

K  
UNIVERSITAT DE VALENCIA  
FACULTAD DE CIENCIAS QUIMICAS

---

"GEPOL: UN METODO PARA EL CALCULO DE  
SUPERFICIES MOLECULARES"

MEMORIA

que, para optar al Grado de Doctor en Ciencias  
Químicas, presenta el Licenciado en Ciencias Químicas:

Juan Luis Pascual-Ahuir Giner

DEPARTAMENTO DE QUIMICA-FISICA-

Burjassot, Febrero de 1988

UMI Number: U603128

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U603128

Published by ProQuest LLC 2014. Copyright in the Dissertation held by the Author.  
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against  
unauthorized copying under Title 17, United States Code.



ProQuest LLC  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

FACULTAD QUIMICA  
BIBLIOTECA

Cat. 296  
Encuad.  
Ent. 4414 Q  
Mat. 541.2(043)  
Sell.  
Test. T.D.Q  
Top. 340

Ri. 18684932

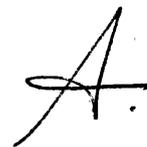
R. 4414

ESTANISLAO B. SILLA SANTOS Y JUAN ANDRES BORT, PROFESORES  
TITULARES DEL DEPARTAMENTO DE QUIMICA-FISICA DE LA UNIVERSITAT DE  
VALENCIA.

CERTIFICAMOS: Que D. Juan Luis Pascual-Ahuir Giner, Licenciado en  
Ciencias Químicas, ha realizado bajo nuestra  
dirección, desde Octubre de 1983 hasta la actualidad,  
el trabajo que se recoge en esta Memoria, para aspirar  
al Grado de Doctor en Ciencias Químicas. Así mismo,  
que los valores numéricos que en ella se consignan,  
responden a las investigaciones que se han realizado  
en el Departamento de Química-Física de esta  
Facultad.

Y para que así conste a los efectos legales,  
presentamos ante el Departamento de Química-Física  
de la Universitat de València dicha Tesis, firmando  
este Certificado en Burjassot, a veintinueve de  
Febrero de mil novecientos ochenta y ocho.

Handwritten signature of Estanislao B. Silla Santos, starting with a stylized 'E' and 'S'.

Handwritten signature of Juan Andres Bort, consisting of a large, stylized 'A'.

En memoria de mi padre.



El presente trabajo se ha realizado bajo la dirección de los Doctores Estanislao B. Silla Santos y Juan Andres Bort, a quienes quiero hacer constar mi agradecimiento por toda la ayuda, consejos y orientación prestada en la consecución del mismo.

Deseo expresar mi gratitud:

A todos los miembros del departamento de Química-Física de esta Facultad que con sus comentarios, actitud y apoyo han hecho más llevadero este trabajo y en especial a Concha Sanz, Eloisa Cordoncillo, Arturo Arnau e Ignacio Nebot.

A todos los miembros del Departamento de Química-Física de la Facultad de Químicas de Tarragona, donde trabaje dos años, el apoyo dado, al permitirme desplazarme a Valencia con frecuencia, para continuar con este trabajo.

A Rosana Bonaccorsi y a Jacopo Tomasi que me iniciaron en el conocimiento del estudio teórico del efecto del disolvente, así como de todos los miembros del Dipartimento di Chimica de la Universidad de Pisa y del Istituto di Chimica Quantistica ed Energetica Molecolare (CNR) de Pisa (Italia), que hicieron encantadora mi estancia allí y de los que guardo un grato recuerdo.

A Orlando Tapia, con el que empecé a adentrarme en el maravilloso y desconcertante mundo de la Biología Molecular, así como a todos los miembros del Departamento de Biología Molecular del Centro Biomédico de Uppsala (Suecia), por su paciencia al descifrar mi Inglés.

A Pedro, Jesús, Juan Carlos que han realizado dibujos y correcciones en esta memoria y a otros tantos amigos que me han ayudado y soportado a lo largo de este período.

Y por último a mi madre y hermanos que, sin su apoyo y cariño, no hubiese acabado esta Tesis.

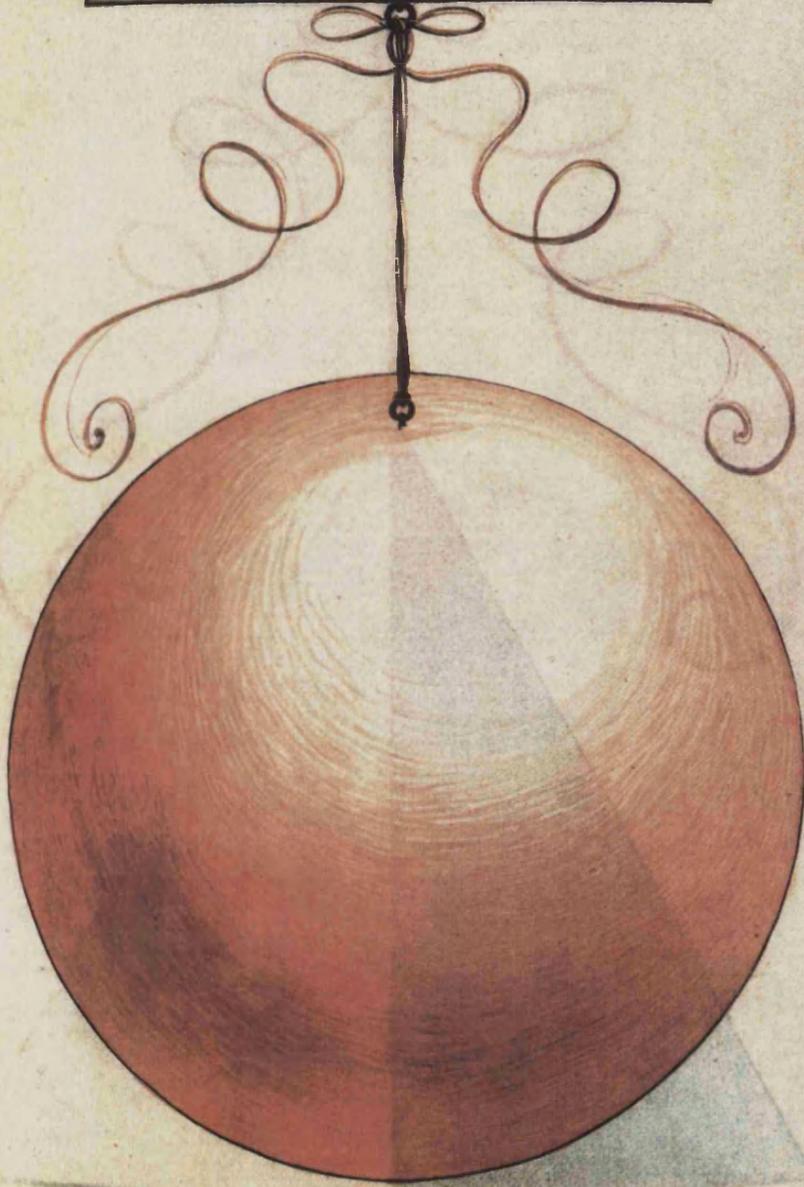
" Lo que quiero decir es simplemente que, cuanto más participe de la perfección sea la ciencia que les es dado a los hombres alcanzar por medio de la demostración y del razonamiento, tanto menor será el número de proposiciones que prometa enseñar y menor aún el de las que demuestre; en consecuencia, tanto menos atractiva resultará y tanto menor será el número de sus seguidores. Por el contrario, la magnificencia de los títulos y la profusión y magnitud de las promesas atraerá la natural curiosidad de los hombres y les mantendrá perpetuamente enmarañados en falacias y quimeras, sin brindarles jamás la posibilidad de conocer esa agudeza de la demostración que pudiera despertar su gusto y hacerles ver cuán insípido es su sustento habitual" (Galileo Galilei, *Il Saggiatore*, VI, 237).

<b><u>1.INTRODUCCION</u></b> .....	1
<b><u>2.METODO GEPOL</u></b> .....	7
<b>2.1 METODO DE SUAVIAZACION</b> .....	10
<b>2.2 DETERMINACION DE LA SUPERFICIE</b> .....	28
2.2.1 Subrutinas TES y DIVIDE.....	31
2.2.2 Subrutina GEOCAV.....	35
<b>2.3 ENTRADA DE DATOS Y SALIDA DE RESULTADOS</b> .....	38
2.3.1 Entrada de datos para GEPOL ( INPUT).....	38
2.3.2 Salida de resultados del GEPOL (OUTPUT).....	40
<b>2.4 CALCULOS DE SUPERFICIES DE VAN DER WAALS</b> .....	42
2.4.1 Comparación con SPHERE. Dos esferas.....	44
2.4.2 Diez esferas.....	46
2.4.3 Electrón solvatado. 19 esferas.....	54
2.4.4 NAD y LOOP. 165 esferas.....	54
2.4.5 CTF. 519 esferas.....	61
2.4.6. RBP. 1438 esferas.....	61
2.4.7 Discusión general.....	67
<b>2.5 CALCULOS DE SUPERFICIES MOLECULARES</b> .....	72
2.5.1 NAD y LOOP. Reducción de RET y FRO a FRADIO.....	72
2.5.2 NAD y LOOP. Variación con OMEGA.....	77
2.5.3 NAD y LOOP. Análisis de resultados.....	77
2.5.4 Electrón solvatado.....	80
2.5.5 RBP.....	83
<b>2.6 EJEMPLOS DE APLICACIONES</b> .....	88
2.6.1 Disociación de una molécula diatómica.....	88
2.6.2 Estudio de una rotación interna del octanol.....	96
2.6.3 Volumen de las aminas.....	99
2.6.4 Estudio de la fractalidad en el RBP.....	106
2.6.5 Estudio de la relación superficie-volumen en el PMMA.....	108
<b><u>3.INTERACCION SOLUTO-DISOLVENTE</u></b> .....	112
<b>3.1 FORMULACION BASICA DEL METODO</b> .....	116
<b>3.2 SUPERFICIE ENVOLVENTE Y ENERGIA</b> .....	125
3.2.1 HF.....	125
3.2.2 H <sub>2</sub> CO.....	127
3.2.3 Rotación del octanol.....	128
<b>3.3 BASICIDAD DE LAS AMINAS EN DISOLUCION</b> .....	130
<b><u>4.CONCLUSIONES</u></b> .....	142
<b><u>5.BIBLIOGRAFIA</u></b> .....	144

## 1.INTRODUCCION

LXXXXI

SPHERA SOLIDA.



XLI

ΣΦΑΙΡΑ ΣΤΕΡΕΑ



Dentro de la Química Física, Langmuir fue el primero en introducir el concepto de área superficial de un soluto para determinar su actividad. El consideraba que el factor predominante en las relaciones de solubilidad era la energía requerida para crear una cavidad en el solvente en el que está sumergido el soluto, considerando que esta energía es proporcional al área del soluto (Langmuir, I., 1925). Posteriormente se han realizado muchos estudios en los que se han correlacionado propiedades como la solubilidad y la hidrofobia de diversas moléculas con el área de éstas. En concreto podemos citar los trabajos de R.B. Herman, 1972, y S. C. Valvani y col, 1976, que estudiaron, respectivamente, la solubilidad de hidrocarburos y alcoholes frente al área de estos. Estos estudios, y otros muchos, han servido para desarrollar un conjunto de modelos y métodos que hacen uso tanto del área y volumen como de la misma superficie para explicar y cuantificar las interacciones de una molécula con sus alrededores (Pierotti, R. A., 1976; Miertus, S. y col, 1981; Sinanoglu, O., 1980, ).

Desde que Linus Pauling postulara que las interacciones entre moléculas de interés biológico, tales como ácidos nucleicos y proteínas, vienen moduladas en primer lugar por su forma y tamaño (Pauling, L., 1948), la Bioquímica ha sido uno de los campos de la química donde ha tenido más eco el estudio de las superficies moleculares. La topología de la superficie de una proteína está íntimamente relacionada con su función. Partes de la superficie están directamente relacionadas con las interacciones con otras moléculas y la interfase disolvente proteína está también relacionada con la estructura de la molécula nativa. Entre las múltiples cuestiones abordadas con ayuda del concepto de superficie molecular están el estudio del plegamiento (folding) de las proteínas y el acoplamiento entre éstas o "docking". Baste recordar el principio tan extendido de la llave y la cerradura.

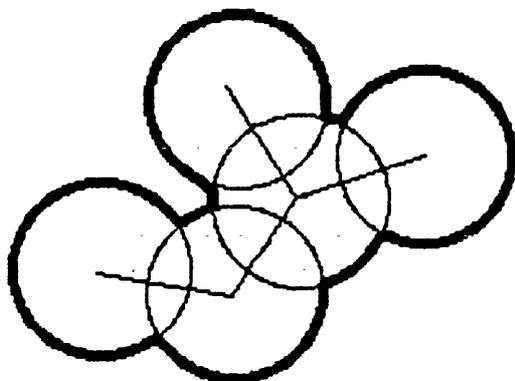
Los dos párrafos anteriores hacen evidente la necesidad de crear métodos fiables que calculen estas superficies, pero sin embargo: ¿podemos afirmar que existen en realidad estas superficies?

Determinar el espacio que ocupe una molécula resulta un problema muy sutil, ya que no existen límites claros de donde empieza y acaba ésta. De hecho la Mecánica Cuántica predice que un electrón perteneciente a una molécula puede encontrarse a una distancia grandísima de ésta,

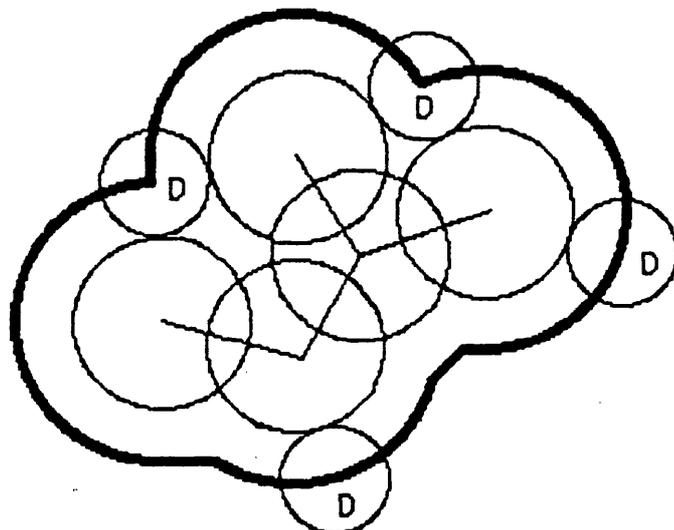
haciéndose nula la probabilidad de encontrarlo, sólo en el infinito. En principio este problema puede resultar limitante, pero se puede encontrar soluciones, tal como sucedió en la definición de la geometría de una molécula, optándose por la geometría de equilibrio o, en otros casos, la promedio. La solución más aceptada es considerar los átomos como esferas rígidas tales que su radio sea la distancia mínima a la que se pueda acercar otro átomo sin llegar a formar un enlace químico. Esta definición, aunque difusa, reduce el campo desde el infinito a un rango de radios llamados de van der Waals.

Una vez definidos los radios de van der Waals, podemos pasar a definir el primer tipo de superficie que es la llamada Superficie Molecular de van der Waals (SMW) y que hemos representado en la **figura 1.a**. Esta superficie está formada por un conjunto de esferas de radios de van der Waals centradas sobre los átomos de la molécula. Sin embargo, esta definición no es del todo correcta, ya que, por ejemplo, en una disolución las moléculas de agua no pueden acceder a ciertas zonas del espacio que sin embargo por la propia definición de la superficie tampoco pertenecen al soluto. Por ello **B. Lee y F.M. Richards** en **1971** definieron la Superficie Molecular Accesible, representada en la **figura 1.b**. Esta superficie está definida por el desplazamiento del centro de una esfera, que llamaron esfera de prueba, al rodar sobre la superficie de van der Waals. Esta esfera de prueba puede, por ejemplo, representar al disolvente. Uno de estos autores redefinió la superficie por una más real, representada en la **figura 1.c**, denominándola Superficie Molecular (SM) (**Richards, F.M., 1977**). El nombre dado a esta última superficie puede crear alguna confusión, ya que coincide con el nombre genérico de las tres. Sin embargo lo hemos mantenido a lo largo de esta tesis, por ser el que se utiliza normalmente en la bibliografía. Para evitar confusiones, a partir de aquí utilizaremos como nombre genérico "la superficie envolvente de una molécula". Es interesante destacar que la definición de las dos últimas superficies conlleva el considerar que la superficie envolvente de una molécula no es una característica propia de ella, sino que puede variar dependiendo de la molécula que se le enfrente (radio de prueba), aunque estas variaciones no sean muy grandes.

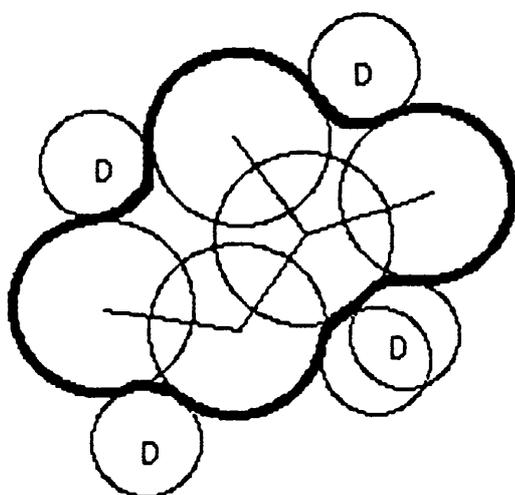
Se han desarrollado algunos métodos de cálculo de superficies pero la mayoría de ellos o



**Figura 1.a** Superficie Molecular de van der Waals.



**Figura 1.b** Superficie molecular Accesible



**Figura 1.c** Superficie molecular

sólo sirven para casos de estructuras sencillas, o sólo calculan el área total o el volumen, o simplemente crean un sistema para poder obtener la superficie en una pantalla grafica. Sin embargo hay dos métodos que han sido ampliamente utilizados, uno es el **R.B. Herman, 1972**, que sólo permite el cálculo del área total de la SMW y de la SMA, el otro es el de **M. L. Connolly, 1983b**, que permite calcular valores del área total y de áreas locales de las 3 superficies, representando la superficie envolvente con un conjunto de puntos sobre ésta. Sin embargo, este último método presenta problemas en la representación de las áreas locales. Tengase en cuenta que muchas de las interacciones moleculares tienen lugar en zonas muy restringidas y, por tanto, es interesante reproducir adecuadamente los valores de las áreas locales.

Así, el principal objetivo de esta tesis es desarrollar un método de cálculo de superficies envolventes que pueda ser aplicado a cualquier molécula y que nos dé el valor de área tanto total como local, el valor del volumen total encerrado por la superficie y que nos describa a ésta de tal manera que podamos conocer su forma. A este método lo hemos denominado GEPOL.

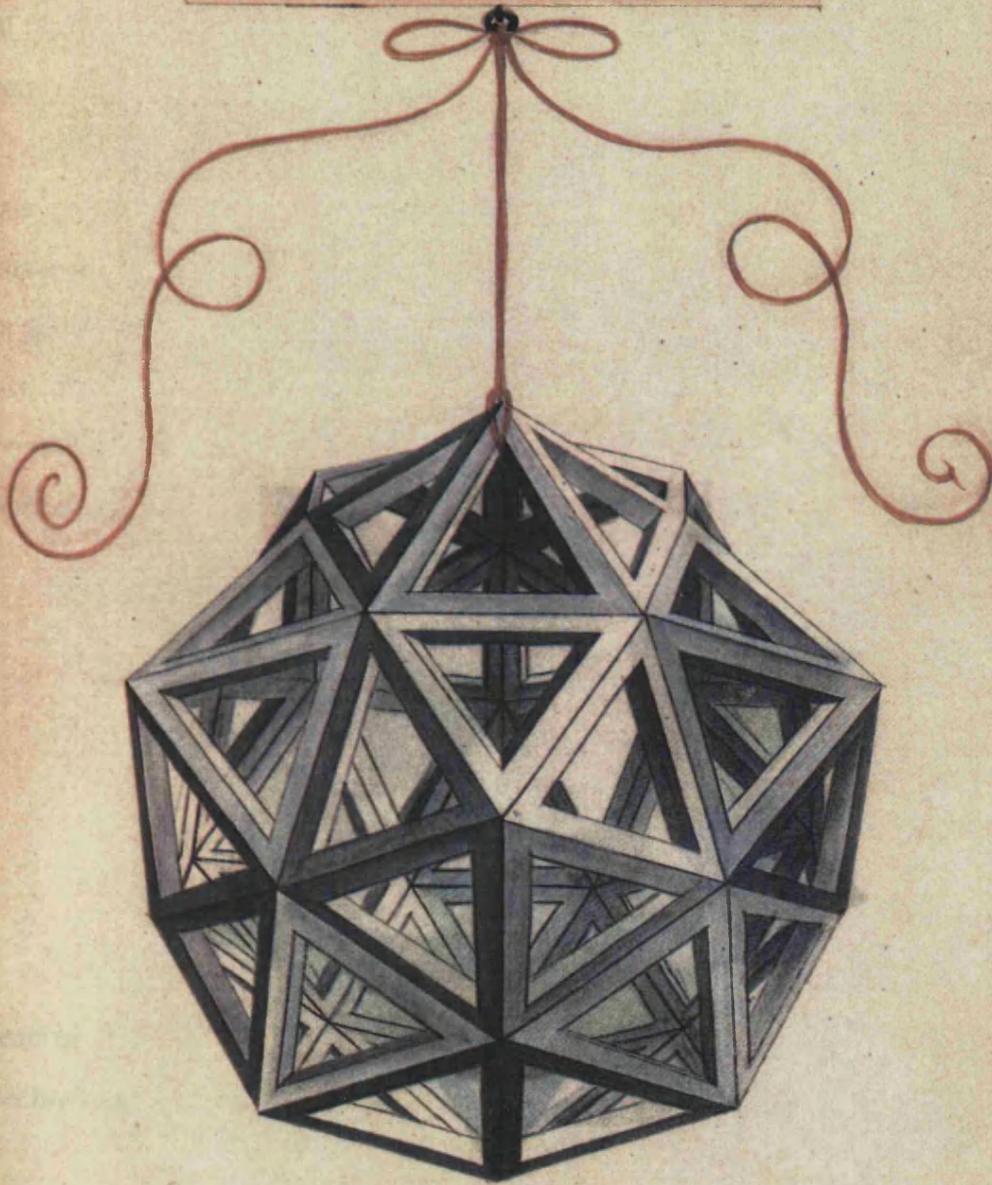
En el primer párrafo de esta introducción hablamos de como se habían correlacionado valores de solubilidad con las áreas de las moléculas. Otro objetivo de la tesis es mostrar como la superficie envolvente puede utilizarse como instrumento para construir un modelo que describa las propiedades de las moléculas en disolución. Para ello utilizamos el modelo de continuo de **Miertus, S. y col, 1981** para el cálculo de la función de onda de una molécula en disolución. Este método ha sido utilizado con otro método de cálculo de superficies, que sólo permite el uso de la SMW. El objetivo en esta tesis ha sido sustituir la parte geométrica del método original de estos autores por la desarrollada por nosotros y analizar las mejoras que introduce el poder realizar los cálculos con una superficie más real como es la SM.

Por último, y como muestra de las posibilidades de utilización del método GEPOL, queremos indicar que nuestro grupo, en colaboración con el grupo que dirige el Prof. J. Tomasi en Pisa, ha desarrollado un método de cálculo de la energía de dispersión que utiliza el método GEPOL para describir la superficie de la molécula, y que actualmente está investigando las posibilidades de utilizarlo para el cálculo de la energía de cavitación.

## 2. METODO GEPOL

. CVII .

DVODECEDRON ELEVA  
TVS VACVVS.

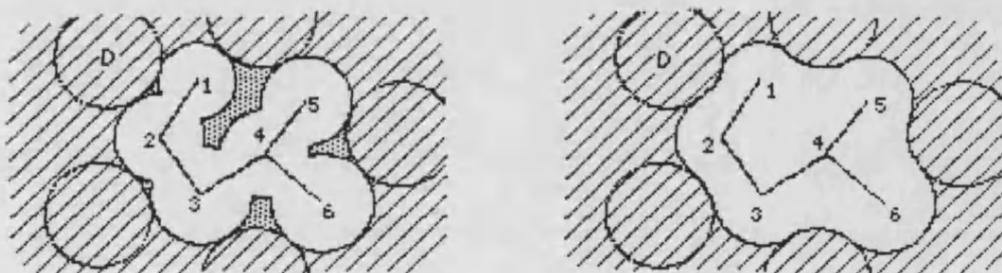


XXXII

Ααλεκεδρον εἰς τὸν κενόν.

El objetivo del programa es doble, por un lado obtener una distribución de puntos sobre la hipotética superficie molecular, de tal manera que ésta quede definida por ellos, y por otro lado tener unos buenos valores del área y el volumen de la molécula. Para obtener esto el método GEPOL, al igual que los métodos antes mencionados, toma como punto de partida de los cálculos un conjunto de esferas, de radio de van der Waals, centradas sobre cada átomo o grupo de átomos.

El programa GEPOL se basa principalmente en dos algoritmos. El primero corresponde a lo que denominamos suavización de la superficie. Como vimos en el capítulo anterior el concepto de superficie no es nada estricto y se pueden definir diferentes tipos de superficie en función de la propiedad en estudio. Así el algoritmo de suavización, partiendo de las esferas de van der Waals utilizadas como punto de partida, genera un nuevo conjunto de esferas, que, llenando los intersticios no accesibles al disolvente, formará lo que se denomina superficie molecular. En la **figura 2. a** se muestra este proceso.



**Figura 2. a**

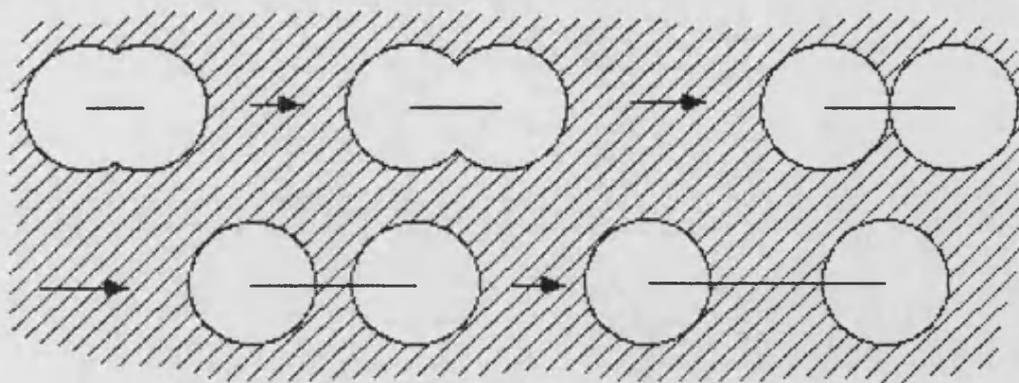
El segundo algoritmo es el que realiza propiamente el cálculo geométrico, partiendo de las esferas de van der Waals, si queremos conocer esta superficie, o partiendo de todas las esferas, incluyendo las creadas por el algoritmo de suavización, para obtener la superficie molecular.

Los apartados de este capítulo forman dos bloques, en el primero de ellos mostramos el funcionamiento de los algoritmos mencionados, así como la entrada y salida del programa (Apartados 2.1, 2.2 y 2.3). En el segundo aplicamos este método a diferentes estructuras moleculares con el fin de analizar su fiabilidad (Apartados 2.4, 2.5 y 2.6).

## 2.1 METODO DE SUAVIZACION

Las consecuencias de una buena suavización de la cavidad son importantes siempre, pero hay algunos casos en los que juega un papel privilegiado, dadas las características del estudio que se realiza. Así, al investigar la rotación interna de una molécula grande, la variación de la superficie y volumen de la cavidad con el ángulo de rotación va a influir apreciablemente sobre las diferentes componentes energéticas (electrostática, cavitación, dispersión,.....). Un estudio de este tipo se realiza en un apartado posterior.

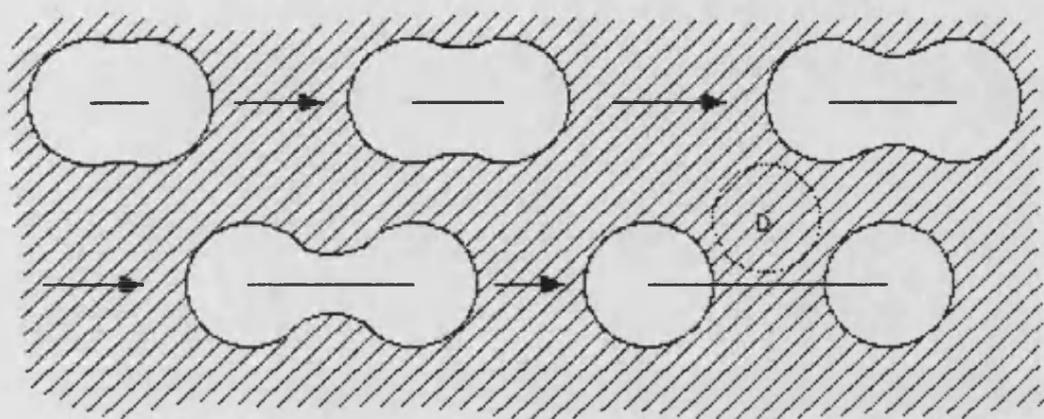
Otra situación donde la forma de la cavidad podría cambiar con la suavización, corresponde al estudio de una reacción de disociación. Así, por ejemplo, en las **figuras 2.1 a** y **b** se da el proceso de separación de dos esferas, que podría representar la disociación de una molécula diatómica. La partición de la cavidad en dos no puede ser tan brusca como aparece en la **figura 2.1.a**, sino que el proceso debe estar mejor representado en la **figura 2.1.b**, en la que se tiene en cuenta la accesibilidad del disolvente.



**Figura 2.1 a**

Dada la gran diversidad de casos y que muchos, por su complejidad, no se podrían tratar singularmente, es decir indicando desde el input su geometría adecuada, ideamos un algoritmo que partiendo de la definición normal de esferas interconectadas, modificaba automáticamente la forma de la cavidad. Para poder obtener este automatismo había que resolver de una manera sencilla dos

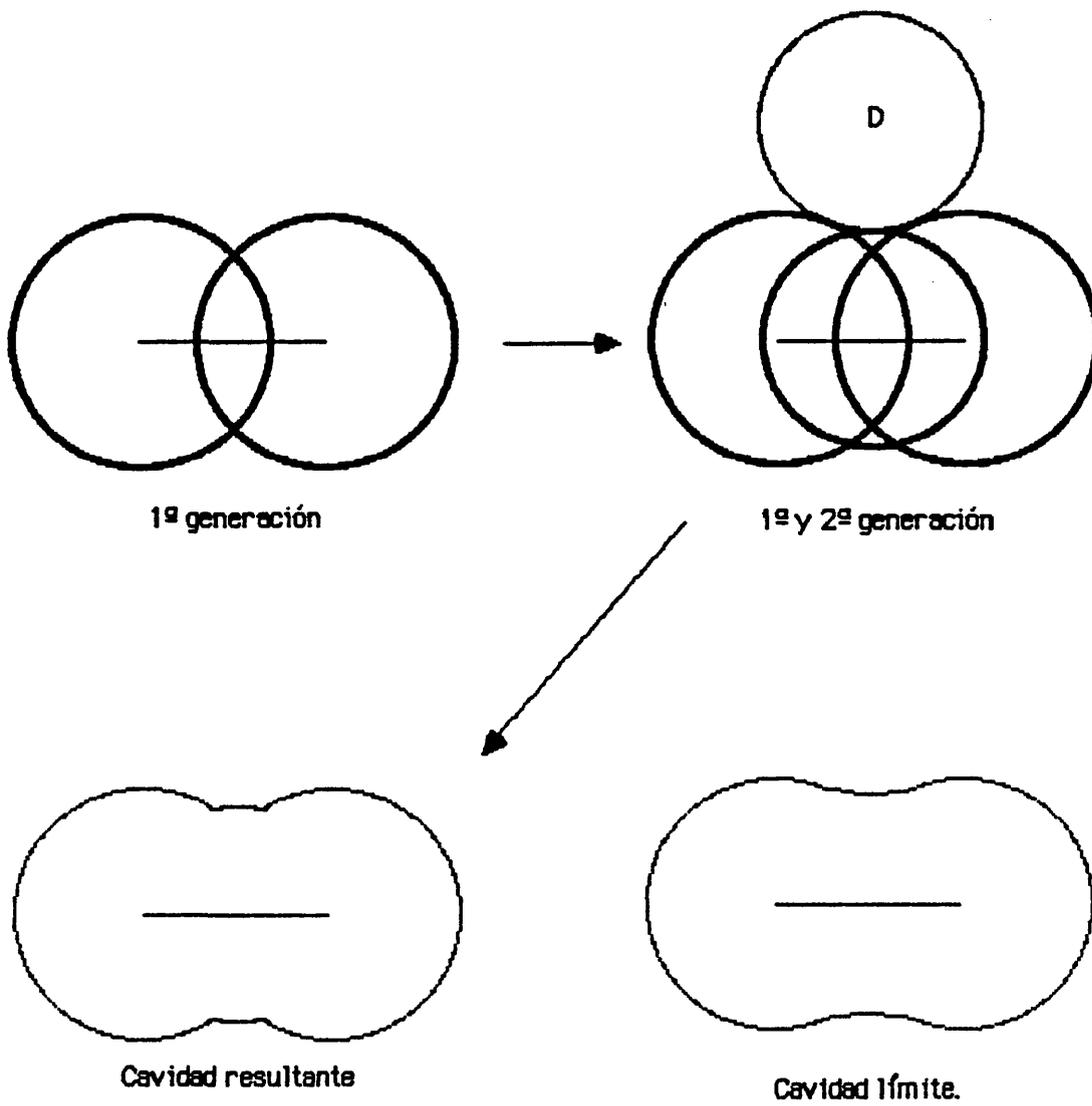
cuestiones. La primera es como detecta el algoritmo las zonas de la cavidad que deben ser modificadas, y segundo, una vez detectadas, como se deben de modificar. A continuación se dará una explicación general, para posteriormente pasar a explicar detalladamente cada uno de los pasos.



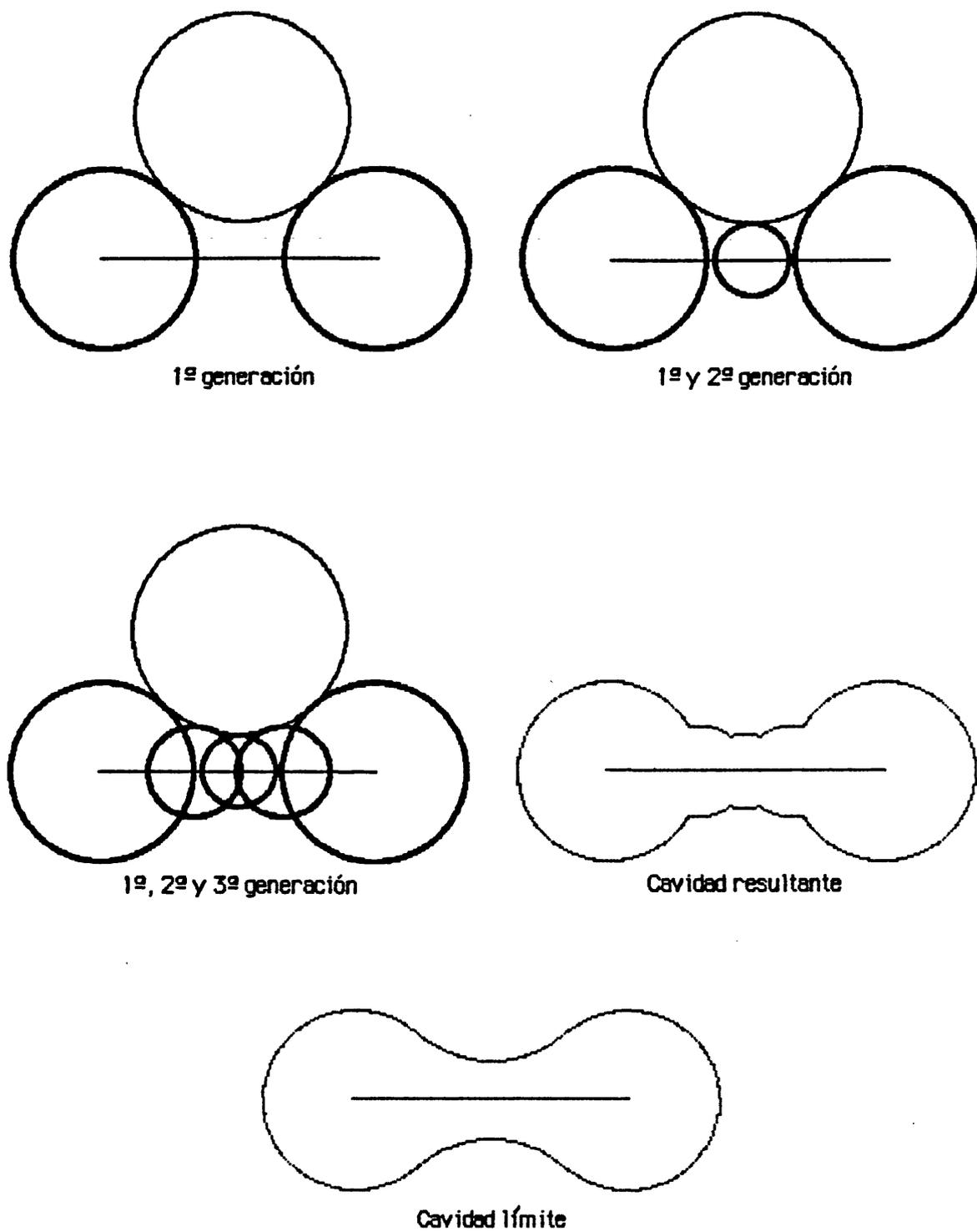
**Figura 2.1 b**

El análisis se extiende a todos los pares de esferas que forman la cavidad (en la **figura 2.a** tenemos 6 esferas y por tanto 15 pares diferentes), determinándose si es necesaria una modificación al comparar la distancia que separa las dos esferas de cada par con sus radios y con el tamaño del disolvente. De esta manera se pueden encontrar tres tipos de pares entre los que se crearán nuevas esferas a los que denominaremos A, B y C y que se muestran en las **figuras 2.1 c, d y e**. Cada vez que el programa detecte alguno de estos casos creará una esfera nueva entre las esferas del par, teniendo en cuenta el tamaño del disolvente. Una vez que ha analizado todos los posibles pares iniciales (dados en el INPUT), suma las esferas nuevas a las iniciales y reinicia el proceso de localización y creación. Este proceso se repite hasta que no se crea ninguna esfera más. Nosotros llamamos a las esferas iniciales, esferas de primera generación. A las primeras que se crean entre éstas de segunda generación, a las siguientes de tercera generación y así sucesivamente.

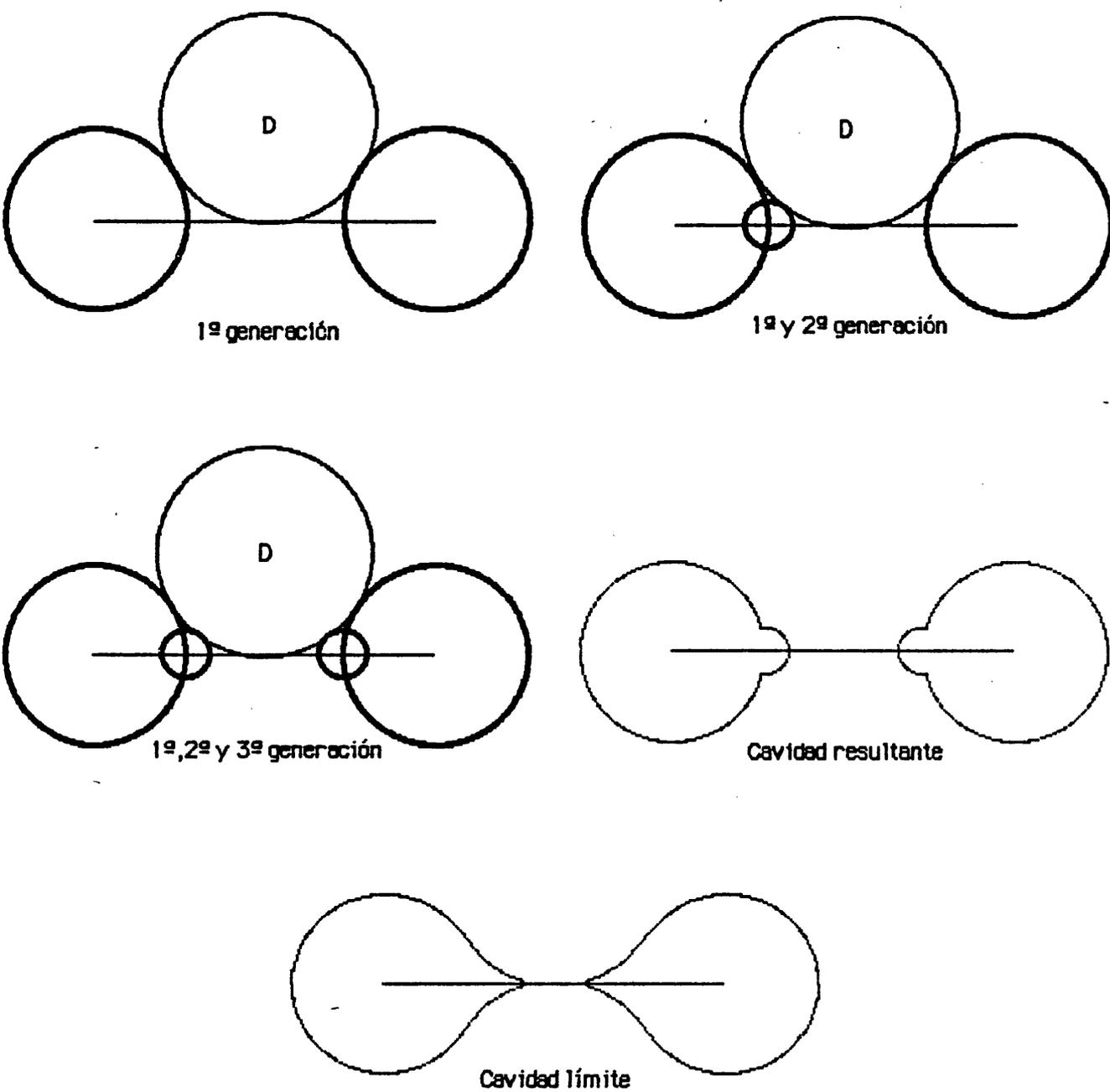
Tal como se ha descrito el proceso hasta ahora, éste nos llevaría a la creación de infinitas esferas, dando lugar a una cavidad que denominaremos límite. Logicamente esto es irrealizable, y por tanto hay que truncarlo de alguna manera. Para ello existen una serie de frenos, que se pueden



**Figura 2.1.c.** Par de esferas del tipo A



**Figura 2.1.d.** Cavidad formada por un par de esferas del tipo B.



**Figura 2.1.e.** Cavidad formada por un par del tipo C.

controlar desde el INPUT, y que limitan el número de esferas.

En las figuras 2.1 c, d y e se puede observar el proceso seguido para llegar a una cavidad definitiva partiendo de cavidades formadas por pares del tipo A, B y C respectivamente, mostrándose generación a generación las esferas que se van creando. También se muestra la cavidad que resultaría de un proceso de creación de infinitas esferas. Los pares de tipo A son aquellos en que las esferas, aun solapándose, no lo hacen suficientemente. Los de tipo B y C están formados por esferas que no se tocan. Se diferencian en que en las primeras el tamaño del disolvente permite crear una nueva esfera en un punto equidistante de las dos superficies iniciales, y en las de tipo C no.

En el esquema 2.1 a, se indica en líneas generales el flujo que sigue un par de esferas, para determinar si es necesaria la creación de otra entre ellas. Este proceso se ha resumido en 8 fases. Si en alguna de las fases se determina que no hay que crear ninguna esfera entre ese par, se dirá que el par es eliminado.

### FASE 1.

Consiste simplemente en seleccionar el par de esferas que se va a estudiar, a la vez se determina qué esfera del par es la grande y cual la pequeña, denominándolas G y P respectivamente. En fases posteriores se justificará esta diferenciación.

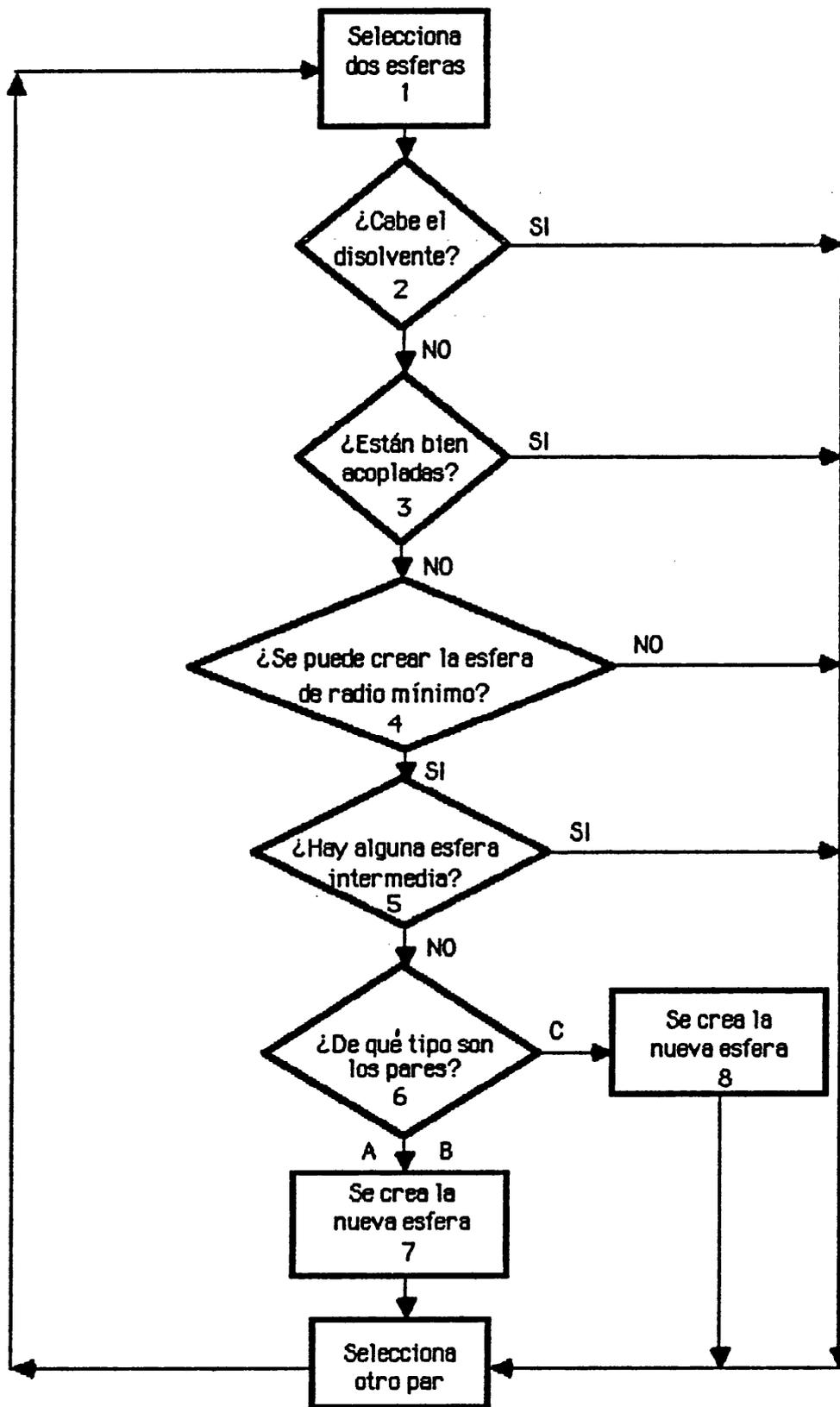
### FASE 2.

Determina si el disolvente puede pasar entre las dos esferas G y P. Para ello calcula la distancia,  $d_{PG}$ , entre los centros de las dos esferas y se compara con la suma de los radios del par ( $R_P$  y  $R_G$ ) más el diámetro del disolvente ( $R_D$ ). Si la distancia es mayor o igual, resulta que el disolvente puede pasar entre ellas. No hace falta ninguna modificación y se elimina el par.

$$d_{PG} \geq R_P + R_G + 2R_D$$

En el caso contrario este par pertenece necesariamente a uno de los tres tipos mencionados anteriormente.

Cuando se habla del diámetro del disolvente, se está considerando implícitamente que éste

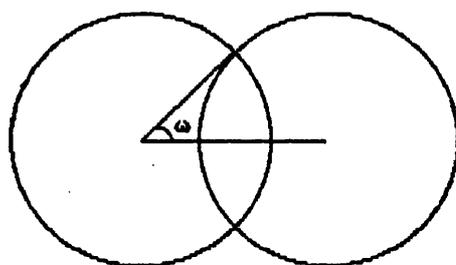


**Esquema 2.1.a** Diagrama general de la suavización.

viene representado por una esfera rígida. Esto puede ser bastante correcto para el caso del agua, dado su tamaño y geometría, pero en el caso de moléculas de disolvente con estructura distante a la esférica, la decisión es más difícil y habrá que estudiar detenidamente la morfología de ésta.

### FASE 3

En esta fase se introduce uno de los frenos antes comentados. En ella se determina si este par de esferas están bien solapadas y por tanto no es necesario la creación de una esfera entre ellas. El solapamiento se valora comparando el ángulo  $\omega$ , indicado en la **figura 2.1.f**, que forman las dos esferas, con un ángulo óptimo mínimo de solapamiento  $\omega_m$ , dado en el INPUT. Cuanto mayor sea  $\omega_m$  mayor será la exigencia de solapamiento y mayor será el número de esferas que habrá que crear. Este ángulo puede variar entre 0 y 90°. Hay que indicar que el ángulo está definido desde la esfera pequeña, como se puede ver en la **figura 2.1.f**. De otra manera aparecerían tantos casos particulares que harían inoperativo este freno.

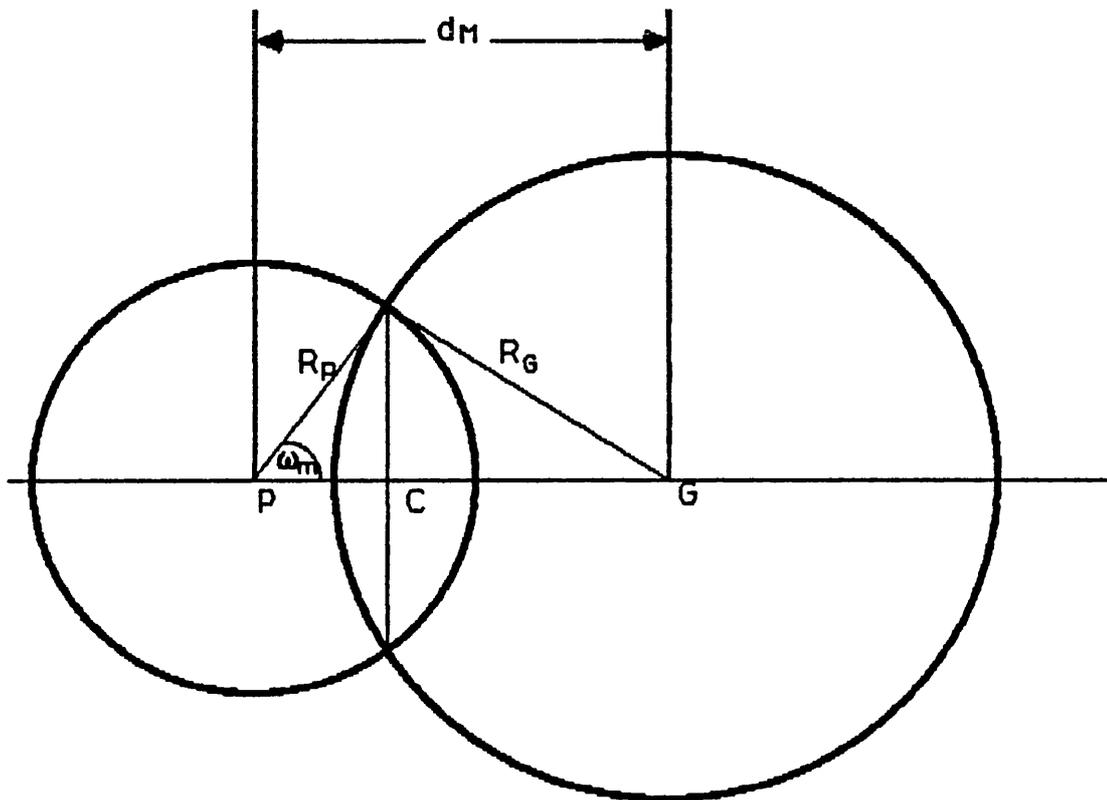


**Figura 2.1.f.** Angulo de solapamiento  $\omega$ .

El funcionamiento real de esta fase tiene lugar comparando distancias, como a continuación veremos.

Primero calcula la distancia óptima máxima de solapamiento,  $d_M$ , a la que deberían estar las dos esferas para formar un ángulo  $\omega$  igual a  $\omega_m$ . Para el cálculo de esta distancia se hace uso de trigonometría elemental como se muestra en la **figura.2.1.g**

Una vez conocida esta distancia la compara con la distancia real,  $d_{PG}$ , entre las dos esferas.



$$d_M = \overline{PC} + \overline{CG}$$

$$\overline{PC} = R_p \cos \omega_m$$

$$\overline{CG} = \sqrt{R_g^2 - R_p^2 \sin^2 \omega_m}$$

$$d_M = R_p \cos \omega_m + \sqrt{R_g^2 - R_p^2 \sin^2 \omega_m}$$

**Figura 2.1.a.** Cálculo de la distancia máxima óptima de solapamiento.

Sí  $d_{PG}$  es menor o igual que  $d_M$  las dos esferas están bien solapadas y este par es eliminado ya que no necesita ninguna modificación. En caso contrario el par pasa a la siguiente fase.

#### FASE 4

Otra manera de frenar la creación de esferas, consiste en delimitar el tamaño de las nuevas esferas, es decir, no dejar que se creen esferas con un radio menor que un radio mínimo,  $R_m$ , dado en el INPUT.

Primero se calcula la distancia máxima,  $d_M$ , a la que deben de estar las dos esferas para permitir, teniendo en cuenta el tamaño del disolvente, crear una esfera de radio mayor que  $R_m$  (ver **figura 2.1.h**). Si la distancia real  $d_{PG}$  es mayor que  $d_M$ , no se podrá crear esta esfera y por tanto se eliminará el par. En el caso contrario el par pasará a la siguiente fase.

Por motivos de economía de cálculo, en esta fase se considera que las dos esferas tienen el mismo radio y que éste es igual al radio de la más grande  $R_G$ . Esto hace que la eliminación de pares que no cumplen esta condición no sea total, creándose algunas esferas de radio menor que  $R_m$ , las cuales son eliminadas en la fase 8.

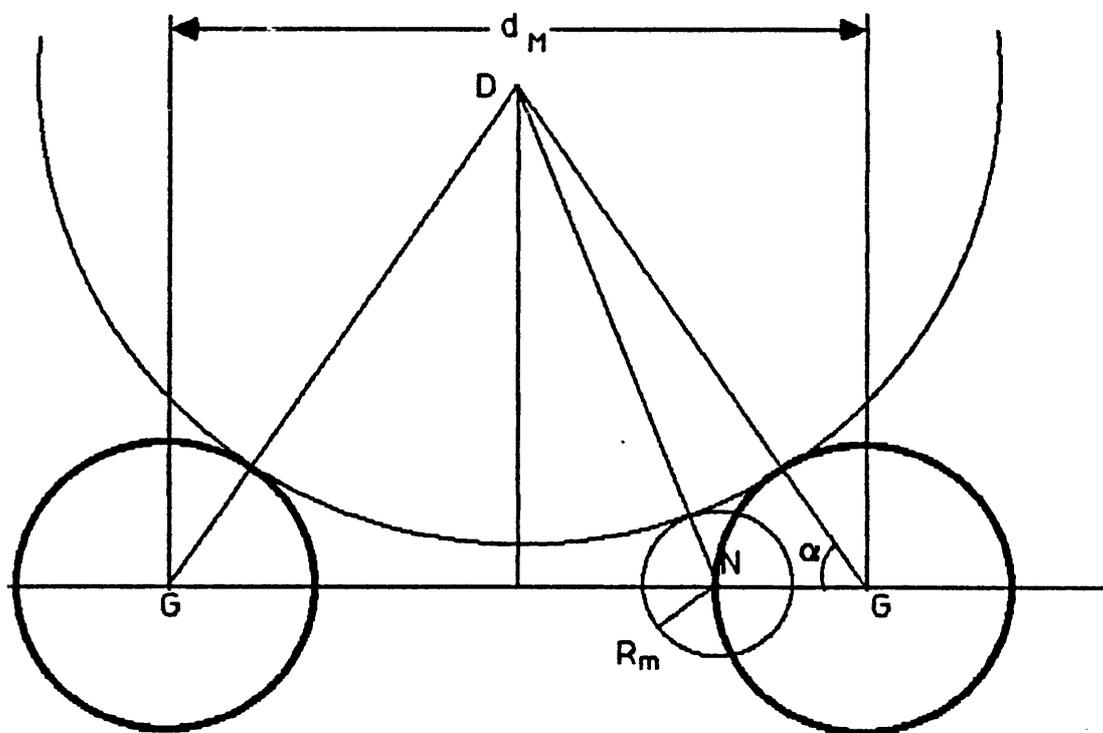
#### FASE 5

En ésta se determina si se ven los centros de las esferas del par en estudio, o por el contrario si existe alguna esfera intermedia,  $I$ , entre ellas, en cuyo caso no se realiza ninguna modificación.

En general este proceso se realiza comparando la distancia  $H$ , que hay desde cualquier esfera  $I$  al eje que une los centros de  $P$  y  $G$ , con el radio de  $I$ ,  $R_I$  (ver **figura 2.1.i**). Si ésta distancia es menor que  $R_I$  quiere decir que la esfera  $I$  oculta a  $P$  de  $G$  o viceversa, por lo cual no se creará ninguna esfera entre ellas.

