19	11.96
142	Àcid propanoic <sup>a,c</sup>	AN	22.04	22.16	22.28
143	Àcid propanoic <sup>a,c</sup>	DMA	12.05	12.24	12.29
144	Àcid propanoic <sup>a,c</sup>	DMF	13.95	13.63	13.71
145	Àcid propanoic <sup>a,c</sup>	DMSO	12.52	12.24	12.28
146	Àcid propanoic <sup>a,c</sup>	EtOH	10.60	10.67	10.36
147	Àcid propanoic <sup>b,c</sup>	MeOH	9.71	9.75	9.84
148	Àcid propanoic <sup>d</sup>	Ac	18.74	18.44	18.28
149	Àcid propanoic <sup>d</sup>	H <sub>2</sub> O	4.88	4.59	4.71
150	Àcid tricloroacètic <sup>a,c</sup>	AN	10.57	10.58	10.61
151	Àcid tricloroacètic <sup>a,c</sup>	H <sub>2</sub> O	0.70	0.96	0.94
152	Àcid tricloroacètic <sup>a,c</sup>	MeOH	4.90	4.85	4.66

#	Àcid carboxílic	Solvent	Exp.	BPNN	SVM
153	Àcid tricloroacètic <sup>a,c</sup>	NM	7.27	7.18	7.51
154	Àcid tricloroacètic <sup>a,c</sup>	<sup>t</sup> BuOH	8.90	8.87	8.66
155	Àcid vinilacètic <sup>a,c</sup>	<sup>i</sup> PrOH	11.20	11.30	11.44
156	Àcid vinilacètic <sup>a,c</sup>	H <sub>2</sub> O	4.34	4.31	4.39
157	Àcid vinilacètic <sup>a,c</sup>	MeOH	9.50	9.50	9.60
158	Àcid vinilacètic <sup>d</sup>	EtOH	10.20	10.40	9.94

<sup>a</sup> Training set de BPNN. <sup>b</sup> Cross-validation set de BPNN.

<sup>c</sup> Training set de SVM. <sup>d</sup> Prediction set de BPNN i SVM.

#### 4.2.5. $pK_b$ d'anilines

Les anilines són la quarta i última família en què s'ha aplicat la metodologia QSPR per calcular les constants de dissociació. Cal dir que, en aquest cas, es calcularà el  $pK_b$  i no el  $pK_a$  de les espècies orgàniques. La primera intenció va ser calcular la constant d'acidesa d'anilines protonades però, per evitar treballar amb compostos carregats, que donen problemes amb el programa CODESSA, al final es va decidir que es calcularia la constant de basicitat. Així, els valors de  $pK_a$  que s'havien buscat en la bibliografia pels cations aniloni es van convertir en valors de  $pK_b$  per les respectives anilines mitjançant la constant d'autoprotòlisi de cada solvent.

La família de les anilines presenta analogies semblants a les que es donen en les famílies dels fenol i dels àcids benzoics. Tots els compostos presenten un anell aromàtic però, a la vegada, el seu comportament químic varia molt degut als diferents substituents que es poden trobar en les posicions possibles de substitució de l'anell. Com a novetat per aquesta família, hi ha la possibilitat de substitució sobre l'àtom de nitrogen, les anilines *N*-substituïdes. Degut a aquestes raons el rang de  $pK_b$  per aquesta família s'estén a més de 23 unitats. D'acord amb els bons resultats aconseguits amb la resta de famílies d'àcids orgànics, s'espera poder aplicar la metodologia QSPR al càlcul de les constants d'associació d'aquesta família de compostos amb models no lineals desenvolupats amb BPNN i SVM.

#### *Conjunt de dades*

La recerca bibliogràfica va proporcionar 108 valors experimentals de  $pK_b$  a partir de 41 anilines en set solvents diferents: aigua, metanol, acetona, acetonitril, DMSO, DMF i nitrometà. Els valors experimentals en aigua i metanol s'han obtingut, majoritàriament, dels treballs de Bosch i col·laboradors,<sup>172,192</sup> alguns valors s'han tret d'altres publicacions;<sup>223,224</sup> els valors en solvents apròtics s'han extret preferentment de la compilació d'Izutsku<sup>183</sup> encara que alguns valors han estat trobats en altres fonts.<sup>225-</sup>

229



Tot i l'elevat nombre de posicions de substitució que presenten les anilines, ha estat difícil trobar valors de  $pK_b$  per a anilines amb més de tres posicions substituïdes. D'entre les 41 anilines, sense tenir en compte l'anilina mateixa, s'han trobat valors de  $pK_b$  de 26 anilines monosubstituïdes, 3 di- i 3 trisubstituïdes i 8 *N*-substituïdes. Els substituents trobats són els ja habituals en aquestes famílies de compostos: halògens, hidroxils, grups alquílics i alcoxi, etc. Com en la resta de famílies estudiades, el nombre de valors en cada solvent és molt variable. És relativament senzill trobar més valors d'aquesta propietat en aigua però no s'han inclòs tots en la base de dades per tal de no afavorir l'aparició de models que funcionin bé per predir el  $pK_b$  en aquest solvent però que no ho facin per la resta de solvents. Així, s'han triat 28 valors de  $pK_b$  en aigua, 26 en metanol, 8 en acetona, 11 en acetonitril, 6 en DMSO, 5 en DMF i 24 en nitrometà; per tant, es tenen 54 valors en solvents pròtics i 54 en apròtics. En aquest cas la substitució en la posició *orto* de l'anell aromàtic pot ser important i condicionar el valor de  $pK_b$  mesurat, per aquesta raó durant l'estudi s'analitzaran amb detall els resultats que s'obtenen per les anilines que estan o no substituïdes en aquesta posició. En total hi ha 25 anilines que tenen la posició *orto* substituïda i 83 que no la tenen. El rang de  $pK_b$  que cobreixen els valors seleccionats és molt ampli: abasta gairebé 23 unitats logarítmiques, des de 7.43 per a la *N,N*-dietilanilina en aigua fins a 30.8 per a la 2,6-dicloro-4-nitroanilina en acetonitril. Tots els valors experimentals i calculats de  $pK_b$  es troben en la Taula 4.43.

La divisió del conjunt de dades en els *tset* i *pset* necessaris per establir el model amb SVM proporciona subconjunts de 88 i 20 compostos, respectivament. A l'hora de fer l'estudi amb les xarxes neuronals implementades pel programa ADAPT es necessita construir el *cvset* amb aproximadament un 10% del *tset*. Així, en l'estudi BPNN, els tres sets *tset*, *cvset* i *pset* constaran de 78, 10 i 20 valors de  $pK_b$ , respectivament.

### **Model no lineal: Xarxes neuronals BPNN**

El programa CODESSA calcula un total de 565 descriptors moleculars de les anilines. Després d'afegir-hi els 37 descriptors dels solvents el nombre total de descriptors és de 602. El mètode heurístic implementat per CODESSA és l'encarregat de fer la primera reducció en el nombre de descriptors mitjançant els seus criteris. Després d'aquest procediment el nombre total de descriptors s'ha reduït a 180, 163 dels àcids carboxílics i 17 dels solvents. La segona reducció en el nombre de descriptors, utilitzant criteris objectius, que fa el programa ADAPT rebaixa aquest nombre fins a 90, 80 de les anilines i 10 dels solvents. L'arquitectura de les xarxes neuronals és, en aquest cas, 7-4-1. Amb aquesta arquitectura s'obté un nombre de paràmetres ajustables de 37, el *ratio* de compostos del *tset*, 78, sobre el de paràmetres ajustables és de 2.11, encara per sobre del 2.0 recomanat.

El model que s'obté amb les xarxes neuronals d'arquitectura 7-4-1 conté els descriptors que es mostren en la Taula 4.44. Com en la resta de famílies de compostos que s'han tractat abans, el model presenta cinc descriptors dels soluts i dos descriptors dels solvents.

Taula 4.44. Descriptors continguts en el model de les anilines amb BPNN.

<b>Descriptor</b>	<b>Component</b>	<b>Classe</b>
Càrrega parcial mínima d'un àtom de nitrogen	Solut	Electrostàtic
Àrea superficial positiva fraccionada, FPSA-3	Solut	Electrostàtic
Energia HOMO	Solut	Quàntic
Índex màxim de reacció nucleòfila d'un àtom de nitrogen	Solut	Quàntic
Energia màxima d'intercanvi d'un enllaç C-N	Solut	Quàntic
Capacitat de donació d'enllaç d'hidrogen de Kamlet i Taft, $\alpha$	Solvent	-
Energia de vaporització molar, $\Delta_{\text{vap}}U^{\circ}$	Solvent	-

Com es pot comprovar, els descriptors obtinguts són del mateix tipus que els que s'han trobat anteriorment amb les altres famílies de compostos. Hi ha descriptors de dues classes diferents de les anilines: dos electrostàtics i tres quàntics. Els descriptors del solvent són els mateixos que s'han trobat per als àcids benzoics i els àcids carboxílics:  $\alpha$  de Kamlet i Taft<sup>160</sup> i l'energia de vaporització molar,  $\Delta_{\text{vap}}U^{\circ}$ .

Els descriptors electrostàtics, que es calculen a partir de la distribució de càrregues parcials calculada pel mètode de Zefirov,<sup>73</sup> són la càrrega parcial mínima sobre un àtom de nitrogen i l'àrea superficial positiva fraccionada, FPSA-3. La càrrega parcial mínima sobre un àtom de nitrogen se situa sempre sobre el nitrogen anilínic, que és el punt de la molècula on s'ha d'enllaçar el protó en la reacció estudiada. El descriptor FPSA-3, de tipus CPSA, es calcula de la següent manera:

$$FPSA-3 = \frac{\sum_A q_A \cdot S_A}{TMSA}$$

on  $q_A$  són les càrregues atòmiques parcials positives,  $S_A$  és la superfície total accessible al solvent carregada positivament i  $TMSA$  és la superfície total de la molècula. Els descriptors quàntics, que s'obtenen directament del càlcul AM1 implementat per MOPAC, són: l'energia de l'orbital molecular HOMO, l'índex màxim de reacció nucleòfila d'un àtom de nitrogen i l'energia màxima d'intercanvi d'un enllaç C-N. Aquests dos últims descriptors se situen sempre sobre el nitrogen anilínic i sobre l'enllaç C-N que uneix aquest àtom de nitrogen amb l'anell aromàtic, respectivament. Així doncs, sembla que els descriptors obtinguts poden ser útils per interpretar la reacció que s'està estudiant. La Figura 4.29 mostra les posicions que ocupen els descriptors que es poden fixar sobre l'estructura molecular.

El descriptor  $\alpha$  de Kamlet i Taft mesura la capacitat de les molècules de solvent per establir enllaços donadors d'hidrogen amb els soluts. L'energia de vaporització molar del solvent,  $\Delta_{\text{vap}}U^\circ$ , és una mesura de la capacitat del solvent per formar les cavitats en les quals s'han d'encabir els soluts.

