

UCH-FQF  
DOC-62  
A478  
C.1

# UNIVERSIDAD DE CHILE



## ESTUDIOS COMPUTACIONALES DE LA INTERACCIÓN PROTEINA-LIGANDO EN EL SISTEMA CDK2/CICLINA A

Tesis entregada a la Universidad de Chile para optar al grado de  
Doctor en Química

Por:

**JANS HUMBERTO ALZATE MORALES**

DIRECTORES DE TESIS

**Dr. Renato Contreras R.**

Departamento de Química  
Facultad de Ciencias  
Universidad de Chile

**Dr. Bruce Cassels**

Departamento de Química  
Facultad de Ciencias  
Universidad de Chile

Santiago, Chile  
Diciembre, 2006

**UNIVERSIDAD DE CHILE**  
**INFORME DE APROBACIÓN**  
**TESIS DE DOCTORADO**

Se informa a la Escuela de Graduados de la Facultad de Ciencias Químicas y Farmacéuticas que la Tesis de Doctorado presentada por el candidato

**JANS HUMBERTO ALZATE MORALES**

Ha sido aprobada por la Comisión Evaluadora de la Tesis como requisito para optar al grado de Doctor en Química, en el examen de Defensa de Tesis rendido el día 29 de diciembre de 2006.

Directores de Tesis

  
\_\_\_\_\_  
**Dr. RENATO CONTRERAS**

  
\_\_\_\_\_  
**Dr. BRUCE CASSELS**

**Comisión Evaluadora de la Tesis**  
**Dr. CLAUDIO OLEA (Presidente)**

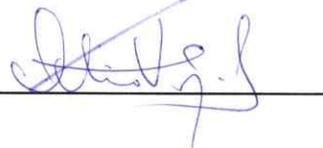
**Dra. MARÍA CECILIA ROJAS**

**Dra. PATRICIA PÉREZ**

**Dr. MARCOS CAROLI**

**Dr. OCTAVIO VÁSQUEZ**

  
\_\_\_\_\_  
  
\_\_\_\_\_  
  
\_\_\_\_\_

  
\_\_\_\_\_  
  
\_\_\_\_\_

## DEDICATORIA

*"A Maribel por su gran cariño y comprensión durante toda esta etapa de mi vida y porque siempre ha estado a mi lado, no importando a donde vayamos."*

*"A mis padres, hermana y sobrina por todo el amor que me han dado, por su apoyo a la distancia y porque han sido el motor emocional para realizar muchos de mis sueños."*

## AGRADECIMIENTOS

Quiero agradecer de todo corazón a los Profesores Renato Contreras y Bruce Cassels por su ayuda incondicional, por su paciencia, disposición incansable, apoyo continuo y amistad durante todo el desarrollo de la tesis.

A los Profesores Joel Ireta y Matthias Scheffler del Departamento de Teoría del Fritz Haber Institut der Max Planck por su colaboración y enseñanzas durante la estadía realizada.

A los Profesores Estanislao Silla e Iñaki Tuñon del Grupo "Efectos del medio" de la Universidad de Valencia, por su hospitalidad, colaboración y generosidad durante mi estadía.

A la comisión evaluadora de la tesis por su continua retroalimentación y valiosas críticas y sugerencias aportadas a esta tesis.

Quiero dar mis más sinceros agradecimientos al gobierno Alemán por la beca para realizar estudios de doctorado en la Universidad de Chile, otorgada a través del Servicio Alemán de Intercambio Académico (DAAD).

Al Departamento de Postgrado y Postítulo de la Universidad de Chile por la beca de financiamiento parcial de tesis PG/95/2004.

Al Núcleo Milenio de Mecánica Cuántica Aplicada y Química Computacional, beca P02-004-F, Mideplan-Conicyt.

Finalmente agradezco a todos los compañeros que hicieron y que hacen parte del Grupo de Química Teórica y que me han brindado su apoyo, moral y académico, para avanzar hasta el final y con los que además pude compartir grandes e inolvidables momentos.

# TABLA DE CONTENIDO

<b>RESUMEN .....</b>	<b>XV</b>
<b>ABSTRACT.....</b>	<b>XVI</b>
<b>CAPÍTULO I.....</b>	<b>1</b>
I.1. INTRODUCCIÓN.....	1
I.2. BIBLIOGRAFÍA .....	7
<b>CAPÍTULO II.....</b>	<b>10</b>
<b><u>ESTUDIO DE LAS INTERACCIONES DE PUENTE DE HIDRÓGENO ENTRE DIFERENTES LIGANDOS Y LA REGIÓN BISAGRA DE LA CDK2 CON EL MÉTODO ONIOM .....</u></b>	<b>10</b>
II.1 INTRODUCCIÓN .....	10
II.2. MODELOS Y MÉTODOS.....	13
II.2.1 Sistemas estudiados.....	13
II.2.2 Cálculos ONIOM .....	13
II.2.3 Cálculo de los índices de reactividad global. ....	15
II.3. RESULTADOS Y DISCUSIÓN. ....	16
II.3.1. Aplicabilidad del método ONIOM a la estructura del compuesto NU2058 .	16
II.3.2 Estructura y energía de interacción del complejo entre el compuesto NU2058 y la región bisagra de la enzima CDK2. ....	17
II.3.3 El complejo Inhibidor-Leu83 .....	23
II.3.4 Geometría de los puentes de hidrógeno.....	28
II.3.5 Cálculo de los índices de reactividad global. ....	28
II.4. CONCLUSIONES.....	30
II.5. BIBLIOGRAFÍA.....	34
<b>CAPÍTULO III.....</b>	<b>37</b>
<b><u>LA ENERGÍA DE INTERACCIÓN QM/MM COMO ÍNDICE DE LA ACTIVIDAD BIOLÓGICA.....</u></b>	<b>37</b>
III.1 INTRODUCCIÓN .....	37
III.2 MODELOS Y MÉTODOS .....	44
III.2.1 Simulaciones .....	44
III.2.2 Partición de la energía.....	46

<b>III.3 RESULTADOS Y DISCUSIÓN</b> .....	<b>48</b>
III.3.1 Interacciones Ligando-Sitio activo .....	48
III.3.1.1 Compuestos 2 (Nu6094) y 25 .....	48
III.3.1.2 Compuesto 3 (Nu6102).....	51
III.3.1.3 Compuestos 9 y 7.....	52
III.3.2 Análisis de las energías de interacción.....	54
<b>III.4 CONCLUSIONES</b> .....	<b>59</b>
<b>III.5. BIBLIOGRAFÍA</b> .....	<b>62</b>

**CAPÍTULO IV..... 68**

**ESTUDIO TEÓRICO EN INHIBIDORES DE CDK2 UTILIZANDO UNA BLANDURA GLOBAL OBTENIDA A PARTIR DE LA DENSIDAD DE ESTADOS..... 68**

<b>IV.1 INTRODUCCIÓN</b> .....	<b>68</b>
<b>IV.2. ECUACIONES MODELO Y DETALLES COMPUTACIONALES.</b> .....	<b>72</b>
IV.2.1. Índices de reactividad obtenidos a partir de la aproximación del orbitales moleculares de frontera. ....	72
IV.2.2. Índices de reactividad obtenidos a partir de la aproximación de la densidad de estados (DOS). ....	75
<b>IV.3. RESULTADOS Y DISCUSIÓN.</b> .....	<b>77</b>
IV.3.1. Índices globales, locales y regionales de reactividad calculados a partir de la aproximación de orbitales moleculares de frontera. ....	77
IV.3.2. Índices globales de reactividad calculados a partir de la DOS. ....	83
IV.3.2. Análisis del Potencial Electroestático Molecular. ....	90
<b>IV.4. CONCLUSIONES</b> .....	<b>93</b>
<b>IV.5. BIBLIOGRAFÍA</b> .....	<b>96</b>

**CONCLUSIONES GENERALES Y PROYECCIONES FUTURAS..... 100**

**APÉNDICE 1..... 102**

<b>A.1.1. INTRODUCCIÓN</b> .....	<b>102</b>
<b>A.1.2. LA ENERGÍA COMO FUNCIÓN DE PARÁMETROS DE GEOMETRÍA INTERNOS.</b> .....	<b>103</b>
A.1.2.1. Términos enlazantes de la función de energía. ....	104
A.1.2.1.1. Energía de enlace .....	105
A.1.2.1.2. Energía de ángulo de enlace.....	105
A.1.2.1.3. Energía asociada a ángulos diedros.....	106
A.1.2.1.4. Energía asociada a un ángulo impropio .....	107
A.1.2.2. Términos no-enlazantes de la función de energía. ....	109
A.1.2.2.1. Energía electrostática.....	109
A.1.2.2.2. Energía de Lennard-Jones .....	110
A.1.2.2.3. Otros aspectos a considerar .....	113
<b>A.1.3. MÉTODOS HÍBRIDOS DE CÁLCULO QM/MM</b> .....	<b>114</b>

A.1.4. BIBLIOGRAFÍA.....	118
<b>APÉNDICE 2.....</b>	<b>119</b>
A.2.1. INTRODUCCIÓN.....	119
A.2.2. LAS BASES DE LA DFT.....	119
A.2.2.1. Los teoremas de Hohenberg y Kohn.....	119
A.2.2.2. Las ecuaciones de Kohn-Sham.....	122
A.2.3. CONCEPTOS QUÍMICOS Y SU DEFINICIÓN.....	123
A.2.3.1. La electronegatividad y el potencial químico electrónico.....	123
A.2.3.2. La dureza y la blandura globales.....	124
A.2.3.3. Electrofilia Global.....	126
A.2.3.4. La función de Fukui y la blandura local y grupal.....	128
A.2.4. BIBLIOGRAFÍA.....	133
<b>ANEXO 1.....</b>	<b>137</b>
<b>ANEXO 2.....</b>	<b>157</b>

## ÍNDICE DE TABLAS

TABLA 2.1. DISTANCIAS DE ENLACE SELECCIONADAS DEL ANILLO PIRIMIDÍNICO DE LA MOLÉCULA NU2058, QUE HA SIDO OPTIMIZADA UTILIZANDO DIFERENTES MÉTODOS. ....	20
TABLA 2.2. DISTANCIAS ENTRE LOS ÁTOMOS DADORES Y ACEPTORES DE PUENTE DE HIDRÓGENO SELECCIONADAS OPTIMIZADAS Y ENERGÍAS DE ENLACE ( $-\Delta E$ , KCAL/MOL) Y SUS COMPONENTES PARA EL MODELO NU2058-LEU83 ESTUDIADO CON LOS MÉTODOS ONIOM2 Y ONIOM3. ....	21
TABLA 2.3. ENERGÍAS DE INTERACCIÓN OBTENIDAS EN LOS NIVELES DE CÁLCULO MP2, B3LYP Y HF PARA LOS DIFERENTES MODELOS DE INHIBIDORES DE TIPO PURINA N-2 SUSTITUIDOS ESTUDIADOS. ....	23
TABLA 2.4. ENERGÍAS DE INTERACCIÓN ONIOM3 (MP2/6-31G*: HF/3-21G: PM3) OBTENIDAS PARA LOS DIFERENTES MODELOS DE INHIBIDORES TIPO PURINA C-2 SUSTITUIDOS ESTUDIADOS. ....	25
TABLA 2.5. DISTANCIAS INTERATÓMICAS ENTRE LOS ÁTOMOS DADOR Y ACEPTOR DE PUENTE DE HIDRÓGENO EN LOS DIFERENTES MODELOS ONIOM3 ESTUDIADOS. ....	26
TABLA 2.6. ÁNGULO DE ENLACE ENTRE LOS ÁTOMOS QUE PARTICIPAN EN LOS PUENTES DE HIDRÓGENO EN LOS DIFERENTES MODELOS ONIOM3. ....	27
TABLA 2.7. VALORES DE LOS DESCRIPTORES DE REACTIVIDAD GLOBAL, LA ACTIVIDAD BIOLÓGICA EXPRESADA COMO CONCENTRACIÓN INHIBITORIA MEDIA Y LA ENERGÍA DE INTERACCIÓN DE PH DE LOS INHIBIDORES CON EL AMINOÁCIDO LEUCINA 83 (LEU83) OBTENIDA CON EL MÉTODO ONIOM3 (MP2/6-31G*:HF/3-21G:PM3).....	27

TABLA 3.1. TAMAÑO DEL SISTEMA DE ACUERDO AL INHIBIDOR ESTUDIADO.....	46
TABLA 3.2. DISTANCIAS PROMEDIO DE ENLACE DE HIDRÓGENO (Å) ENTRE EL ANILLO PURINA DE CADA LIGANDO Y LA REGIÓN BISAGRA DEL SITIO ACTIVO DE LA CDK2. LOS VALORES DENTRO DEL PARÉNTESIS CORRESPONDEN A LA DESVIACIÓN ESTÁNDAR.....	50
TABLA 3.3. PROMEDIOS DE LAS ENERGÍAS DE INTERACCIÓN QM/MM Y SUS COMPONENTES (KCAL/MOL) PARA LOS INHIBIDORES DE CDK2 ESTUDIADOS. LAS ACTIVIDADES BIOLÓGICAS SON REPORTADAS COMO EL LOGARITMO COMÚN DE IC <sub>50</sub> (MICROMOLAR).....	54
TABLA 4.1. VALORES DE LOS ÍNDICES GLOBALES DE REACTIVIDAD CALCULADOS A PARTIR DEL FMO Y CLASIFICACIÓN SUGERIDA DE LOS COMPUESTOS, DE ACUERDO AL GRUPO SUSTITUYENTE EN LA POSICIÓN C2 DEL ANILLO PURINA.....	80
TABLA 4.2. VALORES DE BLANDURA GRUPAL Y COMPLEMENTARIEDAD (MATCH) PARA LOS INHIBIDORES Y EL MODELO DE SITIO ACTIVO.....	82
TABLA 4.3. DESCRIPTORES GLOBALES DE REACTIVIDAD CALCULADOS A PARTIR DE LA APROXIMACIÓN DE LA DOS PARA EL CONJUNTO DE INHIBIDORES DE CDK2 ESTUDIADOS Y SUS RESPECTIVOS MODELOS DEL SITIO ACTIVO.....	85

## ÍNDICE DE FIGURAS

- FIGURA 1.1. ESTRUCTURA MONOMÉRICA DE LA CDK2. LAS ALFA HÉLICES DEL LÓBULO CARBOXI-TERMINAL ESTÁN REPRESENTADAS POR CILINDROS Y LAS LÁMINAS BETA DEL LÓBULO AMINO-TERMINAL POR FLECHAS INDICANDO LA DIRECCIÓN EN QUE SE ORIENTAN LAS LÁMINAS BETA. EN EL CENTRO DE LOS DOS LÓBULOS ESTÁ EL INHIBIDOR NU6102 REPRESENTADO EN SUPERFICIE DE VAN DER WAALS..... 2
- FIGURA 2.1 ESTRUCTURAS DE LOS INHIBIDORES DE CDK ESTUDIADOS. .... 12
- FIGURA 2.2 MODELOS DE LOS COMPUESTOS Y AMINOÁCIDOS ESTUDIADOS CON EL MÉTODO ONIOM. .... 12
- FIGURA 2.3 EL SISTEMA MODELO ESTUDIADO QUE CONSTA QUE 5 RESIDUOS DE AMINOÁCIDO PERTENECIENTES A LA REGIÓN BISAGRA DEL SITIO ACTIVO DE LA ENZIMA CDK2 Y EL INHIBIDOR. LA PARTICIÓN DEL MODELO ES MOSTRADA PARA EL MÉTODO DE CÁLCULO ONIOM2 [A ES LA CAPA INTERIOR, Y B+C ES LA CAPA EXTERIOR] Y ONIOM3 [A ES LA CAPA INTERIOR, B ES LA CAPA MEDIA Y C ES LA CAPA EXTERIOR]..... 14.
- FIGURA 2.4 GRÁFICO DEL  $\Delta E^{\text{ONIOM3}}$  VS DISTANCIA ENTRE INHIBIDOR-REGIÓN BISAGRA EN Å. SE REALIZÓ EL CÁLCULO DE ENERGÍA ONIOM3 A VARIAS DISTANCIAS INHIBIDOR-REGIÓN BISAGRA PARA DETERMINAR LA DISTANCIA ÓPTIMA PARA REALIZAR LOS CÁLCULOS DE ENERGÍA EN LOS DEMÁS SISTEMAS..... 16
- FIGURA 2.5. GEOMETRÍA DEL INHIBIDOR NU2058 Y NÚMERO Y TIPO DE ÁTOMOS QUE SON TRATADOS EN UN NIVEL CUÁNTICO ALTO EN LOS CÁLCULOS ONIOM. LOS ÁTOMOS EN COLOR VERDE HACEN PARTE DEL NIVEL DE CÁLCULO MEDIO (ONIOM3) O BAJO (ONIOM2)..... 18

FIGURA 2.6. MODELO DEL COMPUESTO NU2058 Y DEL AMINOÁCIDO LEU83 EN INTERACCIÓN, OPTIMIZADOS CON EL MÉTODO ONIOM3 (MP2/6-31G*: HF/3-21G: PM3). LOS ÁTOMOS COLOREADOS EN VERDE HACEN PARTE DEL MODELO INTERMEDIO EN LOS CÁLCULOS ONIOM3 Y LOS QUE NO ESTÁN NUMERADOS CORRESPONDEN A LOS ÁTOMOS QUE UNEN LAS CAPAS INTERIOR Y MEDIA (ÁTOMOS LINK).....	19
FIGURA 2.7. ESTRUCTURA DEL COMPLEJO ENTRE NU2058 Y LEU83. (A) OBTENIDA A PARTIR DE LA OPTIMIZACIÓN TOTAL EN EL VACÍO AL NIVEL DE CÁLCULO MP2/6-31+G*. Y (B) OPTIMIZADA CON EL MÉTODO ONIOM3 (MP2/6-31G*: HF/3-21G: PM3). .....	22
FIGURA 2.8. ESTRUCTURA DEL COMPLEJO ENTRE NU2017 Y LEU83 OBTENIDA A PARTIR DE LA OPTIMIZACIÓN TOTAL EN EL VACÍO AL NIVEL DE CÁLCULO ONIOM3 (MP2/6-31G*: HF/3-21G: PM3).....	26
FIGURA 2.9. GRÁFICA DE CORRELACIÓN ENTRE LA ACTIVIDAD MEDIA INHIBITORIA (IC <sub>50</sub> ) Y EL POTENCIAL QUÍMICO ELECTRÓNICO DE LA SERIE DE INHIBIDORES ESTUDIADOS. ....	30
FIGURA 2.10. GRÁFICA DE CORRELACIÓN ENTRE EL VALOR DE LA CONCENTRACIÓN MEDIA INHIBITORIA (IC <sub>50</sub> ) Y LA ELECTROFILIA ( $\omega$ ) DE LA SERIE DE INHIBIDORES ESTUDIADOS. ....	31
FIGURA 3.1. PUENTES DE HIDRÓGENO TÍPICOS FORMADOS ENTRE LOS INHIBIDORES DE CDK'S Y LA REGIÓN BISAGRA DE LA ENZIMA.....	39
FIGURA 3.2. CAVIDAD POCO PROFUNDA FORMADA POR LA CADENA LATERAL DEL AMINOÁCIDO PHE80 EN EL SITIO ACTIVO DE LA CDK2. LOS AMINOÁCIDOS DE LA REGIÓN BISAGRA ESTÁN REPRESENTADOS EN FORMA DE SUPERFICIE Y EL INHIBIDOR NU6102 EN ESFERAS Y BARRAS. ....	40

FIGURA 3.3 ESTRUCTURAS DE LOS DERIVADOS DE LAS PURINAS O <sup>6</sup> -CICLO-HEXILMETIL N <sup>2</sup> -SUSTITUIDAS DE HARDCASTLE <i>ET AL.</i> TAMBIÉN SE MUESTRA LA NUMERACIÓN DE LOS ÁTOMOS DEL FRAGMENTO PURINA.....	42
FIGURA 3.4. MODELO FINAL DEL SISTEMA. LA CDK2 ESTA REPRESENTADA EN FORMA DE CILINDROS (HÉLICES ALFA) Y EN FLECHAS PLANAS AMARILLAS (LÁMINAS BETA). LA CICLINA A ESTA REPRESENTADA EN FORMA DE CILINDROS (HÉLICES ALFA). POR ÚLTIMO EL INHIBIDOR NU6102 Y LA THR160 ESTÁN REPRESENTADOS EN FORMA DE SUPERFICIE DE VAN DER WAALS. ....	45
FIGURA 3.5. FOTOGRAFÍA DE LA ESTRUCTURA CDK2-Cp2. LOS RESIDUOS SELECCIONADOS DEL SITIO ACTIVO DE LA CDK2 SON MOSTRADOS EN LA REPRESENTACIÓN "LICORICE", MIENTRAS QUE EL Cp2 ES PRESENTADO EN LA REPRESENTACIÓN "ESFERAS Y BARRAS", CON LOS ÁTOMOS DE CARBONO COLOREADOS EN AMARILLO (CDK2) Y CYAN (INHIBIDOR). LOS ENLACES DE HIDRÓGENO SON ESQUEMATIZADOS COMO LÍNEAS DISCONTINUAS NEGRAS.....	49
FIGURA 3.6. FOTOGRAFÍA DE LA ESTRUCTURA DE CDK2-Cp25. LA REPRESENTACIÓN ES REALIZADA COMO EN LA FIGURA 3.5. ....	51
FIGURA 3.7 FOTOGRAFÍA DE LA ESTRUCTURA CDK2-Cp3. LA REPRESENTACIÓN ES REALIZADA COMO EN LA FIGURA 3.5.....	53
FIGURA 3.8. GRÁFICA DE LOGIC <sub>50</sub> VERSUS ENERGÍA DE INTERACCIÓN QM/MM (KCAL/MOL) PARA LOS CINCO DERIVADOS DE PURINA O <sup>6</sup> -CICLO-HEXILMETIL N <sup>2</sup> -SUSTITUIDOS ESTUDIADOS Y LA ECUACIÓN OBTENIDA CON EL AJUSTE DE CUADRADOS MÍNIMOS. ....	55
FIGURA 3.9. GRÁFICA DE LOS VALORES CALCULADOS DE LOGIC <sub>50</sub> (OBTENIDOS POR MEDIO DE LA ECUACIÓN (3.5)) VERSUS LOS VALORES EXPERIMENTALES	

