

Obtener espectros UV-Vis a diferentes pHs:

- Usar los valores de pka experimentales.
- Considerar la ley de Beer
- Nivel de teoría B3LYP/6-31+G(d), solvente continuo SMD (agua)
- Palabras claves importantes: TD(NStates=6), scrf=(smd,solvent=water)

Comparar con espectros experimentales

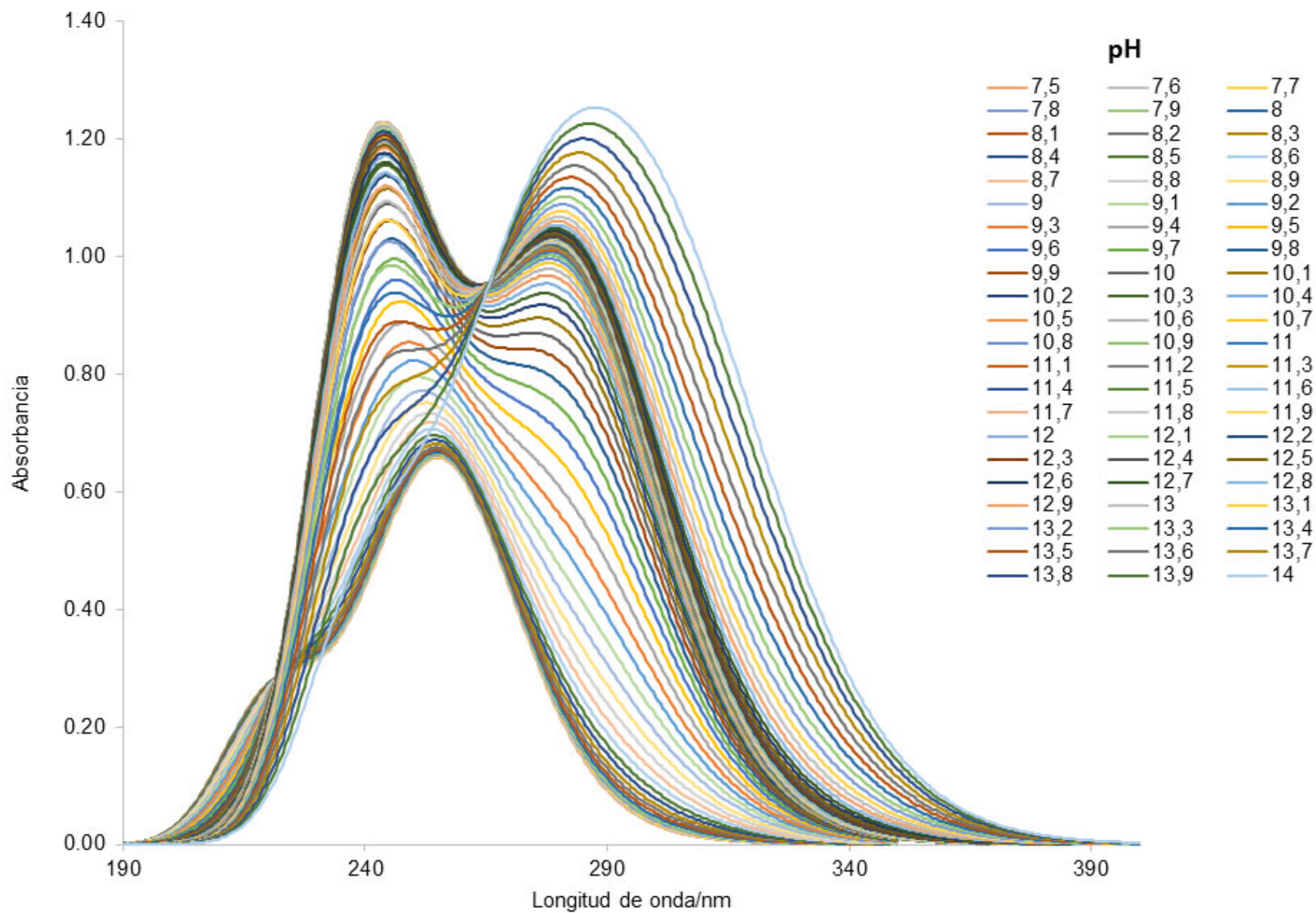


Figura10. Conjunto de espectro de absorción UV en un intervalo de pH de 7.5-14.0 de la molécula 4-metil-catecol calculada con el funcional B3LYP con un nivel de teoría 6-31+G(d) con SMD para modelar medio acuoso.

Constantes de Acidez

$$K_a = e^{-\Delta G_s / RT}$$

1M

$$pK_a = -\log(K_a)$$

$$K_P = K_C (V_M)^{\Delta n}$$

Δn stands for the mole change through reaction (i.e. $\Delta n = -1$ for bimolecular reactions), V_M represents the molar volume, and K_C and K_P are the equilibrium constants in concentration and pressure units, respectively.

Then: $K_C = K_P V_M$

$$e^{-\frac{\Delta G^{1M}}{RT}} = e^{-\frac{\Delta G^{1atm}}{RT}} V_M$$

$$\Delta G^{1M} = \Delta G^{1atm} - RT \ln(V_M)$$

This conversion lowers the ΔG s in 1.89 kcal/mol, for bimolecular reactions, at 298K.



$$\Delta n = +1$$

$$\Delta G^{1M} = \Delta G^{1atm} + RT \ln(V_M)$$

Increases ΔG s in 1.89 kcal/mol, at 298K.

*Estrategias computacionales para el
cálculo de pK_a s confiables.
¿Dónde estamos y hacia dónde vamos?*



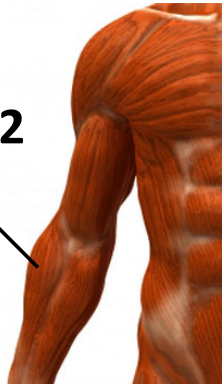


¿Por qué seguir torturándonos?

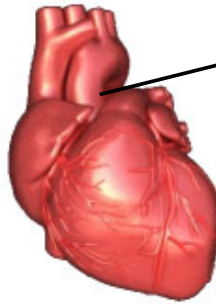
- Muchos compuestos presentan equilibrios ácido-base en solución acuosa.
- La forma predominante puede cambiar la reactividad.
- Fármacos (acción, solubilidad, velocidad de absorción, dosis y formas de administración)

pH en diferentes órganos y fluidos humanos

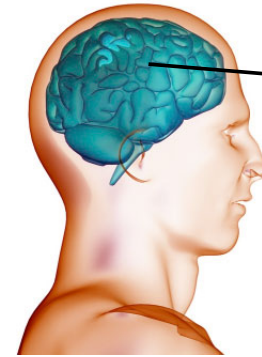
Músculos
pH = 6.9-7.2



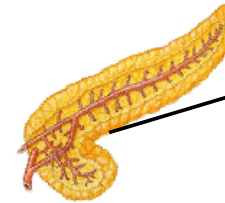
Corazón
pH = 7.0-7.4



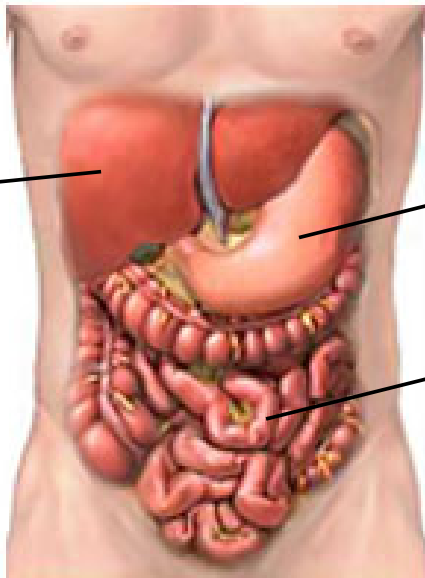
Cerebro
pH = 7.1



Páncreas
pH = 8.0-8.3



Hígado
pH = 7.2



Estómago
pH = 1.0-3.5

Intestinos
delgado pH = 5-7
Grueso pH = 7.5-8



Orina
pH = 4.5-8.0

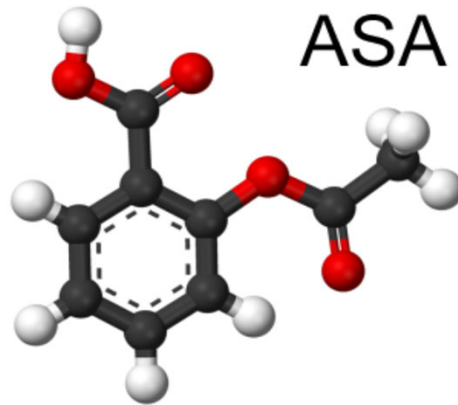


Saliva
pH = 6.0-7.4



Sangre
pH = 7.35-7.45

Ej:



La aspirina tiene un $pK_a = 3.5$

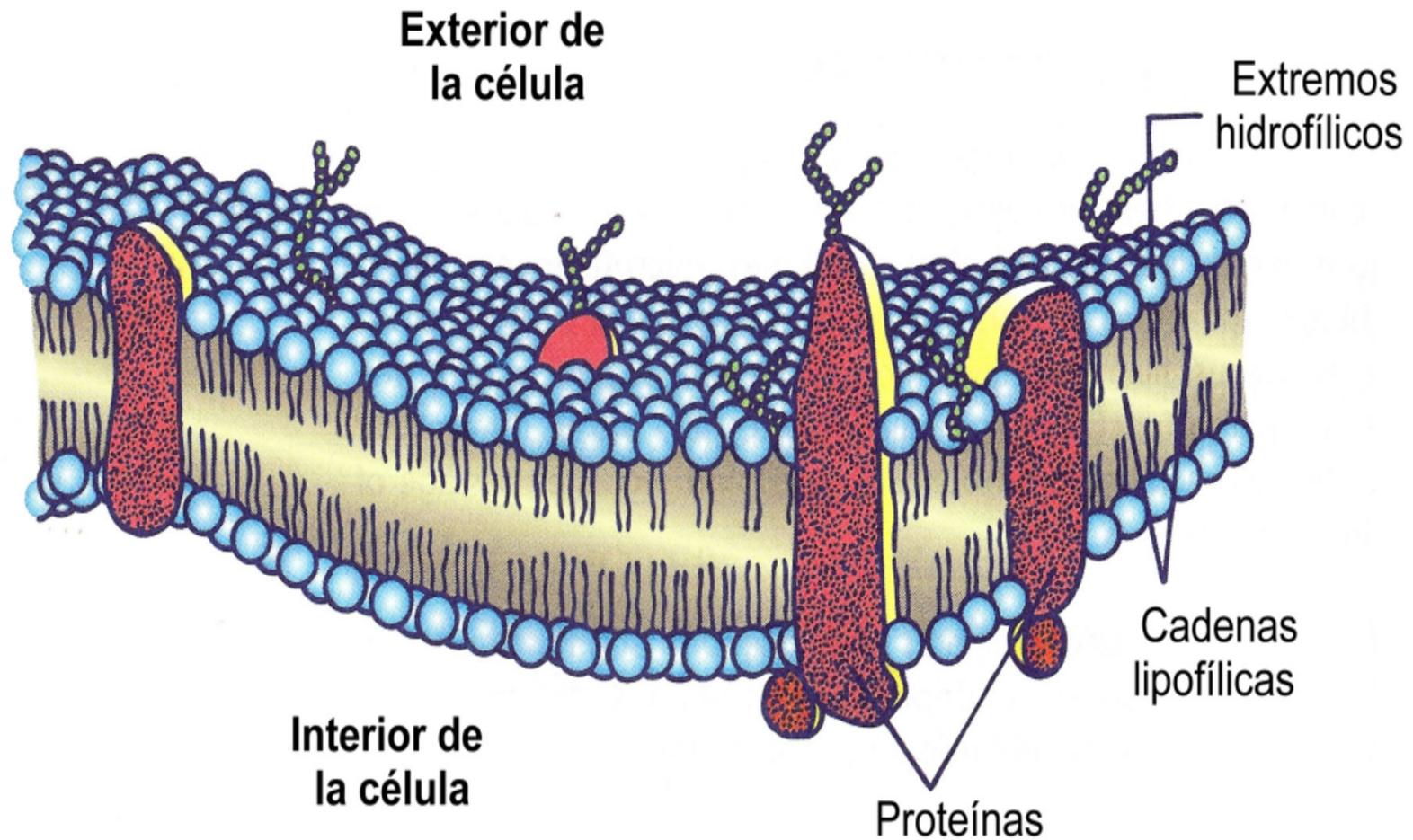
Su forma protonada (mayoritaria a pHs menores que 3.5) corresponde a una especie neutra, mientras que su forma deprotonada (mayoritaria a pHs mayores que 3.5) corresponde a un monoanión.

El pH fisiológico es 7.35-7.45

Por tanto la aspirina estará casi completamente ionizada en estas condiciones.

