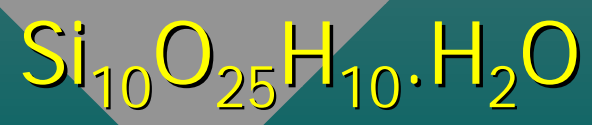
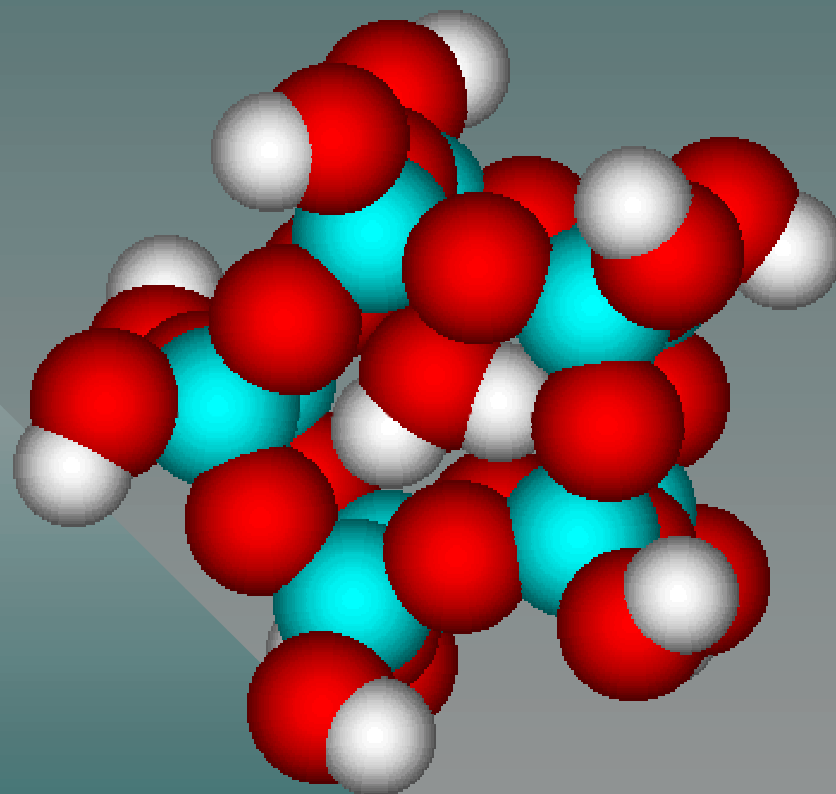


*La modelización de
moléculas y
biomoléculas:
Principios básicos*

Luis A. Montero
Laboratorio de Química Computacional y
Teórica
Universidad de La Habana

La química computacional es la "*química del laboratorio seco*" porque se propone obtener **modelos teóricos** de objetos al nivel molecular en un ordenador.

De esta forma se pueden representar muchos fenómenos químicos y físicos asociados con tales objetos, tanto de especies estables como transientes.



El principal propósito es apoyar a la ciencia experimental y hasta reemplazar algunos experimentos muy caros con modelos computacionales más baratos

Entre las propiedades más importantes que pueden ser calculadas están:

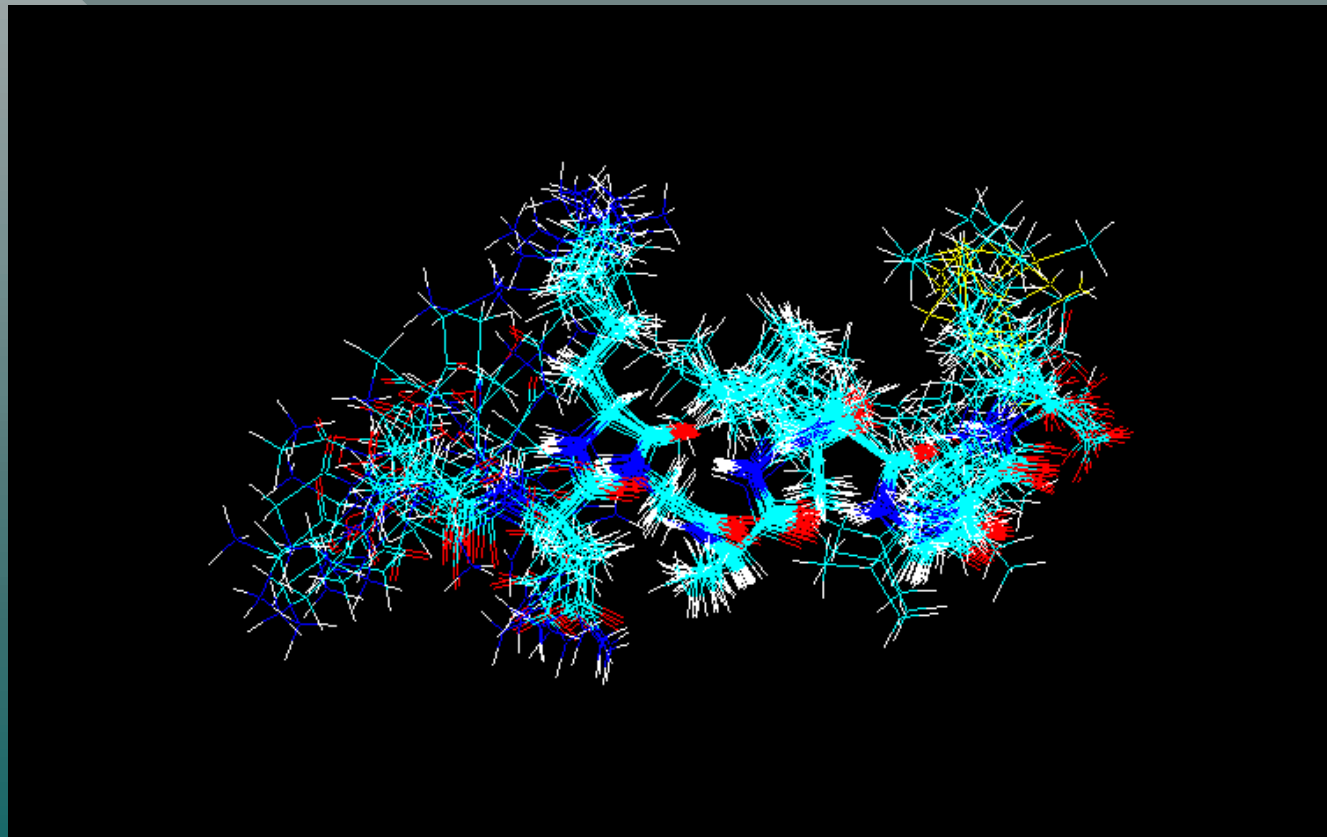
- La geometría molecular
- Las densidades electrónicas y cargas parciales
- Espectros ópticos y de resonancia magnética nuclear
- Propiedades termodinámicas

Muchas de estas propiedades se pueden usar como **descriptores** de propiedades moleculares que pueden ser correlacionados con la actividad biológica o ciertos comportamientos macroscópicos de materiales.

En la medida en que la **selección de descriptores** sea acertada y el **método teórico** utilizado para calcular el sistema molecular sea bueno, se podrá predecir la actividad deseada en sistemas poliatómicos solo existentes en el ordenador mediante correlaciones estadísticas simples.