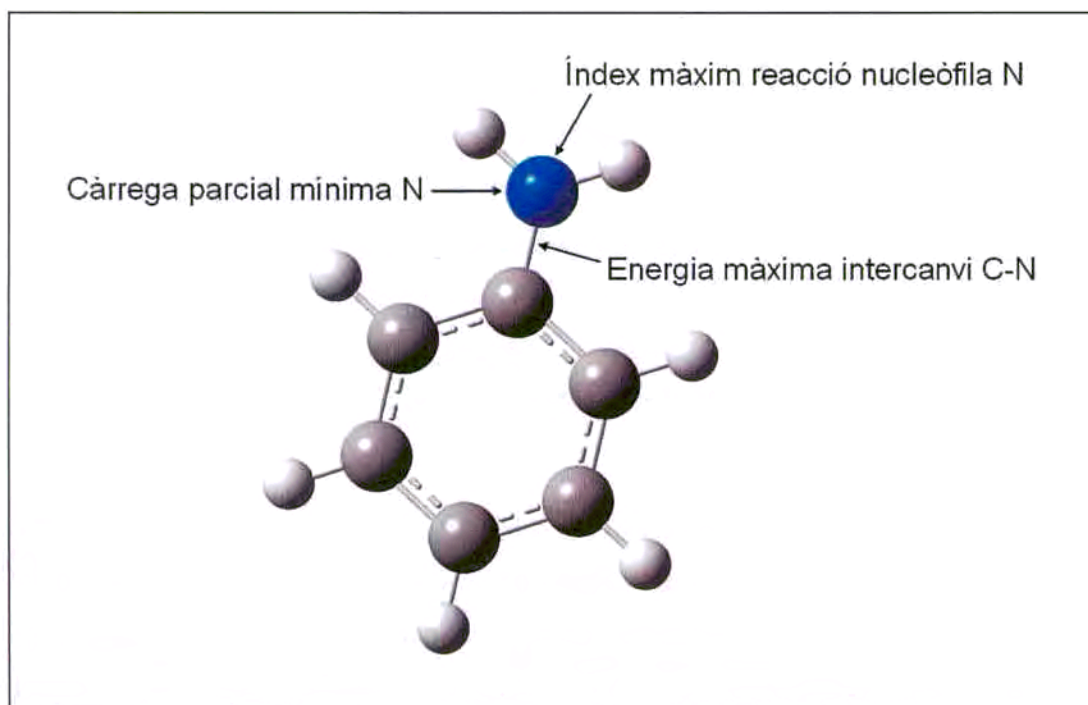


Figura 4.29. Localització dels descriptors del model sobre les anilines.

Els resultats trobats, Taula 4.45, són realment molt bons per a tots els subconjunts de valors: *tset*, *cvset* i *pset*. Els valors de  $pK_b$  calculats es poden trobar en la Taula 4.43. La representació gràfica dels valors calculats enfront dels experimentals es pot veure en la Figura 4.30.

Taula 4.45. Paràmetres estadístics obtinguts amb el model BPNN.

Conjunt	n	$R^2$	F	rmse
<i>tset</i>	78	0.999	61358.3	0.24
<i>cvset</i>	10	0.999	8116.8	0.26
<i>pset</i>	20	0.999	13529.6	0.28

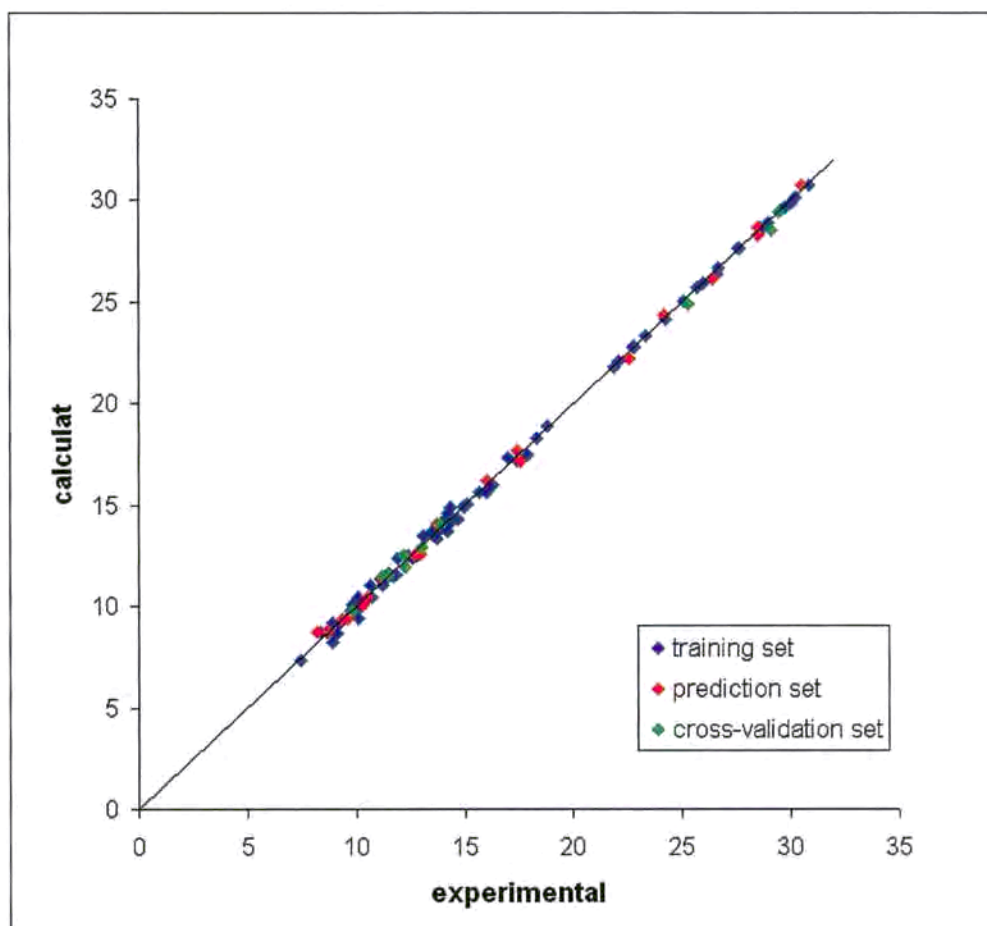


Figura 4.30. Representació gràfica de  $pK_b$  calculat vs. experimental per a les anilines amb BPNN.

De la mateixa manera que amb les altres famílies, els descriptors continguts en el model es poden relacionar amb la reacció estudiada en base a les interaccions solut/solvent. Els descriptors de les anilines: càrrega parcial mínima d'un àtom de nitrogen, FPSA-3 i energia màxima d'intercanvi d'un enllaç C-N fan referència a les interaccions solut/solvent no específiques, a l'igual que el descriptor de solvent energia de vaporització molar ( $\Delta_{\text{vap}}U^\circ$ ). Per la seva banda, les interaccions solut/solvent específiques venen referides pels descriptors energia de l'orbital HOMO i l'índex màxim de reacció nucleòfila d'un àtom de nitrogen de les anilines i per  $\alpha$  de Kamlet i Taft dels solvents.

El descriptor càrrega parcial mínima sobre un àtom de nitrogen se situa sempre sobre el nitrogen anilínic i presenta valors negatius per totes les anilines. Aquest descriptor mesura la densitat electrònica sobre el nitrogen que ha de cedir el seu parell



d'electrons no enllaçant per formar el catió aniloni. Les anilines que presenten substituents electroatracients sobre l'anell aromàtic tenen valors més alts (menys negatius) per aquest descriptor mentre que les anilines *N*-substituïdes o substituïdes amb grups donadors o alquílics tenen valors més petits. El descriptor energia màxima d'intercanvi en un enllaç C-N, que se situa sempre en l'enllaç entre el nitrogen anilínic i l'anell aromàtic, es pot relacionar amb els efectes que produeixen els substituents sobre l'anell. Els compostos que tenen substituents electroatracients en l'anell presenten valors més alts mentre que els que tenen substituents alquílics o alcoxi presenten valors més baixos. El descriptor FPSA-3 fa referència a la capacitat de l'anilina per establir interaccions electrostàtiques i polars amb les molècules del solvent. El descriptor de solvent  $\Delta_{\text{vap}}U^\circ$  és una mesura de l'energia necessària per formar les cavitats en el solvent on es produirà el procés de dissolució de les molècules de solut.

Les interaccions solut/solvent específiques, aquelles que es relacionen amb els aspectes acceptor/donador o d'àcid/base de Lewis, es poden relacionar amb els dos descriptors que resten de les anilines i el descriptor  $\alpha$  de Kamlet i Taft dels solvents. Aquest últim descriptor és, per definició, idoni per a aquest estudi ja que és una mesura de la capacitat del solvent per establir enllaços d'hidrogen per donació a les molècules de solut i, tal i com s'ha exposat al principi, en la reacció que s'està estudiant el solvent ha de donar un protó a l'anilina per formar el corresponent catió aniloni. El descriptor índex màxim de reacció nucleòfila d'un àtom de nitrogen, derivat de la teoria de Fukui,<sup>74</sup> se situa també sobre el nitrogen anilínic i mesura la facilitat amb la qual aquest àtom cedeix el seu parell d'electrons en la reacció estudiada. Les anilines *N*-substituïdes presenten els valors més alts per aquest descriptor, demostrant que són les que cediran el parell no enllaçant amb més facilitat. Per últim, el descriptor energia de l'orbital HOMO es correspon al primer potencial d'ionització de l'anilina i es pot relacionar directament amb el caràcter de base de Lewis d'aquests compostos.

La importància relativa dels descriptors continguts en el model es pot mesurar amb el mètode de Jurs. Aquest mètode, que es basa en l'increment de *rmse* que es produeix en el model quan es fan permutes dels descriptors, dona els resultats que es troben en la Taula 4.46.

Taula 4.46. Increments de *rmse* per la permuta de cada descriptor.

Descriptor	Increment <i>rmse</i>	Importància
Càrrega parcial mínima d'un àtom de nitrogen	0.90	6
Àrea superficial positiva fraccionada, FPSA-3	1.37	3
Energia HOMO	1.30	5
Índex màxim de reacció nucleòfila d'un àtom de nitrogen	0.49	7
Energia màxima d'intercanvi d'un enllaç C-N	1.33	4
Capacitat de donació d'enllaç d'hidrogen de Kamlet i Taft, $\alpha$	7.65	1
Energia de vaporització molar, $\Delta_{\text{vap}}U^\circ$	3.24	2

Com en tots els casos anteriors, els dos descriptors més importants són els dels solvents:  $\alpha$  de Kamlet i Taft i  $\Delta_{\text{vap}}U^\circ$ . Tot just després es troben tres descriptors de les anilines que presenten un grau d'importància gairebé idèntic: FPSA-3, energia màxima d'intercanvi d'un enllaç C-N i energia de l'orbital HOMO.

El valor de *rmse* que s'obté amb el model per a tots els compostos estudiats és de 0.25. S'han trobat cinc compostos que presenten un residual dues vegades més gran que aquest *rmse*, dins d'ells no se n'ha trobat cap amb residual tres vegades superior a aquest valor; la Taula 4.47 mostra aquests compostos. No sembla existir una tendència estructural en base a l'anilina ja que tots cinc valors corresponen a cinc anilines diferents, dues d'elles substituïdes en *orto*- i dues *N*-substituïdes. Dins dels tres solvents, els residuals són sempre positius per a les mesures en aigua, 2, i negatius per a les mesures en nitrometà, 2; tot i així, la resta d'anilines en aquests dos solvents presenten residuals positius i negatius, sense cap tendència aparent en aquest sentit.

Taula 4.47. Prediccions amb valors més grans que dues vegades el *rmse*.

Solut	Solvent	Exp.	Calc.	residual
3-metilnilina	NM	14.3	14.88	-0.58
<i>N,N</i> -dietilnilina	NM	11.89	12.40	-0.51
4-cloro-2-nitroanilina	AN	29,1	28,51	0.59
2,6-dimetilnilina	H <sub>2</sub> O	10.11	9.44	0.67
<i>N,N</i> -dimetilnilina	H <sub>2</sub> O	8.93	8.24	0.69

Per comprovar la robustesa del model establert, s'han estudiat els paràmetres estadístics que s'obtenen per a tots els subconjunts d'estructures en què es pot dividir el conjunt inicial. La Taula 4.48 mostra tots aquests subconjunts i els paràmetres estadístics que s'obtenen per a cadascun d'ells. Els resultats per a tots els subconjunts són bons, sobretot tenint en compte el gran desequilibri que hi ha entre el nombre de valors en solvents pròtics i apròtics.