Para flexibilizar un poco este proceso y la vez poderlo controlar desde el INPUT, se introdujeron los siguientes aspectos:

a) Para agilizar los cálculos no se considerarán en esta fase aquellas esferas  $I$  que se



$$\frac{d_M}{2} = \overline{GD} \cos \alpha$$

$$\cos \alpha = \frac{\overline{GD}^2 + R_G^2 - \overline{ND}^2}{2 \overline{GD} R_G}$$

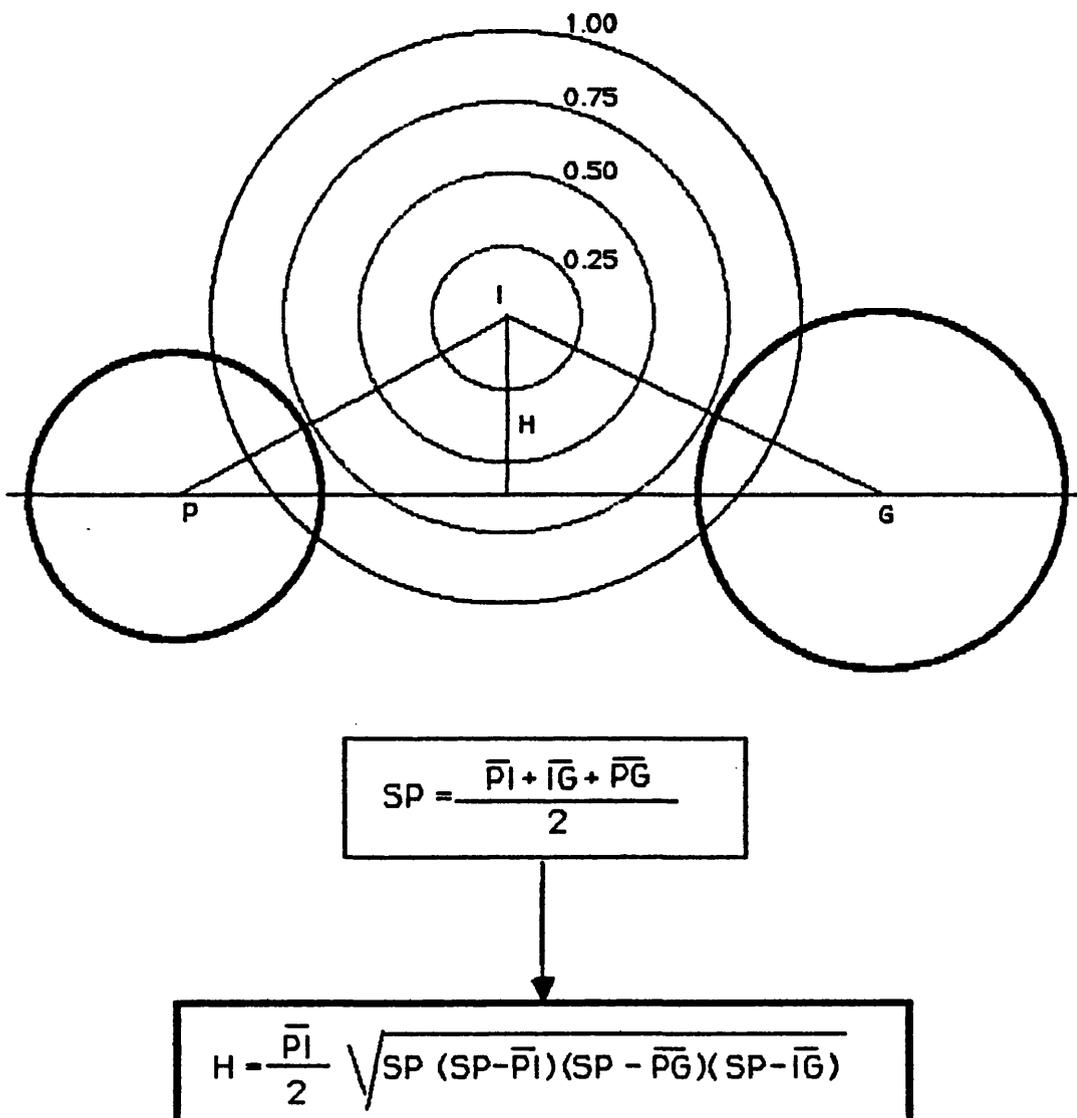
$$d_M = \frac{\overline{GD}^2 + R_G^2 - \overline{ND}^2}{R_G}$$

$$\overline{GD} = R_D + R_G$$

$$\overline{ND} = R_D + R_m$$

$$d_M = \frac{(R_D + R_G)^2 + R_G^2 - (R_D + R_m)^2}{R_G}$$

**Figura 2.1.h.** Cálculo de la distancia óptima máxima para que se pueda crear una esfera con radio mínimo,  $R_m$  (suponiendo que el par de esferas en estudio tienen el mismo radio,  $R_G$ )



**Figura 2.1.i.** Ocultación de esferas. Cálculo de la distancia del centro de la esfera I al eje PG. Diferentes radios de ocultación para factores de 0.25, 0.50, 0.75 y 1.00.

encuentren a una distancia de P o de G mayor que la distancia entre ellas,  $d_{PG}$ .

b) Se dividen las esferas I de la cavidad en dos grupos. El primero incluye todas las esferas de generaciones anteriores a las que se están creando, y el segundo corresponde a las esferas de la nueva generación.

c) A las esferas del primer grupo se les asocia, a efecto de este test, un radio que llamaremos  $R_0$  (radio de ocultación) y que será igual a su radio real  $R_1$  por un factor FRO que se da desde el INPUT. Este factor, que puede variar entre 0 y 1, nos servirá para controlar desde el exterior del programa la dureza de este test. Cuanto más pequeño sea este factor mejor se "verán" las esferas y por tanto más esferas nuevas se crearán. En la **figura 2.1.i** se puede observar que dependiendo del valor utilizado de FRO, las esferas P y G están ocultas por I o no.

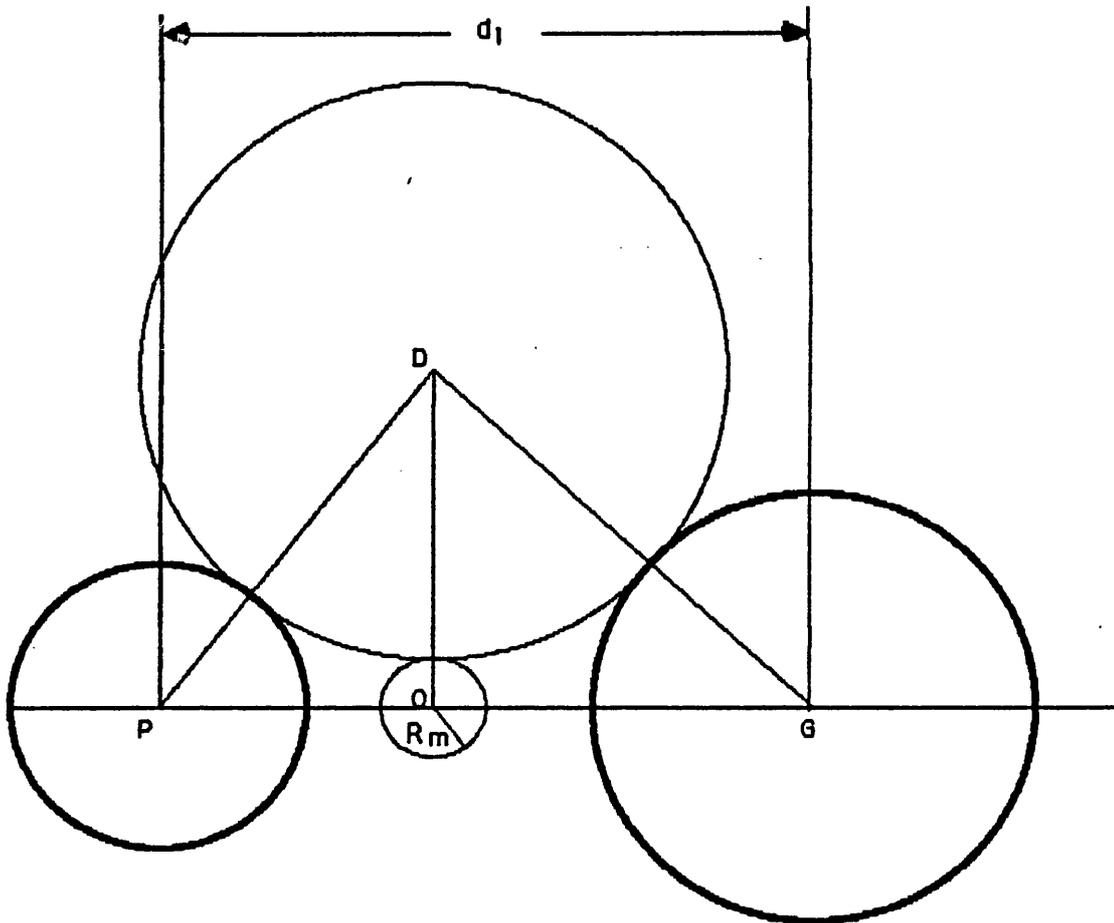
d) Las esferas del segundo grupo solo ocultarán a P y G si están situadas sobre el eje que une a estas dos esferas. Si no se introduce esta limitación, sucede que, cuando se parte de una cavidad simétrica, el resultado después de la creación de las nuevas esferas puede ser una cavidad asimétrica.

## FASE 6

Los pares que llegan a este punto son del tipo A, B o C y además han pasado los frenos que se han puesto y por tanto hay que crear una esfera entre ellas de tal manera que quede modificada la cavidad. Este test diferencia el tipo de par de que se trata para que sigan caminos diferentes. Así, si el par correspondiente es del tipo A o B ira a la fase 7, y si es del tipo C a la 8.

La pregunta que se hace es si en el punto del eje PG más cercano a la superficie del disolvente, éste permite la creación de una esfera con un radio mayor que el radio mínimo óptimo,  $R_m$ . Si es así este par será del tipo A o B y si no será del C.

Para realizar esta pregunta, se calcula una distancia llamada distancia límite,  $d_l$ , que correspondería a la que debería haber entre P y G para que justo se crease una esfera de radio  $R_m$  en las condiciones antes mencionadas (ver **figura 2.1.j**). Entonces si  $d_{PG}$  es menor que  $d_l$  el par será del tipo A o B, y en el caso contrario será del C.



$$d_1 = \overline{PO} + \overline{OG}$$

$$\overline{PO} = \sqrt{(R_P + R_D)^2 - (R_D + R_m)^2}$$

$$\overline{OG} = \sqrt{(R_G + R_D)^2 - (R_D + R_m)^2}$$

$$d_1 = \sqrt{(R_P + R_D)^2 - (R_D + R_m)^2} + \sqrt{(R_G + R_D)^2 - (R_D + R_m)^2}$$

**Figura 2.1.i.** Distancia l mite entre las esferas de los tipos A y B y las del tipo C.

### FASE 7

Para este tipo de pares, A y B, se creará una esfera nueva centrada sobre un punto del eje PG equidistante de las superficies de las esferas que forman el par y con un radio tal que sea tangente a la esfera de disolvente. En las figuras 2.1.k y l se muestran los dos tipos de pares así como las fórmulas utilizadas para la determinación de las coordenadas y radio de la nueva esfera.

### FASE 8

Para los pares tipo C, la nueva esfera está centrada sobre el punto de intersección del eje PG con la superficie de la esfera más grande, G, y con un radio tal que sea tangente al disolvente. Puede suceder, como comentamos en la fase 4, que en esta fase aparezca alguna esfera con un radio menor que  $R_m$ , en tal caso no se crearía la nueva esfera (figura 2.1.m).

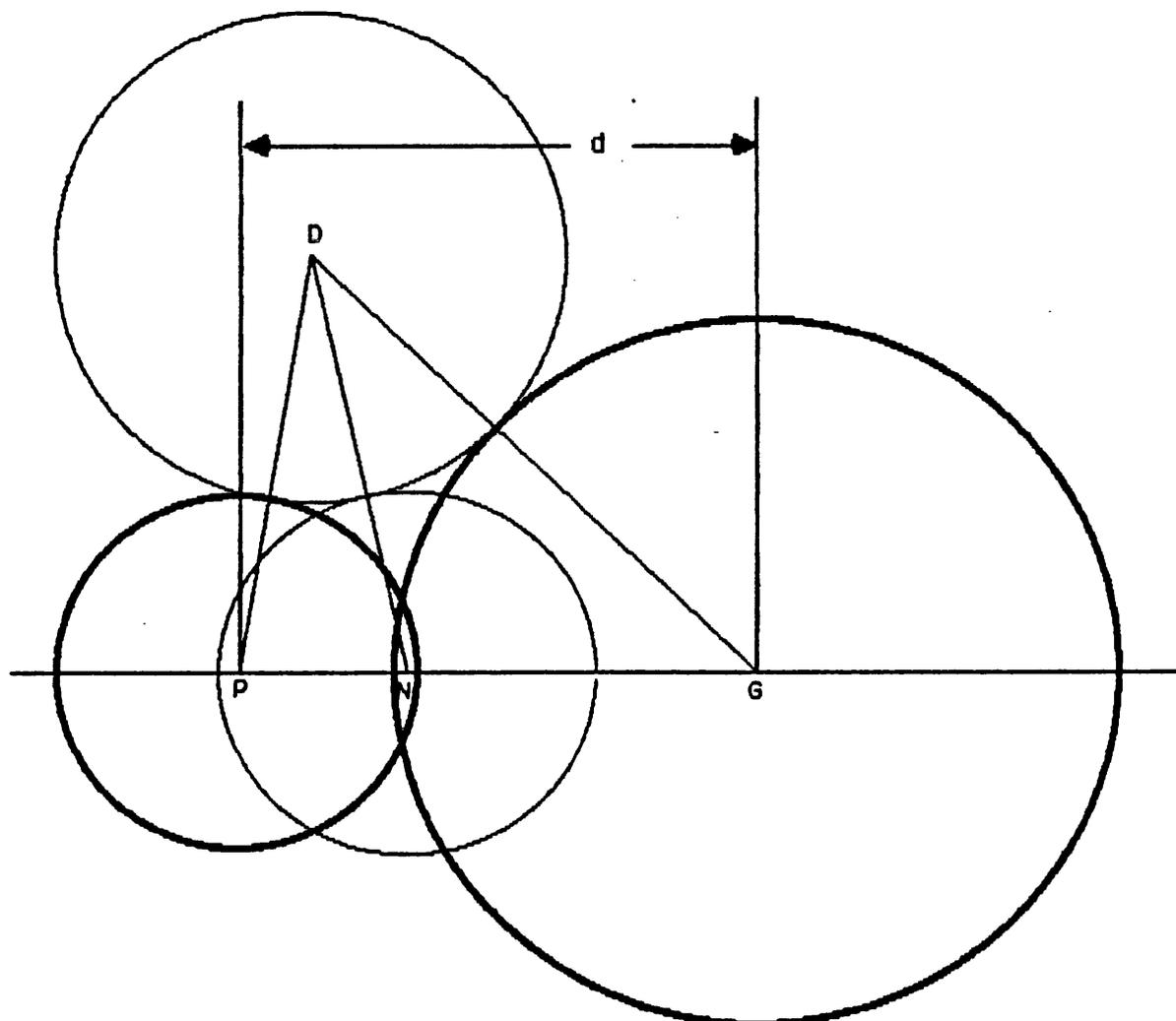
Una vez realizadas todas estas fases para un par determinado se vuelve a la fase 1 para estudiar otro par.

Como hemos visto este algoritmo depende de tres parámetros, OMEGA, RET y FRO que se indicarían desde el INPUT. Resulta evidente que la utilización conjunta de estos parámetros en el cálculo de una superficie puede resultar altamente tedioso y por ello realizamos la siguiente simplificación, que será debidamente justificada en el apartado 2.5:

-Primero el valor de OMEGA se da desde dentro del programa asignándole un valor de  $50.0^\circ$ .

-Segundo, los parámetros FRO y RET se reducen a un solo parámetro llamado FRADIO que puede tomar valores entre 0.0 y 1.0. De tal manera que asigna el valor de FRADIO a FRO. A continuación calcula el valor promedio de todos los radios de las esferas que se han dado en el INPUT y multiplica este valor por FRADIO, asignándole el valor obtenido a RET.

De esta manera desde el INPUT solo hay que introducir un solo valor, hecho que simplifica el uso del programa, aunque internamente siga utilizando los parámetros iniciales.



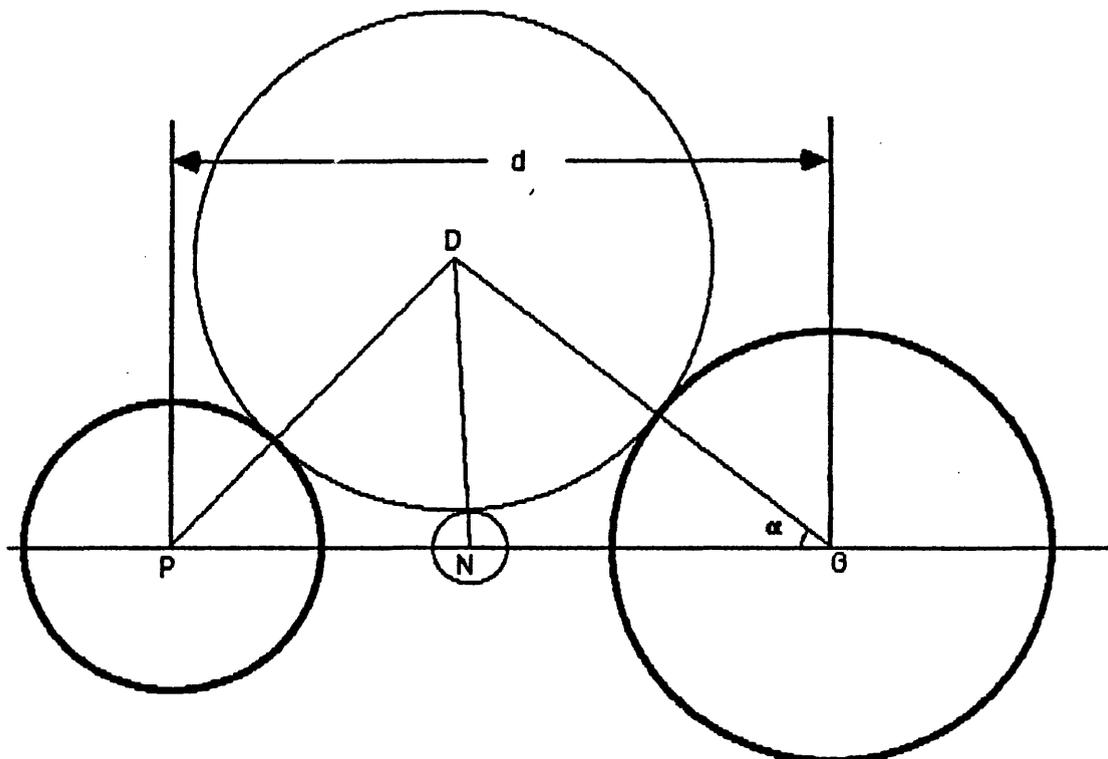
$$\overline{PN} = R_P + \frac{d - (R_P + R_G)}{2} = \frac{d + R_P - R_G}{2}$$

$$\overline{GN} = R_G + \frac{d - (R_P + R_G)}{2} = \frac{d + R_G - R_P}{2}$$

$$\lambda = \frac{\overline{GN}}{\overline{PN}} = \frac{d + R_G - R_P}{d + R_P - R_G}$$

$$x_N = \frac{x_G + \lambda x_P}{1 + \lambda} \quad y_N = \frac{y_G + \lambda y_P}{1 + \lambda} \quad z_N = \frac{z_G + \lambda z_P}{1 + \lambda}$$

**Figura 2.1.k.** Par de esferas del tipo A. Cálculo de las coordenadas de la nueva esfera.  
Las mismas fórmulas son válidas para los pares del tipo B.

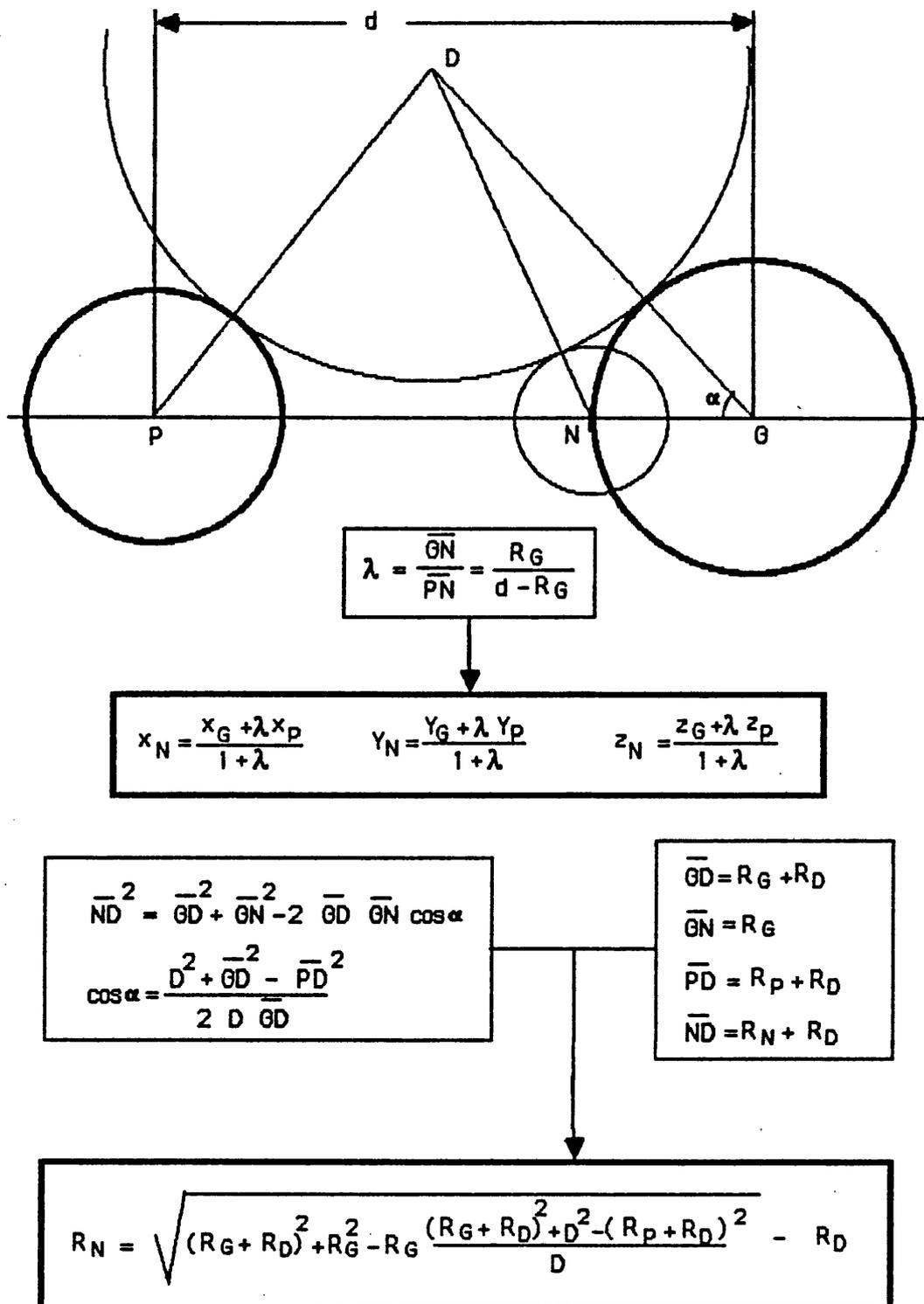


$$\begin{aligned} \overline{GD} &= R_G + R_D \\ \overline{GN} &= \frac{d + R_G - R_P}{2} \\ \overline{PD} &= R_P + R_D \\ \overline{ND} &= R_N + R_D \end{aligned}$$

$$\begin{aligned} \overline{ND}^2 &= \overline{GD}^2 + \overline{GN}^2 - 2 \overline{GD} \overline{GN} \cos \alpha \\ \cos \alpha &= \frac{D^2 + \overline{GD}^2 - \overline{PD}^2}{2 D \overline{GD}} \end{aligned}$$

$$R_N = \sqrt{(R_G + R_D)^2 + \frac{d + R_G - R_P}{2} \left( \frac{d + R_G - R_P}{2} - \frac{(R_G + R_D)^2 + D^2 - (R_P + R_D)^2}{D} \right)} - R_D$$

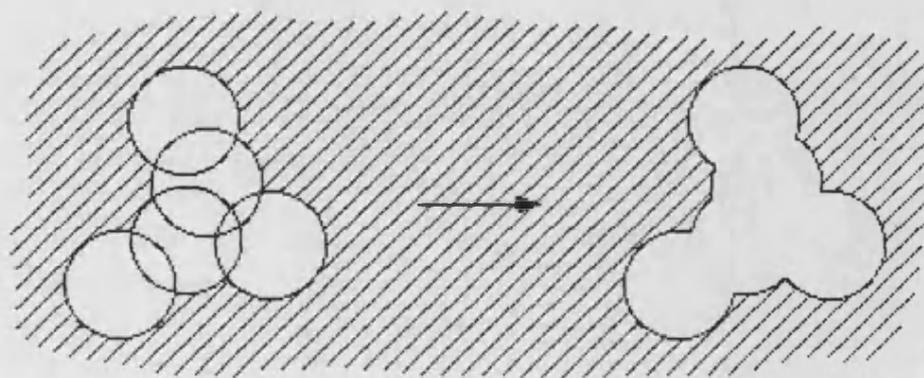
**Figura 2.1.1.** Par de esferas de tipo B. Cálculo del radio de la nueva esfera. Las mismas fórmulas son válidas para las esferas de tipo A.



**Figura 2.1.m.** Cálculo de las coordenadas y radio de la nueva esfera para un par del tipo C.

## 2.2 DETERMINACION DE LA SUPERFICIE.

La superficie, objeto de nuestro estudio, se forma a partir de un conjunto de esferas interconectadas de las cuales nosotros conocemos los radios y las coordenadas de sus centros. Así en la parte izquierda de la **figura 2.2 a** se muestra un conjunto de esferas centradas sobre los átomos de una hipotética molécula, mientras que a la derecha de esta figura vemos la superficie envolvente correspondiente. Puede verse que tenemos definidas las unidades (esferas) que forman la cavidad, pero desconocemos la expresión que nos describe a ésta, es decir, la superficie compuesta por trozos de las esferas iniciales. En este apartado explicaremos el método que hemos desarrollado para obtener la definición de la superficie y que se basa en una idea originalmente desarrollada por **Miertus, S. y col, 1981**. Estos autores dividían la superficie de cada esfera en pequeñas teselas, limitadas por meridianos y paralelos. El número de teselas venía determinado por el número de meridianos y paralelos, los cuales eran controlados desde el INPUT indicando la diferencia angular, en coordenadas polares, entre dos consecutivos de los primeros y dos de los segundos. Una vez divididas las esferas, se eliminaban aquellas teselas cuyos centros se encontraran en el interior de cualquier otra esfera. Este método presentaba algunos problemas derivados de la forma no uniforme de sus teselas y del gran número de éstas que hay que crear para obtener una superficie aceptable.



**Figura 2.2.a.**

En el método GEPOL se divide cada esfera en 60 triángulos esféricos iguales, producidos

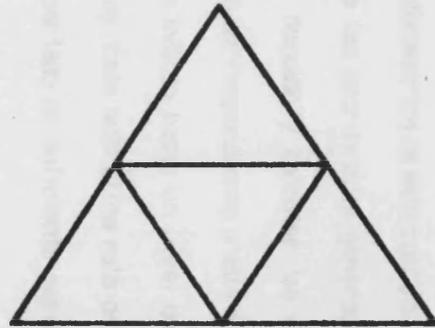
por la proyección del poliedro llamado pentaquisdodecaedro sobre cada esfera. Este poliedro de fórmula geodésica  $\{3,5_+,1,1\}$  (Coexter, H.S.M., 1974), deriva de un dodecaedro al que se le han añadido unas pirámides pentagonales de una altura tal que todos los vértices del cuerpo inscriben a una esfera.

Una vez obtenidos estos 60 triángulos iniciales, que de ahora en adelante denominaremos teselas, se plantea el problema de como eliminar aquellas teselas o trozos de ellas que quedan dentro de lo que será la superficie envolvente. Para resolver esta cuestión hemos ensayado diferentes métodos y algoritmos que no explicaremos en esta memoria, por no hacerla demasiado extensa. Sin embargo estos ensayos, que han requerido muchas horas de cálculo, son los que han permitido obtener el método que a continuación explicaremos.

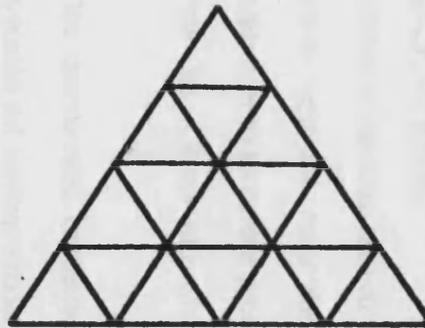
En líneas generales el método actúa de la siguiente manera: Primero divide las teselas en triángulos más pequeños, tal como se ve en la **figura 2.2.b**. Es decir, cada tesela triangular original se divide en cuatro triángulos esféricos. Después se divide cada uno de estos nuevos triángulos en cuatro más, con lo que nos encontraríamos en lo que llamamos nivel 3 de división, repitiéndose este proceso todas las veces que se quiera. El número de divisiones de los triángulos viene dado por el parámetro NDIY que se indica en el INPUT. Una vez que tiene los triángulos tan pequeños como se le han indicado, calcula el centro sobre la superficie de cada uno de ellos determinando si cada centro está fuera o dentro de la superficie envolvente. Los triángulos cuyos centros están dentro son eliminados, el resto son los que formaran esta superficie.

Para evitar un exceso de puntos sobre la superficie molecular que dificulte el uso posterior de ésta, se realiza un conjunto final de operaciones consistentes en reagrupar los triángulos pequeños en sus teselas originales, de tal manera que al final del proceso se obtienen como mucho 60 teselas, de las cuales, las situadas en las zonas de intersección con otra esfera habrán perdido su forma triangular. En el proceso de reagrupación se determinan las coordenadas de los centros de las teselas y a cada uno de los centros se le asocia un área equivalente a la suma de las áreas de los triángulos pequeños que componen esa tesela.

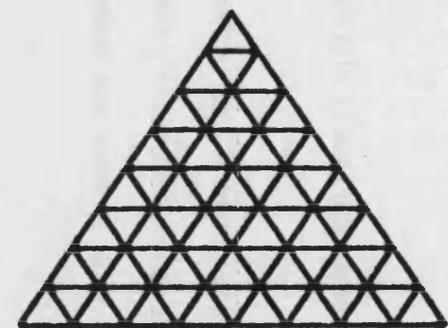
De esta manera tenemos definida la superficie envolvente con las coordenadas y áreas



NDIV=2



NDIV=3



NDIV=4

Figura 2.2.b. División del triángulo inicial para NDIV igual a 2, 3 y 4.

asociadas de un conjunto de puntos sobre ella. El área total de la superficie será simplemente la suma de las áreas de todas las teselas.

En la **figura 2.2. c** se muestra, para el caso de una superficie formada por dos esferas, lo que correspondería a un corte perfecto y el obtenido cuando se utiliza NDIV igual a 1 y 2. Se puede observar con estas figuras que a medida que aumenta NDIV el corte es mejor y que resultaría perfecto cuando NDIV fuese infinito.

Todo este proceso que hemos comentado en líneas generales está distribuido en tres subrutinas que a continuación pasaremos a describir en detalle.

### **2.2.1 Subrutinas TES y DIVIDE.**

El objetivo de estas dos subrutinas es calcular las coordenadas cartesianas de los centros de los triángulos pequeños para una esfera de radio uno, referidas al centro de la esfera (coordenadas locales relativas, CLR), de tal manera que cuando se quiera conocer las coordenadas de los puntos de cualquier esfera, bastará con multiplicar las CLR por el radio de la esfera deseada y sumarles las coordenadas del centro de la esfera.

La subrutina TES calcula las coordenadas cartesianas de los 32 vértices de los 60 triángulos esféricos que forman la proyección del pentaquisdodecaedro sobre una esfera de radio uno, tomando como origen de coordenadas el centro de la esfera. El cálculo se realiza a partir de los ángulos  $\theta$  y  $\phi$  correspondiente a las coordenadas polares de 6 vértices, dados en un DATA, utilizando las fórmulas clásicas de transformación de coordenadas polares a cartesianas y obteniendo el resto de los puntos aprovechando la simetría dodecaédrica de este cuerpo.

La subrutina DIVIDE utiliza las coordenadas generadas por la subrutina anterior para dividir los triángulos hasta el nivel deseado y calcular las CLR de los centros de éstos. En el **esquema 2.2.1 a** se muestra el flujo correspondiente a esta subrutina. En él se puede apreciar el conjunto de DO anidados, en los que se avanza hasta un nivel determinado dependiendo del grado de división de los triángulos que queramos. Esta subrutina está preparada hasta un nivel de NDIV=5, que como veremos en un posterior apartado es suficiente para alcanzar unos resultados más que

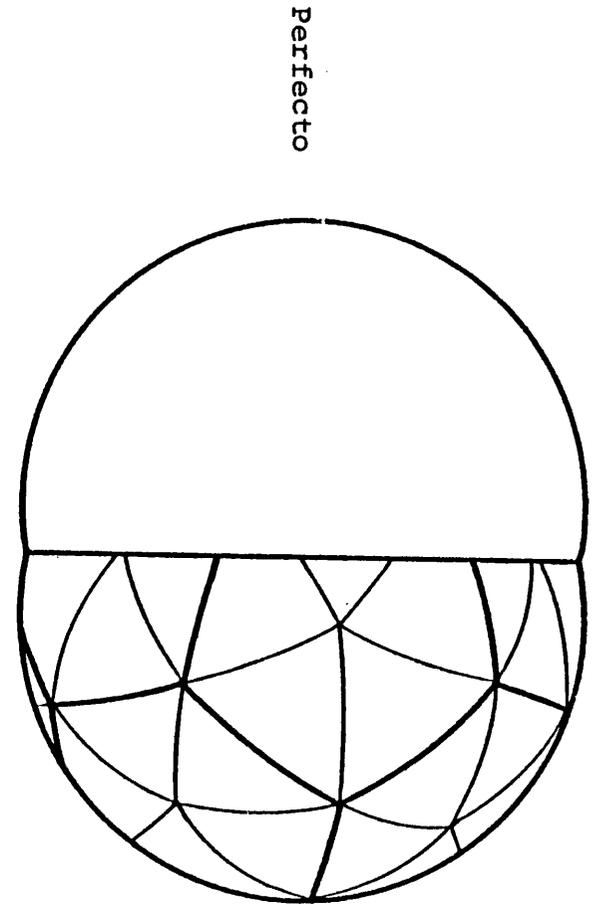
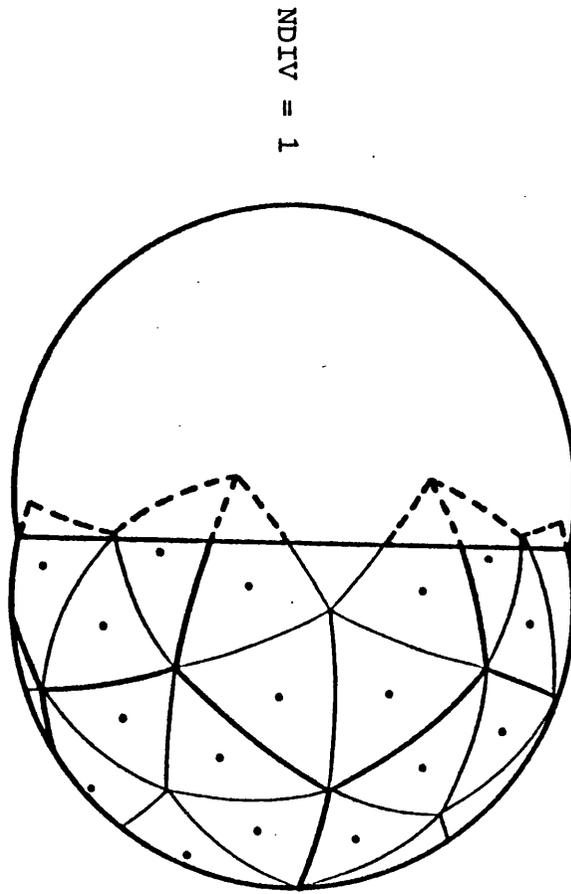
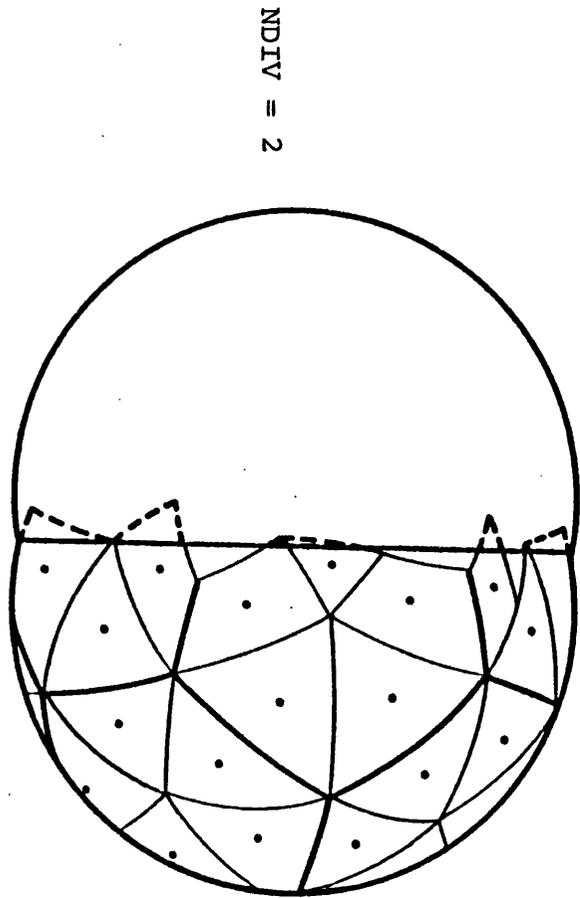
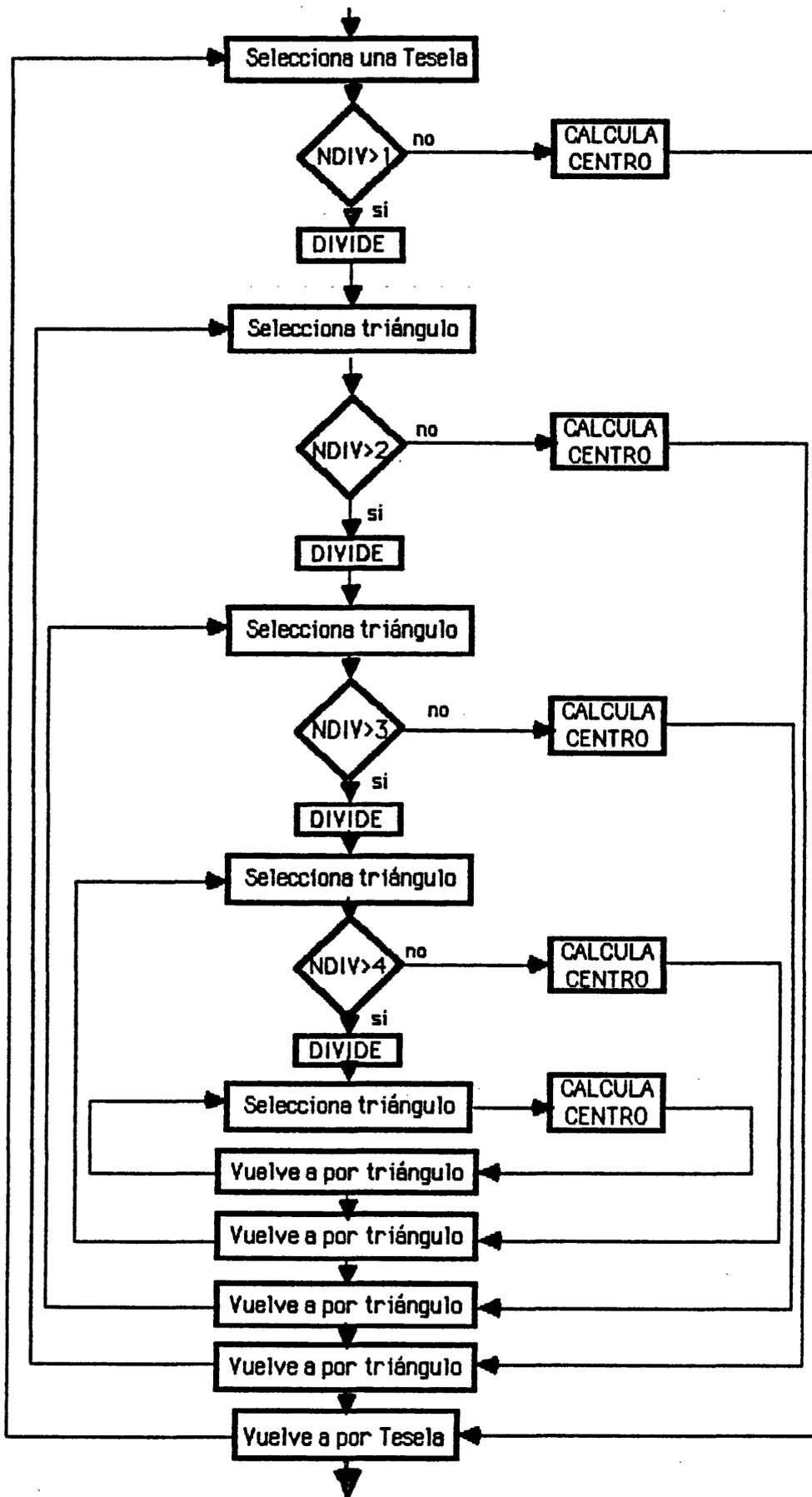


Figura 2.2.c Corte de los triángulos en el caso de dos esferas.



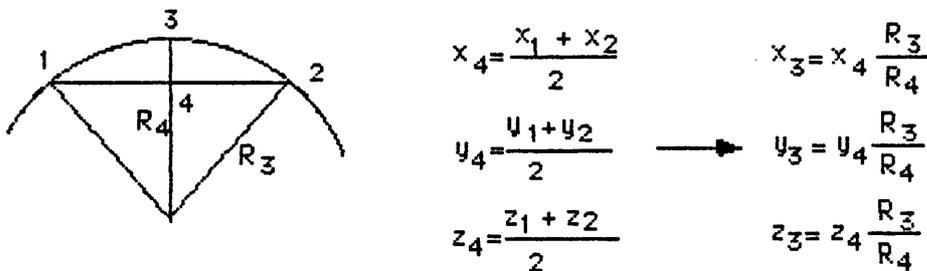
Esquema 2.2.1. a Diagrama de flujo para la subrutina DIVIDE.

optimos. A continuación explicamos como se realiza la división de los triángulos y el cálculo de los centros.

### DIVISION DE UN TRIANGULO

Un triángulo esférico queda identificado por las coordenadas de sus vértices. Para los triángulos originales el programa hace uso de un DATA, donde se indica qué vértices, de los 32 calculados en TES, pertenecen a cada uno de los 60 triángulos. Para los niveles inferiores de división existe otro DATA que permite en cualquier momento relacionar vértices y triángulos.

Para realizar la división de un triángulo en cuatro más, solo hay que calcular las coordenadas de los centros de los arcos que unen los vértices del triángulo a dividir. Así en la **figura 2.2.1 a**, se representa el arco formado por uno de los lados de un triángulo esférico, donde los puntos 1 y 2 corresponden a los vértices del triángulo, 4 es el punto medio de la recta que une 1 y 2, y 3 es el punto que se busca.  $R_3$  vale 1, ya que es el radio de la esfera.



**Figura 2.2.1.a**

### CALCULO DE LAS COORDENADAS DEL CENTRO DE UN TRIANGULO ESFERICO.

El programa necesita para los cálculos posteriores identificar cada triángulo a un punto, para ello realizamos el cálculo del centro del triángulo de la siguiente manera. Primero calcula el baricentro del triángulo plano formado por los vértices del triángulo esférico, para posteriormente proyectarlo sobre la superficie de la esfera. Las fórmulas utilizadas son:

$$\begin{array}{l}
 x_b = \frac{x_1 + x_2 + x_3}{2} \\
 y_b = \frac{y_1 + y_2 + y_3}{2} \\
 z_b = \frac{z_1 + z_2 + z_3}{2} \\
 R_b = \sqrt{x_b^2 + y_b^2 + z_b^2}
 \end{array}
 \longrightarrow
 \begin{array}{l}
 x_c = x_b \frac{R}{R_b} \\
 y_c = y_b \frac{R}{R_b} \\
 z_c = z_b \frac{R}{R_b}
 \end{array}$$

Donde  $x_b$  se refiere al baricentro de triángulo plano,  $x_c$  al centro del triángulo esférico y  $R$  el radio de la esfera que en nuestro caso es 1.

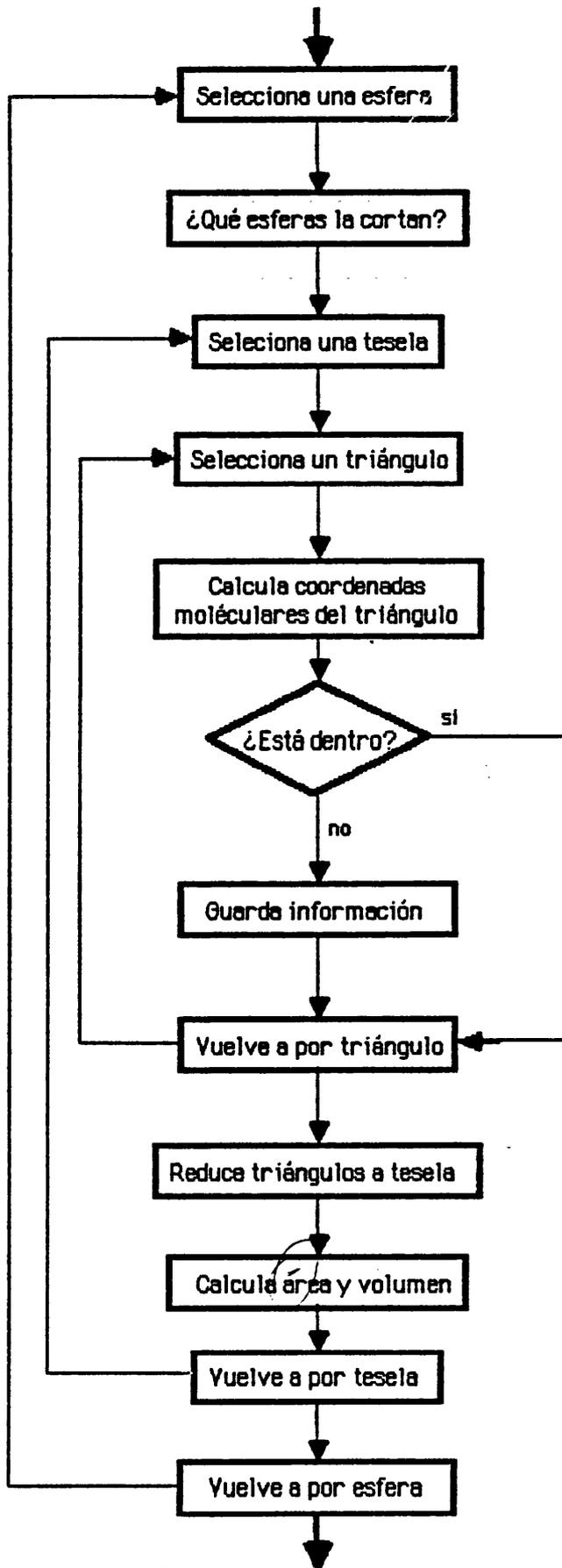
### 2.2.2 SUBROUTINA GEOCAV.

En esta subrutina se realiza el resto de procesos para la definición de la superficie. En el **esquema 2.2.2. a** se muestra el flujo general de ésta. La anidación esta hecha de tal manera que selecciona en el orden: Esfera, tesela y triángulo.

Una vez seleccionada una esfera  $E$  se determina qué otras esferas la cortan. Este cálculo que ahorrará tiempo a la hora de determinar si un triángulo se elimina o no. Para ello el programa calcula la distancia entre el centro de la esfera  $E$  y los centros de las otras esferas. Cuando la distancia es menor que la suma de los radios de las dos esferas, éstas se cortan.

En el paso siguiente primero selecciona una de las 60 teselas que componen la esfera, para pasar a seleccionar uno de los triángulos que componen ésta. Cuando selecciona un triángulo, lo que realmente utiliza el programa son las coordenadas del centro del triángulo calculadas en la subrutina anterior. Para poder continuar, se calcula las coordenadas moleculares del punto, multiplicando las CLR por el radio de la esfera a que pertenece y sumándole las coordenadas moleculares del centro de la esfera.

A continuación se determina si este punto está fuera o dentro de la superficie. Si está



Esquema 2.2.2. • Diagrama de flujo para la subrutina GECCAY

dentro se elimina del cálculo. Un punto está dentro cuando su distancia al centro de cualquier esfera que compone dicha superficie es menor que el radio de la esfera. Un punto forma parte de la superficie envolvente cuando en ningún caso su distancia al centro de cualquier esfera sea menor que el radio de ésta. Este proceso se repite para todos los triángulos que forman una tesela.

Cuando ya se conocen los triángulos que permanecen para una tesela, se procede a reagruparlos, es decir a determinar las coordenadas de un punto que represente a todos. Para ello se utilizan las coordenadas locales de los centros de los triángulos (coordenadas referidas al centro de la esfera a la que pertenecen), entonces se calcula el centro de masas de ese conjunto de puntos y se proyecta ese punto sobre la superficie de la esfera, de tal forma que sumándole las coordenadas del centro de la esfera tendremos las coordenadas cartesianas moleculares de un punto, que representará la tesela original o el trozo que quede de ella. Para que la representatividad de este punto sea total hay que asociarle el valor del área de la tesela. Como se conoce el número de triángulos en que está dividida la superficie y el radio de la esfera, dividiendo el área total de la esfera por el número de triángulos tendremos el área de un triángulo y por tanto para conocer el área de una tesela simplemente se multiplica el área de un triángulo por el número de triángulos que la componen después de la eliminación de los que quedan dentro de la superficie. De tal manera que el área total de la cavidad será la suma del área de todas las teselas.

Por último, el volumen se calcula con la fórmula indicada a continuación, donde el sumatorio se extiende a todas las teselas que forma la superficie,  $S$  es el área de la tesela,  $D$  es la distancia entre el centro de la tesela y el origen de coordenadas moleculares, y  $\alpha$  es el ángulo más pequeño que forman el vector normal a la superficie en el punto que representa a la tesela, y cuyo sentido es hacia el interior de la superficie, con el vector de posición del punto.

$$V = (1/3) \sum S D \cos \alpha$$

## **2.3 ENTRADA DE DATOS Y SALIDA DE RESULTADOS.**

En este apartado mostramos cómo se introducen los datos necesarios para el funcionamiento del programa GEPOL, e indicamos la información recogida en los ficheros que almacenan los resultados.

### **2.3.1 Entrada de datos para GEPOL (INPUT).**

La entrada de datos está formada por cinco grupos, cada uno de los cuatro primeros ocupa una tarjeta o línea, mientras que el quinto incluye una tarjeta o línea por cada átomo del soluto. El programa está construido de tal manera que es transparente al tipo de unidades utilizadas. Con ello queremos decir que se puede utilizar cualquier tipo de unidades, obteniéndose, como es lógico, el resultado en estas mismas unidades. A continuación veremos cada uno de los grupos.

#### **1ª TARJETA**

Esta tarjeta está reservada para el título (con un máximo de 80 caracteres) y sirve para indentificar los datos que vienen a continuación.

#### **2ª TARJETA**

Esta tarjeta está dividida en tres campos, correspondientes a tres variables, separados por un blanco (3(A3,1X)).

El primer campo corresponde a la variable SUAVE, que indicará si se quiere una suavización de la superficie o por el contrario sólo se utilizan las esferas iniciales. Si se opta por lo primero se indicará con la palabra CON y si no se quiere la suavización se pondrá SIN.

El segundo campo corresponde a la variable VECTOR que puede tomar dos valores: VES y VEN. En el primer caso el programa imprime las coordenadas y áreas asociadas a todos los puntos que genera sobre la superficie, dando también las componentes de los vectores de módulo 1 que, partiendo de los puntos anteriores, son normales a la superficie y se dirigen hacia fuera de ésta. En

el caso de que no se quiera utilizar esta información se pondrá VEN. En el interior del programa hay una variable , como un DATA, llamada DVEC que es la que controla el módulo y sentido de estos vectores. En el caso de que DVEC sea positiva los vectores se dirigen de la superficie hacia el interior de ésta.

Como vimos en el capítulo 1, la Superficie Molecular Accesible ( Lee, B. y F.M. Richards, 1971 ) corresponde a sumar el radio del disolvente a las esferas de van der Waals que forman la molécula. Si a la variable SUMA se le da el valor SUM se realiza esta suma, en caso contrario se pondrá NSU. En el caso de que SUMA=SUM, SUAVE tiene que ser igual a SIN.

### 3ª TARJETA

Esta está formada por tres variables: NDIV, RD y FRADIO (12,2F10.5). NDIV corresponde al nivel de división de los triángulos y puede variar entre 1 y 5. RD corresponde al radio del disolvente, dándose en las mismas unidades que las coordenadas de los átomos. FRADIO, como ya indicamos en el apartado anterior controla la subrutina de suavización, y puede variar entre 0.0 y 1.0.

### 4ª TARJETA

Indica el número de átomos que forman la molécula y se da en formato (18).

### 5ª TARJETA y siguientes

En cada tarjeta se indican las coordenadas cartesianas de un átomo, el radio de la esfera que va centrada sobre él y una variable entera, llamada USE, que indica si se utiliza este átomo para formar la cavidad o no, dependiendo que tome el valor 1 ó 0. Se utiliza una tarjeta por átomo con un formato (4F10.5,12).

### **2.3.2 Salida de resultados del GEPOL (OUTPUT).**

La salida de resultados se reparten en cuatro ficheros que a continuación comentamos. En todos los ficheros se imprime el título y todas las variables dadas en el INPUT.

#### **FOR006. Resumen de resultados.**

En este fichero se da un resumen de los resultados obtenidos:

- Radio medio de todas las esferas (RAP)
- RET (obtenido de multiplicar RAP por FRADIO)
- FRO (mismo valor que FRADIO)
- Número de esferas creadas.
- Número de esferas con superficie cero. Estas son las esferas que tienen todos sus triángulos dentro de otras esferas.
- Número de esferas interiores. Son aquellas esferas que están totalmente dentro de otra esfera.
- Área total.
- Volumen total.
- Número de puntos generados sobre la superficie.

#### **FOR007. Información esfera a esfera.**

En este fichero se da información de cada esfera utilizada en el cálculo, tanto las dadas en el input como las creadas en la subrutina de suavización. Los valores que aparecen corresponden a:

- Número de orden de la esfera.
- Área de la esfera después de eliminar los trozos interiores a la superficie molecular.
- Coordenada X del centro de la esfera.
- Coordenada Y del centro de la esfera.
- Coordenada Z del centro de la esfera.
- Radio de la esfera.
- Área inicial de la esfera  $4\pi R^2$ .

-Número de puntos resultantes sobre la esfera.

-En el caso de que sea una esfera creada en la subrutina de suavización, las dos últimas cifras indican el número de orden de las dos esferas que la generaron.

#### FOR008. Información de los puntos y vectores.

En este fichero se encuentra toda la información correspondiente a los puntos que se generan sobre la superficie y sus vectores en el siguiente orden:

- Número de orden de la esfera a que pertenece el punto.
- Coordenada X del punto.
- Coordenada Y del punto.
- Coordenada Z del punto.
- Area asociada al punto.
- Componente X del vector.
- Componente Y del vector.
- Componente Z del vector.

## 2.4 CALCULOS DE SUPERFICIES DE VAN DER WAALS

En este apartado se analiza la bondad del método GEPOL para calcular la superficie resultante de un conjunto de esferas interconectadas. Estudiaremos la superficie de van der Waals de varias estructuras formadas por un número de esferas comprendido entre 2 y 1438.

Algunas de estas estructuras corresponden a modelos geométricos sencillos, sin relación directa con una molécula concreta en los que se pueden conocer su superficie y su volumen exactos y, por tanto, determinar el error del método. Los otros ejemplos corresponden a proteínas o fragmentos de éstas y a alguna estructura de interés químico como el electrón solvatado. Con ellos no se pretende por el momento obtener ninguna conclusión de carácter químico o estructural, sino solo ilustrar la aplicación del método con estructuras reales de diferente complejidad. Debido a esta complejidad es imposible conocer el área exacta calculada analíticamente, y con el fin de analizar la bondad del método, se compara éste con el MS (Conolly, M.L., 1983) del que hablamos en el capítulo I de esta memoria y que es el que está más extendido actualmente para el cálculo de superficies.

La bondad de un método de cálculo de superficie se puede valorar en base a seis aspectos principalmente:

### EXACTITUD

La exactitud sólo se puede determinar cuando se conoce el valor real y esto solo se da en nuestro caso en las estructuras sencillas. Sin embargo se puede extraer alguna idea de la exactitud por comparación con otro método aproximado, siempre que éste sea un método lo suficientemente usado y por tanto merezca un cierto grado de confianza, y que ambos utilicen algoritmos lo más diferentes posibles. Así, si ambos métodos dan resultados similares podemos decir que el valor real se encuentra cerca de los valores de uno o de otro. Pero si los resultados son dispares, habrá que utilizar otra manera de analizarlos para determinar cual de los dos da el valor más correcto. El método que utilizaremos para la comparación, como ya hemos indicado, será el MS.

### PRECISION

GEPOL utiliza como parámetro de cálculo el NDIV, ya definido anteriormente, mientras que el método MS utiliza el parámetro DEN que corresponde a una densidad de puntos por unidad de área. La utilización de un valor u otro para estos parámetros lleva a valores diferentes del área. En esta tesis denominamos precisión al comportamiento del resultado frente al parámetro utilizado. Así se puede dar que unos métodos presenten un comportamiento más convergente en la variación del resultado con el parámetro y otros un comportamiento más bien oscilante.

### INVARIANCIA

La invariancia espacial o dependencia del resultado de la orientación de una estructura frente a los ejes cartesianos, es un aspecto muy importante a la hora de valorar un método. Para poder determinar este aspecto, en varios de los casos que presentamos, hemos calculado la superficie para diferentes disposiciones de las estructuras en el espacio, determinando la desviación estándar para estos cálculos. Esta desviación no hay que tomarla en un sentido estrictamente estadístico, ya que el número de orientaciones será como mucho de tres, pero sin embargo nos puede dar una idea de la invariancia espacial.

### LOCALIDAD

Algunas propiedades moleculares dependen de las superficies locales en regiones específicas de las moléculas y no de la total. Por ello, no solo interesa obtener un valor bueno del área total, sino que el método describa bien las áreas locales.

### TIEMPO DE CALCULO

En muchas ocasiones el tiempo es un factor determinante a la hora de realizar un cálculo. A veces se quiere obtener un buen resultado a costa del tiempo de máquina que haga falta, pero la mayoría de las veces, debido a las limitaciones propias de cada laboratorio o las circunstancias del

estudio que se realiza, se necesiten resultados buenos, o al menos aceptables, en un tiempo no muy grande. Por ello, en muchas de las gráficas se representa la propiedad en estudio frente al tiempo de c.p.u.. Como los cálculos se han realizado con diferentes ordenadores, cada vez que se haga referencia al tiempo se le acompañará de un número romano que indicará con qué máquina se hizo. Al final de la memoria se encuentra la lista de ordenadores utilizados con su clave en números romanos.

### CANTIDAD DE INFORMACION

En los anteriores párrafos hemos hecho referencia a aspectos relacionados con la calidad de un método, pero otro aspecto importante es la cantidad de información que puede dar un programa. Es interesante que el programa utilizado dé la mayor cantidad de información posible con el mismo esfuerzo computacional, o por lo menos que esté estructurado de tal manera que permita obtenerla con simples modificaciones en el algoritmo.

A continuación pasaremos a mostrar los resultados para hacer, al final del apartado, una discusión global basada en los aspectos citados.

#### 2.4.1. Comparacion con SPHERE. Dos esferas

El método SPHERE se construyó originalmente para ser utilizado en el modelo de **Miertus, S. y col., 1981**, para el cálculo de la energía libre de disolución, por eso en el capítulo 3 se analizarán con detalle sus resultados en comparación con el GEPOL. En este capítulo nos limitamos a un pequeño análisis comparativo con un modelo de dos esferas, que pensamos que es suficiente, junto a lo expuesto en el apartado 2.2, para mostrar la superioridad de GEPOL.

En la **tabla 2.4.1.a** se dan los resultados para la cavidad formada por dos esferas de radio  $R_1=R_2=2.4$ , con sus centros situados a una distancia de 1.54 unidades, variando NDIV de 1 a 5 y para dos teselaciones diferentes de SPHERE. Los valores de superficie, volumen y número de puntos resultantes sobre la superficie, son la media de tres cálculos correspondientes a situar el eje de

**Tabla 2.4.1.a.** Area (S), volumen (V) y número de puntos (P) de una superficie formada por dos esferas.

NIVEL	S	V	P
1	95.±3.	84.±2.	79±2
2	95.±2.	85.±1.	86±6
3	94.8±0.6	84.4±0.4	94±5
4	95.5±0.2	85.0±0.1	97±4
5	95.57±0.007	84.98±0.04	97±4
$\Delta\theta=30, \Delta\phi=36$	101.±7	90.±5.	88±8
$\Delta\theta=\Delta\phi=25.71$	94.±5.	86.±2.	134±19
Exacto	95.6049	84.8169	

**Tabla 2.4.1. b.** Area, volumen y número de puntos para la superficie del FH.

NIVEL	S	S <sub>F</sub>	S <sub>H</sub>	V	P
1	38. ±2.	24. ±2.	14. ±1.	22. ±1.	76 ±5
2	38.4 ±0.9	23.8 ±0.6	14.5 ±0.4	21.8 ±0.4	87 ±3
3	38.79 ±0.07	24.0 ±0.1	14.7 ±0.2	22.06 ±0.03	93 ±3
4	38.69 ±0.06	24.18 ±0.07	14.52 ±0.01	22.03 ±0.03	93 ±3
5	38.71 ±0.02	24.22 ±0.02	14.49 ±0.00	22.04 ±0.01	94 ±2
$\Delta\theta=30, \Delta\phi=36$	39. ±1.	24.4 ±0.4	14. ±1.	21.9 ±0.4	79 ±8
$\Delta\theta=\Delta\phi=25.71$	40. ±3.	25. ±2.	15. ±1.	23. ±1.	131 ±6
Exacto	38.7012	24.2188	14.4824	22.0026	

unión de las dos esferas sobre los ejes X, Y y Z. También se da la desviación estándar para estos valores. En la última línea de la tabla aparecen los valores exactos, calculados con fórmulas de geometría elemental, lo que permite determinar la exactitud de cada una de las aproximaciones.