Péptido β amiloide (25-35)

20 conformaciones
experimentales de RMN en
solución



Hipersuperficies Potenciales

- Las *hipersuperficies potenciales* son funciones multidimensionales que expresan la energía potencial total de un sistema molecular a partir de, en principio, las coordenadas de todas las partículas del mismo y de los tipos de núcleos atómicos involucrados:

$$E = E(\vec{r}, Z)$$

El número de dimensiones depende del sistema y de lo que se desee (o se pueda) modelar. Es este uno de los puntos más delicados y que requieren de una mayor profesionalidad: *la selección de las "variables independientes" a considerar.*

El problema general de los métodos de modelación (o modelización) molecular es encontrar la forma funcional de la hipersuperficie del sistema en estudio

La única forma teóricamente consistente para predecir los fenómenos al nivel molecular y macromolecular es mediante el uso de la **mecánica cuántica** y sus procedimientos.

Por lo tanto, la forma funcional aceptable como exacta sería:

$$E = \frac{\hat{H}\Psi(\vec{r}, Z)}{\Psi(\vec{r}, Z)}$$

Para solucionar el problema se procede con diferentes tratamientos a la teoría básica, que se basan alternativamente en:

- *Considerar explícitamente la densidad electrónica de las moléculas*

O también en:

- *Tratar el sistema como un conjunto mecánico de bolas asociadas mediante resortes*

La consideración explícita de la densidad electrónica implica, necesariamente, formular y resolver un hamiltoniano del sistema, aunque sea aproximado.

Esta es la única forma de poder abordar el fenómeno de las interacciones y la reactividad molecular *a priori* y con capacidad interpretativa.

Se trata de los llamados ***métodos cuánticos.***

Todos los métodos cuánticos son aproximados, en mayor o menor medida, porque la solución exacta de un hamiltoniano solo puede obtenerse en el caso de sistemas extremadamente simples.

Se llaman métodos *a priori* aquellos cuyas aproximaciones se introducen de acuerdo con una lógica dada por la propia teoría y sin tener en consideración los resultados que se desea obtener.

Se consideran *a posteriori* aquéllos métodos cuyas aproximaciones y parámetros tienen en cuenta resultados obtenidos en sistemas conocidos y se optimizan para lograr la mayor coherencia entre los datos teóricos y los experimentales esperados.

Entre los *métodos cuánticos simples* están el de *orbitales moleculares de Hückel (HMO)* y la *teoría de Hückel extendida (EHT)*.

Estos hacen grandes simplificaciones para construir el hamiltoniano. El problema se reduce a un simple cálculo de valores y vectores propios para obtener las funciones moleculares. Generalmente contienen parámetros clave que son introducidos *a posteriori*.

Existen también los llamados hamiltonianos **semiempíricos**

Estos se basan en la teoría de Hartree - Fock autoconistente con hamiltonianos simplificados y con parámetros empíricos (seleccionados *a posteriori*). Tienen en cuenta implícitamente la energía de correlación electrónica. Son populares los que niegan la superposición diferencial diatómica (***PPP, CNDO, INDO, NDDO, MNDO, AM1, PM3***).

Son muy usados actualmente los métodos basados en la teoría de los funcionales de la densidad (DFT)

En ellos se simplifica el hamiltoniano con un fundamento teórico consistente. Sus aplicaciones se basan en elaboraciones semiempíricas *a posteriori* en dependencia de las propiedades moleculares que se desea calcular. Existen funcionales para diversos propósitos y algunos de ellos incluyen diferentes efectos de correlación electrónica.

Los más consistentes con la teoría básica son los llamados métodos **ab initio**

Se basan en hamiltonianos construidos rigurosamente para cada sistema poliatómico de acuerdo con la teoría de Hartree - Fock autocoherente. Las integrales son calculadas casi siempre a partir de funciones de base atómicas (seleccionadas *a priori* y con diferentes grados de exactitud). La energía de correlación electrónica debe ser calculada como una corrección a partir de métodos muy elaborados. Puede presentar dificultades en el cálculo de interacciones moleculares.

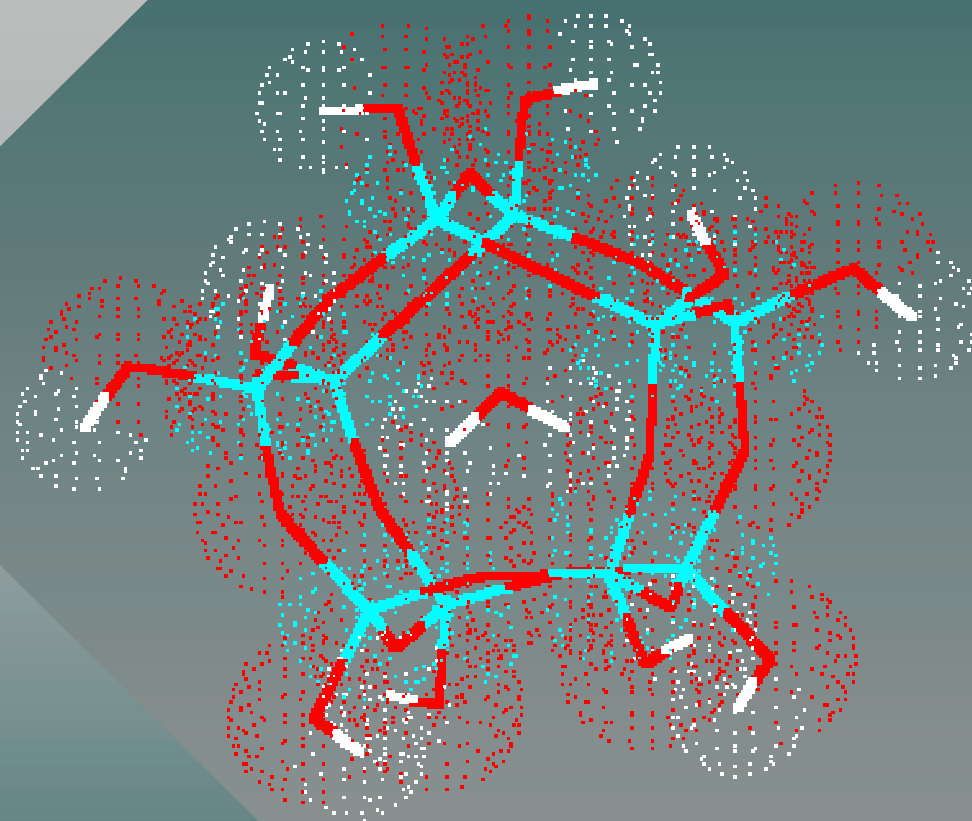
Los métodos más simples consumen el mínimo de tiempo de cálculo y por lo tanto con ellos se puede modelar a los mayores sistemas poliatómicos, en dependencia de la potencia computacional disponible. La capacidad predictiva es limitada en la mayoría de los casos.

Al aumentar la complejidad teórica:

- aumenta la capacidad predictiva de los modelos obtenidos.
- aumenta el tiempo de cálculo
- disminuye la posibilidad de modelar “grandes” sistemas moleculares, aunque sean relativamente sencillos en el mundo de los materiales y de la biología.