Taula 4.48. Estadística dels subconjunts d'anilines.

Subconjunt	n	R <sup>2</sup>	F	rmse
<i>orto</i>	25	0.999	18898.2	0.27
No <i>orto</i>	83	0.999	57739.9	0.25
Pròtics	54	0.986	3598.0	0.26
Apròtics	54	0.998	33348.1	0.25
No <i>N</i> -substituïdes	91	0.999	83532.2	0.23
<i>N</i> -substituïdes	17	0.998	6424.3	0.33
H <sub>2</sub> O	28	0.977	1090.4	0.27
MeOH	26	0.984	1522.5	0.26
DMSO	6	0.961	99.4	0.16
DMF	5	0.991	329.4	0.16
AN	11	0.996	2262.9	0.24
Ac	8	0.992	781.1	0.17
NM	24	0.971	731.6	0.32

Els resultats obtinguts amb el model desenvolupat són molt bons i són millors o iguals que els que es poden trobar en la bibliografia. En un estudi recent es calcula amb mètodes teòrics *ab initio* el p*K*<sub>a</sub> de 20 amines en acetonitril, els resultats obtinguts donen uns errors al voltant d'una unitat de p*K*<sub>a</sub>, per sobre dels resultats aquí obtinguts.<sup>155</sup> Un altre treball, on s'utilitzen els models de Hammet-Taft i Drago per predir el p*K*<sub>a</sub> de 28 anilines protonades en metanol proporciona resultats semblants als obtinguts aquí.<sup>156</sup>

**Model no lineal: Support Vector Machines**

El model SVM utilitza els mateixos descriptors trobats amb els algorismes genètics pel programa ADAPT en el model BPNN. Aquest model es construeix mitjançant l'optimització dels paràmetres  $\gamma$ ,  $C$  i  $\epsilon$  amb el mètode de minimització de l'error de validació creuada per als compostos del *tset*, Figura 4.31. L'optimització del paràmetre  $\gamma$  (amb  $\epsilon=0.1$  i  $C=100$ ) condueix a un valor òptim de 3.4. L'optimització del paràmetre  $\epsilon$  (amb  $\gamma=3.4$  i  $C=100$ ) fixa aquest paràmetre en 0.07. El paràmetre  $C$  s'optimitza amb  $\gamma=3.4$  i  $\epsilon=0.07$ . El valor amb el qual el *rmse* del model és mínim és  $C=150$ .

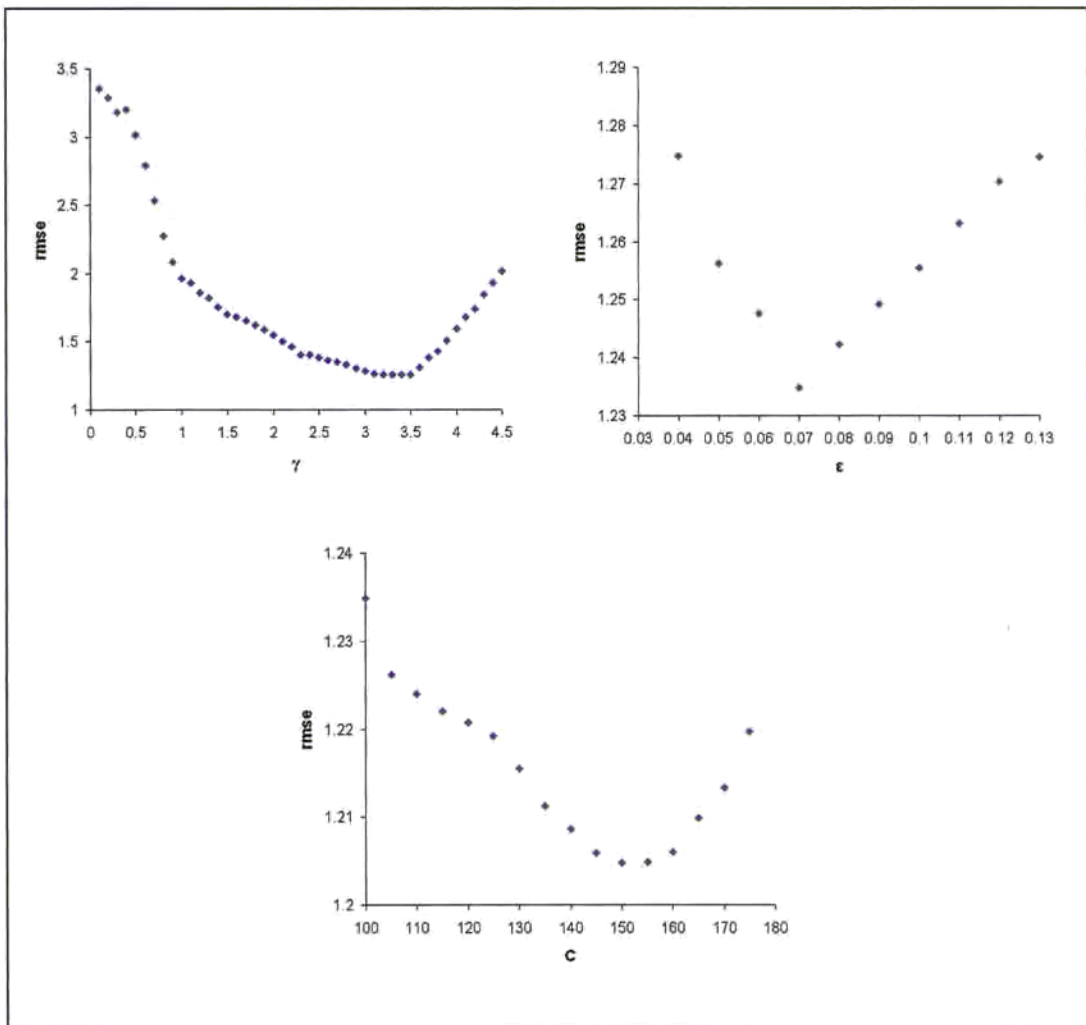


Figura 4.31. Optimització dels paràmetres  $\gamma$ ,  $\epsilon$  i  $C$ .

El model obtingut amb aquests paràmetres es construeix amb 74 vectors suport, escollits d'entre els 88 valors experimentals que conformen el *tset*. Els resultats que s'obtenen per aquest conjunt de dades i per al *pset* són molt bons. En la Taula 4.49 es troben els paràmetres estadístics obtinguts per a aquests dos conjunts de compostos. Els valors de  $pK_b$  calculats amb aquest model es troben a la Taula 4.43 i la seva representació gràfica enfront dels valors de  $pK_b$  experimentals es mostra en la Figura 4.32.

Taula 4.49. Paràmetres estadístics obtinguts amb el model SVM.

Conjunt	n	R <sup>2</sup>	F	rmse
<i>tset</i>	88	0.999	240758.5	0.13
<i>pset</i>	20	0.999	45757.7	0.15

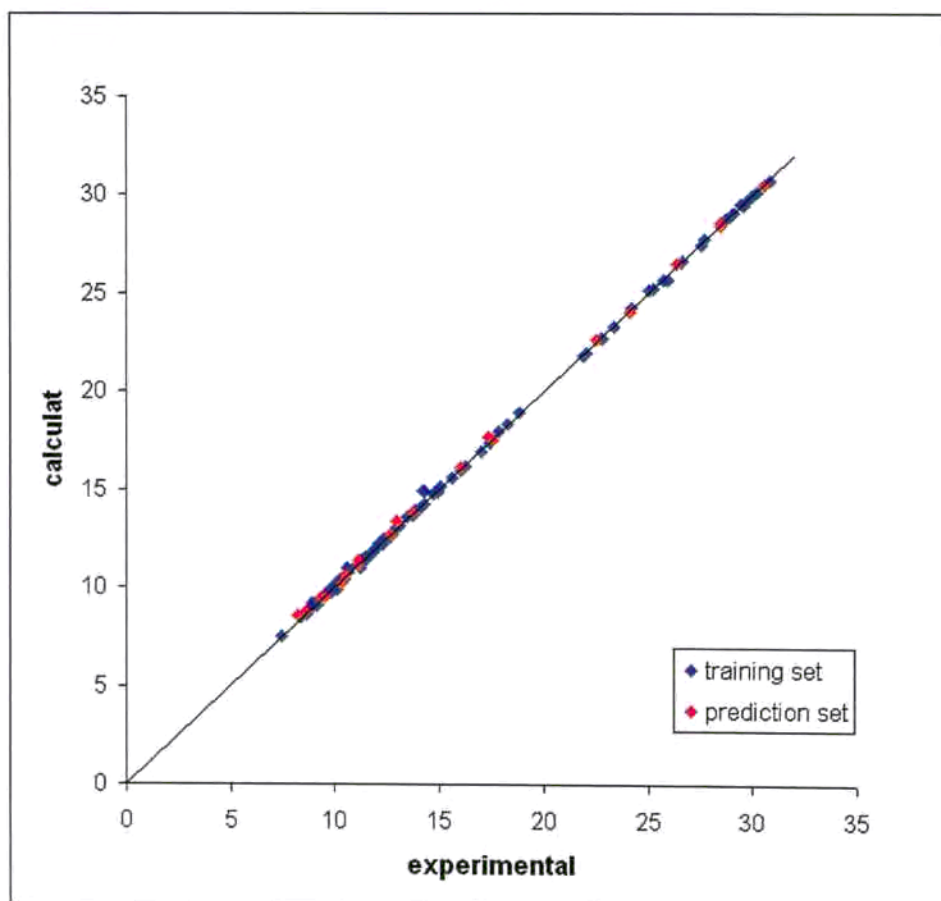


Figura 4.32. Representació gràfica de  $pK_b$  calculat vs. experimental per a les anilines amb SVM.



Els resultats que s'obtenen analitzant els resultats dels subconjunts de compostos de naturalesa anàloga són també molt bons i són encara millors, en tots els casos, que els que s'obtenen amb el mètode BPNN. La Taula 4.50 mostra els paràmetres estadístics que s'obtenen per a tots aquests subconjunts de compostos.

Taula 4.50. Estadística dels subconjunts d'anilines.

<b>Subconjunt</b>	<b>n</b>	<b>R<sup>2</sup></b>	<b>F</b>	<b>rmse</b>
<i>orto</i>	25	1.000	208770.2	0.08
No <i>orto</i>	83	1.000	166035.4	0.15
Pròtics	54	0.997	18032.1	0.12
Apròtics	54	0.999	91939.2	0.15
No <i>N</i> -substituïdes	91	1.000	241998.9	0.14
<i>N</i> -substituïdes	17	1.000	37594.3	0.14
H <sub>2</sub> O	28	0.996	6087.1	0.11
MeOH	26	0.997	6999.9	0.12
DMSO	6	0.993	541.2	0.07
DMF	5	0.997	1191.1	0.08
AN	11	1.000	27186.5	0.07
Ac	8	0.997	1763.8	0.11
NM	24	0.988	1748.5	0.21

Taula 4.43. Valors experimentals i calculats de  $pK_b$  per a la família de les anilines.