En la **tabla 2.4.1.b** se dan los resultados para la superficie de van der Waals del HF ( $R_H=1.44 \text{ \AA}$ ,  $R_F=1.62 \text{ \AA}$  y  $D_{FH}=0.92 \text{ \AA}$ ). La tabla incluye, además, las superficies de cada una de las esferas.

Como se puede observar en ambos casos, a medida que aumenta NDIV los valores de la superficie y del volumen convergen hacia los valores exactos, a la vez que la desviación se hace más pequeña (la dependencia con la orientación es menor). Sin embargo, con el método SPHERE hay un error mayor al aumentar las teselas del HF.

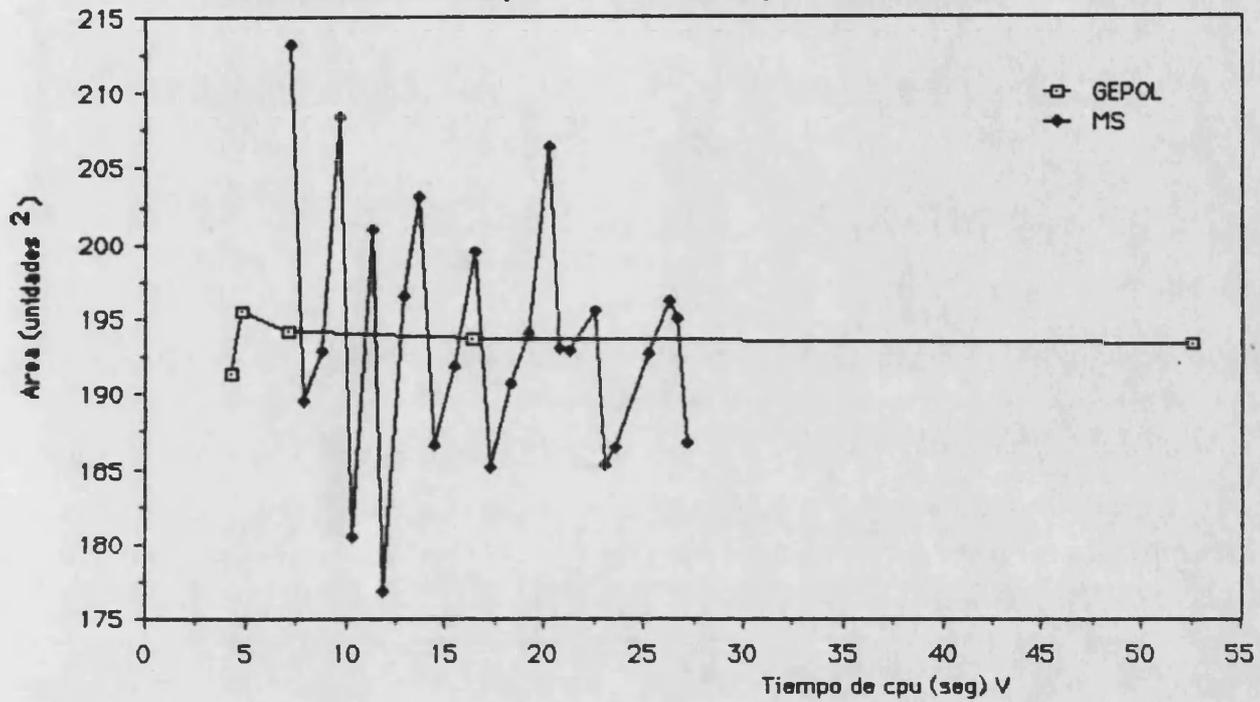
En líneas generales se puede ver en las tablas que el método GEPOL produce menores errores y desviaciones que el método SPHERE. Un aspecto interesante hubiera sido analizar los tiempos de cálculo de un método y otro, pero el programa SPHERE realiza los cálculos geométricos y electrostáticos al mismo tiempo y por tanto no se pueden desglosar uno del otro. En el capítulo 3 se incluye una tabla para el HF, donde aparecen los tiempos totales para el cálculo de la energía electrostática de solvatación.

### **2.4.2 Diez esferas**

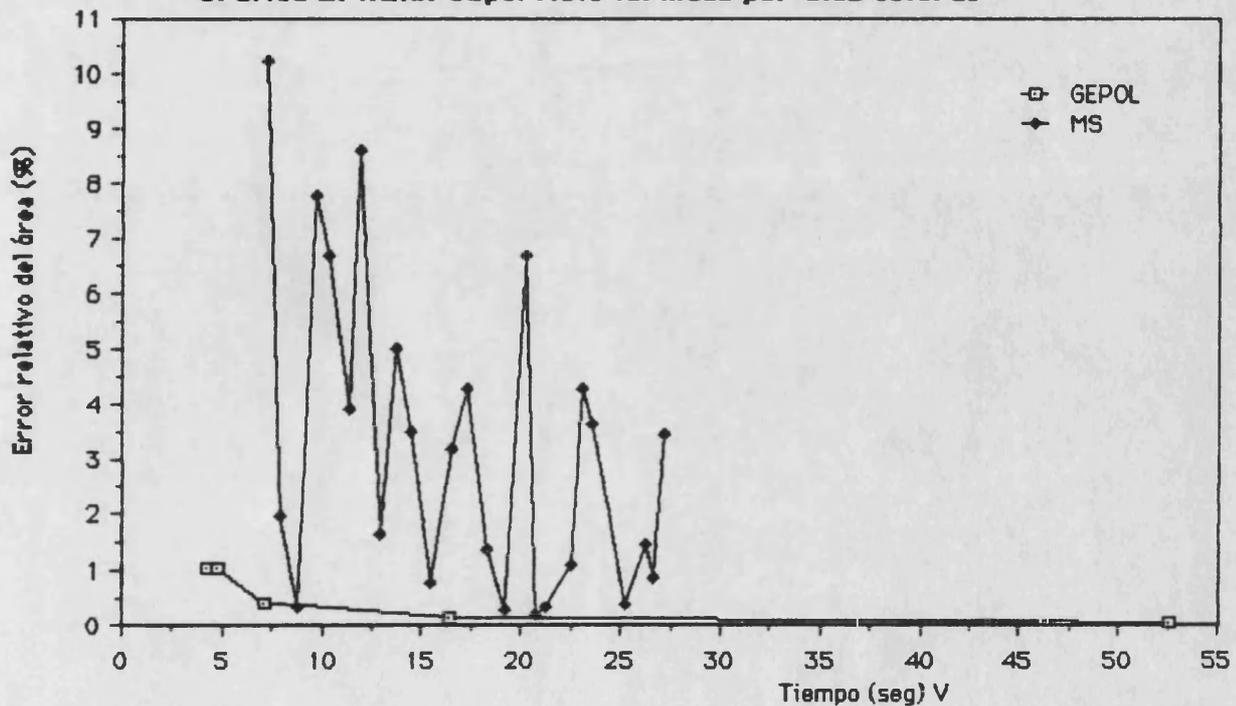
En este apartado presentamos los resultados para una cavidad formada por diez esferas de radio=1.8 y cuyos centros situados sobre un mismo eje equidistan uno del otro 1.5 unidades. Los cálculos se han realizado con los métodos GEPOL y MS, variando el parámetro NDIV (GEPOL) de 1 a 5 y el parámetro DEN (MS) de 1.0 a 25.0. DEN corresponde a una densidad de puntos por unidad de superficie, el valor DEN=25 es el máximo que puede tomar este parámetro en el programa MS para una esfera de radio 1.8 (en este programa no se permite que el número de puntos sobre una esfera sea mayor de 1000). Los cálculos se han realizado para tres orientaciones diferentes, situando los centros de las esferas sobre los ejes X, Y y Z.

En las **gráficas 2.4.2. a, b y c** se representan los valores de la media de la superficie

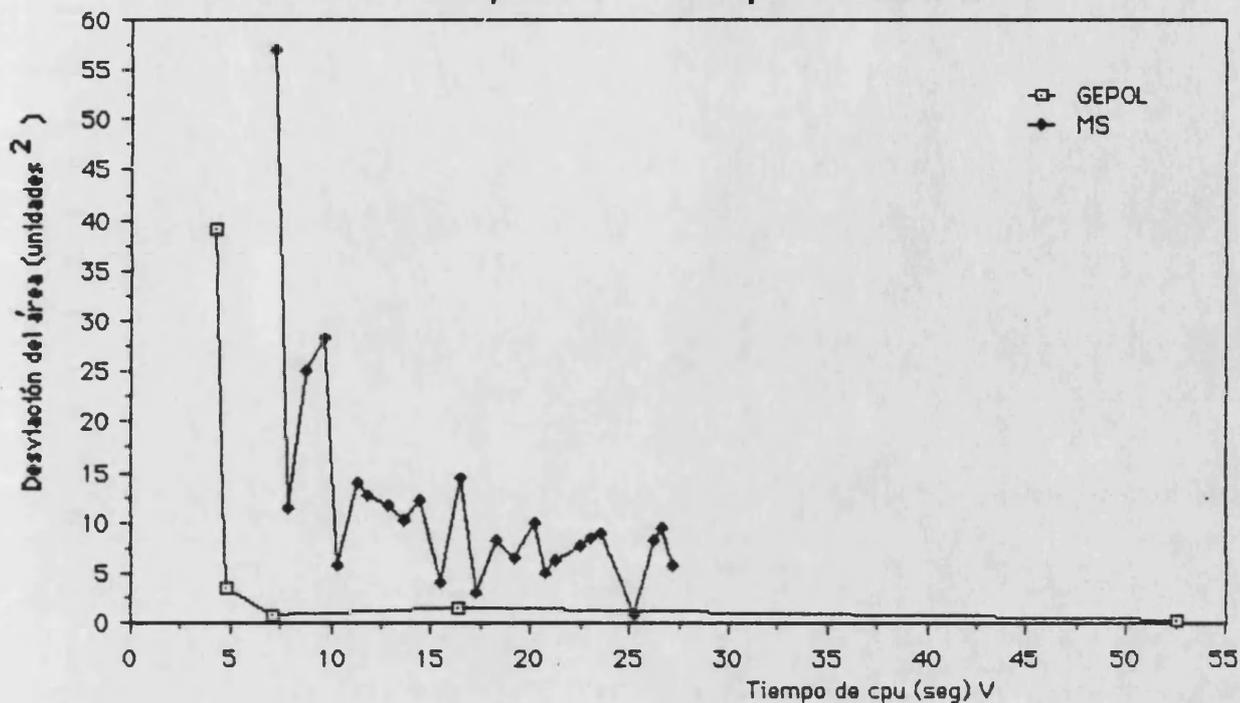
Gráfica 2.4.2.a. Superficie formada por diez esferas



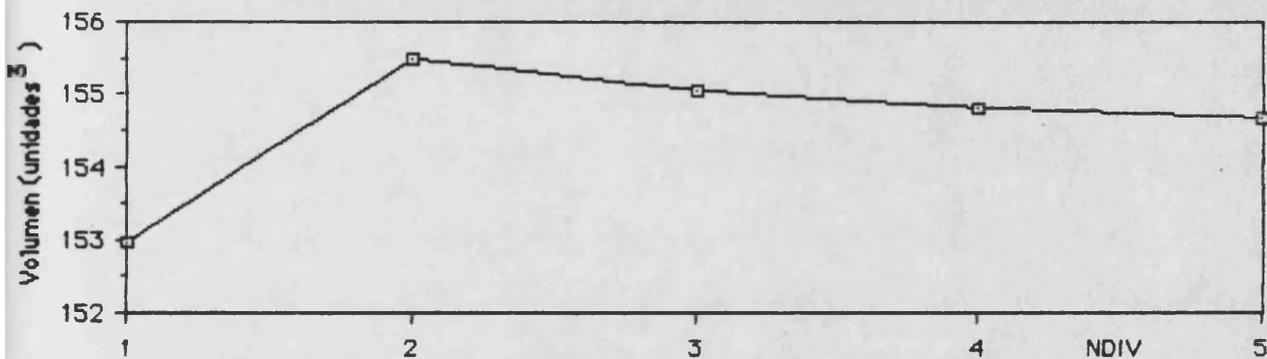
Gráfica 2.4.2.b. Superficie formada por diez esferas



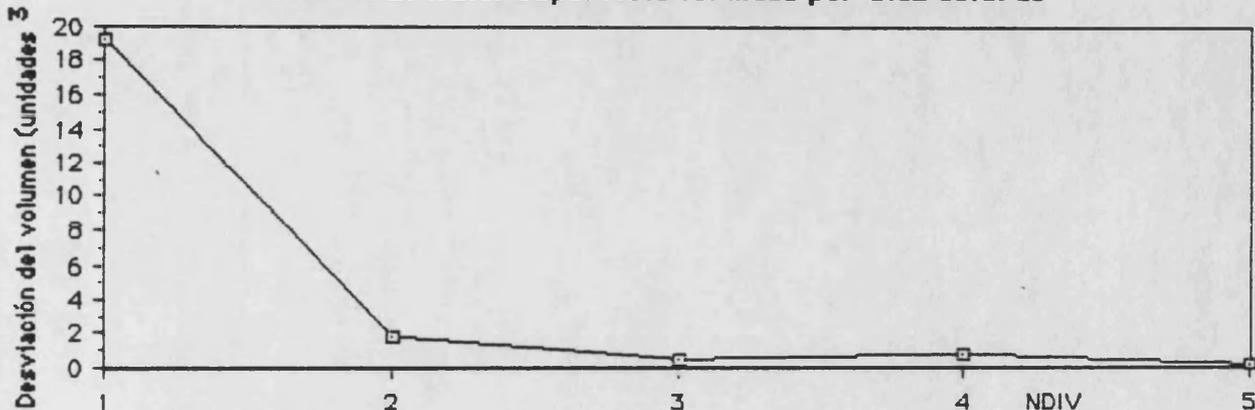
**Gráfica 2.4.2.c. Superficie formada por diez esferas**



**Gráfica 2.4.2.d. Superficie formada por diez esferas**



**Gráfica 2.4.2.e. Superficie formada por diez esferas**



total ( unidades<sup>2</sup>), error relativo en ‰ (superficie exacta=193.396 unidades<sup>2</sup>) y desviación (unidades<sup>2</sup>) frente al tiempo de C.P.U.(seg) respectivamente.

Los diferentes puntos corresponden a los resultados obtenidos con los diferentes valores de NDIY y DEN, los valores más pequeños de estos parámetros suponen tiempos menores y, a medida que aumentan los valores de NDIY y DEN, el tiempo de cálculo se hace mayor.

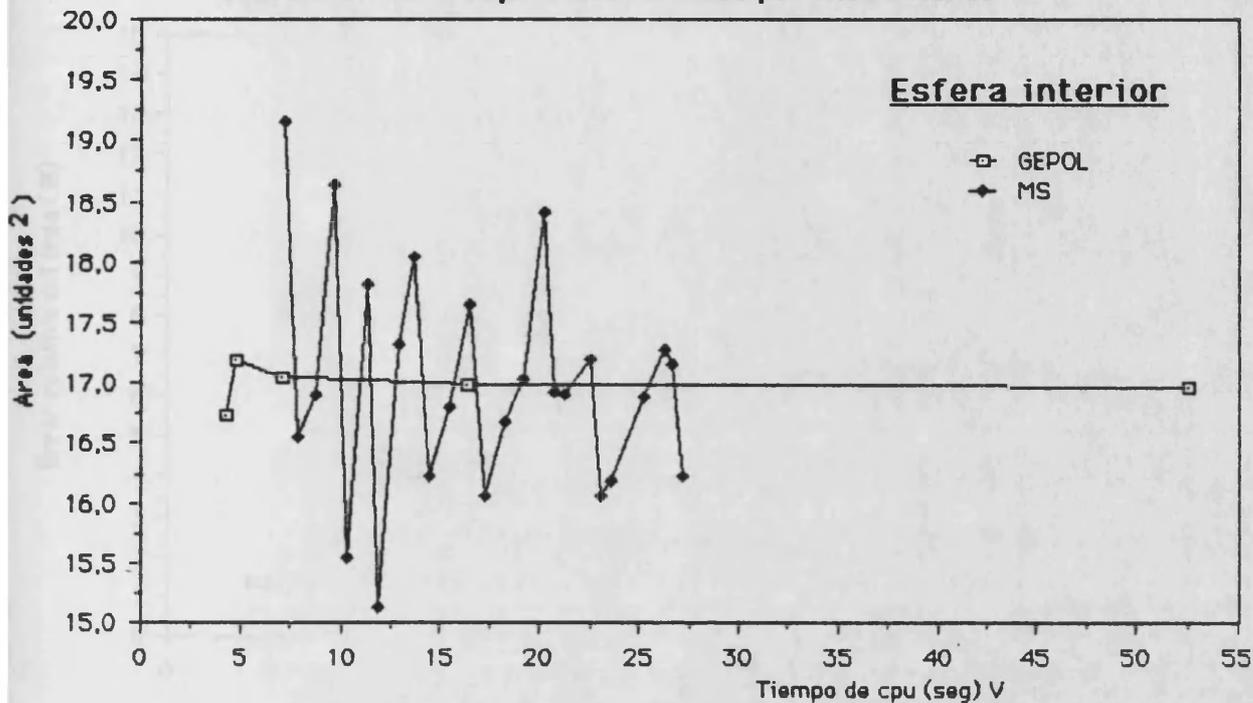
De estas gráficas se deduce un mejor comportamiento de nuestro método frente al MS. El GEPOL tiene errores que no superan el 1.05‰, mientras que la mayoría de los errores del MS se sitúan entre el 1 y el 10‰. Los resultados de nuestro método también son menos dependientes de la orientación espacial, como se puede ver en la representación de la desviación. Además los resultados de GEPOL son más convergentes, ya que el método MS tiene un comportamiento oscilante.

Dado que el método MS no calcula el volumen, en las **gráficas 2.4.2. d y e** representamos el volumen (unidades<sup>3</sup>), y la desviación (unidades<sup>3</sup>) obtenido con el GEPOL frente al parámetro NDIY. Hay que destacar que los errores no superan el 1,05‰, como en el caso de la superficie.

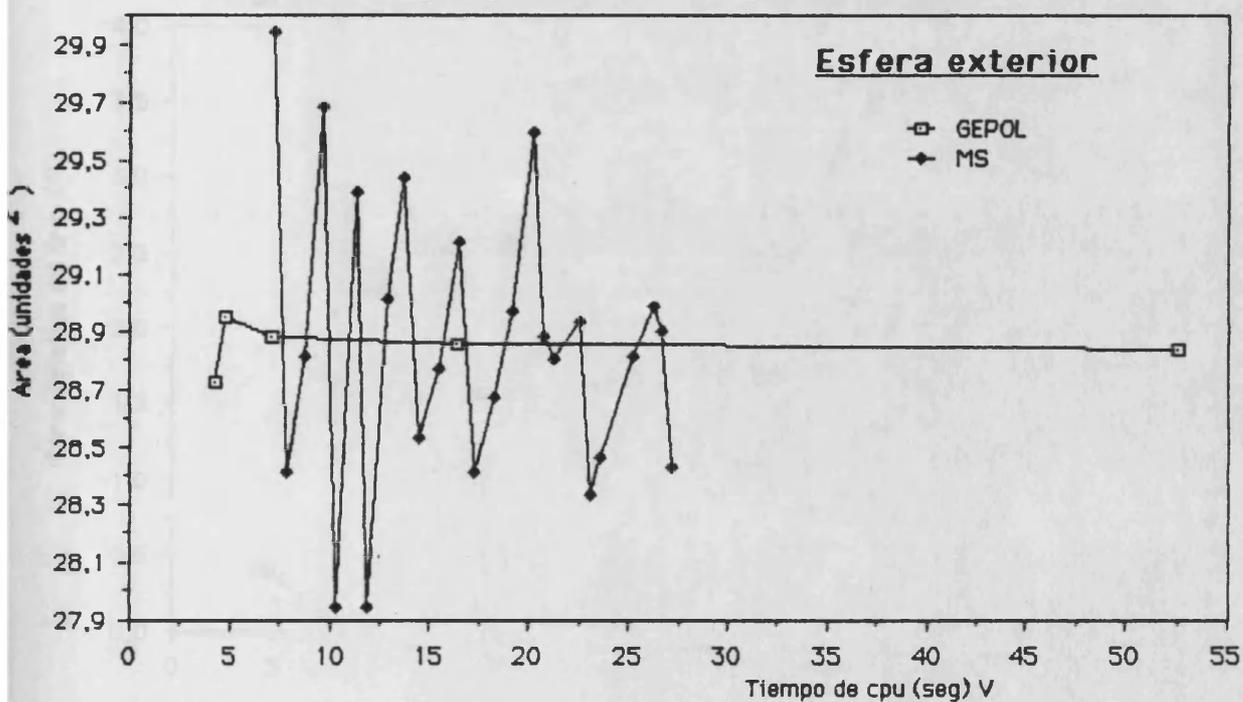
Para poder analizar la calidad de la superficies locales y a la vez determinar como participan en la superficie total, se pueden dividir las diez esferas en dos grupos, las ocho interiores y las dos externas. Así en las **gráficas 2.4.2 f, g,h,i,j y k** se representan de la misma forma que para la superficie total, la superficie, error y desviación de una esfera interna y de una externa. En ellas se puede ver que las esferas internas participan más en el error total. En ambos casos resulta evidente el mejor comportamiento de método GEPOL.

Un aspecto interesante es observar como se puede producir una compensación de los errores de las áreas locales de tal manera que el área total resulta más exacta que alguna local. Así en las **gráficas 2.4.2 l y m** hemos representado el error en el área total, esfera exterior y esfera interior, para los dos métodos en función de sus parámetros. En ellas se puede ver claramente que el error de la superficie total es menor que el de la esfera interna en ambos métodos. Esto se debe a que los errores absolutos de las áreas de las esferas internas y externas poseen signos contrarios.

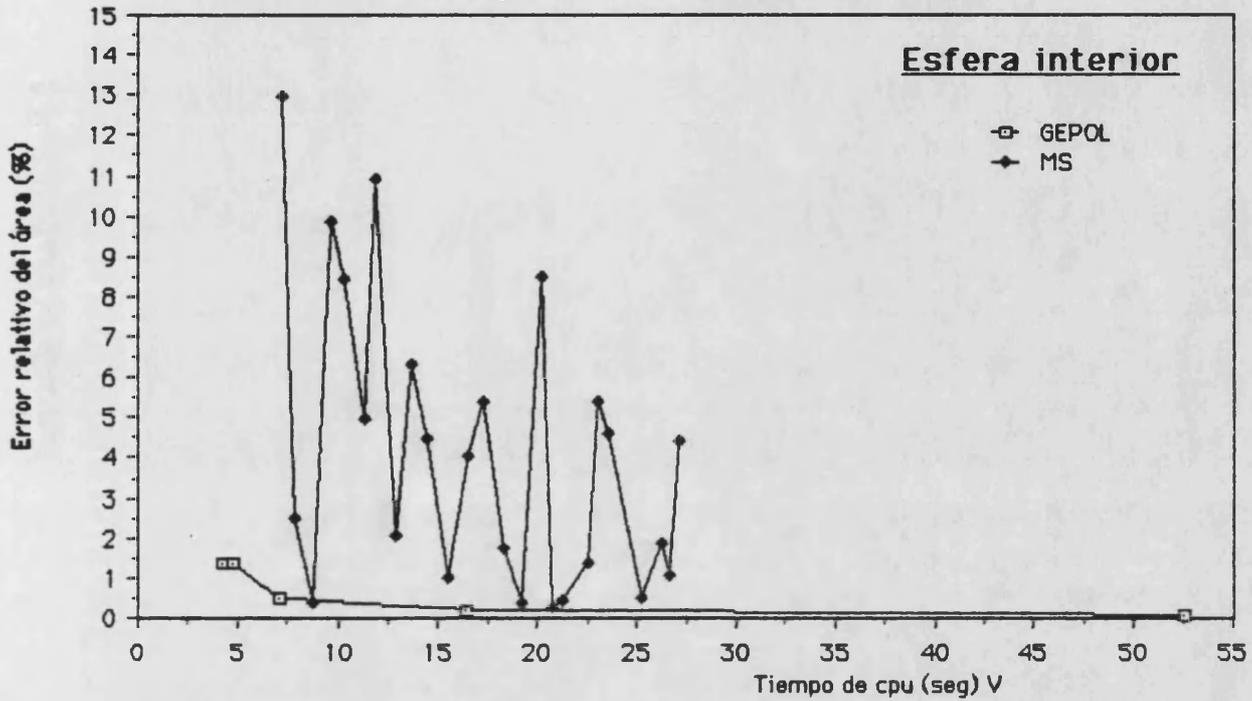
Gráfica 2.4.2.f. Superficie formada por diez esferas



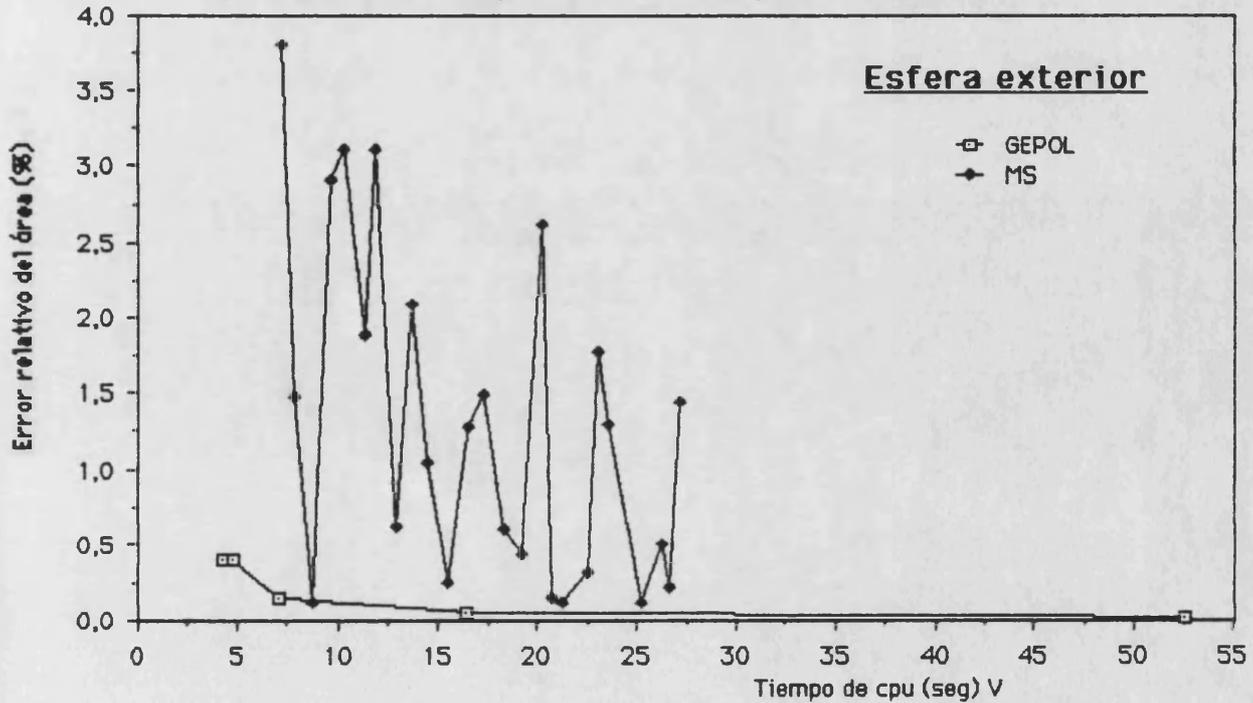
Gráfica 2.4.2.g. Superficie formada por diez esferas



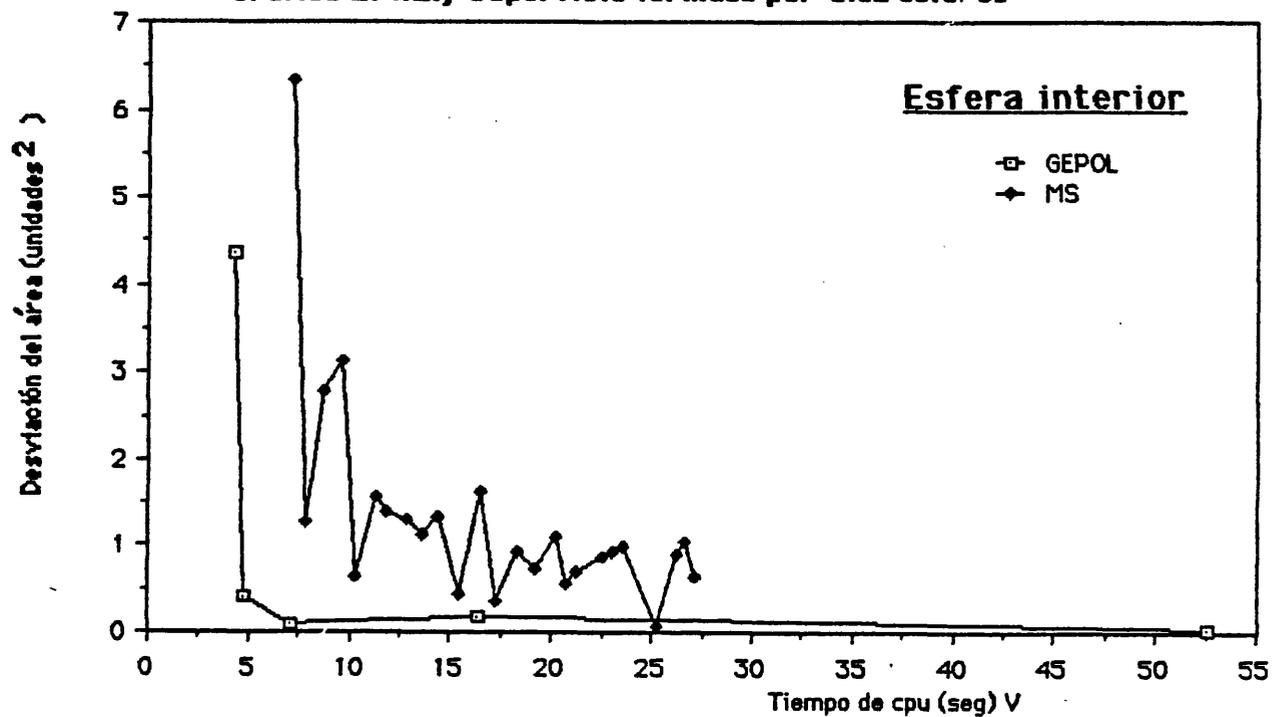
Gráfica 2.4.2.h. Superficie formada por diez esferas



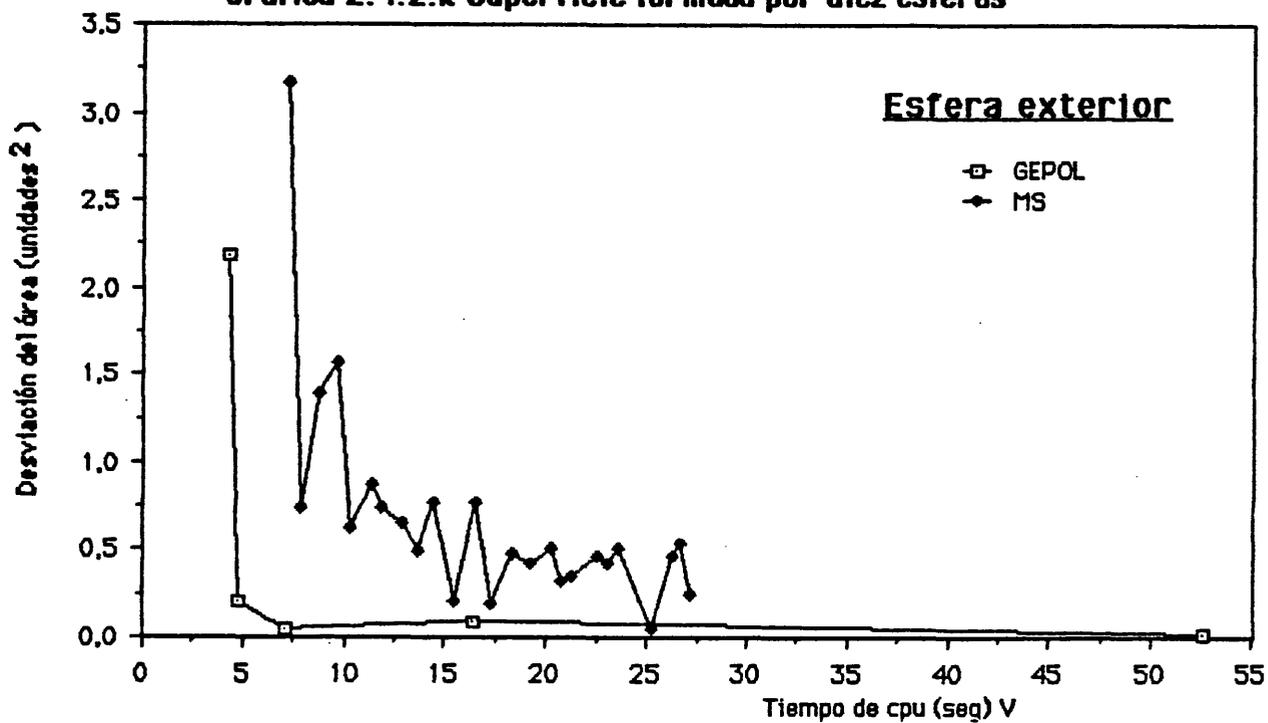
Gráfica 2.4.2.i. Superficie formada por diez esferas



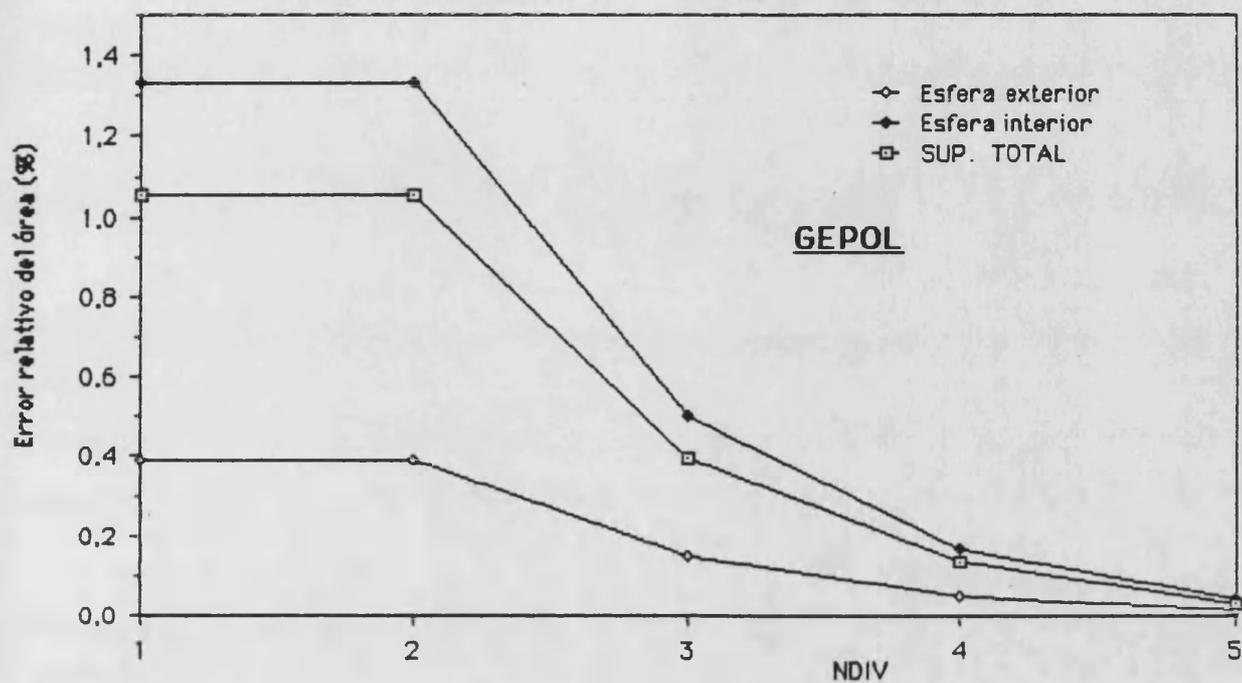
Gráfica 2.4.2.j Superficie formada por diez esferas



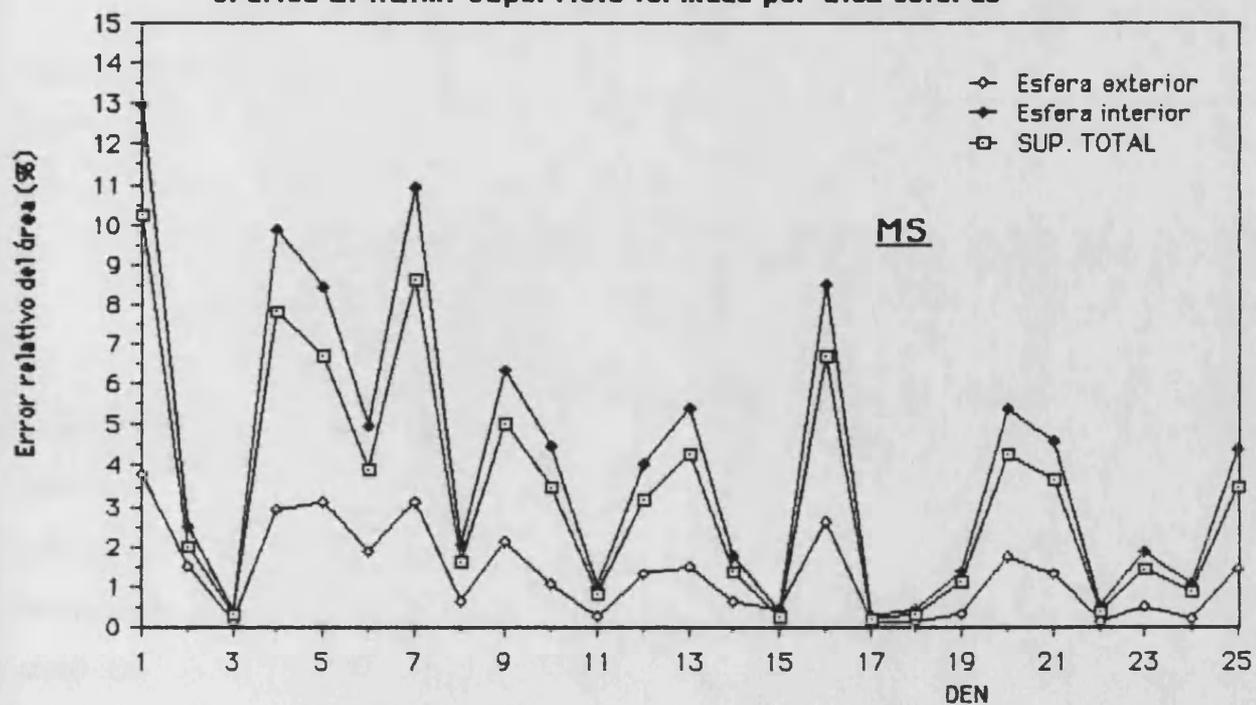
Gráfica 2.4.2.k Superficie formada por diez esferas



Gráfica 2.4.2.1. Superficie formada por diez esferas



Gráfica 2.4.2.m. Superficie formada por diez esferas



### **2.4.3. Electrón solvatado. 19 esferas**

Hemos calculado la superficie correspondiente a una estructura en forma de octaedro en cuyo centro esta el electrón y en cada vertice una molécula de agua, obtenida en una optimización ab-initio con la base 4-31G (Aulló, J.M., et col, 1987). Hemos asignado los valores 1.0, 1.4 y 1.2 Å a los radios de las esferas correspondientes al electrón, los átomos de oxígeno y de hidrógeno respectivamente. Los parámetros han sido variados del mismo modo que en el ejemplo anterior.

En la **gráfica 2.4.3.a** se da el valor del área ( $\text{Å}^2$ ) para esta estructura, viéndose un comportamiento muy similar al caso de diez esferas. En la **gráfica 2.4.3.b** se representan los errores relativos (%), dado que en este caso no conocemos la superficie exacta, los errores se refieren a los valores obtenidos con NDIV=5 y DEN=25. Los resultados muestran que GEPOL tiene unas fluctuaciones despreciables frente al método MS.

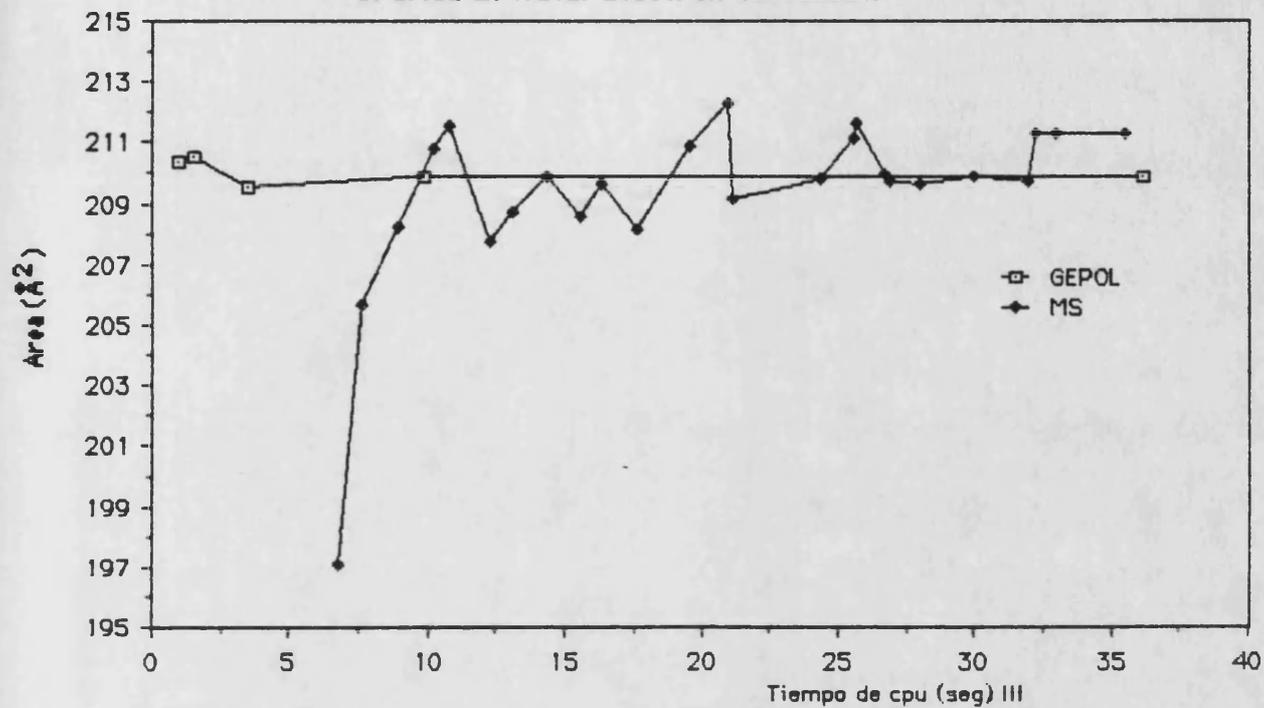
En las **gráficas 2.4.3. c y d** se dan el valor del volumen ( $\text{Å}^3$ ) y el error relativo de ésta. Para el cálculo de el error relativo se ha utilizado el mismo sistema que en el área. Los errores son aproximadamente del mismo orden que en la superficie aunque en las gráficas aparezca un poco exagerado debido a las escalas de los ejes.

En la **gráfica 2.4.3. e** se representa el número de puntos resultantes sobre la superficie frente al tiempo de cálculo. En ella se puede ver que el método MS produce muchos más puntos sobre la superficie que el método GEPOL, esto puede resultar incomodo para una posterior utilización de éstos.

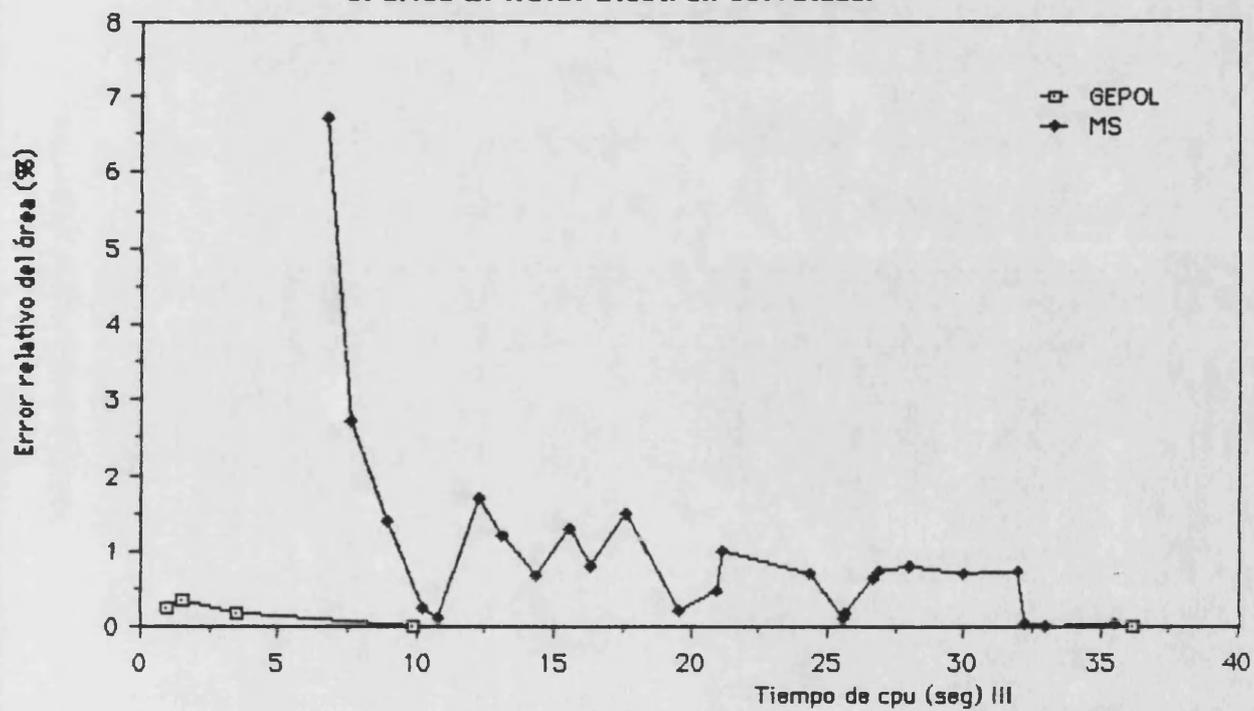
### **2.4.4. NAD y LOOP. 165 esferas**

La LADH es la encima que cataliza la oxidación de los alcoholes primarios y secundarios a aldehidos y cetones respectivamente, con la ayuda de la coencima nicotinamida adenina dinucleótido NAD (Bränden, C.I. y col 1975). Para que la catálisis se realice es necesario que se produzcan ciertos cambios conformacionales en la encima. En concreto el NAD al interaccionar con un fragmento de la LADH, que denominamos LOOP, hace que este se abra o se cierre dejando libre el canal que conduce a la zona activa de la proteína (Tapia, O. y col, 1986). En este apartado

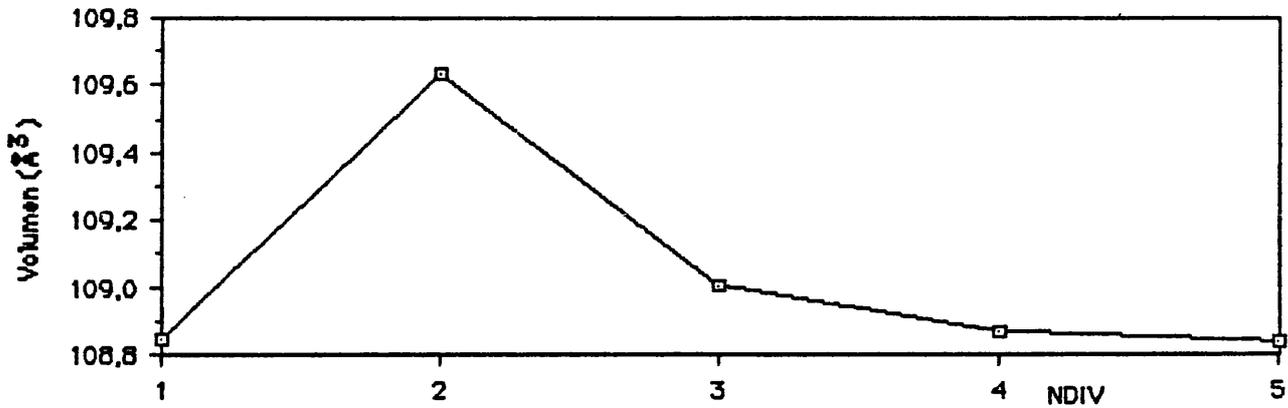
Gráfica 2.4.3.a. Electrón solvatado.



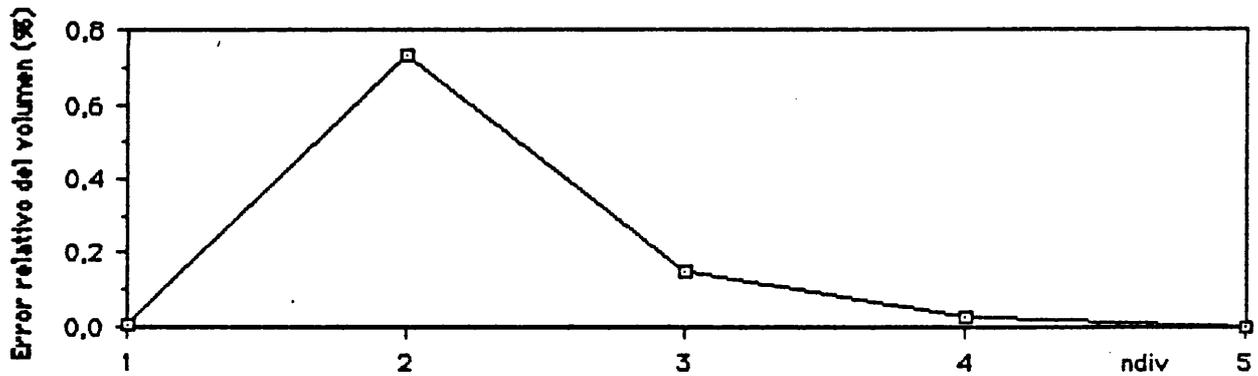
Gráfica 2.4.3.b. Electrón solvatado.



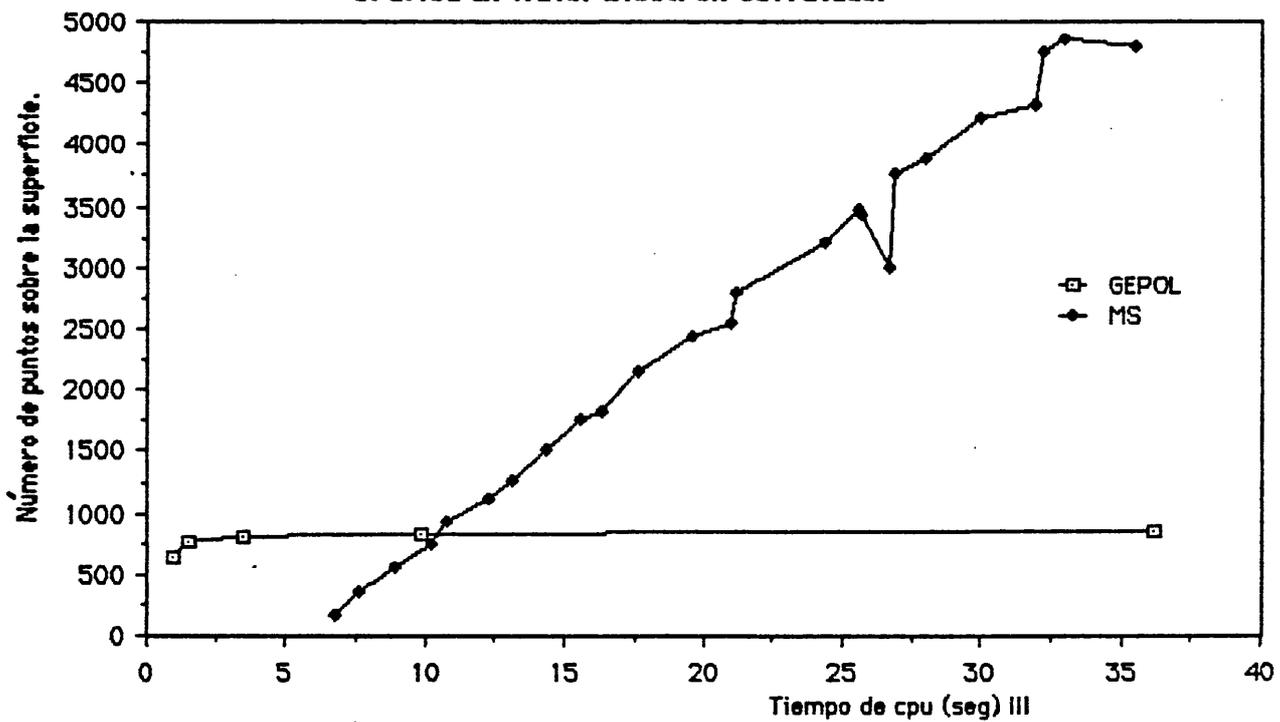
Gráfica 2.4.3.c. Electrón solvatado.



Gráfica 2.4.3.d. Electrón solvatado.



Gráfica 2.4.3.e. Electrón solvatado.



realizaremos cálculos sobre el sistema NAD-LOOP formado por 165 átomos pesados, centrando en cada uno de ellos esferas de un radio igual a 1.8 Å. Las coordenadas cartesianas utilizadas fueron cedidas por O.Tapia y F. Colonna-Cesari y corresponden a un cálculo de dinámica molecular.

Para poder analizar la invariancia espacial en este caso, donde la cavidad no es lineal como en casos anteriores, se eligieron tres orientaciones diferentes por rotación de las coordenadas como se ve en la **tabla 2.4.4. a**

**TABLA 2.4.4. a**

1 <sup>er</sup> cálculo	X	Y	Z
2 <sup>o</sup> cálculo	Y	Z	X
3 <sup>er</sup> cálculo	Z	X	Y

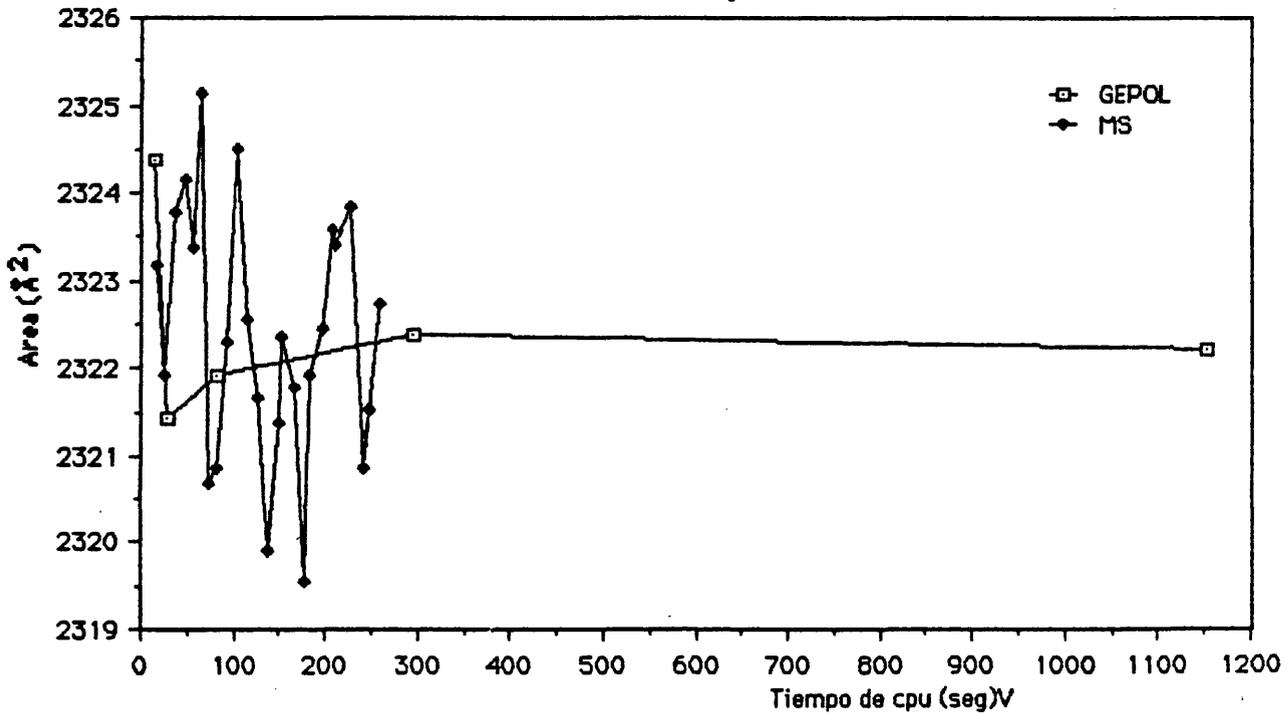
En las **gráficas 2.4.4.a y b** se representa la superficie media y la desviación de estas tres orientaciones frente al tiempo de cpu. Debido a que no se conoce el valor exacto del área se obró como en casos anteriores, calculando el error relativo con relación a los resultados obtenidos para los valores más altos de NDIY y DEN. Así se representan estos errores relativos del GEPOL y del MS frente al tiempo de cpu en las **gráficas 2.4.4 c y d** respectivamente.

Como se deduce de estas gráficas, el método GEPOL sigue resultando más convergente que el MS. En cuanto a la invariancia espacial el método GEPOL resulta mejor sobre todo a partir del nivel NDIY=3 . Los errores en ambos casos son bastante pequeños, pero aun así el comportamiento es mejor en el caso de GEPOL.

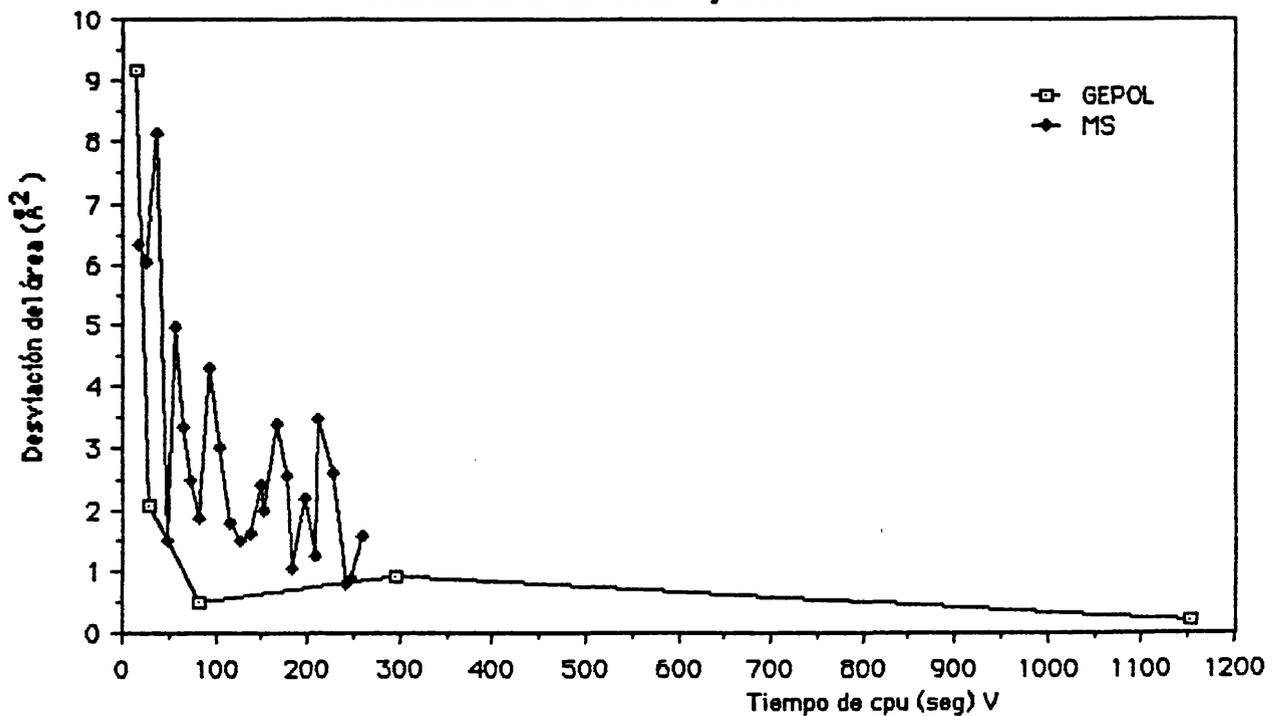
En la **gráfica 2.4.4. e** se representa el número de puntos resultante sobre la superficie para ambos métodos en función del tiempo de cpu.