DE LOGIC <sub>50</sub> PARA LOS CINCO COMPUESTOS ESTUDIADOS Y LA CORRESPONDIENTE ECUACIÓN OBTENIDA POR EL AJUSTE DE MÍNIMOS CUADRADOS. ....	59
FIGURA 4.1. MODELO DE LA REGIÓN BISAGRA DE LA CDK2 EN INTERACCIÓN CON EL INHIBIDOR Cp15. LAS INTERACCIONES DE PUENTES DE HIDRÓGENO SON REPRESENTADAS POR LÍNEAS NEGRAS DISCONTINUAS Y EL ESQUELETO DE LOS AMINOÁCIDOS IMPLICADOS EN DICHAS INTERACCIONES ESTÁ ENCERRADO EN LOS CUADRADOS DELINEADOS Y CORRESPONDEN A LOS RESIDUOS GLU81 Y LEU83.....	73
FIGURA 4.2. MODELO ADOPTADO PARA EVALUAR LAS BLANDURAS REGIONALES EN CADA FRAGMENTO Y EL PORCENTAJE DE SIMILARIDAD ENTRE ELLOS. LOS ÁTOMOS EN VERDE Y DEBIDAMENTE ETIQUETADOS SON LOS IMPLICADOS EN EL ANÁLISIS. ....	75
FIGURA 4.3. CORRELACIÓN ENTRE LA ACTIVIDAD MEDIA INHIBITORIA (IC <sub>50</sub> ) Y LA BLANDURA GLOBAL DE LOS COMPUESTOS DEL GRUPO I.....	78
FIGURA 4.4. GRÁFICA DE LOGIC <sub>50</sub> VS BLANDURA GLOBAL (S <sub>DOS</sub> ) (GRUPO II).....	86
FIGURA 4.5. GRÁFICA DE LOGIC <sub>50</sub> VS BLANDURA GLOBAL (S <sub>DOS</sub> ) (GRUPO I).....	86
FIGURA 4.6. DENSIDAD DE ESTADOS GLOBAL Y CONDENSADA PARA LOS COMPUESTOS: A) 3 Y B) 2.....	89
FIGURA 4.7. POTENCIAL ELECTROSTÁTICO MOLECULAR (PEM) PARA LOS INHIBIDORES DE CDK2 MÁS REPRESENTATIVOS: A) 3 Y B) 2. ....	92
FIGURA 4.8. FOTOGRAFÍA DEL PEM PARA: A) LA INTERACCIÓN DE ENLACE DE HIDRÓGENO ENTRE LA REGIÓN BISAGRA DE LA CDK2 Y EL COMPUESTO 3	

(NU6102) Y B) LA INTERACCIÓN DE ENLACE DE HIDRÓGENO ENTRE EL RESIDUO ASP86 Y EL GRUPO SULFONAMIDA DE NU6102.....	94
FIGURA A.1.1. REPRESENTACIÓN ESQUEMÁTICA DE LAS CUATRO INTERACCIONES CLAVE EN UN CAMPO DE FUERZA MOLECULAR.....	102
FIGURA A.1.2. ARREGLO DE ÁTOMOS EN LA DEFINICIÓN DE UN ÁNGULO DIEDRO IMPROPIO.....	106
FIGURA A.1.3. ENERGÍA POTENCIAL DE LENNARD-JONES PARA UN PAR DE ÁTOMOS DE ARGÓN (TOMADO DE <a href="http://en.wikipedia.org/wiki/Lennard-Jones_potential">HTTP://EN.WIKIPEDIA.ORG/WIKI/LENNARD-JONES_POTENTIAL</a> ).....	109
FIGURA A.1.4. EJEMPLOS DE EXCLUSIONES NO-ENLAZANTES 1-2, 1-3 Y 1-4 PARA UN ÁTOMO DE HIDRÓGENO EN UNA MOLÉCULA DE ETANO.....	111
FIGURA A.1.5. REPRESENTACIÓN ESQUEMÁTICA DE LA DIVISIÓN DE UN SISTEMA MOLECULAR EN SUS ZONAS QM Y MM, ES DECIR, LOS DOS SUBSISTEMAS PRESENTES EN LOS MÉTODOS QM/MM. ....	113

## RESUMEN

En esta tesis se presentan varias aproximaciones computacionales aplicadas al estudio de las interacciones proteína-ligando en el sistema enzimático CDK2/Ciclina A (Cyclin Dependent Kinase 2/Cyclin A). Los métodos computacionales utilizados están dirigidos a estudiar la afinidad, por el sitio activo del ATP (Adenosin Trifosfato) en la CDK2, de un grupo de moléculas derivadas de las O<sup>6</sup>-ciclo-hexil-metil-purinas N<sup>2</sup>-sustituidas. Primero, por medio de cálculos realizados con el método ONIOM (Our own N-layered Integrated molecular Orbital and Molecular mechanics), se han logrado caracterizar energéticamente dos de las tres interacciones de puente de hidrógeno, de un grupo de inhibidores, con la región bisagra de la CDK2. Se discute principalmente la influencia de la geometría de los puentes de hidrógeno en las energías obtenidas y se comparan los resultados obtenidos con el método ONIOM con resultados obtenidos con niveles de cálculo MP2 en los mismos modelos. Luego, la afinidad de un grupo de moléculas de la misma familia se trata de explicar utilizando un descriptor obtenido con los métodos de cálculo híbridos QM/MM (Quantum Mechanics/Molecular Mechanics). Por medio de éste método se logró obtener un descriptor de las interacciones de cada inhibidor dentro del sitio activo de la CDK2 y que puede ser utilizado como un índice de la actividad biológica, al menos dentro del grupo de moléculas estudiadas que pertenecen a una misma clase. A partir de las estructuras obtenidas de las dinámicas moleculares ha sido posible caracterizar interacciones de puente de hidrógeno entre algunos de los inhibidores y los aminoácidos de la "superficie específica" de la CDK2, que antes no habían sido reportadas. Una partición de la energía de interacción en sus componentes de van der Waals y electrostática, permite discutir más detalladamente la importancia de cada componente en la interacción proteína-ligando. Finalmente, con la obtención de estructuras de mínima energía a partir de los cálculos QM/MM y con la construcción de modelos más completos a partir de estas, pudimos aplicar los índices de blandura local y regional en la interacción de cada inhibidor con la región bisagra, los cuales fueron obtenidos con dos aproximaciones diferentes (orbitales moleculares frontera y la densidad de estados), y estos ayudan a clasificar las interacciones de los inhibidores con la región bisagra dentro del principio de ácidos y bases duros y blandas (HSBA). Se espera que estas metodologías puedan ser mejoradas y aplicadas a la interacción proteína-ligando en otros sistemas biológicos de interés en la química medicinal actual.

## ABSTRACT

Some computational approaches to deal with the description of the protein-ligand interactions in the CDK2/Cyclin A (Cyclin Dependent Kinase 2/Cyclin A) are presented. We specifically focus our study on the affinity, for the ATP active site in the CDK2, displayed by a series of derivatives of the N<sup>2</sup>-substituted O<sup>6</sup>-cyclo-hexil-methyl-purines. We first performed calculations using the ONIOM method (Our own N-layered Integrated molecular Orbital and Molecular Mechanics). Three of the two most common H-bond interactions have been characterized between a group of inhibitors and the hinge region of CDK2 at this level of theory. The results including the geometry and energetic of the hydrogen bonds compare well with more the accurate MP2 calculations. The affinity of a subset of compounds belonging to the same group of inhibitors was evaluated using the more accurate QM/MM (Quantum Mechanics/Molecular Mechanics) hybrid method. This methodology provides a more complete description of the protein-ligand interaction and it can be used as a good descriptor of the biological activity. Specifically, we have found new hydrogen bond interactions between the inhibitor and the "specific surface" of CDK2. The minimal energy structures obtained from QM/MM studies were used as input structures for the evaluation of reactivity index derived from the density of states approximation. The global softness and the electrophilicity indexes yield correlations against the biological activity that, are statistically significant.

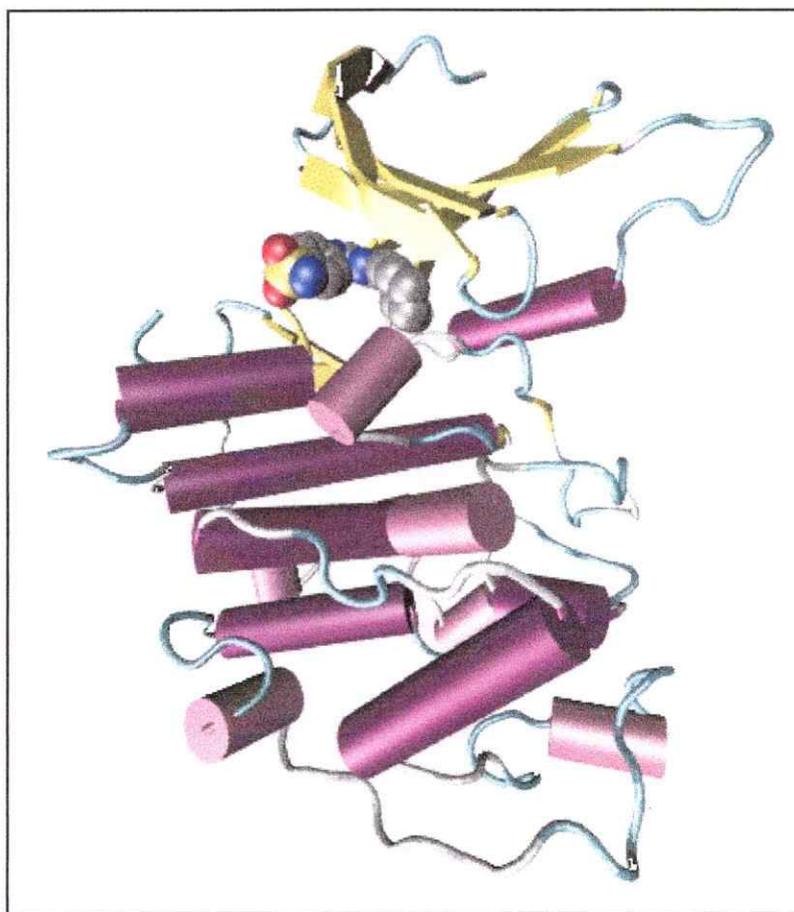
# CAPÍTULO I

## I.1. Introducción

Las quinasas dependientes de ciclina (CDK's) juegan un rol esencial en la regulación del progreso del ciclo celular eucariótico. Están constituidas por un lóbulo amino-terminal rico en láminas- $\beta$  y un lóbulo carboxi-terminal compuesto de hélices- $\alpha$  (Ver Figura 1.1). El ATP (Adenosine 5'-triphosphate) se une en una hendidura profunda formada por los dos lóbulos. Son altamente homólogas (40-75%) y están formadas por aproximadamente 300 residuos de aminoácidos.<sup>1</sup> Las proteínas quinasas son generalmente clasificadas como reguladoras de fase G1, S y G2; debido a que están presentes en varios puntos de control dentro del ciclo celular.<sup>2</sup> Como su nombre lo sugiere, la activación de las CDK's depende de otras proteínas de gran tamaño conocidas como ciclinas. Sólo en forma de complejo, estas proteínas pueden regular apropiadamente el crecimiento celular y la síntesis del ADN. La activación parcial ocurre después de la unión de esas sub-unidades reguladoras positivas; y la activación completa requiere de la fosforilación de la sub-unidad de CDK, en un residuo conservado de Treonina, por medio de la quinasa activadora de CDK. La CDK abordada en el desarrollo de esta tesis es la CDK2, la cual se combina con la ciclina E en el punto de control de la fase S conocido como punto de restricción. De la misma forma, la finalización de la fase S esta sujeta a la formación del complejo entre la CDK2 y la ciclina A. La actividad del complejo CDK-ciclina puede ser reducida por al menos dos mecanismos principales: la fosforilación de la sub-unidad de CDK en los sitios inhibitorios, y la unión de los inhibidores especializados de la proteína y sintetizados por la célula, conocidos como CKI's o inhibidores de quinasas dependientes de ciclina. Estos inhibidores compiten con el ATP por la unión al sitio activo de la CDK. Sin embargo, se ha demostrado que en algunas células del cáncer, los CKI's están expresados en bajas cantidades, y los químicos medicinales han realizado numerosos esfuerzos para reemplazar estos inhibidores naturales con inhibidores de tipo sintético.<sup>3</sup>

Se ha logrado un progreso considerable en la identificación de agentes farmacológicos que tienen como blanco las CDK's.<sup>4</sup>

Se han identificado una cantidad considerable de inhibidores competitivos del ATP y que pertenecen a diferentes clases químicas.<sup>2, 5, 6, 7</sup>



**Figura 1.1.** Estructura monomérica de la CDK2. Las alfa hélices del lóbulo carboxi-terminal están representadas por cilindros y las láminas beta del lóbulo amino-terminal por flechas indicando la dirección en que se orientan las láminas beta. En el centro de los dos lóbulos está el inhibidor Nu6102 representado en superficie de van der Waals.

Entre los intentos por crear tales inhibidores, los más dignos de mencionar pertenecen a una serie de compuestos basados en la estructura O<sup>6</sup>-ciclo-hexilmetil-purina (NU2058), los cuales son inhibidores competitivos del ATP con respecto a la CDK1 y la CDK2, y también muestran buena selectividad sobre la CDK4.<sup>8</sup> Varios autores, con la ayuda del diseño de drogas basado en la estructura (structure-based drug design), han explorado

cuidadosamente este núcleo básico. De esta forma, ha sido posible identificar interacciones importantes para esta clase de compuestos dentro del sitio activo de algunas CDK's.

Además de esto, en los últimos años la búsqueda de compuestos biológicamente activos ha empezado a ser apoyada por los métodos computacionales que fueron desarrollados inicialmente para el estudio de mecanismos de reacción, efectos de sustituyente, superficies de energía potencial; entre otras áreas de interés químico, de compuestos o reacciones que involucren pocos átomos. Con el gran adelanto tecnológico en el diseño de supercomputadores en serie y con la necesidad de extender éstos métodos a problemas de interés biológico, siendo uno de ellos el diseño de nuevos fármacos, se han desarrollado metodologías computacionales que permiten estudiar sistemas en fase condensada que están compuestos por cientos e incluso miles de átomos. Algunos de estas metodologías hacen uso de varios niveles de cálculo que permiten partir el sistema de interés en varias capas y así asignarle a cada una de ellas un nivel de cálculo adecuado para describir las propiedades electrónicas y energéticas de todo el sistema a un costo computacional moderado. Aunque estos métodos "híbridos" aún están en pleno desarrollo, han sido utilizados con éxito en la descripción de varias reacciones en medios enzimáticos<sup>9, 10, 11</sup> y en la determinación de las energías libres de unión de algunos inhibidores por su sitio activo.<sup>12</sup> Las ventajas de utilizar este tipo de métodos en el estudio de sistemas biológicos radican en incluir en los cálculos computacionales efectos tales como la polarización, efectos del medio enzimático, etc., que no son posibles de incluir en los métodos *ab-initio* debido al costo computacional tan elevado cuando se considera un número muy grande de átomos en el modelo a estudiar. Por el contrario, los métodos utilizados actualmente en la industria farmacéutica para la búsqueda de compuestos activos, como por ejemplo el docking,<sup>13</sup> COMFA,<sup>14, 15</sup> COMSIA,<sup>16, 17</sup> etc., carecen de la descripción detallada que ofrece los métodos híbridos QM/MM (Quantum Mechanics/Molecular Mechanics) de los procesos que ocurren en el sitio activo de una enzima. Esto se debe a que no se incorporan los efectos inducidos por el ambiente enzimático y el solvente, pero a su vez brindan descriptores que son mucho más fáciles y rápidos de obtener, y que obviamente reducen el tiempo

computacional y permiten evaluar un mayor número de candidatos. No obstante, lo que se busca con éstos métodos es un balance entre el costo computacional y la descripción lo más detallada posible de las propiedades del sistema de interés, por lo tanto es de esperar que en unos pocos años más se logre conseguir este balance con el desarrollo continuo de nuevos y más veloces computadores y con la depuración de los protocolos de la técnica. Consecuentemente, la obtención de descriptores teóricos a partir de cálculos computacionales de diferente nivel de complejidad y que permitan al menos clasificar una serie de inhibidores de acuerdo a su afinidad por un determinado sitio activo es una de las retos más grandes que se han formulado algunos químicos medicinales y teóricos en los últimos años. En esta tesis hemos realizado un trabajo que va en esa dirección y en el cual tratamos de evaluar la afinidad de una serie de inhibidores de la CDK2 utilizando algunos descriptores derivados de la teoría de funcionales de la densidad (DFT) <sup>18</sup> y de los métodos híbridos QM/MM mencionados anteriormente. Los descriptores derivados de la DFT, en general, permiten definir de una manera cualitativa los fragmentos más reactivos en una molécula y así racionalizar la forma como se produce una interacción entre dos sistemas que reaccionan. Por ejemplo, a partir del principio de ácidos y bases duras y blandas (HSAB) <sup>19</sup> una molécula blanda, o sea con sus electrones más polarizables, va a reaccionar preferentemente con una molécula que tenga las mismas características. De ésta manera se pueden clasificar las interacciones en: electrostáticas o controladas por carga, que corresponderían a interacciones duras y en interacciones de tipo covalente o controladas por orbitales, que corresponden a interacciones blandas. Existen también otros descriptores más generales que se pueden obtener a partir de métodos computacionales menos costosos como son la dinámica molecular en sus diferentes niveles. Un descriptor común derivado de éste tipo de cálculos es la energía libre de Gibbs ( $\Delta G$ ), que tiene la desventaja de ser difícil de obtener debido a la cantidad de cálculos computacionales involucrados. Así la hipótesis propuesta en esta tesis es la posibilidad de obtener un descriptor a partir de los métodos híbridos de cálculo QM/MM, el cual puede predecir de manera correcta la afinidad de los inhibidores de CDK2 estudiados pertenecientes a una misma familia. Además también es

posible utilizar las estructuras de mínima energía resultantes de los cálculos QM/MM para obtener descriptores tradicionales derivados de la DFT en el esquema de la densidad de estados (DOS).

Para este fin hemos utilizado en esta tesis dos métodos híbridos incorporados en diferentes paquetes de cómputo: ONIOM en Gaussian98<sup>20</sup> y QM/MM<sup>21</sup> en DYNAMO<sup>22</sup>, para el estudio computacional del sistema enzimático activo de la CDK2/Ciclina A.

De esta manera en el segundo capítulo discutiremos como con el método ONIOM hemos podido caracterizar energéticamente dos de las tres principales interacciones de puente de hidrógeno que se presentan en varias familias de inhibidores hasta ahora sintetizadas, y que corresponden a las interacciones con la región bisagra de la CDK2, más específicamente con el aminoácido Leu83. De acuerdo a las energías de puente de hidrógeno obtenidas es posible evaluar el efecto de los grupos sustituyentes en la posición C2 del anillo purínico y determinar la importancia y contribución de cada uno de los puentes de hidrógeno formados en la interacción proteína-ligando.

En el tercer capítulo se describe la utilización del método híbrido de cálculo QM/MM en el cálculo de la energía de interacción para una serie de 5 inhibidores de la CDK2 y la descomposición de esta energía en su parte electrostática y de Van der Waals (VdW), como una forma útil y práctica de determinar cual es la componente dominante en la interacción proteína-ligando.

En el cuarto capítulo describimos la obtención, a partir de las estructuras obtenidas con el método híbrido QM/MM, de diferentes descriptores de reactividad que dan cuenta de las interacciones dentro del sitio activo de la enzima CDK2 de una serie de inhibidores derivados del anillo purina y que pertenecen a la familia de las O<sup>6</sup>-ciclo-hexil-metil-purinas N<sup>2</sup>-sustituidas.

De este modo fue posible evaluar en el contexto del principio HSAB (Hard and Soft Acids and Bases) los índices de reactividad local y regional derivados de la teoría de funcionales de la densidad, utilizando dos aproximaciones para el cálculo de los índices, a saber: los orbitales moleculares de frontera (FMO)<sup>23</sup> y la densidad de estados (DOS).<sup>24</sup> A partir de estos índices fue posible modelar mapas de blandura regional en

cada uno de los subsistemas en interacción (proteína y ligando) y determinar el porcentaje de similitud entre los dos fragmentos. Al relacionar estos valores de similitud con las actividades biológicas es posible inferir con un grado de certeza aproximado la afinidad de un inhibidor por el sitio activo de la enzima.

En los Apéndices 1 y 2 se presentan los aspectos teóricos más relevantes de las metodologías empleadas en la ejecución de esta tesis, que son necesarios para entender su desarrollo. En el Apéndice 1 se discute detalladamente la definición de un campo de fuerza típico y además los diferentes términos involucrados en la obtención de la energía QM/MM. En el Apéndice 2 se presenta de manera breve la obtención de los índices globales de reactividad a partir de la DFT y su relación con el principio de HSAB.

Finalmente en el Anexo 1 de esta tesis, se consignan el resto de tablas y gráficos que se han compilado durante el desarrollo de la tesis y que no han podido ser incluidos en el cuerpo principal del manuscrito por motivos de claridad y espacio.

## I.2. Bibliografía

- [1] Morgan, D.O. Principles of CDK regulation. *Nature* **1995**, 374, 131-134.
- [2] Siewicki, T. M.; Boylan, J. F.; Benfield, P. A.; Trainor, G. L. Cyclin-Dependent Kinase Inhibitors: Useful Targets in Cell Cycle Regulation. *J. Med. Chem.* **2000**, 43, 1-18.
- [3] Losiewicz, M. D.; Carlson, B. A.; Kaur, G.; Sausville, E. A.; Worland, P. J. Potent inhibition of CDC2 kinase activity by the flavonoid L86-8275. *Biochem. Biophys. Res. Commun.* **1994**, 201, 589-595.
- [4] Senderowicz, A. M.; Sausville, E. A. Preclinical and clinical development of cyclin-dependent kinase modulators. *J. Natl. Cancer Inst.* **2000**, 92, 376-387.
- [5] Hardcastle, I. R.; Golding, B. T.; Griffin, R. J. Designing inhibitors of cyclin-dependent kinases. *Annu. Rev. Pharmacol. Toxicol.* **2002**, 42, 25-348.
- [6] Knockaert, M.; Greegard, P.; Meijer, L. Pharmacological inhibitors of cyclin-dependent kinases. *Trends Pharmacol. Sci.* **2002**, 23, 417-425.
- [7] Toogood, P. L. Progress toward the development of agents to modulate the cell cycle. *Curr. Opin. Chem. Biol.* **2002**, 6, 472-478.
- [8] Davies, T. G.; Bentley, J.; Arris, C. E.; Boyle, F. T.; Curtin, N. J.; Endicott, J. A.; Gibson, A. E.; Golding, B. T.; Griffin, R. J.; Hardcastle, I. R.; Jewsbury, P.; Johnson, L. N.; Mesguiche, V.; Newell, D. R.; Noble, M. E. M.; Tucker, J. A.; Wang, L.; Whitfield, H. J. Structure-based design of a potent purine-based cyclin-dependent kinase inhibitor. *Nat. Struct. Biol.* **2002**, 9, 745-749.
- [9] Soriano, A.; Silla, E.; Tuñon, I.; Ruiz-López, M. F. Dynamic and Electrostatic Effects in Enzymatic Processes. An Analysis of the Nucleophilic Substitution Reaction in Haloalkane Dehalogenase. *J. Am. Chem. Soc.* **2005**, 127, 1946-1957.