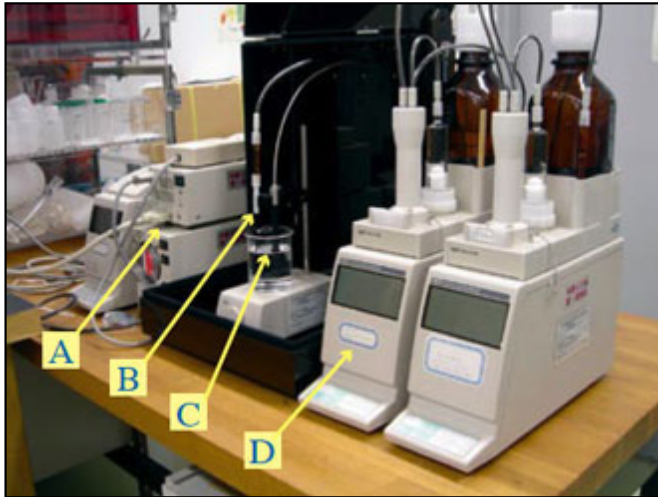
Mientras que debido al pH del estómago 1-3

En estas condiciones estará mayormente no ionizada, incrementando así su facilidad de absorción.



Dependiendo de su forma de administración los fármacos necesitan atravesar una o más membranas celulares para llegar al torrente sanguíneo y alcanzar su sitio de acción. Una característica que tienen en común todas las membranas celulares es la bicapa fosfolipídica, de un grosor aproximado de 10nm. En esta frontera natural los extremos hidrofílicos se acomodan hacia el exterior de la membrana y las cadenas lipofílicas hacia su interior. Esta organización conlleva a una distribución tipo sándwich con dos capas hidrofílicas rodeando la zona central hidrofóbica. Cruzando esta bicapa o unidas a ella se encuentran glicoproteínas, que actúan como canales de iones, receptores, o mensajeros.

¿No es suficiente con las técnicas experimentales?



Valoraciones
espectrofotométricas



Electroforesis
capilar

Dificultades:

- especies de vida media breve
- ácidos muy débiles o muy fuertes
- orden de deprotonación



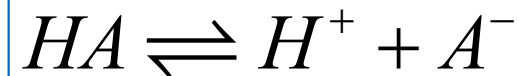
¿Qué pasa con la Química Computacional y los equilibrios de acidez?

“With respect to acid/base equilibria, the prediction of an absolute pK_a by any first-principles theory **remains extraordinarily hard**, in part because of the difficulty in computing accurate gas-phase deprotonation free energies and also because of the large magnitude of solvation effects for charged species.”

“**Typical accuracies in the computed quantities are rarely better than 2-5 kcal/mol**, which results in absolute errors of possibly several log units in pK_a s a *disappointing result given* the accuracy possible from experimental measurement (when such measurements can be made).”

“Fortunately, **trends in pK_a s across similar compounds tend to be much more accurately predicted**”

¿Cuáles son los problemas fundamentales?



- El número y tipo de especies cargadas no se conserva a ambos lados del equilibrio.
- No hay métodos que reproduzcan energías en solución comparables a G_n o CBS para fase gas.
- El tamaño de la mayor parte de los sistemas de interés.
- Modelos continuos de solvente: no interacciones específicas soluto-solvente (corto alcance).
- Modelos discretos: es necesario incluir muchas moléculas de solvente para reproducir interacciones de largo alcance ∴ computacionalmente inviables (con DFT or WFT).

Un error de **1.36 kcal/mol** en $\Delta G \rightarrow$
error de **1** unidad de pK_a

¿Dónde estamos?

Desviaciones de 2 unidades de pKa se consideran con exactitud aceptable.



Búsqueda de estrategias que maximicen la cancelación de errores

- Diferentes formas de modelar el solvente
- Diferente número de moléculas de agua explícitas
- Diferentes esquemas de reacción
- Diferentes niveles de cálculo (DFT, WFT, ...)
- Diferentes funciones de base
- Separación gas-solución (ciclos termodinámicos, Hess)

Número de variaciones $\rightarrow \infty$

??

Divide et impera



¿Dónde estamos?

Solvente

- Modelos Continuos
- Modelos discretos
- Modelos mixtos

Modelos Continuos y Mixtos:

Optimización de Geometrías
Cálculos de Frecuencias
En fase gas
(ΔG gas)
+
Punto simple en solución
(ΔG solvatación)

Ciclos termodinámicos

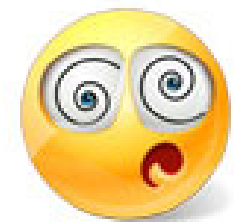
Optimización de Geometrías
Cálculos de Frecuencias
En solución
(ΔG sol)

Algunos Esquemas de reacción :

Scheme	Equilibrium	$pK_a =$
<i>A</i>	$HA \leftrightarrow H^+ + A^-$	$\frac{\Delta G_s}{RT \ln(10)}$
<i>B</i>	$HA + H_2O \leftrightarrow H_3O^+ + A^-$	$\frac{\Delta G_s}{RT \ln(10)} - \log [H_2O]$
<i>C1</i>	$HA + 2H_2O \leftrightarrow H_3O^+ + A^- (H_2O)$	$\frac{\Delta G_s}{RT \ln(10)} - 2\log [H_2O]$
<i>C2</i>	$HA + 3H_2O \leftrightarrow H_3O^+ + A^- (2H_2O)$	$\frac{\Delta G_s}{RT \ln(10)} - 3\log [H_2O]$
<i>C3</i>	$HA + 4H_2O \leftrightarrow H_3O^+ + A^- (3H_2O)$	$\frac{\Delta G_s}{RT \ln(10)} - 4\log [H_2O]$
<i>D</i>	$HA + Ref^- \leftrightarrow A^- + HRef$	$\frac{\Delta G_s}{RT \ln(10)} + pK_a(HRef)$
<i>E1</i>	$HA + OH^- (3H_2O) \leftrightarrow A^- (H_2O) + 3H_2O$	$\frac{\Delta G_s}{RT \ln(10)} + 14 + 3\log [H_2O]$
<i>E2</i>	$HA + OH^- (3H_2O) \leftrightarrow A^- (2H_2O) + 2H_2O$	$\frac{\Delta G_s}{RT \ln(10)} + 14 + 2\log [H_2O]$
<i>E3</i>	$HA + OH^- (3H_2O) \leftrightarrow A^- (3H_2O) + H_2O$	$\frac{\Delta G_s}{RT \ln(10)} + 14 + \log [H_2O]$

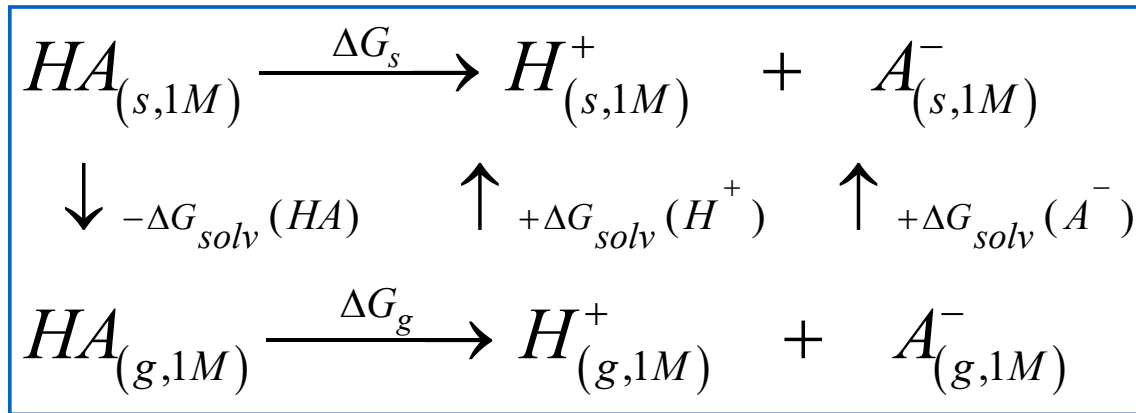
F1	$HA + 4H_2O \leftrightarrow H_3O^+ (3H_2O) + A^-$	$\frac{\Delta G_s}{RT \ln(10)} - 4 \log [H_2O]$
F2	$HA + 5H_2O \leftrightarrow H_3O^+ (3H_2O) + A^- (H_2O)$	$\frac{\Delta G_s}{RT \ln(10)} - 5 \log [H_2O]$
F3	$HA + 6H_2O \leftrightarrow H_3O^+ (3H_2O) + A^- (2H_2O)$	$\frac{\Delta G_s}{RT \ln(10)} - 6 \log [H_2O]$
F4	$HA + 7H_2O \leftrightarrow H_3O^+ (3H_2O) + A^- (3H_2O)$	$\frac{\Delta G_s}{RT \ln(10)} - 7 \log [H_2O]$
CN1	$HA(H_2O) + H_2O \leftrightarrow H_3O^+ + A^- (H_2O)$	$\frac{\Delta G_s}{RT \ln(10)} - \log [H_2O]$
CN2	$HA(H_2O) + 2H_2O \leftrightarrow H_3O^+ + A^- (2H_2O)$	$\frac{\Delta G_s}{RT \ln(10)} - 2 \log [H_2O]$
CN3	$HA(H_2O) + 3H_2O \leftrightarrow H_3O^+ + A^- (3H_2O)$	$\frac{\Delta G_s}{RT \ln(10)} - 3 \log [H_2O]$
FN1	$HA(H_2O) + 4H_2O \leftrightarrow H_3O^+ (3H_2O) + A^- (H_2O)$	$\frac{\Delta G_s}{RT \ln(10)} - 4 \log [H_2O]$
FN2	$HA(H_2O) + 5H_2O \leftrightarrow H_3O^+ (3H_2O) + A^- (3H_2O)$	$\frac{\Delta G_s}{RT \ln(10)} - 5 \log [H_2O]$
FN3	$HA(H_2O) + 6H_2O \leftrightarrow H_3O^+ (3H_2O) + A^- (3H_2O)$	$\frac{\Delta G_s}{RT \ln(10)} - 6 \log [H_2O]$
EN1	$HA(H_2O) + OH^- (3H_2O) \leftrightarrow A^- (H_2O) + 4H_2O$	$\frac{\Delta G_s}{RT \ln(10)} + 14 + 4 \log [H_2O]$
EN2	$HA(H_2O) + OH^- (3H_2O) \leftrightarrow A^- (2H_2O) + 3H_2O$	$\frac{\Delta G_s}{RT \ln(10)} + 14 + 3 \log [H_2O]$
EN3	$HA(H_2O) + OH^- (3H_2O) \leftrightarrow A^- (3H_2O) + 2H_2O$	$\frac{\Delta G_s}{RT \ln(10)} + 14 + 2 \log [H_2O]$

¿y esos no son todos?!!!



Los más usados:

Esquema A:



$$pKa = \frac{\Delta G_s}{2.303RT}$$

$$\Delta G_{sol} = \Delta G_{gas} + \Delta\Delta G_s$$

$$\Delta\Delta G_s = \Delta G_s(A^-) + \Delta G_s(H^+) - \Delta G_s(HA)$$

exp

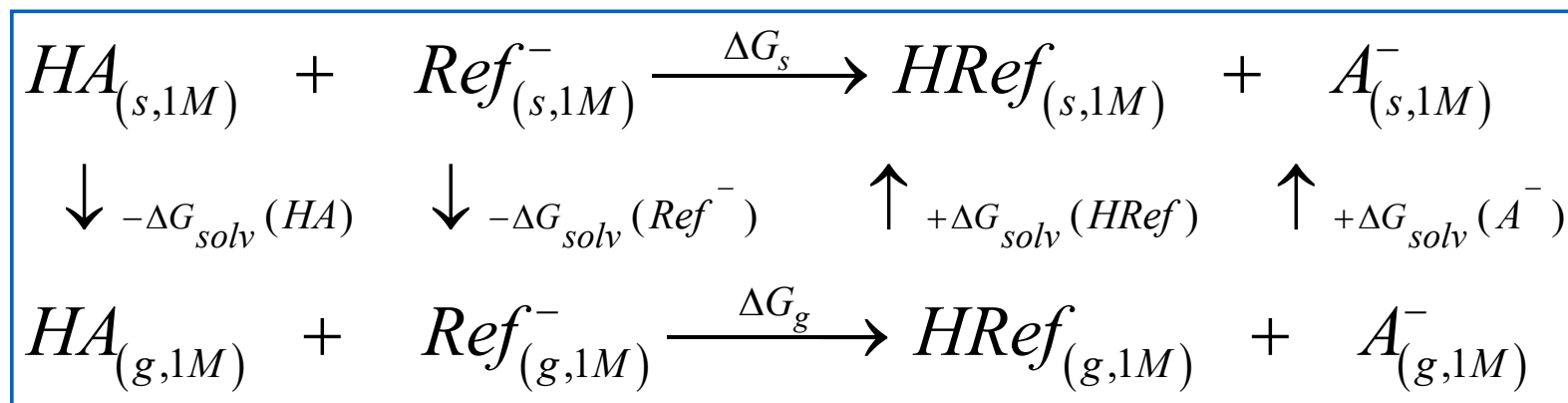
$$\left\{ \begin{array}{l}
 \Delta G_{gas}(H^+) = -6.28 \text{ kcal/mol} \quad (1 \text{ atm}) \\
 \Delta G_{gas}(H^+) = -4.39 \text{ kcal/mol} \quad (1 \text{ M}) \\
 \Delta G_s(H^+) = -264.61 \text{ kcal/mol}
 \end{array} \right. + RT \ln(V_M) = 1.89 \text{ kcal/mol}$$

-259 to -265 kcal/mol

Los más usados:

Esquema D:

Errores pequeños



$$pKa = \frac{\Delta G_s}{2.303RT} + pKa[HRef]$$

Esquemas A y D:

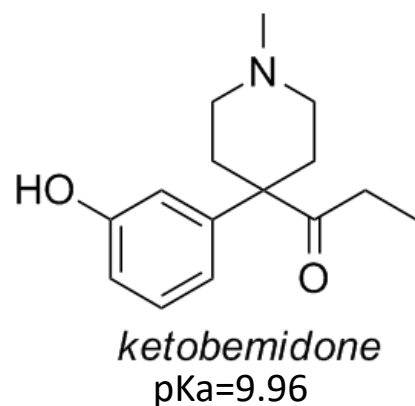
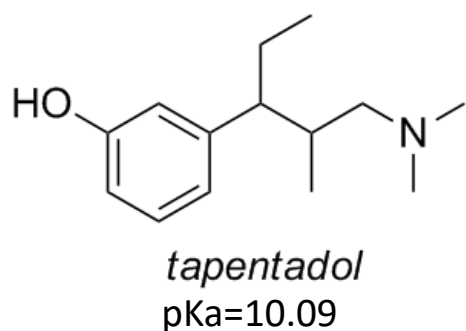
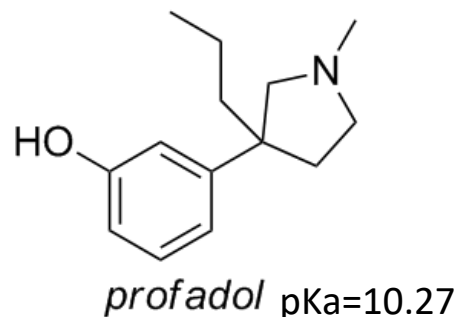
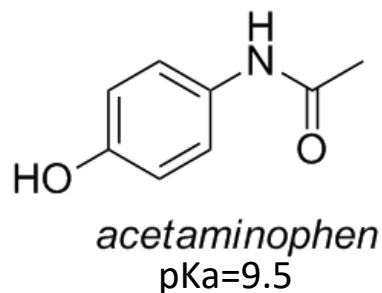
Dan resultados razonablemente buenos, especialmente **D**
(máxima cancelación de errores)

pero...

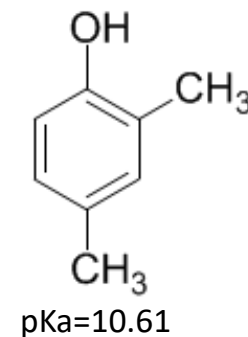
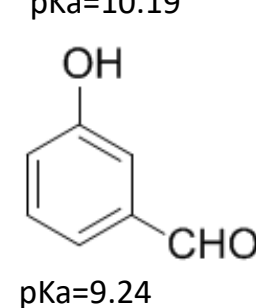
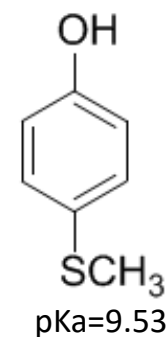
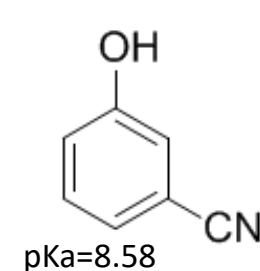
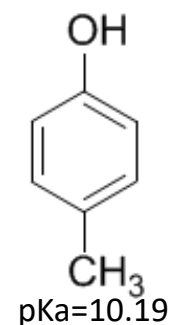
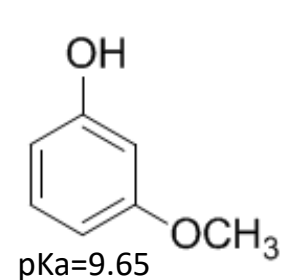
Ambos dependen de datos experimentales

Caso particular: fenoles

Anti-inflamatorios no esteroidales



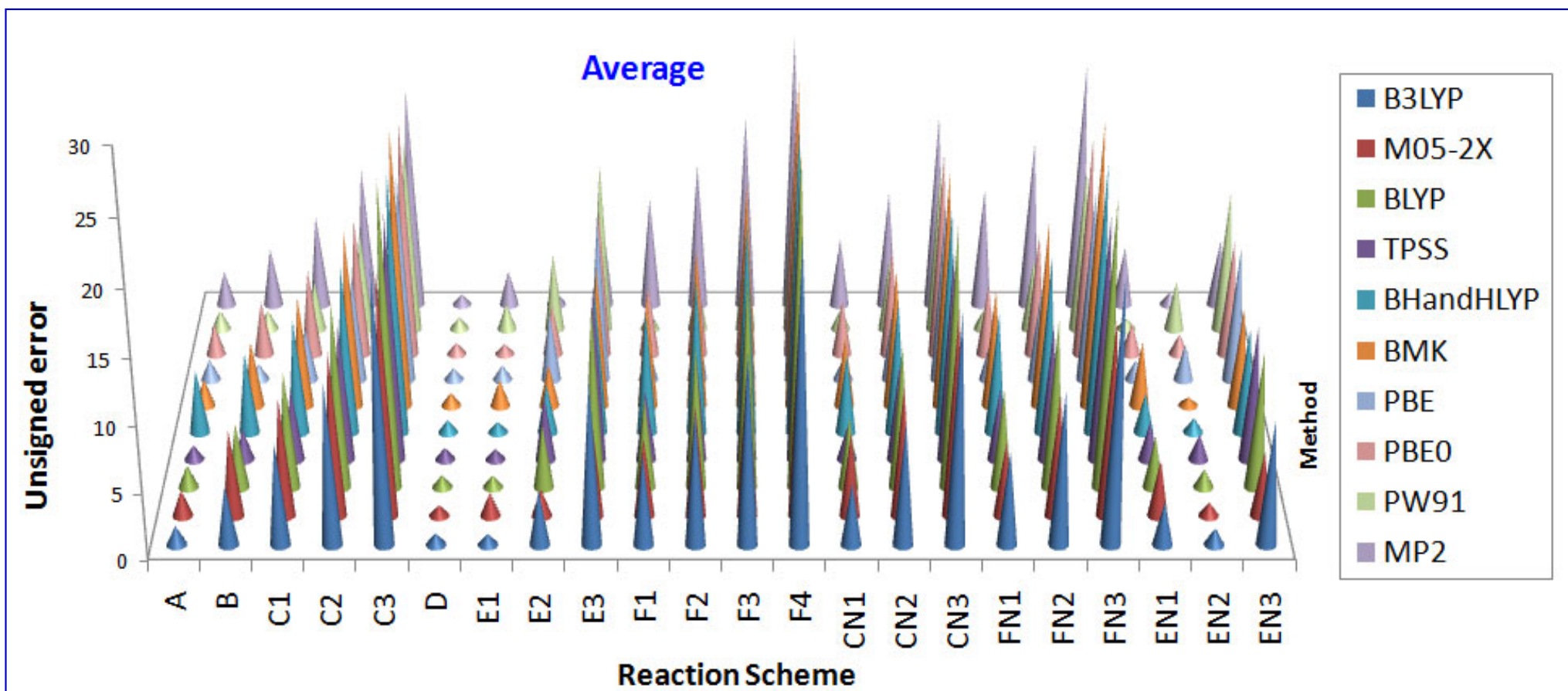
Otros fenoles



Conjunto de funciones de base: 6-311++G(d,p)
Solvente: PCM, radii=UAHF (HF/6-31+g(d))

Ciclos termodinámicos
22 esquemas de reacción

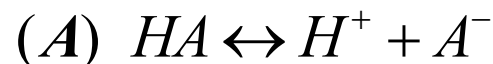
BLYP, B3LYP, BHandHLYP, PBE,
PBE0, PW91, BMK, TPSS,
M05-2x, MP2

Caso particular: fenoles

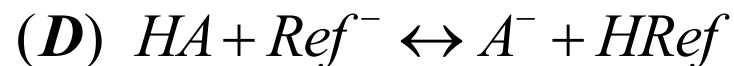
La calidad de los resultados depende más de la selección del esquema de reacción que del método de cálculo

Caso particular: fenoles

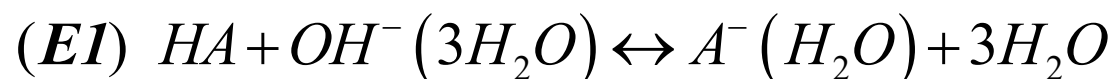
Los mejores resultados se obtuvieron con los esquemas:



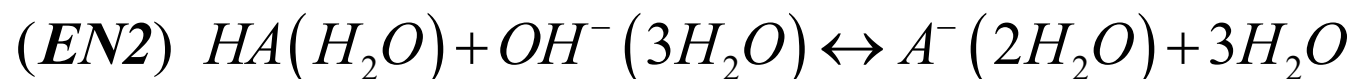
Dep. exp.



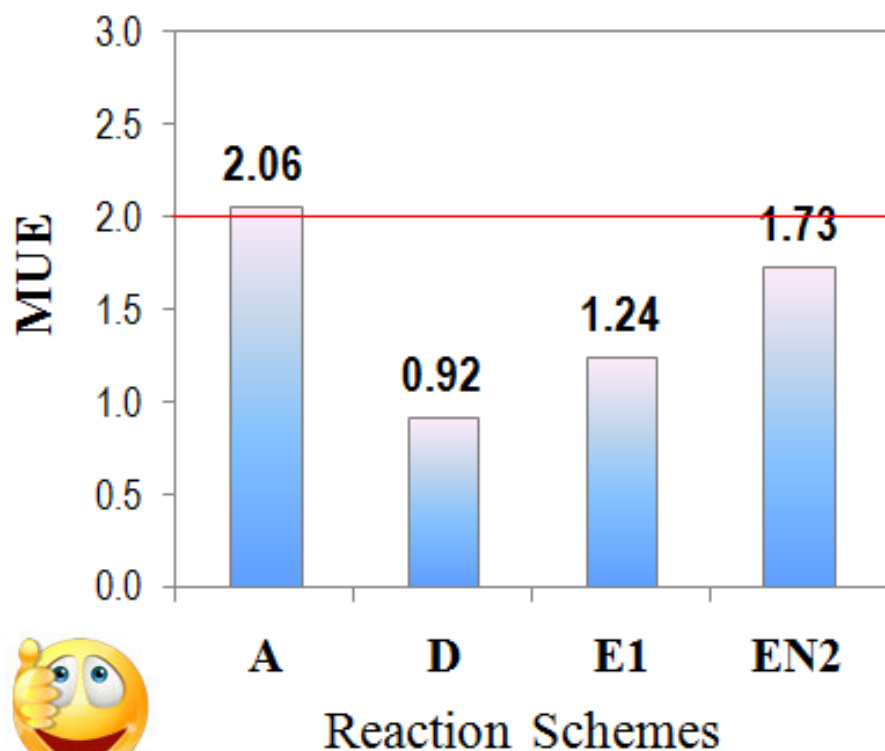
Dep. exp.



Indep. exp.



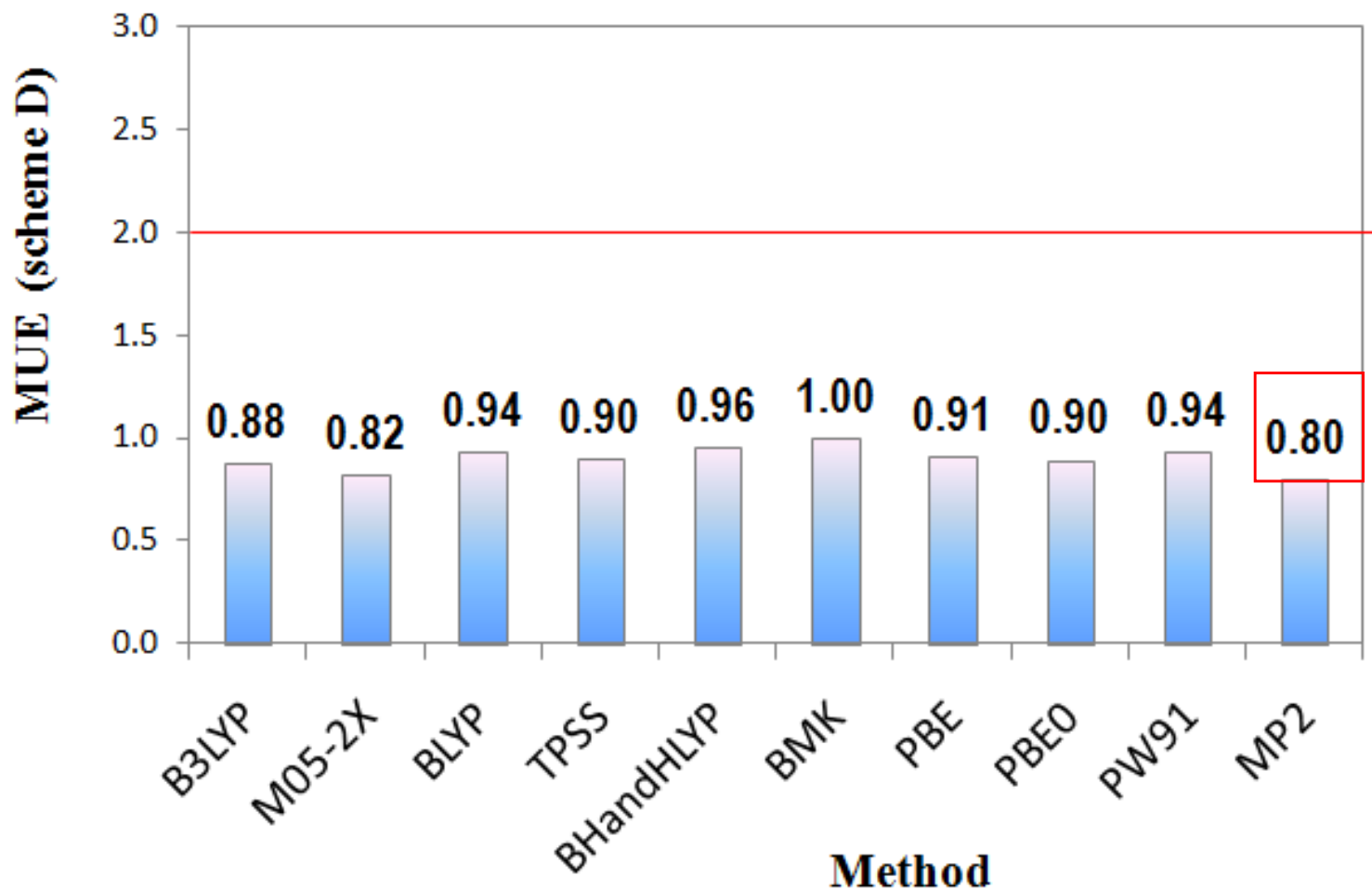
Indep. exp.