En las **gráficas 2.4.4. f, g y h** se representan el volumen, desviación y error relativo, de igual forma que para la superficie en función de NDIY.

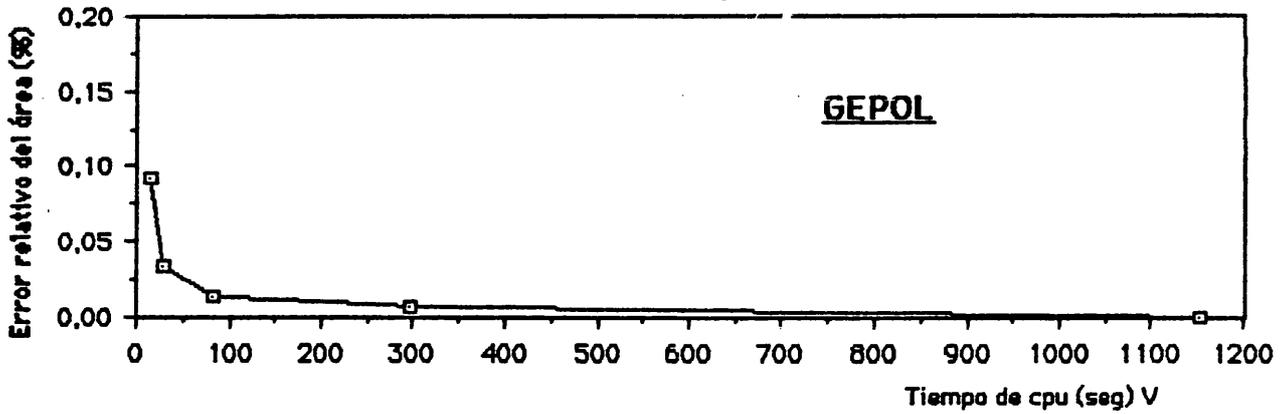
Gráfica 2.4.4.a. NAD y LOOP.



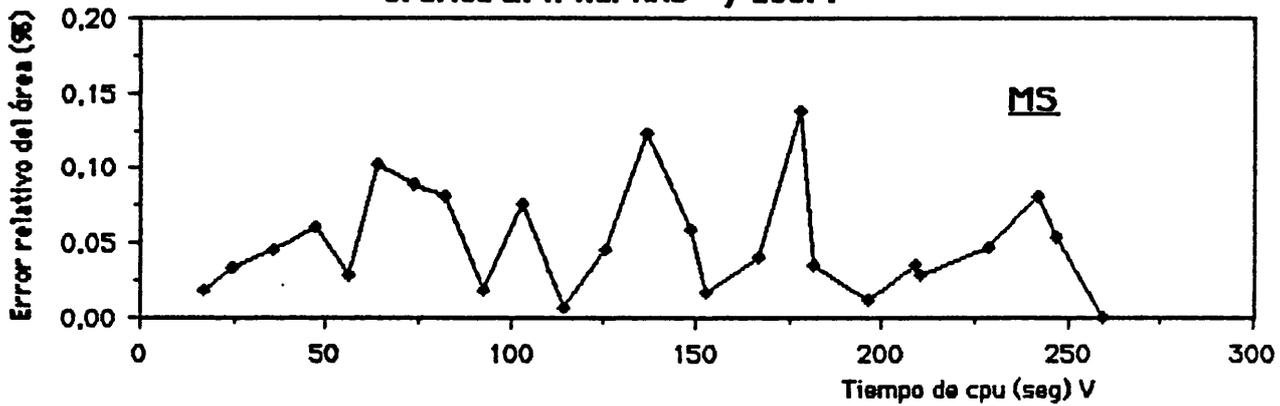
Gráfica 2.4.4.b. NAD y LOOP.



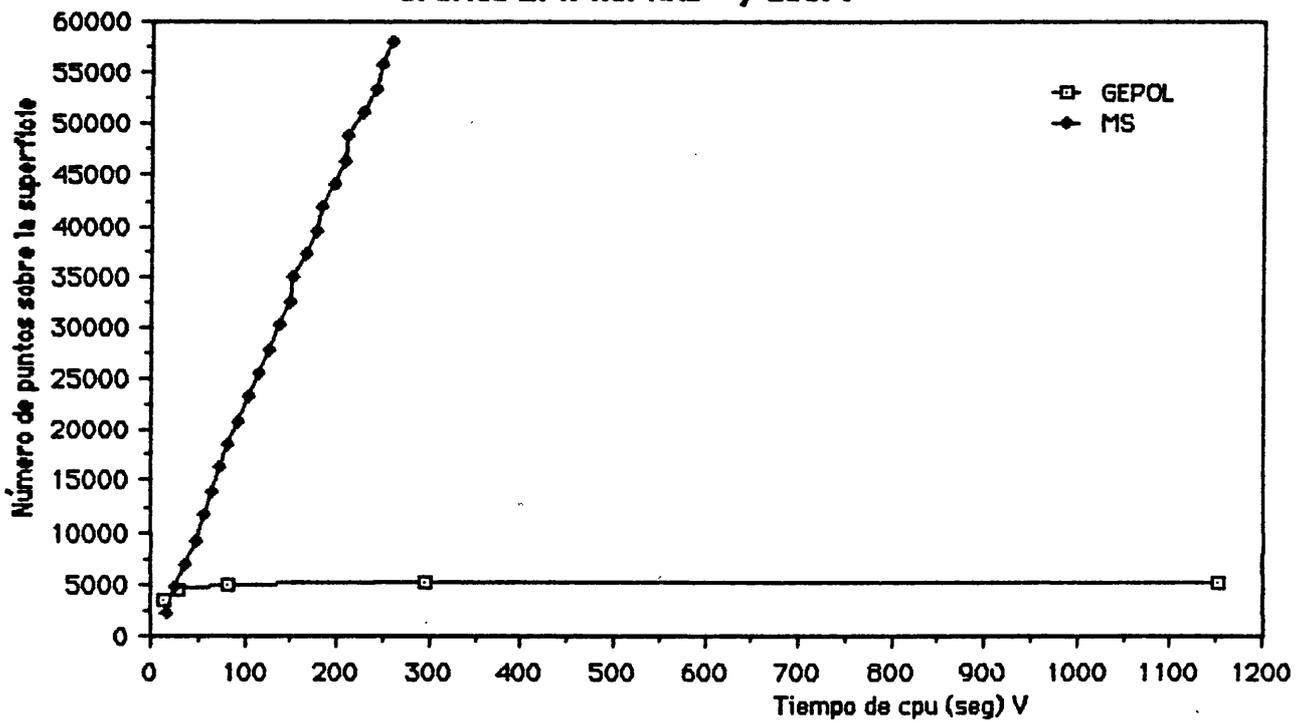
Gráfica 2.4.4.c. NAD y LOOP.



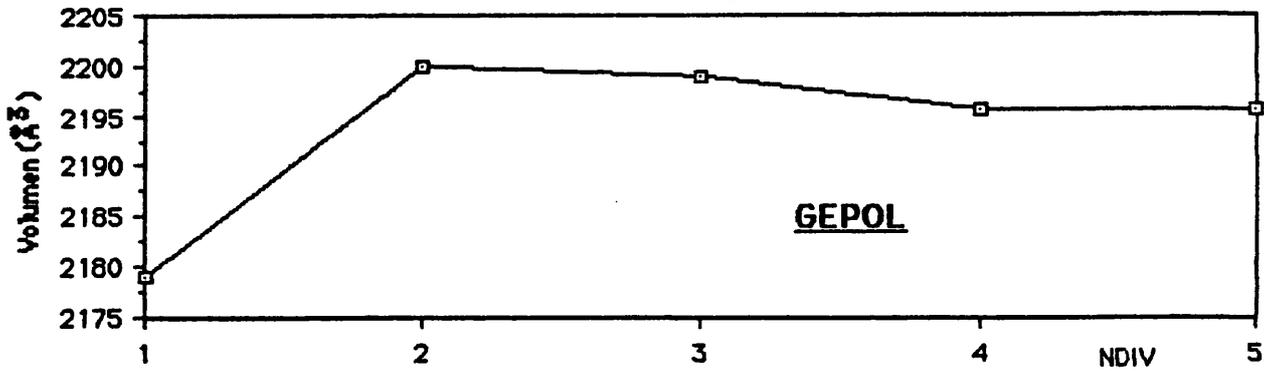
Gráfica 2.4.4.d. NAD y LOOP.



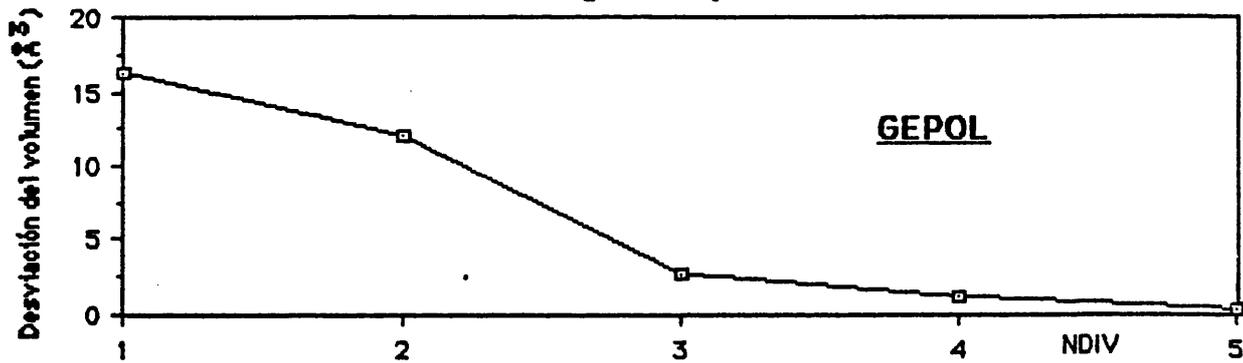
Gráfica 2.4.4.e. NAD y LOOP.



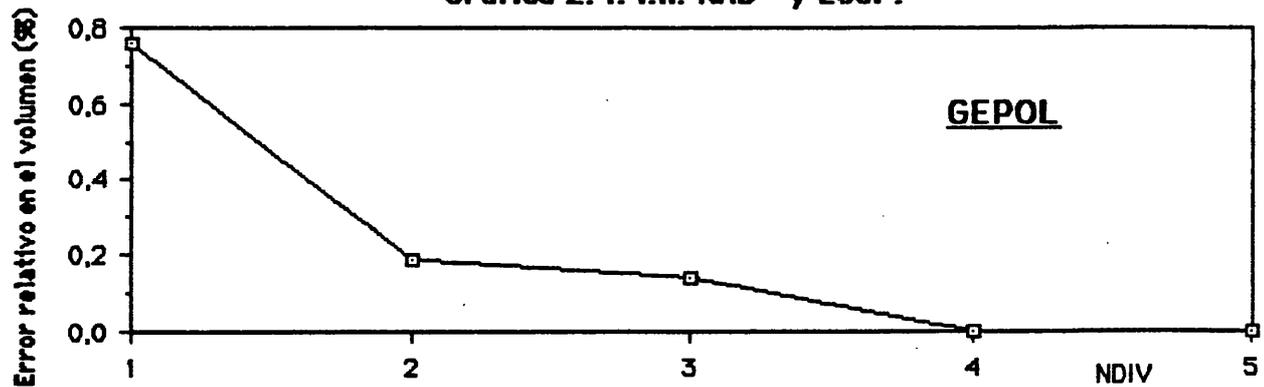
Gráfica 2.4.4.f. NAD y LOOP.



Gráfica 2.4.4.g. NAD y LOOP.



Gráfica 2.4.4.h. NAD y LOOP.



### **2.4.5. CTF. 519 esferas.**

El CTF es el fragmento C-terminal de la proteína ribosómica L7/L12. Esta proteína es esencial para una síntesis eficaz de los polipéptidos en las bacterias (**Möller, W., 1974; Pettersson, I. y Kurland, C. O., 1980**). La estructura cristalográfica tridimensional del CTF está determinada (**Leijonmarck, M. et col., 1980**), sin embargo nosotros hemos utilizado las coordenadas cartesianas producidas en un cálculo de dinámica molecular (**Åquist, J. et col., 1985**) cedidas por los autores. Este fragmento proteínico está formado por 519 átomos pesados a los que les hemos asociado una esfera de radio 1,8 Å.

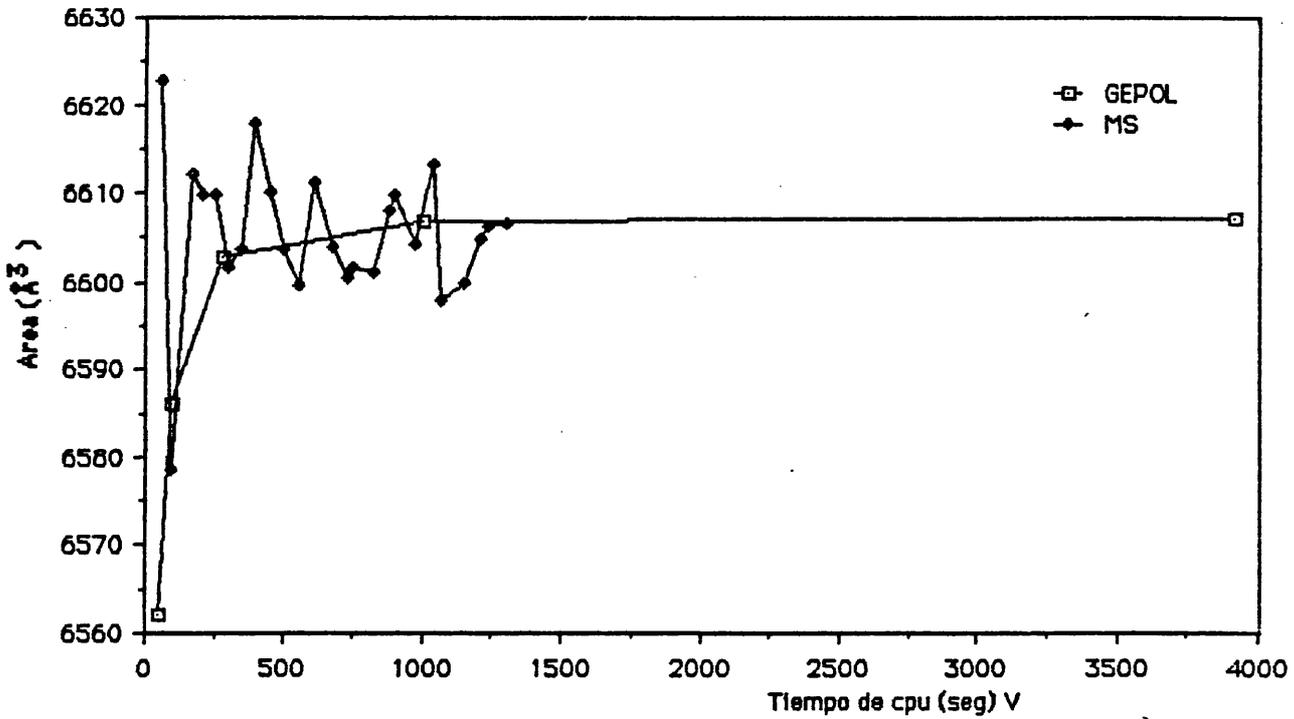
Para esta proteína, así como para la más grande, sólo hemos realizado cálculos en una orientación determinada. Así en la **gráfica 2.4.5. a** se representa el valor del área obtenida frente al tiempo de cálculo, y en las **gráficas 2.4.5 b y c** los errores obtenidos utilizando el sistema antes mencionado. Ambos métodos presentan errores pequeños aunque, como siempre, el método GEPOL tiene un comportamiento más convergente. En este caso las diferencias a nivel de exactitud entre los dos métodos son menores.

En la **gráfica 2.4.5 d** se representa el número de puntos sobre la superficie generados por ambos programas frente al tiempo de cálculo. En las **gráficas 2.4.5. e, f y g** se ha representado el área, volumen, número de puntos y tiempo de cálculo para el cálculo GEPOL en función del parámetro NDIV. En las dos primeras se observa la convergencia de este método en función del factor NDIV. En la representación del tiempo de CPU frente a NDIV se puede ver que la variación es exponencial. Esto, unido a que para el valor de NDIV=5 el error que se comete es totalmente despreciable, nos llevó a no preparar el programa para valores mayores de NDIV.

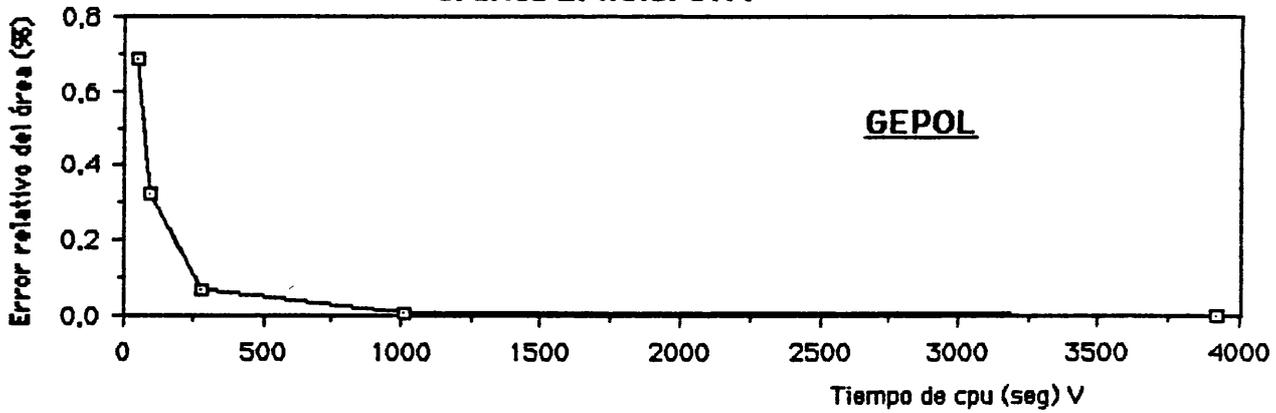
### **2.4.6. RBP.1438 esferas**

Esta proteína no ha podido ser cristalizada y por tanto no se han podido determinar sus coordenadas, por ello hemos utilizado las obtenidas con ayuda de Dinámica Molecular (**Åquist, J. et col., 1986**). Esta proteína está formada por 1438 átomos pesados. Aquí hemos asociado a los átomos de carbono una esfera de radio 1.9, a los de oxígeno una esfera de radio 1.4 y de 1,5 para los

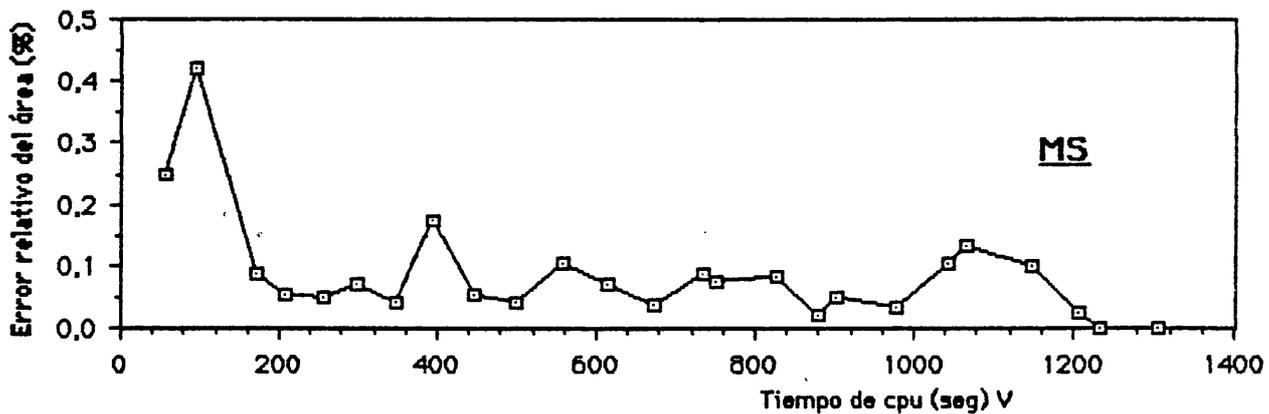
Gráfica 2.4.5.a. CTF.



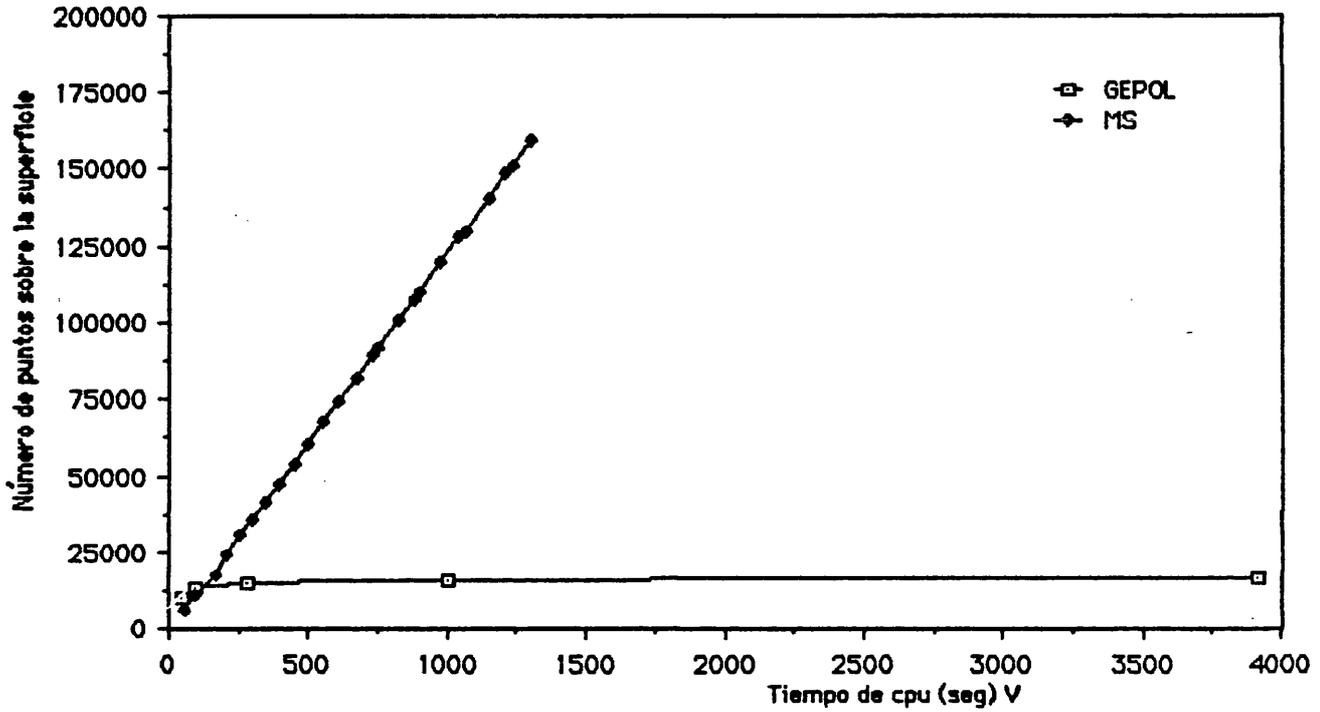
Gráfica 2.4.5.b. CTF.



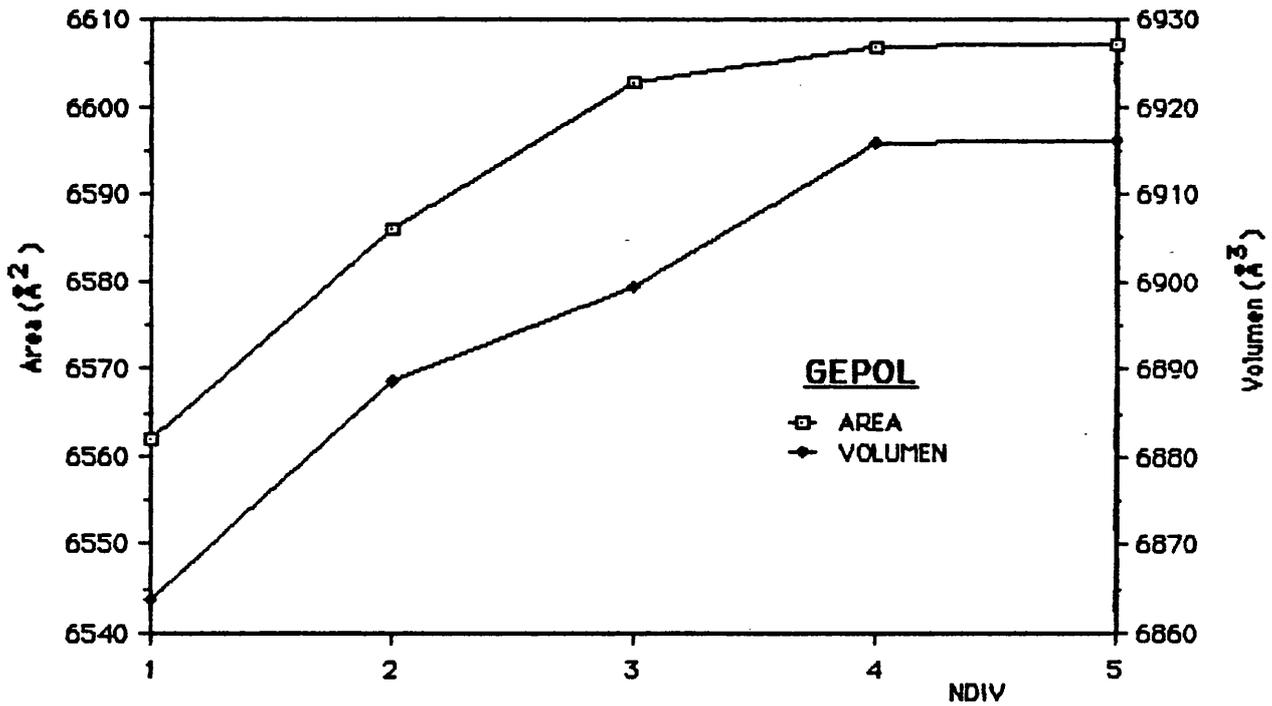
Gráfica 2.4.5.c. CTF.



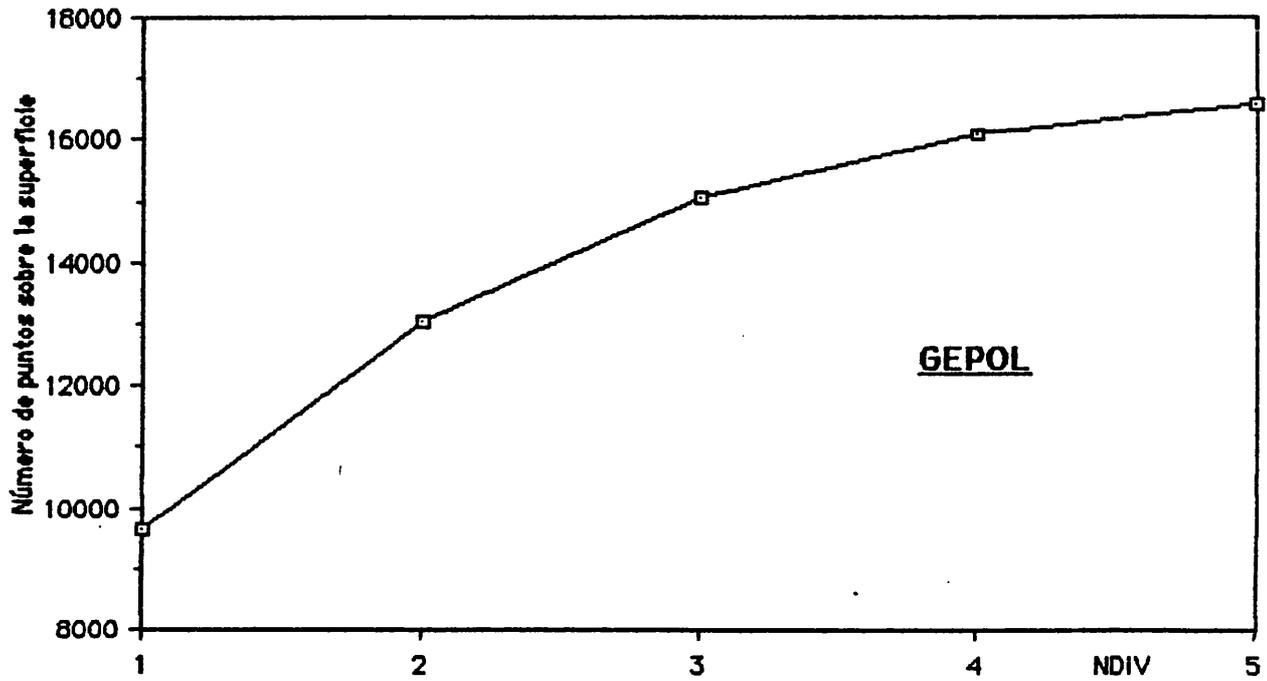
Gráfica 2.4.5.d. CTF.



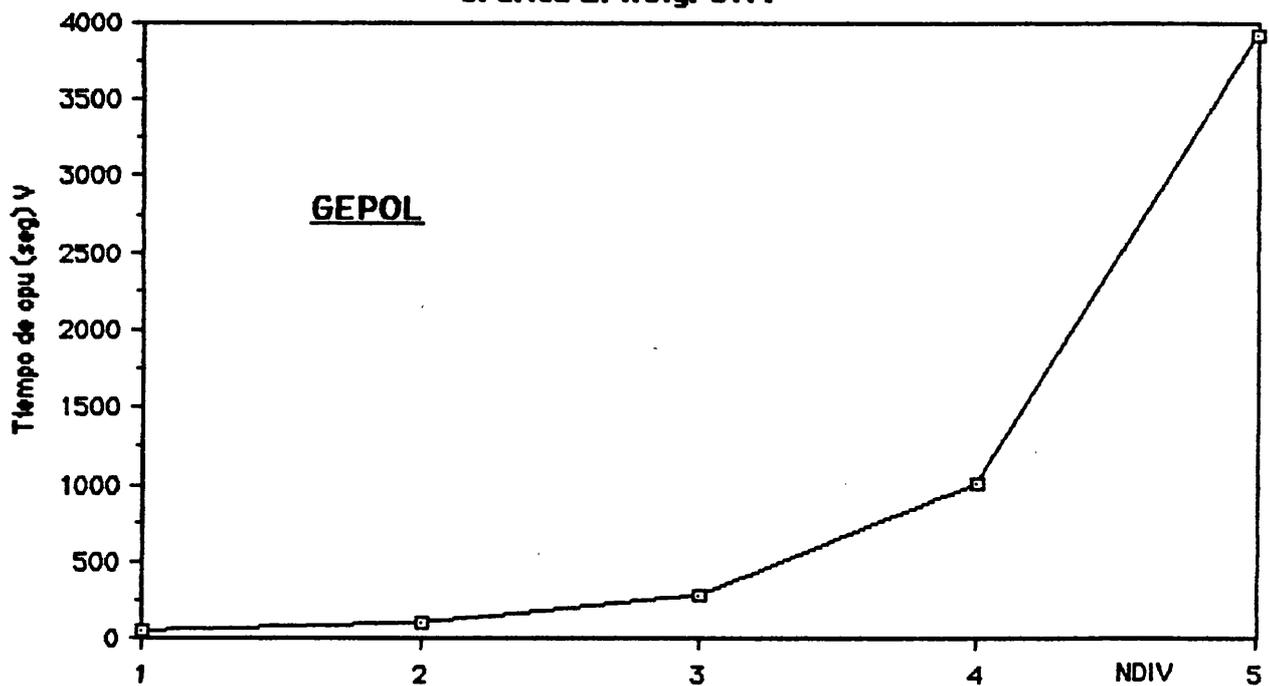
Gráfica 2.4.5.e. CTF.



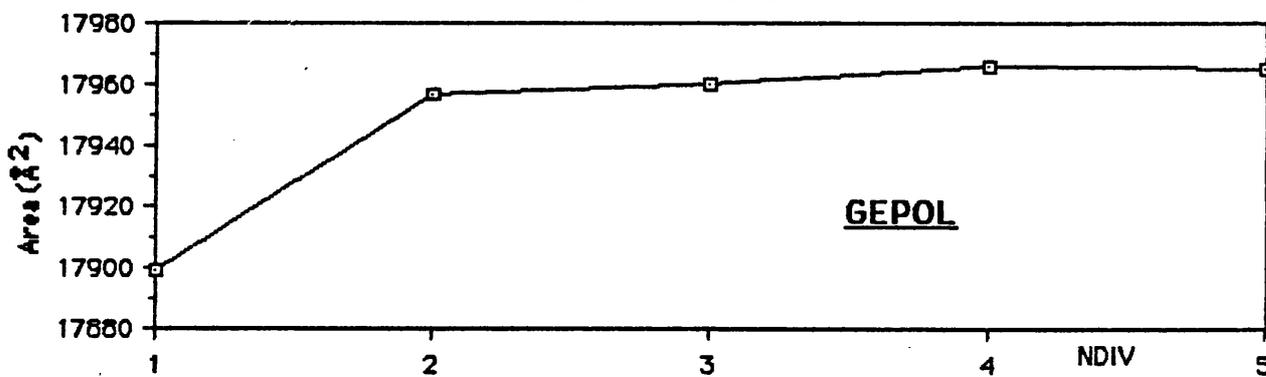
Gráfica 2.4.5.f. CTF.



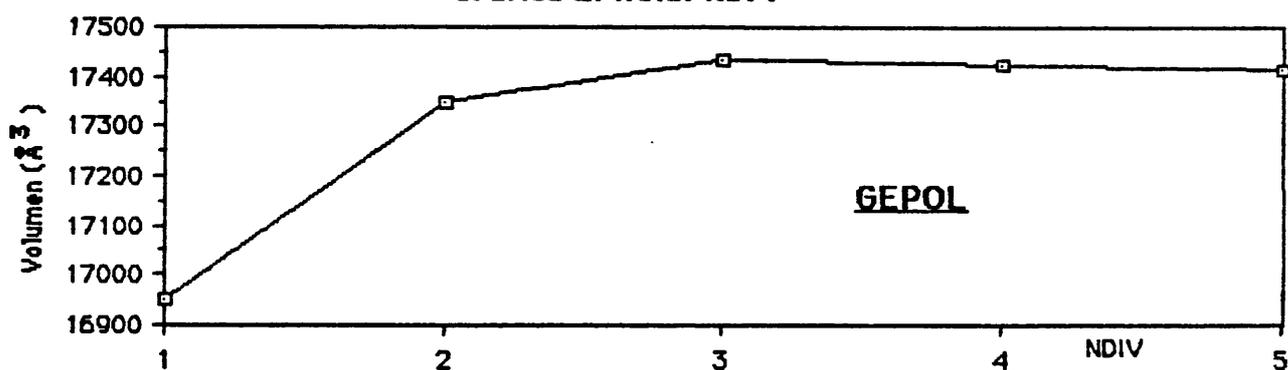
Gráfica 2.4.5.g. CTF.



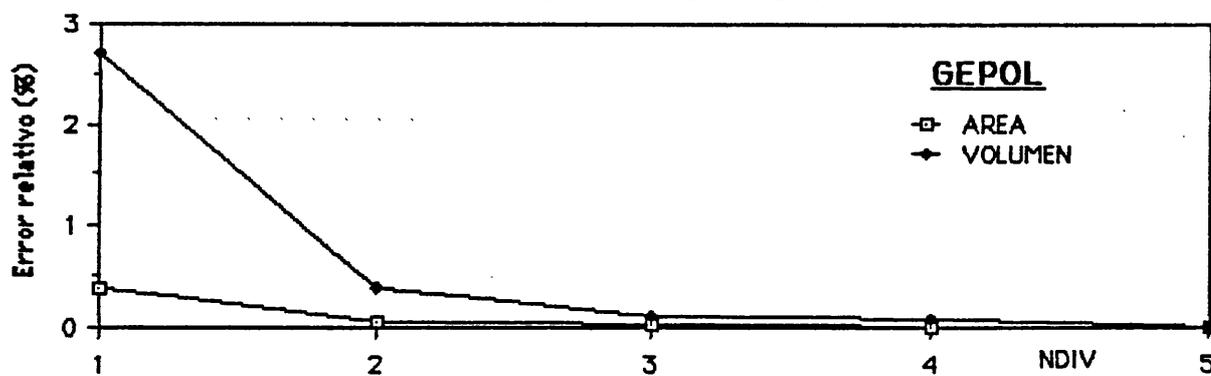
Gráfica 2.4.6.a. RBP.



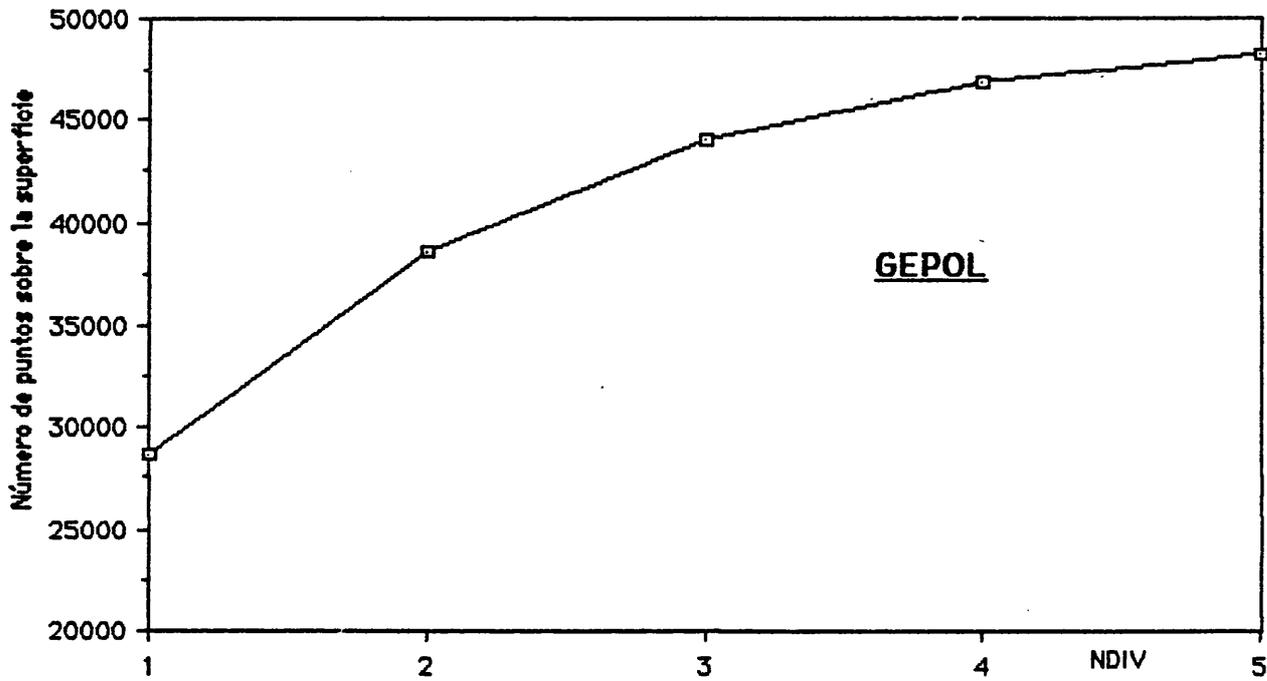
Gráfica 2.4.6.b. RBP.



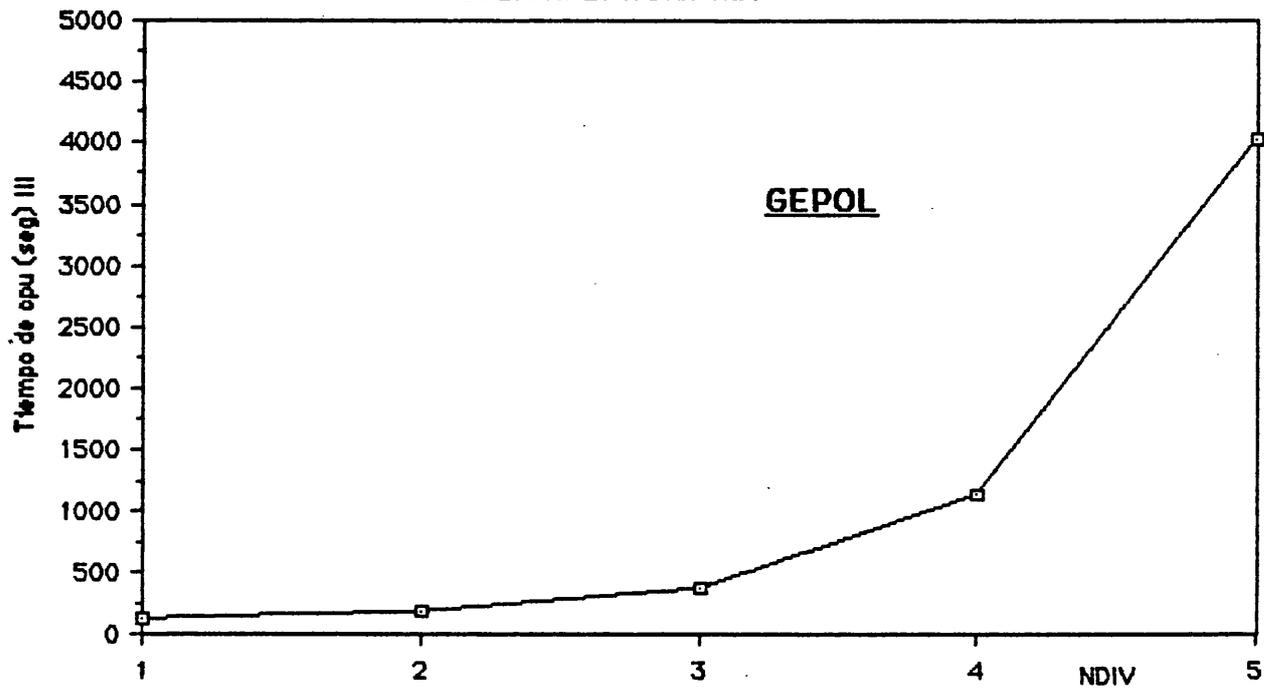
Gráfica 2.4.6.c. RBP.



Gráfica 2.4.6.d. RBP.



Gráfica 2.4.6.e. RBP



nitrógenos.

Para esta molécula sólo hemos realizado cálculos GEPOL. Así en la **gráfica 2.4.6.a y b** se representan el área y el volumen obtenidos en función de NDIV.

En la **gráfica 2.4.6. c** se representa el error relativo, con respecto al resultado obtenido para NDIV=5 , en función de NDIV.

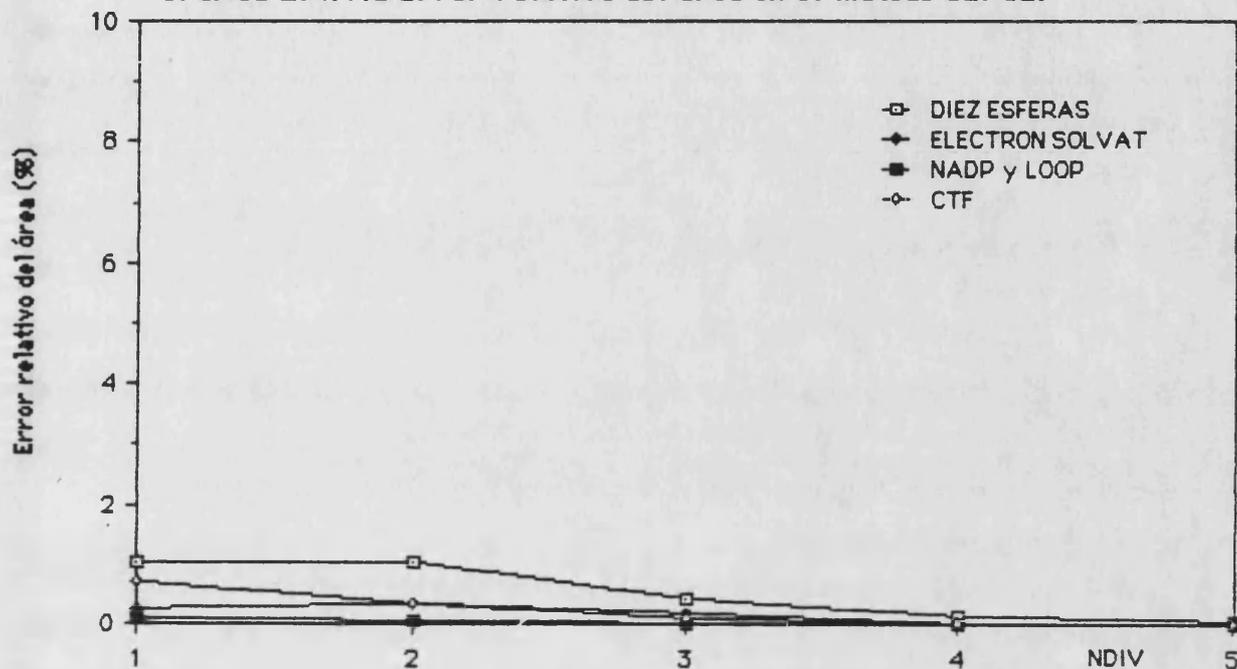
En las **gráficas 2.4.6 d y e** se representa el número de puntos y el tiempo de cálculo en función de NDIV.

### **2.4.7 Discusión general.**

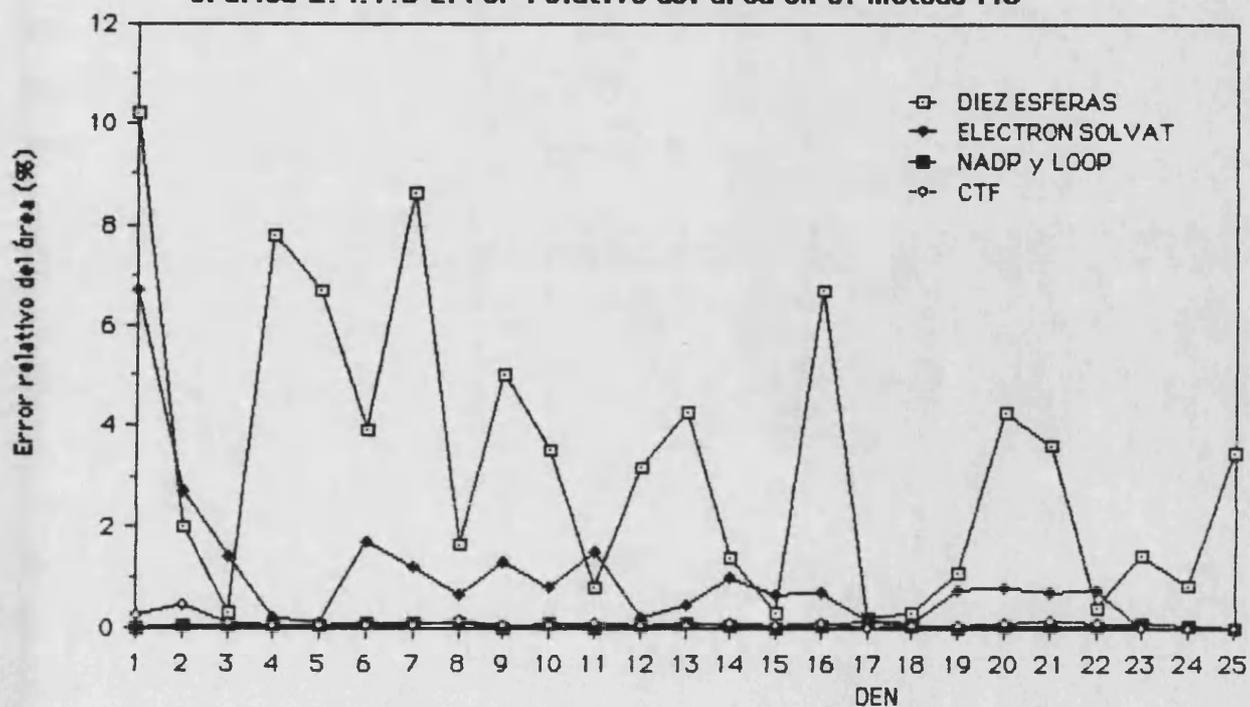
A continuación pasamos a realizar una discusión general de los resultados anteriores. Con respecto a la comparación con el método SPHERE , como ya dijimos , relegamos el análisis a un estudio posterior que realizaremos en el capítulo correspondiente a la energía electrostática de disolución.

A excepción del caso de las diez esferas, no conocemos el valor de área y volumen exactos. Aun así podemos ver, de lo expuesto anteriormente, que el método GEPOL tiene un comportamiento general mejor que el MS. En las **gráficas 2.4.7. a y b** mostramos un extracto de los errores relativos (tal como se definieron para cada caso), para los dos métodos y para las cuatro estructuras en las que se han utilizado ambos, en función de sus parámetros respectivos. En ellas podemos apreciar como el método GEPOL mantiene sus errores siempre dentro un margen del 1%, mientras que en el MS este margen es mucho mayor ,10%. Además el MS disminuye el error en función del tamaño de la estructura, es decir en función del número de esferas, con lo cual se podrían dar situaciones donde el error en la superficie total fuera del 0,1% y en cambio el área sobre un fragmento en particular tuviese un error del 10%, es decir 100 veces más. Sin embargo en el método GEPOL, aunque sucede el mismo fenómeno de compensaciones, la diferencia es mucho menor.

En las gráficas se observa una excepción con respecto a esta disminución del error en función del número de esferas en el caso del NAD , ya que éste mantiene en ambos métodos errores menores que el CTF. Esta aparente discrepancia se debe a que estos valores de NAD corresponden a una

Gráfica 2.4.7.a Error relativo del área en el método  $\Theta$ EPOL.

Gráfica 2.4.7.b Error relativo del área en el método MS

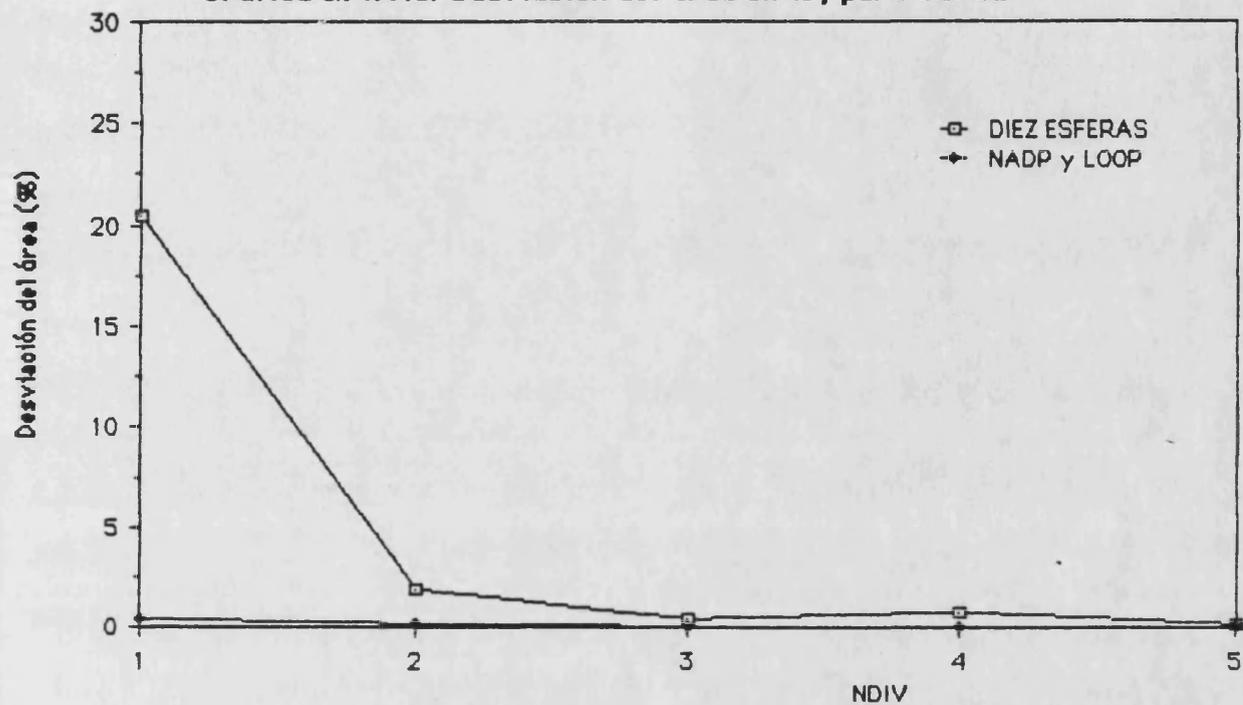


media de tres orientaciones espaciales distintas con lo que produce una compensación de los errores de una con las otras. Por ello para tener un cuadro completo hemos representado en la **gráficas 2.4.7 c y d** las desviaciones del valor medio de la superficie para los dos métodos (para diez esferas y para el NAD) frente a sus respectivos parámetros. Las desviaciones en estas gráficas se han representado en % con respecto a los valores medios de sus respectivas áreas medias con el fin de que se puedan comparar las dos estructuras. En estas gráficas se observa un mismo comportamiento que en las anteriores ya que la desviación se hace menor con el número de esferas de la estructura. Además GEPOL sigue manteniendo un mejor comportamiento que MS, aunque para  $NDIV=1$  presenta una desviación grande.

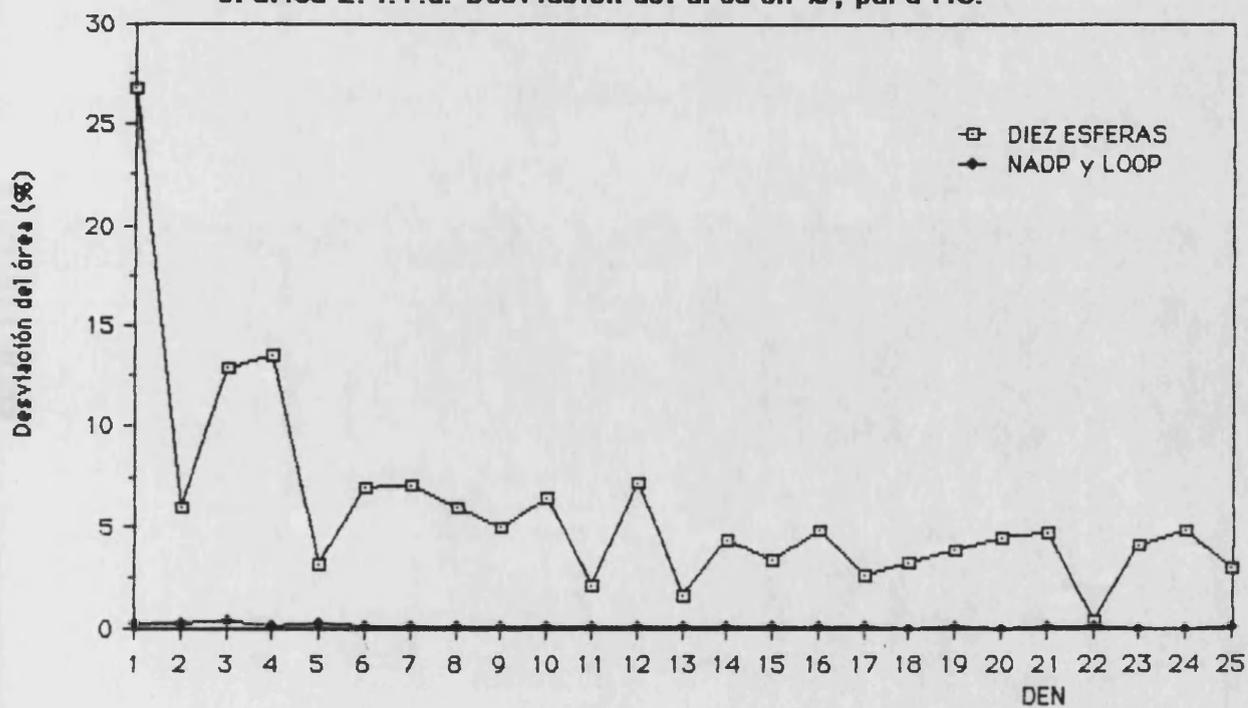
Aquí estamos haciendo un estudio comparativo y por tanto hacemos cálculos variando todos los parámetros, sin embargo en un estudio particular se necesita saber qué parámetro utilizar. Esta decisión viene mediatizada por el tiempo de cálculo disponible y por la naturaleza del estudio que se realiza. Si lo que se quiere es conocer el área total de una proteína, con cualquier de los dos métodos y utilizando cualquier valor para sus parámetros se obtiene un resultado bueno. Pero si lo que se calcula es una estructura pequeña o se quiere realizar un estudio de las áreas locales de una estructura grande la cosa es muy diferente. El método MS tiene un comportamiento bastante oscilante (ver **gráfica 2.4.7. b**) en función de su parámetro y por tanto no se puede indicar un valor. El método GEPOL sin embargo tiene un comportamiento convergente como se puede ver en todas las gráficas que hemos dado, convergencia que se da tanto en los valores del área y del volumen como en el número de puntos que se generan sobre la superficie. Nosotros en los cálculos que realizamos en apartados posteriores utilizamos  $NDIV=3$  porque pensamos que da unos valores buenos en un tiempo razonable.

Con respecto al tiempo, si analizamos las gráficas de apartados anteriores vemos que la relación tiempo frente al error es menor generalmente en el método GEPOL. Aunque se observa que al aumentar el tamaño de las estructuras aumenta más el tiempo de este método que el MS, hay que indicar que en el tiempo que utiliza MS en calcular la superficie el método GEPOL calcula la superficie y el valor del volumen con lo cual da más información.

Gráfica 2.4.7.c. Desviación del área en %, para GEPOL



Gráfica 2.4.7.d. Desviación del área en %, para MS.



Estos métodos de cálculo de superficies no sólo son utilizados para el cálculo de valores del área o del volumen sino que se utilizan los puntos que generan sobre la superficie para realizar otros tipos de cálculos tal y como haremos nosotros en el capítulo 3, o para representar la superficie en una pantalla de ordenador. En estas aplicaciones se necesita que a cada punto le venga asociada una área de tal forma que la suma de todas las áreas de estos puntos sea el área total y que el valor asignado represente bien la zona donde está situado este punto. También es necesario que el número de puntos sobre la superficie no sea excesivo, de tal manera que sean manejables para cálculos posteriores. Con respecto al primer punto, que hace referencia a la exactitud en los cálculos locales, ya hemos hablado en párrafos anteriores. Con respecto al número de puntos que se generan en la superficie podemos ver en las gráficas de apartados anteriores que el número de ellos en el MS aumenta linealmente con el parámetro utilizado, mientras que en el GEPOL el número es más pequeño aparte de variar menos con NDIV, ya que de hecho tiende a un valor asintótico.

## **2.5. CALCULOS DE SUPERFICIES MOLECULARES.**

Como indicamos en el apartado 2.1, aunque el algoritmo de suavización depende internamente de tres parámetros: RET, OMEGA y FRO, desde el INPUT sólo hay que indicar el valor del parámetro FRADIO. En este apartado justificaremos esta manera de actuar, a la vez que analizaremos la influencia del valor de FRADIO sobre las superficies calculadas. También analizaremos los resultados comparándolos con los obtenidos con el programa MS. A no ser que se indique otra cosa, dentro de este apartado siempre utilizaremos NDIV=3 y un radio del disolvente igual a 1.5.

### **2.5.1 NAD y LOOP. Reducción de RET y FRO a FRADIO.**

Esta estructura, como vimos, está formada por 165 átomos pesados en los que hemos centrado esferas de 1,8 Å de radio.