- [10] Roca, M.; Andrés, J.; Moliner, V.; Tuñon, I.; Bertrán, J. On the Nature of the Transition State in Catechol O-Methyltransferase. A Complementary Study Based on Molecular Dynamics and Potential Energy Surface Explorations. *J. Am. Chem. Soc.* **2005**, *127*, 10648-10655.
- [11] Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñon, I.; Bertrán, J. Preorganization and Reorganization as Related Factors in Enzyme Catalysis: The Chorismate Mutase Case. *Chem. Eur. J.* **2003**, *9*, 984-991.
- [12] Gräter, F.; Schwarzl, S. M.; Dejaegere, A.; Fischer, S.; Smith, J. C. Protein/Ligand Binding Free Energies Calculated with Quantum Mechanics/Molecular Mechanics. *J. Phys. Chem. B* **2005**, *109*, 10474-10483.
- [13] Baxter, C. A.; Murray, C. W.; Clark, D. E.; Westhead, D. R.; Eldridge, M. D. Flexible docking using tabu search and an empirical estimate of binding affinity. *Proteins: Structure, Function, and Genetics* **1998**, *33*, 367-382.
- [14] Cramer, R. D.; Patterson, D. E.; Bunce, J. D.; *J. Am. Chem. Soc.* **1988**, *110*, 5959-5967.
- [15] Klebe, G.; Abraham, U.; Mietzner, T.; *J. Med. Chem.* **1994**, *37*, 4130-4146.
- [16] Klebe, G.; Abraham, U. *J. Comput. Aid. Mol. Des.*, **1999**, *13*, 1-10.
- [17] Kaminski, J. J. Computer-assisted drug design and selection. *Advanced Drug Delivery Reviews* **1994**, *14*, 331-337.
- [18] Parr R. G. and Yang W. *Density Functional Theory of Atoms and Molecules* (Oxford University Press, Oxford) **1989**.
- [19] Pearson RG. *J. Am. Chem. Soc.* **1963**; Vol. 85, 3533.

- [20] Gaussian 98 (Revision A.11), M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, P. Salvador, J. J. Dannenberg, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 2001.
- [21] Field, M. J.; Bash, P. A.; Karplus, M. A. A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations. *J. Comput. Chem.* **1990**, *11*, 700-733.
- [22] Field, M. J.; Albe, M.; Bret, C.; Proust-De Martin, F.; Thomas, A. The Dynamo library for molecular simulations using hybrid quantum mechanical and molecular mechanical potentials. *J. Comput. Chem.* **2000**, *21*, 1088-1100.
- [23] C.A. Coulson and H.C. Longuet-Higgins, *Proc. R. Soc. Ser. A*, 191 (1947) 39; 192 (1947) 16; 193 (1948) 447,456; 195 (1948) 188; *Rev. Sci.* 85 (1947) 929.
- [24] Santos, J. C.; Contreras, R.; Chamorro, E.; Fuentealba, P. *J. Chem. Phys.* **2002**, *116*, 4311-4316.

## CAPÍTULO II

### Estudio de las interacciones de puente de hidrógeno entre diferentes ligandos y la región bisagra de la CDK2 con el método ONIOM

#### II.1 Introducción

En investigaciones previas Davies *et al.*<sup>1</sup> han utilizado como compuesto base para su estudio el inhibidor NU2058 (Esquema 4.1, Capítulo IV), el cual es una O<sup>6</sup>-ciclo-hexil-metil-guanina que posee afinidad tanto por la enzima CDK2 como por la CDK1. A partir de este compuesto se han sintetizado varios derivados en los cuales la posición N<sup>2</sup> es sustituida con grupos de diferentes tamaños y características con el fin de mejorar la afinidad y selectividad por la CDK1 y la CDK2.

A partir de estos estudios basados en estructura-actividad se han identificado inhibidores de la CDK2 potentes y selectivos, como por ejemplo el compuesto NU6102 (Esquema 4.1, Capítulo IV). Además se han encontrado relaciones estructura-actividad (SAR) importantes para el diseño de inhibidores más potentes y selectivos.

Sin embargo, a pesar de las largas e intensivas investigaciones experimentales, aún no se tiene una respuesta clara para una pregunta tan importante y de tanta trascendencia como lo es el origen detallado de la interacción enzima-inhibidor.

Un mejor entendimiento de este aspecto es vital para el diseño de nuevos inhibidores que sirvan como compuestos farmacéuticos modelo. Por lo tanto, los trabajos teóricos del tipo presentado en este capítulo, son un método alternativo para estudiar la interacción enzima-inhibidor con más detalle. A pesar de todo, tales trabajos en sistemas moleculares tan grandes como las enzimas son limitados, ya que requieren de un gran esfuerzo computacional. Recientemente, y gracias al desarrollo experimentado por la química computacional, es posible disponer de modelos moleculares bastante

aproximados para moléculas grandes como las involucradas en la biología molecular.<sup>2</sup> Particularmente se ha introducido el método ONIOM (Our own N-layered Integrated molecular Orbital and Molecular Mechanics) y se ha demostrado su eficiencia, la cual ha sido mejorada con el paso de los años.<sup>3</sup>

En la aproximación ONIOM<sup>4, 5</sup> una parte pequeña del sistema, como por ejemplo el inhibidor y un aminoácido reactivo en el sitio activo de la enzima, se estudia a un nivel químico cuántico alto, mientras la región que los rodea es modelada utilizando un nivel de cálculo más bajo.<sup>6, 7</sup>

Como la interacción de la mayoría de los inhibidores de CDK2 y la región bisagra del sitio activo de la enzima (aminoácidos Phe80, Glu81, Phe82, Leu83 e His84) ha sido poco explorada y por tanto permanece como un problema abierto, en esta parte de la tesis aplicamos el método ONIOM para examinar la estructura final de un grupo de inhibidores en interacción con esta región de la enzima CDK2. Analizamos y reportamos una serie de parámetros estructurales y energéticos, entre ellos las energías de interacción de puentes de hidrógeno (PH), para un grupo de compuestos pertenecientes a la serie de las O<sup>6</sup>-ciclo-hexil-metil-purinas-C<sup>2</sup>-sustituidas no aromáticas publicadas por Hardcastle *et al.*<sup>8</sup> y que se esquematizan en las Figuras 2.1a a 2.1d.

Las interacciones en estudio incluyen dos puentes de hidrógeno que se forman entre el aminoácido Leu83 y el anillo pirimidínico de los inhibidores mencionados. El átomo N3 y el grupo 2-NH<sub>2</sub> del anillo pirimidínico acepta y dona un enlace de hidrógeno, respectivamente, a los grupos carbonilo y amido de la cadena principal del aminoácido Leu83. Estas dos interacciones de puente de hidrógeno fueron incluidas en el modelo principal que se estudió con un nivel de cálculo alto. Una tercera interacción de puente de hidrógeno se da entre el grupo NH-9 del anillo pirrol del inhibidor y el grupo carbonilo de la cadena principal del aminoácido Glu81. Esta interacción está incluida en el modelo de cálculo de nivel bajo. Para tal fin se definieron los modelos de cada uno de los compuestos a estudiar en el marco del método ONIOM. Los modelos ONIOM se presentan en la Figura 2.2.

En el presente estudio se plantearon tres objetivos. Primero, establecer la aplicabilidad del método ONIOM a la interacción de los inhibidores estudiados con los aminoácidos

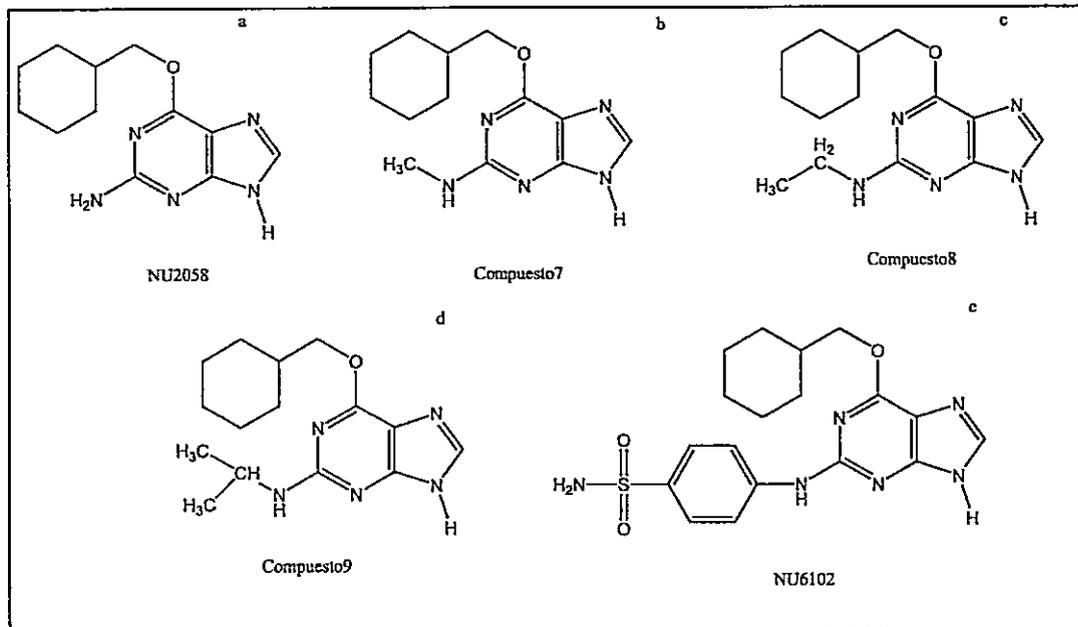


Figura 2.1 Estructuras de los inhibidores de CDK2 estudiados.

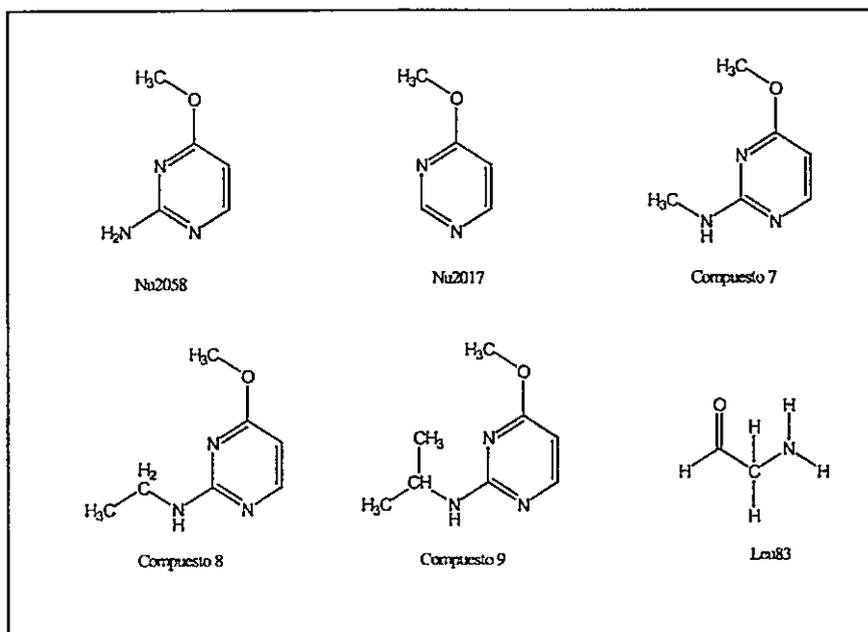


Figura 2.2 Modelos de los compuestos y aminoácidos estudiados con el método ONIOM.

de la región bisagra en el sitio activo de la CDK2. Segundo, estudiar las dos interacciones de puente de hidrógeno mencionadas por medio de un modelo

simplificado de los inhibidores y el aminoácido Leu83 a un nivel de cálculo alto (MP2) y tercero, analizar las interacciones de puente de hidrógeno con la ayuda de los índices globales de reactividad definidos en la teoría de los funcionales de la densidad (DFT). El conocimiento básico de tales interacciones contribuirá a una mejor comprensión de las fuerzas y los mecanismos de unión de los inhibidores de CDK2 con la región bisagra, y su posible utilización en la información estructural disponible para el diseño de inhibidores más potentes y selectivos.

## II.2. Modelos y métodos.

### II.2.1 Sistemas estudiados

Los modelos utilizados en este estudio se obtuvieron a partir de la estructura cristalográfica, resuelta a 2.1 Å y publicada por Davies, T. G. *et al.* del inhibidor NU2058 y la CDK2. (Código en el PDB: 1H1P).

Basados en esta estructura, se adoptó un sistema compuesto por 5 residuos de aminoácido (región bisagra) y el inhibidor. Los residuos de aminoácido incluidos fueron Phe80, Glu81, Phe82, Leu83 e His84. Se asumió que estos aminoácidos están presentes en el sistema en su forma neutra, para lo cual sus extremos y cadenas laterales fueron saturados con un átomo de hidrógeno tanto en el extremo amino-terminal como en el carboxi-terminal. Además se asumió que todos sus ángulos de enlace y de torsión fueran los mismos que en la estructura cristalográfica. Se adicionaron los átomos de hidrógeno faltantes al sistema modelo tomado de la estructura cristalográfica, para generar un sistema estructuralmente completo y neutro. Las posiciones de estos fueron optimizadas con el método semi-empírico PM3.<sup>9, 10</sup> La estructura convergida totalmente fue la utilizada como modelo inicial para realizar los cálculos restantes.

### II.2.2 Cálculos ONIOM

Para investigar la interacción de los inhibidores de CDK2 con los aminoácidos de la región bisagra del sitio activo se utilizaron los métodos ONIOM2 y ONIOM3 que constan de dos y tres capas, respectivamente. Luego se escogió la mejor combinación ONIOM para aplicarla a los otros sistemas estudiados. Como se muestra en la Figura 2.3, en el método ONIOM2, el sistema se divide en dos partes. La capa interior o la región de interacción (Figura 2.3, región A que está formada por los modelos de cada inhibidor y del aminoácido Leu83 presentados en la Figura



2.2) fue tratada a un nivel de cálculo alto, mientras que la capa exterior o la región que rodea el sistema de interés (regiones B y C que está formada por los átomos restantes) fue tratada a un nivel de cálculo bajo.

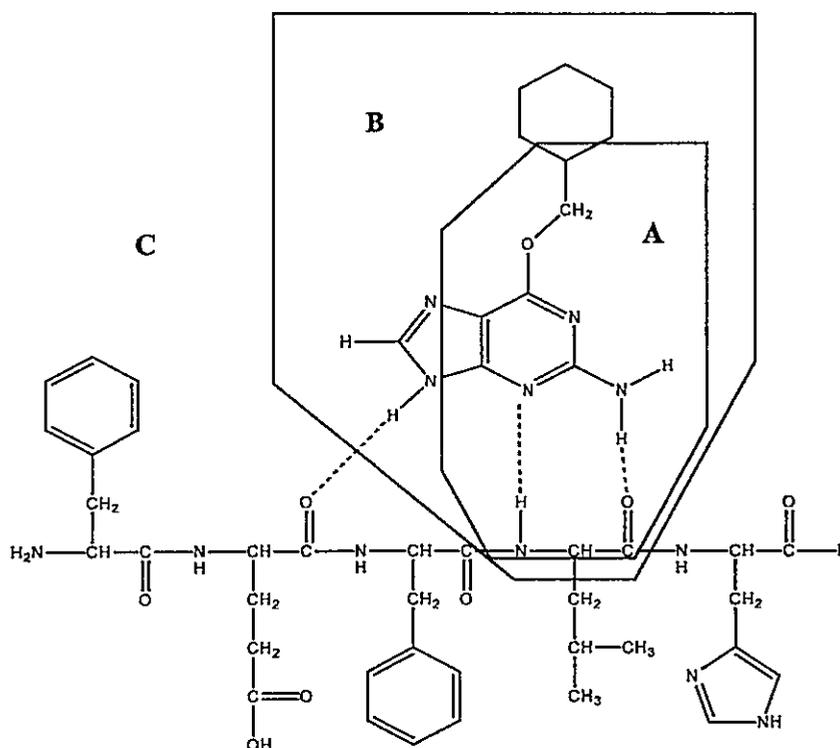


Figura 2.3 El sistema modelo estudiado que consta que 5 residuos de aminoácido pertenecientes a la región bisagra del sitio activo de la enzima CDK2 y el inhibidor. La partición del modelo es mostrada para el método de cálculo ONIOM2 [A es la capa interior, y B+C es la capa exterior] y ONIOM3 [A es la capa interior, B es la capa media y C es la capa exterior].

En el método ONIOM3, el sistema fue dividido en tres regiones, que también se muestran en la Figura 2.3. La energía total ONIOM de los sistemas completos (ABC) se obtuvo a partir de tres y cinco cálculos de energía independientes que fueron realizados con los métodos ONIOM2 y ONIOM3 respectivamente<sup>3</sup>:

$$E^{\text{ONIOM2}} [\text{ABC}] = E[\text{Alto, A}] + E[\text{Bajo, -ABC}] - E[\text{Bajo, A}] \quad (2.1)$$

$$E^{\text{ONIOM3}} [\text{ABC}] = E[\text{Alto, A}] + E[\text{Medio, AB}] + E[\text{Bajo, ABC}] - E[\text{Medio, A}] - E[\text{Bajo, AB}] \quad (2.2)$$

La energía calculada de esta manera se denomina “esquema de extrapolación”, donde no hay necesidad de incluir un Hamiltoniano de interacción especial entre las capas ya que esta interacción es tratada consistentemente en el nivel de teoría bajo. Existe otro esquema llamado “esquema de conexión” en el cual se debe calcular específicamente un término de energía de interacción entre las dos capas.<sup>3</sup> Para calcular las energías ONIOM2 y ONIOM3 descritas anteriormente se deben realizar dos cálculos por separado. Uno en el cual se incluya el inhibidor y los aminoácidos de la región bisagra y otro en el cual el sistema no esté en interacción. Para disminuir el efecto de error por la superposición de bases (BSSE)<sup>11</sup>, se realizaron varios cálculos “single point” a la estructura separándola cada vez más, hasta que la diferencia en las energías ONIOM calculadas variara poco. En la Figura 2.4 se muestra la variación de la energía con respecto a la distancia de separación entre el inhibidor y la región bisagra. Se determinó una distancia de 20 Å como la óptima para realizar el cálculo de “single point” a todos los demás sistemas y poder calcular las diferencias en energía ONIOM.

Todos los cálculos computacionales fueron realizados utilizando el paquete GAUSSIAN98.<sup>12</sup>

### II.2.3 Cálculo de los índices de reactividad global.

Se calcularon los índices de reactividad globales que están definidos en la teoría de funcionales de la densidad (DFT) introducida por Parr y Yang.<sup>13</sup> Los índices utilizados en este estudio están definidos en Apéndice 2. Estos conceptos están bien definidos en la teoría y pueden explicar convenientemente un cuadro completo de reactividad. En nuestro caso se calculó el potencial químico electrónico ( $\mu$ ), la dureza ( $\eta$ ),<sup>14</sup> la blandura ( $S$ )<sup>15</sup> y la electrofilia ( $\omega$ )<sup>16</sup> en base en la aproximación de Koopmans,<sup>17</sup> esto significa que estas cantidades se calculan en términos de las energías de los orbitales frontera HOMO y LUMO.

Estos descriptores se calcularon para los modelos optimizados con el método ONIOM3 (MP2/6-31G\*: HF/3-21G: PM3) y se utilizó el método MP2/6-31+G\* para el cálculo de

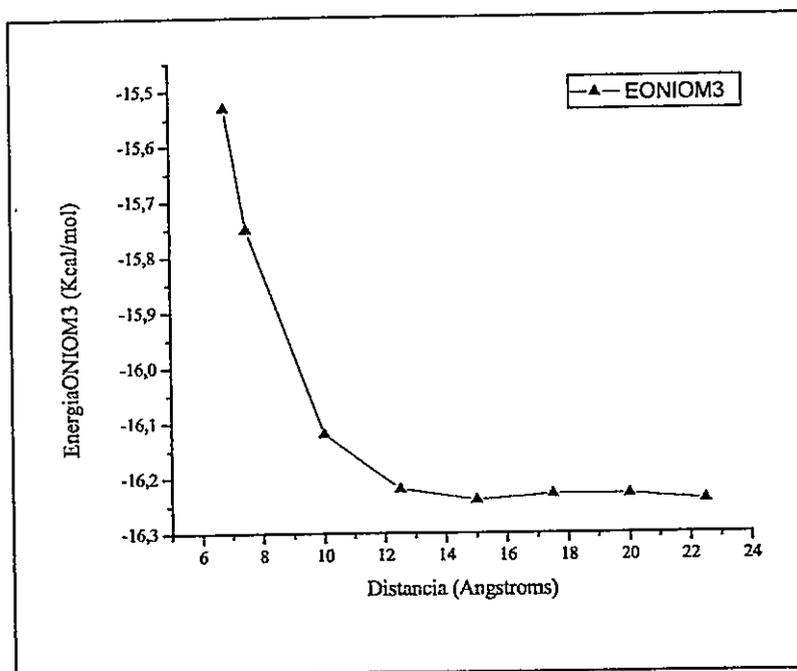


Figura 2.4 Gráfico del  $\Delta E^{\text{ONIOM3}}$  vs Distancia entre Inhibidor-Región bisagra en Å. Se realizó el cálculo de energía ONIOM3 a varias distancias inhibidor-región bisagra para determinar la distancia óptima para realizar los cálculos de energía en los demás sistemas.

los orbitales moleculares del sistema.