#	Anilina	Solvent	Exp.	BPNN	SVM
1	2,6-dicloro-4-nitroanilina <sup>a,c</sup>	AN	30.80	30.80	30.73
2	2,6-dimetilanilina <sup>a,c</sup>	H <sub>2</sub> O	10.11	9.44	9.86
3	2,6-dimetilanilina <sup>a,c</sup>	MeOH	11.70	11.50	11.63
4	2-aminoanilina <sup>a,c</sup>	DMF	24.21	24.19	24.26
5	2-aminoanilina <sup>a,c</sup>	DMSO	28.50	28.66	28.57
6	2-bromoanilina <sup>b,c</sup>	H <sub>2</sub> O	11.47	11.57	11.48
7	2-bromoanilina <sup>d</sup>	MeOH	13.74	14.12	13.78
8	2-cloroanilina <sup>a,c</sup>	H <sub>2</sub> O	11.33	11.38	11.40
9	2-cloroanilina <sup>a,c</sup>	MeOH	13.49	13.69	13.56
10	2-cloroanilina <sup>d</sup>	NM	17.54	17.18	17.47
11	2-hidroxianilina <sup>a,c</sup>	NM	14.15	14.39	14.11
12	2-metilanilina <sup>a,c</sup>	MeOH	11.25	11.43	11.32
13	2-metilanilina <sup>a,c</sup>	NM	14.95	14.92	14.88
14	2-metilanilina <sup>d</sup>	H <sub>2</sub> O	9.57	9.40	9.50
15	2-metoxianilina <sup>a,c</sup>	NM	14.68	14.32	14.75
16	2-nitroanilina <sup>a,c</sup>	Ac	28.73	28.71	28.77
17	2-nitroanilina <sup>a,c</sup>	H <sub>2</sub> O	14.25	14.23	14.18
18	2-nitroanilina <sup>a,c</sup>	MeOH	17.00	17.34	16.93
19	2-nitroanilina <sup>d</sup>	AN	28.45	28.72	28.52
20	3,5-dimetoxianilina <sup>a,c</sup>	H <sub>2</sub> O	10.14	9.97	10.21
21	3,5-dimetoxianilina <sup>a,c</sup>	MeOH	11.54	11.55	11.47
22	3-aminoanilina <sup>a,c</sup>	NM	13.87	13.82	13.94
23	3-aminoanilina <sup>d</sup>	DMF	24.13	24.43	24.06
24	3-bromoanilina <sup>a,c</sup>	H <sub>2</sub> O	10.44	10.37	10.37
25	3-bromoanilina <sup>a,c</sup>	MeOH	12.78	12.55	12.71
26	3-cloroanilina <sup>a,c</sup>	H <sub>2</sub> O	10.49	10.38	10.42
27	3-cloroanilina <sup>a,c</sup>	NM	16.26	16.01	16.19
28	3-cloroanilina <sup>d</sup>	MeOH	12.68	12.50	12.66
29	3-etoxianilina <sup>a,c</sup>	MeOH	11.83	11.55	11.76
30	3-etoxianilina <sup>b,c</sup>	H <sub>2</sub> O	9.82	9.75	9.83
31	3-fluoroanilina <sup>a,c</sup>	MeOH	12.60	12.42	12.53
32	3-fluoroanilina <sup>a,c</sup>	NM	16.18	15.95	16.11
33	3-fluoroanilina <sup>d</sup>	H <sub>2</sub> O	10.50	10.47	10.57
34	3-hidroxianilina <sup>a,c</sup>	H <sub>2</sub> O	9.75	9.76	9.82
35	3-hidroxianilina <sup>b,c</sup>	MeOH	11.21	11.49	11.14
36	3-metilanilina <sup>a,c</sup>	MeOH	11.11	11.31	11.14
37	3-metilanilina <sup>a,c</sup>	NM	14.30	14.88	14.88
38	3-metilanilina <sup>b,c</sup>	DMSO	29.44	29.42	29.51
39	3-metilanilina <sup>d</sup>	Ac	26.35	26.20	26.47
40	3-metilanilina <sup>d</sup>	H <sub>2</sub> O	9.31	9.33	9.38
41	3-metoxianilina <sup>a,c</sup>	DMSO	30.01	29.87	29.94
42	3-metoxianilina <sup>a,c</sup>	H <sub>2</sub> O	9.80	9.67	9.73
43	3-metoxianilina <sup>d</sup>	MeOH	11.16	11.39	11.36
44	3-nitroanilina <sup>a,c</sup>	Ac	27.53	27.64	27.46
45	3-nitroanilina <sup>a,c</sup>	AN	25.70	25.71	25.63
46	3-nitroanilina <sup>a,c</sup>	DMF	26.66	26.68	26.59
47	3-nitroanilina <sup>a,c</sup>	H <sub>2</sub> O	11.49	11.66	11.56
48	3-nitroanilina <sup>a,c</sup>	MeOH	13.74	14.01	13.81
49	3-nitroanilina <sup>d</sup>	NM	17.38	17.74	17.67
50	4,5-dicloro-2-nitroanilina <sup>a,c</sup>	AN	30.20	30.20	30.13

#	Anilina	Solvent	Exp.	BPNN	SVM
51	4,5-dimetil-2-nitroanilina <sup>a,c</sup>	AN	27.65	27.65	27.72
52	4-aminoanilina <sup>a,c</sup>	DMF	22.77	22.78	22.70
53	4-aminoanilina <sup>b,c</sup>	NM	12.15	12.50	12.22
54	4-benzilanilina <sup>a,c</sup>	H <sub>2</sub> O	9.22	9.20	9.15
55	4-benzilanilina <sup>a,c</sup>	MeOH	11.22	11.02	10.96
56	4-bromoanilina <sup>a,c</sup>	H <sub>2</sub> O	10.12	10.44	10.19
57	4-bromoanilina <sup>a,c</sup>	MeOH	12.36	12.52	12.43
58	4-bromoanilina <sup>d</sup>	NM	16.03	16.26	16.10
59	4-bromo- <i>N,N</i> -dimetilanilina <sup>a,c</sup>	NM	14.28	14.11	14.27
60	4-cloro-2-nitroanilina <sup>a,c</sup>	H <sub>2</sub> O	15.07	15.07	15.14
61	4-cloro-2-nitroanilina <sup>a,c</sup>	MeOH	17.87	17.53	17.94
62	4-cloro-2-nitroanilina <sup>b,c</sup>	AN	29.10	28.51	29.07
63	4-cloro-2-nitro- <i>N</i> -metilanilina <sup>a,c</sup>	AN	29.53	29.54	29.46
64	4-cloroanilina <sup>a,c</sup>	H <sub>2</sub> O	10.02	10.03	9.95
65	4-cloroanilina <sup>a,c</sup>	NM	16.01	15.65	15.94
66	4-cloroanilina <sup>b,c</sup>	MeOH	12.25	11.96	12.18
67	4-cloroanilina <sup>d</sup>	DMSO	30.51	30.75	30.44
68	4-cyanoanilina <sup>a,c</sup>	NM	18.28	18.28	18.35
69	4-etoxianilina <sup>a,c</sup>	MeOH	10.28	10.11	10.35
70	4-etoxianilina <sup>d</sup>	H <sub>2</sub> O	8.68	8.73	8.75
71	4-hidroxianilina <sup>a,c</sup>	H <sub>2</sub> O	8.35	8.75	8.42
72	4-hidroxianilina <sup>a,c</sup>	MeOH	9.79	10.12	9.86
73	4-hidroxianilina <sup>a,c</sup>	NM	13.09	13.48	13.16
74	4-metilanilina <sup>a,c</sup>	Ac	25.93	25.97	25.68
75	4-metilanilina <sup>a,c</sup>	AN	22.05	22.06	21.98
76	4-metilanilina <sup>a,c</sup>	DMSO	28.96	28.94	28.89
77	4-metilanilina <sup>a,c</sup>	H <sub>2</sub> O	8.92	9.21	9.21
78	4-metilanilina <sup>a,c</sup>	MeOH	10.63	11.06	10.97
79	4-metilanilina <sup>a,c</sup>	NM	14.20	14.61	14.87
80	4-metoxianilina <sup>a,c</sup>	H <sub>2</sub> O	8.64	8.73	8.57
81	4-metoxianilina <sup>a,c</sup>	NM	13.72	13.36	13.65
82	4-metoxianilina <sup>d</sup>	MeOH	10.31	10.09	10.13
83	4-nitroanilina <sup>a,c</sup>	MeOH	15.65	15.66	15.58
84	4-nitroanilina <sup>a,c</sup>	NM	18.82	18.94	18.89
85	4-nitroanilina <sup>b,c</sup>	H <sub>2</sub> O	13.00	12.92	13.07
86	4-nitroanilina <sup>d</sup>	Ac	28.48	28.29	28.41
87	Anilina <sup>a,c</sup>	Ac	26.58	26.42	26.51
88	Anilina <sup>a,c</sup>	AN	22.74	22.84	22.67
89	Anilina <sup>a,c</sup>	DMF	25.04	25.04	25.11
90	Anilina <sup>a,c</sup>	DMSO	29.70	29.63	29.63
91	Anilina <sup>a,c</sup>	H <sub>2</sub> O	9.40	9.38	9.46
92	Anilina <sup>a,c</sup>	NM	14.93	15.01	14.86
93	Anilina <sup>d</sup>	MeOH	11.15	11.44	11.11
94	<i>N,N</i> -dietilanilina <sup>a,c</sup>	Ac	23.33	23.36	23.26
95	<i>N,N</i> -dietilanilina <sup>a,c</sup>	H <sub>2</sub> O	7.43	7.36	7.50
96	<i>N,N</i> -dietilanilina <sup>a,c</sup>	MeOH	10.28	10.09	10.21
97	<i>N,N</i> -dietilanilina <sup>a,c</sup>	NM	11.89	12.40	11.96
98	<i>N,N</i> -dimetilanilina <sup>a,c</sup>	AN	21.87	21.79	21.80
99	<i>N,N</i> -dimetilanilina <sup>a,c</sup>	H <sub>2</sub> O	8.93	8.24	9.00
100	<i>N,N</i> -dimetilanilina <sup>b,c</sup>	Ac	25.26	24.89	25.19
101	<i>N,N</i> -dimetilanilina <sup>d</sup>	NM	12.96	12.59	13.38

#	Anilina	Solvent	Exp.	BPNN	SVM
102	<i>N</i> -etilanolina <sup>b,c</sup>	NM	13.82	14.08	13.75
103	<i>N</i> -isopropilanolina <sup>a,c</sup>	MeOH	10.70	10.46	10.77
104	<i>N</i> -isopropilanolina <sup>d</sup>	H <sub>2</sub> O	8.23	8.73	8.55
105	<i>N</i> -metil-3-nitroanolina <sup>a,c</sup>	NM	17.41	17.22	17.34
106	<i>N</i> -metilanolina <sup>a,c</sup>	H <sub>2</sub> O	9.15	8.69	9.08
107	<i>N</i> -metilanolina <sup>a,c</sup>	NM	14.20	13.75	14.13
108	<i>N</i> -metilanolina <sup>d</sup>	AN	22.55	22.25	22.62

<sup>a</sup> *Training set* de BPNN. <sup>b</sup> *Cross-validation set* de BPNN.

<sup>c</sup> *Training set* de SVM. <sup>d</sup> *Prediction set* de BPNN i SVM.