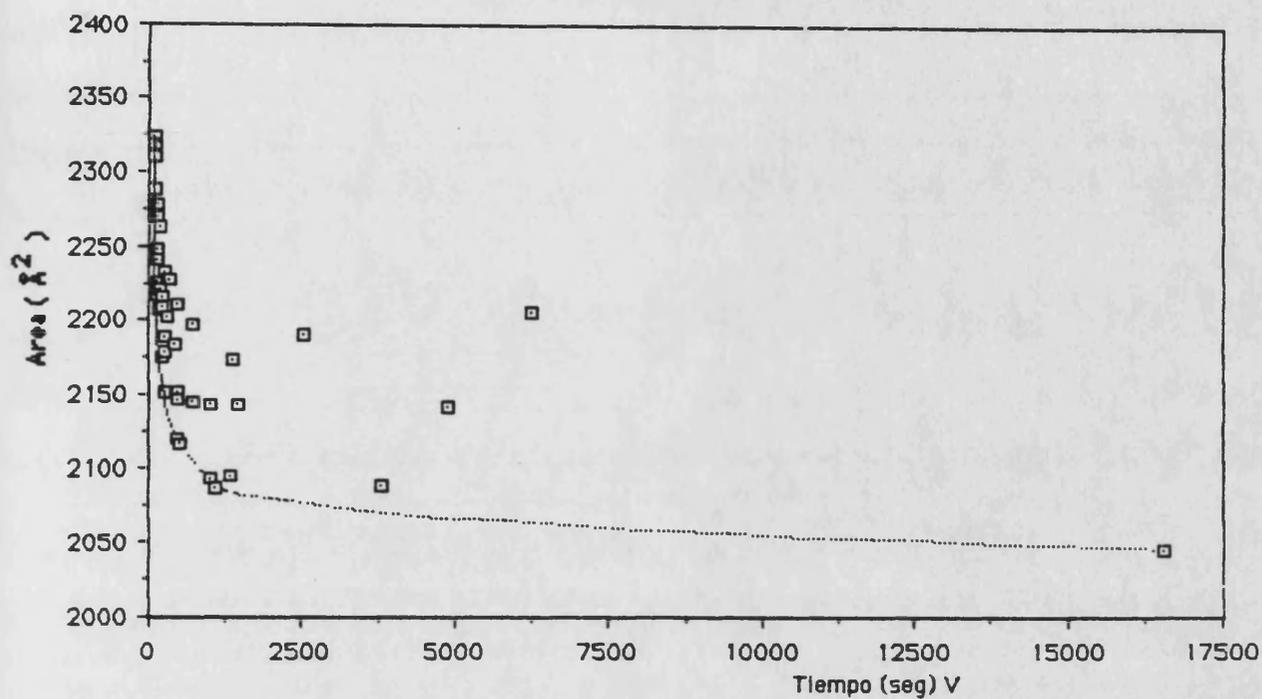
Para realizar el estudio de los tres parámetros, fijamos en principio omega, dándole un valor de 50. Omega puede variar entre 0 y 90. Cuando vale 0 indica que no es necesario que las esferas estén solapadas y a medida que aumenta este valor el solapamiento es mayor. Así, un valor de 50 es una situación intermedia que puede servir para empezar el análisis.

Una vez fijado este parámetro, variamos los dos restantes tal como se muestra en la **tabla 2.5.1 a**, donde con una cruz se indican los cálculos realizados.

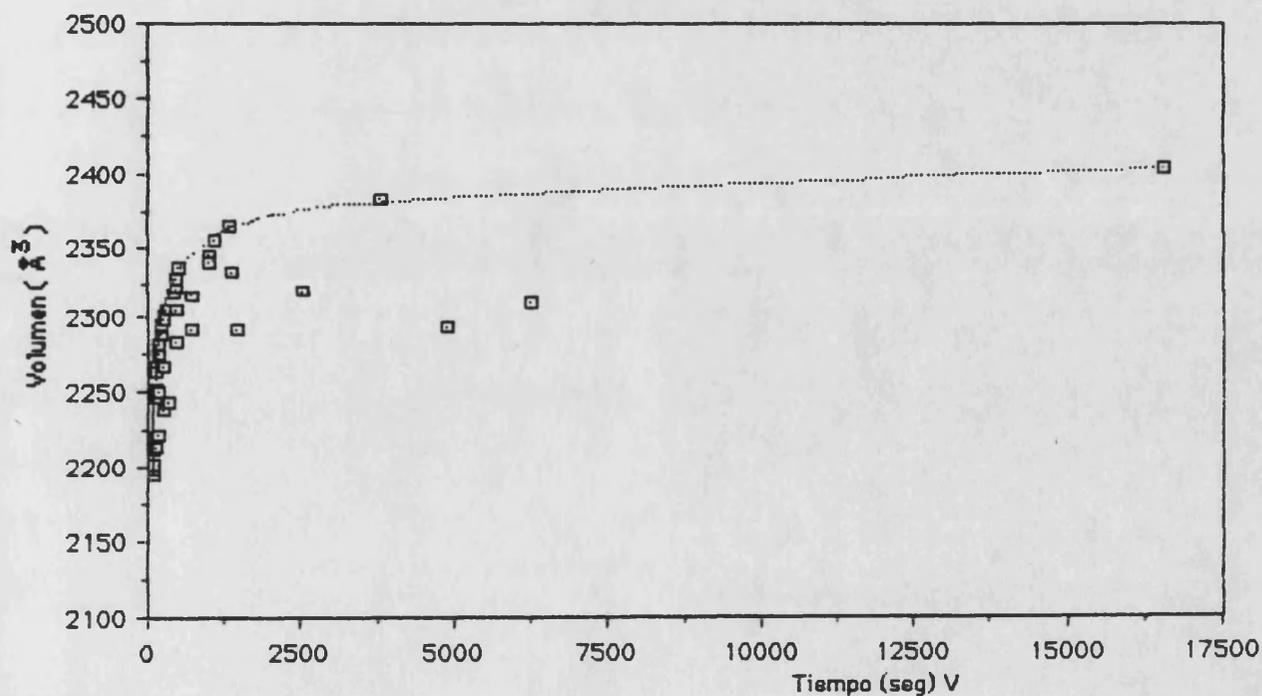
En esta tabla damos los valores de RET en Å, que es como los utiliza el programa, y en tanto por uno con respecto a 1.8 Å, que es el radio de las esferas de partida. En la columna correspondiente a RET=.9 no se realizaron más cálculos ya que no se modificaba para nada la superficie de van der Waals. Tampoco se realizaron cálculos para aquellos valores donde se preveía un tiempo de cálculo excesivo.

La superficie y volumen obtenidos con estos valores se muestran en las **gráficas 2.5.1 a** y **b**. La representación del tiempo, en abcisas, nos permite valorar la eficacia de los diferentes

**Gráfica 2.5.1.a. Área del NAD y LOOP frente al tiempo, para diversos valores de RET y FRO.**



**Gráfica 2.5.1.b. Volumen del NAD y LOOP frente al tiempo, para diversos valores de RET y FRO.**



valores de los parámetros. Sabemos que al pasar de la superficie de van der Waals a la superficie molecular, el valor del área tiene que disminuir mientras que el volumen aumentar (vease la **figura 2. a**) por tanto habrá que utilizar aquellos valores para RET Y FRO que hagan cumplir esto en el mínimo tiempo de cálculo posible. En estas gráficas se puede observar una especie de límite, línea de puntos, correspondiente a la eficacia entre el área y volumen obtenidos y el tiempo de cálculo utilizado. Así, lo interesante sería obtener una serie de parámetros que nos hiciese deslizarnos por esa línea o cerca de ella.

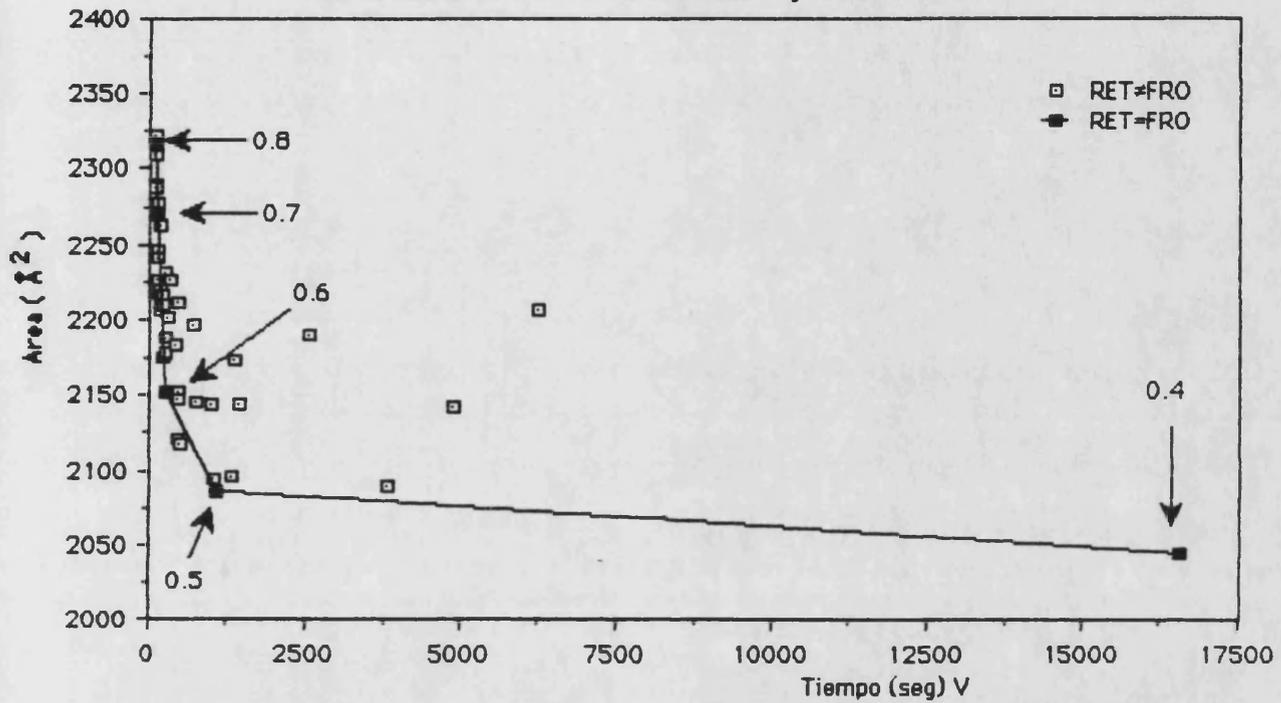
**Tabla 2.5.1.a.** Valores de RET y FRO utilizados.

RET(Tanto por uno)		.1	.2	.3	.4	.5	.6	.7	.8	.9
RET(Å)		0.18	0.36	0.54	0.72	0.9	1.08	1.26	1.44	1.62
FRO	.1	-	-	-	-	-	-	+	+	+
	.2	-	-	-	-	-	-	+	+	-
	.3	-	-	-	-	-	-	+	+	-
	.4	-	-	-	+	-	+	+	+	-
	.5	-	-	-	-	+	+	+	+	+
	.6	-	-	+	+	+	+	+	+	-
	.7	-	-	+	+	+	+	+	+	-
	.8	-	+	+	+	+	+	+	+	-
	.9	+	+	+	+	+	+	+	+	-
	1.0	+	+	+	+	+	+	+	+	-

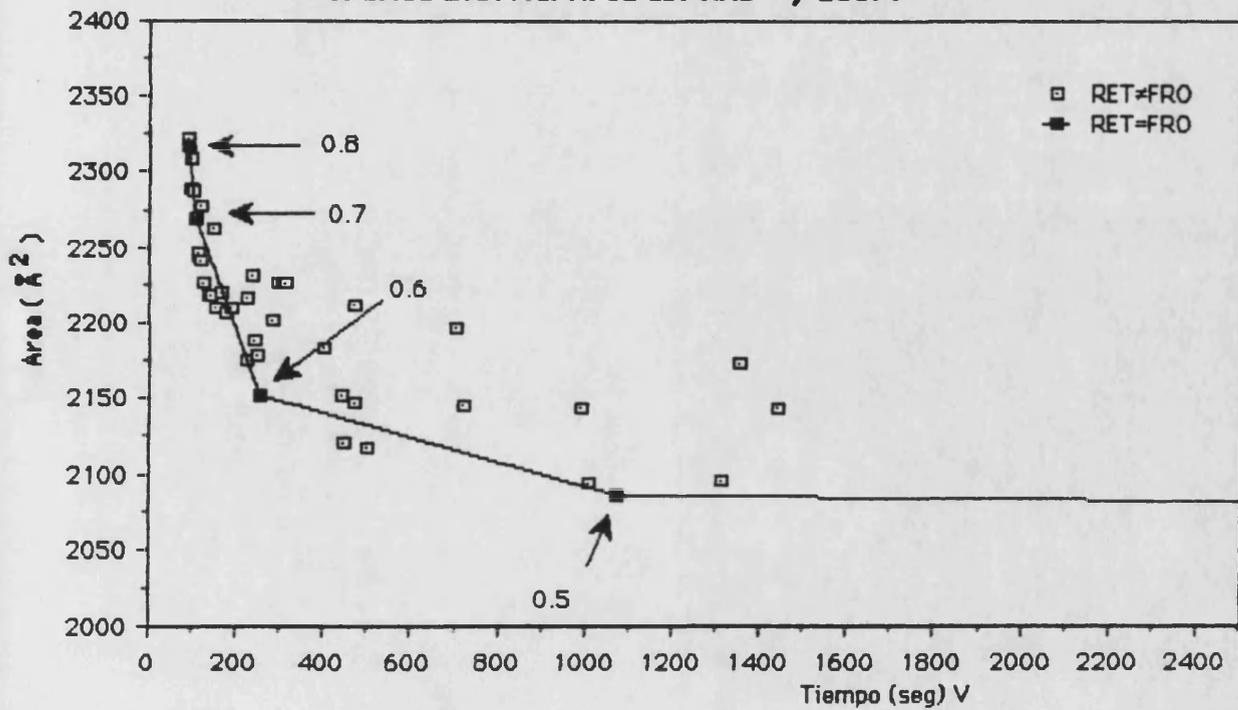
En las **gráficas 2.5.1. c y e** se representan los mismos valores del área y el volumen que en las anteriores y en las **gráficas 2.5.1 d y f** una ampliación de las mismas. La diferencia con las anteriores estriba en que hemos señalado en negro los resultados obtenidos con los valores de los parámetros que se encuentran en la diagonal de la **tabla 2.5.1. a**. Al lado de cada uno de estos puntos se da el valor de RET(en tanto por uno)y FRO que, evidentemente, coinciden. Como se puede ver, estos valores siguen el límite que antes comentábamos. Este resultado es lógico ya que en ellos se cumple que la variación de los dos parámetros resulta compensada. Es decir, los dos parámetros toman valores iguales y por tanto se le van dando exigencias similares al programa.

La observación de este efecto es lo que nos llevo a unificar RET y FRO bajo el parámetro FRADIO.

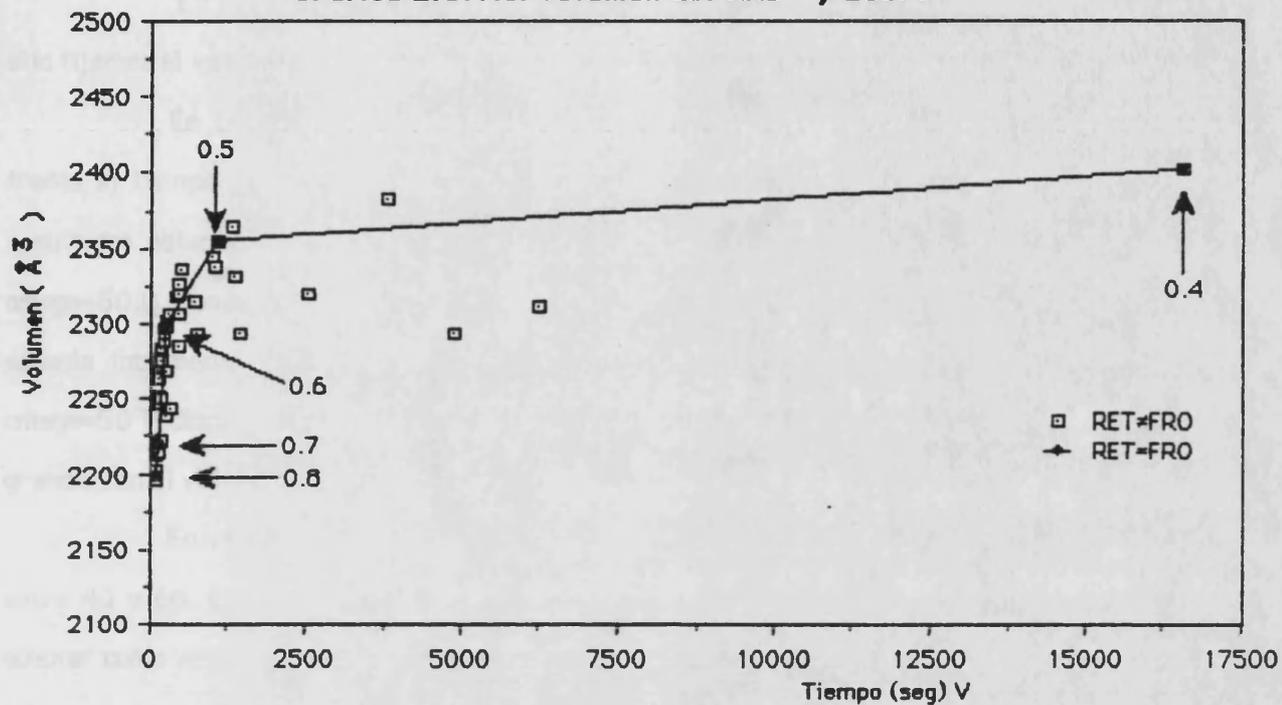
Gráfica 2.5.1.c. Area del NAD y LOOP.



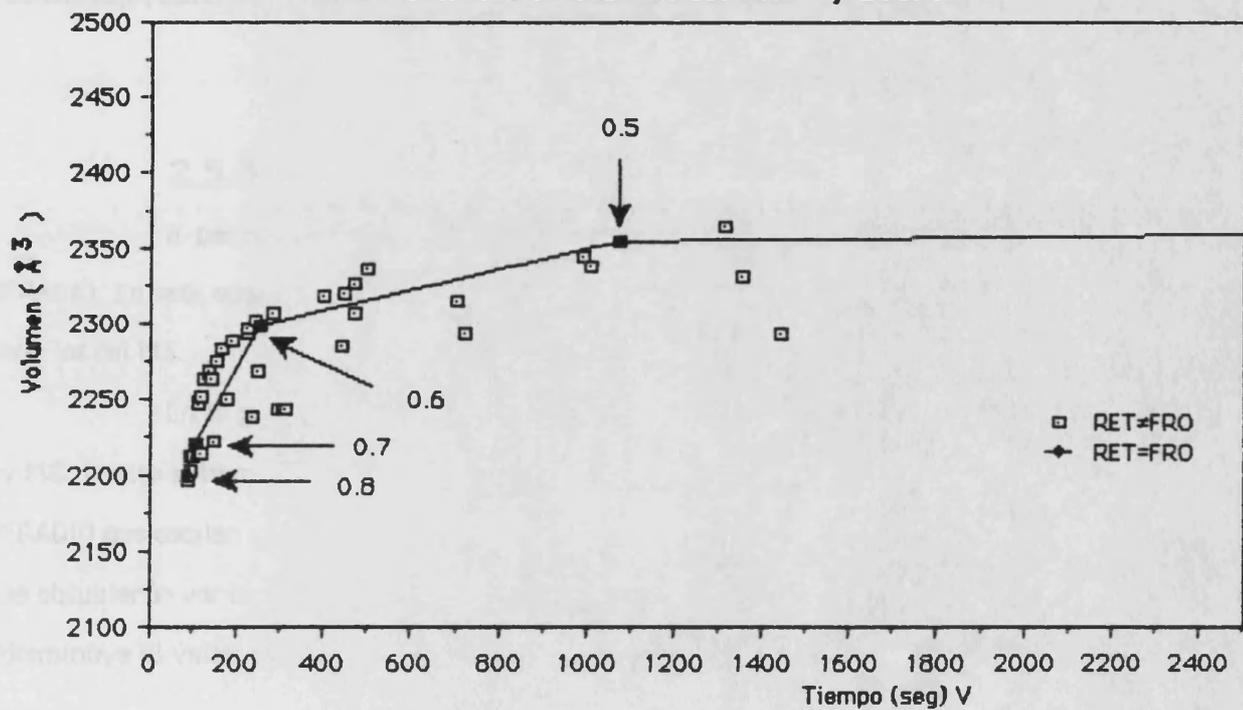
Gráfica 2.5.1.d. Area del NAD y LOOP.



Gráfica 2.5.1.e. Volumen del NAD y LOOP.



Gráfica 2.5.1.f. Volumen del NAD y LOOP.



### **2.5.2 NAD y LOOP. Variación con omega.**

En este apartado analizaremos la variación de la superficie con el valor de omega. Para ello fijamos el valor de FRADIO en 0.6 ( $FR0=0.6$  y  $RET=1.08 \text{ \AA}$ ) y variamos omega entre 20 y 80.

En las **gráficas 2.5.2. a y b** se representan los valores del área y del volumen frente al tiempo de cálculo para los diferentes valores de omega utilizados. También se dan los resultados obtenidos en el apartado anterior para los valores de  $FRADIO=0.8$ ,  $0.7$  y  $0.5$ , con  $omega=50.0$ . Como se puede ver, el valor del área disminuye con el aumento del valor de omega, aspecto totalmente predecible dada la definición de omega. Pero cuando se supera el valor de  $omega=50$  la disminución de superficie viene acompañada de un aumento de tiempo de CPU demasiado grande. Con el volumen sucede algo similar, con la salvedad de que el volumen aumenta.

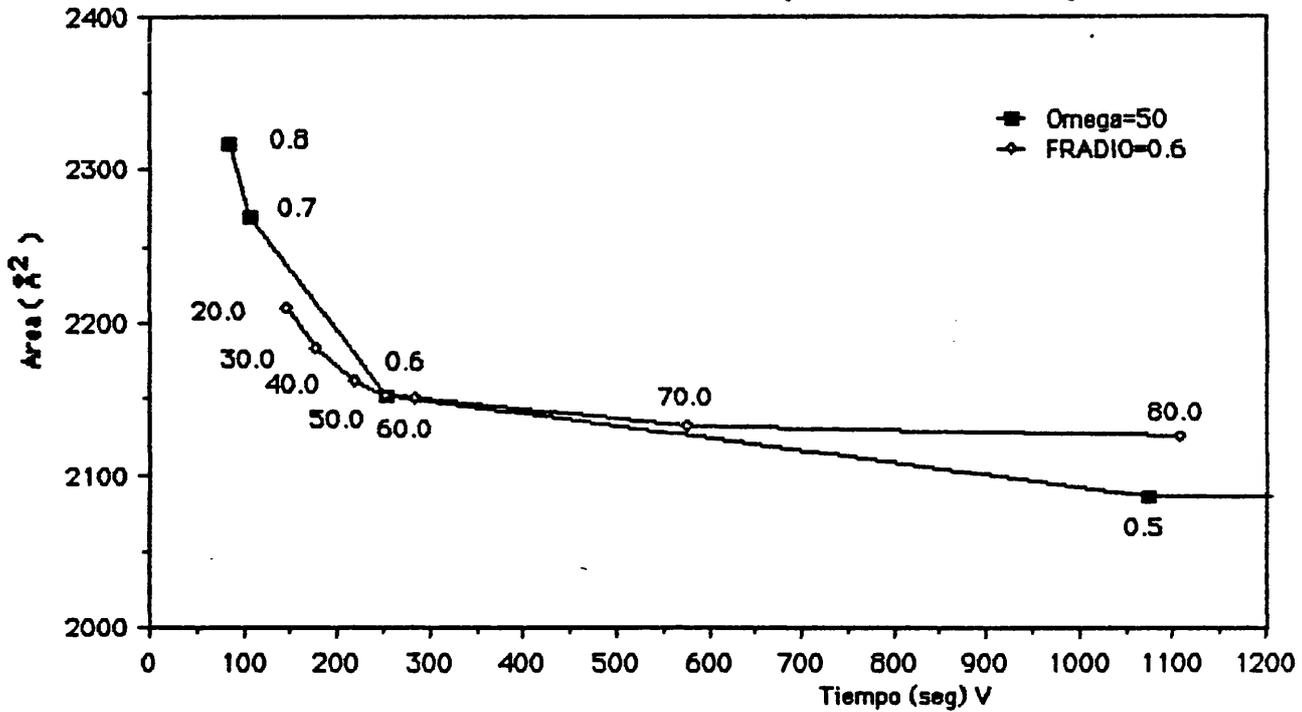
En vista de lo anterior, se puede considerar que un valor bueno de omega se encontrará entre 40 y 60. Esto unido a que la superficie es más sensible al parámetro FRADIO, nos decidió a adoptar como valor de omega el de 50. Así se introdujo en el programa este valor como un DATA, de tal manera que para controlar este algoritmo desde el INPUT sólo se utiliza el parámetro FRADIO. Aunque esto quita libertad de movimiento en el cálculo, aumenta la facilidad de uso del programa y, desde luego, como veremos en los apartados posteriores, no perjudica a los resultados.

### **2.5.3. NAD y LOOP. Análisis de resultados.**

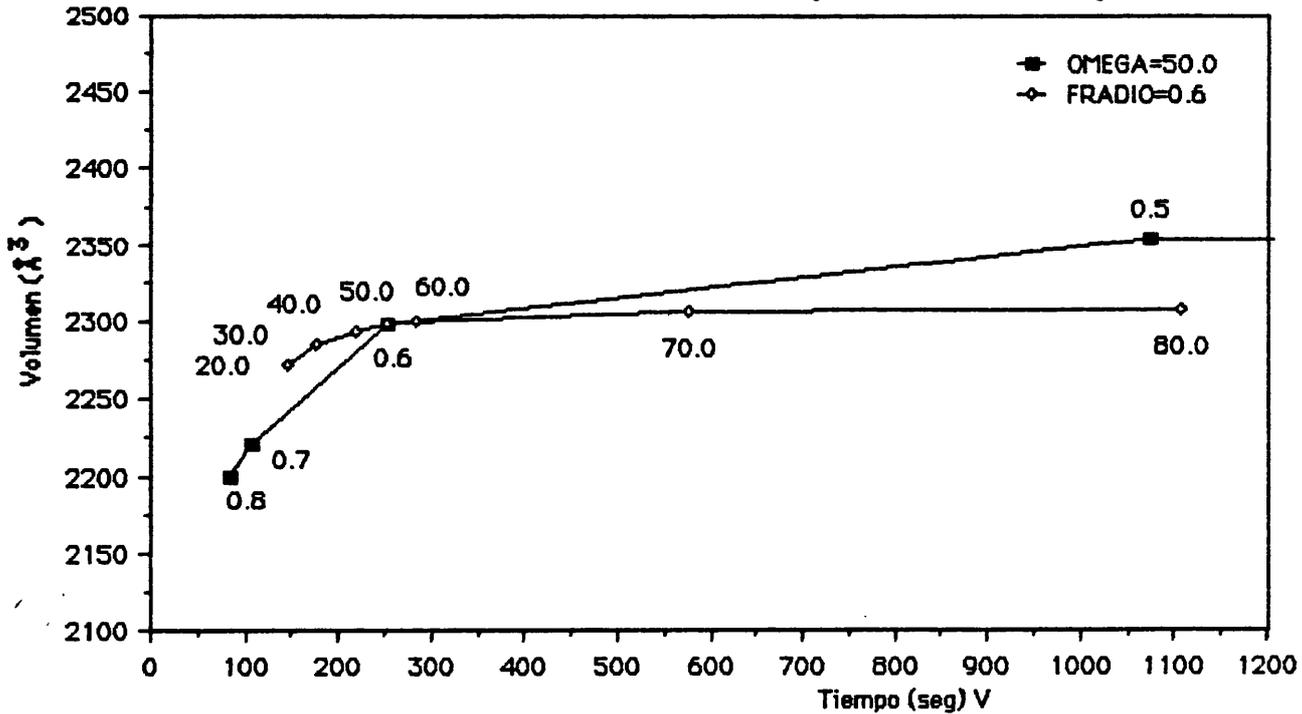
A partir de ahora utilizaremos la versión en la que se indica solamente el valor de FRADIO. En este apartado analizaremos los resultados frente a este parámetro y los compararemos con los del MS.

En la **gráfica 2.5.3. a** se representa el valor del área obtenida con los métodos GEPOL y MS, frente al tiempo de cpu utilizado. Los valores de GEPOL corresponden a cálculos con valores de FRADIO que oscilan entre 0.85 y 0.45, en incrementos de 0.05. Los resultados obtenidos con el MS se obtubieron variando DEN desde 1 a 25 en incrementos de 1. Podemos observar que a medida que disminuye el valor de FRADIO disminuye el valor del área, aunque a partir de un valor de FRADIO

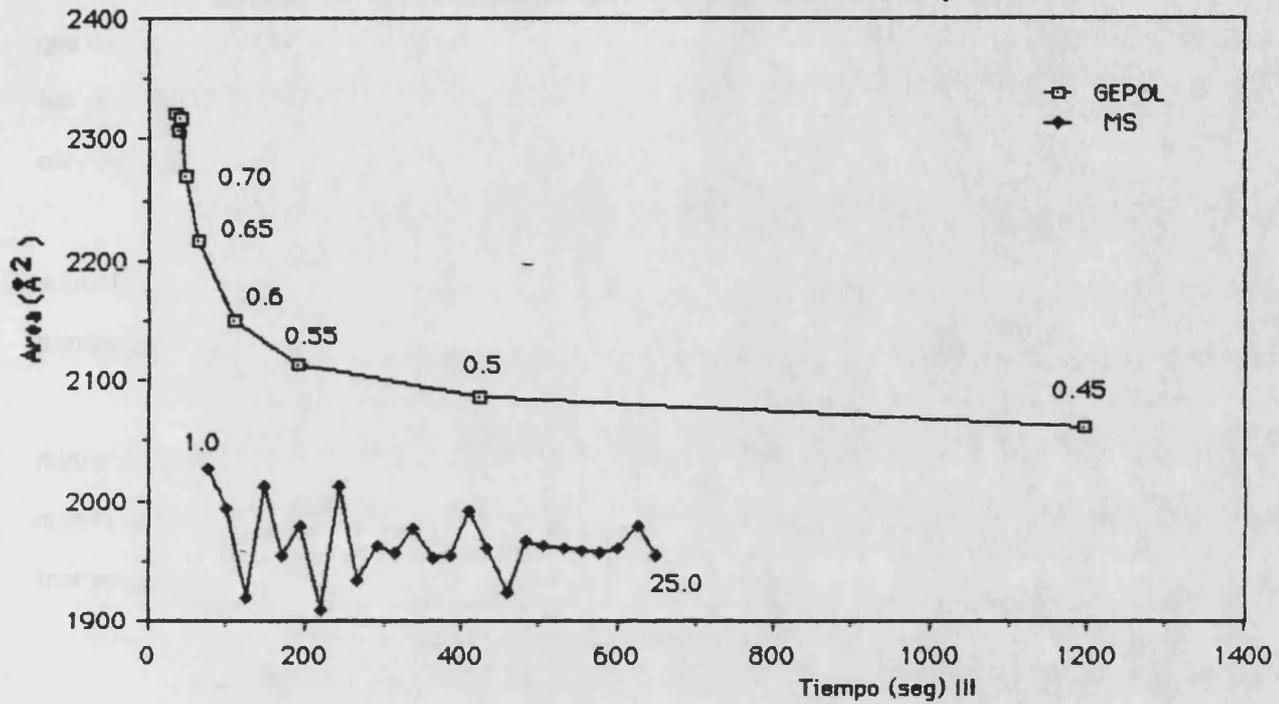
Gráfica 2.5.2.a. Variación del área con OMEGA y FRADIO en NAD y LOOP



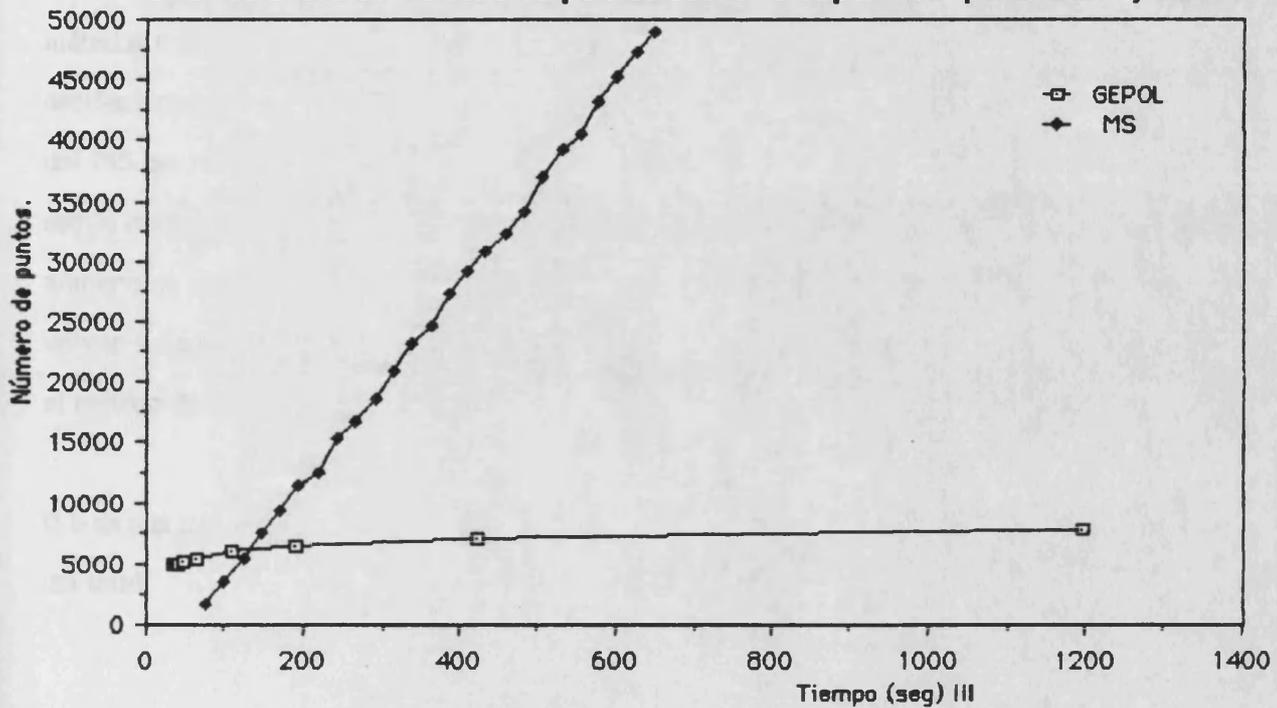
Gráfica 2.5.2.b. Variación del volumen con OMEGA y FRADIO en NAD y LOOP.



Gráfica 2.5.3.a. Area de NAD Y LOOP con GEPOL y MS



Gráfica 2.5.3.b. Número de puntos sobre la superficie para NAD y LOOP.



aproximadamente igual a 0.5, una pequeña variación del área necesita un gran tiempo de cálculo. El comportamiento oscilante del método MS es el ya observado en apartados anteriores. también hay que destacar que ambos métodos tienden a un límite diferente, la diferencia en el valor del área entre los dos métodos para los valores de DEN y FRADIO que utilizan más tiempo es de  $106.8 \text{ \AA}^2$ , que corresponde a un 5% aproximadamente de las superficies respectivas.

En la **gráfica 2.5.3 b** se representa el número de puntos generados sobre la superficie por ambos métodos en función del tiempo de CPU. Como se puede apreciar, el número de puntos generado por el MS aumenta mucho más rápidamente que en el caso del GEPOL.

En las **gráficas 2.5.3. c, d y e** se representan el volumen, el tiempo de cpu y el número de esferas creadas en función del parámetro FRADIO. Al comparar estas últimas puede observarse que el gran incremento de tiempo para valores de FRADIO  $> 0.5$ , coincide con un gran incremento en el número de esferas creadas.

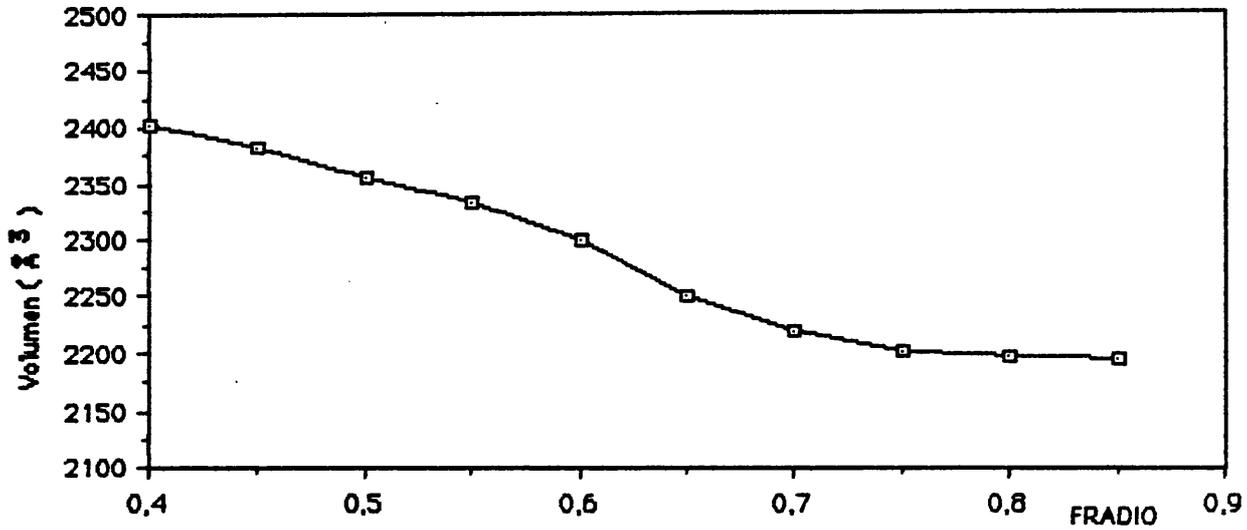
### **2.5.4 Electrón solvatado.**

Para realizar la comparación con el método MS hemos seguido la misma técnica que el apartado anterior. Así en la **gráfica 2.5.4 a** se representa los valores del área que dan ambos métodos frente al tiempo de cálculo. La variación es similar al caso anterior pero, sin embargo, las oscilaciones del método MS resultan mucho mayores. Si se recuerda como variaban las oscilaciones del MS en función del tamaño de la estructura, apartado 2.4, se puede concluir que este método oscila menos cuanto más grande sea la molécula y ello puede significar que a medida que aumenta el número de esferas, se promedian los errores dando la impresión que desaparecen. Aquí queremos volver a hacer hincapié en el concepto de área local, puesto que estos resultados parecen indicar que el método GEPOL reproduce unas áreas locales mucho más precisas que el método MS.

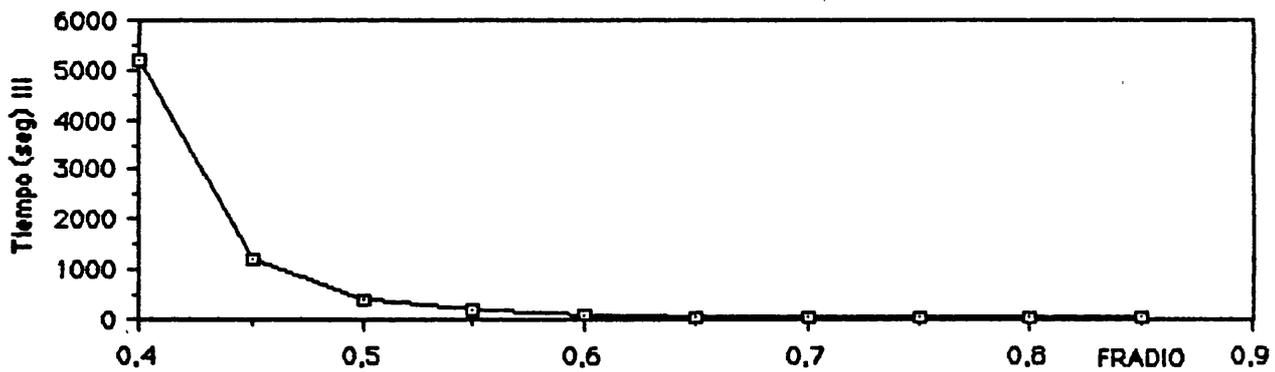
La diferencia entre las áreas calculadas con el MS y el GEPOL para los valores de 25 y 0.5 de sus parámetros respectivos es de  $15.175 \text{ \AA}^2$ , lo que corresponde a un 8.5% aproximadamente del total.

Indicar cual de los dos métodos se acerca más al valor exacto de la superficie resulta

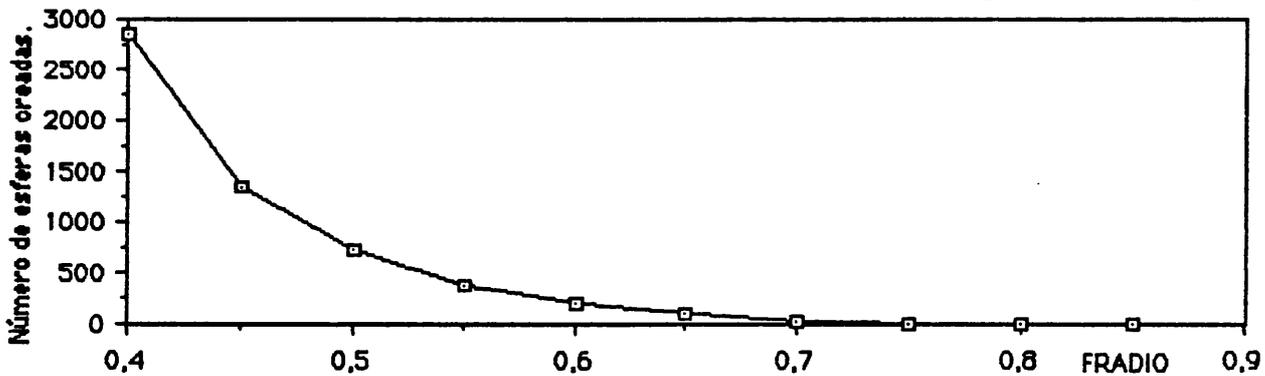
**Gráfica 2.5.3.c. Volumen del NAD y LOOP frente a FRADIO**



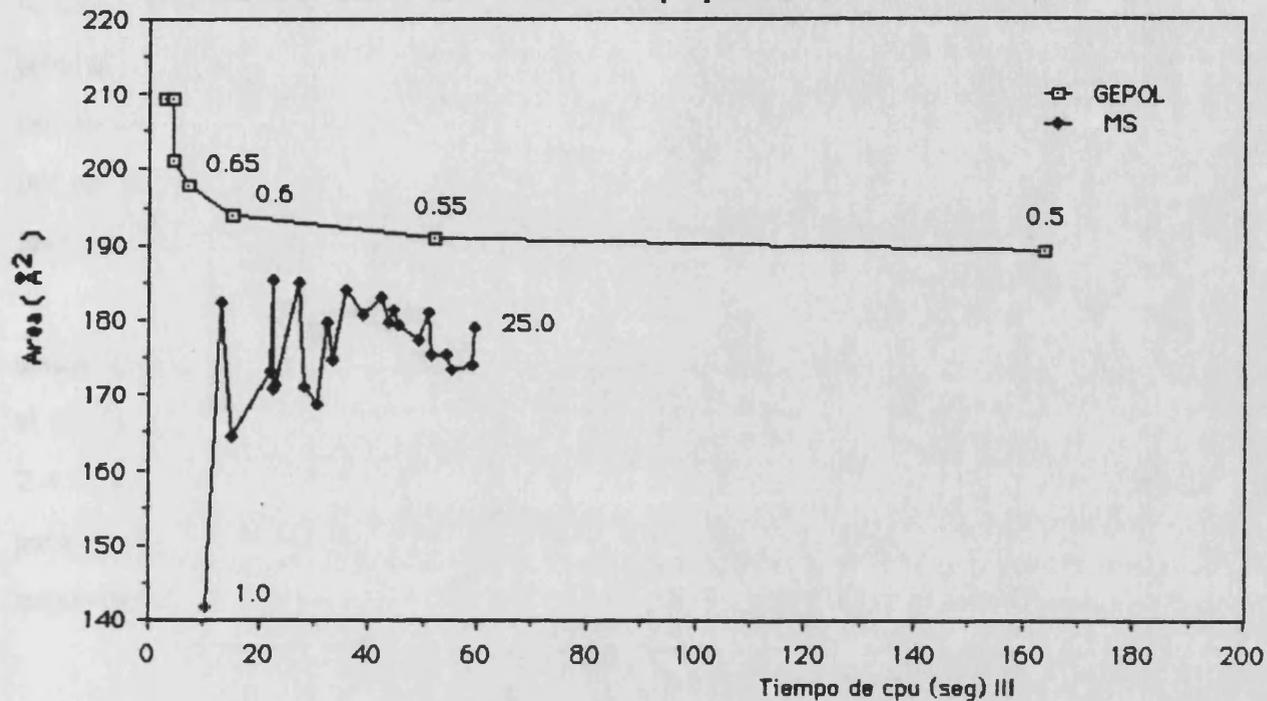
**Gráfica 2.5.3.d. Tiempo de cpu frente a FRADIO para el NAD y LOOP.**



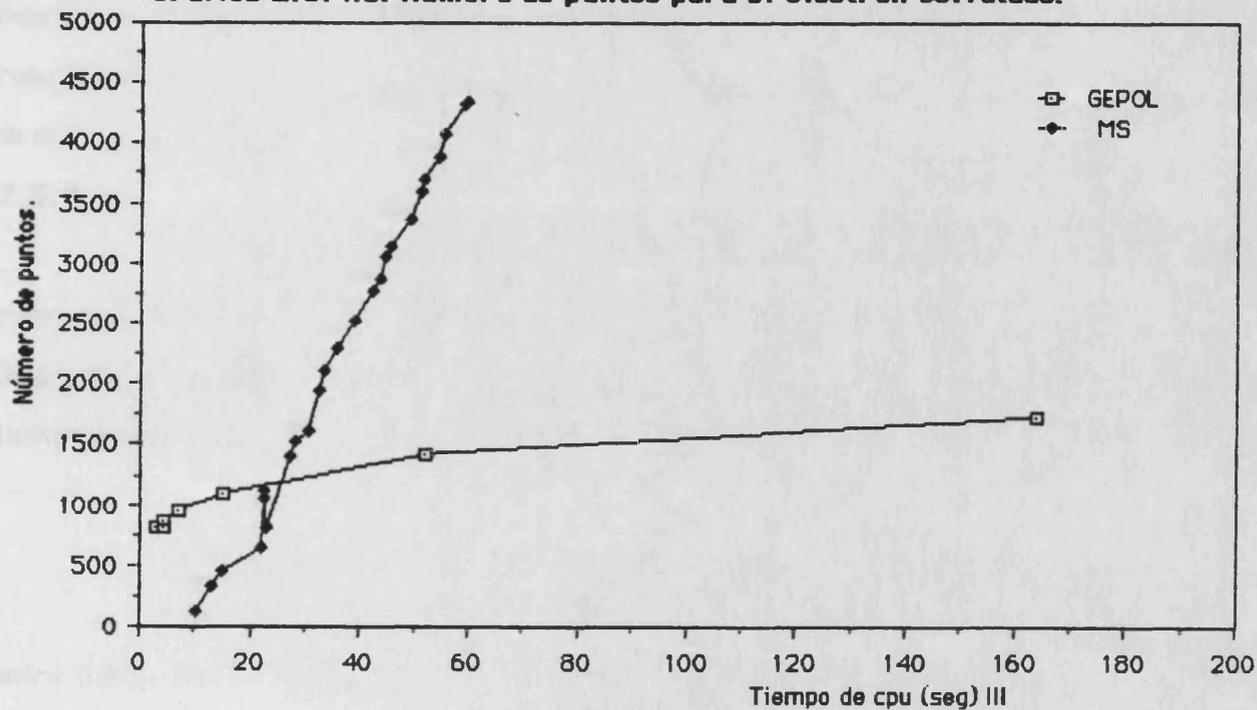
**Gráfica 2.5.3.e. Número de esferas creadas frente a FRADIO para el NAD y LOOP.**



Gráfica 2.5.4.a. Area frente a tiempo para el electrón solvatado.



Gráfica 2.5.4.b. Número de puntos para el electrón solvatado.



difícil, mas aún cuando las superficies que generan ambos métodos tienen una diferencia fundamental; la superficie que quiere calcular el método MS corresponde exactamente a la que generaría la esfera de prueba (esfera de disolvente) al rodar por encima de la molécula, produciendo partes concavas, convexas y de tipo silla. La superficie GEPOL, en cambio, está formada solamente por partes cóncavas. Por eso aunque los dos métodos calcularan exactamente sus superficies tal como las tienen definidas, el método GEPOL obtendría siempre valores mayores que el MS.

En la **gráfica 2.5.4. b** representamos los puntos que generan sobre la superficie ambos métodos frente al tiempo. Como se puede apreciar el método MS genera muchos más puntos que el GEPOL. Esto sin duda es una dificultad más del método MS, ya que, como dijimos en el apartado 2.4.7, uno de los aspectos interesantes de estos métodos es la utilización de los puntos que generan para el cálculo posterior de otras propiedades y para poder realizar esto es necesario que el número de puntos sea manejable, es decir, no muy grande.

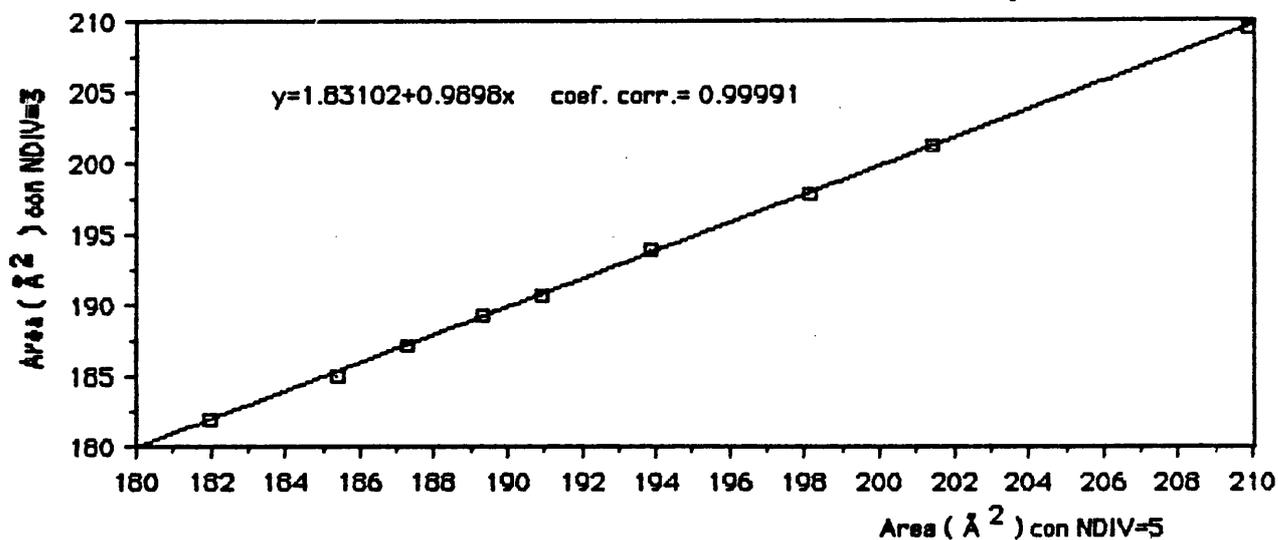
Todos los cálculos GEPOL de este apartado se han realizado con un valor de NDIV igual a 3. Para comprobar que esta opción no introduce cambios sustanciales en los resultados, hemos realizado los mismos cálculos anteriores para NDIV igual a 5. En las **gráficas 2.5.4. c y d** hemos representado los valores del área y volumen cuando se utiliza NDIV=3 frente a cuando vale 5. El resultado de unir los puntos es una recta y por lo tanto podemos decir que la utilización de un valor de NDIV de 3 es adecuada, sobre todo si observamos el ahorro de tiempo que significa (ver **gráfica 2.5.4. e**).

En la **gráfica 2.5.4. f** se representa el valor del volumen en función de FRADIO. La curva resulta continua, tendiendo a un límite para valores pequeños de FRADIO. Sin embargo como se desprende de la **gráfica 2.5.4. g**, la utilización de estos valores pequeños conlleva un gasto de tiempo grande, por lo que no parece valer la pena la utilización de valores de  $FRADIO < 0.4$ .

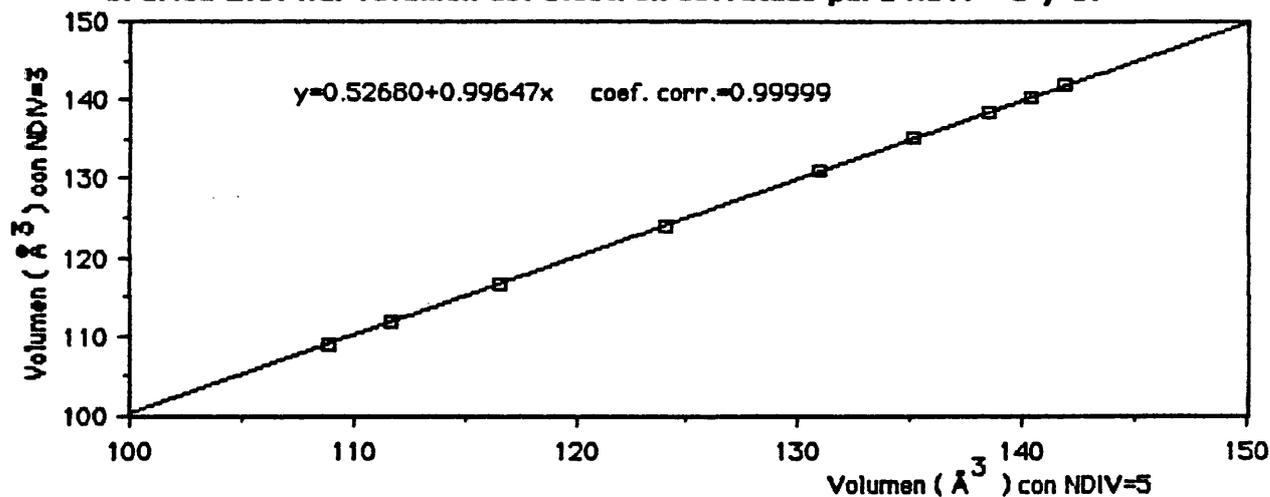
### 2.5.5 RBP.

Para esta molécula sólo hemos realizado cálculos GEPOL variando el parámetro FRADIO entre 0.85 y 0.5. En la **gráfica 2.5.5. a** se representan los valores obtenidos para el área y

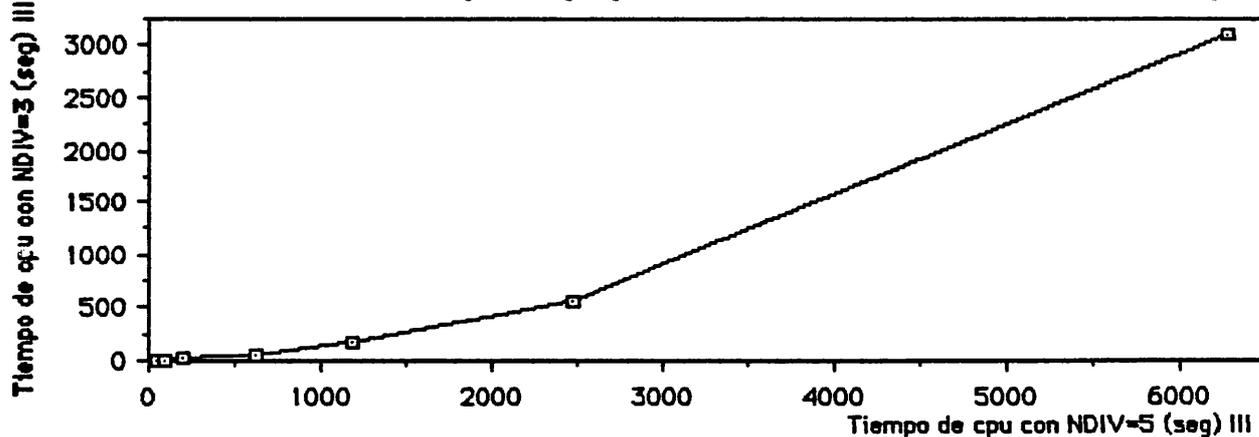
Gráfica 2.5.4.c. Area del electrón solvatado con NDI=3 y 5.

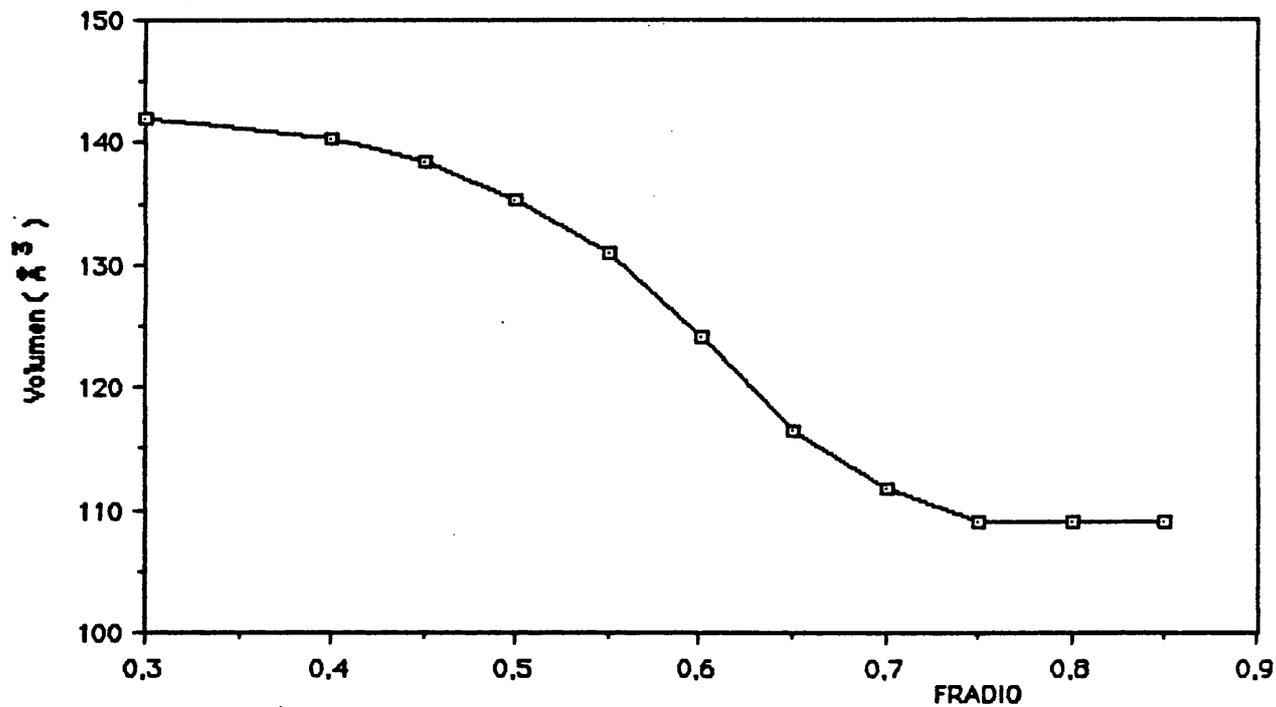
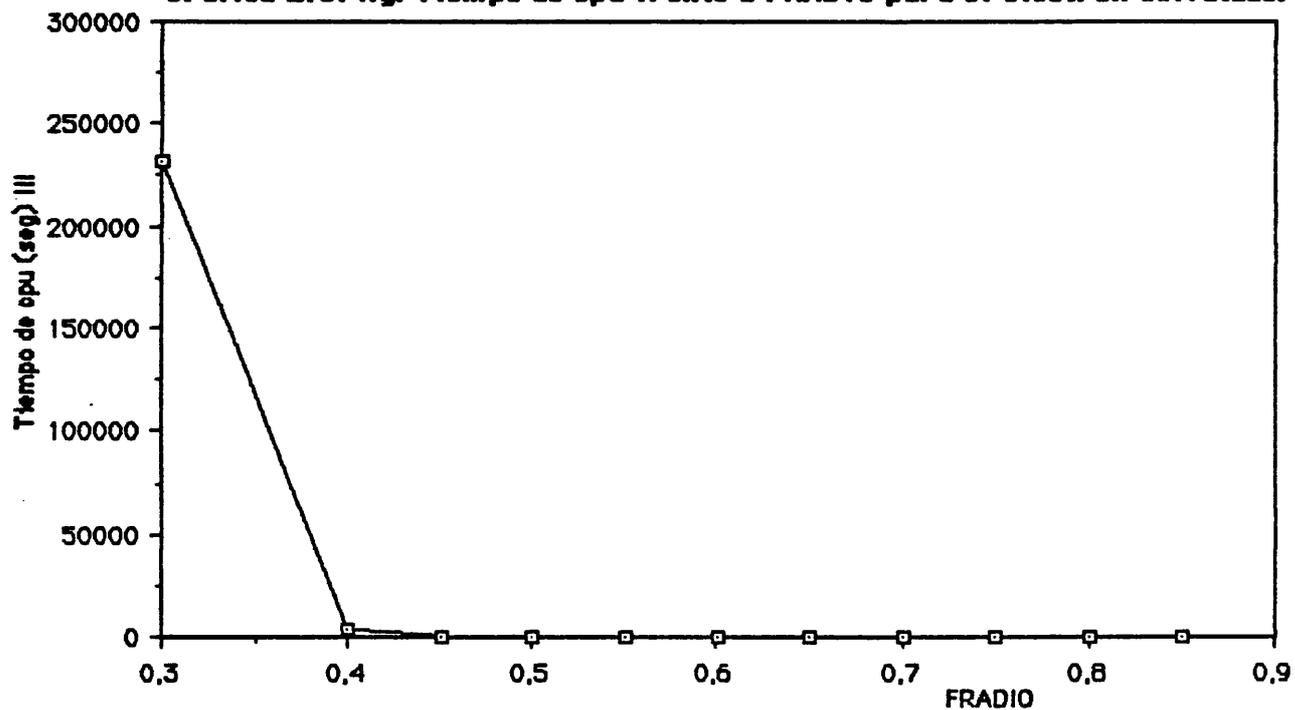


Gráfica 2.5.4.d. Volumen del electrón solvatado para NDI= 3 y 5.