## II.3. Resultados y Discusión.

### II.3.1. Aplicabilidad del método ONIOM a la estructura del compuesto NU2058

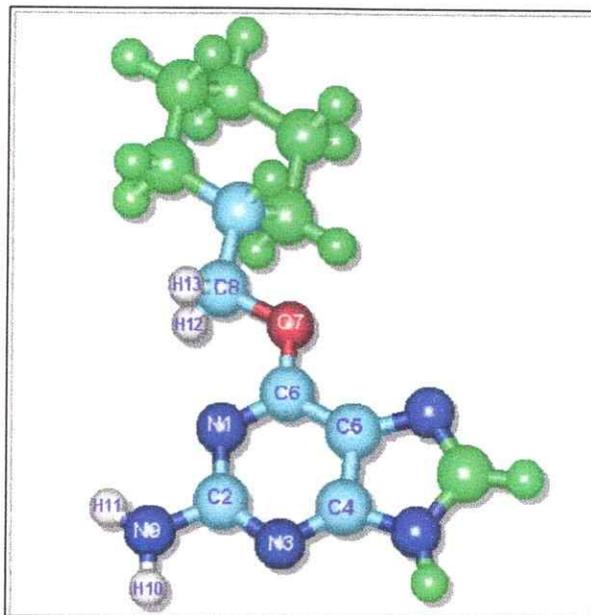
El principal objetivo de este apartado es analizar la interacción de los inhibidores del tipo purinas con la región bisagra del sitio activo de la enzima CDK2 y principalmente con el aminoácido Leu83 que forma parte de dicha región. Sin embargo, debido a que el complejo Inhibidor-Leu83 es muy grande para realizar un cálculo MP2, se ha dividido cada uno de los inhibidores estudiados en dos partes: el anillo pirimidínico y sus sustituyentes en la posición C2 y O6 y el resto de la molécula. Se ha aplicado el método ONIOM al complejo inhibidor-Leu83 en presencia y en ausencia de los otros residuos de aminoácido.

El primer paso a realizar es examinar como se reproduce la estructura del compuesto NU2058, calculada por un método tipo MP2 y por los métodos ONIOM con diferentes niveles de cálculo bajos. La geometría de NU2058 fue optimizada con los métodos y bases HF/6-31G\*, B3LYP/6-31G\* y MP2/6-31G\* como también con el método ONIOM utilizando los anteriores métodos para el nivel de cálculo alto y los métodos HF/3-21G y PM3 para el nivel bajo.

En la Tabla 2.1 se resumen las distancias de enlace seleccionadas optimizadas del anillo pirimidínico de NU2058, con las denominaciones para los átomos mostradas en la Figura 2.5. En general, cualquier combinación de los métodos produce distancias de enlace que están en concordancia con las obtenidas con el método de cálculo de nivel alto MP2.

### **II.3.2 Estructura y energía de interacción del complejo entre el compuesto NU2058 y la región bisagra de la enzima CDK2.**

Se estudió la interacción del complejo NU2058 y la región bisagra de la enzima CDK2, esta última compuesta por cinco residuos de aminoácido. Como se dijo anteriormente, el complejo NU2058 y la región compuesta por los cinco aminoácidos son demasiado grandes para realizar un cálculo con un nivel alto como MP2. Por lo tanto hemos utilizado varios esquemas ONIOM para estudiar dicha interacción. Hemos empleado los métodos MP2<sup>18, 19</sup>/6-31G\*, B3LYP<sup>20, 21, 22</sup>/6-31G\* y HF<sup>23, 24, 25</sup>/6-31G\* para la región interna en interacción (región A, Figura 3), el método HF/3-21G para la región intermedia en el esquema ONIOM3 (región B) y el método semi-empírico PM3<sup>26, 27</sup> para modelar la región exterior (región C en el esquema ONIOM3 y región B+C en el esquema ONIOM2). Se congeló la geometría de los residuos de aminoácido de la región C en las posiciones experimentales reportadas en la estructura de rayos-X y se optimizaron las geometrías de los modelos del aminoácido Leu83 y del compuesto NU2058 para cada combinación ONIOM propuesta.

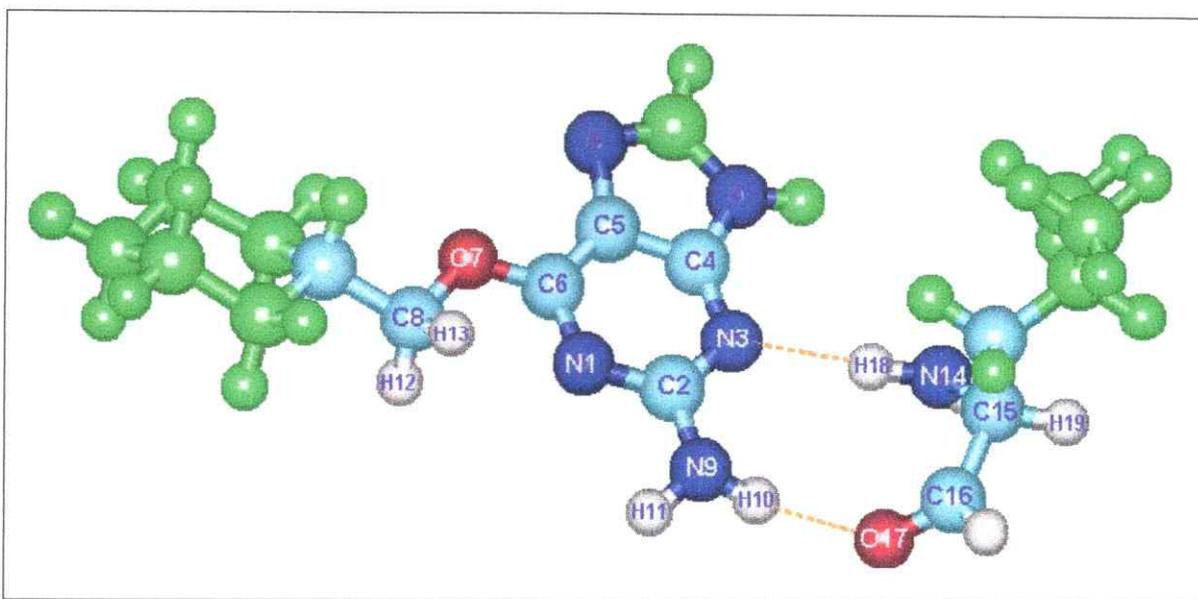


**Figura 2.5.** Geometría del inhibidor NU2058 y número y tipo de átomos que son tratados en un nivel cuántico alto en los cálculos ONIOM. Los átomos en color verde hacen parte del nivel de cálculo medio (ONIOM3) o bajo (ONIOM2).

La Tabla 2.2 muestra las distancias de interacción, entre los átomos dadores y aceptores de puente de hidrógeno, seleccionadas y optimizadas en el modelo de NU2058 y Leu83 en la región interna en estudio. También se calcularon las energías de interacción ( $-\Delta E$ ) de los puentes de hidrógeno en el modelo de NU2058 y Leu83.

La Figura 2.6 muestra la estructura optimizada del complejo al nivel ONIOM3 (MP2/6-31G\*: HF/3-21G: PM3). Consideramos que el método ONIOM3 (MP2/6-31G\*: HF/3-21G: PM3) es el más adecuado para describir las interacciones de puente de hidrógeno en estudio, debido a que el método MP2/6-31G\* utilizado para la región interna en interacción debe ser capaz de dar cuenta no sólo de las interacciones electrostáticas, de polarización y de transferencia de carga, sino además de los efectos de dispersión. La energía de interacción encontrada para el complejo entre el modelo del compuesto NU2058 y el aminoácido Leu83 es substancial y tiene un valor de  $-9.7$  kcal/mol en el nivel de cálculo superior.

Al evaluar las energías de interacción obtenidas con los diferentes métodos ONIOM utilizados, se puede ver el efecto que tiene el método y la base en la exactitud del cálculo de la energía obtenida (Ver Tabla 2.2).



**Figura 2.6.** Modelo del compuesto NU2058 y del aminoácido Leu83 en interacción, optimizados con el método ONIOM3 (MP2/6-31G\*: HF/3-21G: PM3). Los átomos coloreados en verde hacen parte del modelo intermedio en los cálculos ONIOM3 y los que no están numerados corresponden a los átomos que unen las capas interior y media (átomos link).

La comparación de las distancias experimentales, obtenidas por rayos-X, entre los átomos dadores y aceptores de PH y las distancias obtenidas con los diferentes métodos utilizados nos indica la dependencia de la distancia interatómica con la fuerza del enlace en la interacción de puente de hidrógeno. Como es de esperar, a una menor distancia entre los átomos dador y receptor de PH la energía de interacción es mayor.

Además este estudio nos indica como las distancias entre los átomos involucrados en el PH son sobreestimadas por los modelos ONIOM que utilizan el método HF/6-31G\* para el cálculo de las energías de interacción en la región interna y por lo tanto los resultados de las energías no son muy confiables. Por otra parte las distancias mencionadas y las energías obtenidas utilizando el método B3LYP/6-31G\* como nivel de cálculo para la

Tabla 2.1. Distancias de enlace seleccionadas del anillo pirimidínico de la molécula Nu2058, que ha sido optimizada utilizando diferentes métodos.

Métodos	Distancia de enlace (Å)															
	N1- C2	N1- C6	C2- N3	N3- C4	C4- C5	C5- C6	C2- N9	N9- O7	O7- C8	N9- H10	N9- H11	C4- (H*) <sup>a</sup>	C5- (H*) <sup>a</sup>	C8- H12	C8- H13	C8- (H*) <sup>a</sup>
MP2/6-31G*	1.33	1.34	1.36	1.33	1.40	1.40	1.38	1.34	1.45	1.01	1.01	1.37	1.38	1.09	1.10	1.52
(MP2/6-31G*:HF/3-21G)	1.33	1.34	1.36	1.32	1.40	1.40	1.38	1.34	1.45	1.01	1.01	1.37	1.41	1.09	1.09	1.53
(MP2/6-31G*:PM3)	1.33	1.37	1.37	1.32	1.41	1.41	1.38	1.34	1.46	1.01	1.01	1.41	1.41	1.10	1.10	1.52
B3LYP/6-31G*	1.34	1.33	1.35	1.32	1.40	1.40	1.37	1.33	1.44	1.01	1.01	1.37	1.38	1.10	1.10	1.52
(B3LYP/6-31G*:HF/3-21G)	1.34	1.34	1.36	1.32	1.40	1.40	1.37	1.33	1.44	1.00	1.01	1.37	1.41	1.09	1.09	1.54
(B3LYP/6-31G*:PM3)	1.33	1.36	1.37	1.32	1.41	1.42	1.37	1.33	1.45	1.00	1.01	1.41	1.41	1.10	1.10	1.53
HF/6-31G*	1.31	1.32	1.34	1.31	1.40	1.38	1.36	1.31	1.42	0.99	0.99	1.35	1.38	1.08	1.08	1.52
(HF/6-31G*:HF/3-21G)	1.32	1.32	1.34	1.30	1.39	1.38	1.35	1.31	1.42	0.99	0.99	1.37	1.40	1.08	1.08	1.53
(HF/6-31G*:PM3)	1.31	1.35	1.35	1.30	1.41	1.39	1.35	1.30	1.43	0.99	0.99	1.41	1.40	1.09	1.09	1.52
HF/3-21G	1.32	1.33	1.35	1.31	1.39	1.39	1.34	1.33	1.45	0.99	0.99	1.36	1.39	1.08	1.08	1.52

<sup>a</sup> Átomos de hidrógeno que completan la valencia en el cálculo ONIOM y que en el modelo real corresponden a átomos pesados

Distancia de enlace (Å)	Métodos					
	X-ray	MP2/ 6-31G*: HF/ 3-21G:PM3	B3LYP/ 6-31G*: HF/ 3-21G:PM3	HF/ 6-31G*: HF/ 3-21G:PM3	B3LYP/ 6-31G*: PM3	HF/ 6-31G*: PM3
H <sub>2</sub> N-O=C(Leu83)	2.72	2.89	2.85	2.95	2.85	2.95
N3-NH(Leu83)	3.06	3.08	3.08	3.17	3.03	3.13
HN9-O=C(Glu81)*	2.79	2.78	2.79	2.81	2.79	2.79
Energías de interacción (Kcal/mol)						
ΔE	-18.0	-16.2	-15.7	-12.1	-15.4	-11.9
ΔE(High,A)	-10.4	-9.7	-9.4	-6.2	-9.2	-6.1
ΔΔE(Mid, AB-A)	-0.2	-0.4	-0.2	-0.2	---	---
ΔΔE(Low, ABC-AB)	-7.4	-6.2	-6.1	-5.8	-6.2	-5.9

Tabla 2.2. Distancias entre los átomos dadores y aceptores de puente de hidrógeno seleccionadas optimizadas y energías de enlace (-ΔE, kcal/mol) y sus componentes para el modelo NU2058-Leu83 estudiado con los métodos ONIOM2 y ONIOM3. Las energías de interacción para la estructura referencia de rayos-x fueron obtenidas por medio de cálculos "single point" con los métodos ONIOM.

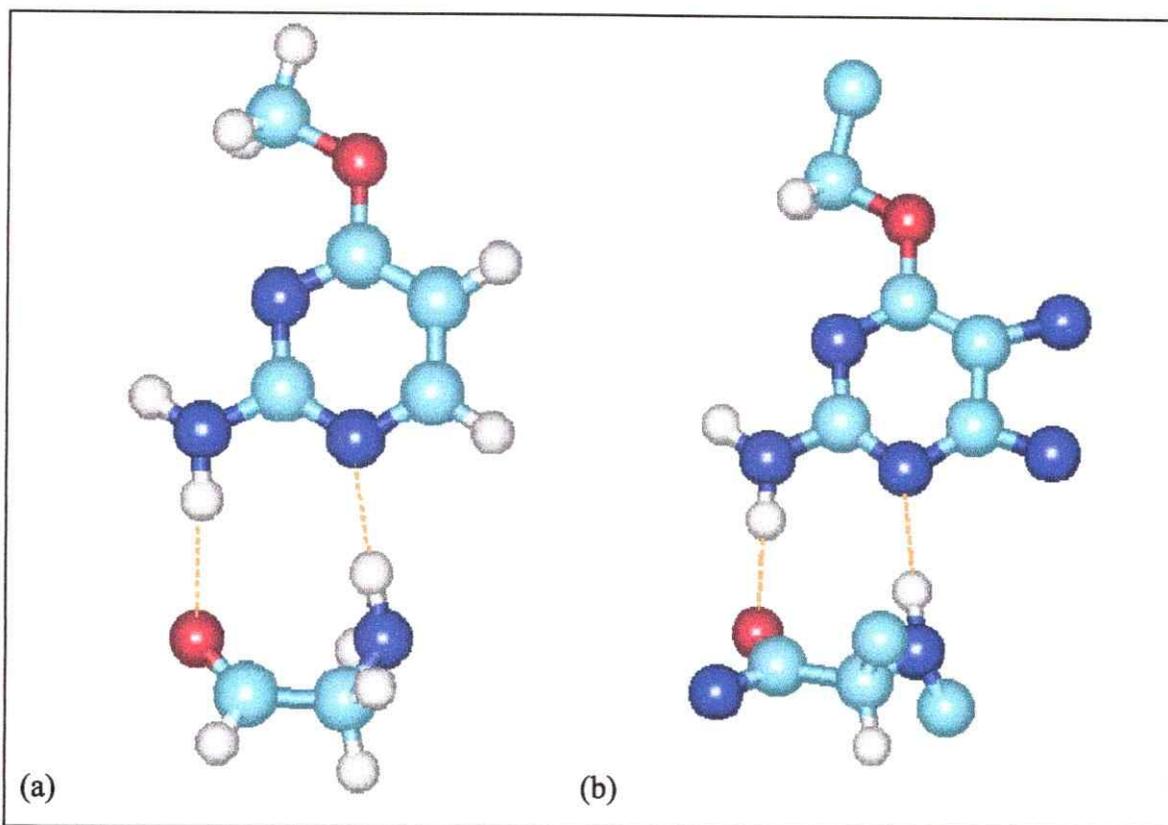
región que interacciona, se aproximan mejor a los datos obtenidos con el método MP2/6-31G\*.

Es importante mencionar también que la energía de interacción de los modelos del compuesto NU2058 y el aminoácido Leu83 es más favorable en la geometría obtenida a partir de la estructura referencia de rayos-X que las energías obtenidas con los diferentes modelos ONIOM utilizados.

La Figura 2.7 muestra las estructuras del complejo NU2058+Leu83 obtenidas a partir de los cálculos MP2/6-31+G\* y ONIOM3 (MP2/6-31G\*: HF/3-21G: PM3). Se puede observar que las geometrías obtenidas con los dos métodos son diferentes, especialmente en la parte de la estructura que pertenece al modelo del aminoácido Leu83.

En la estructura obtenida con el método MP2, la posición que toma el átomo de oxígeno que forma parte del grupo carbonilo del aminoácido está orientada a establecer la interacción con el átomo de hidrógeno a través de sus pares libres.

En la estructura optimizada con el método ONIOM3 (MP2/6-31G\*: HF/3-21G: PM3) la orientación del grupo carbonilo es casi lineal con respecto al hidrógeno con el que forma el PH. Esta orientación es forzada por la conformación del enlace peptídico en la



**Figura 2.7.** Estructura del complejo entre NU2058 y Leu83. (a) obtenida a partir de la optimización total en el vacío al nivel de cálculo MP2/6-31+G\*. (b) y optimizada con el método ONIOM3 (MP2/6-31G\*: HF/3-21G: PM3).

estructura de la enzima. Estas diferencias en la geometría de las estructuras optimizadas se ven reflejadas en las energías de interacción de puente de hidrógeno (Ver Tabla 2.2 y Tabla 2.3). Las energías de PH obtenidas a un nivel MP2 indican una interacción mayor entre los diferentes inhibidores y el aminoácido Leu83, cuando se comparan con las energías de PH obtenidas con el método ONIOM3 empleado. Esta diferencia en energía entre los modelos ONIOM3 y MP2 se puede explicar por el impedimento estérico que se presenta en el modelo ONIOM3, al estar el inhibidor en presencia de los 5 aminoácidos de la región bisagra. En el modelo MP2, el cual es más simplificado, no se presenta este efecto. En la Tabla 2.3 se presentan junto a las energías MP2, las energías determinadas con los métodos B3LYP y HF con la base 6-31+G\*. Los resultados obtenidos con los métodos B3LYP y HF no pueden representar bien las contribuciones de energía de los grupos sustituyentes  $-NH$  y  $-NHCH_3$  a la interacción de PH analizada.

Sin embargo el método MP2 puede reproducir bien las contribuciones, a la energía de interacción, que hace cada grupo.

Energías de interacción (Kcal/mol)	Métodos y bases de cálculo		
	MP2/6-31+G*	B3LYP/6-31+G*	HF/6-31+G*
Modelo1 (NU2058)	-10,1	-7,7	-6,3
Modelo2 (NU2017)	-7,1	-4,2	-3,7
Modelo3 (Comp7)	-10,9	-7,7	-6,3
Modelo4 (Comp8)	-10,8	-7,4	-6,0
Modelo5 (Comp9)	---	-7,2	-5,9

Tabla 2.3. Energías de interacción obtenidas en los niveles de cálculo MP2, B3LYP y HF para los diferentes modelos de inhibidores de tipo Purina N-2 sustituidos estudiados.

### II.3.3 El complejo Inhibidor-Leu83

Con el objetivo de investigar la energía de interacción entre los diferentes inhibidores (Figura 2.1) y el residuo de aminoácido Leu83, se adoptó el método MP2 como el adecuado para describir de una manera óptima la energía de interacción de puente de hidrógeno en estos modelos.

Con el fin de reducir la demanda computacional, se utilizaron las estructuras esquematizadas en la Figura 2.2 como modelo de los diferentes inhibidores y del aminoácido en estudio. Las estructuras modelo de la interacción fueron optimizadas totalmente en el nivel de cálculo MP2/6-31+G\*. Las energías de interacción obtenidas con el método ONIOM3 (MP2/6-31G\*: HF/3-21G: PM3) para los diferentes modelos se presentan en la Tabla 2.4.

Con el objetivo de analizar con detalle la interacción de puente de hidrógeno en estudio, se ha dividido la energía de interacción ONIOM ( $-\Delta E^{\text{ONIOM3}}$ ) del complejo inhibidor-Leu83 en tres contribuciones, así:

$$\begin{aligned}
 \Delta E^{\text{ONIOM3}} &= E(\text{Complejo Optimizado}) \\
 &\quad - [E(\text{Inhibidor Optimizado}) \\
 &\quad + E(\text{Modelo completo})] \\
 \\
 &= \Delta E(\text{Alto, A}) + [\Delta E(\text{Medio, AB}) \\
 &\quad - \Delta E(\text{Medio, A})] + [\Delta E(\text{Bajo, ABC}) \\
 &\quad - \Delta E(\text{Bajo, AB})] \\
 \\
 &= \Delta E(\text{Alto, A}) + \Delta \Delta E(\text{Medio, AB-A}) \\
 &\quad + \Delta \Delta E(\text{Bajo, ABC-AB})
 \end{aligned}$$

Donde  $\Delta E(\text{Alto, A})$  es la energía de interacción en la región A, llamada la interacción entre el modelo del inhibidor y el aminoácido Leu83 evaluada a un nivel alto de cálculo.  $\Delta \Delta E(\text{Medio, AB-A})$  es la energía de interacción entre las regiones A y B evaluada a un nivel de cálculo medio, y  $\Delta \Delta E(\text{Bajo, ABC-AB})$  es la interacción entre las regiones AB y C evaluada a un nivel de cálculo bajo. Como se muestra en la Tabla 2.4, el cálculo ONIOM3 (MP2/6-31G\*: HF/3-21: PM3) para el inhibidor NU2017, por ejemplo, presenta unos valores  $\Delta E(\text{Alto, A}) = -6.0$ ,  $\Delta \Delta E(\text{Medio, AB-A}) = -0.4$ , y  $\Delta \Delta E(\text{Bajo, ABC-AB}) = -5.7$  kcal/mol. Esto significa que la interacción total entre el compuesto NU2017 y el aminoácido Leu83,  $\Delta E(\text{Alto, A}) + \Delta \Delta E(\text{Medio, AB-A}) = -6.4$  kcal/mol. La interacción del complejo modelo proviene de una única interacción de puente de hidrógeno (Figura 2.8) formada entre el átomo N3 del anillo pirimidínico del inhibidor NU2017 y el NH del grupo amido del aminoácido Leu83 y que tiene un valor de energía de  $-6.0$  kcal/mol.