### 4.3. Bibliografia

1. Katritzky, A. R.; Fara, D. C. *Energy & Fuels* **2005**, 19, 922.
2. Suzuki, T.; Ebert, R.-U.; Schüürmann, G. *J. Chem. Inf. Comput. Sci.* **2001**, 41, 776.
3. Chalk, A. J.; Beck, B.; Clark, T. A. *J. Chem. Inf. Comput. Sci.* **2001**, 41, 1053.
4. Halberstam, N. M.; Baskin, I. I.; Palyulin, V. A.; Zefirov, N. S. *Dokl. Chem* **2002**, 384, 140.
5. Halberstam, N. M.; Baskin, I. I.; Palyulin, V. A.; Zefirov, N. S. *Mendeleev Comm.* **2002**, 185.
6. Marino, T.; Russo, N.; Toscano, M. *Inorg. Chem.* **2001**, 40, 6439.
7. Siu, F. M.; Ma, N. L.; Tsang, C. W. *J. Chem. Phys.* **2001**, 114, 7045.
8. Talley, J. M.; Cerda, B. A.; Ohanessian, G.; Wesdemiotis, C. *Chem. Eur. J.* **2002**, 8, 1377.
9. Castleman, A. W.; Keesee, R. G. *Chem. Rev.* **1986**, 86, 589.
10. Cerda, B. A.; Wesdemiotis, C. *J. Am. Chem. Soc.* **1995**, 117, 9734.
11. Marshall, A. G. *Acc. Chem. Res.* **1996**, 29, 307.
12. Rodgers, M. T.; Armentout, P. B. *Mass Spectrom. Rev.* **2000**, 19, 215.
13. Luo, Y. R. *Comprehensive Handbook of Chemical Bond Energies*, CRC Press: Boca Raton, **2007**.
14. Cooks, R. G.; Patrick, J. S.; Kotiaho, T.; McLukey, S. A. *Mass Spectrom. Rev.* **1994**, 13, 287.
15. Cooks, R. G.; Wong, P. H. S. *Acc. Chem. Res.* **1998**, 31, 379.
16. Gapeev, A.; Dunbar, R. C. *Int. J. Mass Spectrom.* **2003**, 228, 825.
17. Lau, J. K. C.; Wong, C. H. S.; Ng, P. S.; Siu, F. M.; Ma, N. L.; Tsang, C. W. *Chem. Eur. J.* **2003**, 9, 3383.
18. Shoeib, T.; Siu, K. W. M.; Hopkinson, A. C. *J. Phys. Chem. A* **2002**, 106, 6121.
19. Gal, J.-F.; Maria, P.-C.; Massi, L.; Mayeux, C.; Burk, P.; Tamniku-Taul, J. *Int. J. Mass Spectrom.* **2007**, 267, 7.
20. Russo, N.; Toscano, M.; Grand, A. *J. Phys. Chem. A* **2003**, 107, 11533.
21. Mercero, J. M.; Fowler, J. E.; Ugalde, J. M. *J. Phys. Chem. A* **1998**, 102, 7006.
22. Mercero, J. M.; Fowler, J. E.; Ugalde, J. M. *J. Phys. Chem. A* **2000**, 104, 7053.
23. Mercero, J. M.; Irigoras, A.; López, X.; Fowler, J. E.; Ugalde, J. M. *J. Phys. Chem. A* **2001**, 105, 7446.



24. Mercero, J. M.; Matxain, J. M.; López, X.; Fowler, J. E.; Ugalde, J. M. *Int. J. Quant. Chem.* **2002**, 90, 859.
25. Marino, T.; Russo, N.; Toscano, M. *J. Mass Spectrom* **2002**, 37, 786.
26. Godbout, N.; Salahub, D. R.; Andzelm, J.; Wimmer, E. *Can. J. Chem. - Revue Canadienne De Chimie* **1992**, 70, 560.
27. Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, 19, 553.
28. Simon, S.; Duran, M.; Dannenberg, J. J. *J. Chem. Phys.* **1996**, 105, 11024.
29. Pulkkinen, S.; Noguera, M.; Rodriguez-Santiago, L.; Sodupe, M.; Bertran, J. *Chem. Eur. J.* **2000**, 6, 4393.
30. Strittmatter, E. F.; Lemoff, A. S.; Williams, E. R. *J. Phys. Chem. A* **2000**, 104, 9793.
31. Hoyau, S.; Ohanessian, G. *Chem. Eur. J.* **1998**, 4, 1561.
32. Halgren, T. A. *J. Comp. Chem.* **1996**, 17, 490.
33. Wavefunction Inc., Spartan 5.0, Irvine, CA, **1995**.
34. Gaussian 03, Revision 6.0, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, **2003**.
35. Marino, T.; Russo, N.; Toscano, M. *J. Inorg. Biochem.* **2000**, 79, 179.
36. Remko, M.; Rode, B. M. *J. Phys. Chem. A* **2006**, 110, 1960.
37. Ai, H. Q.; Bu, Y. X.; Li, P.; Zhang, C. *New J. Chem.* **2005**, 29, 1540.
38. Jensen, F. *J. Am. Chem. Soc.* **1992**, 114, 9533.
39. Abirami, S.; Wong, C. C. L.; Tsang, C. W.; Ma, N. L. *Chem. Eur. J.* **2005**, 11, 5289.
40. Abirami, S.; Xing, Y. M.; Tsang, C. W.; Ma, N. L. *J. Phys. Chem. A* **2005**, 109, 500.

41. Jockusch, R. A.; Price, W. D.; Williams, E. R. *J. Phys. Chem. A* **1999**, 103, 9266.
42. Andersen, U. N.; Bojesen, G. *J. Chem. Soc. Perkin Trans. 2* **1997**, 323.
43. Feng, W. Y.; Gronert, S.; Lebrilla, C. *J. Phys. Chem. A* **2003**, 107, 405.
44. Ryzhov, V.; Dunbar, R. C.; Cerda, B.; Wesdemiotis, C. *J. Am. Soc. Mass Spectrom.* **2000**, 11, 1037.
45. Bojesen, G.; Breindahl, T.; Andersen, U. N. *Org. Mass Spectrom.* **1993**, 28, 1448.
46. Marino, T.; Russo, N.; Toscano, M. *J. Phys. Chem. B* **2003**, 107, 2588.
47. Moision, R. M.; Armentrout, P. B. *J. Phys. Chem. A* **2006**, 110, 3933.
48. Rozman, M.; Bertosa, B.; Klasinc, L.; Srzic, D. *J. Am. Soc. Mass Spectrom.* **2006**, 17, 29.
49. Jockusch, R. A.; Lemoff, A. S.; Williams, E. R. *J. Am. Chem. Soc.* **2001**, 123, 12255.
50. Kish, M. M.; Ohanessian, G.; Wesdemiotis, C. *Int. J. Mass Spectrom.* **2003**, 227, 509.
51. Dunbar, R. C. *J. Phys. Chem. A* **2000**, 104, 8067.
52. Hoyau, S.; Norrman, K.; McMahon, T. B.; Ohanessian, G. *J. Am. Chem. Soc.* **1999**, 121, 8864.
53. Moision, R. M.; Armentrout, P. B. *J. Phys. Chem. A* **2002**, 106, 10350.
54. Klassen, J. S.; Anderson, S. G.; Blades, A. T.; Kebarle, P. *J. Phys. Chem.* **1996**, 100, 14218.
55. Ye, S. J.; Moision, R. M.; Armentrout, P. B. *Int. J. Mass Spectrom.* **2005**, 240, 233.
56. Ye, S. J.; Moision, R. M.; Armentrout, P. B. *Int. J. Mass Spectrom.* **2006**, 253, 288.
57. Wong, C. H. S.; Siu, F. M.; Ma, N. L.; Tsang, C. W. *J. Mol. Struct. Theochem* **2002**, 588, 9.
58. Moision, R. M.; Armentrout, P. B. *Phys. Chem. Chem. Phys.* **2004**, 6, 2588.
59. Hoyau, S.; Ohanessian, G. *C. R. Acad. Sci., Serie Fasc. C-Chimie* **1998**, 1, 795.
60. Cerda, B. A.; Wesdemiotis, C. *Int. J. Mass Spectrom.* **1999**, 187, 107.
61. Hoyau, S.; Ohanessian, G. *J. Am. Chem. Soc.* **1997**, 119, 2016.
62. Bertran, J.; Rodriguez-Santiago, L.; Sodupe, M. *J. Phys. Chem. B* **1999**, 103, 2310.
63. Hoppilliard, Y.; Ohanessian, G.; Bourcier, S. *J. Phys. Chem. A* **2004**, 108, 9687.
64. Luna, A.; Alcamí, M.; Mo, O.; Yanez, M. *Chem. Phys. Lett.* **2000**, 320, 129-138.
65. Lynch, B. J.; Truhlar, D. G. *Chem. Phys. Lett.* **2002**, 361, 251.
66. Rimola, A.; Rodriguez-Santiago, L.; Sodupe, M. *J. Phys. Chem. B* **2006**, 110, 24189.

67. Shoeib, T.; Rodriguez, C. F.; Siu, K. W. M.; Hopkinson, A. C. *Phys. Chem. Chem. Phys.* **2001**, 3, 853.
68. Caraiman, D.; Shoeib, T.; Siu, K. W. M.; Hopkinson, A. C.; Bohme, D. K. *Int. J. Mass Spectrom.* **2003**, 228, 629.
69. Lee, V. W. M.; Li, H. B.; Lau, T. C.; Guevremont, R.; Siu, K. W. M. *J. Am. Soc. Mass Spectrom.* **1998**, 9, 760.
70. Carvajal, M. A.; Novoa, J. J.; Álvarez, S. *J. Am. Chem. Soc.* **2004**, 126, 1465.
71. Amunugama, R.; Rodgers, M. T. *J. Phys. Chem. A* **2002**, 106, 5529.
72. Emsley, J. *The Elements*, Claredon Press, **1989**.
73. Zefirov, N. S.; Kirpichenok, M. A.; Izmailov, F. F.; Trofimov, M. I. *Dokl. Akad. SSSR* **1987**, 296, 883.
74. Fukui, K. *Theory of Orientation and Stereoselection*, Springer-Verlag: Berlin, **1975**.
75. Shannon, R. D. *Acta Crystallogr.* **1976**, A32, 751.
76. Wang, H.; Ulander, J. *Toxicology* **2006**, 2, 139.
77. Reichardt, C. *Solvent and Solvent Effects in Organic Chemistry*, 3<sup>rd</sup> Ed., VCH: Weinheim, **2003**.
78. Albert, A.; Serjeant, E. P. *The determination of ionization constants*, Chapman & Hall: London, **1971**.
79. Cookson, R. F. *Chem. Rev.* **1974**, 74, 5.
80. Qiang, Z.; Adams, C. *Water Res.* **2004**, 38, 2874.
81. Ertekin, K.; Alp, S.; Yalcin, I. *Dyes and pigments* **2005**, 65, 33.
82. Seok, Y.-J.; Yang, K.-S.; Kang, S.-O. *Anal. Chim. Acta* **1995**, 306, 351.
83. Leito, I.; Rodima, T.; Koppel, I. A.; Schwesinger, R.; Vlasov, V. M. *J. Org. Chem.* **1997**, 62, 8479.
84. Saurina, J.; Hernández-Cassou, S.; Tauler, R.; Izquierdo-Ridorsa, A. *Anal. Chim. Acta* **2000**, 408, 135.
85. Szacács, Z.; Hägele, G. *Talanta* **2004**, 62, 819.
86. Szacács, Z.; Hägele, G.; Tyka, R. *Anal. Chim. Acta* **2004**, 522, 247.
87. Hardcastle, J. E.; Jano, E. *J. Chromatogr. B* **1998**, 717, 39.
88. Winczling, P.; Markuszewski, M. J.; Kaliszan, R. *Anal. Chem.* **2004**, 76, 3069.
89. Herrero-Martínez, J. M.; Sanmartín, M.; Rosés, M.; Bosch, E.; Ràfols, C. *Electrophoresis* **2005**, 26, 1886.
90. Caliaro, G. A.; Herbots, C. A. *J. Pharm. And Biomed. Anal.* **2001**, 26, 427.
91. Smith, B. J.; Radom, L. *J. Phys. Chem.* **1991**, 95, 10549.