Los problemas más importantes actualmente para los métodos cuánticos son:

- La consideración de la energía de correlación electrónica
- Los efectos del ambiente molecular para modelar las condiciones reales de existencia de los sistemas poliatómicos
- La relación de escala de las propiedades físicas y químicas calculadas con respecto a las mediciones experimentales



Cálculo semiempírico PM3 del
 $\text{Si}_{10}\text{O}_{25}\text{H}_{10}\cdot\text{H}_2\text{O}$

La modelación mediante bolas y resortes conduce a un sencillo esquema de fórmulas que se denominan ***métodos clásicos***

Casi todas las fórmulas se basan en la ley de Hook o del oscilador armónico

$$E = \frac{1}{2} k \Delta r^2$$

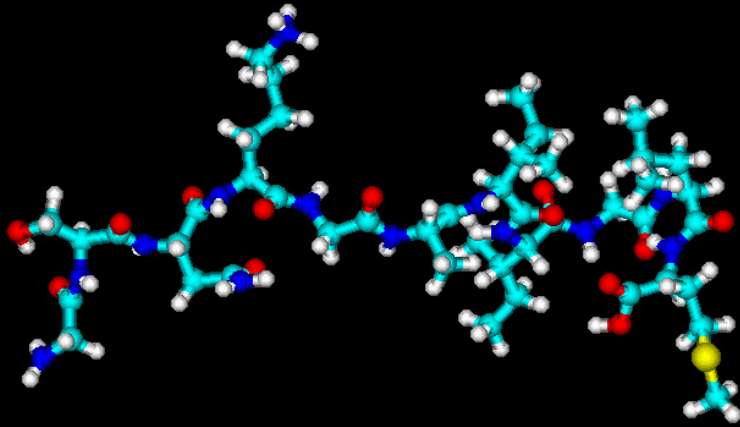
donde r es la distancia entre dos núcleos y la k depende de las masas y de la envoltura electrónica de cada uno de ellos

Los métodos se diferencian entre si por las formulaciones utilizadas, que siempre son *a posteriori*, y en los valores de los parámetros ajustados (MM4, AMBER, etc.).

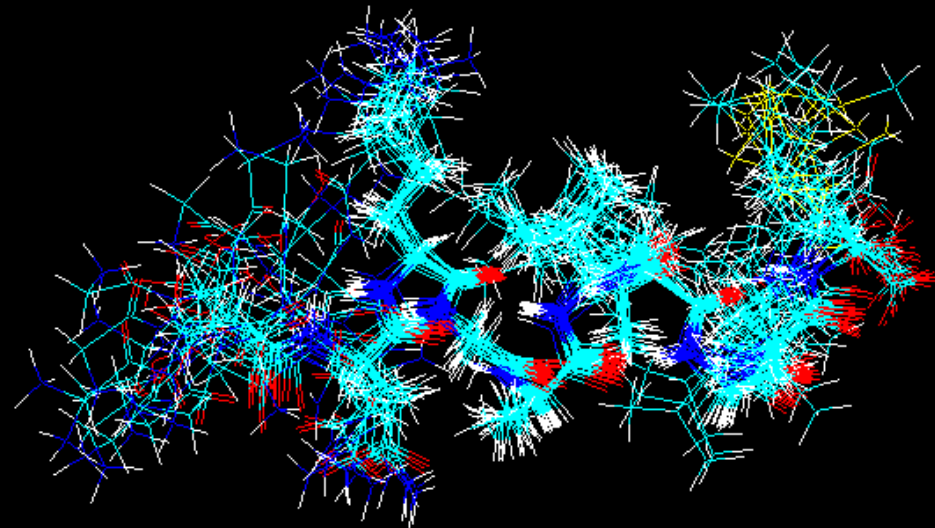
Casi invariablemente se utiliza una tabla periódica agrandada, donde un mismo elemento puede aparecer con diferentes *tipos* en una molécula.

En los casos de moléculas muy complejas o poliméricas, como las proteínas, aún una modelación clásica puede resultar impracticable, por lo que a veces que se suele limitar el número de dimensiones de la hipersuperficie a solo ciertos ángulos relacionados con enlaces significativos.

β amiloide 25-35 (AMBER)



β amiloide 25-35 (exp.)



Un aspecto clave es la exploración de la hipersuperficie en busca de los mínimos de energía potencial.

Esto es importante tanto para los métodos clásicos como cuánticos.

El camino de los gradientes implica determinar las fuerzas que actúan sobre cada núcleo en las tres direcciones, a partir del cálculo de la energía en cada configuración. De acuerdo con esto se modifica sucesivamente la estructura hasta que todas las fuerzas se anulen:

$$\frac{\partial E}{\partial [\vec{r}]} = 0$$

Las técnicas Monte Carlo implican la exploración de toda la hipersuperficie variando desordenadamente (de forma aleatoria) *todas* las coordenadas de *todos* los núcleos en un número significativo de pruebas. De esta forma se llega a una o varias configuraciones de menor energía con respecto a las demás. Tiene la ventaja de "explorar" todo el espacio.

La **dinámica molecular** implica utilizar la participación del tiempo en las ecuaciones de Newton del movimiento

$$-\nabla_i V(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = m_i \frac{d^2 \vec{r}_i}{dt^2}$$

Donde la solución de esta ecuación diferencial para cada componente de masa m_i proporciona la función que determina el cambio de su posición en el tiempo

Si el potencial V_i se calcula por medio de una función armónica o similar, se trata de **Dinámica Molecular Clásica**. Si ese potencial es una energía calculada mediante un hamiltoniano cuántico se trata de **Dinámica Molecular Cuántica**.

Un problema importante: LA POTENCIA COMPUTACIONAL

Si consideramos **potencia computacional** a la rapidez y eficiencia en "masticar" números. *Los requerimientos para la química computacional, así como para otras ciencias, son infinitos.*

¿Como se logra potencia computacional?

- Con procesadores que puedan realizar millones de millones de operaciones por segundo
- Con grandes capacidades de almacenamiento en la memoria volátil
- Con grandes capacidades de almacenamiento y velocidad de acceso aleatorio en la memoria permanente

El surgimiento de muy avanzadas tecnologías de cómputo, accesibles de modo masivo, nos ha llevado a que los medios para realizar investigaciones de alto nivel en química computacional se pueden adquirir hasta en un supermercado y con una considerable potencia computacional.

Trabajos de prueba entre diferentes procesadores*

* Todos los tiempos de utilización del procesador central en segundos y los porcentajes entre paréntesis se calculan con respecto al PPro 200 MHz (s).