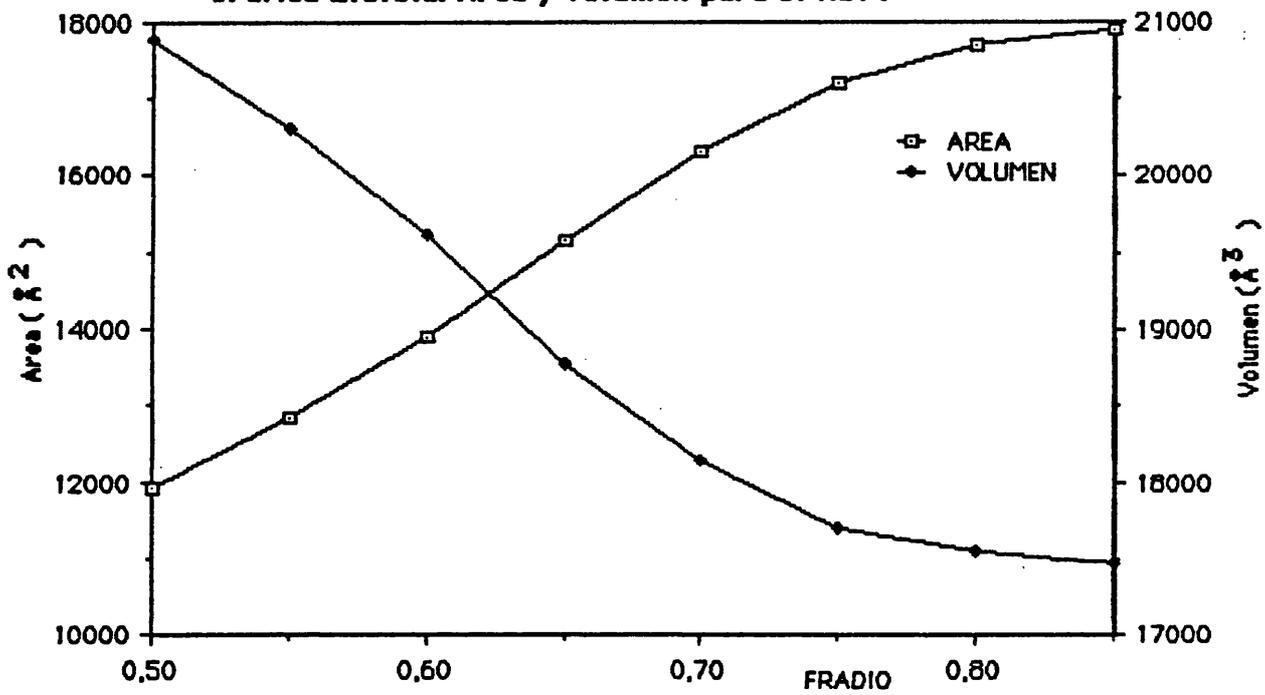


Gráfica 2.5.4.e. Tiempo de cpu para el electrón solvatado con NDI= 3 y 5.

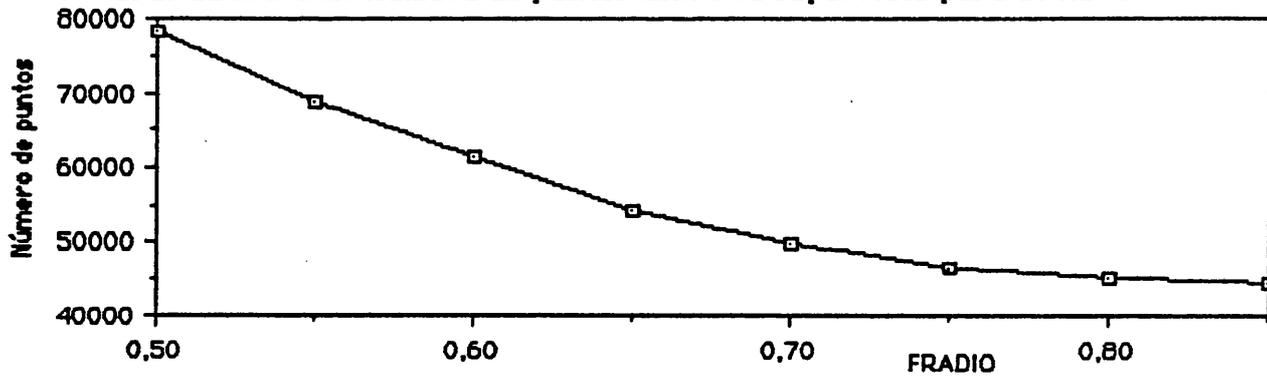


**Gráfica 2.5.4.f. Volumen del electrón solvatado frente a FRADIO.****Gráfica 2.5.4.g. Tiempo de cpu frente a FRADIO para el electrón solvatado.**

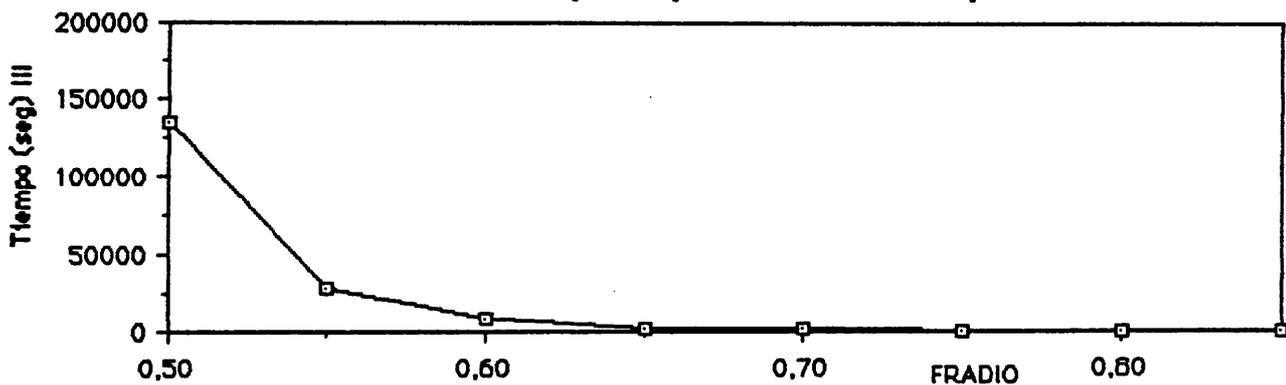
Gráfica 2.5.5.a. Area y volumen para el RBP.



Gráfica 2.5.5.b. Número de puntos sobre la superficie para el RBP.



Gráfica 2.5.5.c Tiempo de cpu frente a FRADIO para el RBP.



volumen en función de FRADIO. En ella se ve un comportamiento similar a los casos anteriores.

En las **gráficas 2.5.5. b y c** se representan el número de puntos de la superficie y tiempo de cpu frente a los valores de FRADIO utilizados, no observándose tampoco grandes diferencias respecto a lo obtenido en casos anteriores.

## **2.6 EJEMPLOS DE APLICACIONES.**

Hasta aquí hemos analizado como depende el resultado de GEPOL de los parámetros utilizados. En este apartado mostraremos el comportamiento del método en un conjunto de aplicaciones, las cuales no se analizarán en profundidad, sino que únicamente nos servirán para ilustrar las posibilidades del método.

La interacción de una molécula con sus alrededores depende, como indicamos en el capítulo 1 y como mostraremos en el capítulo 3, de la superficie que expone al exterior. Por otro lado, esta superficie no es estática, sino que depende de la disposición geométrica de los átomos. Por ello iniciaremos el apartado analizando la variación de la superficie en una disociación y una rotación interna. A continuación compararemos los resultados obtenidos para el volumen con el volumen molar en una serie de aminas.

El estudio de la fractalidad de una biomolécula es una de las aplicaciones más recientes del cálculo de superficies. De ella se puede deducir la capacidad reactiva global de la biomolécula, así como qué partes de ella pueden ser más reactivas. En el cuarto apartado calcularemos la fractalidad global de una molécula de este tipo.

### **2.6.1 Disociación de una molécula diatómica.**

Hasta ahora hemos hecho referencia a dos definiciones de superficie, la superficie molecular de van der Waals (SMW) y la superficie molecular (SM). Y, como dijimos en el capítulo 1, la segunda de ellas da una representación más realista de la superficie de la molécula. De hecho, en muchos casos el comportamiento de una u otra superficie no solo se diferencia cuantitativa sino cualitativamente. Uno de estos casos es el que nos ocupa en este subapartado.

Para representar una disociación molecular hemos elegido dos esferas de 2.4 Å de radio que parten de una distancia entre sus centros de 1.5 Å. El disolvente está representado por una esfera con un radio de 1.5 Å.

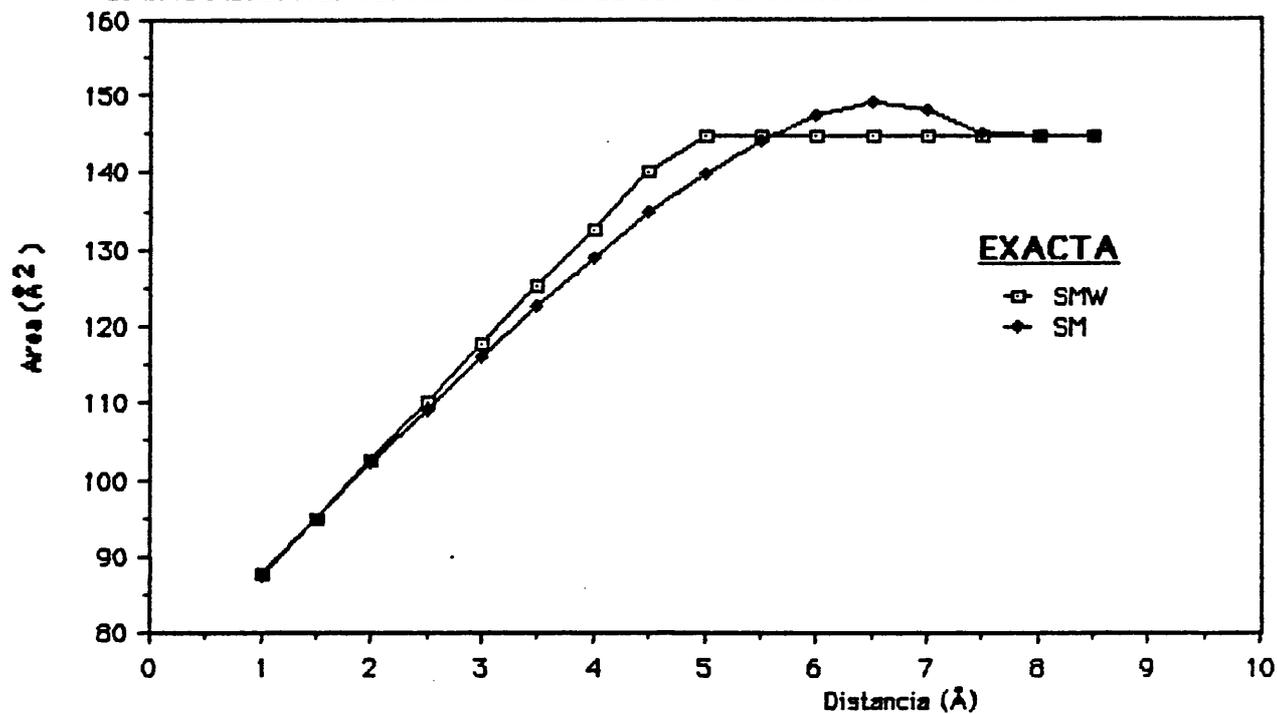
Para mostrar el comportamiento de ambas superficies y poder analizar los resultados obtenidos con nuestro método, se calculó el valor exacto del área y volumen (J.C.Gavilan, 1987) para todo el proceso de disociación, representándose estos resultados en las **gráficas 2.6.1.a y b** respectivamente. Como se puede ver, el área correspondiente a la SM es menor que la SMW hasta una distancia de aproximadamente  $5.5 \text{ \AA}$ , donde toma valores mayores hasta alcanzar un máximo y posteriormente decrece hasta coincidir de nuevo con el valor del área de la SMW. Esta diferencia de comportamiento es muy importante ya que la interacción con el disolvente se da a través de la superficie y por tanto la variación de ésta se reflejará en la interacción. El volumen de la SM también tiene un máximo, pero sus valores son siempre mayores que los correspondientes a la SMW.

En las **gráficas 2.6.1.c y d** se ha representado el valor del área y volumen obtenido con el método GEPOL para las dos superficies. En estos cálculos se ha utilizado un valor para NDIV de 5, y un FRADIO de 0,2. Como se puede ver de la comparación de las anteriores gráficas, los resultados son excelentes a la hora de describir el comportamiento diferente de ambas superficies. Los errores relativos referidos a los valores exactos no superan el 1.5%.

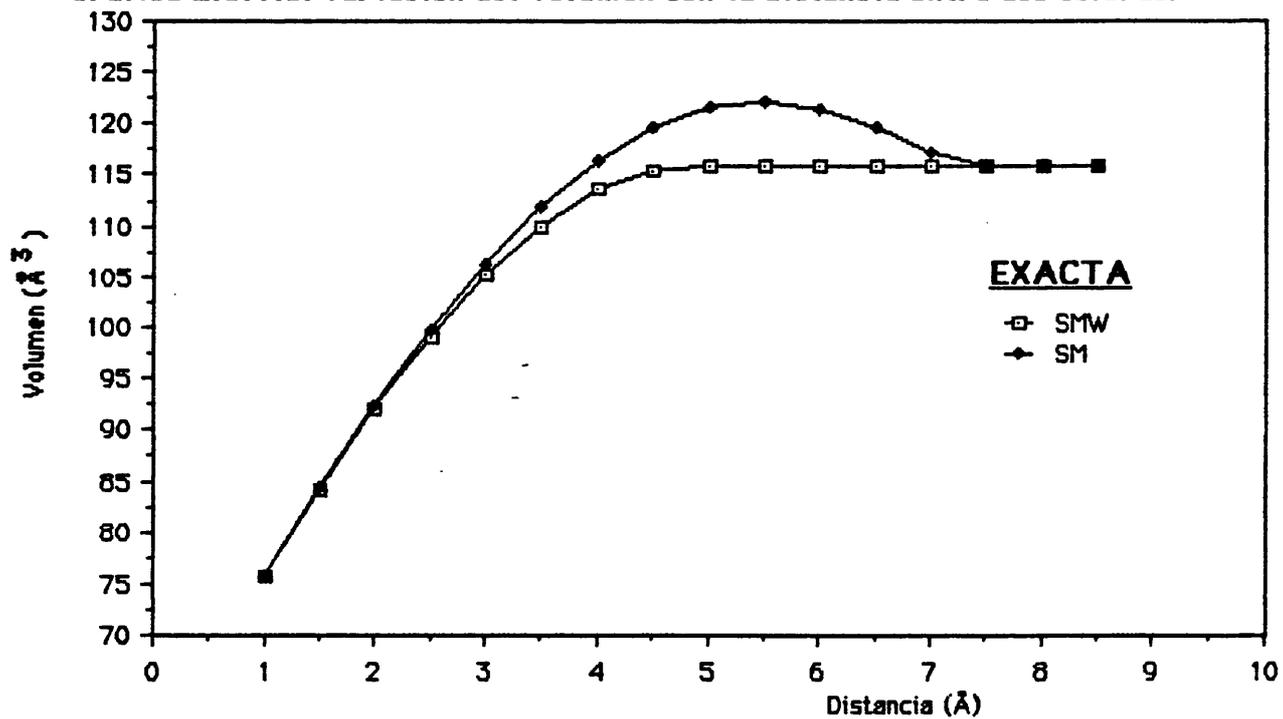
Estas diferencias se pueden observar mejor en las **gráficas 2.6.1. e y f**, donde hemos representado respectivamente la SMW y la SM, para los valores exactos, para los obtenidos con el GEPOL y para los obtenidos con el método MS. En el caso de la SMW los diferencias entre el valor exacto y el GEPOL son casi nulas, mientras sí se observan desviaciones en el caso del MS. Para la SM nuestro método siempre sobreestima el área, aunque mantiene una variación muy similar a la exacta. Mientras el método MS, aun dando una descripción similar a los otros dos, oscila por encima y por debajo del valor exacto.

Que nuestro método sobreestime el valor del área y por otra lado subestime el valor del volumen es un comportamiento esperable, dado el tipo de método que hemos utilizado para representar la SM. En la **figura 2.6.1 a** se muestra el proceso de disociación de las dos esferas y cómo nuestro algoritmo de suavización crea nuevas esferas que simulan la superficie exacta, representada en la **figura 2.6.1 b**. De ellas se deduce claramente que la superficie formada por un conjunto de esferas, y por tanto cóncava en todas sus partes, siempre será mayor que una

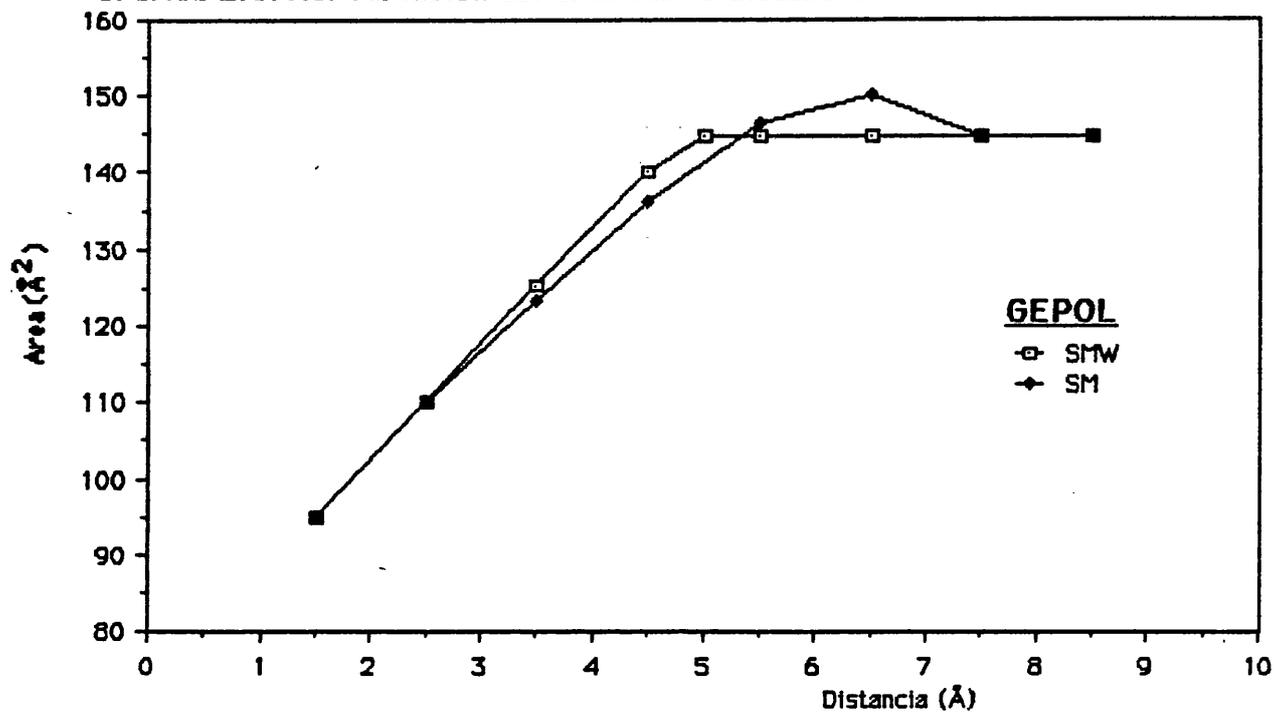
Gráfica 2.6.1.a. Variación del área con la distancia entre dos esferas



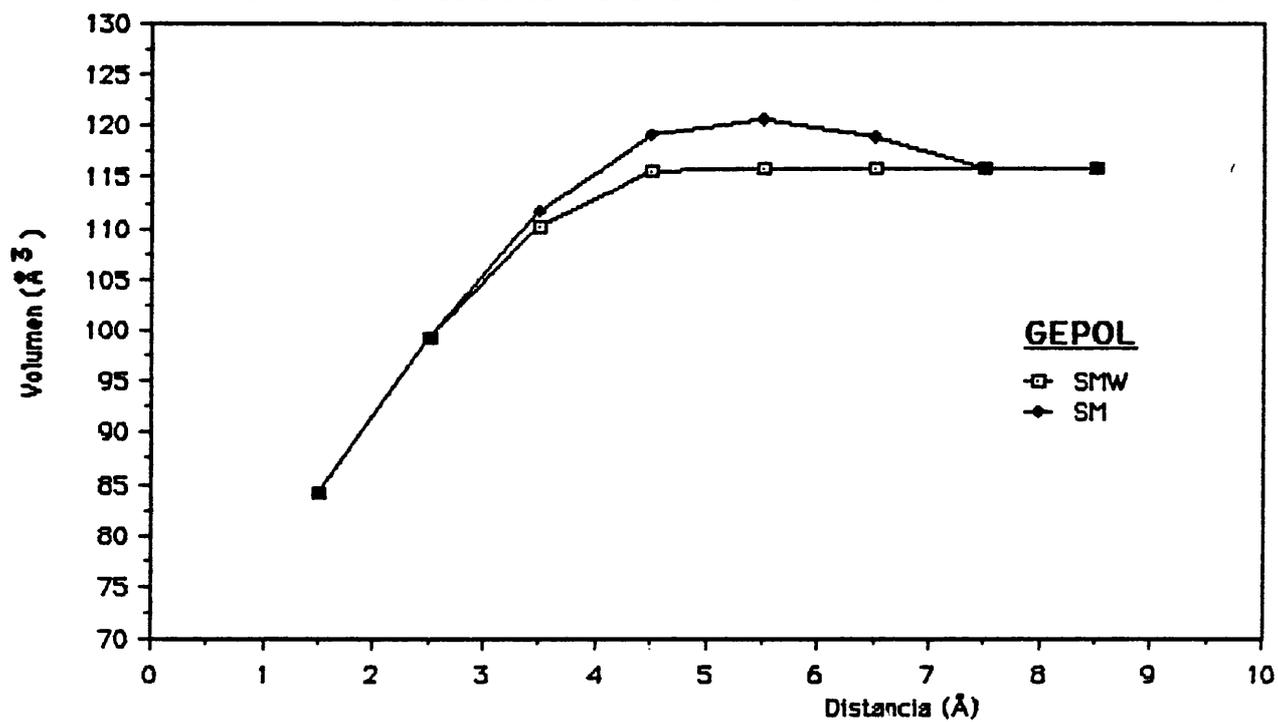
Gráfica 2.6.1.b. Variación del volumen con la distancia entre dos esferas.



Gráfica 2.6.1.c. Variación del área con la distancia entre dos esferas

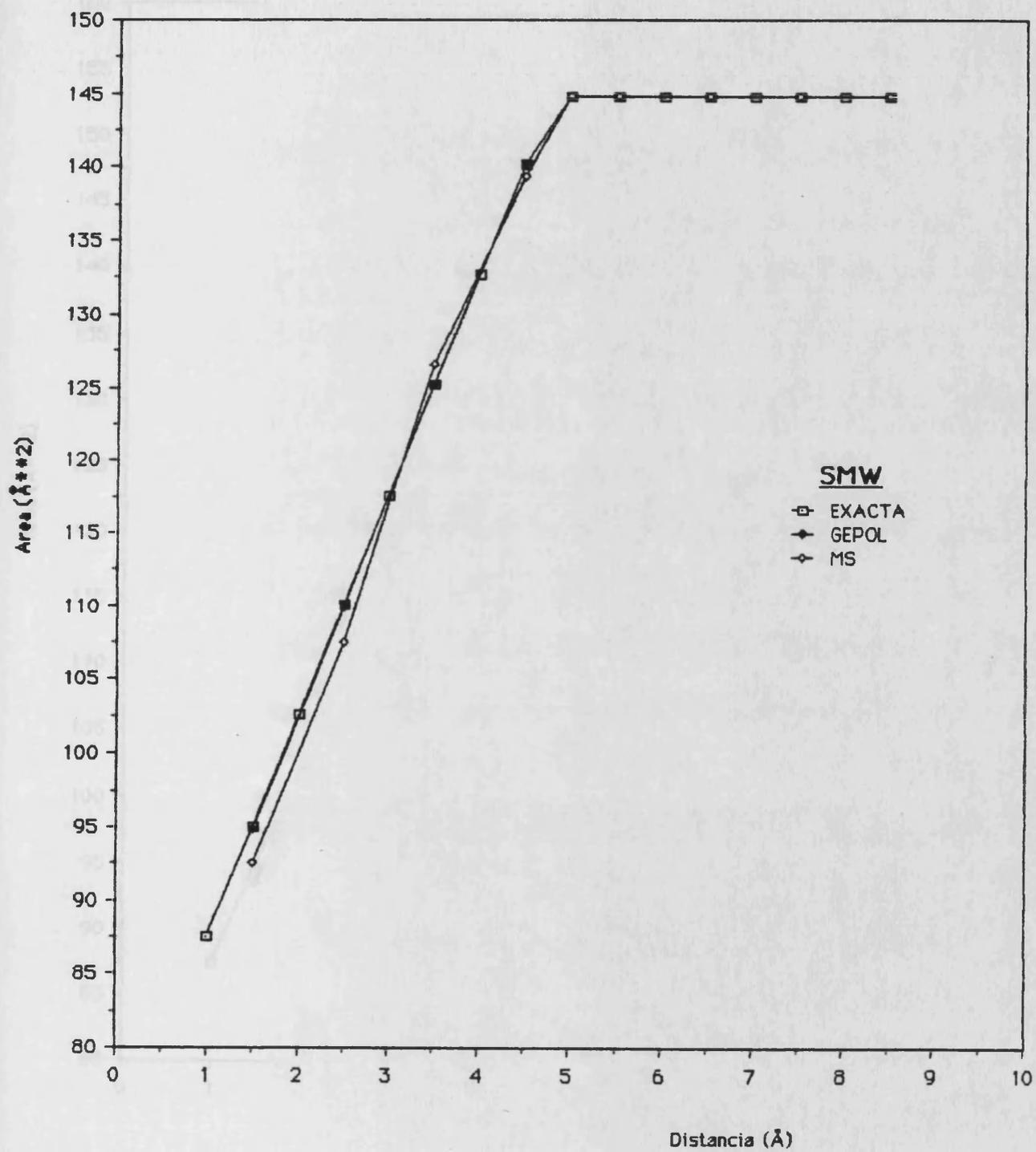


Gráfica 2.6.1.d. Variación del volumen con la distancia entre dos esferas.

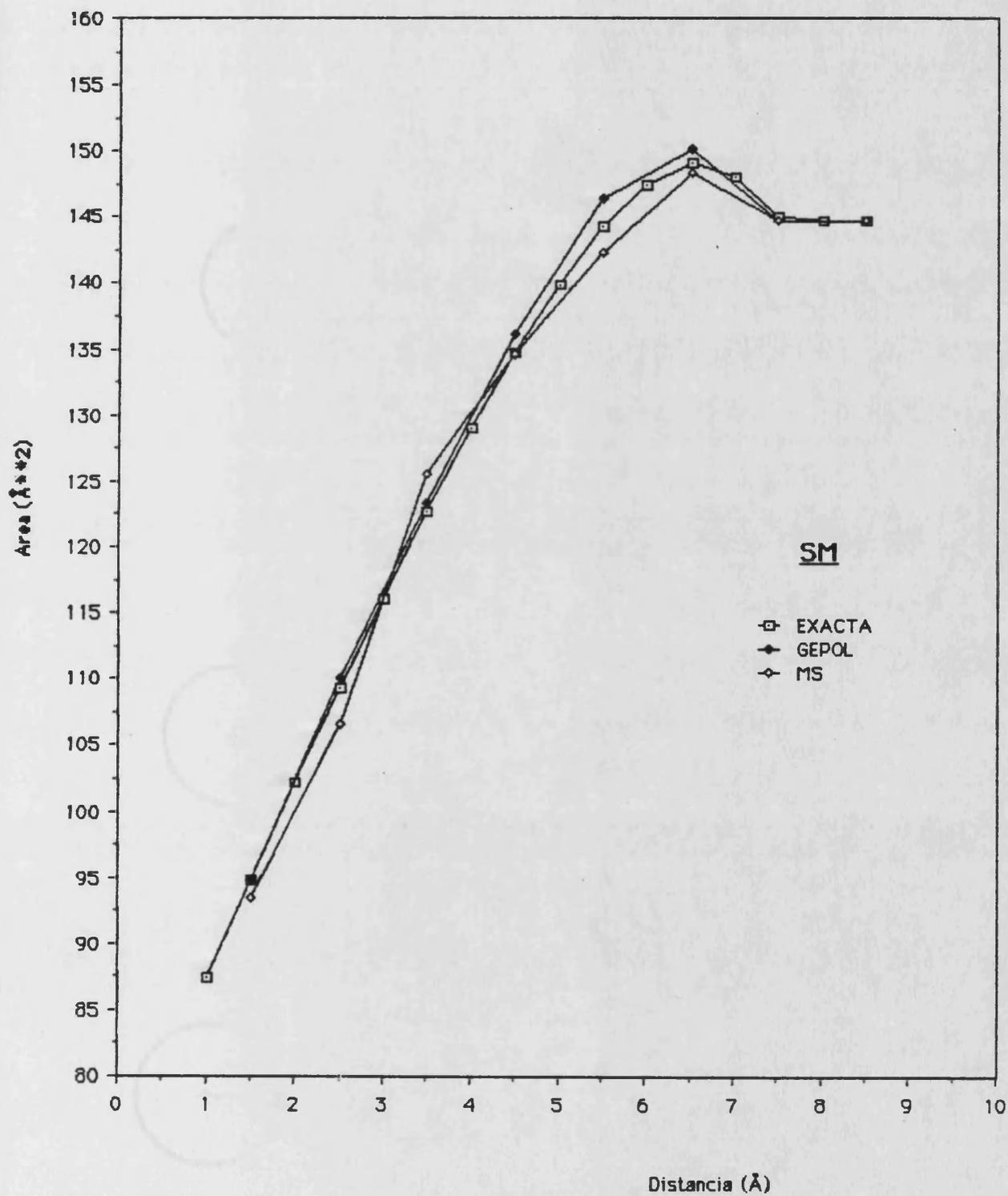


Gráfica 2.6.

Gráfica 2.6.1.e. Variación del área con la distancia entre dos esferas.



Gráfica 2.6.1.f. Variación del área con la distancia entre dos esferas.



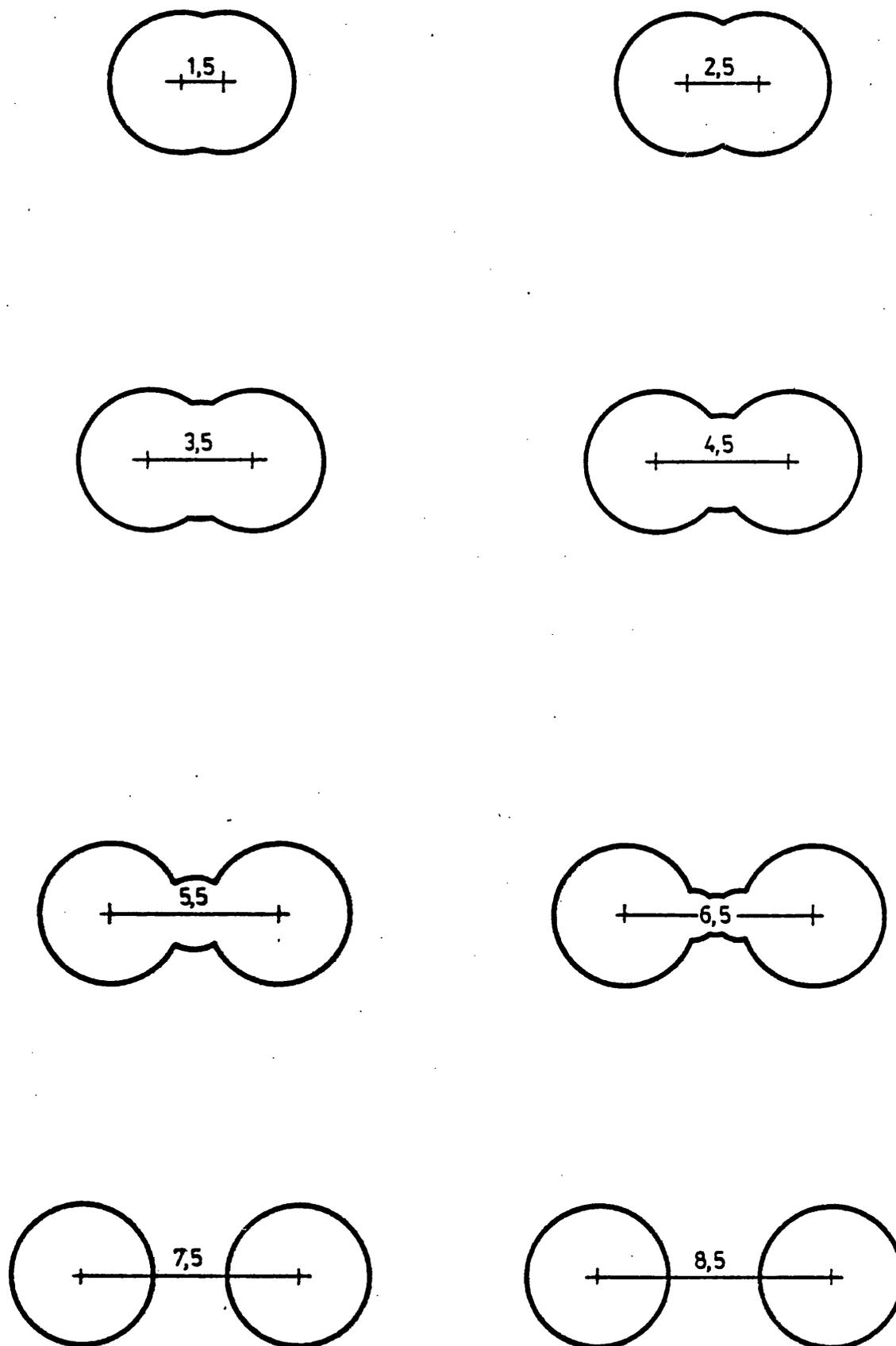


Figura 2.6.1.a. Disociacion de dos esferas con GEPOL.

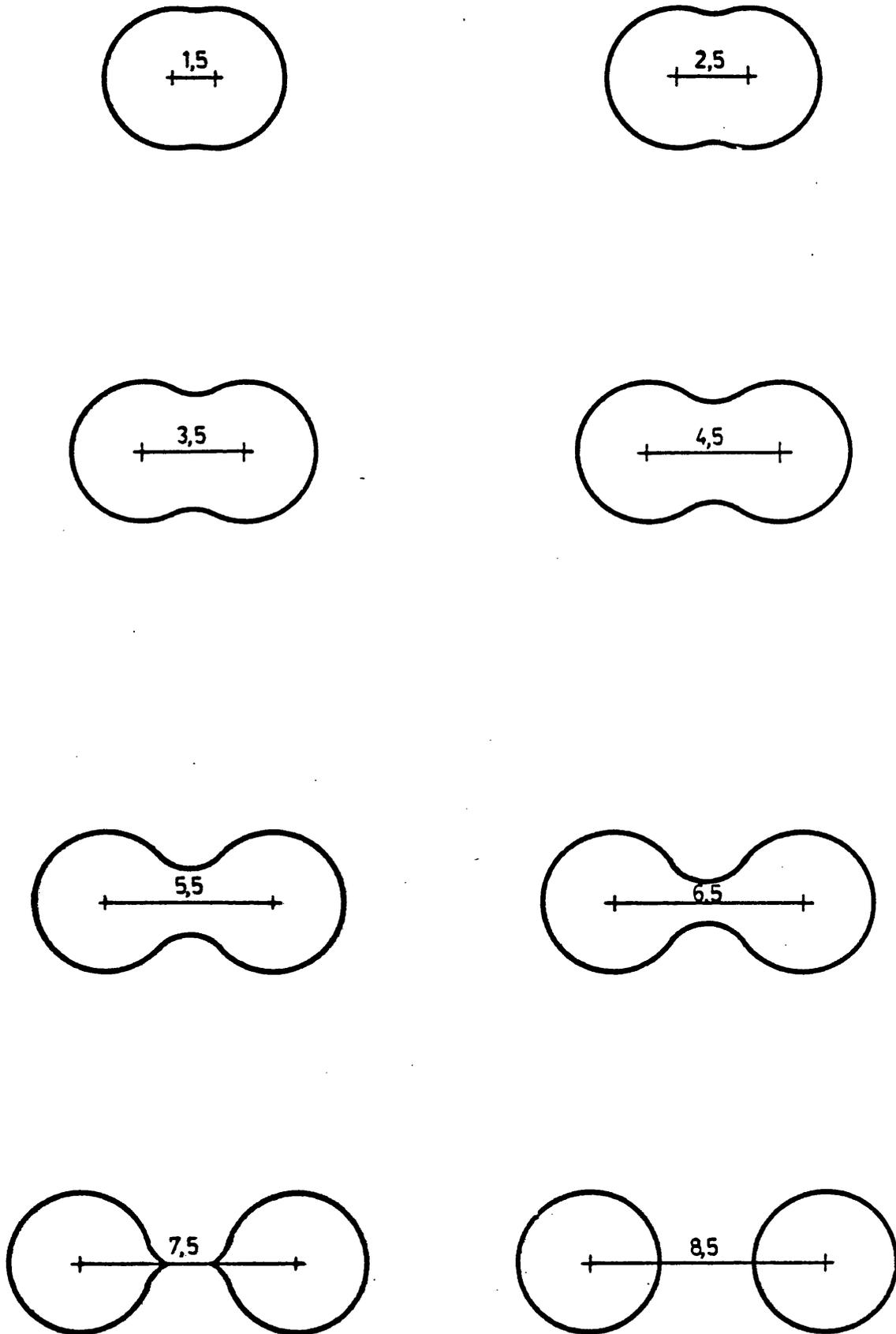
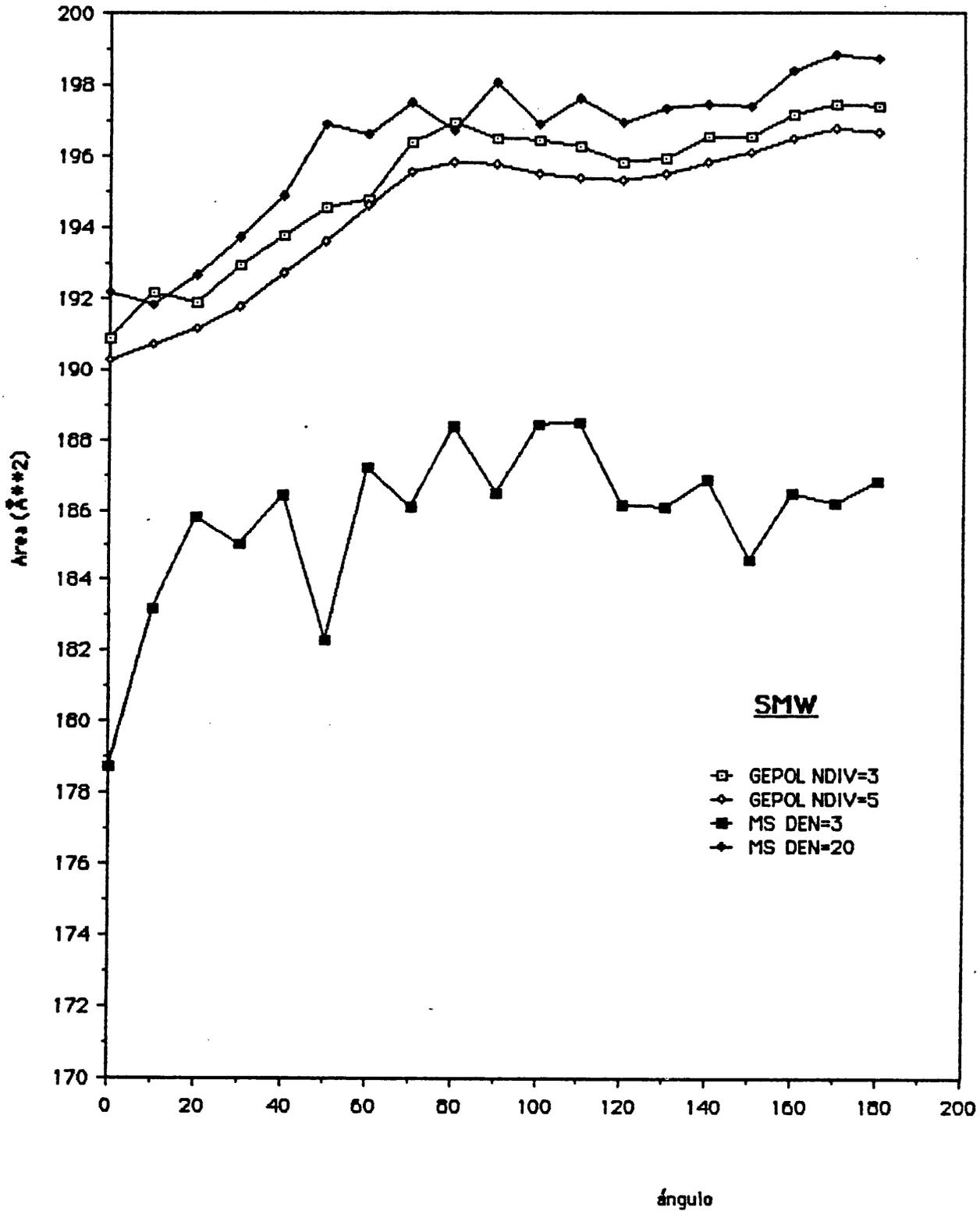


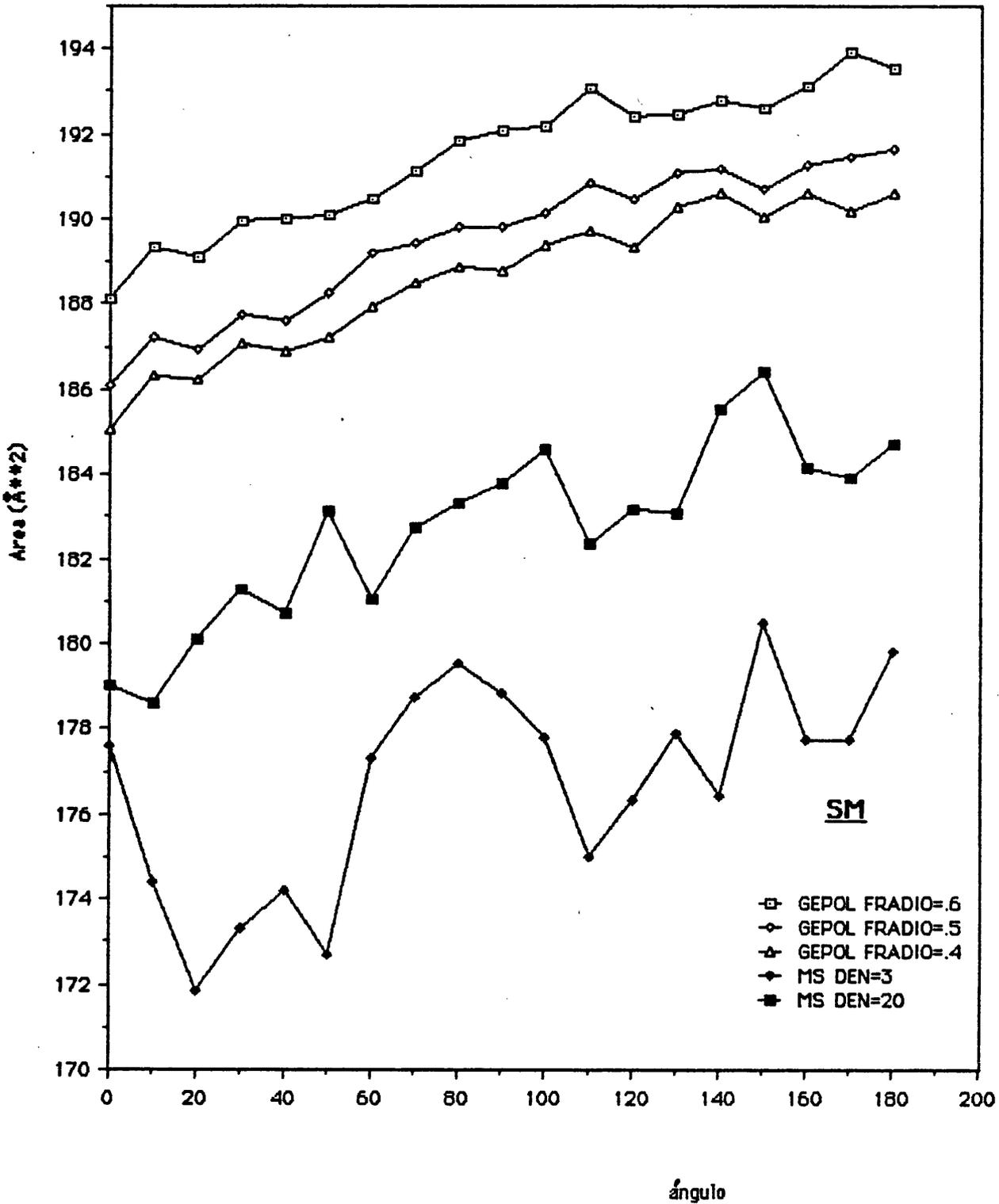
Figura 2.6.1.b Disociación exacta de dos esferas.



Gráfica 2.6.2.a. Variación del área en la rotación interna del octnoI.



Gráfica 2.6.2.b. Variación del área en la rotación interna del octanol.



comparación con las del MS), todas dan un perfil similar. En cuanto a la diferencia entre los valores absolutos de estas curvas, se puede ver como disminuye el valor del área a medida que disminuye el valor de FRADIO y como esta disminución cada vez es menor, indicando una convergencia hacia un valor límite.

Con el fin de determinar la diferencia entre la SMW y SM hemos representado en las **gráficas 2.6.2 c y d**, para ellas dos, el valor del área y del volumen respectivamente (SMW NDIV=5, SM NDIV=5 y FRADIO= 0.5). Obsérvese que las escalas de ordenadas se han ampliado respecto de los casos anteriores y, sin embargo, la continuidad de las curvas es perfecta. En la gráfica que se representa el área se puede observar como al pasar de la SMW a la SM se produce una especie de suavización de la curva perdiéndose un máximo que se encontraba aproximadamente a  $TETA=80^\circ$  y, por tanto, la molécula no va a tener el mismo comportamiento frente a los alrededores si se considera una superficie u otra. El volumen tiene un comportamiento inverso al del área, como se puede ver en la gráfica el volumen correspondiente a la SM es mayor que el de la SMW, recalándose además dos máximos, que en la SMW eran muy pequeños.

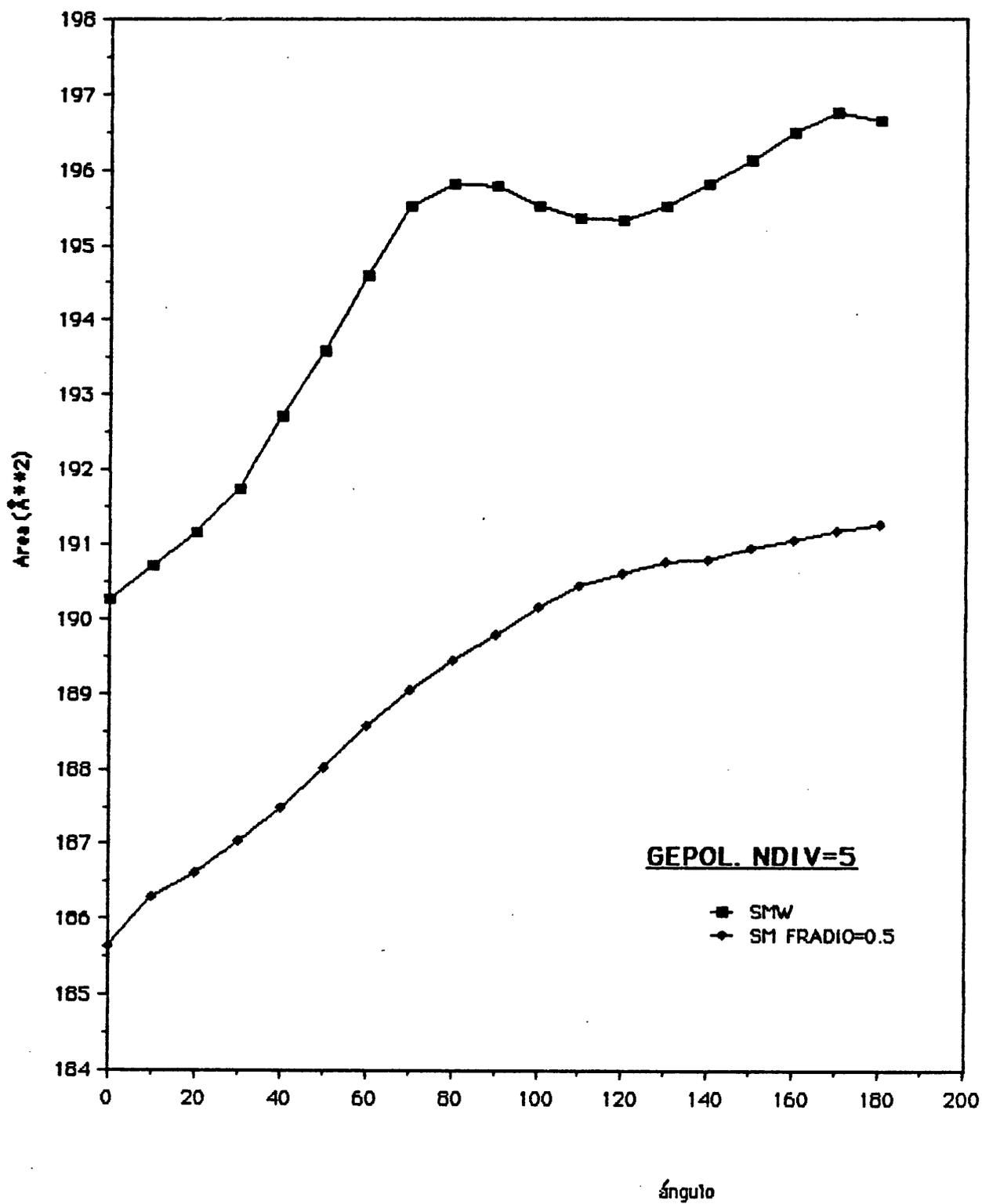
### **2.6.3 Volumen de las aminas**

Para determinar la validez de los valores del volumen obtenido con el método GEPOL, y a falta de poder compararlos con otro método, hemos hecho una serie de cálculos sobre un conjunto de aminas, comparando los volúmenes obtenidos con los volúmenes molares parciales en disolución acuosa correspondientes, aun siendo conscientes de que el volumen molar parcial es un valor termodinámico sin una correspondencia directa en el mundo microscópico.

Las aminas utilizadas han sido el amoníaco, la metilamina, la dimetilamina y la trimetilamina y las correspondientes especies protonadas. Las geometrías fueron obtenidas realizando una optimización con una base 4-31G (ver nota al final del subapartado). Los datos geométricos así como los radios de las esferas se dan en la **tabla 2.6.3. a**.

Los cálculos se realizaron sin y con suavización de la superficie, es decir, obtuvimos la

Gráfica 2.6.2.c. Variación del área en la rotación interna del octanol.



Gráfica 2.6.2.d. Variación del volumen en la rotación interna del octanol.

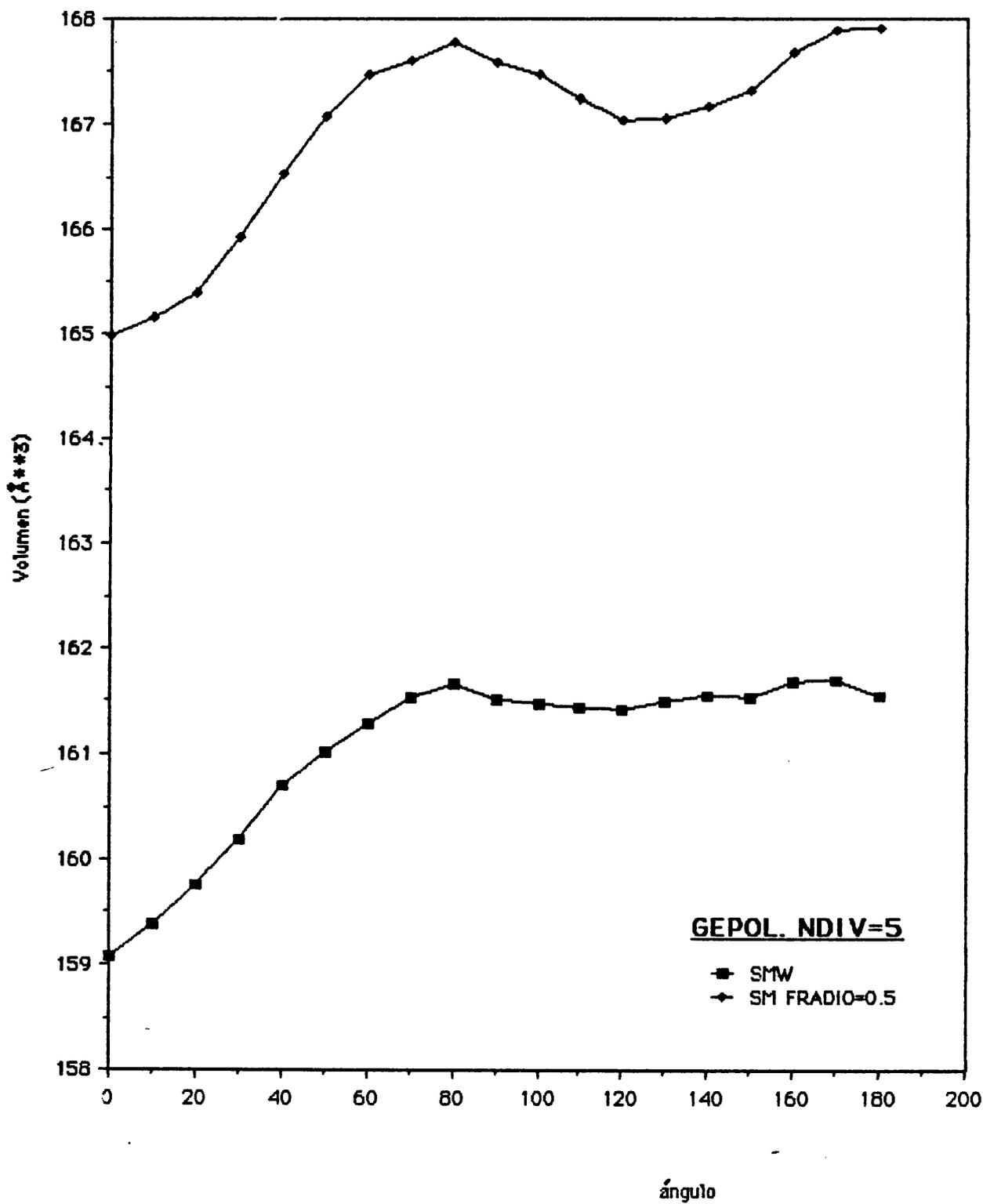


Tabla 2.6.3.a. Datos geométricos de las Aminas en R

=====				
NH3				
natom	dist.	fi	teta	radio
0				
0	1 1.00			
7	2 1.00	1 90.00		1.86
1	3 0.9912	2 101.9616	1 0.0	1.44
1	3 0.9912	2 101.9616	1 120.0	1.44
1	3 0.9912	2 101.9616	1 240.0	1.44
=====				
+NH4				
natom	dist.	fi	teta	radio
0				
1	1 1.00			1.44
7	2 1.0116	1 90.00		1.86
1	3 1.0116	2 109.4714	1 0.0	1.44
1	3 1.0116	2 109.4714	1 120.0	1.44
1	3 1.0116	2 109.4714	1 240.0	1.44
=====				
NH2-Me				
natom	dist.	fi	teta	radio
0				
0	1 1.00			
7	2 1.00	1 90.00		1.86
6	3 1.4506	2 97.6046	1 0.0	1.92
1	3 0.9931	2 104.9381	1 120.0	1.44
1	3 0.9931	2 104.9381	1 240.0	1.44
1	4 1.0838	3 111.1926	2 60.0	1.44
1	4 1.0838	3 111.1926	2 180.0	1.44
1	4 1.0838	3 111.1926	2 300.0	1.44
=====				
+NH3-Me				
natom	dist.	fi	teta	radio
0				
1	1 1.00			1.44
7	2 1.0090	1 90.00		1.86
6	3 1.5306	2 109.8739	1 0.0	1.92
1	3 1.0090	2 108.5524	1 120.0	1.44
1	3 1.0090	2 108.5524	1 240.0	1.44
1	4 1.0756	3 108.2093	2 60.0	1.44
1	4 1.0756	3 108.2093	2 180.0	1.44
1	4 1.0756	3 108.2093	2 300.0	1.44
=====				
NH-2Me				
natom	dist.	fi	teta	radio
0				
0	1 1.00			
7	2 1.00	1 90.00		1.86
6	3 1.4511	2 104.0043	1 0.0	1.92
6	3 1.4511	2 104.0043	1 120.0	1.92
1	3 0.9947	2 102.8127	1 240.0	1.44
1	4 1.0848	3 111.0241	2 60.0	1.44
1	4 1.0848	3 111.0241	2 180.0	1.44
1	4 1.0848	3 111.0241	2 300.0	1.44
1	5 1.0848	3 111.0241	2 60.0	1.44
1	5 1.0848	3 111.0241	2 180.0	1.44
1	5 1.0848	3 111.0241	2 300.0	1.44
=====				

=====				
+NH2-2Me				
natom	dist.	fi	teta	radio
0				
1	1 1.00			1.44
7	2 1.0078	1 90.00		1.86
6	3 1.5173	2 108.2603	1 0.0	1.92
6	3 1.5173	2 108.2603	1 120.0	1.92
1	3 1.0078	2 108.1139	1 240.0	1.44
1	4 1.0764	3 108.5886	2 60.0	1.44
1	4 1.0764	3 108.5886	2 180.0	1.44
1	4 1.0764	3 108.5886	2 300.0	1.44
1	5 1.0764	3 108.5886	2 60.0	1.44
1	5 1.0764	3 108.5886	2 180.0	1.44
1	5 1.0764	3 108.5886	2 300.0	1.44
=====				
N-3Me				
natom	dist.	fi	teta	radio
0				
0	1 1.00			
7	2 1.00	1 90.00		1.86
6	3 1.4511	2 104.6101	1 0.0	1.92
6	3 1.4511	2 104.6101	1 120.0	1.92
6	3 1.4511	2 104.6101	1 240.0	1.92
1	4 1.0854	3 110.8362	2 60.0	1.44
1	4 1.0854	3 110.8362	2 180.0	1.44
1	4 1.0854	3 110.8362	2 300.0	1.44
1	5 1.0854	3 110.8362	2 60.0	1.44
1	5 1.0854	3 110.8362	2 180.0	1.44
1	5 1.0854	3 110.8362	2 300.0	1.44
1	6 1.0854	3 110.8362	2 60.0	1.44
1	6 1.0854	3 110.8362	2 180.0	1.44
1	6 1.0854	3 110.8362	2 300.0	1.44
=====				
+NH-3Me				
natom	dist.	fi	teta	radio
0				
1	1 1.00			1.44
7	2 1.0080	1 90.00		1.86
6	3 1.5086	2 107.3974	1 0.0	1.92
6	3 1.5086	2 107.3974	1 120.0	1.92
6	3 1.5086	2 107.3974	1 240.0	1.92
1	4 1.0770	3 108.8325	2 60.0	1.44
1	4 1.0770	3 108.8325	2 180.0	1.44
1	4 1.0770	3 108.8325	2 300.0	1.44
1	5 1.0770	3 108.8325	2 60.0	1.44
1	5 1.0770	3 108.8325	2 180.0	1.44
1	5 1.0770	3 108.8325	2 300.0	1.44
1	6 1.0770	3 108.8325	2 60.0	1.44
1	6 1.0770	3 108.8325	2 180.0	1.44
1	6 1.0770	3 108.8325	2 300.0	1.44
=====				

SMW y SM. En la **tabla 2.6.3. b** se dan los volúmenes con ambas superficies para las cuatro aminas neutras y protonadas. Dado el tamaño y forma de estas moléculas es predecible que no hayan cambios muy grandes de una superficie a otra. De hecho en las dos primeras aminas (tanto neutras como protonadas) no hay variación.