D, *E1* y *EN2* conservan el número y tipo de especies cargadas a ambos lados del equilibrio.

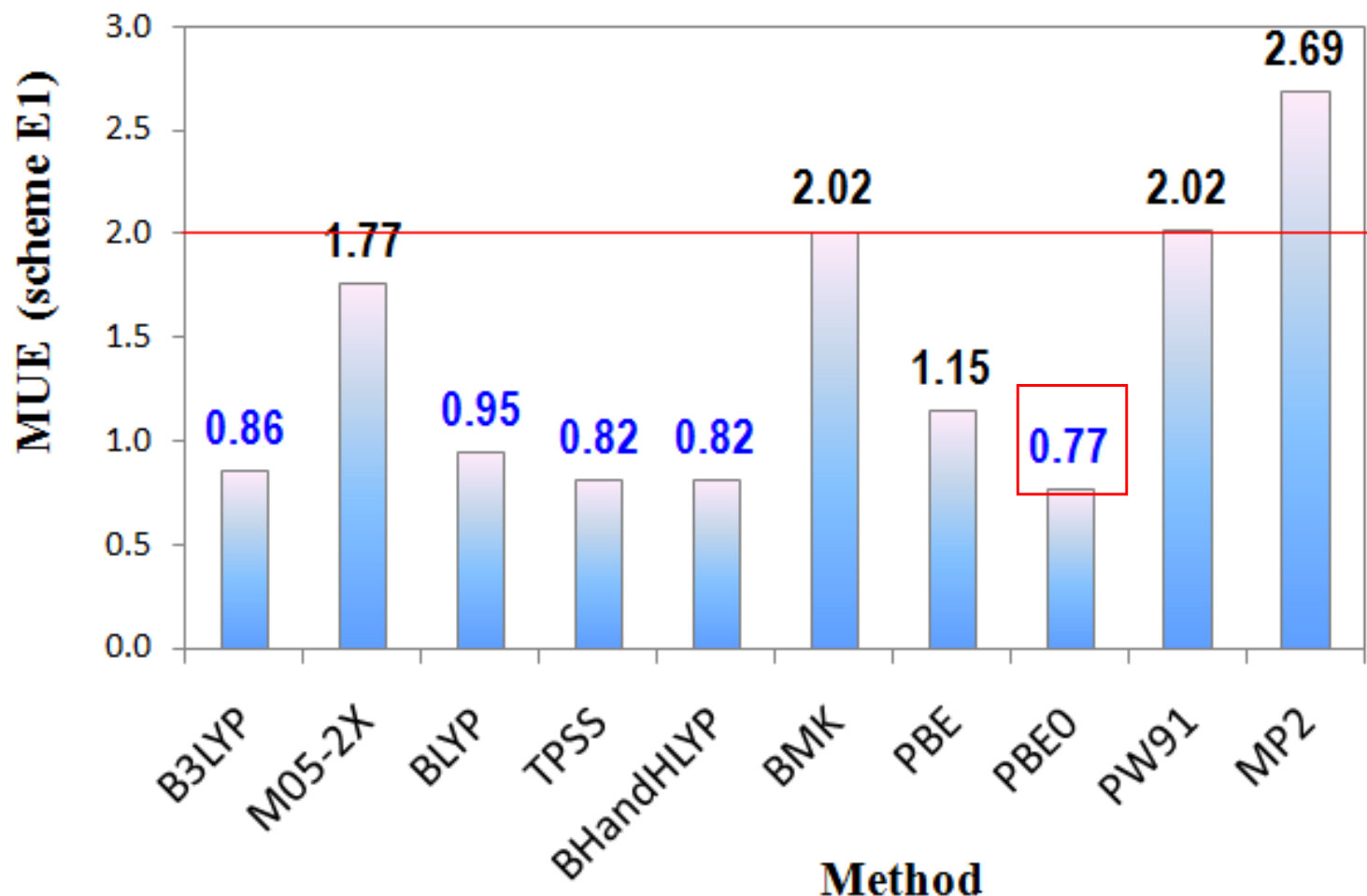
Esta parece ser la característica más importante a tener en cuenta al seleccionar el esquema de reacción.



Caso particular: fenoles

La calidad de los resultados con el esquema D es prácticamente independiente del método de cálculo utilizado y en todos los casos $MUE \leq 1$.

Es muy dependiente de la exactitud del pK_a experimental de Href y de cuan parecidos sean el HRef y el que se está estudiando.

Caso particular: fenoles

Es más sensible al método de cálculo utilizado.

MP2 → los peores resultados

Los mejores resultados ($MUE \leq 1$) se obtuvieron con PBE0, TPSS, BHandHLYP, B3LYP y BLYP, en se orden.

Caso particular: fenoles

Conclusiones:

- La calidad de los resultados obtenidos depende fuertemente del esquema de reacción utilizado. Los mejores resultados se obtienen para esquemas donde el tipo y número de especies cargadas se conserva a ambos lados del equilibrio.
- Incluir al menos una molécula de agua explícita (anion) mejora los resultados, pero incrementar su número no necesariamente mejora los resultados.
- Los dos esquemas con los que se obtienen las menores desviaciones con respecto a los datos experimentales son ***D*** y ***E1***.
- El esquema ***D*** depende de datos experimentales (pKa de un sistema similar al estudiado).
- Se recomienda el uso del esquema ***E1*** (independiente de los experimentos), en combinación con los funcionales PBE0, TPSS o BHandHLYP. Con PBE0 se obtuvo un valor de MUE = 0.77, muy inferior a la exactitud aceptada actualmente (2 unidades de pKa) y menor al obtenido con el esquema ***D***.
- Con el uso de esta estrategia (PBE0+E1) se pueden obtener valores de pKas confiables para fenoles de tamaño mediano, a costos computacionales razonables.
- Estas conclusiones son válidas para PCM (UAHF) y ciclos termodinámicos.

¿Hacia dónde vamos?

2011

A corto y mediano plazos (CANCELACIÓN DE ERRORES):

- Con la misma estrategia utilizada para fenoles (PCM-UAHF, ciclos termodinámicos) identificar los esquemas de reacción y funcionales que reproducen mejor los valores de pK_a para Aminas y Ácidos carboxílicos.
- Cambiando la estrategia a todo en solución (SMD), identificar las mejores combinaciones esquema-funcional .
- Hacer generalizaciones, identificando protocolos con errores menores a 1 unidad de pK_a (con los métodos disponibles actualmente)
- Encontrar otros protocolos computacionales que no se basen en esquemas de reacción.

A largo plazo (MEJORES MÉTODOS Y MODELOS):

- Mejorar los métodos de solvente continuo.
- Mejorar la eficiencia de los códigos (o la capacidad de cómputo)

LO MÁS IMPORTANTE:



No desanimarse,



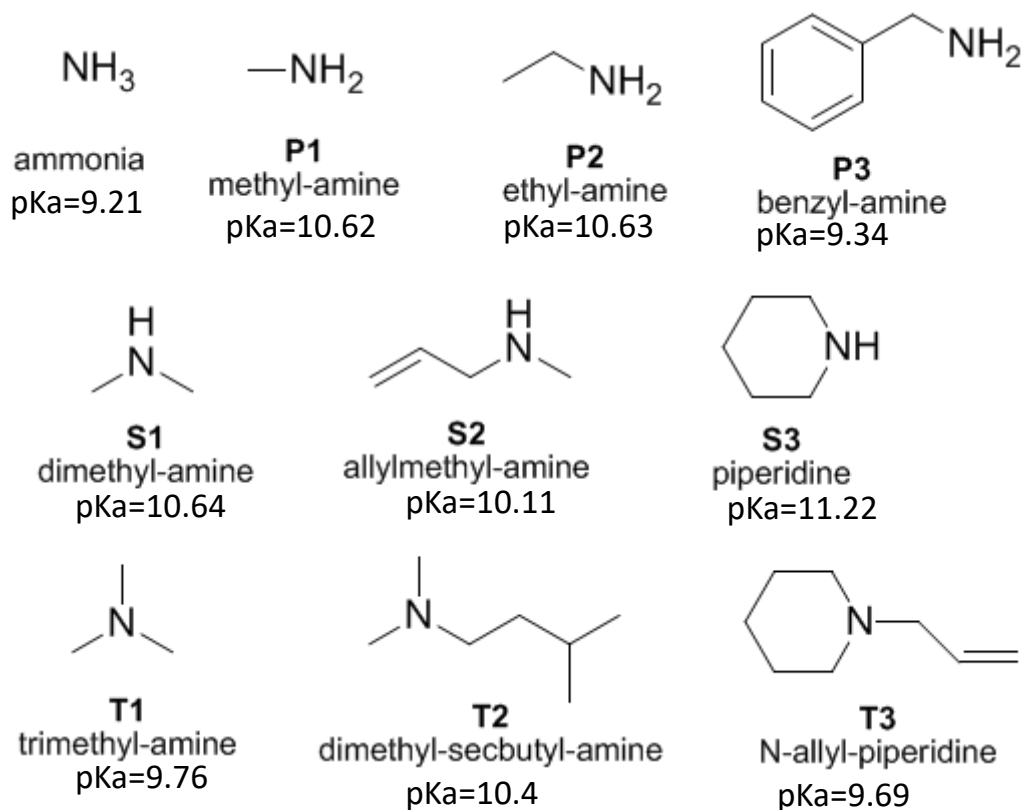
No desesperarse,

El único modo de lograr el éxito es seguir intentando!

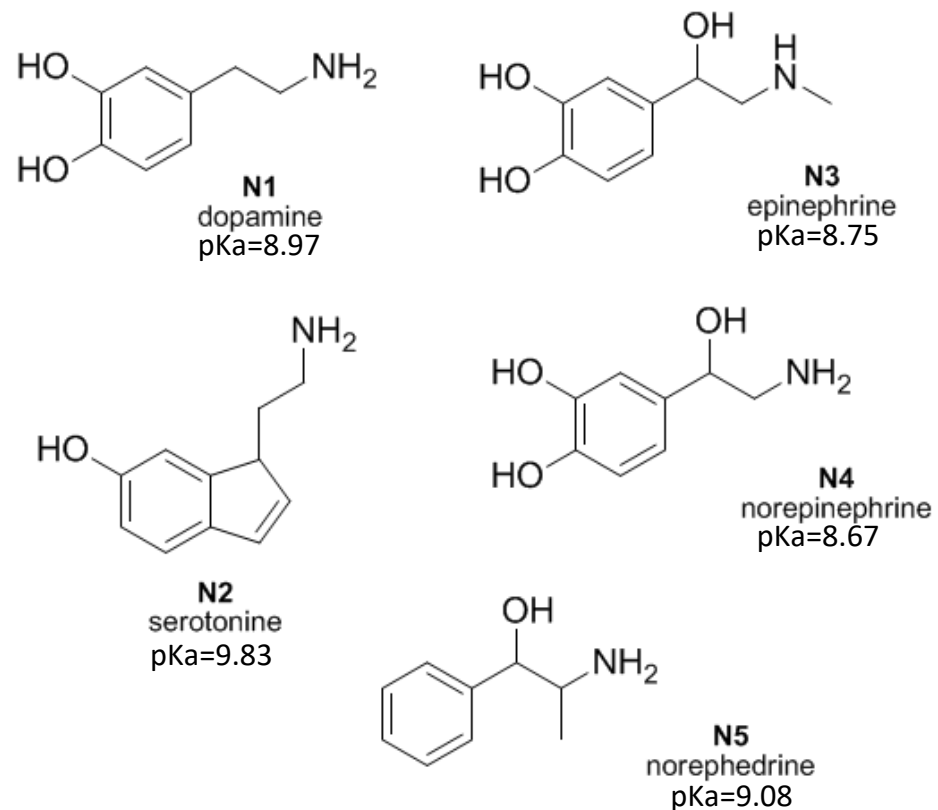


Aminas

Aminas primarias, secundarias y terciarias



Neurotransmisores



Conjunto de funciones de base: 6-311++G(d,p)