Caso 1. Formaldehído [RHF 6-311G(d,p) MP2 (optimización) en Gaussian 94 bajo Linux 2]

PPro	PPro	PII
200 MHz (s)	200 MHz (d)	400 MHz (d)
405.9 (100)	251.2 (162)*	115.1 (353)*

*Tanto el kernel como el g94 fueron recompilados para SMP

Caso 2. Formaldehído [RHF 6-311G(d,p) MP2 (cálculo de vibraciones internas) en Gaussian 94 bajo Linux 2]

PPro	Ppro	PII
200 MHz (s)	200 MHz (d)	400 MHz (d)
1136.9 (100)	551.9 (206)*	260.2 (437)*

*Tanto el kernel como el g94 fueron recompilados para SMP

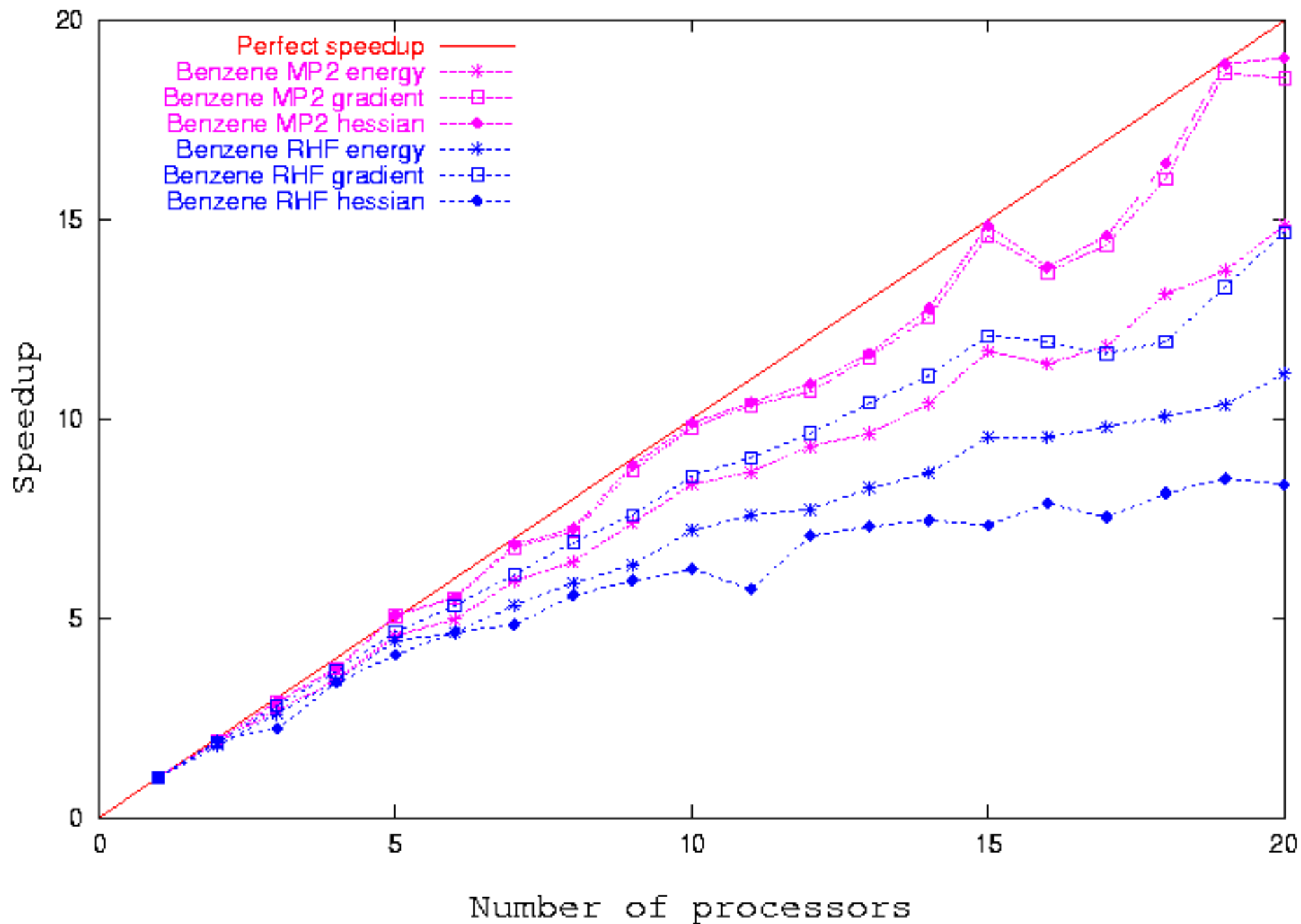
Caso 3. t-Glioxal [RHF 6-31G(d,p) (optimización) en Gaussian 94 bajo Linux 2]

PPro 200 MHz (s)	PPro 200 MHz (d)	AMD-K6 350 MHz (s)	PII 400 MHz (d)	PIV 1.4 GHz (s)
715.0 (100)	258.7 (276)*	242.8 (294)	116.4 (614)*	33.6 (2128)

*Tanto el kernel como el g94 fueron recompilados para SMP

Claustro de ordenadores personales para modelación molecular en la Universidad de Adelaida (Australia)



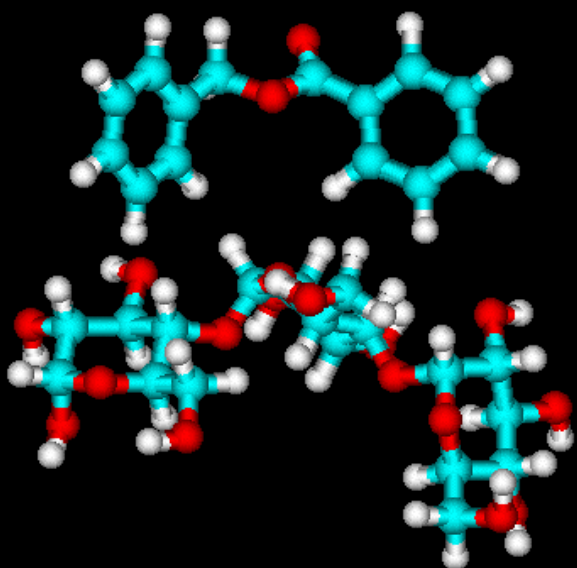


K.A. Hawick, D.A. Grove, P.D. Coddington, M.A. Buntine, DHPC
 Technical Report DHPC-073, 21 January 2000 with GAMESS-US

Veamos un ejemplo...

La amilosa es un soporte común en cromatografía y podemos simularla con una cadena de tres glucosas dispuestas adecuadamente.