**Tabla 2.6.3.b.** Volúmenes experimentales y teóricos en  $\text{cm}^3/\text{mol}$  para un conjunto de aminas neutras y protonadas. Los volúmenes experimentales corresponden a volúmenes molares parciales. Aminas neutras (Terry, B. y col, 1981). Aminas cargadas (Aue, D.H. y col, 1976b).

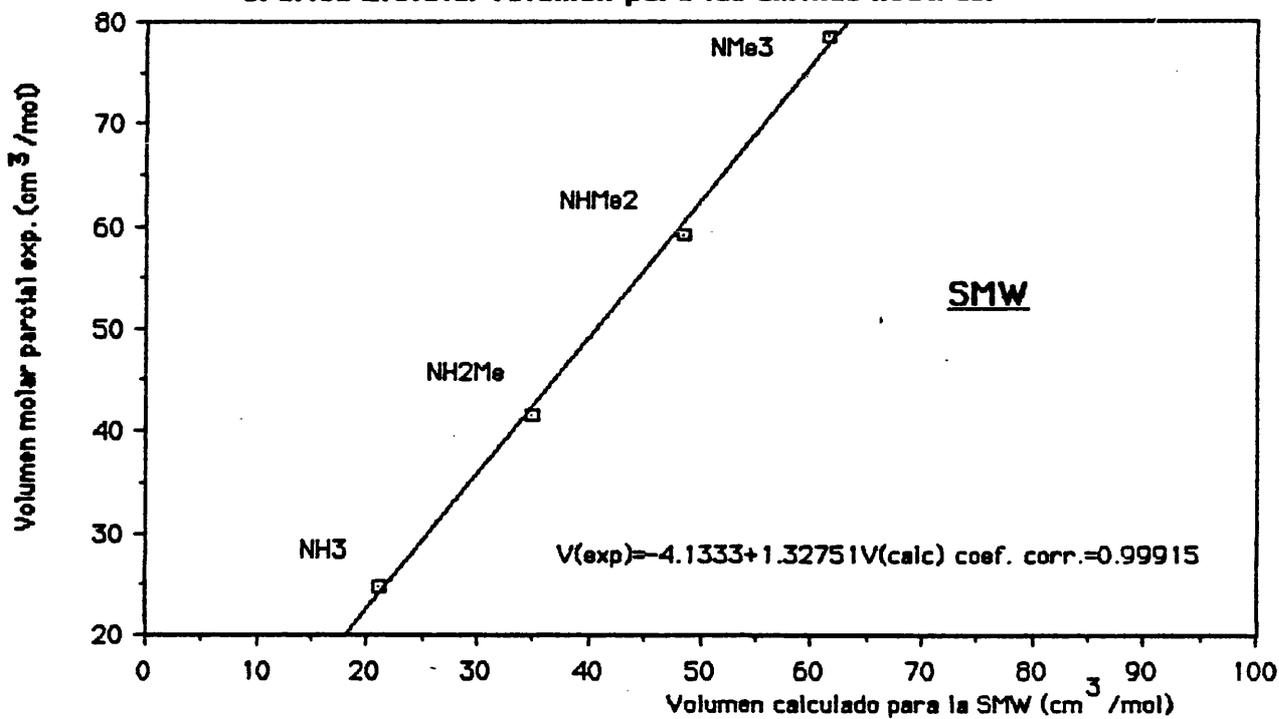
Molécula	V(exp)	V(SMW)	V(SM)
$\text{NH}_3$	24.6	21.1	21.1
$\text{NH}_2\text{-Me}$	41.6	34.9	34.9
$\text{NH-2Me}$	59.1	48.4	48.9
$\text{N-3Me}$	78.4	61.5	63.0
$+\text{NH}_4$	12.5	23.1	23.1
$+\text{NH}_3\text{-Me}$	30.7	37.0	37.0
$+\text{NH}_2\text{-2Me}$	49.2	50.5	51.0
$+\text{NH-3Me}$	67.3	63.8	65.2

En las **gráficas 2.6.3. a y b** se ha representado el volumen de la SMW y SM para las aminas neutras, con respecto al volumen molar parcial. En las mismas gráficas se indica el resultado de la recta de regresión que se ajusta a estos puntos. De igual manera se ha hecho para las aminas protonadas en las **gráficas 2.6.3. c y d**. Todas las rectas presentan un buen coeficiente de correlación, con lo que se puede asegurar una buena correspondencia entre los datos teóricos y experimentales. Las diferencias entre la SMW y SM son mínimas, pero es interesante observar que el coeficiente de correlación para las rectas de la SM son ligeramente mejores que para la SMW.

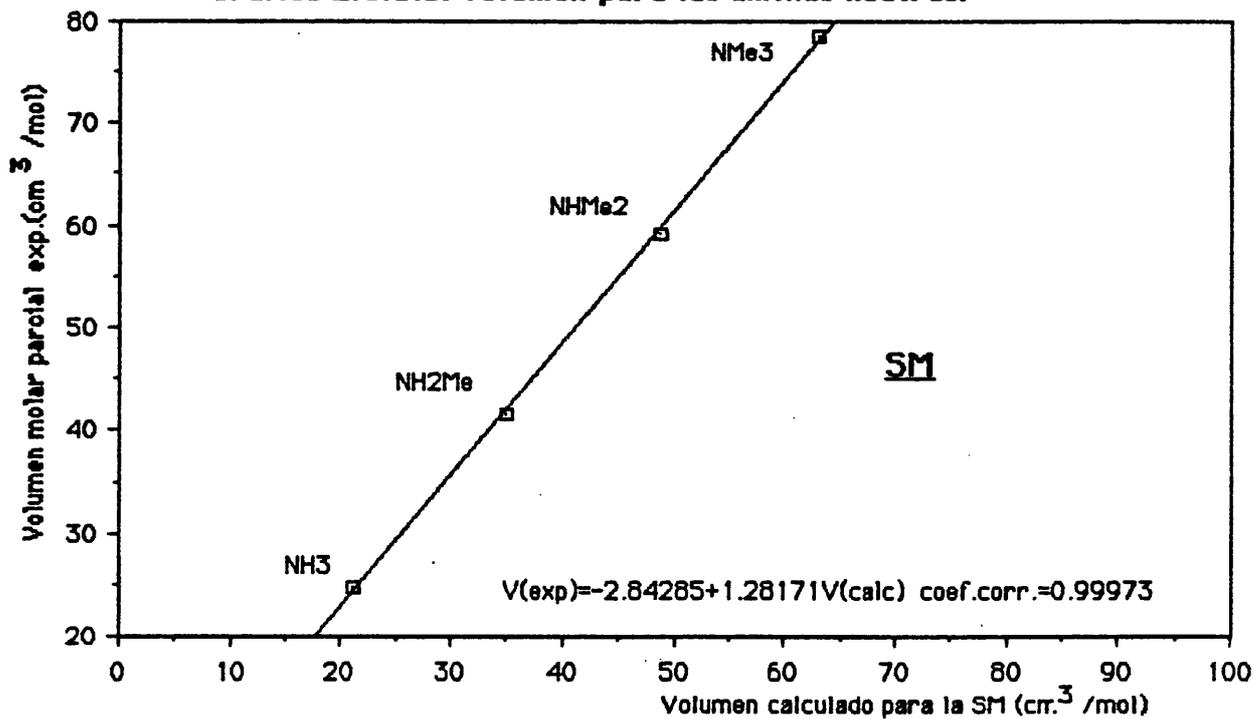
**Nota:**

La optimización total de las diferentes geometrías se ha realizado con el programa MONSTERGAUSS (Peterson, M.R. y col, 1980) utilizando la base doble zeta 4-31G (Ditchfield, R. y col, 1971) y con el método de gradientes analíticos propuesto por Davidon (Davidon, N.C., 1975). Las optimizaciones se finalizaron cuando la norma del gradiente total es menor que  $5 \cdot 10^{-4}$  mdinas o mdinas  $\text{\AA}/\text{rad}$ .

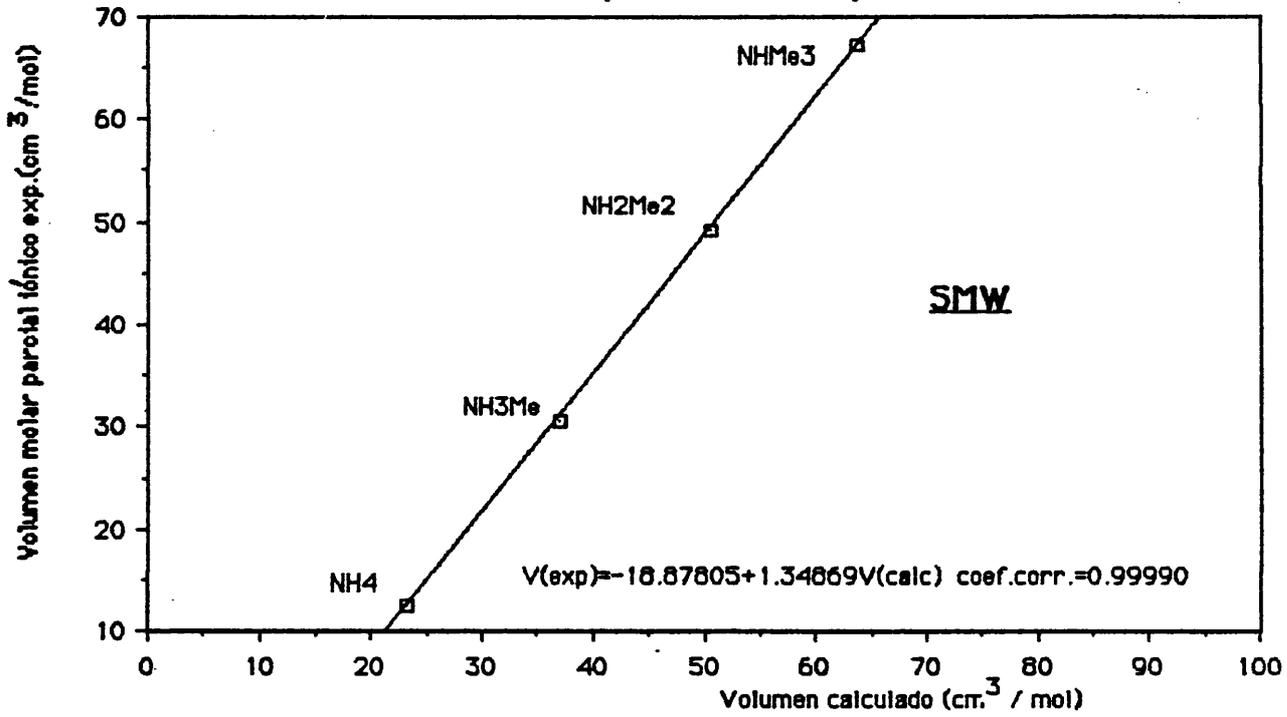
Gráfica 2.6.3.a. Volumen para las aminas neutras.



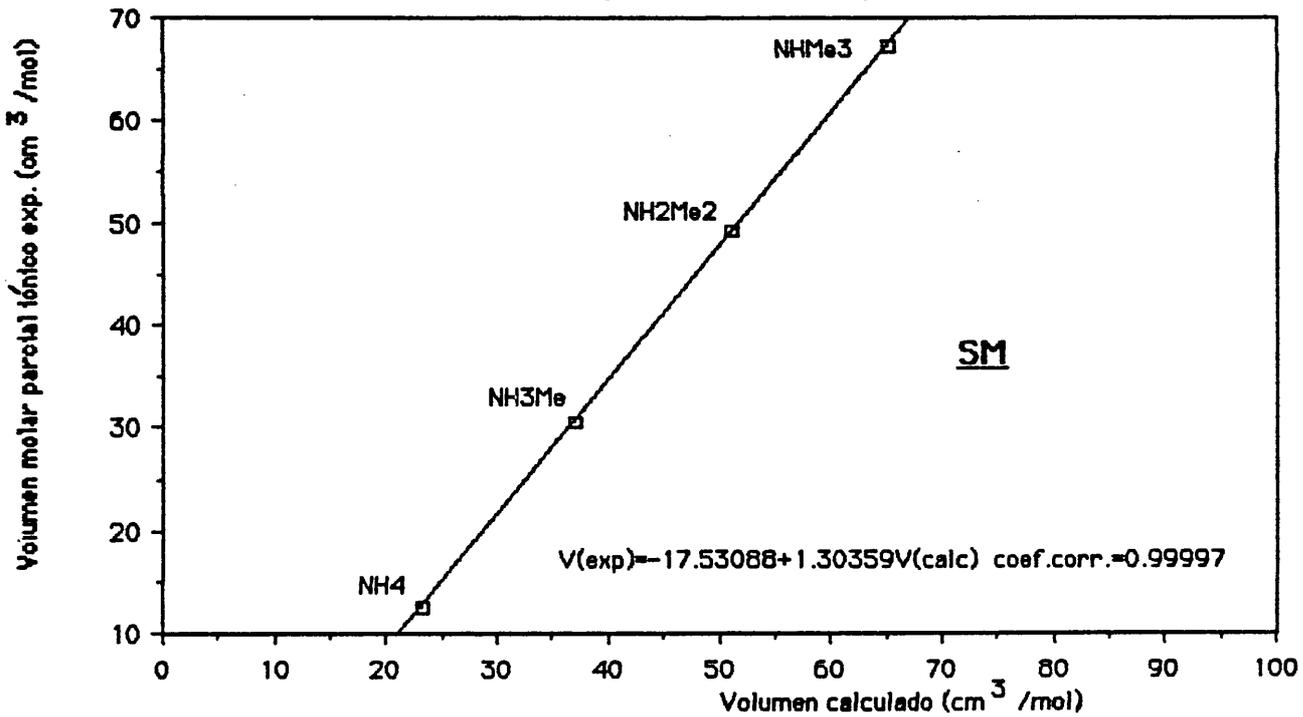
Gráfica 2.6.3.b. Volumen para las aminas neutras.



Gráfica 2.6.3.c. Volumen para las aminas protonadas.



Gráfica 2.6.3.d. Volumen para las aminas protonadas



### **2.6.4 Estudio de la fractalidad en el RBP.**

Una de las características más importantes de los biopolímeros es la especificidad que presentan al reaccionar con otras moléculas. Esta especificidad se refiere tanto a la zona de la molécula causante de la reacción como a con quién va a reaccionar. La fractalidad (Mandelbrot, B.B, 1983) nos proporciona un medio para caracterizar las irregularidades de la superficie, las cuales están relacionadas con los lugares activos de las biomoléculas (Lewis, M. y col, 1985).

Según M. Lewis y D. C. Rees (Lewis, M. y col 1985) el índice fractal de una proteína se puede determinar con la siguiente fórmula:

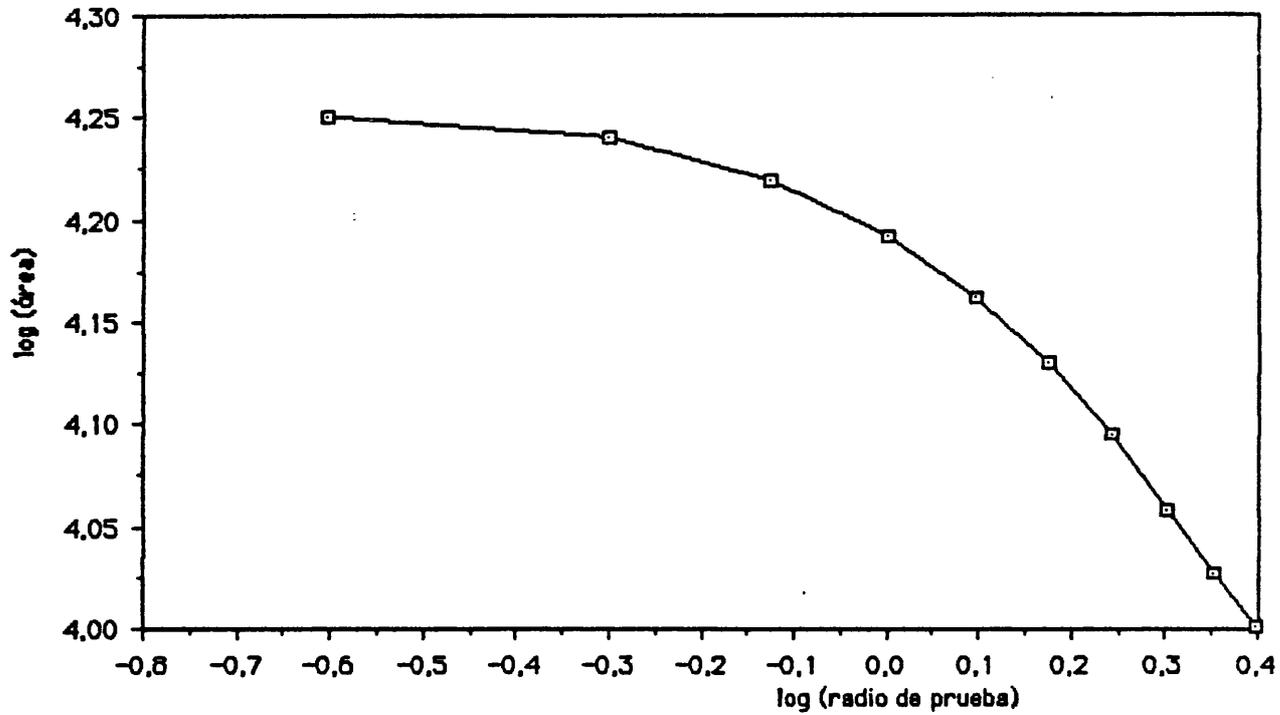
$$2 - D = \frac{d \log (A_s)}{d \log (R)}$$

Donde D es el índice de fractalidad y A es el valor del área que se obtiene para la MS cuando se utilizan esferas de prueba de diferente radio R.

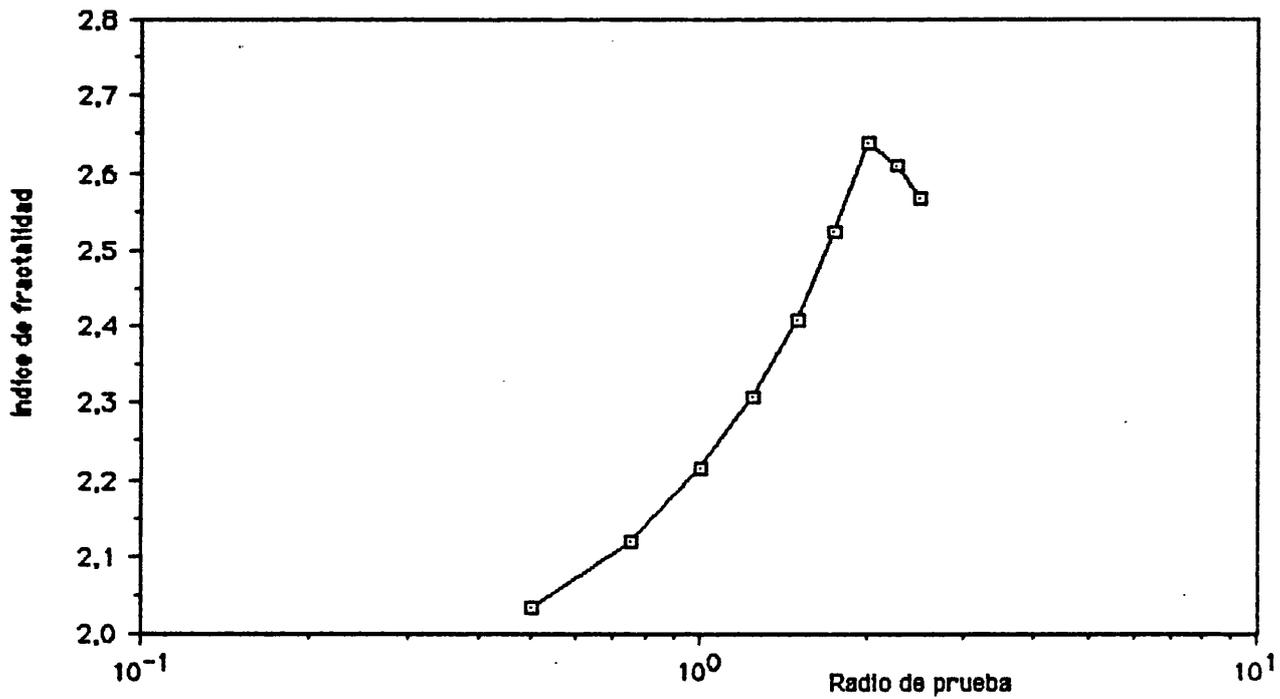
Nosotros hemos calculado el área de la proteína RBP con diferentes radios de prueba a fin de obtener la fractalidad de esta molécula. Los valores de los parámetros fueron NDIV=3, RET=1.0 y FRO=0.6. Los radios son los mismos que los dados en el subapartado 2.4.6..

En la **gráfica 2.6.4.a** se representa, en escala logarítmica, el valor del área frente al radio de la esfera de prueba, representándose en la **gráfica 2.6.4.b** la fractalidad obtenida frente al radio de la esfera. Estas gráficas presentan un aspecto similar a las obtenidas por Lewis y Rees para otras proteínas. Sin embargo Åquist, J y O. Tapia, 1987, indican que para que una superficie sea fractal y por tanto puedan ser utilizadas las leyes correspondientes, el índice fractal tiene que permanecer constante para cualquier valor del radio de prueba. Nosotros no vamos a entrar a discutir este concepto, ya que con este apartado solo hemos querido indicar la posibilidad de utilizar nuestro método para el cálculo de dicha propiedad.

**Gráfica 2.6.4.a. Variación del área del RBP con el radio de prueba.**



**Gráfica 2.6.4.b. Variación del índice de fractalidad con el radio de prueba.**



### **2.6.5 Estudio de la relación superficie-volumen en el PMMA**

En la teoría de Flory-Prigogime-Patterson (Pouchly, J. y col, 1976; Horta, A., 1979, 1985), las derivadas segundas de la energía libre de una mezcla ternaria polímero-dos disolventes a dilución infinita de polímero, que predicen los coeficientes de solvatación preferencial y total, vienen en función del parámetro  $S = S_3/S_{12}$ , donde  $S_{12}$  es el promedio de las relaciones superficie volumen de las moléculas de disolvente y  $S_3$  es la relación superficie volumen del polímero. Esta magnitud suele tomarse como una constante, relacionada con magnitudes geométricas de la cadena, cuando las expresiones teóricas donde aparece, predicen magnitudes termodinámicas en condiciones en que la macromolécula esta solvatada de forma diferente por la mezcla de disolventes. Es decir en condiciones donde la superficie expuesta al disolvente varía y por tanto es lógico pensar que también varíe la relación superficie-volumen.

Así nosotros con el fin de ilustrar otra posibilidad del cálculo de superficies, hemos realizado un estudio de la relación superficie volumen para el Polimetacrilato de metilo (PMMA) en una mezcla de MeCN y BuCl como disolventes, con el fin de analizar el grado de fiabilidad de esta aproximación. Sobre este sistema Horta, A. y col, 1981, han realizado un estudio experimental en el que asigna un valor constante para  $S_3$  de 0,59.

Los problemas que se plantean para la utilización de nuestro método en este tipo de moléculas, son el tamaño y la geometría de éstas. Los polímeros no son moléculas de un tamaño determinado sino que son una mezcla con cadenas de diferente longitud y por tanto las propiedades características de estos son propiedades promedio. Para poder realizar nuestros cálculos hicimos las siguientes consideraciones:

a) Tomamos el valor promedio del peso molecular del PMMA, como el representativo del tamaño de la molécula. De tal manera que dividiendo éste por el peso molecular de un monómero, obtuvimos el número de monómeros que lo forman, siendo éste de 2370 que corresponde a 16590 átomos pesados.

b) Dado que no se conocen las coordenadas cartesianas de estos átomos, construimos un

programa que utilizaba ángulos y distancias estándar y generaba aleatoriamente los ángulos diedros. Esta generación aleatoria estaba controlada por un parámetro que dabamos desde el exterior y que desechaba los angulos diedros que se apartasen un valor  $\pm \beta$  del de  $180^\circ$ , de tal manera que se podía controlar la "expansión" de la molécula. Por último el programa daba como resultado, además de las coordenadas cartesianas de los átomos, el radio de giro de la molécula.

c) La selección de las coordenadas a utilizar se realizó haciendo variar  $\beta$  entre 0 y  $180^\circ$ , eligiéndose aquellas coordenadas que tuvieran un radio de giro igual a los deducidos de los datos experimentales de Horta (**Campos, A., 1988**), representados en la **gráfica 2.6.5.a.**

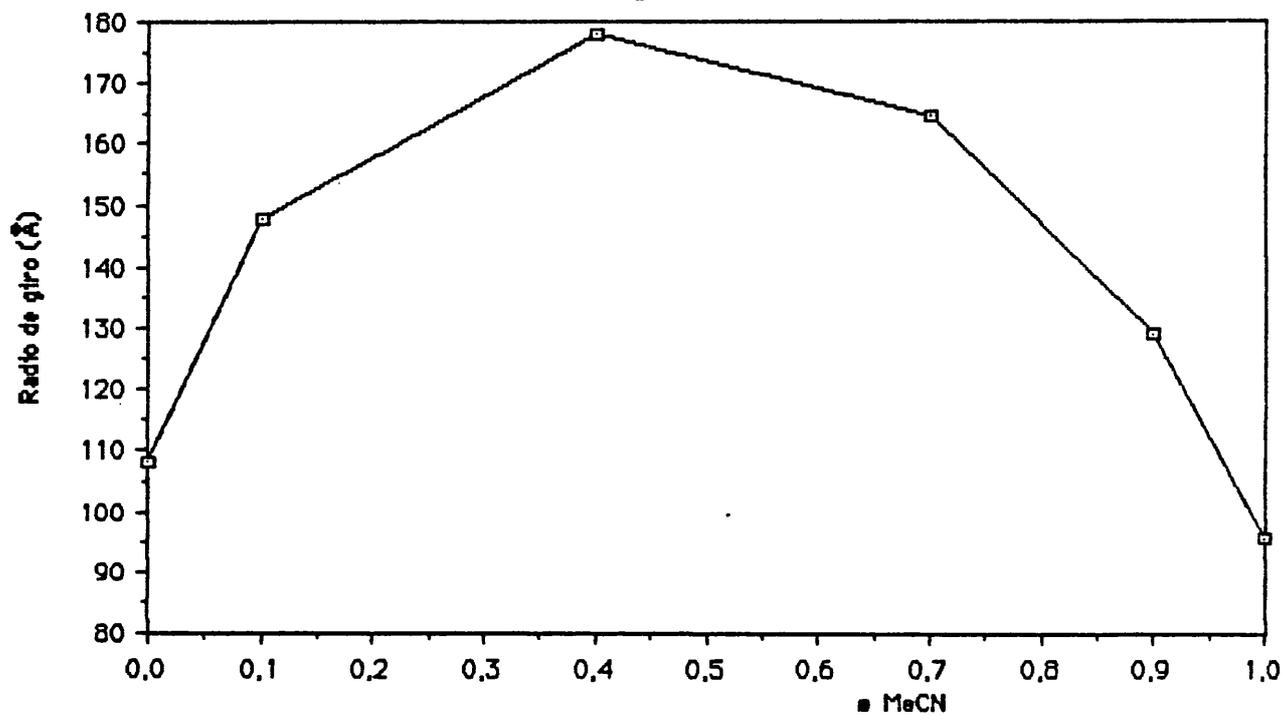
d) Los radios de las esferas utilizados fueron de 1.8 Å, centrando esferas sólo sobre los átomos pesados.

e) Dado que en este caso trabajamos con dos disolventes, había que tomar alguna decisión con respecto al radio a utilizar. Para ello calculamos con GEPOL el volumen de cada uno de los disolventes, utilizando coordenadas estandar y radios de van der Waals, definiendo un volumen promedio a partir de las fracciones molares,  $V_p = x_1V_1 + x_2V_2$ , considerando que este volumen correspondía a una esfera de radio R, que era el que utilizaríamos en los cálculos.

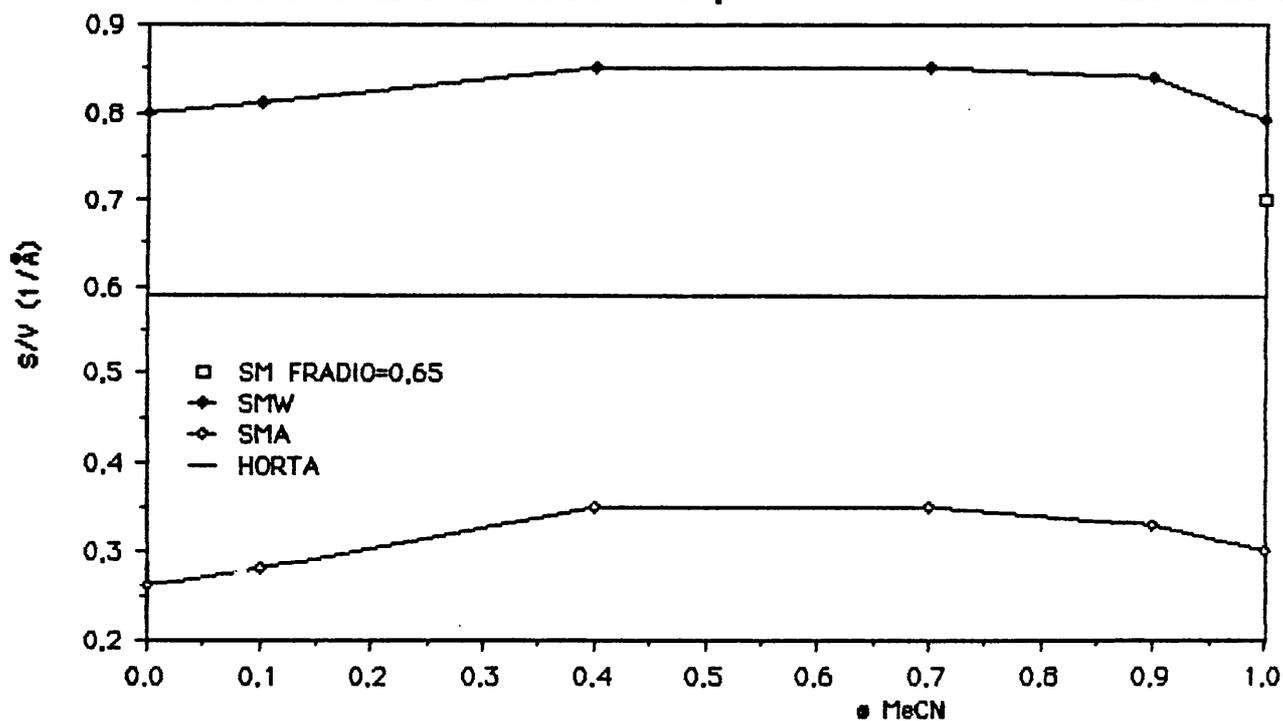
A lo largo de esta tesis estamos viendo que la superficie más real es la SM, sin embargo dado el número de átomos de estas estructuras, un cálculo de este tipo tiene un tiempo de cpu tan grande que se nos hace casi prohibitivo el realizarlo. Aun así, realizamos un cálculo para el polímero con un radio de giro de 95,5 Å, que correspondía a una fracción molar de 1.0 en MeCN y utilizando un valor de FRADIO de 0.65, obteniendo un factor superficie volumen de 0.7. El tiempo de cpu de este cálculo fue de 5 días, 14 horas.

Dados estos tiempos de cálculo decidimos utilizar para el análisis la SMW y la SMA. Los resultados obtenidos de la relación superficie volumen para éstas, se han representado en la **gráfica 2.6.5.b.**, representando también el único valor obtenido para la SM y el indicado por Horta. Como se puede ver la relación, en el caso de la SMW y de SMA, varía con la fracción molar, pero sin embargo esta variación no es muy grande. Con ello podemos decir, que la aproximación realizada al considerar  $S_3$  como constante no es muy grave. Sin embargo, sería interesante realizar los cálculos

Gráfica 2.6.5.a Variación del radio de giro del PMMA con la fracción molar del MeCN



Gráfica 2.6.5.b Variación de la relación superficie volumen con la fracción molar.

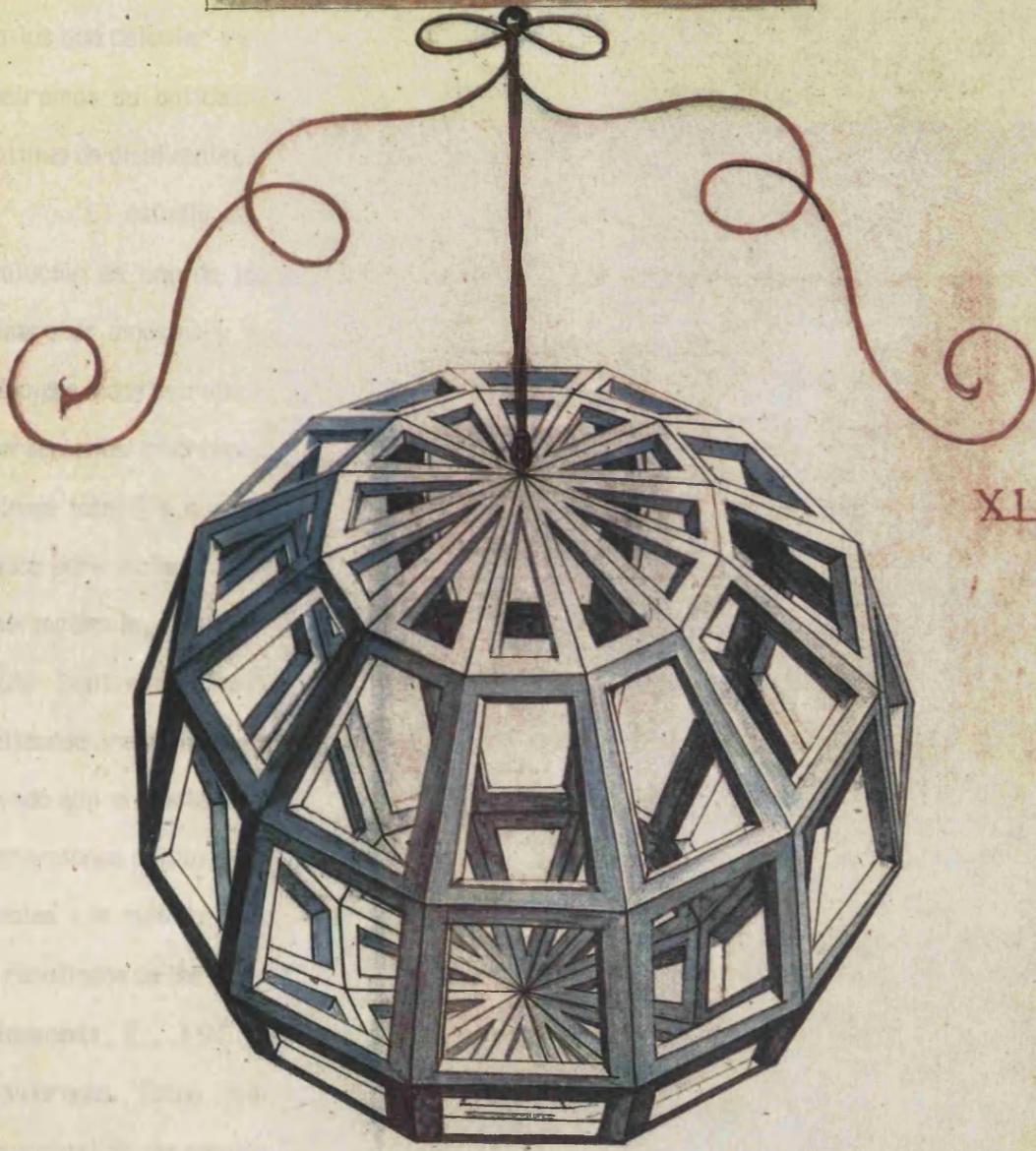


de estas propiedades termodinámicas de los polímeros con un valor variable de  $S_3$ . Los valores absolutos que hemos obtenido de  $S_3$  se diferencian del valor de 0.59 ajustado por Horta, pero pensamos que realizando cálculos con la SM los resultados serían más cercanos. Obsérvese en la **gráfica 2.6.5.b** como el valor obtenido con la SM se acerca al valor de Horta, aun habiendo utilizado un valor de FRADIO grande.

**3. INTERACCION SOLUTO-DISOLVENTE.**

.CXI.

SEPTVAGINTA DVARVM  
BASIVM VACVVM



XL

*Handwritten text in a cursive script, possibly a signature or note.*

Como comentamos en el capítulo 1, la determinación de la superficie envolvente de una molécula no solo nos permite obtener relaciones empíricas de su área y volumen con diversas propiedades moleculares, sino que además puede ser utilizada para construir modelos microscópicos con los que calcular valores de magnitudes como las energías libres de disolución. En este apartado mostramos su aplicación al cálculo de la interacción electrostática de un soluto con un modelo continuo de disolvente (Miertus, S. y col, 1981).

El estudio de la variación de las propiedades de una molécula al pasar de fase gas a disolución es uno de los grandes temas de la Química Teórica, habiéndose desarrollado un gran número de modelos y métodos (Pullman, B. ed., 1976). Sin embargo todos ellos pueden ser reducidos a dos estrategias principales: La primera de ellas fija la atención principalmente sobre las interacciones microscópicas del soluto con un número limitado de moléculas del disolvente; el sistema total (la supermolécula) es estudiado con los métodos normales de la mecánica cuántica usados para moléculas simples. De tal manera que, introduciendo más moléculas de disolvente en la supermolécula, se puede aumentar el grado de información y detalle de las interacciones soluto-disolvente (Pullman, B. y col, 1979). Aun así, esta adición de moléculas no puede realizarse indefinidamente, no solo por el gran gasto de tiempo de cálculo necesario, que resulta elevado aún en sistemas pequeños, sino porque desprecia uno de los aspectos más relevantes de las interacciones soluto disolvente, como es la baja especificidad y direccionalidad que tienen gran parte de estas. Los métodos estadísticos, como son los de Dinámica Molecular y Monte Carlo, hacen uso de los resultados de los de supermolécula para construir los potenciales intermoleculares que necesitan (Clementi, E., 1980) y permiten un notable incremento en el número de moléculas de disolvente consideradas. Estos métodos han sido utilizados para obtener información termodinámica y estructural de las disoluciones, pero sin embargo su información no es lo suficientemente explícita como para ser utilizada en muchos casos de interés químico.

La segunda estrategia intenta introducir directamente el efecto promedio del disolvente cambiando la descripción microscópica de éste por un medio continuo con unas adecuadas propiedades macroscópicas tales como la constante dieléctrica, el coeficiente de expansión térmica, etc.. La gran

simplicidad y el fácil análisis que permite este modelo ha hecho que se desarrollen una gran cantidad de métodos basados en esta aproximación (**Tapia, O., 1982**).

Entre todos los métodos de continuo existentes hemos elegido el de Miertus y col, porque pone de manifiesto la importancia de la superficie molecular tal como es, ya que en otros métodos esta superficie es sustituida por una esfera o elipsoide. Este modelo ha sido usado con éxito en diversos problemas de interés químico como conformaciones moleculares, interacciones moleculares, procesos tautoméricos, reactividad química, procesos fotoquímicos, etc... (**Bonaccorsi, R. y col, 1982, 1983a, 1983b, 1984a, 1984b, 1984c, 1985, 1986; Miertus, S. y col, 1982; Duben, A. y col, 1982; Cimiraglia, R. y col, 1981; Persico, M. y col, 1984; Alagona, G. y col 1986a, 1986b**).

En el primer apartado mostraremos la formulación básica del método de Miertus y col. En el segundo apartado estudiaremos las ventajas de sustituir el algoritmo geométrico original, denominado SPHERE por el GEPOL y analizaremos las diferencias de usar la superficie de van der Waals o la superficie molecular. En el tercer apartado mostramos un ejemplo de aplicación, estudiando la inversión de la basicidad de las aminas al pasar de fase gas a disolución.

### 3.1. FORMULACION BASICA DEL METODO

En los modelos del continuo el soluto es representado por una distribución de carga  $\rho(r)$  en el interior de una cavidad dentro de un medio dieléctrico polarizable e infinito, con una permitividad  $\epsilon$ . La distribución de carga molecular induce en el dieléctrico un campo de reacción, el cual actúa sobre la misma distribución de carga, produciendo un cambio en esta (con  $\rho(r)$  indicamos la distribución de carga total, incluyendo electrones y núcleo, siendo  $\rho^0(r)$  la distribución de carga en el vacío). Una de las formas usadas para obtener  $\rho(r)$  consiste en considerar al dieléctrico polarizado como una perturbación del hamiltoniano del soluto en el vacío  $H^0$ , resolviendo posteriormente la nueva ecuación de Schrödinger independiente del tiempo.

La polarización del dieléctrico infinito se puede representar por un sistema de cargas virtuales sobre la superficie de la cavidad con una densidad  $\sigma(\mathbf{s})$  (se utiliza el vector  $\mathbf{s}$  para hacer hincapié en que la distribución de carga es diferente de cero solamente en la superficie de la cavidad  $S$ ).

Para poder determinar  $\sigma(\mathbf{s})$  se necesita conocer el potencial que actúa sobre la superficie de la cavidad, siendo este:

$$V(r) = V_{\rho}(r) + V_{\sigma}(r) \quad \text{Fórmula 3.1.a}$$

donde  $V_{\rho}(r)$  y  $V_{\sigma}(r)$  corresponden a las contribuciones al  $V$  por parte de la distribución de carga del soluto y de la misma distribución de carga  $\sigma(\mathbf{s})$ . Según la electrostática clásica el potencial total  $V(r)$ , se puede obtener resolviendo la ecuación de Laplace

$$\nabla^2 V(r) = 0, \quad \text{Fórmula 3.1.b}$$

en la parte libre de carga del sistema ( C.J.F. Böttcher, 1952) utilizando las adecuadas condiciones de contorno sobre la superficie de la cavidad

$$V(s)_- = V(s)_+ \quad \text{Fórmula 3.1.c}$$

$$\left(\frac{\partial V(s)}{\partial n}\right)_{s^-} = \epsilon \left(\frac{\partial V(s)}{\partial n}\right)_{s^+}$$

$$\text{Fórmula 3.1.d}$$

y en el infinito

$$\lim_{r \rightarrow \infty} rV(r) = \alpha$$

**Fórmula 3.1.e**

$$\lim_{r \rightarrow \infty} r^2 \left[ \frac{\partial V(r)}{\partial r} \right] = \beta$$

**Fórmula 3.1.f**

Con los índices - y + se quiere indicar puntos situados a una infinitésima distancia de la superficie hacia dentro y hacia fuera de la cavidad y con  $n$  el vector normal a la superficie en el punto  $s$  de esta. En algunas cavidades, como son la esfera y los elipsoides, la laplaciana es separable, pudiéndose resolver analíticamente la ecuación de Laplace, pero sin embargo para cavidades con una forma más general, uno tiene que utilizar soluciones numéricas.

El modelo de Miertus y col fue creado con el fin de poder calcular la función de onda de una molécula en disolución, con una cavidad formada por un conjunto de esferas interconectadas, centradas sobre los núcleos atómicos de la molécula, es decir, cavidades que no permiten resolver la ecuación de Laplace. En líneas generales el modelo se basa en el siguiente proceso interactivo:

La molécula de soluto crea un campo eléctrico que polariza al dieléctrico, la polarización del dieléctrico se manifiesta por la aparición de una distribución de carga continua  $\sigma(\mathbf{s})$ . Con el fin de poder trabajar con  $\sigma(\mathbf{s})$  el modelo realiza su principal aproximación consistente en sustituir a esta por una distribución de carga discreta. Para ello se divide la superficie de la cavidad en un conjunto de teselas de área  $\Delta S$ , situándose en el centro de cada una de ellas una carga. Estas cargas se calculan utilizando el campo eléctrico que crea el soluto, que en un primer paso corresponde al soluto no polarizado. A continuación estas cargas se autopolarizan, es decir, sobre cada carga actúa un campo eléctrico debido al resto de las cargas de la cavidad y a sí misma. Este proceso de autopolarización parte de las cargas creadas por el campo del soluto y se repite hasta que no se

observa variación en la carga total de la cavidad. Una vez alcanzada esta convergencia se introduce el potencial que crean estas cargas como un operador de perturbación en el hamiltoniano del soluto, obteniendo una nueva función de onda. Entonces esta nueva función crea un nuevo campo eléctrico que actúa sobre el dieléctrico, apareciendo una nueva distribución de carga. Todo el proceso se repite tantas veces como sea necesario para que la energía del soluto polarizado no varíe de una iteración a otra. De esta manera se obtiene la función de onda del soluto en disolución de la cual podremos extraer toda la información necesaria para el estudio de la molécula.

Este proceso iterativo está basado en la siguiente formulación electrostática:

La distribución de carga  $\sigma(\mathbf{s})$  está directamente relacionada con la componente normal a la superficie de la cavidad del vector de polarización  $\mathbf{P}$  del dieléctrico, por:

$$\sigma(\mathbf{s}) = -\mathbf{P}_{n+} = \frac{\epsilon - 1}{4\pi} \left( \frac{\partial \psi}{\partial n} \right)_{s+}$$

**Fórmula 3.1.g**

pudiéndose escribir, si se tienen en cuenta las condiciones de contorno antes indicadas

$$\sigma(\mathbf{s}) = \frac{\epsilon - 1}{4\pi\epsilon} \left( \frac{\partial \psi}{\partial n} \right)_{s-}$$

**Fórmula 3.1.h**

Donde el gradiente del potencial es calculado en la parte interior de la cavidad, estando este potencial compuesto por el potencial que crean los electrones y núcleos del soluto, y por el potencial que crea la misma distribución de carga  $\sigma(\mathbf{s})$ , haciéndose por tanto necesario un proceso iterativo para obtener el valor de  $\sigma(\mathbf{s})$ . Dado que lo que se calculan son realmente las cargas discretas la fórmula utilizada es:

$$q_k = \sigma(\mathbf{s}_k) \Delta S_k$$

**Fórmula 3.1.i**

donde  $\sigma(\mathbf{s}_k)$  es la densidad de carga calculada en el centro de la tesela  $k$  ( $\mathbf{s}_k$  corresponde al vector de posición de la carga  $k$ ).

En el **esquema 3.1.a** se muestran los diferentes pasos del proceso iterativo necesario para obtener la función de onda del soluto  $M$  en disolución. A continuación explicaremos cada uno de los pasos de este proceso.

### **Determinación geométrica de la cavidad.**

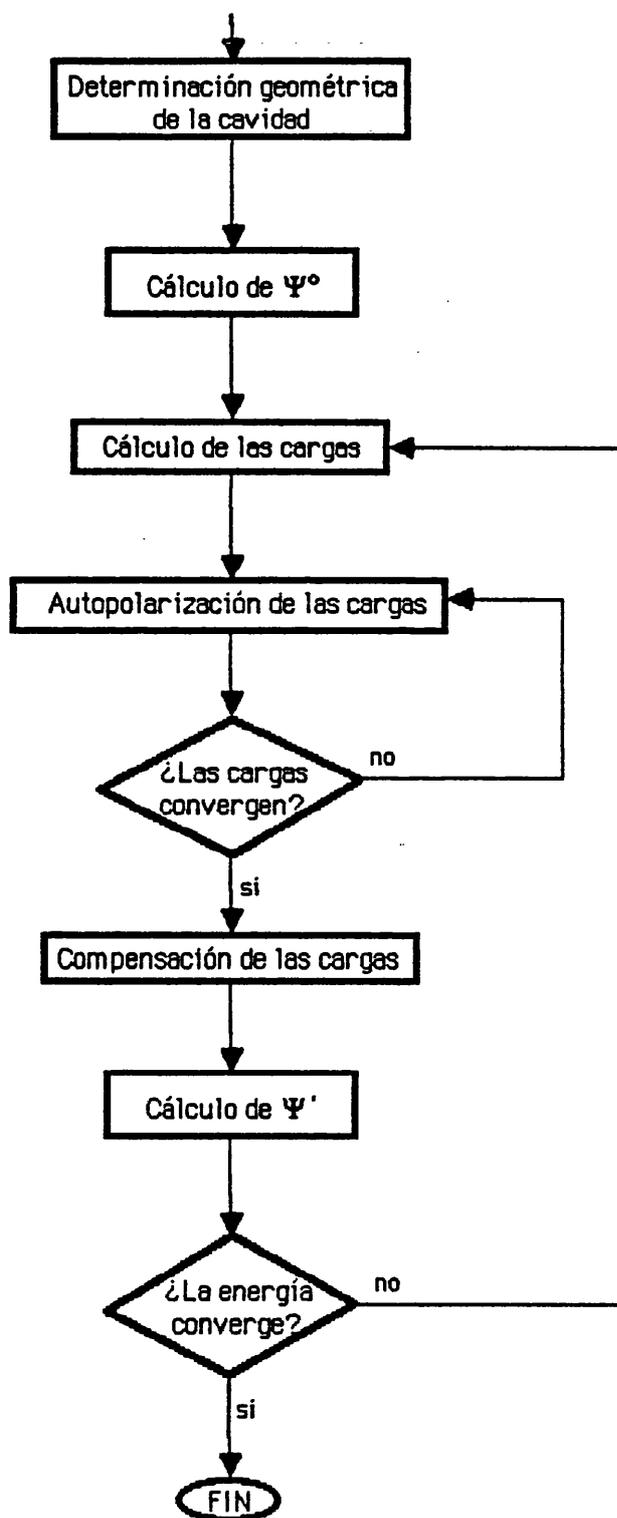
El primer paso del proceso consiste en definir la superficie de la cavidad, determinando los puntos de esta donde van a ir situadas las cargas, que representan la polarización del dieléctrico, así como el área de la tesela asociada a este punto. También se necesita calcular las coordenadas de unos puntos situados en el interior de la cavidad, a una distancia muy pequeña de los centros de las teselas y de tal manera que la recta que une estos puntos con sus respectivos centros de tesela, es perpendicular a la superficie. Estos puntos son utilizados para calcular de una manera numérica el gradiente del potencial.

El cálculo de la superficie de la cavidad y la definición de las teselas puede ser obtenida de dos maneras, indicadas en el capítulo anterior: La primera que corresponde a la forma original utilizada por Miertus y col, hace uso de un conjunto de meridianos y paralelos. Y la segunda utilizando el método GEPOL.

Si se utiliza el método original, que denominaremos SPHERE, se puede controlar el número de teselas y por tanto el número de cargas sobre cada esfera, variando el número de meridianos y paralelos. Sin embargo, este número no puede ser muy grande ya que acarrearía un volumen elevado de cálculos. En el caso de GEPOL el número máximo de cargas por esfera siempre es 60. En el apartado 3.2 analizaremos las diferencias entre uno y otro método.

### **Cálculo de $\Psi^0$**

El cálculo de la función de onda en el vacío es la que permite iniciar el proceso interactivo, ya que con ella se conoce la distribución de carga  $\rho$ . El modelo original utilizaba únicamente funciones de onda del tipo ab-initio, introduciendo nosotros la posibilidad de realizar cálculos del tipo CNDO, con el fin de poder trabajar con moléculas de gran tamaño.



**Esquema 3.1.a** Proceso iterativo para la obtención de la función de onda de un soluto en disolución según el modelo de Miertus et col.

### **Cálculo de las cargas**

Como se pueda deducir de las **Fórmulas 3.1.h e i**, para la obtención de las cargas se necesita conocer la posición de éstas, el área que representan y la variación del potencial sobre la normal a la superficie y hacia dentro de ésta. Aquí analizaremos como se calcula la derivada del potencial puesto que los aspectos geométricos ya los comentamos anteriormente.

Para realizar este cálculo se determina el potencial en el centro de la tesela y en un punto situado en el interior de la cavidad, a una distancia pequeña  $\delta$  de la superficie, hallándose la diferencia entre éstos y dividiendo a ésta por  $\delta$ . Para el cálculo del potencial los autores utilizaron el desarrollado por **E. Scrocco y col, 1972**, para funciones ab-initio. En nuestros calculos CNDO hemos utilizado el indicado como número 2 en el artículo de **C. Giessner-Prettre y col de 1972**.

Este cálculo se realiza de igual forma con la función de onda no perturbada que con las perturbadas.

### **Autopolarización de las cargas.**

El campo eléctrico que actúa en el centro de un elemento  $\Delta S_k$  proviene tanto del producido por la distribución de carga del soluto  $\rho^n$  (con  $n$  indicamos el número de iteración a que corresponde), como del resto de cargas  $q$  de la cavidad. La contribución de  $\rho^n$  ya la hemos tenido en cuenta en el párrafo anterior, y producía un conjunto de cargas sobre la superficie  $q^{n0}$  (el segundo índice hace referencia al ciclo de iteración de la autopolarización en que se encuentran las cargas, en este caso es cero ya que corresponden a las cargas cuando no han sufrido aun un proceso de autopolarización).

El valor de una carga en la superficie después de una iteración es:

$$q_k^{nm} = q_k^{n0} - \frac{\epsilon - 1}{4\pi\epsilon} \frac{\Delta S_k}{|\delta|} \sum_{l \neq k} q_l^{n, m-1} \left[ \frac{1}{|(r_k + \delta) - r_l|} - \frac{1}{|r_k - r_l|} \right] + \frac{\epsilon - 1}{2\epsilon} q_k^{n, m-1} \left[ 1 - \sqrt{\frac{\Delta S_k}{4\pi R_k^2}} \right]$$

**Fórmula 3.1.j**

donde el segundo término corresponde a la polarización producida por todas las cargas de la superficie excepto la k, y el tercer término la producida por k sobre sí misma.

### Compensación de las cargas.

Una vez terminado el proceso de autopolarización la suma de todas las cargas de la superficie tendría que dar cero en el caso de que el soluto fuera neutro y  $-Q_M[(\epsilon - 1)/4\pi\epsilon]$  para solutos con una carga igual a  $Q_M$ . Sin embargo el resultado no es éste. Esta discrepancia entre el valor obtenido y lo que se debiera obtener está producida al considerar que la distribución de carga del soluto  $p$  está encerrada totalmente dentro de la cavidad, hecho que no es correcto ya que la función de onda se extiende más allá de la superficie. Para corregir esto se distribuye la carga extra  $-Q$  sobre todas las cargas  $Q_k$  de la superficie, de acuerdo a la magnitud y signo que tienen, usando para las cargas positivas el factor

$$g^+ = 1 - \frac{Q}{2Q_+}$$

**Fórmula 3.1.k**

y para las cargas negativas

$$g^- = 1 - \frac{Q}{2Q_-}$$

**Fórmula 3.1.l**

**Cálculo de  $\Psi$** 

Una vez que se tiene la distribución de carga sobre la superficie, se introduce el potencial que crea esta en el hamiltoniano del soluto obteniendo la nueva función de onda. Este potencial es de la forma:

$$V_{\sigma}(r) = \sum_k \frac{q_k}{|r - s|} \frac{1}{\Delta S_k}$$

**Fórmula 3.1.m**

Una vez calculada la nueva función de onda se repite el proceso hasta una convergencia de la energía.

Con la energía obtenida en el proceso final se puede obtener la componente electrostática de la energía libre del sistema (i.e., el trabajo necesario para reunir las cargas del soluto en disolución a temperatura y volumen constante):

$$G = E_{\text{tot}}^{\circ} - \frac{1}{2} \int \rho \cdot V_{\sigma} dr$$

**Fórmula 3.1.n**

Resultando la parte electrostática de la energía libre de disolución:

$$\Delta G_{\text{el}} = G - E_{\text{tot}}^{\circ}$$

**Fórmula 3.1.ñ**

Como vemos, con este modelo sólo podemos calcular la componente electrostática de la energía libre de disolución y, para tener la total, habría que tener en cuenta otras contribuciones como son las limitaciones de movimiento que sufre el soluto en fase condensada, términos relacionados con las diferencias de los grados de libertad rotacionales en gas y disolución, términos de dispersión y términos relacionados con la formación de la cavidad en la disolución. En los siguientes apartados utilizaremos en algunos casos la energía de cavitación calculada por el método de **R. A. Pierotti, 1976**, de tal manera que si incluimos este término la energía libre de disolución

quedará como:

$$\Delta G'_{\text{sol}} = \Delta G_{\text{el}} + G_{\text{cav}}$$

**Fórmula 3.1.0**

donde  $G_{\text{cav}}$  depende de la superficie y del volumen de la cavidad así como del disolvente.

## 3.2 SUPERFICIE ENVOLVENTE Y ENERGIA

Como vimos en el apartado anterior, la interacción del soluto con el disolvente depende de la superficie que separa a ambos y, por tanto, se hace necesario utilizar un método lo mejor posible para definir ésta. En los dos primeros subapartados analizaremos el comportamiento de GEPOL y SPHERE frente a diferentes orientaciones espaciales de las moléculas. En el tercero analizaremos las diferencias al usar la SMW o la SM en el estudio conformacional del octanol.

### 3.2.1 HF

En este primer caso hemos realizado el estudio sobre el HF. Para ello hemos usado una distancia intermolecular de 0.92 Å (valor estandar, **Pople, J. A. y col, 1970**), centrandó sobre los átomos de H y F esferas con un radio de 1.44 Å y 1.62 Å respectivamente. Los cálculos se han realizado con una función de onda tipo CNDO.

Para GEPOL se realizaron cálculos variando NDIV desde 1 hasta 5 de uno en uno. Para el SPHERE se realizaron cálculos con dos teselaciones diferentes: En la primera se utilizó para los paralelos un incremento de 30°,  $\Delta\theta=30^\circ$ , y para los meridianos de 36°,  $\Delta\phi=36^\circ$ , tomando en la segunda incrementos iguales,  $\Delta\theta=\Delta\phi=25,71^\circ$ . Estos valores de los incrementos dejan dividida cada esfera en 60 y 98 teselas respectivamente.

Con el fin de analizar la invarianza de los resultados frente a la orientación espacial de la molécula, se repitieron los cálculos situando el eje molecular sobre los ejes X, Y y Z.

En la **tabla 3.2.1.a** se dan los valores promedios y desviaciones, de las tres orientaciones para el área total y parcial de la cavidad(A), el valor del volumen total de la cavidad(V), el número de puntos sobre la superficie resultante (P), la contribución electrostática ( $\Delta G_{elec}$ ) y de cavitación ( $G_{cav}$ ) a la energía libre de disolución y el tiempo de cpu utilizado (ordenador DPS8/52 BULL).