Solvente: PCM, radii=UAHF (HF/6-31+g(d))

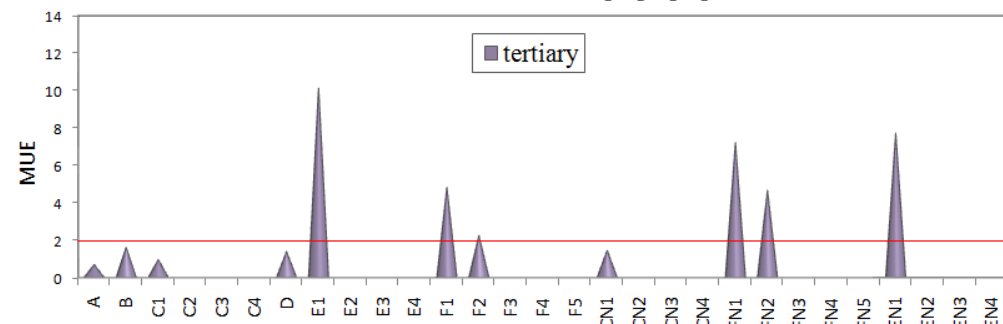
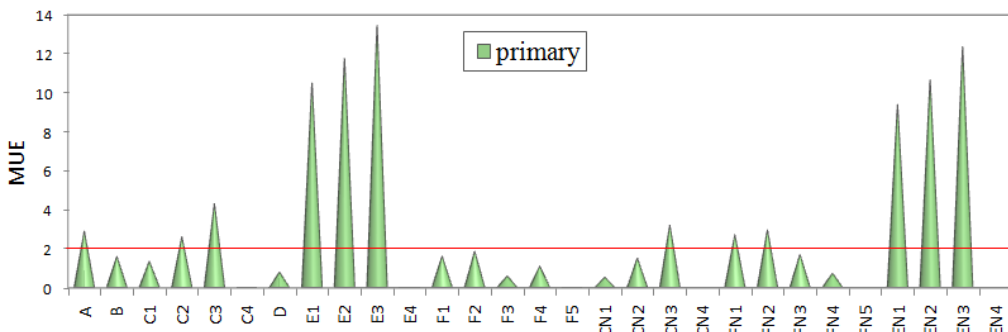
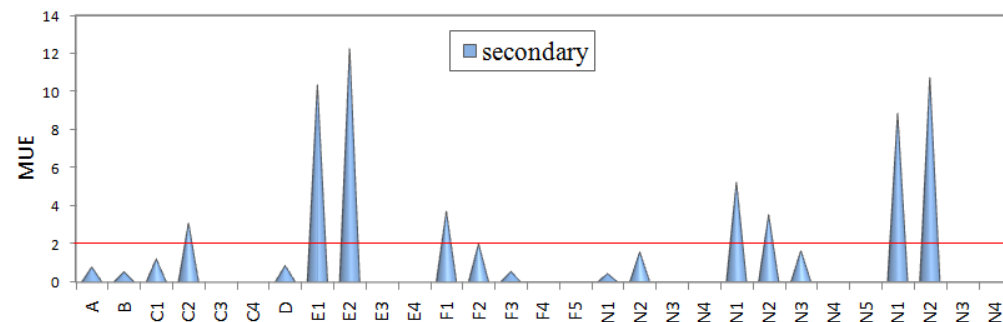
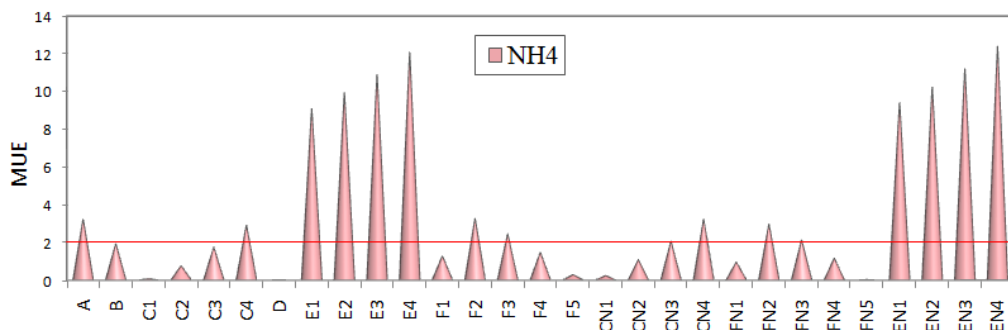
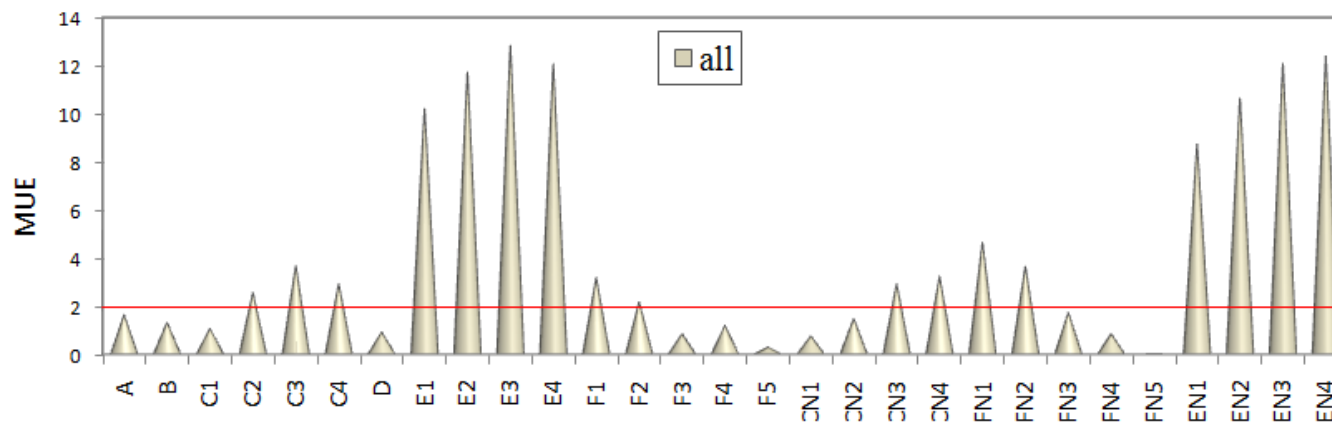
Ciclos termodinámicos

29 esquemas de reacción

BHandHLYP, BLYP, B3LYP, PBE,

PBE0, PW91, BMK,

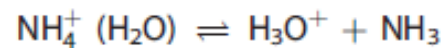
TPSS, M05-2X, and MP2

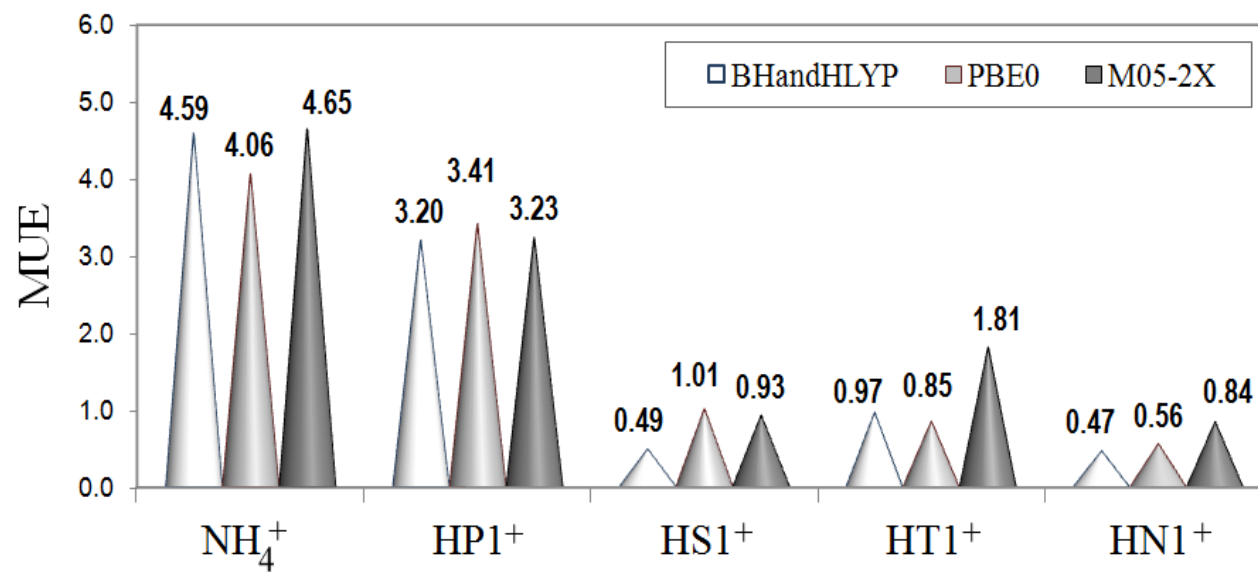
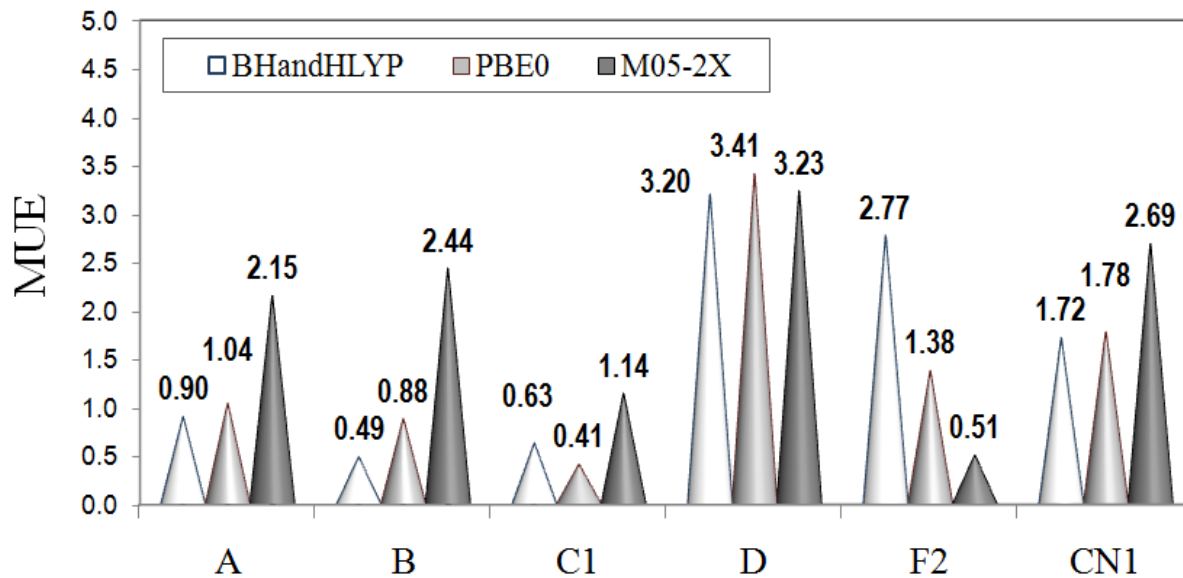


Los mejores (aplicables a cualquier amina): **C1,**

D,

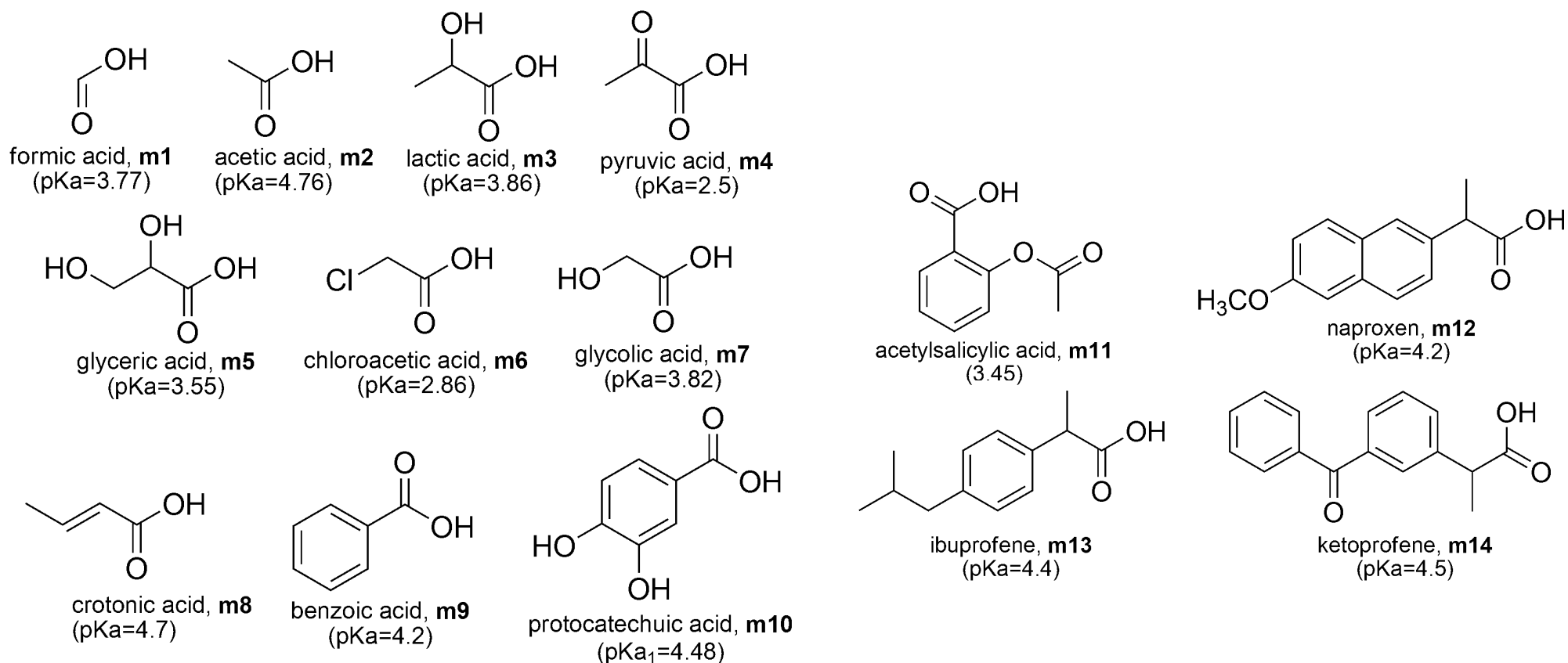
y CN1





Los mejores esquemas son diferentes a los de fenoles y D puede fallar cuando se complica el sistema....