Es importante resaltar que la distancia entre el átomo aceptor y el dador, es la más corta en la serie de compuestos estudiados (Ver Tabla 2.5). A su vez la distancia entre los átomos de O, que pertenece al grupo carbonilo del backbone del aminoácido Leu83, y el átomo C2, perteneciente al anillo pirimidínico del compuesto NU2017 es más larga de lo

Energías (kcal/mol)	Sistemas Modelo				
	Modelo1 (Nu2058)	Modelo2 (Nu2017)	Modelo3 (Comp7)	Modelo4 (Comp8)	Modelo5 (Comp9)
$\Delta E$	-16.2	-12.1	-16.7	-16.3	-15.8
$\Delta E(\text{Alto, A})$	-9.7	-6.0	-10.6	-10.4	-9.7
$\Delta\Delta E(\text{Medio, AB-A})$	-0.4	-0.4	-0.3	-0.8	-1.1
$\Delta\Delta E(\text{Bajo, ABC-AB})$	-6.2	-5.7	-5.8	-5.1	-4.9
$IC_{50}$ ( $\mu\text{M}$ )	12.0	16.0	5.0	2.8	1.2

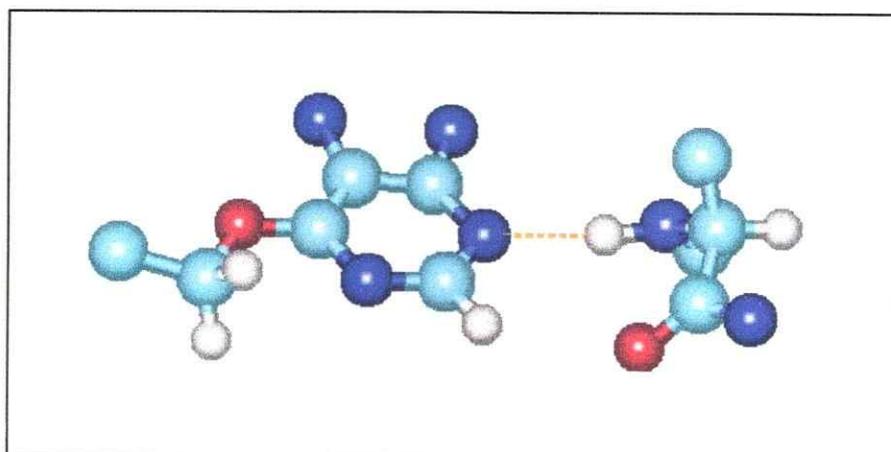
Tabla 2.4. Energías de interacción ONIOM3 (MP2/6-31G\*: HF/3-21G: PM3) obtenidas para los diferentes modelos de inhibidores tipo Purina C-2 sustituidos estudiados.

esperado (3.437 Å), para enlaces de hidrógeno normales. Las restantes distancias interatómicas reportadas pertenecen a los modelos de inhibidores estudiados y presentan un comportamiento que varía de acuerdo al tamaño del grupo sustituyente en la posición C2 del anillo pirimidínico del inhibidor. A un mayor tamaño de grupo hay una mayor separación interatómica entre los átomos dador y aceptor de puente de hidrógeno.

La interacción del inhibidor NU2017 con toda la región bisagra de la enzima tiene una energía  $\Delta\Delta E$  (Bajo, ABC-AB) = -5.7 kcal/mol, o sea que esencialmente más de la mitad de la energía de interacción de este inhibidor con la región bisagra de la enzima CDK2 ( $\Delta E^{\text{ONIOM3}}$ ) proviene de la interacción de puente de hidrógeno mencionada.

Si se analizan de una manera similar las diferentes componentes de la energía ONIOM3 para los demás modelos de inhibidor se puede notar como las energías  $\Delta E^{\text{ONIOM3}}$  y  $\Delta\Delta E$  (Bajo, ABC-AB) presentan una variación dependiente del tamaño del grupo sustituyente en la posición C2 del anillo pirimidínico del modelo de inhibidor. Los grupos sustituyentes pequeños como -H, -NH y -NHCH<sub>3</sub> van favoreciendo, en ese orden, la energía de interacción del inhibidor con la región bisagra de la enzima CDK2.

Los grupos más grandes como -NHCH<sub>2</sub>CH<sub>3</sub> y -NHCH<sub>2</sub>-(CH<sub>3</sub>)<sub>2</sub> desfavorecen la interacción del inhibidor con esta región y por lo tanto presentan energías de interacción menos favorables. Este comportamiento, puede deberse a efectos de tipo estérico con



**Figura 2.8.** Estructura del complejo entre NU2017 y Leu83 obtenida a partir de la optimización total en el vacío al nivel de cálculo ONIOM3 (MP2/6-31G\*: HF/3-21G: PM3).

Distancias (Angstroms)	Sistemas Modelo				
	Modelo1 (Nu2058)	Modelo2 (Nu2017)	Modelo3 (Comp7)	Modelo4 (Comp8)	Modelo5 (Comp9)
O-N	2.891	3.437*	2.866	2.952	3.021
N-N	3.080	3.059	3.098	3.097	3.103

\* La distancia interatómica reportada es del tipo O-C.

**Tabla 2.5.** Distancias interatómicas entre los átomos dador y aceptor de puente de hidrógeno en los diferentes modelos ONIOM3 estudiados.

grupos tan voluminosos como el isopropilo, que provocan que la geometría de los PH no sea la óptima y por lo tanto que las distancias y los ángulos entre los átomos dadores y aceptores de PH varíen con el tamaño del grupo sustituyente (Tabla 2.6). A pesar de esto las energías de interacción de PH se ven favorecidas hasta cierto punto con el carácter y tamaño del grupo sustituyente en la posición C2 del anillo pirimidínico del inhibidor.

Los valores de  $\Delta E$  (Alto, A) +  $\Delta\Delta E$  (Medio, AB-A), que corresponden a la energía de interacción de los diferentes inhibidores con el aminoácido Leu83, se presentan en la Tabla 2.7. Es importante destacar que los valores de la energía de PH corresponden muy bien con los valores de la actividad biológica reportada para estos compuestos, a

Ángulos (Degrees)	Sistemas Modelo				
	Modelo1 (Nu2058)	Modelo2 (Nu2017)	Modelo3 (Comp7)	Modelo4 (Comp8)	Modelo5 (Comp9)
C=O---H	127.412	133.278	130.841	144.230	151.589
N-H---O	157.708	130.077*	154.984	160.655	145.390
N---H-N	175.298	171.996	178.096	172.075	170.412

\*El ángulo de enlace reportado en este modelo es del tipo C-H---O.

Tabla 2.6. Ángulo de enlace entre los átomos que participan en los puentes de hidrógeno en los diferentes modelos ONIOM3.

excepción del compuesto 9, el cual tiene el grupo sustituyente de tipo isopropilo y que debe presentar problemas de impedimento estérico en la región bisagra de la enzima CDK2. Estos resultados indican que un cálculo más preciso de las interacciones de PH en esta clase de inhibidores podría ser útil en el estudio e identificación de nuevas moléculas con características similares.

Compuesto Modelo	IC <sub>50</sub> ( $\mu$ M)	$\Delta E$ (Alto, A) + $\Delta\Delta E$ (Medio, AB-A) (kcal/mol)	Electrofilia (eV)	Blandura Global (a.u.)	Dureza (eV)	Potencial Químico (eV)
Nu2017	16,0	-6.4	0,680	2,314	11,758	-4,000
Nu2058	12,0	-10.1	0,607	2,611	10,420	-3,556
Comp7	5,0	-10.9	0,542	2,654	10,255	-3,334
Comp8	2,8	-11.2	0,510	2,633	10,333	-3,245
Comp9	1,2	-10.8	0,522	2,641	10,302	-3,280

Tabla 2.7. Valores de los descriptores de reactividad global, la actividad biológica expresada como Concentración Inhibitoria media y la energía de interacción de PH de los inhibidores con el aminoácido Leucina 83 (Leu83) obtenida con el método ONIOM3 (MP2/6-31G\*:HF/3-21G:PM3).

### II.3.4 Geometría de los puentes de hidrógeno

Al analizar con más detalle la Tabla 2.6; en la cual se reportan los ángulos de los puentes de hidrógeno presentes en la interacción de los inhibidores y del aminoácido Leu83, se puede ver la influencia que tiene la geometría del PH en la energía de interacción

estudiada. Normalmente el ángulo N-H---O es mayor que  $150^\circ$  y el ángulo C=O---H tiene un rango más amplio que va desde los  $100^\circ$  hasta los  $180^\circ$ .<sup>28</sup>

Los ángulos obtenidos a partir de la optimización con el método ONIOM3 muestran que a medida que se aumenta el tamaño del grupo sustituyente en la posición C2 del anillo pirimidínico, el ángulo C=O---H aumenta su valor, o sea que toma una dirección cada vez más extendida lo cual favorece la interacción de PH.

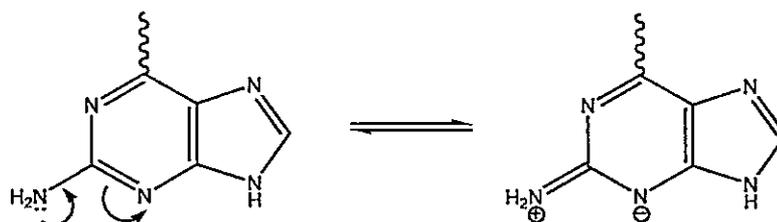
Por otra parte, el ángulo N-H---O tiene un valor menor que  $150^\circ$  en la interacción del compuesto NU2017 con Leu83; lo cual indica claramente la ausencia de uno de los PH en este complejo. Igualmente el compuesto 9 (Modelo5) presenta un valor para este ángulo de aproximadamente  $145^\circ$ , y aunque es un valor que está por debajo del óptimo, es posible formar un puente de hidrógeno para este compuesto pero con un valor de energía menos favorable que sus análogos con sustituyentes de tipo metilo y etilo. Esto se puede explicar por el volumen que tiene el sustituyente de tipo isopropilo en la posición C2 del anillo pirimidínico.

La tendencia a la linealidad del PH expresada en el valor del ángulo N---H-N es favorecido por los sustituyentes -H, -NH y -NHCH<sub>3</sub>, expresándose en el aumento del valor del ángulo que tiende cada vez a ser más extendido. Con sustituyentes más grandes del tipo etil e isopropil, el valor del ángulo empieza a disminuir, provocando con esto que la interacción de puente de hidrógeno sea menos efectiva. La geometría de los dos puentes de hidrógeno, en conjunto, determina los valores de energía de interacción en el complejo modelo Inhibidor-Leu83.

### II.3.5 Cálculo de los índices de reactividad global.

En la Tabla 2.7 se resumen los índices globales de reactividad calculados para los diferentes inhibidores bajo estudio. El valor del potencial químico electrónico ( $\mu$ ) calculado para el modelo del aminoácido Leu83 tiene un valor de aproximadamente -4.16 eV. El potencial químico electrónico de cada uno de los inhibidores estudiados predice correctamente la dirección en que fluyen los electrones en la interacción estudiada, o sea desde cada uno de los inhibidores hacia el backbone del aminoácido Leu83.

Presumiblemente puede existir un mayor flujo de electrones en la región de interacción que se encuentra entre los átomos N3 del anillo pirimidínico del inhibidor y el grupo NH del backbone de la Leu83. La otra interacción pertenece al grupo N<sup>2</sup>H del anillo pirimidínico con el grupo C=O de la Leu83, en la cual el flujo de electrones debiera ocurrir en forma inversa y se debería favorecer con sustituyentes que sean electroattractores y mejoren el flujo de electrones desde el backbone de la Leu83 hacia el inhibidor. Si se considera la resonancia electrónica en el anillo pirimidínico de cada inhibidor como se muestra a continuación:



Podemos ver que ocurre un comportamiento contrario para los inhibidores estudiados. Los grupos metilo y etilo ejercen un efecto dador de electrones, hacia el átomo N<sup>2</sup> del anillo pirimidínico del inhibidor, producido por el efecto inductivo que poseen estos grupos alquílicos, lo cual estabiliza la carga positiva sobre este átomo en la estructura resonante. Con un grupo sustituyente como el isopropilo (Comp9) se pierde este efecto inductivo y además no hay una geometría óptima del enlace de hidrógeno. Esto influye en la energía de interacción de puente de hidrógeno. En la Figura 2.9 se representa la correlación entre la actividad inhibitoria media (IC<sub>50</sub>) y el potencial químico electrónico de cada uno de los inhibidores. A partir de este gráfico se puede concluir que hay una buena correlación entre los dos valores, indicando que los grupos sustituyentes alquílicos en la posición N<sup>2</sup> del anillo pirimidínico estabilizan la carga positiva sobre este átomo y además potencian la transferencia de electrones entre el átomo N3 del inhibidor y el grupo NH del backbone del aminoácido Leu83. Con sustituyentes mayores como el grupo isopropil, priman otras fuerzas como la geometría óptima de puente de hidrógeno y la pérdida del efecto inductivo.

Los valores de la electrofilia también se presentan en la Tabla 2.7. A partir de los resultados numéricos de este descriptor para los diferentes inhibidores, se puede confirmar el comportamiento esperado en la interacción de puente de hidrógeno con respecto a los grupos sustituyentes en la posición C2 del anillo pirimidínico y a los valores de  $\mu$  para cada inhibidor. Si el potencial químico de cada uno de los modelos de los inhibidores estudiados es mayor que el valor de  $\mu$  para el modelo del aminoácido Leu83, significa que los electrones fluyen desde el inhibidor hacia el a.a. Leu83.

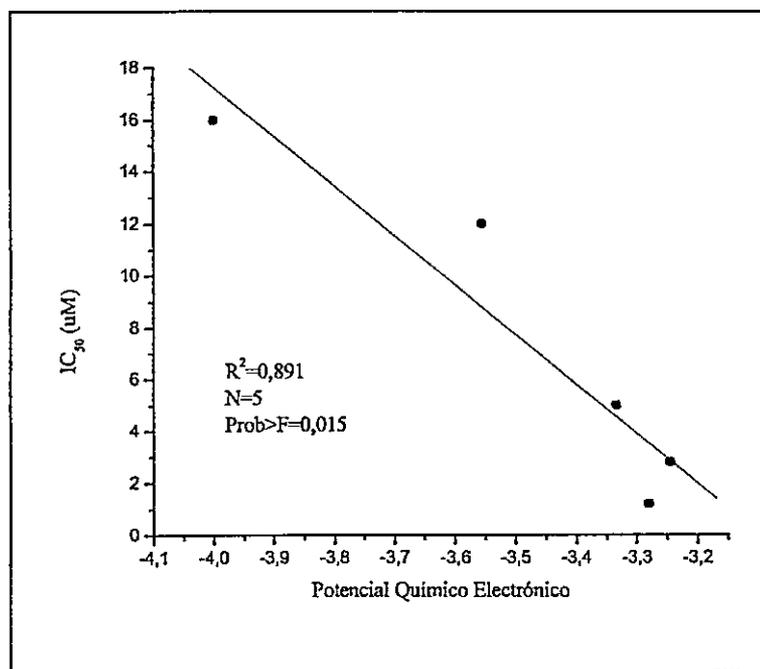


Figura 2.9. Gráfica de correlación entre la actividad media inhibitoria ( $IC_{50}$ ) y el potencial químico electrónico de la serie de inhibidores estudiados.

Los valores de electrofilia confirman este comportamiento, ya que los inhibidores se hacen menos electrofílicos a medida que estabilizamos la carga sobre el átomo  $N^2$  en la estructura resonante con grupos electrodonadores. En la Figura 2.10 se presenta la gráfica de correlación entre la actividad inhibitoria media ( $IC_{50}$ ) de los inhibidores y la electrofilia calculada para los mismos. Los valores de blandura global no muestran una clara correlación con la  $IC_{50}$  reportada para cada inhibidor.

## II.4. Conclusiones.

Los cálculos ONIOM2 y ONIOM3 en los cuales se utilizaron los métodos MP2/6-31G\*, B3LYP/6-31G\* y HF/6-31G\* en el nivel alto y los métodos HF/3-21G y PM3 en los niveles medio y bajo, dieron como resultado estructuras esencialmente iguales para el compuesto NU2058.

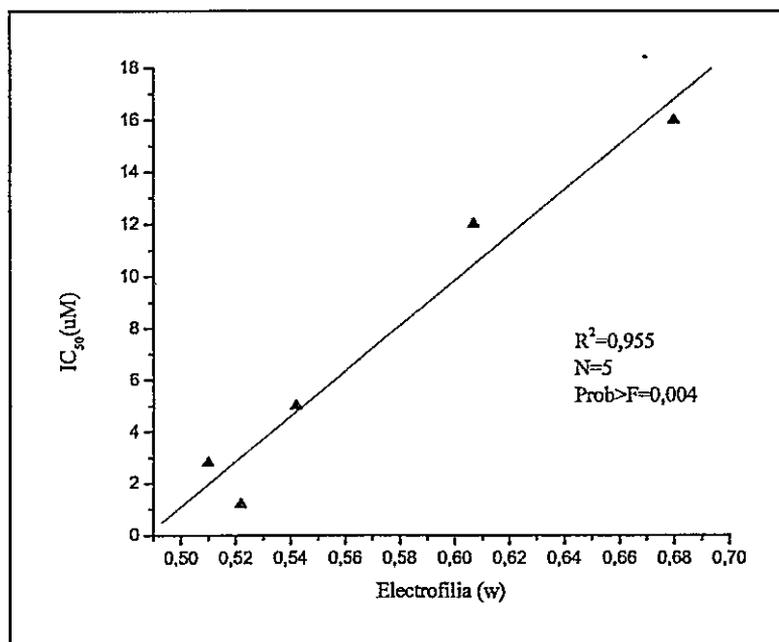


Figura 2.10. Gráfica de correlación entre el valor de la concentración media inhibitoria ( $IC_{50}$ ) y la electrofilia ( $w$ ) de la serie de inhibidores estudiados.

Los complejos modelo inhibidor-Leu83 optimizados al nivel de cálculo MP2/6-31+G\* presentan energías de interacción de puente de hidrógeno más favorables (más negativas) que los modelos optimizados con el método ONIOM3 (MP2/6-31G\*: HF/3-21G: PM3). Esto debido a que en el complejo calculado con el método MP2/6-31+G\* no existe ningún impedimento, ejercido por la conformación del backbone de la proteína, que evite que el modelo del aminoácido Leu83 se mueva libremente.

Las contribuciones de energía de los dos puentes de hidrógeno, analizadas en todos los modelos estudiados, permite concluir que gran parte de la interacción entre la región bisagra y la serie de inhibidores proviene de la interacción de puente de hidrógeno entre

el anillo pirimidínico y el backbone del aminoácido Leucina83 (Leu83). La otra parte de la energía que favorece la interacción del inhibidor con esta parte del sitio activo de la enzima CDK2 proviene de la interacción del resto de la estructura del inhibidor con los otros aminoácidos que pertenecen a la región bisagra.

La geometría de los puentes de hidrógeno está estrechamente relacionada con la energía de interacción de los mismos. Se puede ver claramente a partir de estos resultados, como variaciones en los ángulos de solo algunos grados y en las distancias entre los átomos aceptores y dadores de puente de hidrógeno de solo unas centésimas de Ångstrom, pueden cambiar las energías de interacción en cantidades que están entre 0.5-1.0 kcal/mol.

La energía de interacción de puente de hidrógeno se ve afectada por sustituyentes electrodonadores en la posición N<sup>2</sup> del anillo pirimidínico. Sustituyentes del tipo isopropilo cambian marcadamente la geometría de los puentes de hidrógeno y por tanto hacen menos favorable la interacción del inhibidor con la región bisagra de la enzima. Estas interacciones desfavorables deben ser compensadas con otro tipo de interacción en alguna parte de la enzima diferente a la región bisagra y pueden ser del tipo puente de hidrógeno o interacciones de tipo electrostático con otros aminoácidos del sitio activo de la enzima.

Según Arris C. E. *et al.*<sup>29</sup>, el puente de hidrógeno formado por el grupo 2-NH<sub>2</sub> no es indispensable, debido a que se midió la actividad inhibitoria del compuesto NU2017, el cual carece de este grupo en la posición C2, conservando alguna actividad inhibitoria contra la CDK2. De acuerdo los resultados obtenidos, el puente de hidrógeno efectivamente no se forma, debido a la geometría inadecuada que tienen de los átomos aceptor y dador para la interacción de este compuesto con el aminoácido Leu83. Sin embargo el compuesto forma el otro enlace de hidrógeno a partir de la interacción entre el átomo N3 (del anillo pirimidínico) y el grupo NH del backbone del aminoácido Leu83. Esta interacción tiene un valor de energía de -6.0 kcal/mol, que es notablemente más favorable que la interacción de un puente de hidrógeno perteneciente a la molécula NU2058 que tiene una energía de interacción para los dos PH que forma con el a.a. Leu83 de -9.7 kcal/mol. Suponiendo que las dos interacciones fueran equivalentes,

podría decirse que para un solo PH la energía de interacción sería de aproximadamente -4.8 Kcal/mol. Esta diferencia en la energía de interacción se debe principalmente a efectos de tipo electrónico y estérico que provoca cualquier sustituyente en la posición C2 del anillo pirimidínico, haciendo que la geometría de ese PH cambie y por tanto la energía de interacción se torne menos favorable. Las energías totales de interacción para estos dos compuestos con los aminoácidos de la región bisagra tienen un valor de -16.2 (Nu2058) y -12.1 (NU2017) kcal/mol, que son valores que concuerdan con la actividad inhibitoria reportada para los dos compuestos.

Los resultados obtenidos a partir de los valores de electrofilia y del potencial químico electrónico sugieren que puede existir una relación entre el carácter electrofílico de un compuesto y su actividad biológica medida como IC<sub>50</sub>. En este caso los valores de electrofilia para cada inhibidor se relacionan de una manera inversa con respecto a la IC<sub>50</sub>, indicando esto que grupos sustituyentes electrodonores en la posición C2 del anillo pirimidínico de los inhibidores mejoran la reactividad del compuesto hacia la región conformada por el aminoácido Leu83.

Los valores de electrofilia global ( $\omega$ ) para el aminoácido Leu83 (no incluidos en la Tabla 2.7) son mayores que los valores de electrofilia para cada inhibidor. Esto permite deducir que el comportamiento que tienen los inhibidores con respecto a los sitios de interacción de PH corresponde más bien a uno del tipo nucleofílico. Esto explica de cierta forma el hecho de que los inhibidores menos electrofílicos sean los que presenten una energía de interacción de PH más favorable.

La definición de un modelo del inhibidor y el sitio activo que incluya una cantidad de átomos mayor y que sea por lo tanto más representativo, permitirá caracterizar por medio de este método de cálculo la energía de interacción entre el inhibidor y otros aminoácidos importantes dentro del sitio activo. Los descriptores globales y locales de reactividad aplicados a las estructuras de menor energía, obtenidas con el método ONIOM, nos permitirá describir las interacciones entre el inhibidor y el sitio activo de la CDK2 en el contexto del principio HSAB.