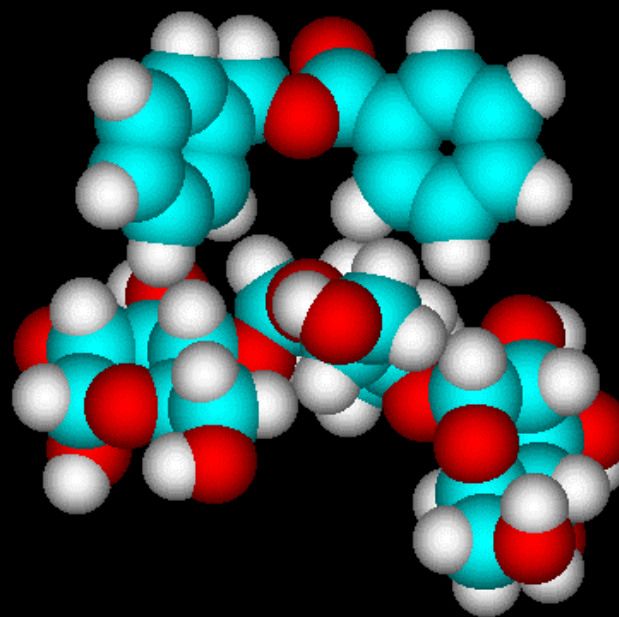
Explorando todo el espacio de las posibles asociaciones entre este trímero de la amilosa con el benzoato de bencilo, se halló mediante el método semiempírico PM3 que una de las asociaciones más favorables es:

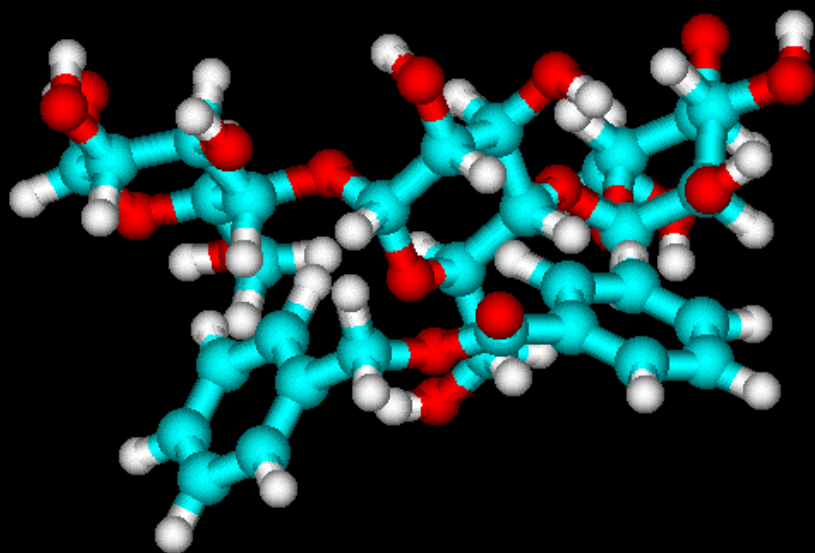


Vista A

Benzoato de
bencilo sobre
amilosa

($E = -9757.12310$ EV)

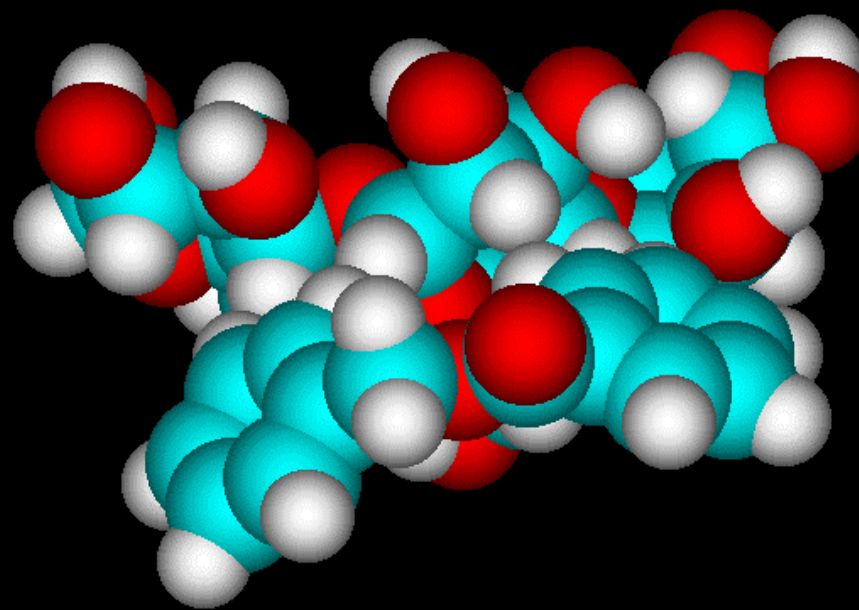




Vista B

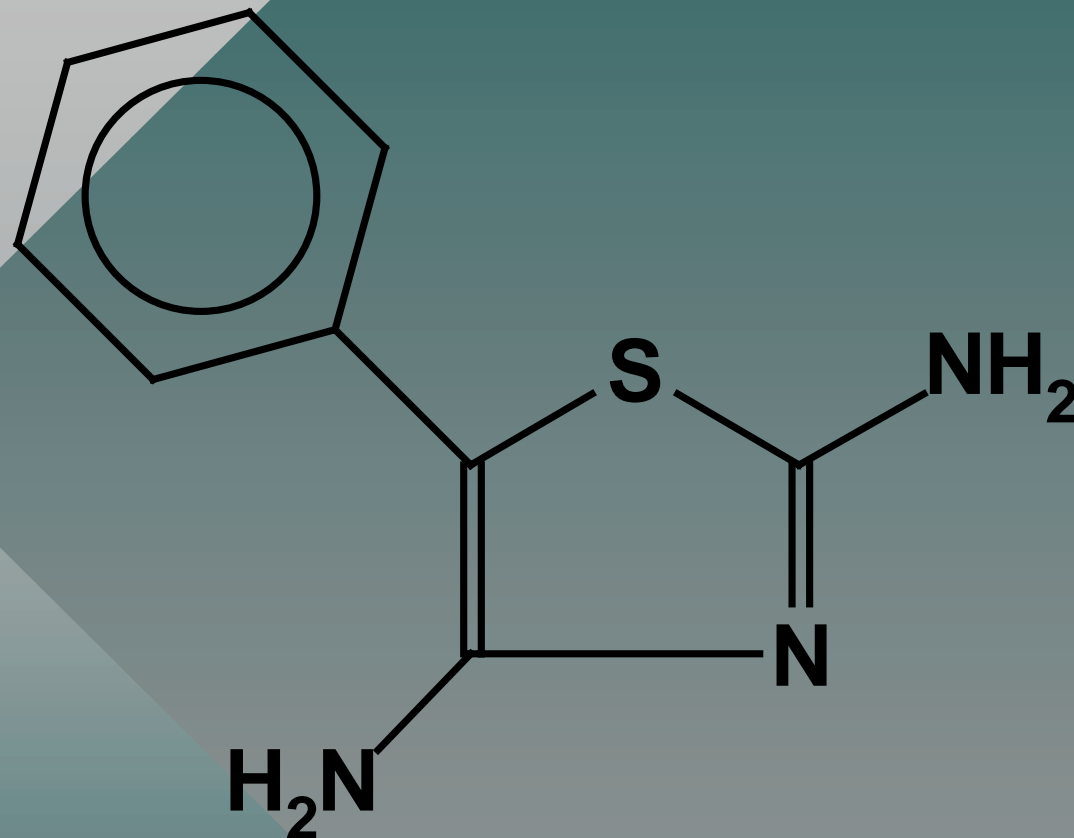
Benzoato de
bencilo sobre
amilosa

($E = -9757.12310$ EV)



Veamos el caso de una droga en solución...