92. Smith, B. J.; Radom, L. *J. Am. Chem. Soc.* **1993**, 115, 4885.
93. Ochterski, J.; Petersson, G. A.; Wiberg, K. B. *J. Am. Chem. Soc.* **1995**, 117, 11299.
94. Bagno, A.; Scorrano, G. *J. Phys. Chem.* **1996**, 100, 1536.
95. Martin, J. M. L.; Lee, T. J. *Chem. Phys. Lett.* **1996**, 258, 136.
96. Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, 109, 7764.
97. Peterson, K. A.; Xantheas, S. S.; Dixon, D. A.; Dunning, T. H., Jr. *J. Phys. Chem. A* **1998**, 102, 2449.
98. Hammerum, S. *Chem. Phys. Lett.* **1999**, 300, 529.
99. Burk, P.; Koppel, I. A.; Koppel, I.; Leito, I.; Travnikova, O. *Chem. Phys. Lett.* **2000**, 323, 482.
100. Seo, Y.; Kim, Y.; Kim, Y. *Chem. Phys. Lett.* **2001**, 340, 186.
101. Pokon, E. K.; Liptak, M. D.; Feldgus, S.; Shields, G. C. *J. Phys. Chem. A* **2001**, 105, 10483.
102. Ervin, K. M.; DeTuri, V. F. *J. Phys. Chem. A* **2002**, 106, 9947.
103. Wiberg, K. B. *J. Org. Chem.* **2002**, 67, 4787.
104. Remko, M. *J. Phys. Chem. A* **2002**, 106, 5005.
105. Jorgensen, W. L.; Briggs, J. M.; Gao, J. *J. Am. Chem. Soc.* **1987**, 109, 6857.
106. Rullmann, J. A. C.; van Duijnen, P. Th. *Mol. Phys.* **1988**, 63, 451.
107. Guissani, Y.; Guillot, B.; Bratos, S. *J. Chem. Phys.* **1988**, 88, 5850.
108. Jorgensen, W. L.; Briggs, J. M. *J. Am. Chem. Soc.* **1989**, 111, 4190.
109. Tanaka, Y.; Shiratori, Y.; Nakagawa, S. *Chem. Phys. Lett.* **1990**, 169, 513.
110. Gao, J.; Li, N.; Freinforf, M. *J. Am. Chem. Soc.* **1996**, 118, 4912.
111. Figueirido, F.; Del Buono, G. S.; Levy, R. M. *J. Phys. Chem.* **1996**, 100, 6389.
112. Byun, K.; Mo, Y.; Gao, J. *J. Am. Chem. Soc.* **2001**, 123, 3974.
113. Florian, J.; Warshel, A. *J. Phys. Chem. B* **1997**, 101, 5583.
114. Lim, C.; Bashford, D.; Karplus, M. *J. Phys. Chem.* **1991**, 95, 5610.
115. Kawata, M.; Ten-no, S.; Kato, S.; Hirata, F. *Chem. Phys.* **1996**, 203, 53.
116. Kawata, M.; Ten-no, S.; Kato, S.; Hirata, F. *J. Phys. Chem.* **1996**, 100, 1111.
117. Sato, H.; Hirata, F. *J. Phys. Chem. A* **1998**, 102, 2603.
118. Sato, H.; Hirata, F. *J. Phys. Chem. B* **1999**, 103, 6596.
119. Grüber, C.; Buß, V. *Chemosphere* **1989**, 19, 1595.
120. Schüürmann, G. *Quant. Struc-Act. Relat.* **1996**, 15, 121.
121. Güven, A.; Öğretir, C. *Journal of Molecular Structure (Theochem)* **1998**, 427, 65.

122. Civcir, P.Ü.; Öğretir, C. *Journal of Molecular Structure (Theochem)* **2000**, 507, 39.
123. Topol, I. A.; Tawa, G. J.; Caldwell, R. A.; Eissenstat, M. A.; Burt, S. K. *J. Phys. Chem. A* **2000**, 104, 9619.
124. Jang, Y. H.; Sowers, L. C.; Cagin, T.; Goddard, W. A., III. *J. Phys. Chem. A* **2001**, 105, 274.
125. Klicic, J. J.; Friesner, R. A.; Liu, S.-Y.; Guida, W. C. *J. Phys. Chem. A* **2002**, 106, 1327.
126. Lopez, X.; Schaefer, M.; Dejaegere, A.; Karplus, M. *J. Am. Chem. Soc.* **2002**, 124, 5010.
127. Adam, K. R. *J. Phys. Chem. A* **2002**, 106, 11963.
128. Cramer, C. J.; Truhlar, D. G. *ReV. Comput. Chem.* **1995**, 6, 1.
129. Cramer, C. J.; Truhlar, D. G. *Chem. ReV.* **1999**, 99, 2161.
130. Cammi, R.; Mennucci, B.; Tomasi, J. *J. Phys. Chem. A* **1998**, 102, 870.
131. Cammi, R.; Mennucci, B.; Tomasi, J. *J. Phys. Chem. A* **2000**, 104, 4690.
132. Schüürmann, G.; Cossi, M.; Barone, V.; Tomasi, J. *J. Phys. Chem. A* **1998**, 102, 6706.
133. da Silva, C. O.; da Silva, E. C.; Nascimento, M. A. C. *J. Phys. Chem. A* **1999**, 103, 11194.
134. da Silva, C. O.; da Silva, E. C.; Nascimento, M. A. C. *J. Phys. Chem. A* **2000**, 104, 2402.
135. Liptak, M. D.; Shields, G. C. *J. Am. Chem. Soc.* **2001**, 123, 7314.
136. Toth, A. M.; Liptak, M. D.; Phillips, D. L.; Shields, G. C. *J. Chem. Phys.* **2001**, 114, 4595.
137. Liptak, M.; Shields, G. C. *Int. J. Quantum Chem.* **2001**, 85, 727.
138. Liptak, M. D.; Gross, K. C.; Seybold, P. G.; Feldgus, S.; Shields, G. C. *J. Am. Chem. Soc.* **2002**, 124, 6421.
139. Namazian, M.; Heidary, H. *Theochem* **2003**, 620, 257.
140. Saracino, G. A. A.; Improta, R.; Barone, V. *Chem. Phys. Lett.* **2003**, 373, 411.
141. Pliego, J. R., Jr.; Riveros, J. M. *J. Phys. Chem. A* **2001**, 105, 7241
142. Pliego, J. R., Jr.; Riveros, J. M. *J. Phys. Chem. A* **2002**, 106, 7434.
143. Wiberg, K.B.; Castejon, H.; Keith, T.A. *J. Comput. Sci.* **1996**, 17, 185.
144. Chipman, D.N. *J. Phys. Chem A* **2002**, 106, 7413.



145. Almerindo, G. I.; Tondo, D. W.; Pliego Jr., J. R. *J. Phys. Chem A* **2004**, 108, 166.
146. Fu, Y.; Liu, L.; Li, R.-Q.; Liu, R.; Guo, Q.-X. *J. Am. Chem. Soc.* **2004**, 126, 6421.
147. Tehan, B. G.; Lloyd, E. J.; Wong, M. G.; Pitt, W. R.; Montana, J. G.; Manallack, D. T.; Gancia, E. *Quant. Struct.-Act. Relat.* **2002**, 21, 457.
148. Tehan, B. G.; Lloyd, E. J.; Wong, M. G.; Pitt, W. R.; Montana, J. G.; Manallack, D. T.; Gancia, E. *Quant. Struct.-Act. Relat.* **2002**, 21, 473.
149. Citra, J.M. *Chemosphere* **1999**, 38, 191.
150. Habibi-Yangjeh, A.; Danandeh-Jenagharad, M.; Nooshyar, M. *J. Mol. Model.* **2006**, 12, 338.
151. Zhang, J.; Kleinöder, T.; Gasteiger, J. A. *J. Chem. Inf. Model.* **2006**, 46, 2256.
152. Ghasemi, J.; Saaidpour, S.; Brown, S. D. *Journal of Molecular Structure: THEOCHEM* **2007**, 805, 27.
153. Jelfs, S.; Ertl, P.; Selzer, P. *J. Chem. Inf. Model.* **2007**, 47, 450.
154. Gieleciak, R.; Polanski, J. *J. Chem. Inf. Model.* **2007**, 47, 547.
155. Li, J.-N.; Fu, Y.; Liu, L.; Guo, Q.-X. *Tetrahedron* **2006**, 62, 11801.
156. Bosch, E.; Rived, F.; Rosés, M.; Sales, J. *J. Chem. Soc. Perkin Trans 2* **1999**, 1953.
157. Marcus, Y. *The properties of solvents*, John Wiley and Sons: New York, **1999**.
158. Abboud, J.-L. M.; Notario, R. *Pure Appl. Chem.* **1999**, 71, 645.
159. Katritzky, A. R.; Tamm, T.; Wang, Y.; Sild, S.; Karelson, M. *J. Chem. Inf. Comput. Sci.* **1999**, 39, 684.
160. Taft, R. W.; Kamlet M. J. *J. Am. Chem. Soc.* **1976**, 98, 2886.
161. Kamlet M. J.; Taft, R. W. *J. Am. Chem. Soc.* **1976**, 98, 377.
162. Kamlet, M. J.; Abboud, J.-L. M.; Taft, R. W. *J. Am. Chem. Soc.* **1977**, 99, 6027.
163. Dimroth, K.; Reichardt, C.; Siepmann, T.; Bohlmann, F. *Liebigs Ann. Chem.* **1963**, 661, 1.
164. Koppel, I. A.; Palm, V. A. In N.B. Chapman, J. Shorter (eds) *The Influence of Solvent on Organic Reactivity in Advances in Linear Free Energy Relationships* Chapter 5, 204-280, Plenum Press: London, **1972**.
165. Koppel, I. A.; Paju, A. I. *Org. Reactivity* **1974**, 1, 121.
166. Gutmann, V.; Wychera, E. *Inorg. Nucl. Chem. Lett.* **1966**, 2, 257.
167. Mayer, U.; Gutmann, V.; Gerger, W. *Monatsh. Chem.*, **1975**, 1235.
168. Kosower, E. M. *J. Am. Chem. Soc.* **1958**, 80, 3253.