## II.5. Bibliografía.

- [1] Davies, T.G.; Bentley, J.; Arris, C.E.; Boyle, F.T.; Curtin, N.J.; Endicott, J.A.; Gibson, A.E.; Golding, B. T.; Griffin, R. J.; Hardcastle, I. R.; Jewsbury, P.; Johnson, L. N.; Mesguiche, V.; Newell, D. R.; Noble, M. E.M.; Tucker, J. A.; Wang, L. and Whitfield, H.J. Structure-based design of a potent purine-based cyclin-dependent kinase inhibitor *Nature Structural Biology*, **2002**, 9(10), 745-749.
- [2] Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, 16, 1170.
- [3] Dapprich, S.; Komaromi, I.; Byun, K.S.; Morokuma, K.; Frisch, M. J. *J. Mol. Struct. (THEOCHEM)* **1999**, 461-462, 1.
- [4] Humbel, S.; Sieber, S. and Morokuma, K. *J. Chem. Phys.* **1996**, 105, 1959.
- [5] Matsubara, T.; Sieber, S. and Morokuma, K. *Int. J. Quant. Chem.* **1996**, 60, 1101.
- [6] Koga, N. and Morokuma, K. *Chem. Rev.* **1991**, 91, 823.
- [7] Veillard, A. *Chem. Rev.* **1991**, 91, 743.
- [8] Hardcastle, I. R.; Arris, C. E.; Bentley, J.; Boyle, F. T.; Chen, Y.; Curtin, N. J.; Endicott, J. A.; Gibson, A. E.; Golding, B. T.; Griffin, R. J.; Jewsbury, P.; Menyerol, J.; Mesguiche, V.; Newell, D. R.; Noble, M. E. M.; Pratt, D. J.; Wang, L. Z.; Whitfield, H. J. *J. Med. Chem.* **2004**, 47, 3710-3722.
- [9] Stewart, J. J. P. *J. Comp. Chem.* **1989**, 10, 209.
- [10] Stewart, J. J. P. *J. Comp. Chem.* **1989**, 10, 221.

- [11] Boys, S.F. and Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
- [12] Gaussian 98 (Revision A.11), M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, P. Salvador, J. J. Dannenberg, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 2001.
- [13] Parr R. G. and Yang W. *Density Functional Theory of Atoms and Molecules* (Oxford Univ. Press, Oxford) **1989**.
- [14] Parr, R. G.; Pearson, R. G. *J. Am. Chem. Soc.* **1983**, *105*, 7512.
- [15] Yang, W.; Parr, R. G. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 6723.
- [16] Parr, R. G.; von Szentpaly, L.; Liu, S. *J. Am. Chem. Soc.* **1999**, *121*, 1922.
- [17] Koopmans, T. A. *Physica* **1933**, *1*, 104.
- [18] Head-Gordon, M.; Pople, J. A. and Frisch, M. J. *Chem. Phys. Lett.* **1988**, *153*, 503.

- [19] Frisch, M. J.; Head-Gordon, M. and Pople, J. A. *Chem. Phys. Lett.* **1990**, *166*, 275.
- [20] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
- [21] Lee, C.; Yang, W. and Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- [22] Michlich, B.; Savin, A.; Stoll, H. and Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200.
- [23] Roothan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69.
- [24] Pople, J. A. and Nesbet, R. K. *J. Chem. Phys.* **1954**, *22*, 571.
- [25] McWeeny, R. and Dierksen, G. *J. Chem. Phys.* **1968**, *49*, 4852.
- [26] Stewart, J. J. P. *J. Comp. Chem.* **1989**, *10*, 209.
- [27] Stewart, J. J. P. *J. Comp. Chem.* **1989**, *10*, 221.
- [28] Taylor, R.; Kennard, O.; Versichel, W. *J. Am. Chem. Soc.* **1983**, *105*, 5761-5766.
- [29] Arris C.E. *et al.* Identification of Novel Purine and Pyrimidine Cyclin-Dependent Kinase Inhibitors with Distinct Molecular Interactions and Tumor Cell Growth Inhibitions Profiles. *J. Med. Chem.* **2000**, *43*, 2797-2804.

## CAPÍTULO III

### La Energía de Interacción QM/MM como Índice de la Actividad Biológica

#### III.1 Introducción

Al final del capítulo anterior se expresó la necesidad de utilizar un método de cálculo en el cual se pudiera considerar un modelo del sistema mucho más grande y por lo tanto más representativo del sistema "real". Algunos de los problemas encontrados al aplicar el método ONIOM a nuestros sistemas radicaron en dos puntos principalmente:

1. La imposibilidad de lograr converger los cálculos de energía en modelos que incluyeran más aminoácidos de los que consideramos en el sistema final. Esto pudo deberse a la falta de la capacidad de cálculo apropiada para el tratamiento de este tipo de sistemas.
2. La necesidad de mantener los átomos, que pertenecen a la región con nivel de cálculo más bajo, congelados en sus posiciones cristalográficas con el fin de que la estructura del modelo se mantenga lo más cercana al sistema de partida. Esto puede conducir a problemas de convergencia de la energía en la región QM y a valores de distancias de enlace, ángulos de enlace y ángulos diedros ligeramente distorsionados. Sin embargo, Kuno *et al.*<sup>1</sup> han aplicado exitosamente el método en el estudio de la interacción de la nevirapina con la enzima transcriptasa reversa del HIV-1.

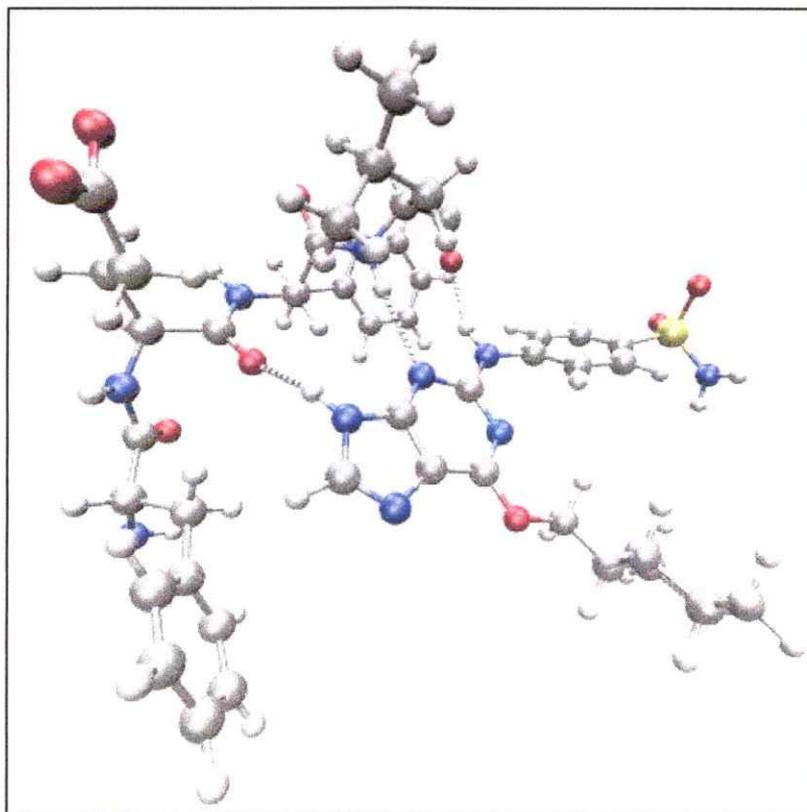
Además de estos dos problemas que son básicamente técnicos, es bien sabido que existen otros inconvenientes más importantes con el método ONIOM. Uno de ellos es que este método no incluye efectos de polarización de la parte cuántica QM debidos a los átomos o más bien residuos pertenecientes a la parte descrita por mecánica molecular MM.<sup>2</sup> Para tratar de solucionar estos problemas se decidió aplicar otro método híbrido de cálculo, el cual permitiera incluir los efectos de polarización mencionados y en

general los efectos del medio enzimático y del solvente, como también incluir todos los residuos y subsistemas pertenecientes a la enzima en estudio. Este método es el llamado Mecánica Cuántica/Mecánica Molecular (Quantum Mechanics/Molecular Mechanics (QM/MM)) que nos permitió resolver los problemas antes mencionados.

Por lo tanto en este capítulo se discuten los resultados obtenidos al aplicar el método híbrido de cálculo QM/MM a un conjunto de 5 inhibidores de la CDK2 que pertenecen a la familia de las 6-ciclo-hexil-metoxi-purinas N<sup>2</sup>-sustituidas. Varios autores, con la ayuda del diseño de drogas basado en la estructura, han explorado cuidadosamente este núcleo básico. De esta forma, ha sido posible identificar interacciones importantes para esta clase de compuestos dentro del sitio activo de algunas CDK's. Las características más importantes encontradas hasta ahora son: (i) El triplete de enlaces de hidrógeno formado entre los diferentes compuestos estudiados y la región bisagra presente en la CDK1 y la CDK2. Esta característica confiere a los compuestos una orientación diferente dentro del sitio activo de la enzima con respecto a la orientación que toman otros compuestos como el flavopiridol y la olomousina, y por consiguiente tiene consecuencias directas sobre la inhibición de la enzima y del crecimiento celular (Figura 3.1).<sup>3</sup> (ii) La CDK2 tiene una cavidad poco profunda en la parte posterior del bolsillo de unión del ATP, que esta definida por la cadena lateral del a.a. Phe80 (Figura 3.2). Aunque ésta no es explotada por el sustrato natural ATP, algunos inhibidores, incluyendo la staurosporina<sup>4,5</sup>, indirubina<sup>6</sup>, e hymenialdisina<sup>7</sup> pueden empaquetar los lados de sus anillos hidrofóbicos en este sitio. El anillo de cinco miembros de varias familias de purinas, está próximo al a.a. Phe80. Por lo tanto, queda un espacio que puede ser aprovechado para extender el sistema de anillos con grupos hidrofóbicos pequeños. Por ejemplo la olomousina tiene un grupo metilo en la posición N9<sup>11</sup>, y los compuestos roscovitina, OL567 y purvanalol poseen un grupo isopropilo en la misma posición.

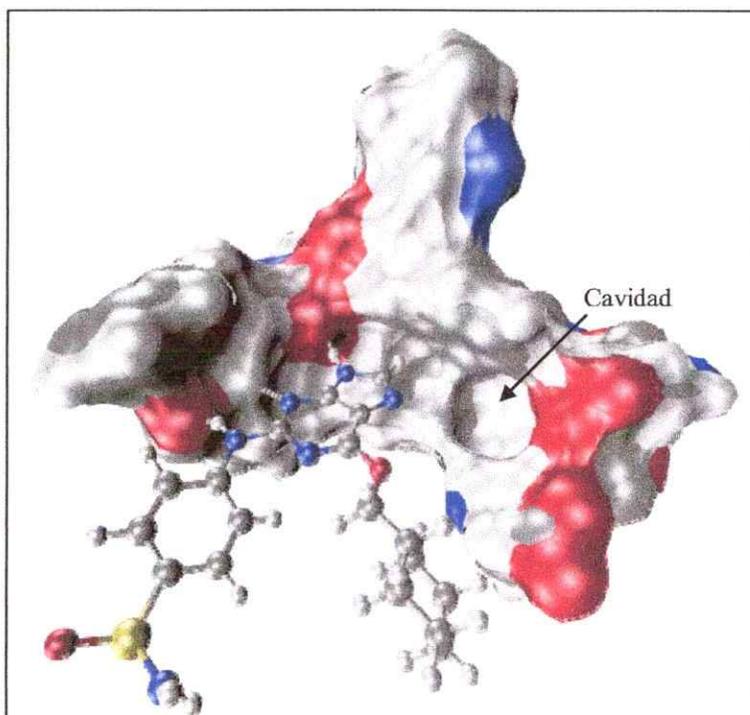
(iii) Una unión óptima se presenta con un sustituyente alifático O<sup>6</sup> de tamaño moderado que se empaqueta fuertemente contra una zona de carácter hidrofóbico que está presente en el lazo rico en residuos de Glicina, y que está centrado en el aminoácido Val18. Ésta es una interacción promovida por las restricciones conformacionales impuestas en el

anillo tipo ciclohexilmetil o ciclohexenilmetil. Así, el compuesto original O<sup>6</sup>-ciclohexilmetilguanina (NU2058) es el punto de partida preferido para explorar otras áreas del sitio activo de la quinasa <sup>8</sup>, y finalmente, (iv) las interacciones establecidas con la “superficie específica”. Estas tienen que ver con residuos que se encuentran fuera del bolsillo de unión del ATP que se considera altamente conservado.<sup>9, 10</sup> Existen secuencias diferentes de aminoácidos en esta región entre los diferentes miembros de la familia de CDK's y se ha dicho que aprovechar esas diferencias en el diseño de nuevas moléculas les podría aportar selectividad por los diferentes tipos de CDK. Por ejemplo, los residuos His84, Gln85 y Lys89 de la CDK2 son reemplazados por ácido aspártico, glutamina y treonina, respectivamente, en la CDK4. Estudios con la olomoucina <sup>11</sup> y roscovitina <sup>12</sup> han demostrado que son posibles grandes mejoras en potencia, además de especificidad, cuando se explota esta “superficie específica”.



**Figura 3.1.** Puentes de hidrógeno típicos formados entre los inhibidores de CDK's y la región bisagra de la enzima.

Adicionalmente, estudios de análisis estructural han indicado que un anillo aromático en la posición N2 de NU2058 podría mejorar la actividad inhibitoria con respecto a la CDK1 y la CDK2 (Figura 3.3).<sup>13</sup> Este fue el caso encontrado con el compuesto resultante de estos estudios y el cual mantiene especificidad para la inhibición del par CDK1/CDK2 sobre la CDK4. Se puede obtener potencia adicional con la presencia de un grupo en la posición 4' en el anillo aromático que sea capaz de donar un enlace de hidrógeno. El compuesto resultante de todas estas investigaciones previas basadas en la estructura es el NU6102 (Figura 3.3). Este inhibidor es selectivo y uno de los inhibidores más activos de CDK2 descritos hasta ahora.<sup>13, 14</sup> A pesar de este significativo progreso en la relaciones estructura-actividad para esta clase de compuestos anticáncer, permanece la necesidad de encontrar un inhibidor de amplio espectro que pueda, por ejemplo, inhibir selectivamente y a bajas concentraciones  $IC_{50}$ , del rango nanomolar, las enzimas CDK1, CDK2, CDK4 y CDK6.

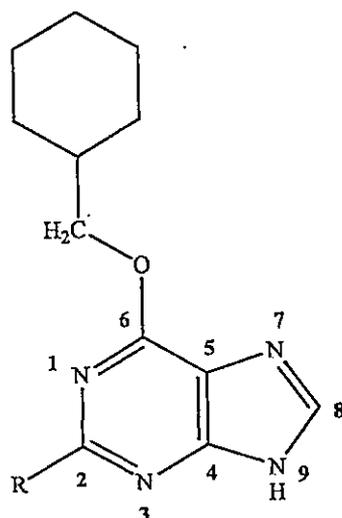


**Figura 3.2.** Cavidad poco profunda formada por la cadena lateral del aminoácido Phe80 en el sitio activo de la CDK2. Los aminoácidos de la región bisagra están representados en forma de superficie y el inhibidor NU6102 en esferas y barras.

Adicionalmente se han reportado pocos estudios teóricos sobre las CDK's basados en cálculos clásicos ó mecánico-cuánticos. Por ejemplo, Cavalli *et al.*, han desarrollado parámetros de campos de fuerzas basados en cálculos mecánico-cuánticos para un sistema modelo, útiles para la simulación de la unión del ATP a las CDK's.<sup>15</sup> Las simulaciones de dinámica molecular realizadas con los nuevos parámetros calculados mostraron que la CDK1 debe ser más estable que la CDK2, implicando esto que las diferencias en la flexibilidad de la estructura de la proteína entre las CDK's deben ser consideradas en el diseño de nuevos inhibidores selectivos. Estudios computacionales más recientes sobre la CDK2 realizados por Sims *et al.*, han establecido un modelo de energía libre de unión para inhibidores de CDK2 dentro del marco del modelo de solvente continuo.<sup>16</sup> Esta aproximación mostró ser exitosa en la reproducción de las actividades inhibitorias relativas de compuestos análogos del flavopiridol. Sobre esta base, ellos sugirieron nuevos inhibidores eventualmente más potentes que el compuesto original.

En un artículo más reciente, los mismos autores implementaron el modelo previamente desarrollado de optimización de cargas basado en un solvente representado mediante un dieléctrico continuo con un algoritmo de programación cuadrática simple y la ecuación de Poisson-Boltzmann.<sup>17</sup> Este método les permitió calcular el mejor juego de cargas puntuales, para un ligando o para una región del ligando basada en la forma del receptor y del ligando, y las cargas parciales del receptor; por medio de la optimización de la energía libre de unión obtenida a partir del modelo de solvente continuo.<sup>16</sup> Sin embargo, existen numerosas limitaciones para la energía libre de unión del modelo de solvente continuo.<sup>16, 18, 19</sup> En su caso particular, ellos utilizaron un modelo de conformación fija en el cual se asume que el ligando y el receptor no experimentan un cambio conformacional en la unión. Por tanto, el considerar la flexibilidad de la proteína es indispensable en la evaluación acertada de la afinidad de unión del ligando<sup>20, 21, 22, 23</sup>. Además es conocido que la plasticidad conformacional es un sello distintivo del dominio catalítico de las proteínas quinasas.<sup>24</sup> Como siempre, es inevitable un sacrificio entre la exactitud y el gasto computacional. Las funciones empíricas (Empirical scoring functions) diseñadas para la estimación rápida de afinidades de unión y utilizadas en

monitoreo virtual de alto rendimiento (high-throughput virtual screening) frecuentemente dejan



COMPUESTO 2 R = 2-anilino

COMPUESTO 3 R = 2-(4'-sulfamoil-anilino) (Nu6102)

COMPUESTO 7 R = 2-metil-amino

COMPUESTO 9 R = 2-(1-metil-etilamino)

COMPUESTO 25 R = 2-(4'-hidroxi-anilino)

Figura 3.3 Estructuras de los derivados de las purinas O<sup>6</sup>-ciclo-hexilmetil N<sup>2</sup>-sustituidas de Hardcastle *et al.*<sup>14</sup> También se muestra la numeración de los átomos del fragmento purina.

de lado la flexibilidad conformacional de la proteína, los efectos entrópicos y la (de)solvatación durante el proceso de unión.<sup>25</sup> En el otro extremo del rango de exactitud están las técnicas basadas en las simulaciones de dinámica molecular, introduciendo explícitamente moléculas de solvente, que hacen uso de los métodos de perturbación de la energía libre (FEP), integración termodinámica (TI) y energía de interacción lineal (LIE).<sup>26, 27</sup> Con un muestreo amplio, los métodos de dinámica molecular pueden calcular la energética de las interacciones proteína/ligando con alta precisión, incluyendo

todos los efectos entrópicos y de solvente como también la flexibilidad del receptor. Los métodos híbridos de mecánica-cuántica/mecánica molecular (QM/MM) han sido desarrollados para cálculos con proteínas y son ahora ampliamente utilizados para el análisis de las reacciones de tipo enzimático,<sup>28, 29, 30</sup> especialmente en el contexto de la catálisis enzimática, la cual tiene que ver con los cambios en la energía libre de unión de los reactivos, del estado de transición y del producto. Entonces, existe una clara analogía en los métodos computacionales utilizados en el estudio de las reacciones enzimáticas y en el estudio de diseño de drogas. En los métodos híbridos QM/MM, las especies de tipo ligando/sustrato son tratadas explícitamente por medio de un modelo QM,<sup>31, 32, 33</sup> el cuál en general, es sistemático y puede incluso llegar a ser exacto. El ambiente de la proteína y el solvente son representados por campos de fuerza MM, los cuales son computacionalmente eficientes. Tratar al ligando con el modelo QM y la proteína con el modelo MM tiene varias ventajas: (i) no se requiere la derivación de parámetros para cada nuevo ligando, (ii) se incluyen efectos cuánticos tales como la polarización del ligando después de la unión, (iii) se describe gran parte del sistema clásicamente, permitiendo con esto realizar cálculos rápidos y un muestreo suficiente, y (iv) el modelo QM puede ser sistemáticamente mejorado. El último aspecto es particularmente importante porque si se estima que el modelo semi-empírico QM no es aplicable a un compuesto en particular, siempre existe un camino alternativo en los métodos *ab initio* de la mecánica cuántica para mejorar la exactitud de los resultados a través del uso de la teoría de funcionales de la densidad (DFT)<sup>34</sup> o la teoría de los orbitales moleculares. Existen estudios computacionales recientes que aplican exitosamente los métodos híbridos QM/MM a la investigación de las interacciones proteína-ligando en el sistema de la proteasa del HIV-1<sup>35</sup> y en el sistema de la enzima Tripsina.<sup>36</sup> Nosotros investigamos en esta parte de la tesis si es posible evaluar por medio de los métodos híbridos QM/MM, la correlación entre las energías de interacción y los datos biológicos de actividad inhibitoria IC<sub>50</sub> de un grupo de inhibidores de CDK2. Se presenta también un análisis detallado de las interacciones de cada inhibidor dentro del sitio activo de la enzima CDK2. Finalmente, se realiza una partición de la energía de interacción en sus componentes que son la energía electrostática y la energía de van der

Waals y por medio de un análisis de regresión múltiple a los datos de energía de interacción y sus componentes se puede comprobar que el modelo podría predecir nuevos y más potentes inhibidores.