El amifenazol es un estimulante de la respiración celular prohibido por el Comité Olímpico Internacional para los atletas en competencia



Forma neutra del amifenazol:

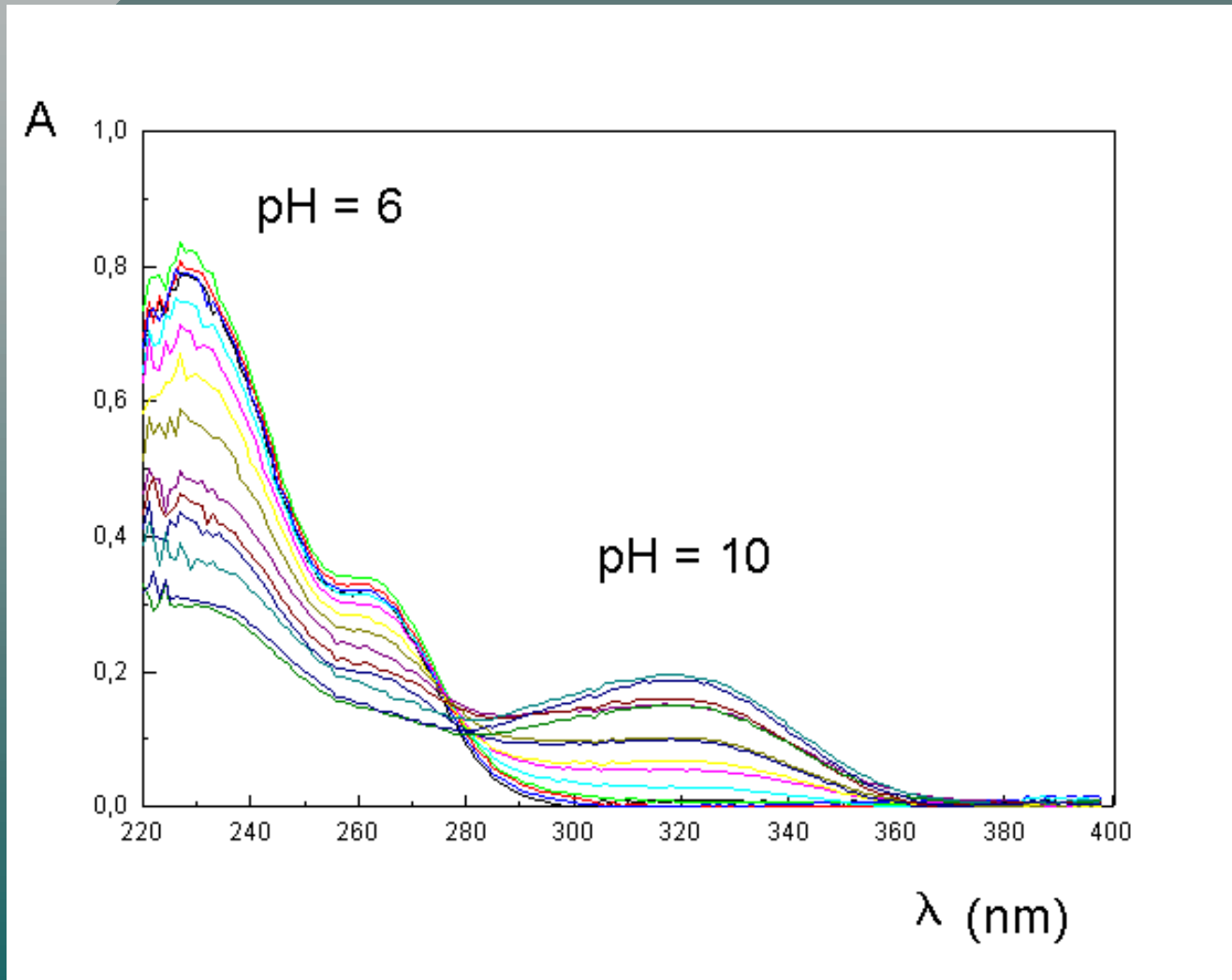
- rotación interna
- cuatro sitios básicos
- tautomerías

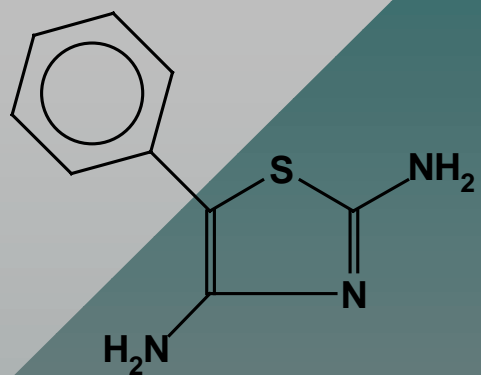
Las evidencias experimentales de interés estructural son:

1. Tanto las formas neutras como las protonadas presentan una señal de RMN protónica sobre carbono cuaternario si el solvente es dimetil sulfóxido.

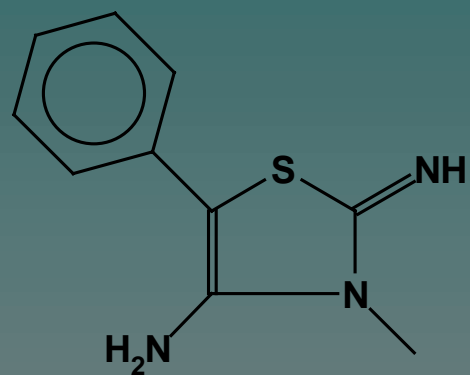
2. Tanto las formas neutras como las protonadas **no** presentan la señal de RMN protónica sobre carbono cuaternario si el solvente es agua deuterada.

3. El espectro UV-Visible en diferentes pH es:

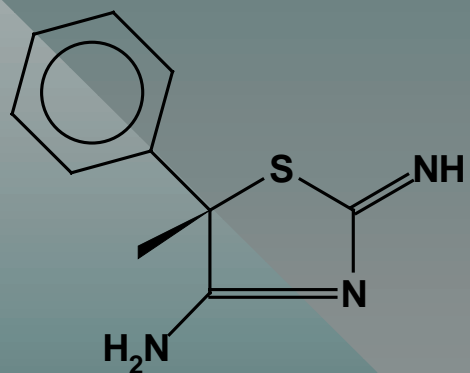




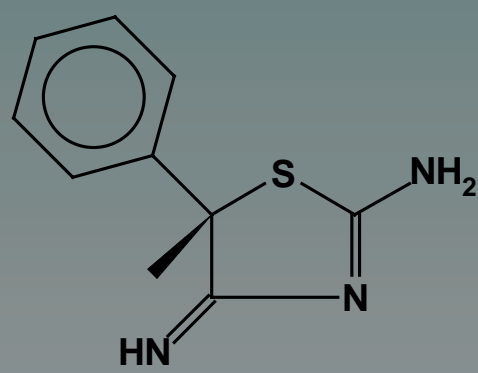
diamina (AMF)



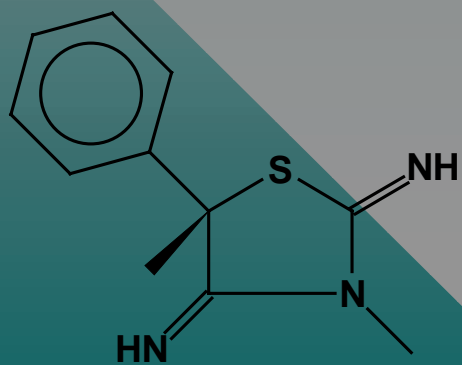
monoimina en N2
variante A (AMI2A)



monoimina en N2
variante B (AMI2B)



monoimina en N4
(AMI4)



diimina (AMII)

Los equilibrios posibles entre estas especies son los siguientes:



Donde X y Y pueden ser, indistintamente, AMF, AMI2A, AMI2B, AMI4, o AMII.