169. Catalán, J.; Díaz, C. A. *Liebigs Ann./Recueil* **1997**,1941.
170. Catalán, J.; Díaz, C.; López, V.; Pérez, P.; Paz J. L. G.; Rodríguez, J. G. *Liebigs Ann.* **1996**, 1785.
171. Catalán, J.; López, V.; Pérez, P.; Martín-Villamil, R.; Rodríguez, J. G. *Liebigs Ann.* **1995**, 241.
172. Swain, C. G.; Swain, M. S.; Powell, A. L.; Alumni, S. *J. Am. Chem. Soc.* **1983**, 105, 502.
173. Drago, R. S. *J. Chem. Soc. Perkin Trans. 2* **1992**, 1827.
174. Drago, R. S. *Applications of Electrostatic-Covalent Models in Chemistry*. Surfside: Gainesville, **1994**.
175. Browstein, S. *Can. J. Chem.* **1960**, 38, 1590.
176. Janowski, A.; Turowska-Tyrk, I.; Wrona, P. K. *J. Chem. Soc., Perkin Trans. 2* **1985**, 821.
177. Fowler, F. W.; Katritzky, A. R.; Rutherford, R. J. D. *J. Chem. Soc. B* **1971**, 460.
178. Bosch, E.; Ràfols, C.; Rosés, M. *Talanta* **1989**, 36, 1227.
179. Bosch, E.; Ràfols, C.; Rosés, M. *Anal. Chim. Acta* **1995**, 302, 109.
180. Ràfols, C.; Rosés, M.; Bosch, E. *Anal. Chim. Acta* **1997**, 338, 127.
181. Rived, F.; Rosés, M.; Bosch, E. *Anal. Chim. Acta* **1998**, 374, 309.
182. Rosés, M.; Rived, F.; Bosch, E. *J. Chromatograph. A* **2000**, 867, 45.
183. Izutsu, K. *Acid-base dissociation constants in dipolar aprotic solvents*, Blackwell Scientific Publications, **1990**.
184. Maran, F.; Celadon, D.; Severin, M. G.; Vianello, E. *J. Am. Chem. Soc.* **1991**, 113, 9320.
185. Bosque, R.; Sales, J. *J. Chem. Inf. Comput. Sci.* **2003**, 43, 637.
186. Pytela, O.; Kulhánek, J. *Collect. Czech. Chem. Commun.* **2002**, 67, 596.
187. Pytela, O.; Kulhánek, J.; Ludwig, M.; Ríha, V. *Collect. Czech. Chem. Commun.* **1994**, 59, 627.
188. Pytela, O.; Kulhánek, J.; Ludwig, M. *Collect. Czech. Chem. Commun.* **1994**, 59, 1637.
189. Pytela, O.; Kulhánek, J. *Collect. Czech. Chem. Commun.* **1995**, 60, 829.
190. Pytela, O.; Prusek, O. *Collect. Czech. Chem. Commun.* **1999**, 64, 1617.
191. Pytela, O.; Kulhánek, J. *Collect. Czech. Chem. Commun.* **1997**, 62, 913.
192. Rived, F.; Canals, I.; Bosch, E.; Rosés, M. *Anal. Chim. Acta* **2001**, 439, 315.
193. Chantooni, M. K.; Kolthoff, I. M. *Anal. Chem.* **1979**, 51, 133.

194. Kolthoff, I. M.; Chantooni, M. K. *J. Am. Chem. Soc.* **1971**, 93, 3843.
195. Kulhánek, J.; Exner, O. *J. Chem. Soc. Perkin Trans. 2* **1998**, 1397.
196. Kortüm, G.; Vogel, W.; Andrussow, K. *Dissociation Constants of Organic Acids in Aqueous Solution*, Butterworths, London, **1961**.
197. Strong, L. E.; Blubaugh, D. J.; Cavalli, C. R. *J. Solution Chem.* **1981**, 10, 811.
198. Dippy, J. F. J.; Hughes, S. R. C.; Laxton, J. W. *J. Chem. Soc.* **1958**, 1470.
199. Menger, F. M.; Mounier, C. E. *J. Am. Chem. Soc.* **1993**, 115, 12222.
200. Dippy, J. F. J.; Hughes, S. R. C.; Laxton, J. W. *J. Chem. Soc.* **1956**, 2995.
201. Aksoy, M. S.; Oezer, U. *Chem. Pharm. Bull.* **2004**, 52, 1280.
202. Dey, B. P.; Biswas, P. K.; Lahiri, S. C. *J. Ind. Chem. Soc.* **2003**, 80, 271.
203. Peltier *Bull. Soc. Sci. Bretagne* **1956**, 31, 9.
204. Sotomatsu, T.; Shigemura, M.; Murata, Y.; Fujita, T. *Bull. Chem. Soc. Jpn.* **1992**, 65, 3157.
205. Dash, U. N. *J. Ind. Chem. Soc.* **1994**, 71, 65.
206. Hoek *Sven. Kem. Tidskr.* **1953**, 65, 182.
207. Mock, W. L.; Morsch, L. A. *Tetrahedron* **2001**, 57, 2957.
208. Peltier *Bull. Soc. Chim. Fr.* **1958**, 994.
209. Bordfoss *Z. Phys. Chem. Stoechiom. Verwandtschaftsl.* **1922**, 102, 46.
210. Jones, B.; Speakman, J. C. *J. Chem. Soc.* **1944**, 19.
211. Garrett, E. R. *J. Am. Chem. Soc.* **1957**, 79, 3401.
212. Porras, S. P.; Riekkola, M.-L.; Kenndler, E. *Electrophoresis* **2001**, 22, 3798.
213. Exner, O.; Fiedler, P.; Budesmsky, M.; Kulhanek, J. *J. Org. Chem.* **1999**, 64, 3513.
214. Headley, A. D.; Starnes, S. D.; Wilson, L. Y.; Famini, G. R. *J. Org. Chem.* **1994**, 59, 8040.
215. Minnick, I. J.; Kilpatrick, M. *J. Phys. Chem.* **1939**, 43, 259.
216. Henne, A. L.; Fox, C. J. *J. Am. Chem. Soc.* **1954**, 76, 479.
217. Sigel, H.; Malini-Balakrishnan, R.; Haering, U. K. *J. Am. Chem. Soc.* **1985**, 107, 5137.
218. Viryulina, V. N.; Chupakhina, R. A.; Serebrennikov, V. V. *J. Gen. Chem. USSR (Engl. Transl.)* **1981**, 51, 1244.
219. Chawla, B.; Mehta, S. K. *J. Phys. Chem.* **1984**, 88, 2650.
220. Bartnicka, H.; Bojanowska, I.; Kalinowski, M. K. *Aust. J. Chem.* **1991**, 44, 1077.

221. Shannon, C.; Weaver, W. *The Mathematical Theory of Communication*, University of Illinois Press: Urbana (ILL), **1949**.
222. Bonchev, D. *Information Theoretic Indices for Characterization of Chemical Structures*, Research Studies Press: Chichester (UK), **1983**.
223. Porras, S. P.; Riekkola, M.-L.; Kenndler, E. *Journal of Chromatography A* **2001**, 905, 259.
224. Korzhenevskaya, N. G.; Mestechkin, M. M.; Chotii, K. Y.; Abdusalamov, A. B.; Ulanenko, K. B.; Rybachenko, V. I. *J.Gen.Chem.USSR (Engl.Transl.)* **1992**, 62, 939.
225. Kaljurand, I.; Kuett, A.; Soovaeli, L.; Rodima, T.; Maemets, V.; Leito, I.; Koppel, I. A. *J. Org. Chem.* **2005**, 70, 1019.
226. Crampton, M. R.; Rabbitt, L. C. *J. Chem. Soc., Perkin Trans. 2* **1999**, 1669.
227. Milyaev, Y. F.; Khorishko, S. A.; Balyatinskaya, L. N. *Gen.Chem.USSR (Engl.Transl.)* **1981**, 51, 168.
228. Kashik, T. V.; Usov, V. A.; Ponomareva, S. M.; Timokhina, L. V.; Voronkov, M. G. *J.Org.Chem.USSR (Engl.Transl.)* **1985**, 21, 549.
229. Pawlak, Z.; Kuna, S.; Richert, M.; Giersz, E.; Wisniewska, M. *J. Chem. Thermodynamics* **1991**, 23, 135.

## **Capítol 5. Conclusions i resum**



S'ha aplicat amb èxit la metodologia QSPR a un seguit de compostos i sistemes que fins ara no hi havien estat tractats:

- La viscositat,  $\eta$ , la tensió superficial,  $\sigma$ , i l'entalpia de formació en fase gasosa,  $\Delta_f H^\circ$ , de compostos organometàl·lics, de fórmula general,  $MR_nX_m$ , on M són metalls, semimetalls i fins i tot no metalls dels grups 12 al 16 de la Taula periòdica; R són grups orgànics habituals com alquil, aril, etc.; i els àtoms terminals X són els halògens Cl, Br i I o bé el H.
- La basicitat catiònica, en concret front el  $Li^+$ , d'un conjunt de molècules orgàniques de naturalesa química molt variada, és a dir l'energia associada a la formació dels compostos de coordinació tipus  $[LiL]^+$ .
- Sistemes multicomponents, en els que la propietat estudiada depèn, a més dels compostos que formen el conjunt analitzat, d'altres factors com la naturalesa dels cations –en el cas de l'afinitat i la basicitat catiòniques dels vint aminoàcids habituals– o dels solvents quan s'estudia la constant d'acidesa de fenols, àcids benzoics, àcids carboxílics alifàtics i anilines en diferents solvents, de naturalesa pròtica o apròtica.

En el tractament dels sistemes multicomponents s'han incorporat descriptors externs per caracteritzar els metalls i els solvents. Aquests descriptors són propietats físiques com potencial d'ionització, radi, electronegativitat, densitat, moment dipolar, etc.; per als solvents, també s'han utilitzat els paràmetres de les escales de polaritat més habituals, Kamlet i Taft, Reichardt, Koppel i Palm, Gutmann, Drago, entre d'altres.

En la deducció dels models s'han emprat mètodes lineals, mitjançant les tècniques de regressió multilineal de mínims quadrats, i mètodes no lineals. En aquest segon cas s'ha fet ús de dos tipus de xarxes neuronals: les BPNN i les RBFNN conjuntament amb tècniques SVM. Dels resultats obtinguts se'n dedueix que ambdós tipus són igualment apropiats. En la utilització d'aquestes xarxes neuronals s'han seguit els procediments habituals, amb les BPNN s'ha fet ús del *cross-validation set* per tal de controlar el procés d'ensinistrament de la xarxa i s'ha respectat que la relació entre el nombre compostos del *tset* i el de paràmetres ajustables hagi estat sempre per sobre del

valor mínim recomanat de 2. Per tal d'obtenir els models SVM, s'ha fet l'ajust dels tres paràmetres característics,  $\gamma$ ,  $\epsilon$  i  $C$ , lligats a aquest procediment.

Cal destacar, també, la utilització del mètode heurístic-SVM proposat per primera vegada en aquest treball, que permet seleccionar de manera no lineal els descriptors del model mitjançant els algorismes propis dels SVM.

En general, els valors dels paràmetres estadístics són excel·lents, tant per al *tset* com per al *pset* amb unes  $R^2 \geq 0.95$  i uns *rmse* bons comparats amb les mitjanes dels valors experimentals. La bondat d'aquests resultats es ratifica amb la representació gràfica dels valors calculats vs. experimentals, en què s'obtenen rectes de pendents propers a 1 i ordenades a l'origen molt petites comparades amb la mitjana dels valors experimentals. Els resultats obtinguts amb models no lineals són sempre millors que amb el lineal, tal com era d'esperar atesa la naturalesa de les propietats estudiades.

La validació del model, que és un dels punts clau en la metodologia QSPR, s'ha fet en tots els casos mitjançant un conjunt de valors extern, *prediction set*, format per compostos que no han participat en la seva deducció.

La robustesa del model s'ha comprovat mitjançant l'anàlisi dels factors estadístics obtinguts per a subconjunts determinats com, per exemple, aquells que contenen compostos d'unes característiques donades: tenir o no substituents en posició *orto*-, derivats d'un determinat element, en un mateix solvent, etc. En tots aquests casos els coeficients de determinació,  $R^2$ , i els *rmse* són molt semblants entre si i també als del *training set* i al conjunt de tots els compostos estudiats. El nombre de *outliers* per a cada cas és relativament baix, i no s'ha trobat cap relació evident amb les estructures moleculars; pot concloure's, doncs, que són veritables *outliers* i que la diferència amb els valors experimentals no és deguda a limitacions del model deduït.

Els descriptors que formen els models proposats contenen en general informació que permet una interpretació fisicoquímica raonable i prou detallada de la propietat experimental estudiada, tal com es comenta en detall en els punts posteriors. A assenyalar, també, que els models derivats per l'aproximació lineal i les dues no lineals emprades, contenen descriptors de les mateixes classes i sovint hi ha descriptors iguals

en models deduïts de manera diferent, fets que reafirmen la capacitat dels mètodes QSPR per calcular i interpretar les propietats tractades.

En general els models proposats contenen set descriptors, perquè a més de donar els millors resultats estadístics, és un nombre prou petit i capaç, alhora, d'incorporar descriptors moleculars i d'externs.