Conjunto de funciones de base: 6-311++G(d,p)

Solvente: PCM, radii=UAHF (HF/6-31+g(d))

Ciclos termodinámicos

22 esquemas de reacción

BLYP, B3LYP, BHandHLYP, PBE,

PBE0, PW91,

BMK, TPSS, M05-2X

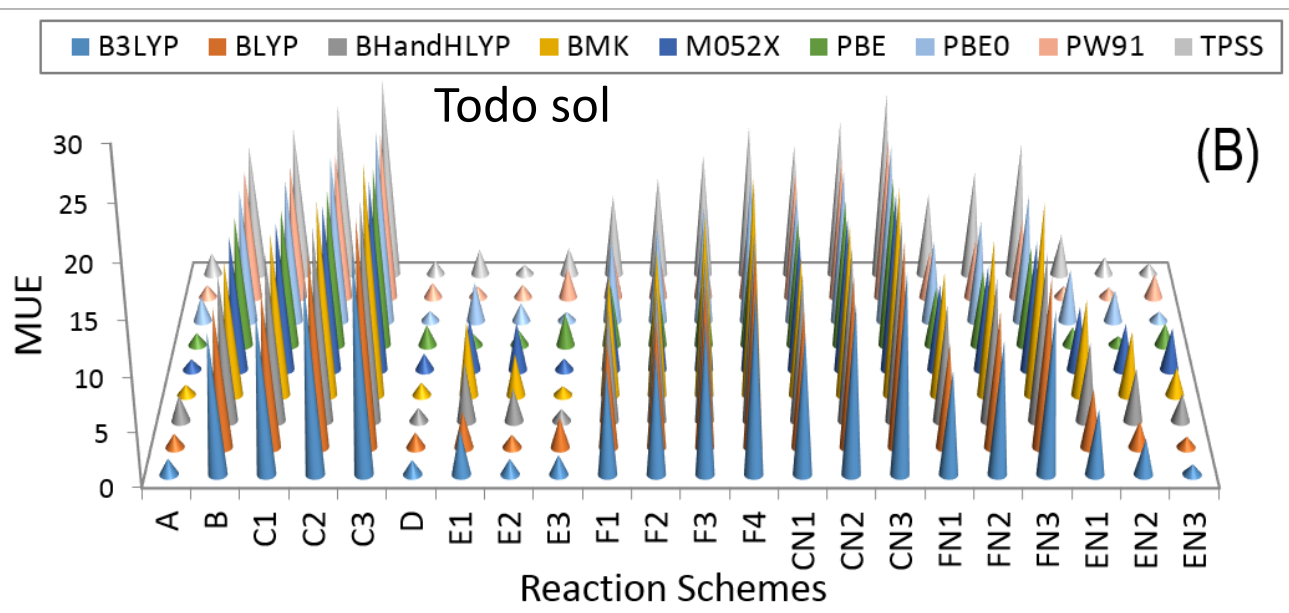
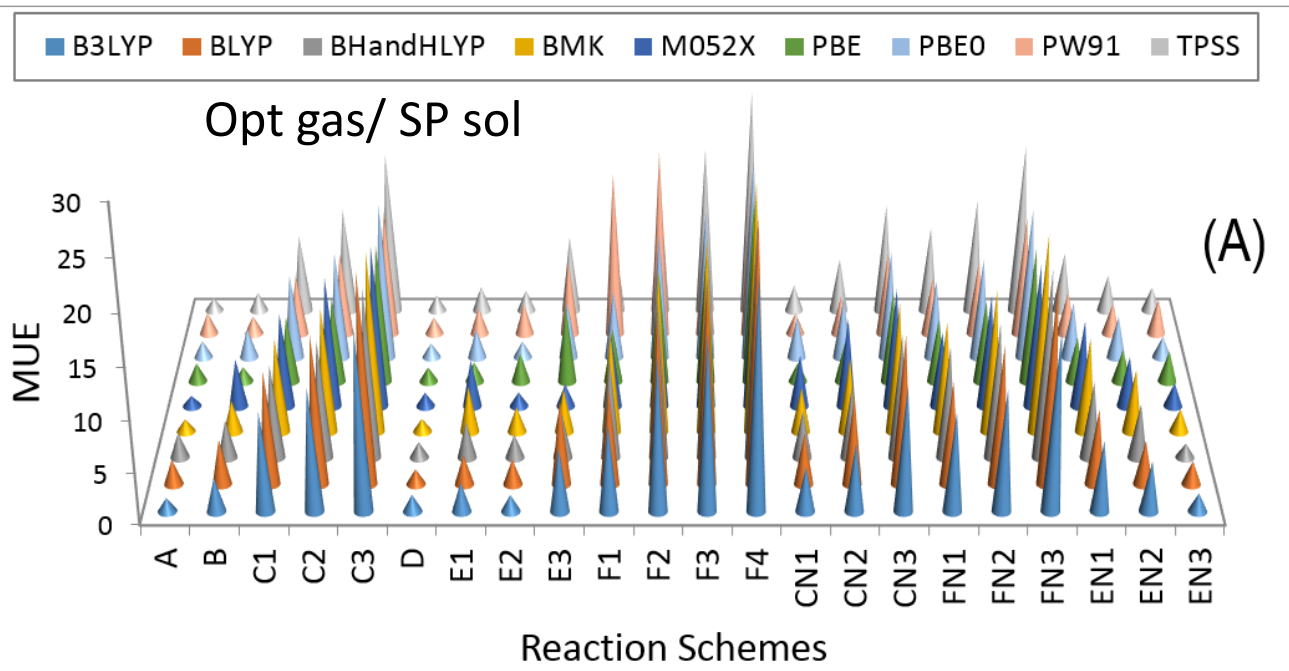
Ácidos Carboxílicos

Constantes de Acidez

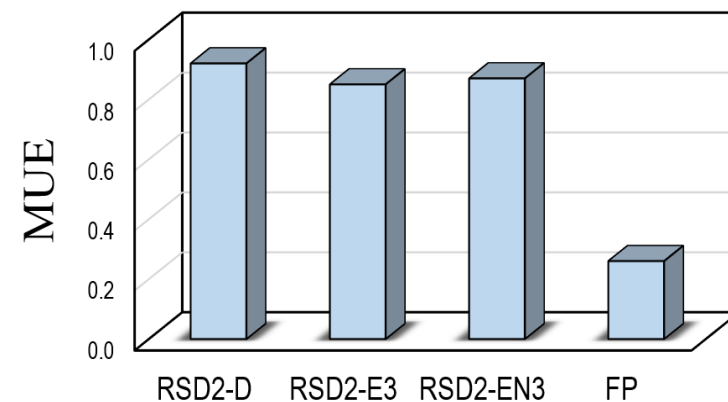
Los mejores esquemas son similares a los de fenoles

Fenoles: A, D, E1 y EN2 G/SP sol
Ác. Carb: A, D, E2 y EN3 G/SP sol
Ác. Carb: A, D, E3 y EN3 Todo sol

Se necesitan más aguas explícitas para ácido: lógico



SMD, PBE0/6-311++G(d,p)



Se ve prometedor



Usando la aproximación de Parámetros Ajustados:

Se ha usado con diferentes tipos de datos calculados, por ejemplo pKa, entalpía, energía libre de Gibbs y cargas.

$$pK_a^{\text{exp}} = m X^{\text{calc}} + C_0 \quad \text{Usando un } \mathbf{training\ set}$$

$$pK_a^{\text{calc,corr}} = m X^{\text{calc}} + C_0 \quad \text{También se usa para validación usando un } \mathbf{test\ set} \text{ que no debe incluir ninguna de las moléculas en el } \mathbf{training\ set}.$$

$$pK_a^{\text{exp}} = m \Delta G_{s(\text{BA})} + C_0$$

$$\Delta G_{s(\text{BA})} = G_{s(\text{B})} - G_{s(\text{A})}$$

PBE, PBE0, LC- ω PBE, TPSS, TPSSh, B3LYP, cam-B3LYP, BHandHLYP, BMK, MPWB1K, M05-2X, M06-2X, M06, ω B97X ω B97X-D, B2PLYP, and B2PLYP-D; one wave function method (MP2), one complete basis set method (CBS-QB3), and one complex method for computing very accurate energies (G3)

Basis sets 6-31+G(d), 6-311+G(d), 6-311+G(d,p) and 6-311++G(d,p)

Training set: 20 moléculas

Test set: 10 moléculas

Training set with experimental pKa ranging: phenols (6.33–10.31), carboxylic acids (-0.2–4.87), amines (0.97–11.125) and thiols (4.7 – 11.2).

Table 1. Training Sets Used to Obtain the m and C_0 Parameters and Experimental pK_a Values

phenol	pK_a^{exp}	carboxylic acid	pK_a^{exp}	amine	pK_a^{exp}
phenol	9.997 ^a	fomic	3.744 ^a	methylamine	10.632 ^a
3-aminophenol	10.02 ^a	acetic	4.757 ^a	methoxyamine	4.62 ^a
4-aminophenol	10.30 ^b	chloroacetic	2.862 ^a	2-propen-1-amine	9.52 ^a
2-chlorophenol	8.531 ^a	oxalic	1.250 ^a	aminoacetonitrile	5.15 ^a
3-chlorophenol	9.125 ^a	methoxyacetic	3.570 ^a	dimethylamine	10.774 ^a
4-chlorophenol	9.426 ^a	prop-2-enoic	4.258 ^a	morpholine	8.492 ^a
<i>o</i> -cresol	10.31 ^a	trichloroacetic	0.660 ^b	piperidine	11.125 ^a
<i>m</i> -cresol	10.095 ^a	<i>trans</i> -2-butenoic	4.686 ^a	trimethylamine	9.799 ^a
<i>p</i> -cresol	10.269 ^a	3-butenoic	4.342 ^a	1-(2-propenyl)piperidine	9.68 ^c
2-nitrophenol	7.230 ^a	benzoic	4.202 ^a	phenilamine	4.601 ^a
3-nitrophenol	8.360 ^a	4-nitrobenzoic	3.442 ^a	1,4-benzenediamine	6.06 ^a
4-nitrophenol	7.150 ^a	2-methylbenzoic	3.906 ^a	2-chloroaniline	2.653 ^a
resorcinol	9.32 ^b	4-methylbenzoic	4.370 ^a	3-hydroxyaniline	4.27 ^a
hydroquinone	9.85 ^a	2-chlorobenzoic	2.925 ^a	4-chloroaniline	3.979 ^a
3-methoxyphenol	9.652 ^a	4-chlorobenzoic	3.986 ^a	pyridine	5.2 ^a
4-methoxyphenol	10.21 ^a	4-aminobenzoic, pKa2	4.870 ^a	pyrazine	0.97 ^a
4-hydroxy-benzonitrile	7.97 ^a	trifluoroacetic	-0.2 ^a	benzylamine	9.34 ^b
4-hydroxy-benzaldehyde	7.617 ^a	cyanoacetic	2.472 ^a	<i>N</i> -methylallylamine	10.11 ^d
2,3-dichlorophenol	7.696 ^a	nitroacetic	1.68 ^a	4-methoxypyridine	6.47 ^a
4-nitrosophenol	6.33 ^a	4-hydroxybenzoic	4.577 ^a	4-aminotoluene	5.08 ^a

Table 2. Test Sets Used to Test the Performance of the FP Method and Experimental pK_a Values

phenol	pK_a^{exp}	carboxylic acid	pK_a^{exp}	amine	pK_a^{exp}
3,4-dihydroxybenzoic acid, pKa2	8.75 ^a	mercaptoacetic	3.68 ^b	glycine, methyl ester	7.66 ^a
2-fluorophenol	8.705 ^a	dichloroacetic	1.35 ^b	1-butanamine	10.640 ^a
2-chloro-4-nitrophenol	5.45 ^c	phenylacetic	4.31 ^a	<i>O</i> -methylhydroxylamine	4.75 ^c
2-(methylthio)phenol	9.53 ^b	fumaric, pKa1	3.02 ^a	dimethyl- <i>N</i> -propylamine	9.99 ^b
4-hydroxy-3-methoxybenzaldehyde	7.396 ^a	fumaric, pKa2	4.38 ^b	4-fluoroaniline	4.654 ^a
4-hydroxy-3-methoxybenzoic acid, pKa2	9.39 ^c	3,4-dihydroxybenzoic, pKa1	4.30 ^a	3-methoxy-benzenamine	4.204 ^a
1-(4-hydroxyphenyl)-ethanone	8.05 ^a	4-methoxybenzoic	4.475 ^a	imidazole	6.993 ^a
4-hydroxy-benzenemethanol	9.82 ^c	4-hydroxy-3-methoxybenzoic, pKa1	4.31 ^d	hydrazine	7.98 ^a
4-trifluoromethyl-phenol	8.675 ^a	dimethylpropanoic	5.032 ^a	ethylamine	10.673 ^a
3-fluorophenol	9.206 ^a	3-nitrobenzoic	3.449 ^a	<i>N</i> -methylaniline	4.832 ^a