Tautómeros más probables calculados después de la protonación del AMF^a

X	Y	PM3 $\Delta H^0(\text{anh.})$ (kJ/mol) ^b	MMH $\Delta H^0(\text{hydr.})$ (kJ/mol) ^c	MMH $\Delta G^0(\text{hydr.})$ (kJ/mol) ^c	SCF MP2 $\Delta E^0(\text{anh.})$ (kJ/mol) ^e	SCF MP2 $\Delta E^0(\text{solv.})$ (kJ/mol) ^f
AMF	AMF(H3) ⁺	-261.1	-111.9 ^d	-83.9 ^d	-247.2	-60.1
AMF	AMF(H4) ⁺	-263.5	-109.3 ^d	-83.8 ^d	-224.0	-62.4
AMF	SZW ⁺	-264.5	-95.2	-67.2	-298.0	-123.1

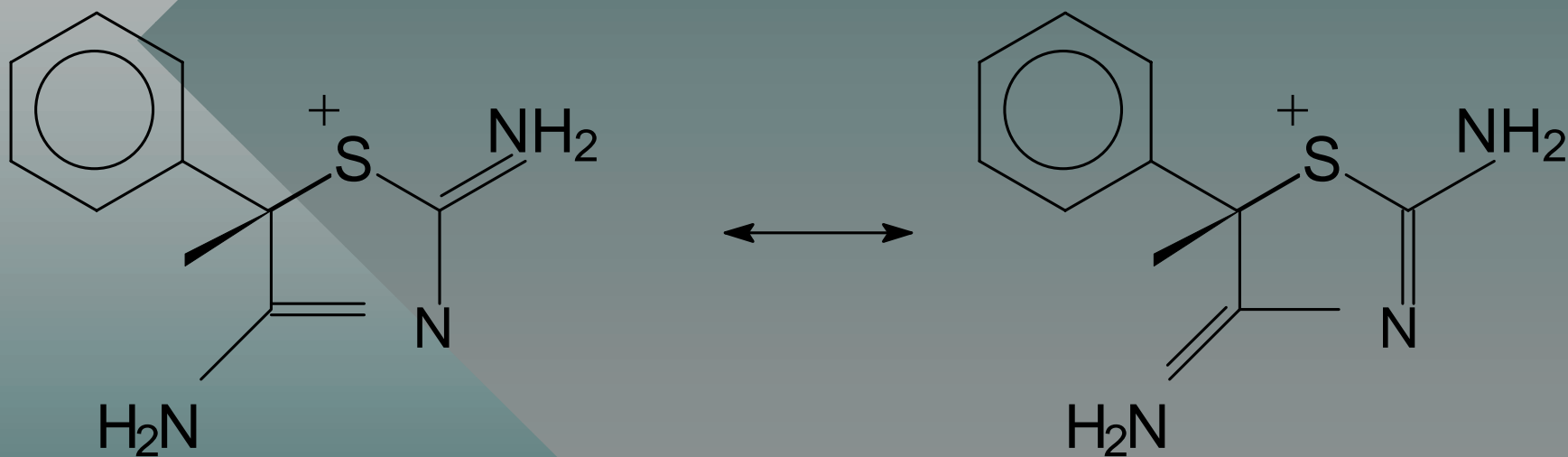
a) Energías en kJ/mol. La lista está en orden ascendente de energías libres calculadas, ΔG^0 , de las especies hidratadas. Otras reacciones no mostradas resultaron en valores positivos de ΔG^0 . Los sitios de protonación se denotan por *Hm*, siendo *m* el número de los átomos pesados en el anillo tiazólico con respecto al S1.

b) Resultados MNDO-PM3 para la reacción [3].

c) Reacción [4].

d) Calculados sobre la base de 45 celdas, en lugar de 15, como se usa en otros casos.

e) HF SCF 6-311G** MP2 | 4-31G** para la reacción [3]. $\Delta E^0(\text{anh.}) + \Delta E^0(\text{SCRFF})$. SCRFF de Tomasi, con la constante dieléctrica del agua.



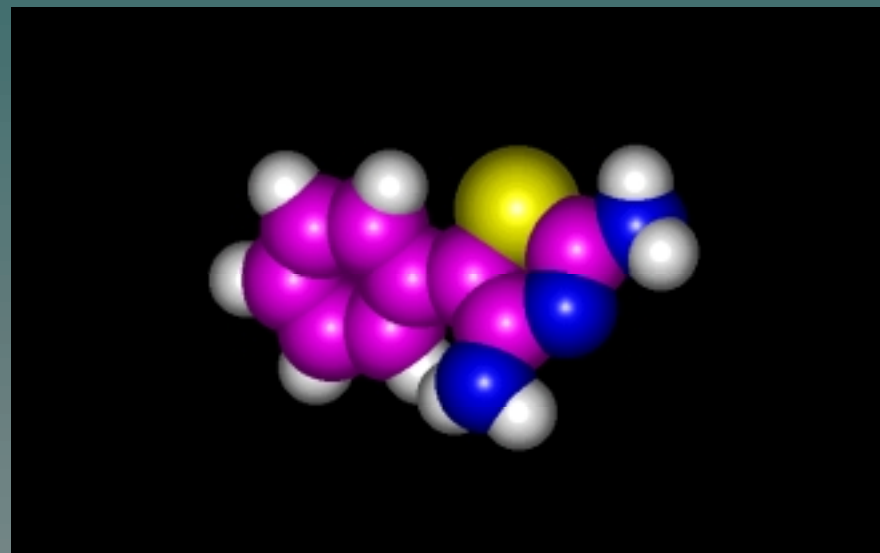
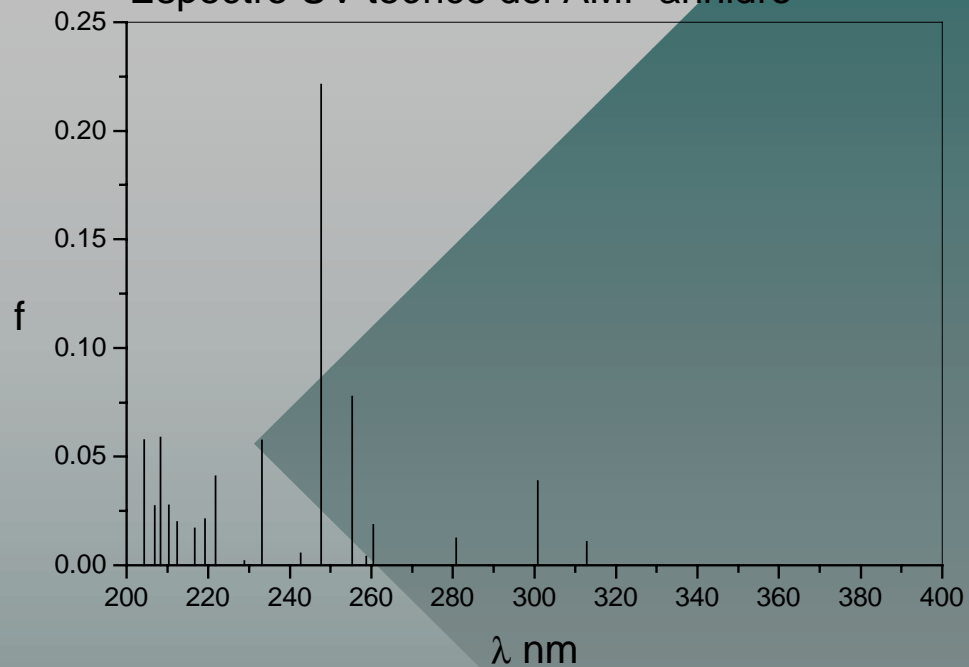
super zwitterion (SZW⁺)