En relació als diferents temes estudiats es pot destacar:

### ***Viscositat i Tensió superficial de compostos organometàl·lics***

Els resultats obtinguts amb les dades determinades al voltant de 20°C, són comparables als trobats amb conjunts de compostos orgànics, en què els valors experimentals són més acurats, mostrant així la bona capacitat de l'aproximació QSPR.

Els millors resultats es troben amb els models no lineals tant BPNN com SVM; amb valors de  $R^2 \approx 0.94$  i  $0.98$ ; i de  $rmse \approx 0.10$  i  $1$ , per a la viscositat,  $\log \eta$ , i la tensió superficial,  $\sigma$ , respectivament que són clarament millors als obtinguts amb el model lineal. Els models proposats contenen descriptors de tipus CPSA: DPSA, FNSA, FPSA, PNSA, WNSA, HDSA, RPCG i RPCS, entre d'altres, que expliquen molt bé les interaccions intermoleculars, principals responsables de la viscositat i la tensió superficial. Aquests tipus de descriptors es troben també en els models reportats per als conjunts de compostos orgànics

Quan s'amplien els conjunts estudiats amb els valors experimentals a diferents temperatures, els resultats són bons i anàlegs al cas anterior. Tant la deducció lineal com la no lineal incorporen la temperatura –el seu invers en el cas de la viscositat– al model proposat. La resta de descriptors són igualment del tipus CPSA: DPSA, FPSA, PNSA, HDCA, RPCG i RPCS, semblants als de l'estudi a 20°C. Les  $R^2$  i  $rmse$  són també del mateix ordre; per a la viscositat:  $0.95$  i  $0.1$  i per a la tensió superficial:  $0.97$  i  $1$ . Tant a 20°C com a temperatura variable, els dos tipus de xarxes neuronals emprades donen resultats pràcticament idèntics, fent-los equivalents a efectes de càlcul.

### ***Entalpies de formació de compostos organometàl·lics***

Aquest és el primer estudi en què s'ha aplicat la metodologia QSPR a l'estimació de  $\Delta_f H^p$  a un conjunt de compostos organometàl·lics. El model lineal proposat conté nou descriptors, dos dels quals són del metall: electronegativitat de Pauling i la relació càrrega/radi iònic; els altres set són descriptors moleculars de tipus constitucional, excepte un de topològic: índex de Kier & Hall d'ordre 0. Aquest últim descriptor permet la diferenciació entre compostos isomèrics de cadena. S'obté una bona  $R^2 = 0.99$  i un *rmse* al voltant dels 29 kJ/mol, que comparat amb el dels valors experimentals és acceptable. A remarcar que tots els descriptors poden calcular-se directament i de manera immediata a partir de la fórmula molecular, fet que fa el mètode proposat com el més senzill en l'estimació de  $\Delta_f H^p$ , de compostos organometàl·lics de les característiques dels aquí estudiats. La robustesa del model es manifesta en el fet que els paràmetres estadístics per als dos subconjunts analitzats, els que contenen halògens i hidrògens terminals i els que no en contenen, són molt semblants.

La capacitat dels mètodes QSPR en la determinació d'entalpies de formació s'ha comprovat amb l'estudi fet amb un conjunt de compostos orgànics, de naturalesa semblant als grups presents en els derivats organometàl·lics, amb un model que conté només descriptors constitucionals, topològics i geomètrics, amb què s'obtenen estimacions comparables a les donades per mètodes additius tan consolidats com el de Pedley.

### ***Afinitat i basicitat catióniques de compostos orgànics***

Aquest estudi es pot dividir en dues parts. En la primera s'estudia la basicitat catiónica d'un bon nombre de compostos de naturalesa química molt variada amb el catió liti. Els models, que contenen set descriptors, s'han deduït amb tres mètodes: lineal, híbrid lineal-no lineal i no lineal tipus BPNN. Els resultats obtinguts amb tots tres models són bons, especialment amb el no lineal:  $R^2 = 0.94$  i *rmse* = 7.3 kJ/mol.

La comparació dels resultats obtinguts amb els calculats per mètodes *ab initio* i DFT trobats en la bibliografia permet constatar que l'aproximació QSPR proporciona, en aquest cas, estimacions similars a les que es poden obtenir amb metodologies molt més complexes.



En la segona part s'han estudiat la basicitat i l'afinitat catióniques dels 20 aminoàcids habituals amb els cations  $\text{Li}^+$ ,  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Cu}^+$  i  $\text{Ag}^+$ . Inicialment, s'ha desenvolupat una metodologia, basada en la teoria del funcional de la densitat, que permet el càlcul d'aquestes dues magnituds. La metodologia proposada, B3LYP/DZVP amb correcció del BSSE, proporciona resultats comparables a aquells obtinguts amb mètodes molt més costosos.

S'han explorat els diferents modes de coordinació dels aminoàcids amb els metalls, s'han buscat les conformacions més estables per als aminoàcids lliures i per als compostos que formen amb els cations metàl·lics i s'han obtingut valors d'afinitat i basicitat molt semblants als experimentals.

En particular, ha estat possible establir comparacions entre els diferents cations, donat que els càlculs es realitzen en les mateixes condicions per a tots ells. Així, els cations alcalins  $\text{Li}^+$ ,  $\text{Na}^+$  i  $\text{K}^+$  segueixen una tendència similar, amb algunes excepcions com els aminoàcids que contenen sofre. En canvi, els cations  $\text{Cu}^+$  i  $\text{Ag}^+$  segueixen una tendència similar entre si, però diferent als anteriors. En general, les afinitats catióniques segueixen la seqüència  $\text{Cu}^+ > \text{Li}^+ > \text{Ag}^+ > \text{Na}^+ > \text{K}^+$ .

L'esquema de descomposició de l'energia electrònica proposat ha permès explorar les tendències en l'energia de preorganització del lligand i en la contribució dels diferents heteroàtoms a l'energia d'interacció entre l'aminoàcid i el metall. Així mateix, possibilita conèixer la importància d'aquests dos factors en l'estabilitat dels diferents complexos catió/aminoàcid, en funció de la naturalesa del lligand i del metall.

El conjunt de dades que s'obté a partir dels càlculs teòrics, més complet i coherent que els que es poden construir a partir de les referències bibliogràfiques, ha estat utilitzat posteriorment per establir un model QSPR que permet estimar l'afinitat catiónica dels 20 aminoàcids amb els cations monovalents triats.

El model desenvolupat no linealment conté set descriptors. La bondat i la robustesa que presenta són molt elevades, obtenint-se uns coeficients de determinació,  $R^2$ , i uns valors de *rmse* al voltant de 0.99 i 3 kJ/mol per al conjunt complet i, també, per als subconjunts dels diferents cations. Dels set descriptors que formen el model, cinc són dels aminoàcids i dos dels metalls. Els primers es poden relacionar amb l'afinitat catiónica ja que es refereixen als heteroàtoms de l'aminoàcid que s'enllacen al catió metàl·lic i als àtoms donadors que es troben sobre la cadena lateral. Els descriptors dels cations, electronegativitat de Pauling i relació càrrega/radi iònic de Shannon, es poden



relacionar amb la naturalesa de l'enllaç catió-aminoàcid que es forma, el qual és determinant per al valor de la propietat estudiada.

### ***Constant d'acidesa de compostos orgànics en diferents solvents***

S'ha aplicat per primera vegada l'aproximació QSPR a sistemes multicomponents constituïts per fenols, àcids benzoics, àcids carboxílics alifàtics i anilines en onze solvents diferents.

Els models no lineals deduïts contenen set descriptors, dos dels quals són descriptors externs dels solvents i els altres cinc dels soluts. Els paràmetres estadístics obtinguts són molt bons, amb  $R^2 \approx 0.98$  i  $rmse \approx 0.2-1$  unitats de  $pK_a$ , en totes les famílies.

Els descriptors dels solvents són els mateixos en pràcticament tots els casos:  $\alpha$  de Kamlet i Taft i  $\Delta_{vap}U^\circ$ ; els descriptors dels soluts són semblants en les quatre famílies estudiades, i hi predominen els de tipus quàntic i/o electrostàtic. Aquests descriptors són de tal naturalesa que permeten una bona interpretació de les interaccions solut/solvent que actuen en el procés de dissociació. Així, les interaccions de tipus específic venen descrites per descriptors com l'energia dels orbitals moleculars HOMO i LUMO+1 dels soluts, i la  $\alpha$  de Kamlet i Taft dels solvents; mentre que les de tipus no específic es poden relacionar amb els descriptors energètics dels enllaços O-H que es dissocien, la càrrega parcial sobre l'hidrogen àcid i l'energia de vaporització molar,  $\Delta_{vap}U^\circ$ , del solvent.

La bondat dels models proposats ve recolzada pels resultats trobats amb els diferents subconjunts analitzats: compostos *orto*-substituïts o no; en solvents pròtics o apròtics; i en cadascun dels onze solvents estudiats. En tots els casos tant  $R^2$  com el  $rmse$  són del mateix ordre i iguals als dels conjunts més nombrosos, els corresponents *tsets*. Fins i tot, l'anàlisi particular dels 38 àcids benzoics dels quals es disposa de valors de  $pK_a$ , en almenys sis solvents, proporciona paràmetres estadístics igual de bons.

Els descriptors del model deduït per als àcids benzoics són adequats també per als fenols i, amb el conjunt que conté els dos tipus –793 valors de 228 compostos en deu solvents– els factors estadístics tenen valors anàlegs als obtinguts per a cada família per separat. Això fa que aquest model serveixi per a l'estimació de  $pK_a$  d'un gran nombre de fenols i d'àcids benzoics –aquells que tinguin el mateix tipus de substituents– en els solvents estudiats.

Com a punt final d'aquestes conclusions és adient fer uns comentaris generals sobre la metodologia QSPR i les seves limitacions, tot recordant-ne el seu objectiu fonamental: predir i interpretar els fets experimentals a partir de descriptors moleculars.

Les interpretacions incompletes que sovint dona l'aproximació QSPR no s'han d'atribuir sempre i de manera sistemàtica a limitacions intrínseques de la pròpia metodologia o a la incapacitat dels descriptors per explicar la realitat.

Cal tenir present que uns quants descriptors –que no deixen de ser uns valors numèrics– difícilment poden representar prou bé tota la complexitat molecular i interpretar les propietats físiques i químiques, conseqüència de múltiples interaccions intra- i/o intermoleculars.

L'amplitud del problema fa que les explicacions assignades a un descriptor determinat puguin ser, sovint, provisionals i, aparentment, sense gaire base; tot i així, la bona capacitat predictiva justifica la seva presència en el model. Per altra banda, descriptors amb poc poder predictiu es mantenen en el model, quan són capaços d'interpretar la propietat estudiada.

No s'ha d'oblidar tampoc, que la dificultat de reduir els conceptes químics a uns descriptors no és pot imputar només a la debilitat dels descriptors sinó a què, potser, calen altres maneres de formular aquests conceptes.

S'ha de tenir en compte també, que la bondat i homogeneïtat dels valors experimentals disponibles de les propietats a estudiar no és sempre la millor, fet que en dificulta significativament la seva anàlisi.

En resum, pot concloure's que els objectius proposats: aplicació de la metodologia QSPR a nous conjunts de compostos i a sistemes multicomponents s'han aconseguit de manera prou reeixida i que els models presentats, per a cadascun dels casos descrits, compleixen amb escreix els requisits d'un bon estudi QSPR: contenen un nombre petit de descriptors, són capaços d'estimar amb prou precisió la propietat estudiada i són, també, aptes per donar-ne una interpretació fisicoquímica prou convincent.