## III.2 Modelos y Métodos

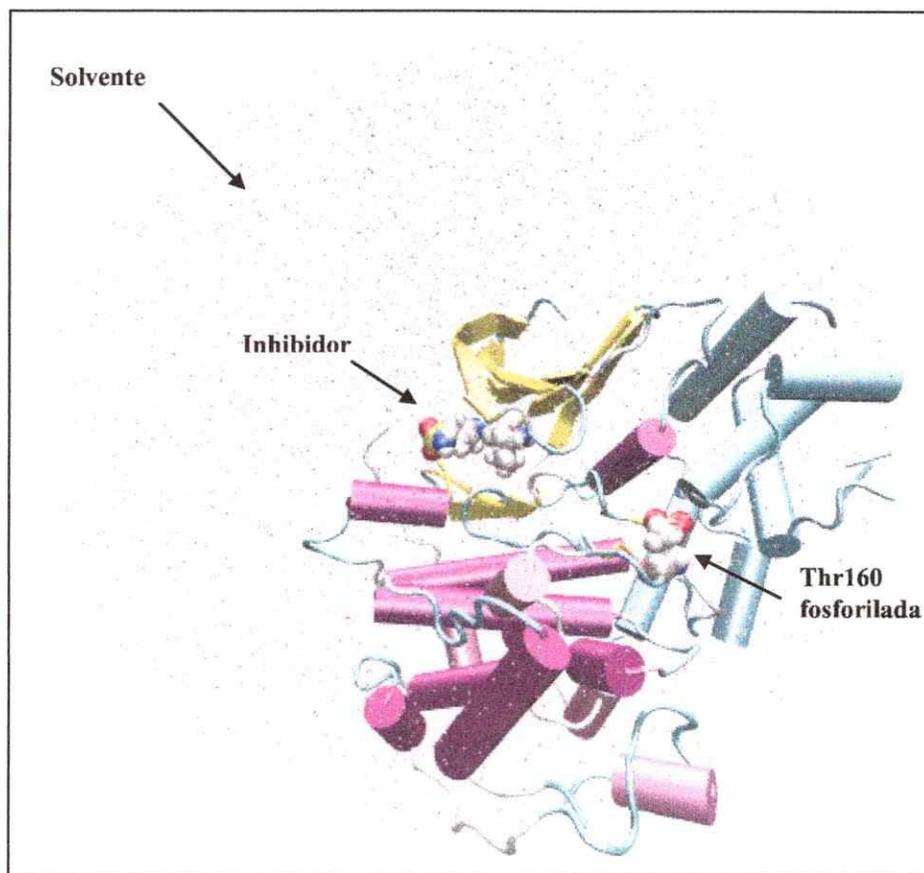
### III.2.1 Simulaciones

Todas las simulaciones fueron realizadas utilizando el programa DYNAMO.<sup>37</sup> Las coordenadas iniciales para los cálculos QM/MM fueron tomadas de la estructura cristalográfica de rayos X del inhibidor Nu6102 unido al complejo Thr160-CDK2-ciclina A fosforilado (T160pCDK2-ciclinaA) (Código en el Protein Data Bank: 1H1S).<sup>14</sup> Los aminoácidos Arg297 y Leu298 fueron adicionados con ayuda del programa MOLDEN<sup>38</sup> a la subunidad de CDK2 de la estructura cristalográfica inicial, con el fin de obtener un sistema de la proteína completo. Para modelar el ambiente acuoso se utilizó una caja de 79.5 Å por lado de moléculas de agua equilibradas, centrada en el centro geométrico de cada inhibidor, para solvatar el sistema T160pCDK2-ciclinaA-ligando. Los residuos de aminoácido ubicados a una distancia mayor de 25 Å del centro de la caja fueron fijados durante la simulación. De esta forma, se obtuvieron sistemas de diferente tamaño de acuerdo al inhibidor y a los átomos congelados a partir de dicha distancia (Tabla 3.1).

Los inhibidores dentro del sitio activo del ATP se eligieron como el subsistema QM, descrito a un nivel semi-empírico AM1,<sup>39</sup> mientras que el resto del complejo T160pCDK2-ciclinaA y las moléculas pertenecientes a las aguas de cristalización de la estructura y aquellas que forman parte de la caja de solvente se describen como subsistema MM utilizando el potencial OPLS-AA<sup>40</sup> y el potencial flexible TIP3P,<sup>41</sup> respectivamente. No existen átomos link entre los subsistemas QM y MM y por ello no necesitamos considerar ningún tratamiento especial para completar la valencia de los átomos frontera de la parte QM.<sup>42</sup> El modelo final se muestra en la Figura 3.4.

El resto de los sistemas estudiados se obtuvieron a partir de este modelo inicial, solo removiendo unos pocos átomos del inhibidor Nu6102 (Figura 3.1). Inicialmente, los

átomos de hidrógeno del sistema fueron relajados utilizando la subrutina de gradientes conjugados<sup>43</sup> implementada en el programa DYNAMO. El número de iteraciones fue establecido en 6000 y el gradiente de tolerancia en 1.0 kJ/mol. En seguida el sistema fue calentado hasta una temperatura de 300 K por medio de una secuencia de simulaciones de dinámica molecular.



**Figura 3.4.** Modelo final del sistema. La CDK2 está representada en forma de cilindros (hélices alfa) y en flechas planas amarillas (láminas beta). La CiclinaA está representada en forma de cilindros (hélices alfa). Por último el inhibidor Nu6102 y la Thr160 están representados en forma de superficie de van der Waals.

Después el sistema fue equilibrado adicionalmente durante 100 picosegundos (ps) utilizando el ensemble NVT a 300 K. Los datos para el análisis de la energía de interacción fueron recogidos en una dinámica molecular posterior de 250 ps a 300 K, utilizando un intervalo de tiempo de 1.0 femtosegundo (fs) para resolver las ecuaciones

de movimiento. Se utilizó una distancia de corte de 15 Å para calcular las interacciones de energía no enlazantes con una función de cambio en la región de 10.0 a 13.5 Å para llevar la energía de interacción a cero en esta región. Para realizar un análisis posterior estructural y de energías se archivó una configuración proteína-ligando cada 15 intervalos de tiempo. El procedimiento computacional descrito se aplicó a todos los 5 ligandos investigados.

Sistema	Nº Átomos totales	Nº Átomos totales congelados
T160pCDK2- ciclinaA+TIP3P+Compuesto2	26 648	15 134
T160pCDK2- ciclinaA+TIP3P+Compuesto3	27 159	15 280
T160pCDK2- ciclinaA+TIP3P+Compuesto7	25 882	14 750
T160pCDK2- ciclinaA+TIP3P+Compuesto9	26 472	15 028
T160pCDK2- ciclinaA+TIP3P+Compuesto25	26 676	14 917

Tabla 3.1. Tamaño del sistema de acuerdo al inhibidor estudiado.

### III.2.2 Partición de la energía

La energía para un sistema QM/MM puede ser obtenida como:

$$E = \langle \Psi | \hat{H}_0 | \Psi \rangle + \langle \Psi | \hat{V}_{QM/MM} | \Psi \rangle + E_{mm} \quad (3.1)$$

Donde  $\hat{H}_0$  es el Hamiltoniano en el vacío para el método QM seleccionado,  $\Psi$  es la función de onda polarizada,  $E_{mm}$  representa la energía del campo de fuerza y  $\hat{V}_{QM/MM}$  es el operador de acoplamiento entre los subsistemas QM y MM e incluye los términos de energía electrostática y de van der Waals (vdW).

La ecuación (3.1) puede ser escrita de una forma diferente con el fin de definir la energía de interacción QM/MM. En nuestro esquema de partición la energía total se puede escribir como la suma de la energía en el vacío del subsistema QM, la energía de interacción QM/MM y la energía MM:

$$E = E_{QM} + E_{QM/MM} + E_{MM} = \langle \Psi^0 | \hat{H}_0 | \Psi^0 \rangle + \left[ \langle \Psi | \hat{H}_0 | \Psi \rangle + \langle \Psi | \hat{V}_{QM/MM} | \Psi \rangle - \langle \Psi^0 | \hat{H}_0 | \Psi^0 \rangle \right] + E_{mm} \quad (3.2)$$

donde  $\Psi^0$  da cuenta de la función de onda no polarizada (fase gaseosa) del subsistema QM. La energía de interacción QM/MM puede ser descompuesta adicionalmente considerando que el operador de acoplamiento entre los subsistemas QM y MM incluye un término de energía electrostática y uno de energía de vdW. Ésta última contribución es evaluada usualmente utilizando la expresión típica de Lennard-Jones:

$$E_{vdW, QM/MM} = \sum_{i \in QM} \sum_{j \in MM} 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{R_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{R_{ij}} \right)^6 \right] \quad (3.3)$$

donde  $R_{ij}$  es la distancia entre los centros QM y MM que interaccionan. Como esta última expresión no involucra coordenadas electrónicas, este término de energía no necesita incluirse en la evaluación del ciclo SCF (Self Consistent Field) de la función de onda del subsistema QM.

### III.3 Resultados y Discusión

Inicialmente, en esta sección presentaremos las características estructurales de cada complejo CDK2-inhibidor y después analizaremos las energías de interacción y su relación con la actividad biológica (medida como IC<sub>50</sub>). Los datos de geometrías y los términos de energía han sido promediados durante las dinámicas de producción y la desviación estándar se reporta como un número dentro de un paréntesis acompañando el valor promedio de cada cantidad.

#### III.3.1 Interacciones Ligando-Sitio activo

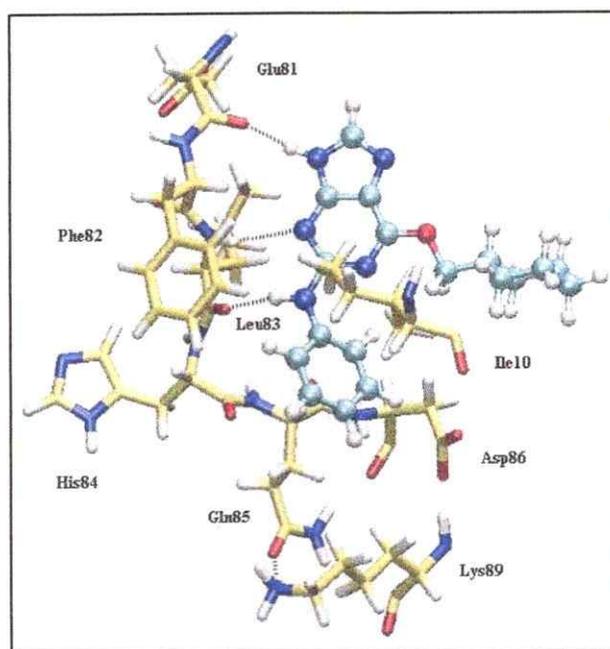
##### III.3.1.1 Compuestos 2 (Nu6094) y 25

Se han desarrollado varios compuestos que compactan su estructura aromática contra la superficie específica de la CDK2, tales como los derivados del purvalanol<sup>44</sup> y la indirubina.<sup>6, 45</sup> El compuesto 2 (Cp2) fue sintetizado inicialmente tomando en cuenta este aspecto. Así, de acuerdo a Davies *et al.*, el grupo anilino de Cp2 se proyecta fuera del sitio de unión del anillo de adenina a través de un túnel hidrofóbico formado por las cadenas laterales de los residuos Phe82 e Ile10, y se compacta contra la superficie de la quinasa, formando una interacción de apilamiento (stacking)  $\pi$ - $\pi$  con el esqueleto peptídico entre los residuos Gln85 y Asp86.<sup>13</sup> En la Figura 3.5 se muestra la estructura final obtenida para este compuesto de las simulaciones de dinámica molecular. Como característica común para los 5 inhibidores analizados aquí, esta la formación de un triplete de enlaces de hidrógeno dentro del sitio de unión del ATP en la CDK2. El grupo NH-9 actúa como un dador de enlace de hidrógeno al grupo carbonilo del enlace peptídico del residuo Glu81, y los grupos N-3 y 2-NH<sub>2</sub>, aceptan y donan un enlace de hidrógeno a los grupos amido y carbonilo, respectivamente, del residuo Leu83.

En la Tabla 3.2 reportamos el valor promedio de las distancias para el triplete de enlaces de hidrógeno descritas anteriormente y obtenidas durante la dinámica molecular completa para cada compuesto. Los valores son muy similares para los 5 compuestos.

A partir del análisis de los datos geométricos obtenidos de las simulaciones de dinámica molecular pudimos observar otras interacciones débiles entre **Cp2** y la proteína, siendo la más significativa la interacción débil C-H $\cdots$  $\pi$  entre el carbono  $\alpha$  del residuo Gln85 y la nube  $\pi$  del anillo anilino del ligando. La distancia desde el átomo de hidrógeno hasta el centroide del anillo es aproximadamente 2.90 Å. Existe otra interacción débil del tipo C-H $\cdots$ O=C entre uno de los hidrógenos del anillo anilino y el grupo carboxilato de residuo Asp86 (**Cp2** Ph-H a Asp86 OD2 = 3.35 Å ( $\pm$  0.49 Å)). Vale la pena mencionar que en esta estructura la posición del residuo Lys89 está inclinada lejos del anillo anilino del ligando y el grupo NH $_3^+$  realiza una interacción de enlace de hidrógeno estabilizante con el grupo carboxilato de la cadena lateral del residuo Gln85 (Figura 3.5).

La potencia de la serie de ligandos pertenecientes a la familia de las C $^2$ -Anilino-O $^6$ -ciclo-hexilmetil-purinas 4'-sustituidas fue evaluada por Hardcastle *et al.*<sup>14</sup> Ellos encontraron que la sustitución 4'-hidroxi en esta serie generó el compuesto 25, el cual

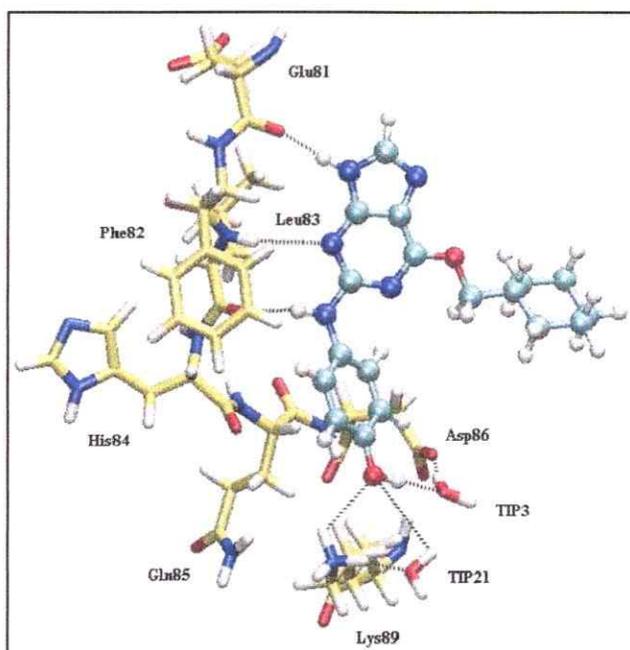


**Figura 3.5.** Fotografía de la estructura CDK2-Cp2. Los residuos seleccionados del sitio activo de la CDK2 son mostrados en la representación “licorice”, mientras que el Cp2 es presentado en la representación “esferas y barras”, con los átomos de carbono coloreados en amarillo (CDK2) y cian (inhibidor). Los enlaces de hidrógeno son esquematizados como líneas discontinuas negras.

Fragmento	Cp 2	Cp 25	Cp 3	Cp 9	Cp 7
Glu81 C=O---H-N9	1.96 (± 0.16)	1.88 (± 0.14)	1.92 (± 0.15)	1.90 (± 0.15)	1.91 (± 0.15)
Leu83 N-H---N3	2.49 (± 0.25)	2.62 (± 0.21)	2.46 (± 0.21)	2.60 (± 0.23)	2.72 (± 0.30)
Leu83 C=O---H-N2	1.94 (± 0.17)	1.94 (± 0.17)	1.90 (± 0.17)	1.97 (± 0.19)	2.01 (± 0.21)

Tabla 3.2. Distancias promedio de enlace de hidrógeno (Å) entre el anillo purina de cada ligando y la región bisagra del sitio activo de la CDK2. Los valores dentro del paréntesis corresponden a la desviación estándar.

resultó substancialmente más potente que el compuesto 2 original. Esto sugirió que existía un aceptor de enlace de hidrógeno en la región de la proteína a la cual tenía acceso el grupo 2-anilino. Un candidato obvio que se propuso fue el residuo Asp86. Los datos obtenidos de nuestras simulaciones de dinámica molecular (ver Figura 3.6) muestran que el grupo hidroxilo del compuesto 25 interactúa con la estructura carboxilato del residuo Asp86 a través de un enlace de hidrógeno mediado por una molécula de agua (Cp25 O-TIP3-Asp86 OD2 = 3.29 Å/2.62 Å (± 0.42 Å/± 0.10 Å)), lo cual está de acuerdo con los datos cristalográficos reportados por Davies *et al.*, no obstante las distancias de enlace de hidrógeno son diferentes de aquellas reportadas por estos autores.<sup>13</sup>



**Figura 3.6.** Fotografía de la estructura de CDK2-Cp25. La representación es realizada como en la Figura 3.5.

Hemos observado además dos enlaces de hidrógeno adicionales formados con el residuo Lys89. Uno de ellos es formado entre el átomo de oxígeno del grupo hidroxilo y el grupo  $\text{NH}_3^+$  del residuo Lys89 ( $\text{Cp25 O a Lys89 HZ} = 3.79 \text{ \AA} (\pm 0.66 \text{ \AA})$ ). El grupo hidroxilo también interactúa con el residuo Lys89 a través de un enlace de hidrógeno mediado por una molécula de agua ( $\text{Cp25 O-TIP21-Lys89 HZ} = 3.31 \text{ \AA}/2.94 \text{ \AA} (\pm 0.37 \text{ \AA}/\pm 0.51 \text{ \AA})$ ). Todas las interacciones juntas dan soporte a la energía de interacción obtenida para este compuesto a partir de los datos de la dinámica molecular.

### III.3.1.2 Compuesto 3 (Nu6102)

Este compuesto muestra un comportamiento diferente dentro del sitio activo de la enzima CDK2 debido a la presencia del grupo sulfonamida. Hardcastle *et al.* han sugerido que este compuesto interactúa con el aminoácido Asp86 a través del grupo  $\text{NH}_2$  de la sulfonamida el cual dona un enlace de hidrógeno al oxígeno de la cadena lateral de Asp86 ( $\text{Cp3 N26 a Asp86 OD2} = 2.9 \text{ \AA}$ ) y a través de un oxígeno de la sulfonamida que

acepta un enlace de hidrógeno del nitrógeno del enlace peptídico de Asp86 (Cp3 O24 a Asp86 NH = 3.1 Å).<sup>14</sup> Ellos han propuesto que la alta actividad de Cp3 se debe a los dos enlaces de hidrógeno formados con el residuo Asp86. Sin embargo, nosotros hemos encontrado que las interacciones formadas por este compuesto dentro del sitio activo de la CDK2 son marcadamente diferentes de las reportadas anteriormente.

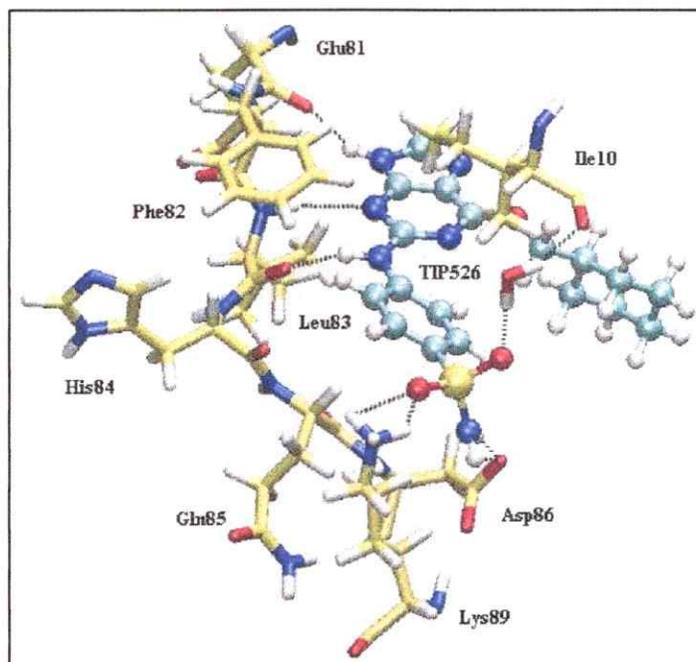
En la Figura 3.7 se muestra un esquema representativo de este complejo proteína-inhibidor. Este se obtuvo a partir de nuestras simulaciones de dinámica molecular (DM). Adicional al triplete de enlaces de hidrógeno, establecidos con la región bisagra de la CDK2, reportados en la Tabla 3.2, nosotros hemos encontrado otros enlaces de hidrógeno entre la CDK2 y Cp3. El grupo NH<sub>2</sub> de la sulfonamida dona un enlace de hidrógeno bifurcado a cada uno de los oxígenos de la cadena lateral del residuo Asp86 (Cp3 O<sub>2</sub>S-N a Asp86 OD1 = 3.78 Å (± 0.40 Å) y Cp3 O<sub>2</sub>S-N a Asp86 OD2 = 3.38 Å (± 0.45 Å)). Un átomo de oxígeno de la sulfonamida acepta un enlace de hidrógeno del grupo NH<sub>3</sub><sup>+</sup> de la cadena lateral del residuo Lys89 (Cp3 S=O27 a Lys89 HZ = 2.80 Å (± 0.15 Å)). De acuerdo con Davies *et al.*<sup>13</sup>, aunque el grupo anilino de Cp3 se compacta muy cerca con la superficie específica, el grupo sulfonamida no forma el enlace de hidrógeno propuesto con el residuo Lys89.

Nosotros hemos encontrado que la interacción ocurre y que el grupo sulfonamida está en una posición de óptima geometría para formar este enlace de hidrógeno. Finalmente, la última interacción de enlace de hidrógeno encontrada para este compuesto y tampoco aún reportada, ocurre entre el otro oxígeno del grupo sulfonamida y el grupo carbonilo del enlace peptídico del residuo Ile10. El enlace de hidrógeno está mediado por una molécula de agua perteneciente a la caja de solvente (Cp3 S=O28-TIP526-Ile10 C=O = 3.07 Å/3.39 Å (± 0.66 Å/± 0.76 Å)).

### III.3.1.3 Compuestos 9 y 7

Estos dos compuestos pertenecen a la serie no aromática estudiada por Hardcastle *et al.*<sup>14</sup> Estos también demostraron ser más potentes, como inhibidores de la CDK1 y la CDK2, que el compuesto original Nu2058,<sup>3</sup> y ambos forman el triplete de enlaces de hidrógeno dentro del sitio de unión del ATP en la CDK2 (Ver Tabla 3.2). El compuesto

**9 (Cp9)** es el derivado N<sup>2</sup>-isopropil mono-sustituido de la serie estudiada y forma una interacción débil a través de uno de los hidrógenos del grupo isopropilo, del tipo



**Figura 3.7** Fotografía de la estructura CDK2-Cp3. La representación es realizada como en la Figura 3.5.

C-H $\cdots$  $\pi$ , con la estructura peptídica entre los residuos Gln85 y Asp86. El complejo proteína-ligando se muestra en la Figura A.3.1 (ver en Anexo 1). El grupo isopropilo forma una interacción de carácter hidrofóbico con la cadena lateral del residuo Ile10 y una interacción de enlace de hidrógeno muy débil; del tipo C-H $\cdots$ O=C, con uno de los oxígenos de la cadena lateral del residuo Asp86 (**Cp9** isopropil-H a Asp86 OD2 = 3.78 Å ( $\pm$  0.49 Å)). Por su parte, el compuesto **7 (Cp7)** es el derivado N<sup>2</sup>-metil mono-sustituido de esta serie. Este compuesto forma una interacción hidrofóbica con la cadena lateral del residuo Ile10 (ver Figura A.3.2 en Anexo 1) y el típico triplete de enlaces de hidrógeno con la región bisagra de la CDK2 (Tabla 3.2). La ausencia de un grupo voluminoso, como el isopropilo en **Cp9**, causa una disminución en la fuerza de la interacción hidrofóbica con el residuo Ile10.