Structure	Names	Exp. pK_a
	ethanethiol	10.61 ^a
	propane-2-thiol	10.86 ^a
	1,1-dimethyl-1-propanethiol	11.21 ^b
	1-propene-2-thiol (2-mercaptopropene)	7.86 ^c
	ethane-1,2-dithiol (pK_{a1})	9.00 ^a
	ethane-1,2-dithiol (pK_{a2})	10.43 ^a
	2-mercaptoethanol	9.7 ^a
	2,2-bis(hydroxymethyl)-3-mercaptopropanol (monothiopenta erythritol)	9.89 ^a
	mercaptoacetic acid (thioglycolic acid)	10.61 ^a
	2-ethoxyethanethiol (2-mercaptoethyl ether)	9.38 ^d
	mercaptoacetic acid methyl ester	7.99 ^a
	4-mercapto-N-methylpiperidine	10.35 ^a
	benzenethiol (thiophenol)	6.615 ^a
	4-methylbenzenethiol (<i>p</i> -methylthiophenol)	6.820 ^a
	2-chlorobenzenethiol (<i>o</i> -chlorothiophenol)	5.675 ^a
	2-aminebenzenethiol (<i>o</i> -aminothiophenol)	6.59 ^d
	4-nitrobenzenethiol (<i>p</i> -nitrothiophenol)	4.715 ^a
	4-methoxybenzenethiol (<i>p</i> -methoxythiophenol)	6.775 ^a
	4-acetylbenzenethiol (<i>p</i> -acetylthiophenol)	5.330 ^a
	benzyl mercaptan (benzenemethanethiol) (benzylthiol)	9.43 ^b

tioles

Structure	Names	Exp. pK_a
	1-butanethiol (butyl mercaptan)	10.66 ^b
	2-methylpropane-2-thiol	11.22 ^a
	2-propene-1-thiol (allyl mercaptan)	9.96 ^b
	mercaptoacetic acid ethyl ester	7.95 ^c
	4-chlorobenzenethiol (<i>p</i> -chlorothiophenol)	6.135 ^a
	3-nitrobenzenethiol (<i>m</i> -nitrothiophenol)	5.241 ^a
	DL-2-mercaptopropanoic acid (hiolactic acid)	10.70 ^a
	2-mercapto-propionylglycine (2-sulfanyl(propionylamino)acetic acid)	8.87 ^a
	2-amino-3-sulfhydrylpropanoic acid (cysteine)	8.36 ^a
	2-methylthiobenzenethiol (<i>o</i> -methylthiothiophenol)	5.754 ^a

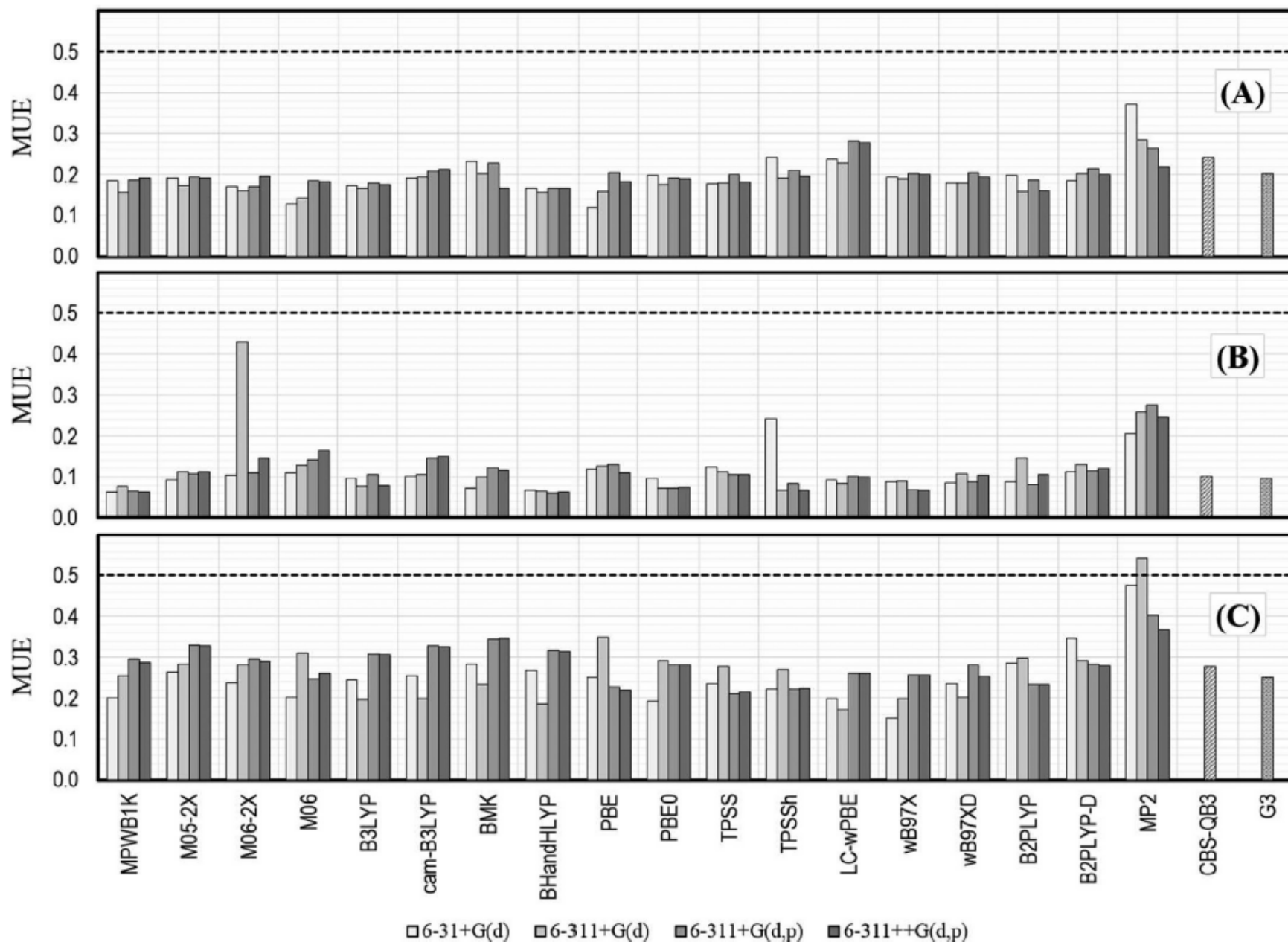


Figure 1. Mean unsigned errors (MUE), in pK_a units, obtained for the test sets of phenols (A), carboxylic acids (B), and amines (C). The dashed line represents the target value.

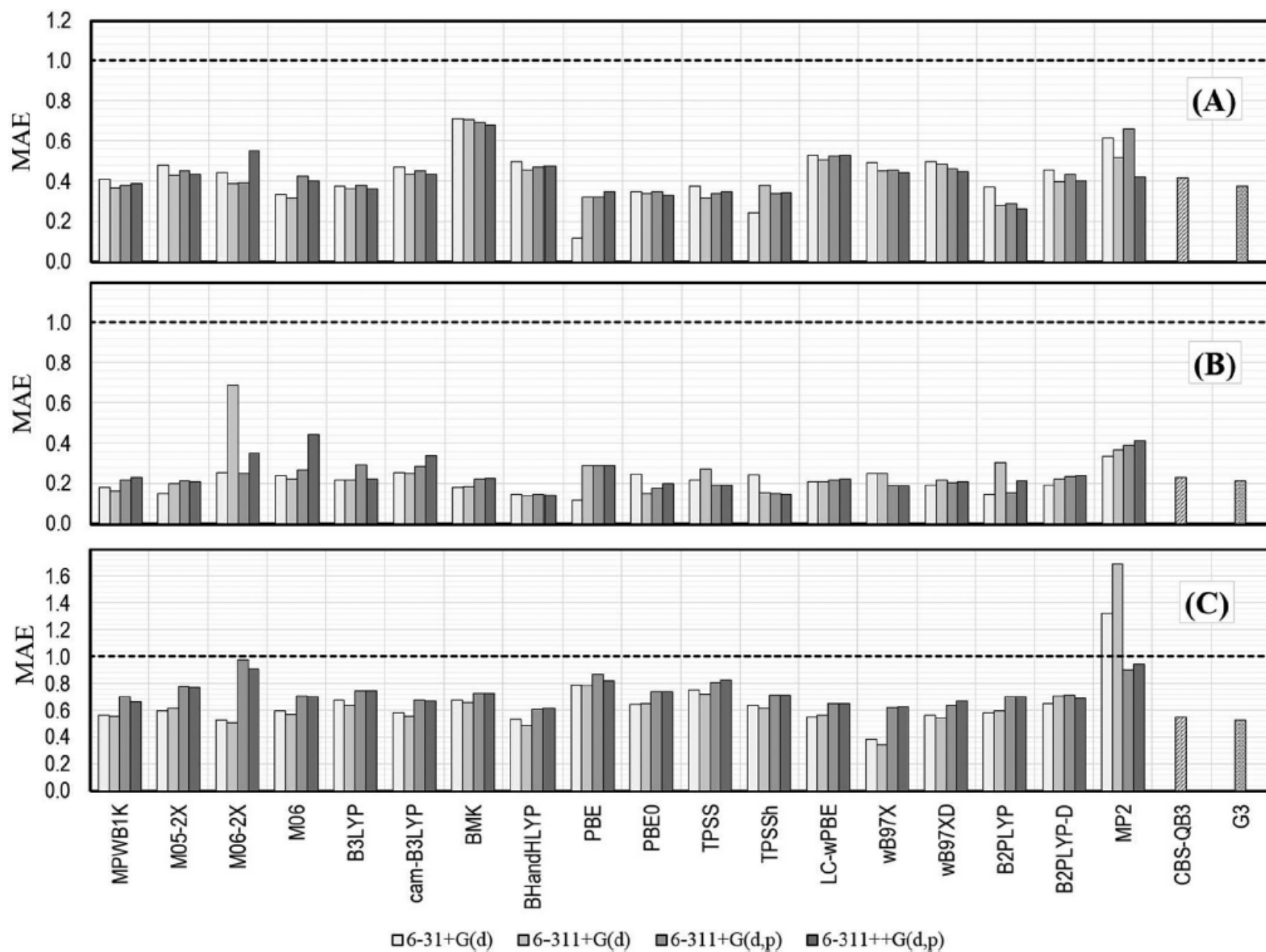


Figure 3. Maximum absolute errors (MAE), in pK_a units, obtained for the test sets of phenols (A), carboxylic acids (B), and amines (C). The dashed line represents the target value.

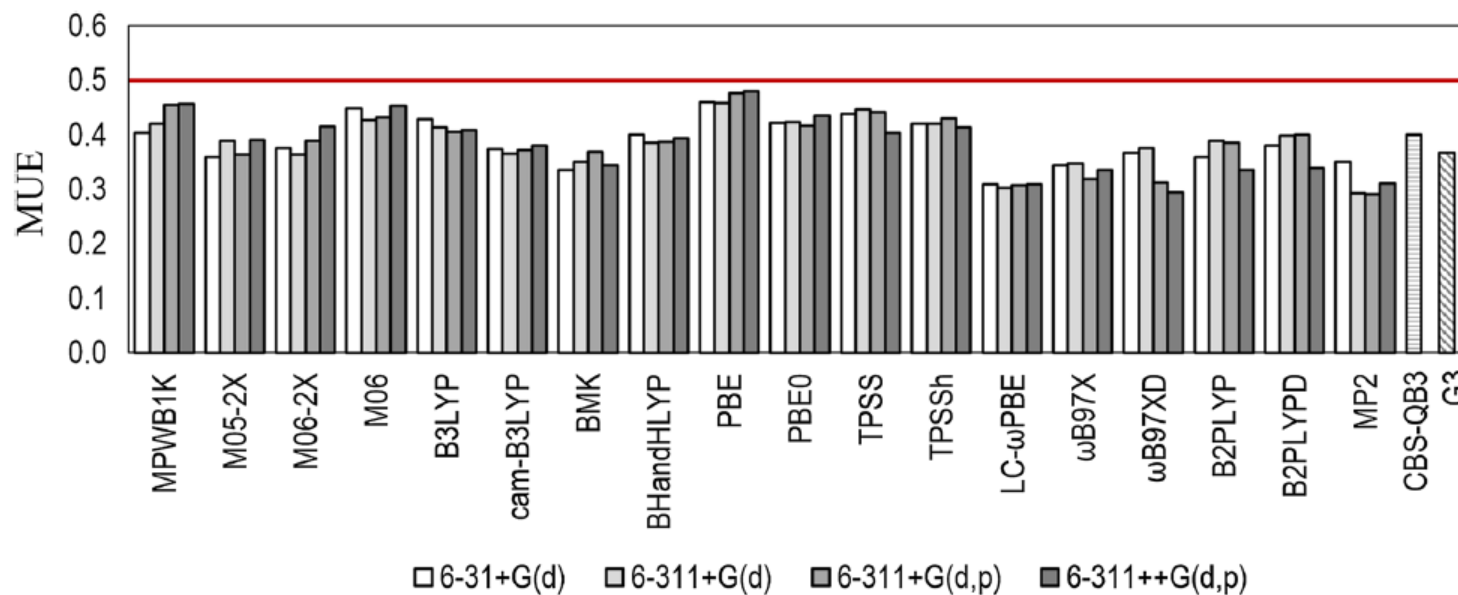


Fig. 1 Mean unsigned errors (MUEs), in pK_a units, obtained for the test sets of thiols. The horizontal line marks the target value