**Tabla 3.2.1.a.** Area, volumen, número de puntos, energía libre de interacción y cavitación para el HF en una cavidad de dos esferas.  $A(\text{Å}^2)$ ,  $V(\text{Å}^3)$ ,  $G(\text{kcal/mol})$ , tiempo(segundos, DPS8/32 BULL).

nivel	$A_T$	$A_F$	$A_H$	$V$	$P$	$\Delta G_{\text{elec}}$	$G_{\text{cav}}$	Tiempo
1	38, ±2,	24, ±2,	14, ±1,	22, ±1,	76 ±5	-1,51 ±0,04	4,7 ±0,2	9 ±3
2	38,4 ±0,9	23,8 ±0,6	14,5 ±0,4	21,8 ±0,4	87 ±3	-1,53 ±0,01	4,78 ±0,09	14 ±1
3	38,79 ±0,07	24,0 ±0,1	14,7 ±0,2	22,06 ±0,03	93 ±3	-1,54 ±0,00	4,82 ±0,01	15 ±2
4	38,69 ±0,06	24,18 ±0,07	14,52 ±0,01	22,03 ±0,03	93 ±3	-1,54 ±0,00	4,81 ±0,01	15 ±2
5	38,71 ±0,02	24,22 ±0,02	14,49 ±0,00	22,04 ±0,01	94 ±2	-1,53 ±0,00	4,81 ±0,00	16 ±2
$\Delta\theta=30^\circ$	39, ±1,	24,4 ±0,4	14, ±1,	21,9 ±0,4	79 ±8	-1,52 ±0,03	4,8 ±0,1	11 ±2
$\Delta\phi=36^\circ$								
$\Delta\theta=25,71^\circ$	40, ±3,	25, ±2,	15, ±1,	23, ±1,	131 ±6	-1,58 ±0,08	4,9 ±0,3	29 ±2
$\Delta\phi=25,71^\circ$								
exacto	38,7012	24,2188	14,4824	22,0026				

Si observamos los resultados obtenidos con GEPOL vemos que a medida que se aumenta el valor de NDIY, los valores de todas las propiedades dadas en la tabla tienden a un valor y, además, la desviación se hace más pequeña. Sin embargo casi todos los valores obtenidos con SPHERE dan una mayor desviación al aumentar el número de puntos sobre la superficie.

Con respecto al tiempo de cálculo es interesante destacar que el aumento, al pasar de un nivel de división a otro, es mayor en el caso del SPHERE que en el de GEPOL. Obsérvese que el tiempo está relacionado con el número de puntos resultante sobre la superficie.

### 3.2.2 H<sub>2</sub>CO

Para el H<sub>2</sub>CO hemos realizado un conjunto de cálculos con las mismas teselaciones que en el caso anterior. Para el análisis de la invarianza espacial hemos repetido los cálculos para 6 orientaciones diferentes. La geometría utilizada es D<sub>H<sub>2</sub>C</sub>=1.08 Å, D<sub>CO</sub>=1.22 Å y <HCH=120°, estando la cavidad formada por cuatro esferas de radio R<sub>H</sub>=1.44 Å, R<sub>C</sub>=1.92 Å y R<sub>O</sub>=1.68 Å.

En la **tabla 3.2.2.a** se dan los resultados obtenidos con el mismo formato que en el caso anterior. Como se puede observar el comportamiento de ambos métodos es similar al caso del HF.

**Tabla 3.2.2.a.** Área, volumen, número de puntos, energía libre de interacción y cavitación para el H<sub>2</sub>CO en una cavidad de cuatro esferas. A(Å<sup>2</sup>), V(Å<sup>3</sup>), G(kcal/mol), tiempo(segundos, DPS8/52 BULL).

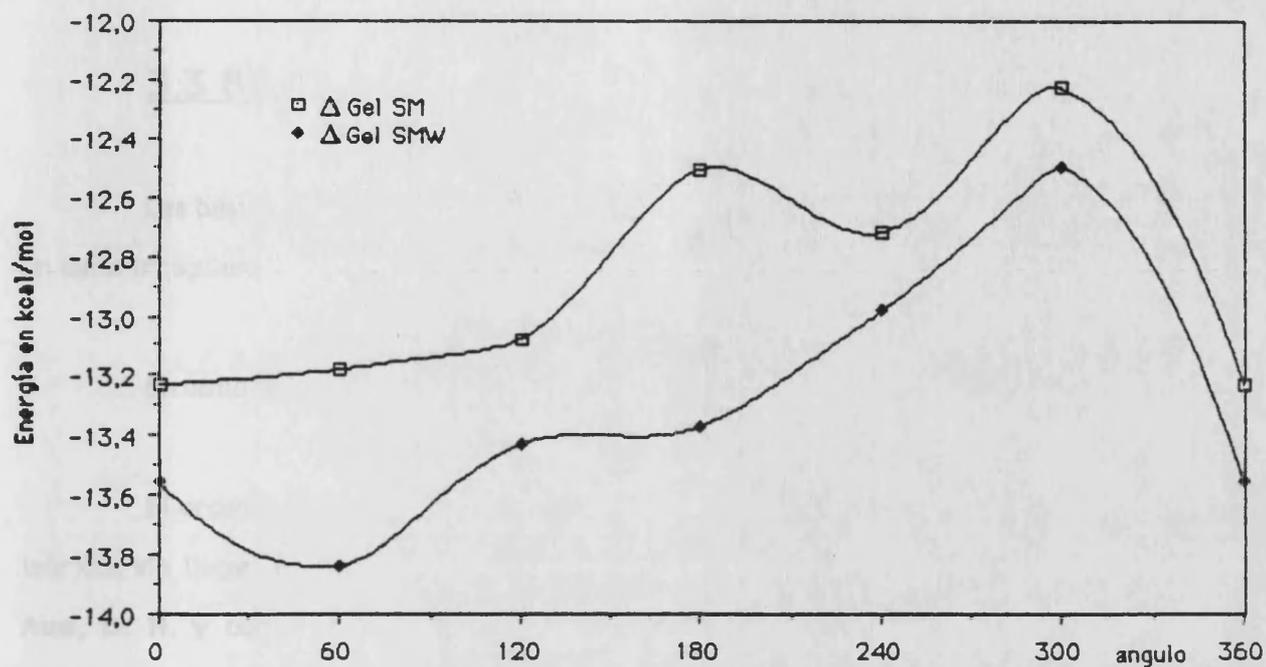
nivel	A <sub>T</sub>	A <sub>C</sub>	A <sub>O</sub>	A <sub>H</sub>	V	P	ΔG <sub>el</sub>	G <sub>cav</sub>	T
1	61. ±4.	19. ±3.	19. ±2.	11,3 ±0,5	42. ±2.	109 ±5	-1,23 ±0,03	7,0 ±0,4	22 ±2
2	62,3 ±0,7	20,1 ±0,3	20,0 ±0,7	11,1 ±0,1	43,0 ±0,4	137 ±5	-1,20 ±0,01	7,12 ±0,07	34 ±2
3	62,4 ±0,1	19,7 ±0,1	20,36 ±0,06	11,17 ±0,06	42,98 ±0,08	150 ±2	-1,19 ±0,01	7,13 ±0,01	41 ±1
4	62,3 ±0,1	19,55 ±0,08	20,46 ±0,04	11,17 ±0,03	42,95 ±0,07	154 ±3	-1,19 ±0,01	7,12 ±0,01	44 ±2
5	62,24 ±0,04	19,49 ±0,05	20,42 ±0,01	11,16 ±0,01	42,89 ±0,03	156 ±3	-1,19 ±0,00	7,11 ±0,00	47 ±2
Δθ=30°	61. ±2.	19. ±3.	20. ±2.	11,2 ±0,8	42,0 ±0,7	111 ±7	-1,21 ±0,05	7,0 ±0,1	23 ±3
Δφ=36°	63. ±3.	20. ±2.	20,9 ±0,7	10,9 ±0,4	43. ±2.	185 ±9	-1,20 ±0,05	7,2 ±0,3	59 ±6
Δθ=25,71°									
Δφ=25,71°									
exacto	62,2650	19,5122	20,4349	11,1589	42,7599				

### 3.2.3 Rotación de n-octanol.

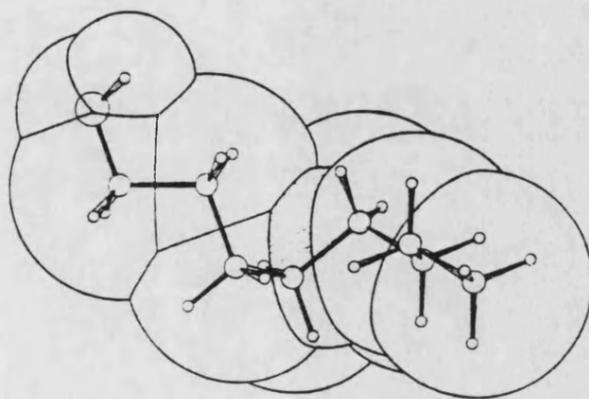
Con el fin de comprobar la influencia de la SM sobre la energía libre de disolución realizamos una serie de cálculos sobre una de las posibles rotaciones del n-octanol. Los cálculos se hicieron utilizando una base STO-3G y utilizando tanto la SM como la SMW. Los resultados, tanto geométricos como energéticos se dan en la **tabla 3.2.3.a**, representándose en la **gráfica 3.2.3.a** la variación de la energía libre de disolución en función del ángulo para la SM y la SMW. Las diferencias de energía entre una curva y otra son pequeñas debido a que, en estas conformaciones, la molécula resulta bastante compacta (**figura 3.2.3.a**). Sin embargo, en compuestos con cadenas más largas las diferencias pueden ser mucho mayores.

**Tabla 3.2.3.a** Area, volumen, y valores energeticos para una rotación del n-octanol, obtenidos con GEPOL con la SM y con la SW. A(Å<sup>2</sup>), V(Å<sup>3</sup>), energías en kcal/mol.

ángulo	SM					SW			
	A	V	E*	$\Delta\theta_{el}$	$\theta_{cav}$	A	V	$\Delta\theta_{el}$	$\theta_{cav}$
0	172,69	162,65	58,24	-13,23	17,37	173,47	161,63	-13,56	17,44
60	180,06	166,53	2,36	-13,18	18,04	181,49	165,76	-13,84	18,17
120	182,13	166,96	3,70	-13,08	18,23	183,35	166,12	-13,43	18,34
180	182,59	168,05	0,00	-12,51	18,27	183,56	166,50	-13,37	18,35
240	173,73	165,81	20,40	-12,72	17,46	176,54	164,78	-12,98	17,72
300	166,31	159,34	147,07	-12,23	16,79	166,99	159,09	-12,50	16,85



**Gráfica 3.2.3.a** Variación de la energía libre de solvatación en una rotación del  $\text{CH}_3\text{-CH}_2\text{-CH}_2\text{-CH}_2\text{-CH}_2\text{-CH}_2\text{-CH}_2\text{-OH}$ . El conformero utilizado es  $(-90, 90, \beta, -60, 0, 150)$ . Los cálculos se realizaron con una base STO-3G, con una geometría estándar y centrando esferas sobre el  $\text{CH}_3$ ,  $\text{CH}_2$ , O, y H (radios de van Waals)



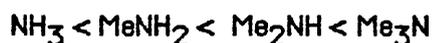
**Figura 3.2.3.a** Dibujo del conformero del octanol utilizado.

### 3.3 BASICIDAD DE LAS AMINAS EN DISOLUCION

Las basicidades del amoniaco y de las metilaminas en disolución acuosa presentan un orden un tanto irregular:



En cambio, en fase gas, siguen el siguiente orden:



El orden de estas basicidades ha dado pie a múltiples trabajos tanto experimentales como teóricos sin llegar a determinar la causa exacta de esta inversión (Arnett, E. M. y col, 1972; Aue, D. H. y col, 1976; Terryn, B. y col, 1981; Mac Dónaill, D. A. y col, 1984; Galera, S. y col, 1984). En este apartado pretendemos analizar el problema con el método de Miertus y col y determinar la importancia del uso de una superficie envolvente adecuada.

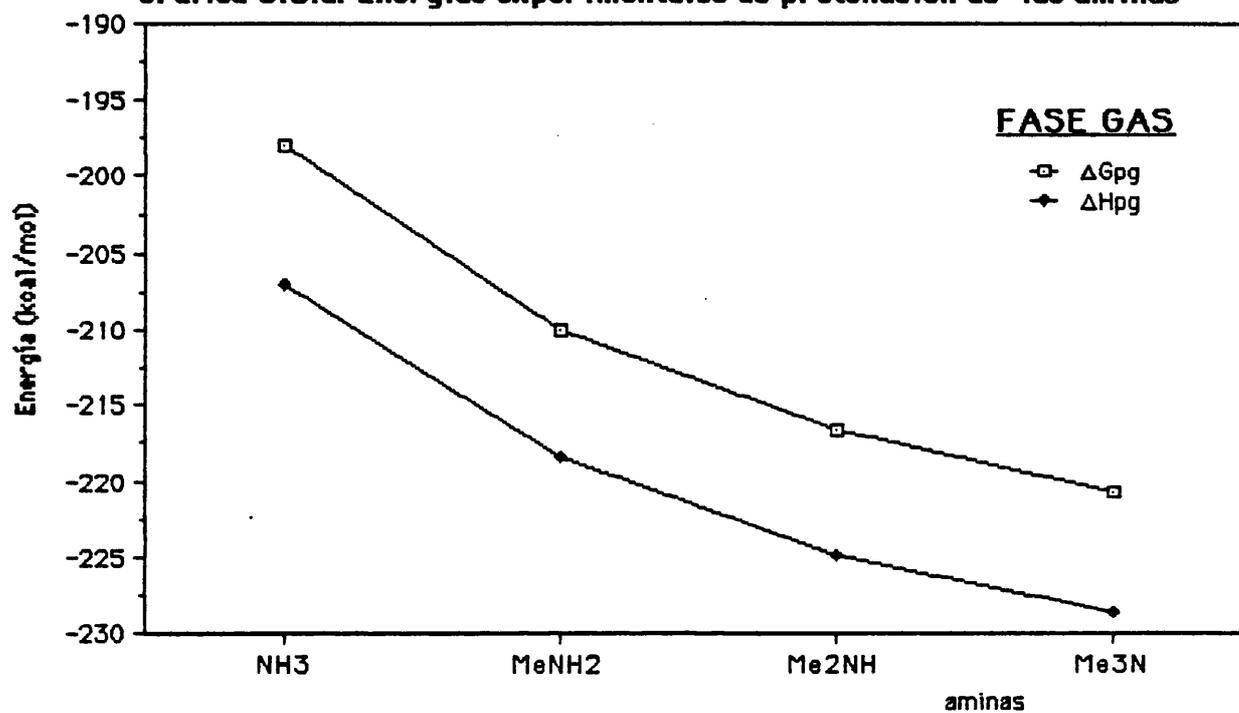
En la **tabla 3.3.a** se dan los valores termodinámicos (Aue, D.H. y col, 1976), que nos servirán de referencia para nuestro estudio, representándose en las **gráficas 3.3.a y b** las energías libres y entalpías de protonación en fase gas y disolución respectivamente.

**Tabla 3.3.a** Valores termodinámicos de las aminas, a 25°, en kcal/mol.  
El valor  $\Delta G_s(\text{H}^+) = -260.5$  kcal/mol se tomó de los mismos autores citados (Aue, D.H., 1976).

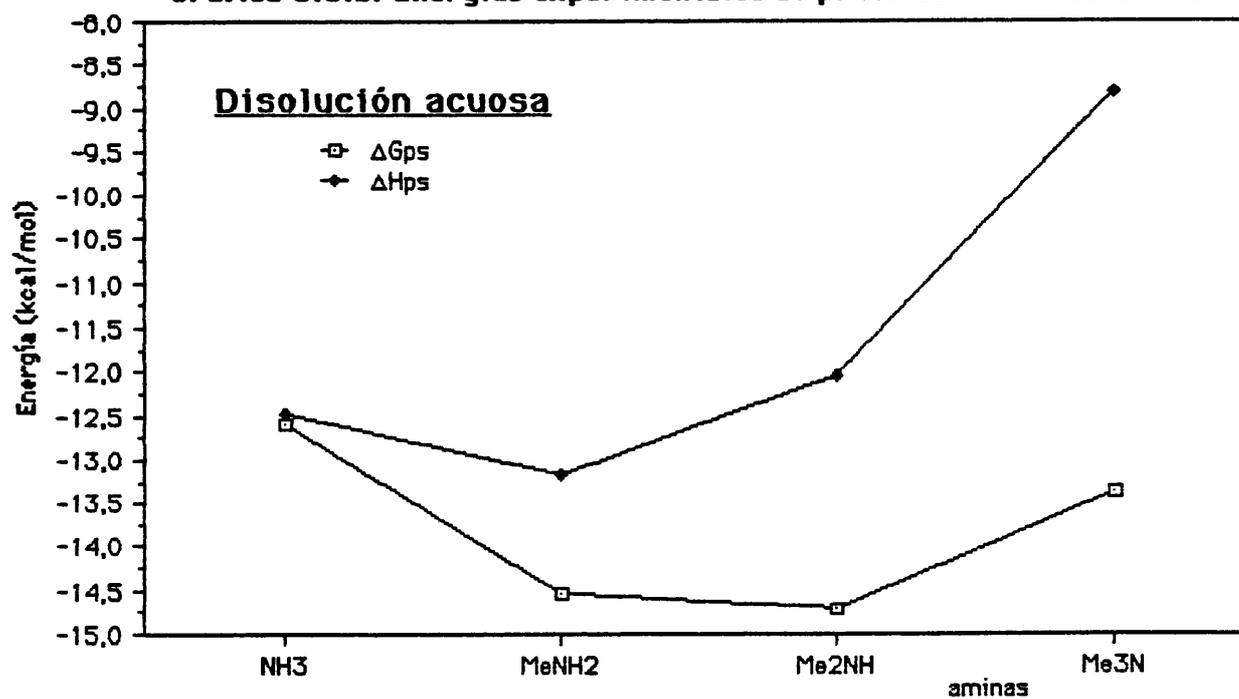
amina	$\Delta G_{pg}$	$\Delta G_{ps}$	$\Delta G_s(B)$	$\Delta G_s(\text{BH}^+)$	$\Delta H_{pg}$	$\Delta H_{ps}$	$-\text{T}\Delta S_{pg}$	$-\text{T}\Delta S_{ps}$
$\text{NH}_3$	-198,	-12,61	-2,41	-78,	-207,	-12,49	8,57	-0,12
$\text{MeNH}_2$	-210,	-14,53	-2,68	-67,7	-218,4	-13,18	8,40	-1,35
$\text{Me}_2\text{NH}$	-216,6	-14,70	-2,41	-61,0	-224,8	-12,04	8,16	-2,66
$\text{Me}_3\text{N}$	-220,8	-13,37	-1,34	-54,4	-228,6	-8,82	7,75	-4,55

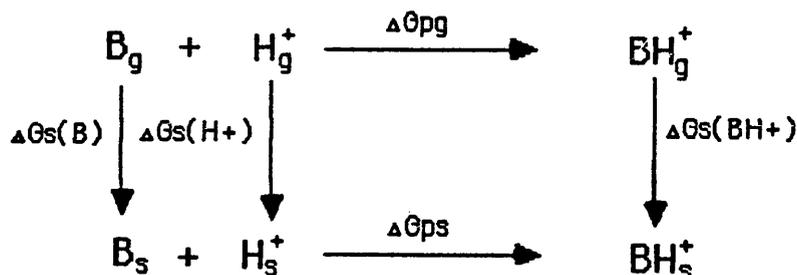
Para el cálculo de la energía libre de protonación en disolución acuosa hemos tenido en cuenta el siguiente ciclo termodinámico:

Gráfica 3.3.a. Energías experimentales de protonación de las aminas



Gráfica 3.3.b. Energías experimentales de protonación de las aminas





resultando que la energía libre de protonación (que es la que nos dará la basicidad) se puede calcular como la suma de los siguientes términos:

$$\Delta G_{ps} = \Delta G_{pg} + \Delta G_s(\text{BH}^+) - \Delta G_s(\text{B}) - \Delta G_s(\text{H}^+) \quad \text{Fórmula 3.3.a}$$

El primer término corresponde a la energía libre de protonación en fase gas y se puede desglosar en una parte entálpica,  $\Delta H_{pg}$ , y otra entropica  $-T\Delta S_{pg}$ . La parte entropica, que no determinaremos teóricamente, no presenta variaciones sustanciales de una amina a otra por lo que no será un factor decisivo en la inversión de la basicidad (ver **tabla 3.3.a**). La parte entálpica se puede obtener como la suma siguiente (**Eades, R.A. y col, 1981**):

$$\Delta H_{pg} = \Delta E_{pg} + \Delta ZPE - (5/2)RT \quad \text{Fórmula 3.3.b}$$

El primer término corresponde a la diferencia de energía interna entre productos y reactivos en su estado fundamental, y corresponde al valor que podemos calcular con los métodos normales de química cuántica. El segundo término es la diferencia de energías en el punto cero entre las especies protonadas y neutras. **Eades y col, 1981** han calculado este término para la protonación del amoníaco resultando de 7.86 kcal/mol. Como se puede observar, este término no se puede despreciar en la determinación de las basicidades absolutas. Sin embargo, es razonable pensar que el resto de las aminas van a tener valores similares, por lo que no influirá decisivamente en el

orden de basicidades. El último término corresponde al balance energético del cambio del número de grados de libertad, que será el mismo para todas las aminas. Así pues, en nuestros cálculos consideraremos  $\Delta H_{pg} = \Delta E_{pg}$ .

El segundo y tercer término de la **Fórmula 3.3.a** corresponden a las energías libres de disolución de las moléculas protonada y neutra. Estos términos se calculan utilizando el formalismo explicado en el apartado 3.1 y están formados por la componente electrostática,  $\Delta G_{se}$ , y por la de cavitación,  $\Delta G_{sc}$ , que incluyen tanto la parte entálpica como la entrópica.

El último término de la **Fórmula 3.3.a** corresponde a la energía libre de disolución del protón y su valor lo hemos tomado del trabajo de **Aue, D.H. y col, 1976b**. Este valor, al ser constante, no afectará al orden de basicidades, sin embargo, lo hemos incluido para trabajar con valores absolutos cercano a los reales.

Así, la energía que calculamos está formada por los siguientes factores:

$$\Delta G_{ps} = \Delta E_{pg} - T\Delta S_{pg} + \Delta G_{sc}(BH^+) + \Delta G_{se}(BH^+) - \Delta G_{sc}(B) - \Delta G_{se}(B) - \Delta G_s(H^+)$$

### Fórmula 3.3.c

Los cinco últimos términos en algunos casos los reuniremos, por comodidad, bajo la denominación de energía libre de disolución total,  $\Delta G_s(T)$ .

Antes de pasar a analizar los resultados teóricos, haremos dos comentarios sobre los datos experimentales. El primero es indicar como el disolvente, además de invertir el orden de las basicidades, amortigua la diferencia entre éstas, ya que las diferencias entre las aminas más extremas pasan de ser de aproximadamente 25 a 2 kcal/mol. El segundo es que la inversión se produce tanto en la energía libre como en la entalpía, siendo mayor en la última. Esto es debido a que los términos entrópicos juegan un papel contrario a la inversión.

Los cálculos sobre las aminas se realizaron con una base 4-31G, efectuando una optimización total en el vacío (ver nota en el apartado 2.6.3.). Para el cálculo de solvatación se centraron esferas sobre todos los átomos, con radios de  $R_H = 1.44$ ,  $R_C = 1.92$  y  $R_N = 1.86$  Å, que

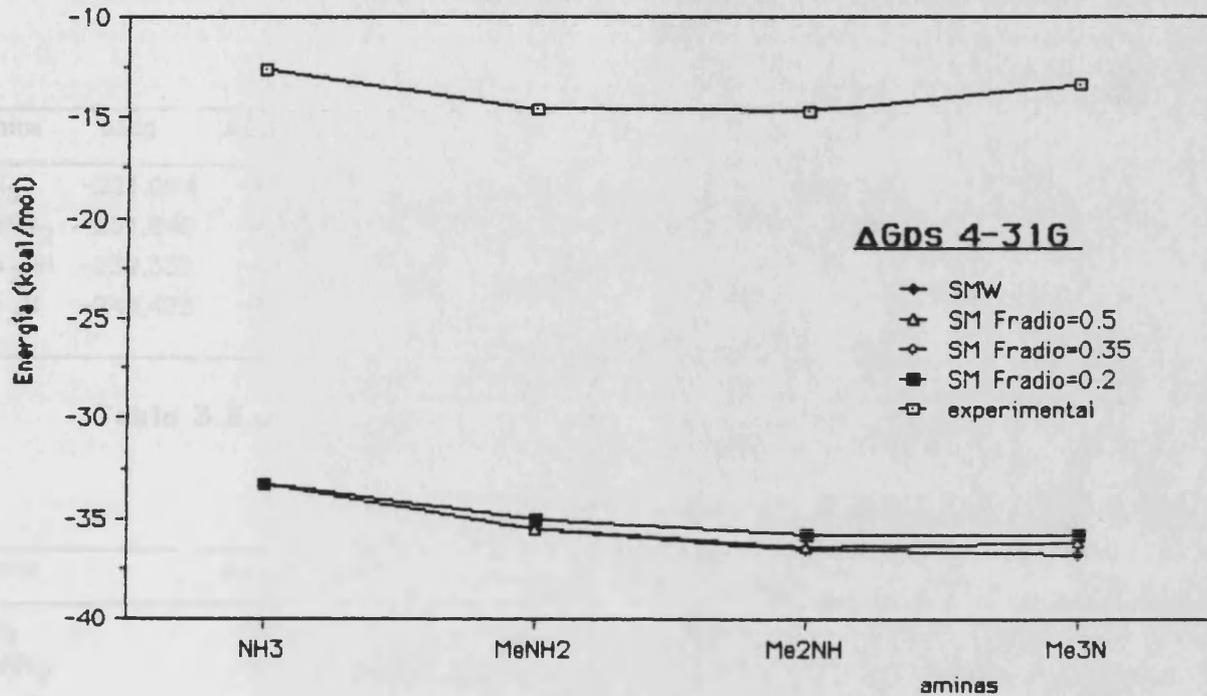
corresponden a los radios de van der Waals más un 20% (Miartus, S. y col, 1981). Para la constante dieléctrica se utilizó un valor de 78.3 que corresponde a la del agua a 298.15°K, utilizándose en el cálculo de la superficie NDIV=5. Todos los cálculos se realizaron para la SMW y para la SM, usando para esta última un radio de prueba para el agua de 1.5 Å y valores de FRADIO de 0.5, 0.35 y 0.2 (FRADIO controla la creación de esferas que nos dirige a la obtención de la SM)

En la **tabla 3.3.b** damos los valores termodinámicos obtenidos cuando se realiza el cálculo con la SMW y con una SM con un valor de FRADIO de 0.2, y en la **tabla 3.3.c** los valores correspondientes con una SM con FRADIO 0.5 y 0.35. Con estos valores y la **Fórmula 3.3.c** obtuvimos las energías libres de protonación en fase acuosa que, junto con el correspondiente valor experimental, representamos en las **gráficas 3.3.c y d**, correspondiendo la primera a los valores absolutos y la segunda relativos al  $\text{NH}_3$ .

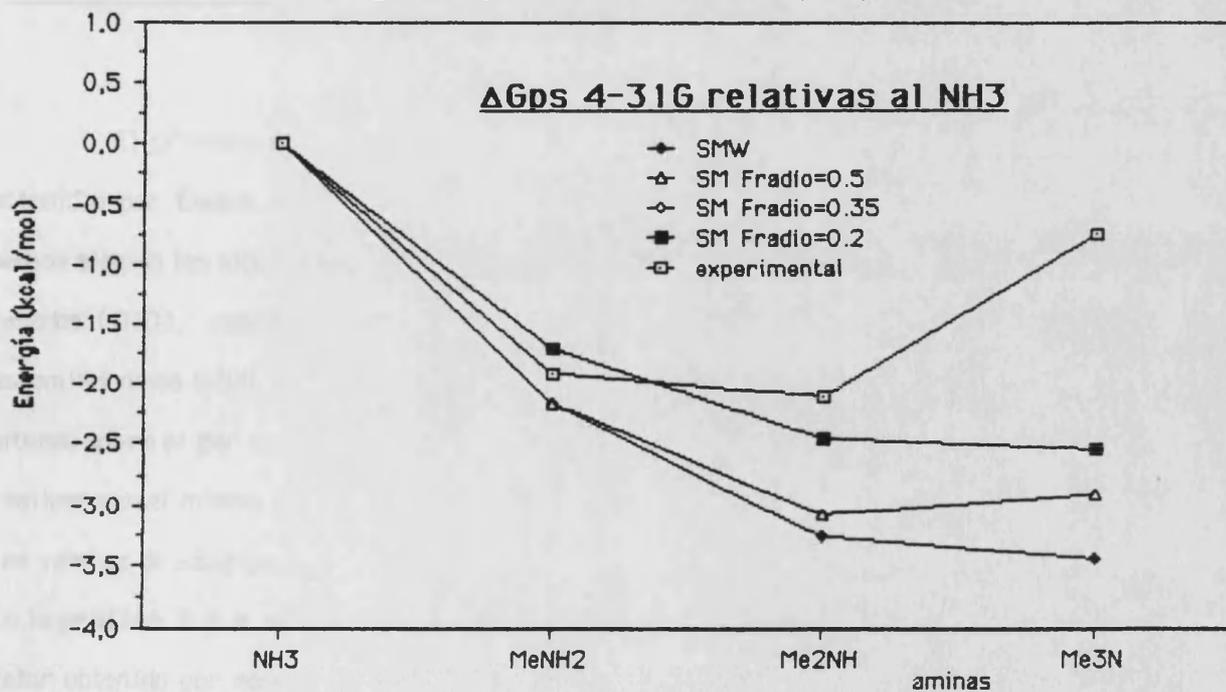
De la observación de las tablas y gráficas anteriores se puede ver que, las energías absolutas calculadas difieren en aproximadamente 20 kcal/mol de las experimentales. Esta diferencia viene determinada principalmente por el valor asignado a  $\Delta\text{Hpg}$ , ya que consideraremos que  $\Delta\text{Hpg}=\Delta\text{Epg}$  eliminando los dos últimos términos de la **Fórmula 3.3.b**. Además, como veremos unos párrafos más abajo, la función de onda utilizada no reproduce adecuadamente los valores absolutos de  $\Delta\text{Epg}$ . Sin embargo, nuestro estudio se dirige más al análisis de los valores relativos de la basicidad, así en la **gráfica 3.3.d** podemos ver, en primer término, que los valores relativos de  $\Delta\text{Gps}$  calculados, aún teniendo unos valores del orden de los experimentales, no siguen su mismo orden. Otro aspecto interesante es ver la influencia de la superficie utilizada en el valor de las basicidades obtenidas. Así, al pasar de la SMW a la SM los valores de las energías calculadas se acercan a los valores experimentales, confirmando la conveniencia de usar unas superficies envolventes lo más adecuadas. También es interesante observar que los valores obtenidos con FRADIO igual a 0.35 y 0.2 coinciden, con lo que podemos decir que estamos en el límite de la corrección que puede introducir la SM.

Sin embargo, como hemos indicado y como se puede observar, los valores no siguen el mismo orden experimental. Para analizar las causas de esta diferencia tomamos como base la

Gráfica 3.3.c. Energías de protonación teóricas y experimental en disolución.



Gráfica 3.3.d. Energías de protonación teóricas y experimental en disolución



Fórmula 3.3.c e indagamos en cada uno de los términos que la componen.

**Tabla 3.3.b** Valores termodinámicos de las aminas a 25° en kcal/mol, calculados por nosotros usando la SMW y la SM con FRADIO igual 0.2.

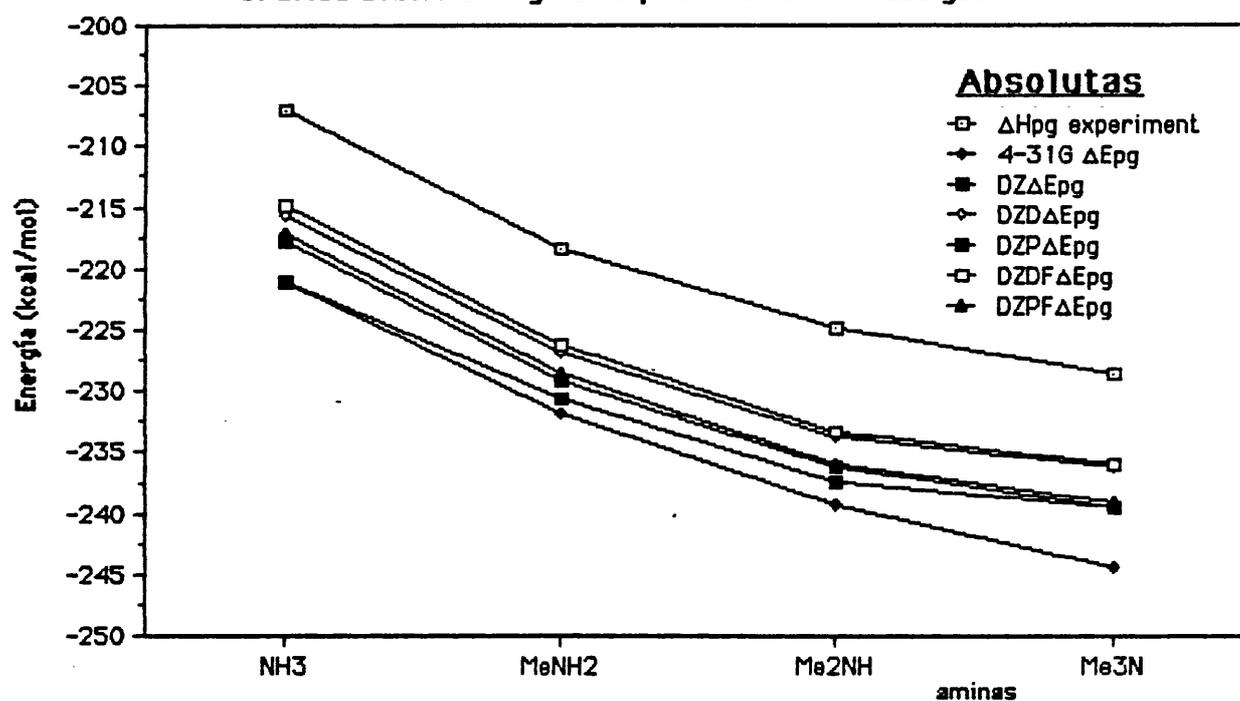
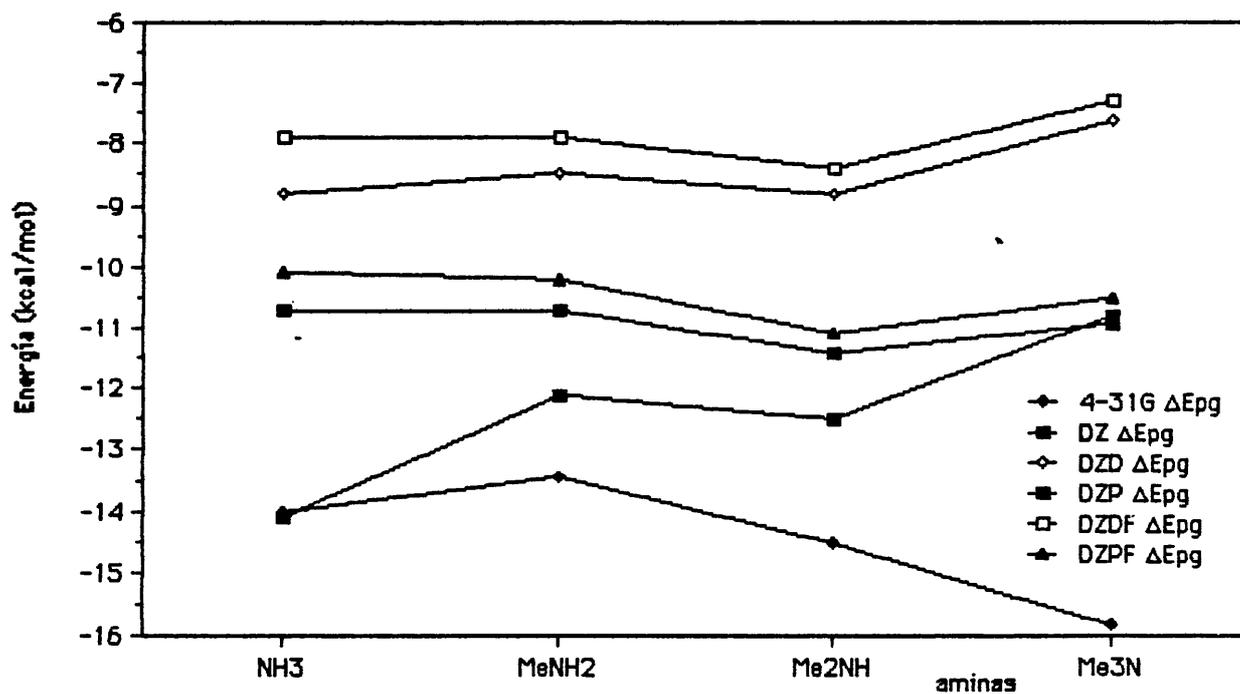
amina	$\Delta E_{pg}$	SMW				SM FRADIO=0.2			
		$\Delta G_{se}(B)$	$\Delta G_{sc}(B)$	$\Delta G_{se}(BH+)$	$\Delta G_{sc}(BH+)$	$\Delta G_{se}(B)$	$\Delta G_{sc}(B)$	$\Delta G_{se}(BH+)$	$\Delta G_{sc}(BH+)$
NH <sub>3</sub>	-221,004	-4,703	6,294	-77,904	6,691	-4,703	6,294	-77,904	6,691
MeNH <sub>2</sub>	-231,840	-4,610	8,775	-69,145	9,172	-4,572	8,737	-68,598	9,093
Me <sub>2</sub> NH	-239,332	-4,668	11,014	-62,747	11,358	-4,552	10,836	-61,757	11,116
Me <sub>3</sub> N	-244,423	-4,633	13,024	-57,768	13,325	-4,257	12,611	-56,394	12,834

**Tabla 3.3.c** Valores termodinámicos de las aminas a 25° en kcal/mol, calculados por nosotros usando la SM con FRADIO igual a 0.5 y 0.35.

amina	SM FRADIO= 0.5				SM FRADIO=0.35			
	$\Delta G_{se}(B)$	$\Delta G_{sc}(B)$	$\Delta G_{se}(BH+)$	$\Delta G_{sc}(BH+)$	$\Delta G_{se}(B)$	$\Delta G_{sc}(B)$	$\Delta G_{se}(BH+)$	$\Delta G_{sc}(BH+)$
NH <sub>3</sub>	-4,703	6,294	-77,904	6,691	-4,703	6,294	-77,904	6,691
MeNH <sub>2</sub>	-4,610	8,775	-69,145	9,172	-4,572	8,737	-68,598	9,093
Me <sub>2</sub> NH	-4,597	10,889	-62,472	11,227	-4,553	10,840	-61,769	11,125
Me <sub>3</sub> N	-4,350	12,634	-56,923	12,924	-4,257	12,616	-56,411	12,846

El primero de ellos es la  $\Delta E_{pg}$ . Para ello comparamos nuestro resultado 4-310 con los obtenidos por Eades, R.A. y col en 1981. De las bases que utilizaron estos autores nosotros hemos elegido las siguientes: doble  $\zeta$  (DZ), doble  $\zeta$  más funciones de polarización sobre los átomos pesados (DZD), doble  $\zeta$  más funciones de polarización sobre todos los átomos (DZP) y las que denominaremos DZDF y DZPF, que corresponden a las DZD y DZP más un "ghost orbital" que iría situado sobre el par solitario del nitrógeno en las especies neutras, de tal forma que los cálculos se realicen con el mismo número de funciones, tanto en las estructuras neutras como en las protonadas. Los valores de  $\Delta E_{pg}$  que obtuvieron dichos autores con estas bases se encuentran en la tabla 3.3.d. En la gráfica 3.3.e se han representado los valores absolutos de  $\Delta E_{pg}$  para estas bases junto con el valor obtenido por nosotros con la 4-310 y el valor experimental de  $\Delta H_{pg}$ . Con el fin de analizar

Gráfica 3.3.e. Energías de protonación en fase gas.

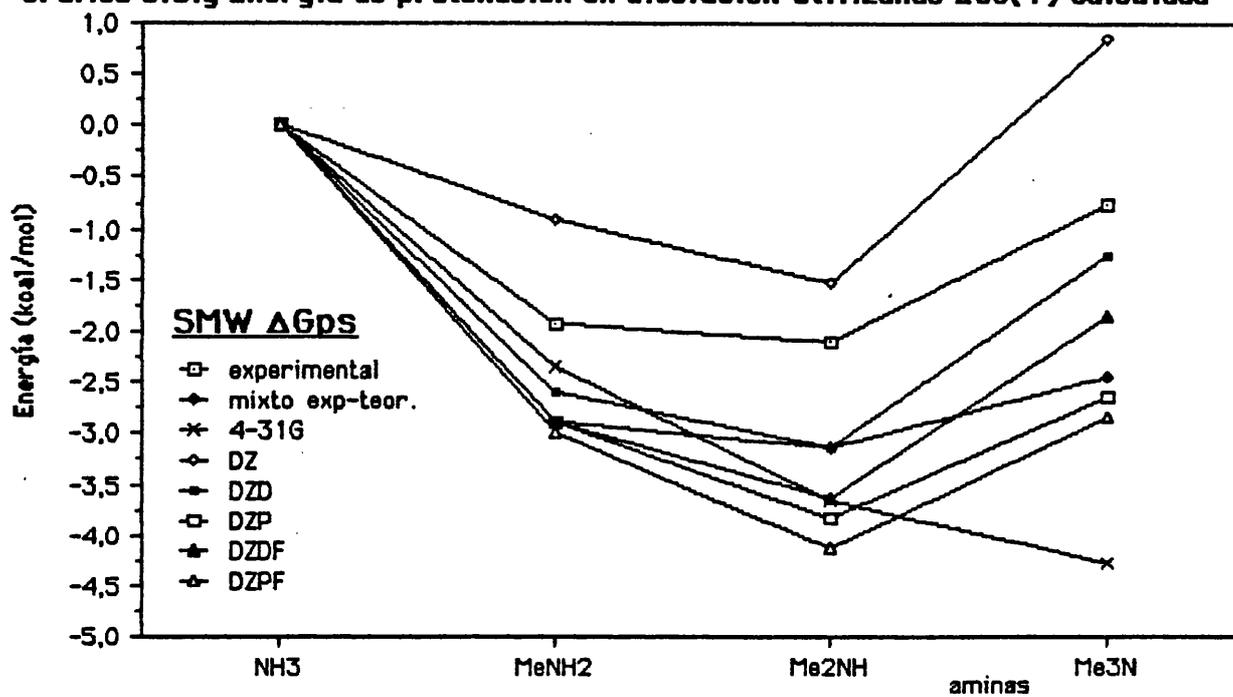
Gráfica 3.3.f. Valores de  $\Delta E_{pg}$ (teórico)- $\Delta H_{pg}$ (experimental) para varias bases

mejor las diferencias de una base a otra representamos en la **gráfica 3.3.f**, los valores correspondientes a las diferencias entre  $\Delta E_{pg}$ (teórico) y  $\Delta H_{pg}$  (experimental). Lo primero que podemos decir es que las bases 4-310 y DZ son las que dan peor resultado tanto a nivel absoluto como relativo. En este momento queremos indicar que el haber trabajado con una base 4-310 se debe a que el programa de solvatación que poseemos solo permite utilizar bases estandar, siendo la 4-310 de la más usadas actualmente y por tanto da mayor fiabilidad. También es interesante destacar que la introducción de orbitales fantasmas no introduce cambios sustanciales en las energías relativas. Por último indicar que las diferencias entre la energía obtenida con la base DZDF y la entalpía experimental son del mismo orden que los términos despreciados, al considerar que  $\Delta H_{pg} = \Delta E_{pg}$ , correspondiendo el resto de la diferencia a que evidentemente la base no es perfecta y que no se ha considerado la energía de correlación.

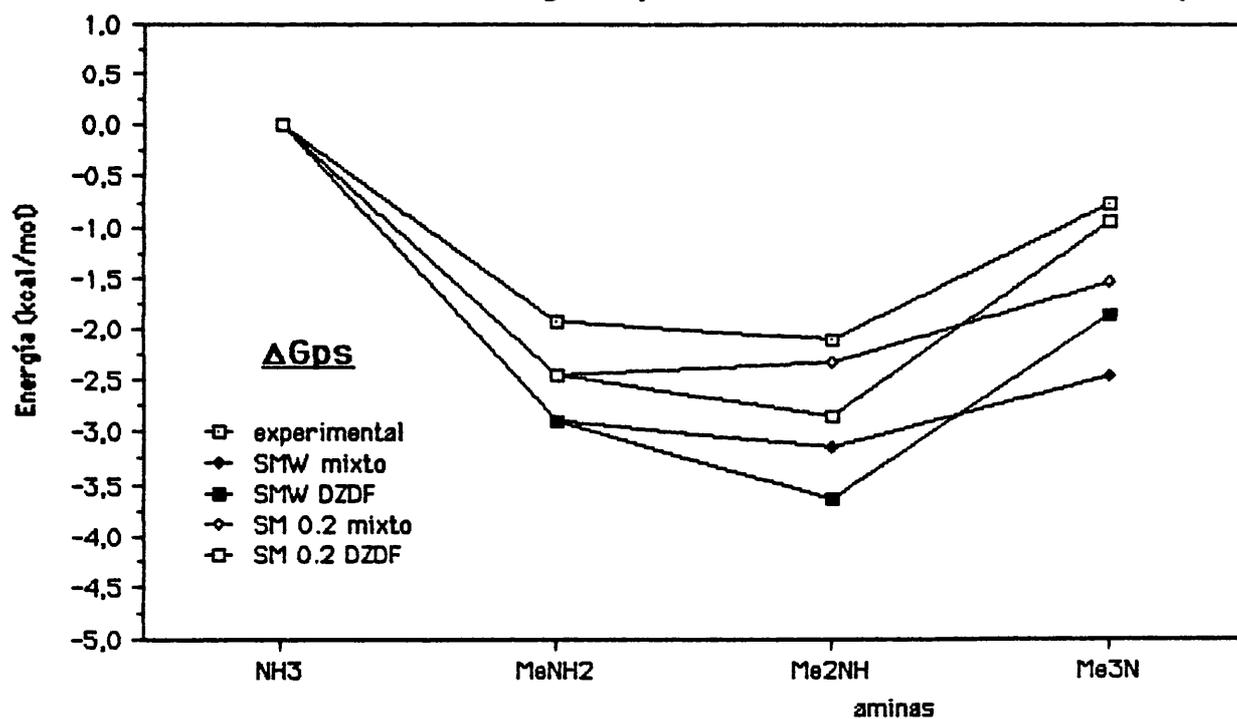
**Tabla 3.3.d** Valores de  $\Delta E_{pg}$  obtenidos por Eades, R.A. y col. 1981.  
En kcal/mol

amina	DZ	DZD	DZP	DZDF	DZPF
NH <sub>3</sub>	-221,1	-215,8	-217,7	-214,9	-217,1
MeNH <sub>2</sub>	-230,5	-226,9	-229,1	-226,3	-228,6
Me <sub>2</sub> NH	-237,3	-233,6	-236,2	-233,2	-235,9
Me <sub>3</sub> N	-239,4	-236,2	-239,5	-235,9	-239,1

Para ver el efecto que produciría el uso de estas bases en el cálculo de la energía de solvatación realizamos la siguiente operación: Calculamos la energía de protonación en disolución utilizando los valores termodinámicos de disolución obtenidos por nosotros usando la SMW y los diferentes valores  $\Delta E_{pg}$  obtenidos con las bases anteriores, utilizando también para uno de estos cálculos el valor experimental de  $\Delta H_{pg}$ . En la **gráfica 3.3.g** se muestran estos resultados, donde aparece también el valor experimental de  $\Delta \theta_{ps}$ . Como se puede observar, aparece la inversión, por lo menos a nivel cualitativo, con casi todas las bases. Este resultado nos muestra la importancia de la selección de la base.

Gráfica 3.3.g Energía de protonación en disolución utilizando  $\Delta G_s(T)$  calculada

Gráfica 3.3.h. Variación de la energía de protonación en disolución con la superficie.



Para analizar la influencia de la superficie envolvente seleccionada sobre la energía de protonación hemos calculado esta utilizando las  $\Delta E_{pg}$  obtenidas con la base DZDF y con los valores termodinámicos obtenidos por nosotros para la SMW y para la SM con  $FRADIO=0.2$ . También hemos hecho un tanto de lo mismo con la  $\Delta H_{pg}$ . Los resultados se muestran en la **gráfica 3.3.h** junto al valor experimental de  $\Delta G_{ps}$ . Como se puede apreciar la influencia de la superficie molecular es decisiva acercando los valores a los experimentales.

En la **gráfica 3.3.h** podemos ver que el modelo de solvatación utilizado, aun describiendo la inversión de la basicidad, difiere de los valores experimentales. Estas diferencias están causadas por los siguientes aspectos:

- Los valores termodinámicos de disolución se han obtenido utilizando una función de onda 4-31G, que, como hemos visto, no es la más adecuada para este caso y por tanto la utilización de una base DZDF presumiblemente mejoraría los resultados.

- El modelo de solvatación utilizado, por considerar al disolvente como un continuo no tiene en cuenta interacciones del tipo discreto como son las de puente de hidrógeno.

- El método de cálculo de la energía de cavitación utilizado considera al soluto como una esfera rígida y, después de vista la importancia de la definición de la superficie molecular, resulta contradictoria su utilización.

- Dentro del modelo de continuo se pueden considerar otras energías que no hemos tenido en cuenta en este trabajo, como son la energía de dispersión y las de tipo repulsivo.

Sin embargo, aun con todas estas aproximaciones, hemos podido ver que el modelo es capaz de describir el proceso de inversión. Además de todo lo dicho hasta ahora se pueden extraer tres conclusiones importantes:

- El método de solvatación utilizado reproduce, al menos a un nivel cualitativo, el orden irregular de la basicidad de las aminas en disolución.

- Aunque el disolvente pone de manifiesto las irregularidades en el orden de la basicidad de las aminas, en contra de lo admitido de forma general (Aue, D. H. y col, 1976), no son sólo las energías de solvatación las responsables de la inversión de las basicidades. Esto ha quedado puesto de

manifiesto ya que utilizando los mismos valores de solvatación, la inversión dependía de la  $\Delta E_{pg}$  utilizada.

- La superficie envolvente de una molécula juega un papel importantísimo en la interacción con el disolvente, describiendo mejor esta interacción las SM que la SMW.

## **4.CONCLUSIONES**

1. Hemos construido un método de cálculo de superficies moleculares aplicable a cualquier estructura y que nos permite conocer:
  - a. -La forma de la superficie.
  - b. -El área total de la superficie.
  - c. -El volumen total encerrado.
  - d. -El área sobre cada átomo.
  
2. El método es aplicable a las:
  - a. -Superficie Molecular de van der Waals.
  - b. -Superficie Accesible.
  - c. -Superficie Molecular.
  
3. Entre las cualidades del método se encuentran:
  - a. -Su fácil utilización.
  - b. -Un error que no supera el 1,5% en los casos más extremos.
  - c. -Una buena invariancia espacial.
  - d. -Una variación predecible del resultado con sus parámetros.
  - e. -Una reproducción fiel de la superficie.
  - f. -Un tiempo de cálculo proporcionado.
  - g. -Que reproduce sin irregularidades los pequeños cambios de la superficie con la variación de las coordenadas atómicas.
  - h. -Puede ser aplicado en múltiples problemas de interés químico
  
4. Hemos utilizado con éxito GEPOL en la parte geométrica del método de solvatación de Miertus y col, siendo las principales conclusiones:
  - a. -El método GEPOL mejora los resultados obtenidos con el SPHERE.
  - b. -Queda demostrada la diferencia de resultados de usar la SM o la SMW para representar la interfase soluto-disolvente, resultando los valores mejores con la SM.
  - c. -El método de solvatación de Miertus y col. describe bien a nivel cualitativo el orden irregular de las basicidades de las aminas en disolución, aunque el orden irregular de estas basicidades no solo viene determinado por la interacción con el disolvente sino por las características intrínsecas de las moléculas.

## 5.BIBLIOGRAFIA

- Alagona, G., R. Bonaccorsi, C. Ghio y J. Tomasi (1986a) *J. Mol. Struct. (Theochem)*, **135**, 39
- Alagona, G., R. Bonaccorsi, C. Ghio y J. Tomasi (1986b) *J. Mol. Struct. (Theochem)*, **137**, 263
- Åquist, J., W. F. van Gunsteren, M. Leijonmarck y O. Tapia (1985). *J. Mol. Biol.* **183**, 461.
- Åquist, J., P. Sandblom, T. A. Jones, M. E. Newcomer, W. F. van Gunsteren, y O. Tapia (1986). *J. Mol. Biol.* **192**, 593.
- Åquist, J. y O. Tapia (1987) *J. Mol. Graphics*, **5**, 30.
- Arnett, E. M., F. M. Jones, M. Taagepera, W. G. Henderson, J. L. Beauchamp, D. Holtz y R.W. Taft (1972) *J. Am. Chem. Soc.*, **94**, 4724
- Aue D. H., H. M. Webb y M. T. Bowers (1976a) *J. Am. Chem. Soc.* **98**, 311
- Aue D. H., H. M. Webb y M. T. Bowers (1976b) *J. Am. Chem. Soc.* **98**, 318
- Aulló, J. M. y O. Tapia (1987). Comunicación personal.
- Bonaccorsi, R., C. Ghio y J. Tomasi (1982). *Studies in Physics and Theoretic Chemistry*, R. Carbo, ed., Elsevier, Vol. 21, pag. 407
- Bonaccorsi, R., R. Cimiraglia, S. Miertus y J. Tomasi (1983a). *J. Mol. Struct. (Theochem)* **94**, 11
- Bonaccorsi, R., R. Cimiraglia y J. Tomasi (1983b). *Chem. Phys. Letters*, **99**, 77
- Bonaccorsi, R., R. Cimiraglia, P. Palla y J. Tomasi (1984a). *J. Am. Chem. Soc.*, **106**, 1945.
- Bonaccorsi, R., R. Cimiraglia y J. Tomasi (1984b). *J. Mol. Struct. (Theochem)*, **107**, 197
- Bonaccorsi, R., C. Ghio y J. Tomasi (1984c). *Int. J. Quant. Chem.*, **26**, 637
- Bonaccorsi, R., E. Scrocco y J. Tomasi (1985) *J. Biosc.*, **8**, 627
- Bonaccorsi, R., E. Scrocco y J. Tomasi (1986). *Int. J. Quant. Chem.*, **29**, 717
- Bränden, C.I., H. Jörnvall, H. Eklund y B. Furugren (1975). *Alcohol dehydrogenase*, in the *Enzymes*, vol 11, 3<sup>rd</sup> ed., Boyer, P.D., Ed., Academic Press, New York, 104.
- Campos, A. (1988) Comunicación personal.
- Cimiraglia, R., S. Miertus y J. Tomasi (1981). *Chem. Phys. Letters*, **80**, 286
- Clementi, E. y C. Corongiu (1980). *J. Chem. Phys.* **72**, 3979.

- Connolly, M. L. (1983a). *Science*, **221**, 709.
- Connolly, M. L. (1983b). *J. Appl. Cryst.* **16**, 548.
- Coexter, H. S. M. (1974), *Regular Complex Polytopes*, Cambridge University Press, Cambridge, England.
- Ditchfield, R., W. J. Hehre y J. A. Pople (1971) *J. Chem. Phys.* **54**, 724.
- Duben, A. y S. Miertus (1982) *Chem. Phys. Letters*, **88**, 395
- Davidon, N. C. (1975) *Math. Prog.*, **9**, 1.
- Eades, R. A., D. A. Weil, D. A. Dixon y C. H. Douglass (1981). *J. Phys. Chem.*, **85**, 981
- Galera, S., A. Oliva, J. M. LLuch y J. Bertrán (1984) *J. Mol. Struct. (Theochem)*, **110**, 15
- Gavilan, J. C. (1987). Comunicación personal.
- Giessner-Prettre, C. y A. Pullman (1972) *Theoret. Chim. Acta*, **25**, 83
- Hermann, R. B. (1972). *J. Phys. Chem.* **76**, 2754
- Horta, A. (1979) *Macromolecules*, **12**, 785
- Horta, A. y I. Fernandez Pierola (1981) *Macromolecules*, **14**, 152
- Horta, A. (1985) *Macromolecules*, **18**, 2498
- Langmuir, I. (1925). *Colloid. Symp. Monogr.*, **3**, 3
- Lee, B. y F. M. Richards (1971). *J. Mol. Biol.*, **55**, 379.
- Leijonmarck, M., S. Eriksson y A. Lilijas (1980). *Nature (London)*, **286**, 824.
- Lewis, M y D. C. Rees (1985) *Science*, **230**, 1163.
- Mac Dónaill, D. A., D. A. Morton-Blake (1984) *Theoret. Chim. Acta*, **65**, 13
- Mandelbrot, B.B. (1983) *The Fractal Geometry of Nature*, Freeman, San Francisco.
- Miertus, S., E. Scrocco y J. Tomasi (1981). *Chem. Phys.*, **55**, 117.
- Miertus, S. y J. Tomasi (1982) *Chem. Phys.*, **65**, 239
- Möller, W. (1974). In *Ribosomes* (Nomura, M., Tissières, A. y Lengyel, P., eds), pp. 711-731, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Pauling, L. (1948). *Nature*, **161**, 707
- Persico, M. y J. Tomasi (1984). *Croat. Chem. Acta*, **57**, 1395

- Peterson, M. R. y R. A. Poirier (1980) Universidad de Toronto, Canada.
- Petterson, I. y C. G. Kurland (1980). Proc. Nat. Acad. Sci., U.S.A. **77**,4007.
- Pierotti, R. A. (1976). Chem. Rev. **76**,717
- Pople, J. A. y D. L. Beveridge (1970) Aproximate Molecular Orbital Theory. McGraw-Hill, New York.
- Pouchly, J. y D. Patterson (1976). Macromolecules, **9**,574
- Pullman, B. (1976) ed., Environmental effects on molecular structure and properties (Reidel, Dordrecht)
- Pullman, B., S. Miertus y C. Perahia (1979) Theoret. Chim. Acta, **50**,317
- Richards, F. M. (1977). Ann. Rev. Biophys. Bioeng., **6**,151
- Scrocco, E. y J. Tomasi (1973) Topics in current Chemistry, Vol 42. Springer-Verlag
- Sinanoglu, O. (1980) Int. J. Quantum Chem. **18**, 381
- Terry, B. y J. L. Rivail (1981) J. Chem. Research(M). 1819
- Tapia, O. (1982) en Molecular Interactions Vol 3, H. Ratajczak y W.J. Orwille-Thomas, Eds., Wiley, New York, pag 47.
- Tapia, O., H. Eklund y C.I. Bränden (1986). Esteric aspects of biomolecular interactions. Náray-Szabo, G., y Simin, K., Ed.. CRC Press, Florida, USA.
- Valvani, S. C., S.H. Yalkowsky y G.L. Amidon (1976) J. Phys. Chem. **80**,829

**ORDENADORES UTILIZADOS**

- I** GOULD 32/8705 del Departamento de Química de la Universidad de Pisa, Italia.
- II** BULL, DPS 8/52 de la Universidad de Valencia.
- III** VAX 8300 del Departamento de Química Física de la Universidad de Valencia.
- IV** VAX 750 del Departamento de Biología Molecular del Centro Biomedico de Uppsala, Suecia.
- V** MICRO VAX II del Departamento de Biología Molecular del Centro Biomedico de Uppsala, Suecia.
- VI** IBM 43/61 Modelo 4 del Departamento de Biología Molecular del Centro Biomedico de Uppsala, Suecia.

Reunido el Tribunal que suscribe, en el día de la fecha, acordó otorgar, por unanimidad, a esta Tesis doctoral de D. JUAN-LUIS PASCUAL-AHUIR GIMEX

La calificación de APTO CVM LAUDE

Valencia, a 18 de ABRIL de 1988

El Secretario,

El Presidente



A handwritten signature in blue ink, appearing to read "Juan L. A.", written over a horizontal line.