Se hace evidente que los solventes polares influyen de forma diferente en el equilibrio en dependencia de su carácter prótico

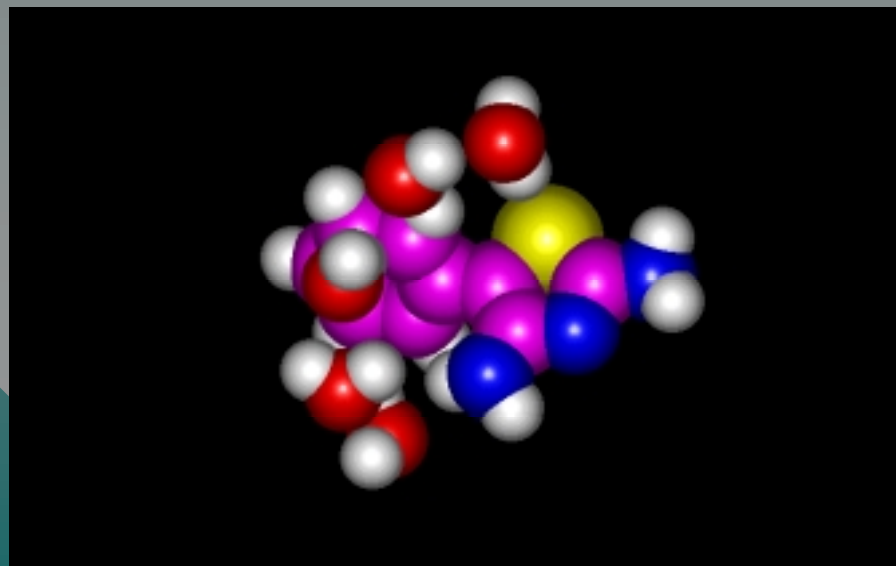
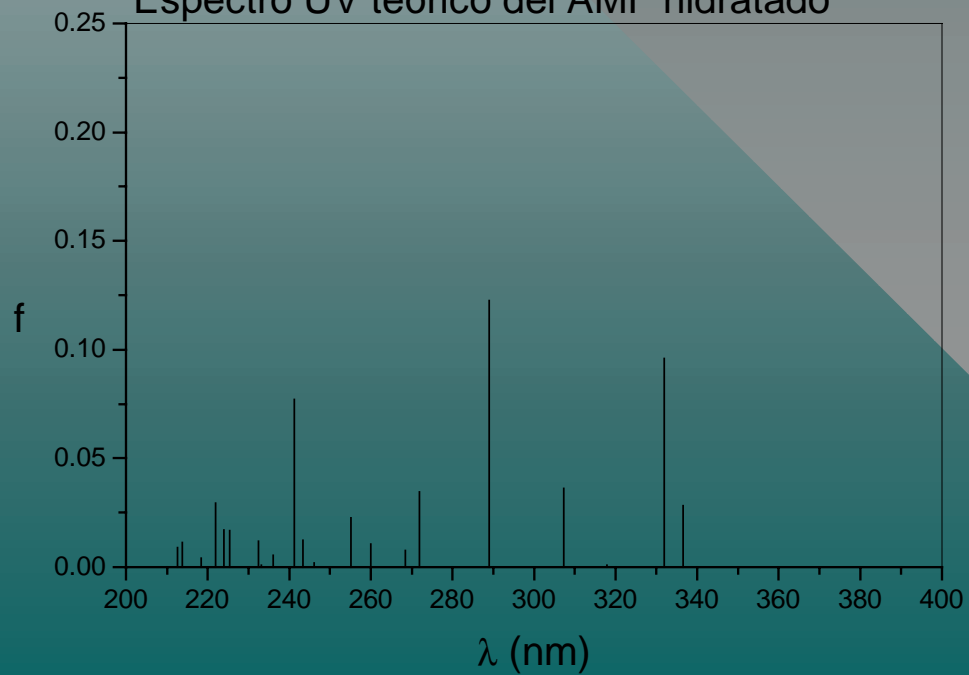
Equilibrios significativos en medio ácido

	ΔH^{aisl}	ΔH^0	ΔG^0
$AMF(H3)^+ \rightarrow AMF(H4)^+$	-2.4	2.3	-2.4 kJ/mol
$SZW^+ \rightarrow AMF(H3)^+$	3.5	-16.4	-16.4 kJ/mol
$SZW^+ \rightarrow AMF(H4)^+$	1.1	-14.1	-18.8 kJ/mol

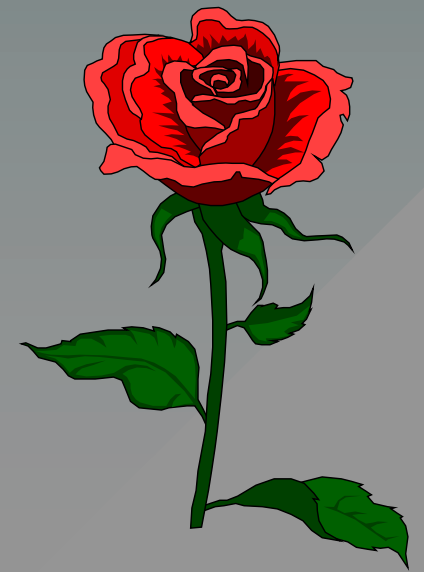
Espectro UV teórico del AMF anhidro



Espectro UV teórico del AMF hidratado



¿conclusiones?



La química computacional puede usarse para modelar cualquier problema que tenga su origen en la estructura molecular

- Es preciso familiarizarse con los métodos, sus ventajas y limitaciones
- Es preciso adquirir cierta habilidad en el uso de las computadoras personales
- La base material puede ser la misma que la de cualquier oficina

Beneficios

Se ahorra tiempo de trabajo experimental para:

- muchas acciones de prueba y error
- comprender lo que está ocurriendo

Esta es una ciencia en la que el valor se crea en el conocimiento y en las técnicas de manipular información

Se trata de buena ciencia a bajo costo...

Es hermoso saber que, en estos tiempos tan materiales, el éxito científico no está preservado solo a la disponibilidad de recursos materiales, sino también al talento, la creatividad, al amor por la ciencia y a otros valores.