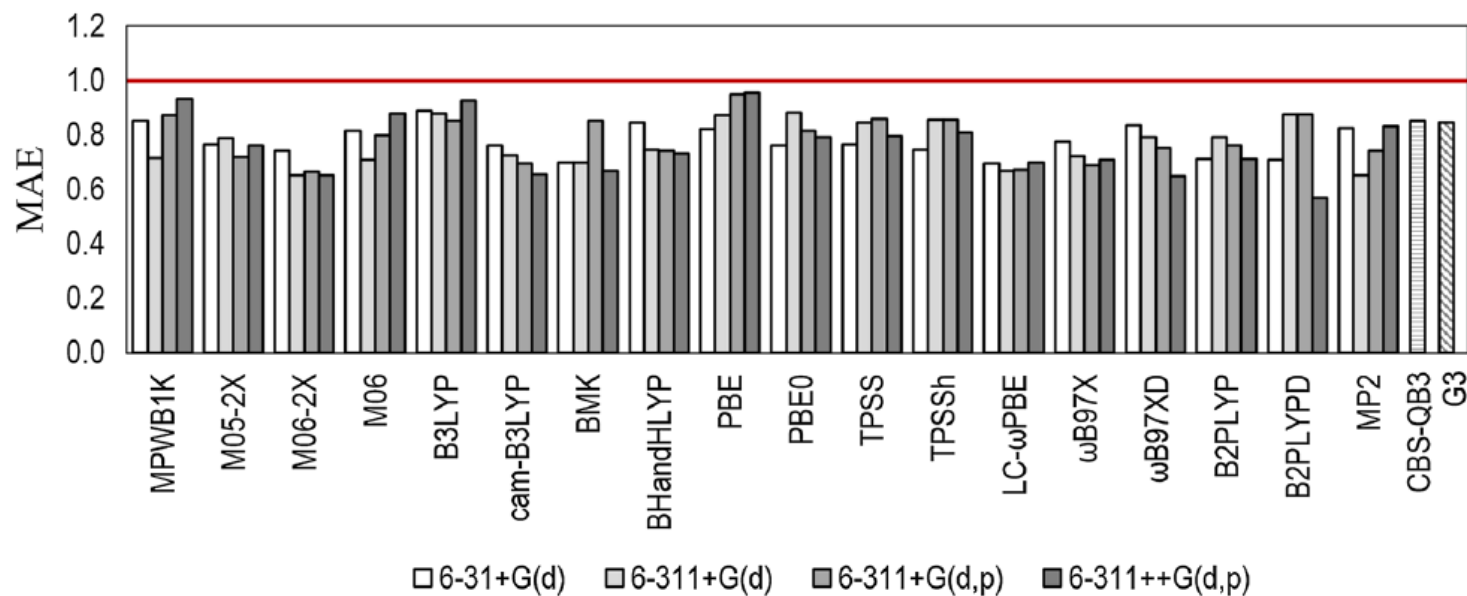


Fig. 2 Maximum absolute errors (MAEs), in pK_a units, obtained for the test sets of thiols. The horizontal line marks the target value

Table 3. Empirically Fitted Parameters (m and C_0) for Calculating pK_a Values of Phenols

	6-31+G(d)		6-311+G(d)		6-311+G(d,p)		6-311++G(d,p)	
	m	C_0	m	C_0	m	C_0	m	C_0
MPWB1K	0.299	-76.603	0.305	-78.215	0.302	-78.493	0.303	-78.564
M05-2X	0.313	-80.020	0.313	-79.435	0.316	-81.497	0.317	-81.693
M06-2X	0.315	-80.680	0.314	-80.073	0.318	-82.209	0.318	-82.349
M06	0.285	-71.924	0.289	-72.746	0.290	-74.064	0.291	-74.411
B3LYP	0.284	-71.669	0.284	-71.465	0.285	-72.933	0.286	-73.092
cam-B3LYP	0.306	-77.994	0.306	-77.676	0.309	-79.509	0.309	-79.644
BMK	0.265	-66.137	0.269	-67.173	0.270	-68.221	0.270	-68.350
BHandHLYP	0.302	-77.406	0.301	-77.230	0.304	-79.183	0.305	-79.473
PBE	0.267	-66.340	0.278	-69.341	0.280	-70.833	0.281	-71.053
PBE0	0.292	-74.192	0.294	-74.605	0.296	-76.202	0.297	-76.472
TPSS	0.280	-70.720	0.280	-70.559	0.282	-71.926	0.282	-72.194
TPSSh	0.269	-67.830	0.285	-72.120	0.286	-73.542	0.287	-73.765
LC- ω PBE	0.328	-85.041	0.325	-83.818	0.327	-85.680	0.328	-85.839
ω B97X	0.328	-84.464	0.326	-83.859	0.332	-86.633	0.332	-86.819
ω B97XD	0.312	-80.511	0.313	-80.509	0.317	-82.746	0.318	-82.998
B2PLYP	0.301	-76.171	0.303	-76.676	0.305	-78.854	0.305	-78.867
B2PLYP-D	0.298	-75.527	0.302	-76.540	0.305	-78.702	0.306	-79.128
MP2	0.363	-93.670	0.352	-90.368	0.358	-94.844	0.363	-96.222
CBS-QB3	0.350	-92.353						
G3	0.342	-90.268						

Table 4. Empirically Fitted Parameters (m and C_0) for Calculating pK_a Values of Carboxylic Acids

	6-31+G(d)		6-311+G(d)		6-311+G(d,p)		6-311++G(d,p)	
	m	C_0	m	C_0	m	C_0	m	C_0
MPWB1K	0.345	-91.274	0.349	-92.266	0.347	-93.017	0.347	-93.123
M05-2X	0.360	-94.513	0.360	-94.101	0.356	-94.380	0.357	-94.622
M06-2X	0.361	-95.155	0.376	-98.630	0.357	-95.033	0.358	-95.120
M06	0.314	-82.101	0.328	-85.970	0.330	-87.368	0.326	-86.411
B3LYP	0.337	-88.582	0.341	-89.657	0.343	-91.387	0.343	-91.399
cam-B3LYP	0.341	-89.049	0.345	-90.158	0.345	-91.268	0.344	-91.194
BMK	0.344	-90.158	0.344	-89.932	0.340	-90.095	0.342	-90.599
BHandHLYP	0.348	-91.988	0.348	-91.988	0.350	-93.812	0.350	-93.858
PBE	0.323	-84.597	0.340	-89.087	0.346	-91.883	0.347	-92.136
PBE0	0.351	-92.737	0.357	-94.106	0.360	-96.221	0.365	-97.737
TPSS	0.319	-84.104	0.324	-85.339	0.326	-86.890	0.325	-86.709
TPSSh	0.322	-84.934	0.336	-88.881	0.340	-90.992	0.340	-91.075
LC- ω PBE	0.377	-99.898	0.377	-99.742	0.379	-101.521	0.379	-101.724
ω B97X	0.366	-96.619	0.371	-97.807	0.371	-99.160	0.371	-99.157
ω B97XD	0.361	-95.530	0.366	-96.990	0.366	-98.307	0.368	-98.617
B2PLYP	0.349	-91.420	0.359	-94.236	0.358	-95.463	0.359	-95.830
B2PLYP-D	0.345	-90.420	0.354	-92.894	0.347	-92.658	0.348	-92.775
MP2	0.376	-97.830	0.396	-103.808	0.401	-107.958	0.403	-108.537
CBS-QB3	0.390	-104.497						
G3	0.392	-105.646						

Table 5. Empirically Fitted Parameters (m and C_0) for Calculating pK_a Values of Amines

	6-31+G(d)		6-311+G(d)		6-311+G(d,p)		6-311++G(d,p)	
	m	C_0	m	C_0	m	C_0	m	C_0
MPWB1K	0.498	-130.930	0.505	-133.064	0.497	-131.441	0.494	-130.452
M05-2X	0.475	-124.072	0.475	-123.807	0.464	-121.000	0.464	-121.235
M06-2X	0.504	-131.574	0.499	-130.033	0.486	-126.867	0.485	-126.727
M06	0.513	-133.920	0.536	-139.968	0.518	-135.294	0.520	-135.911
B3LYP	0.497	-130.452	0.502	-131.801	0.484	-127.244	0.484	-127.509
cam-B3LYP	0.499	-130.528	0.503	-131.693	0.486	-127.438	0.487	-127.546
BMK	0.483	-126.785	0.492	-128.862	0.472	-123.746	0.473	-123.980
BHandHLYP	0.493	-129.951	0.498	-131.390	0.481	-127.277	0.483	-127.691
PBE	0.497	-130.042	0.507	-132.605	0.489	-127.969	0.485	-126.877
PBE0	0.494	-129.783	0.508	-133.801	0.490	-129.155	0.491	-129.461
TPSS	0.496	-131.141	0.505	-133.541	0.492	-130.308	0.495	-131.308
TPSSh	0.494	-130.509	0.503	-133.212	0.490	-129.870	0.491	-130.366
LC- ω PBE	0.497	-131.249	0.494	-130.465	0.491	-129.854	0.492	-130.184
ω B97X	0.506	-133.174	0.514	-135.622	0.502	-132.552	0.503	-132.935
ω B97XD	0.489	-129.398	0.500	-132.565	0.496	-131.715	0.491	-130.301
B2PLYP	0.491	-128.740	0.503	-132.033	0.497	-131.023	0.497	-130.989
B2PLYP-D	0.490	-128.701	0.495	-129.952	0.486	-128.235	0.483	-127.285
MP2	0.493	-128.841	0.495	-129.379	0.522	-138.182	0.528	-139.753
CBS-QB3	0.531	-140.475						
G3	0.517	-136.908						

Table 4 Empirically fitted parameters (m and C_0) for calculating pK_a values of thiols, at different levels of theory

	6-31 + G(d)		6-311 + G(d)		6-311 + G(d,p)		6-311 ++G(d,p)	
	m	C_0	m	C_0	m	C_0	m	C_0
MPWB1K	0.353	- 93.957	0.352	- 93.629	0.356	- 95.629	0.351	- 94.100
M05-2X	0.350	- 91.867	0.353	- 92.371	0.357	- 94.639	0.354	- 93.619
M06-2X	0.353	- 92.784	0.355	- 93.099	0.356	- 94.274	0.356	- 94.378
M06	0.343	- 90.504	0.359	- 95.056	0.360	- 96.049	0.348	- 92.578
B3LYP	0.340	- 90.040	0.342	- 90.208	0.345	- 92.000	0.344	- 91.979
cam-B3LYP	0.369	- 98.029	0.368	- 97.551	0.372	- 99.715	0.380	- 102.014
BMK	0.358	- 94.172	0.363	- 95.488	0.368	- 97.845	0.368	- 97.590
BHandHLYP	0.358	- 95.511	0.357	- 95.106	0.362	- 97.386	0.362	- 97.389
PBE	0.314	- 82.117	0.317	- 82.789	0.317	- 83.419	0.317	- 83.599
PBE0	0.333	- 87.812	0.334	- 88.119	0.335	- 89.319	0.336	- 89.585
TPSS	0.318	- 84.172	0.318	- 83.875	0.321	- 85.688	0.322	- 85.717
TPSSh	0.327	- 86.666	0.329	- 87.092	0.331	- 88.620	0.332	- 88.885
LC- ω PBE	0.378	- 101.825	0.376	- 101.108	0.381	- 103.394	0.381	- 103.420
ω B97X	0.389	- 104.458	0.388	- 103.743	0.396	- 107.176	0.395	- 106.997
ω B97XD	0.376	- 100.881	0.379	- 101.492	0.384	- 104.130	0.382	- 103.603
B2PLYP	0.344	- 90.730	0.344	- 90.767	0.347	- 92.906	0.347	- 93.040
B2PLYP-D	0.343	- 90.738	0.345	- 91.088	0.346	- 92.782	0.349	- 93.658
MP2	0.369	- 97.615	0.360	- 94.837	0.365	- 99.096	0.361	- 97.957
CBS-QB3	0.347	- 93.259						
G3	0.353	- 96.020						

Tarea:

Obtener para el ácido carboxílico, el fenol, la amina y el tiol escogidos los pKas calculados usando los esquemas reaccionales A y D y el método de parámetros ajustados.

Nivel de teoría B3LYP/6-31+G(d), y solvente agua (SMD)

Comparar los resultados contra los correspondientes valores experimentales.