### III.3.2 Análisis de las energías de interacción

En la Tabla 3.3 se reportan las energías de interacción promediadas para la interacción de cada uno de los inhibidores con la CDK2, junto con sus componentes de energía electrostática y de vdW. La actividad biológica se expresa como el logaritmo de  $IC_{50}$ .<sup>14</sup> La potencia de cada ligando está obviamente relacionada con su energía libre de unión siempre y cuando se asuma que, dentro de una familia de compuestos, las energías de solvatación/desolvatación, la energía de deformación de la enzima y los cambios entrópicos son proporcionales a la magnitud de las interacciones establecidas entre el inhibidor y la proteína.

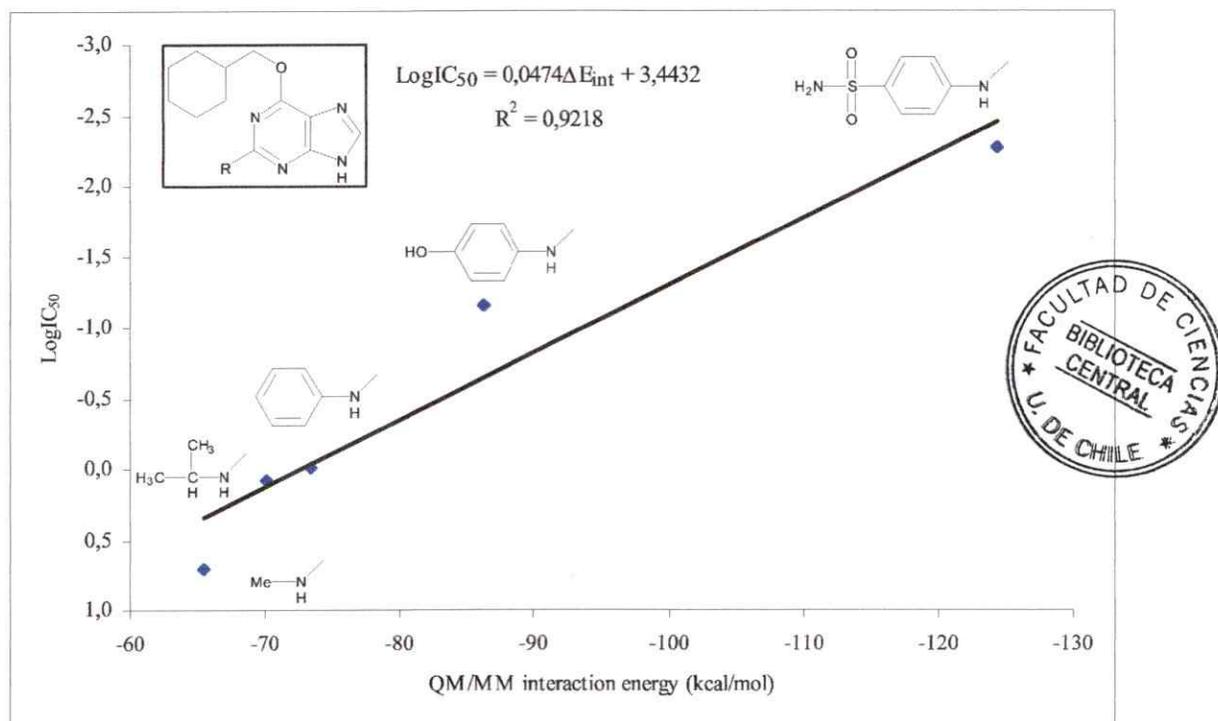
	Compuesto 2	Compuesto 25	Compuesto 3	Compuesto 9	Compuesto 7
Energía de Interacción	-73.4 (± 2.9)	-86.2 (± 3.5)	-124.5 (± 8.2)	-70.2 (± 4.0)	-65.5 (± 3.1)
Energía Electrostática	-17.5 (± 2.4)	-30.2 (± 3.4)	-65.8 (± 9.0)	-23.4 (± 3.5)	-22.8 (± 3.3)
Energía vdW	-55.9 (± 2.2)	-56.1 (± 2.5)	-58.6 (± 3.0)	-46.8 (± 2.4)	-42.7 (± 2.4)
Log $IC_{50}$	-0.013	-1.161	-2.268	0.079	0.699

**Tabla 3.3.** Promedios de las energías de interacción QM/MM y sus componentes (kcal/mol) para los inhibidores de CDK2 estudiados. Las actividades biológicas son reportadas como el logaritmo común de  $IC_{50}$  (micromolar)<sup>15</sup>.

De esta forma, se puede expresar la actividad como una función de la energía de interacción:

$$\text{LogIC}_{50} = \alpha \Delta E_{\text{QM/MM}} + \beta \quad (3.4)$$

En la Figura 3.8 se muestra el resultado del ajuste de cuadrados mínimos de la ecuación (3.4) utilizando los datos consignados en la Tabla 3.3. El alto valor obtenido para el coeficiente de correlación ( $R = 0.96$ ) valida nuestro procedimiento de simulación. Además, esto indica claramente que las energías de interacción pueden ser utilizadas con confianza como índices de la actividad biológica de un inhibidor dado, al menos cuando se comparan para una serie de compuestos que pertenezcan a una misma familia química. El interés que suscita esta clase de correlaciones es obvio, considerando la gran diferencia en esfuerzo computacional que se necesita para evaluar las energías libres de unión o las energías de interacción. La utilización de los métodos QM/MM tiene la ventaja adicional de no requerir la reparametrización extensiva para cada nuevo ligando que se considere dentro del estudio.



**Figura 3.8.** Gráfica de LogIC<sub>50</sub> versus Energía de Interacción QM/MM (kcal/mol) para los cinco derivados de purina O<sup>6</sup>-ciclo-hexil-metil N<sup>2</sup>-sustituidos estudiados y la ecuación obtenida con el ajuste de cuadrados mínimos.

Se puede obtener una conclusión práctica a partir del estudio de la calidad del ajuste de cuadrados mínimos como una función del tamaño de muestra utilizado en la evaluación de las energías de interacción. Si en lugar de utilizar el promedio de los datos de la simulación completa de 250 ps se utiliza el promedio de los datos de solo los últimos 100 ps, el valor del coeficiente de correlación se reduce sólo a 0.95. Esto indica que el muestreo, y por lo tanto el tiempo computacional, se pueden reducir sin una pérdida significativa de la calidad del modelo.

Es también interesante relacionar las diferentes energías de interacción y sus componentes con las características estructurales de cada complejo proteína-inhibidor. Se puede observar a partir de los datos incluidos en la Tabla 3.3 que la magnitud de la contribución electrostática es similar para los compuesto **9** y **7**, como resultado de la similitud en las interacciones de enlace de hidrógeno presentadas por estos ligandos con la región bisagra de la CDK2. La energía vdW es diferente para estos dos compuestos debido a los patrones de sustitución distintos que se presentan en la posición N<sup>2</sup> del anillo purina. La principal consecuencia que se plantea de esta diferencia es que **Cp9** lleva un grupo isopropilo más polarizable en la posición N<sup>2</sup> que **Cp7**, el cual tiene un grupo metilo en esa posición, promoviendo de este modo una interacción más favorable con el residuo Ile10. Este ejemplo muestra claramente que la actividad puede ser mejorada con la optimización de las interacciones de van der Waals. Esta mejora en las interacciones de vdW hace posible un ligero aumento en las interacciones electrostáticas entre el anillo purina y la región bisagra de la CDK2 para el **Cp9**. Con los compuestos **2** y **25** ocurre lo contrario. Estos compuestos presentan contribuciones similares de energía de vdW a la energía de interacción, y sus valores de energía electrostática son diferentes (Tabla 3.3). Las diferencias en la energía electrostática para estos compuestos puede ser atribuida a una interacción importante entre el grupo sustituyente en la posición 4' en el anillo fenilo del grupo anilino y algunos aminoácidos en la superficie de especificidad. Para **Cp2** la alta contribución de la componente de vdW (en valor absoluto) está de acuerdo con el patrón de interacción esperado con la superficie de especificidad en el sitio activo de la CDK2, cuando este se compara con los valores correspondientes para los compuestos **9** y **7**. Esto significa que un anillo anilino en la

posición C-2 del anillo purina es más favorable que un sustituyente no aromático como el metil o isopropil en Cp7 y Cp9, respectivamente. La componente de energía electrostática para Cp2 es más pequeña que los valores correspondientes presentados para Cp7 y Cp9. Esto puede atribuirse a una interacción menos favorable entre el anillo purina de este compuesto y la región bisagra de la CDK2. Sin embargo, esta baja contribución es compensada por un aumento en la componente de energía de vdW.

Cp25, por otro lado, presenta un aumento en la componente de la energía electrostática debido a las interacciones de enlace de hidrógeno formadas por el grupo -OH en la posición 4' del anillo anilino con los residuos de la superficie de especificidad de la enzima. De esta forma, es posible facilitar un empaquetamiento ligeramente más efectivo entre el anillo fenilo y los aminoácidos hidrofóbicos en el sitio activo de la CDK2, lo cual puede verse reflejado en la diferencia en los valores de la energía vdW para estos dos compuestos. Es interesante mencionar que Cp3 es el único compuesto que tiene una componente de energía electrostática mayor que la componente de vdW. Este resultado se puede explicar por las múltiples interacciones de enlace de hidrógeno que este compuesto puede establecer con los residuos Asp86, Lys89 e Ile10, los cuales están presentes en la superficie de especificidad de la CDK2. Al comparar los valores de la energía de vdW de este compuesto con los evaluados para Cp2 y Cp25, se puede concluir que la componente de energía de vdW tiene un peso similar en estos compuestos, además esta ligera mejora en la componente de energía se debe a la interacción formada con los aminoácidos de la superficie de especificidad y, consecuentemente, a un empaquetamiento más efectivo con la superficie hidrofóbica dentro del sitio activo de la CDK2. El resultado final para el compuesto 3 es una interacción óptima y por lo tanto una actividad inhibitoria mayor.

Finalmente hemos realizado un análisis de regresión múltiple de dos variables, con el propósito de evaluar si nuestro esquema de partición de la energía de interacción es un modelo confiable para representar la interacción de cada ligando dentro del sitio activo de la CDK2. Las dos variables independientes son las componentes de energía electrostática y de vdW, y la variable dependiente es el logaritmo de la concentración

inhibitoria  $IC_{50}$  que se obtiene a partir de los experimentos reportados en la literatura. La ecuación que resulta de la regresión es:

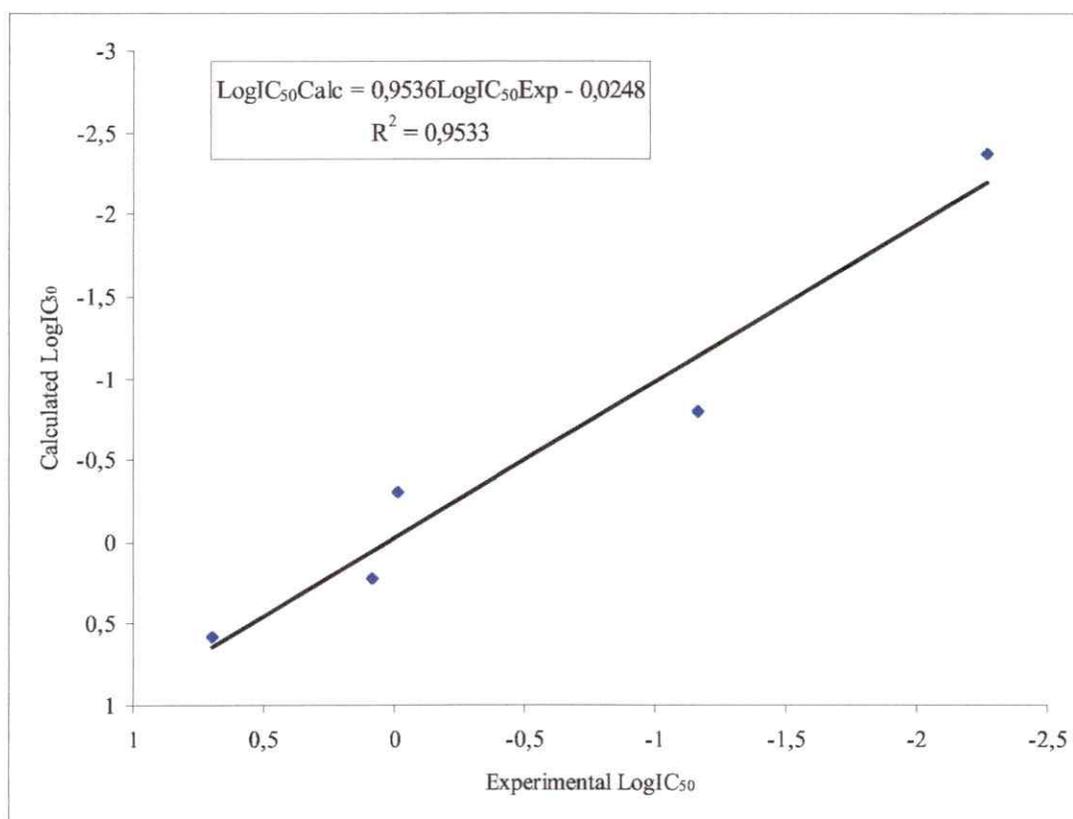
$$\text{Log}IC_{50} = 4.97298 + 0.03804 * \Delta E_{\text{elec}} + 0.08251 * \Delta E_{\text{vdW}} \quad (3.5)$$

donde  $\Delta E_{\text{elec}}$  y  $\Delta E_{\text{vdW}}$  son las componentes de energía electrostática y de vdW, respectivamente. Esta ecuación se puede usar para obtener los valores teóricos de  $\text{Log}IC_{50}$ , los cuales se muestran en la Tabla A.3.1 (Ver Anexo 1). Hemos graficado esos valores teóricos de  $\text{Log}IC_{50}$  contra los valores experimentales de  $\text{Log}IC_{50}$  reportados en la literatura para los compuestos estudiados. La comparación da un coeficiente de correlación  $R=0.976$  (ver Figura 3.9).

Interesantemente, la Ecuación (3.5) muestra valores bastante diferentes para las pendientes asociadas con las energías electrostáticas y de van der Waals, lo cual sugiere que una mejora de esta última contribución podría ser más efectiva, en términos de inhibición, que una mejora energética similar en la componente electrostática. Obviamente, otro asunto es, que obtener diferencias en las energías de interacción de vdW puede ser más difícil que para la parte electrostática, como lo sugieren los rangos cubiertos por esas dos componentes de la energía de interacción en los cinco casos analizados aquí. Como se muestra en la Tabla 3.3, la diferencia máxima en la energía de interacción de vdW entre los inhibidores estudiados en esta parte de la tesis es 16.9 kcal/mol, mientras que en el caso de la energía de interacción electrostática, esta diferencia máxima llega a 48.3 kcal/mol.

Si graficamos  $\text{Log}IC_{50}$  vs  $\Delta E_{\text{elec}}$  o  $\text{Log}IC_{50}$  vs  $\Delta E_{\text{vdW}}$ , los coeficientes de correlación obtenidos son 0.89 y 0.82, respectivamente. Esto sugiere que hacer correlaciones utilizando estos descriptores por separado no es suficiente para describir adecuadamente el comportamiento de los inhibidores dentro del sitio activo de la CDK2, debido al papel tan importante que tiene cada propiedad en la interacción global entre el inhibidor y los aminoácidos del sitio activo. En vez de utilizar esos descriptores por separado, nosotros hemos utilizado el método de regresión múltiple para predecir los valores de  $IC_{50}$  y

LogIC<sub>50</sub> para el conjunto de inhibidores de CDK2 estudiados (ver Tabla A.3.1). Los valores obtenidos a partir de la Ecuación (3.5) tienen una buena concordancia con los valores experimentales y, en principio, el modelo podría ser utilizado de una manera cuantitativa para predecir el valor de IC<sub>50</sub> para un nuevo compuesto relacionado con la estructura del núcleo químico estudiado.



**Figura 3.9.** Gráfica de los valores calculados de LogIC<sub>50</sub> (obtenidos por medio de la ecuación (3.5)) versus los valores experimentales de LogIC<sub>50</sub> para los cinco compuestos estudiados y la correspondiente ecuación obtenida por el ajuste de mínimos cuadrados.

### III.4 Conclusiones

En este capítulo hemos demostrado que los métodos híbridos QM/MM ofrecen una aproximación útil para calcular las interacciones establecidas por una serie de compuestos que pertenecen a la familia de las purinas O<sup>6</sup>-ciclo-hexil-metoxi-N<sup>2</sup>-

sustituídas con el sitio activo de la CDK2. Este método nos permite describir, con algún detalle, la naturaleza de las interacciones específicas involucradas en la unión proteína-ligando. Este hecho hace que el método pueda ser extendido a otros grupos de moléculas inhibitoras de la CDK2, y de hecho, a otros sistemas biológicos de interés en la química actual. Dentro de la presente aproximación es posible, además, obtener un cuadro representativo y aceptable del sistema real, en el cual la evaluación de las diferentes contribuciones del ambiente enzimático y del solvente a la unión proteína-ligando en el sitio activo de la proteína, se hacen una tarea simple de realizar. El considerar la flexibilidad de casi toda la proteína y el ligando en los cálculos realizados con los métodos híbridos QM/MM nos permite conseguir un muestreo más amplio del espacio conformacional, y de esta forma se pueden incorporar al análisis de las simulaciones de dinámica molecular un enorme conjunto de estructuras posibles a una temperatura relevante.

Con respecto a las interacciones específicas formadas dentro de sitio activo de la CDK2 por cada uno de los compuestos estudiados, podemos concluir que el aumento en potencia observado para Nu6102 (Cp3) resulta de las interacciones de enlace de hidrógeno formadas entre el grupo sulfonamida y los residuos Asp86, Lys89 e Ile10. Hay una ganancia adicional en potencia que depende del empaquetamiento favorable entre el anillo aromático anilino y la superficie hidrofóbica presente en el sitio activo de la CDK2. Es digno de mencionar que esta es la primera vez que se reportan las interacciones de un compuesto de esta familia con el residuo Lys89.

La validación estadística del modelo propuesto fue realizada por medio de un análisis de regresión múltiple. Este análisis permite concluir que hay un alto grado de confiabilidad en los datos reportados para las componentes de la energía de interacción de cada inhibidor y que han sido calculadas con los métodos híbridos QM/MM descritos anteriormente.

Tomando en cuenta los valores de las dos variables propuestas, y su correlación con los valores de actividad biológica reportados en la literatura, parece posible la predicción del valor de  $IC_{50}$  para un nuevo que compuesto que pertenezca a la familia de compuestos estudiados en esta parte de la tesis. La evaluación del promedio de las energías de

interacción nos ha permitido demostrar que existe una fuerte correlación de ésta con la actividad biológica para esta serie de inhibidores de CDK2. De hecho, la energía de interacción QM/MM puede ser utilizada como un índice de la actividad biológica. La utilización de las energías de interacción QM/MM tiene una doble ventaja: i) se evita la reparametrización extensiva de los nuevos ligandos que se sometan al estudio y ii) el esfuerzo computacional es considerablemente menor que el necesario para la evaluación de las energías libres de unión. Finalmente, la partición de la energía muestra que en todos los compuestos, excepto para Cp3, la mayor contribución a la energía de interacción total la realiza la componente de energía de van der Waals. En algunos compuestos, los enlaces de hidrógeno facilitan las interacciones con la superficie hidrofóbica de la enzima.

A partir de los resultado obtenidos en esta parte de la tesis, se puede afirmar que la región de la CDK2 llamada "superficie específica" es la responsable de la interacción más favorable entre la CDK2 y el inhibidor Nu6102 y que es la enzima la que debe ser explorada con más detalle para conseguir inhibidores con más selectividad y potencia. Finalmente es importante resaltar que en los cálculos realizados con este método se ha utilizado el modelo de la enzima CDK2 unida a su unidad activadora, Ciclina A, y fosforilada en el aminoácido Thr160; lo cual permite trabajar con un sistema enzimático en su estado activo y de esta forma asegurar una evaluación más acertada de las propiedades que se quieran medir a partir de los cálculos computacionales.

### III.5. Bibliografía

- [1] Kuno, M.; Hannongbua, S.; Morokuma, K. Theoretical investigation on nevirapine and HIV-1 reverse transcriptase binding site interaction, based on ONIOM method *Chemical Physics Letters* **2003**, *380*, 456-463.
- [2] Martí, S.; Moliner, V.; Tuñón, I.; Williams, I. H. Computing kinetic isotope effects for chorismate mutase with high accuracy. A new DFT/MM strategy. *J. Phys. Chem. B* **2005**, *109*, 3707-3710.
- [3] Arris, C. E.; Boyle, F. T.; Calvert, A. H.; Curtin, N. J.; Endicott, J. A.; Garman, E. F.; Gibson, A. E.; Golding, B. T.; Grant, S.; Griffin, R. J.; Jewsbury, P.; Johnson, L. N.; Lawrie, A. M.; Newell, D. R.; Noble, M. E. M.; Sausville, E. A.; Schultz, R.; Yu, W. Identification of novel purine and pyrimidine cyclin-dependent kinase inhibitors with distinct molecular interactions and tumour cell growth inhibition profiles. *J. Med. Chem.* **2000**, *43*, 2797-2804.
- [4] Lawrie, A. M.; Noble, M. E. M.; Tunnah, P.; Brown, N. R.; Johnson, L. N. and Endicott, J. A. Protein kinase inhibition by staurosporine revealed in details of the molecular interaction with CDK2. *Nat. Struct. Biol.* **1997**, *4*, 796-801.
- [5] Omura, S.; Sasaki, Y.; Iwai, Y. and Takeshima, H. Staurosporine, a potentially important gift from a microorganism. *J. Antibiot.* **1995**, *48*, 535-548.
- [6] Hoessel, R.; Leclerc, S.; Endicott, J. A.; Noble, M. E. M.; Lawrie, A. M.; Tunnah, P.; Leost, M.; Damiens, E.; Marie, D.; Marko, D.; Niederberger, E.; Tang, W.; Eisenbrand, G. and Meijer, L. Indirubin, the active constituent of a Chinese antileukaemia medicine, inhibits cyclindependent kinases. *Nat. Cell Biol.* **1999**, *1*, 60-67.

- [7] Meijer, L. Chemical inhibitors of cyclin-dependent kinases. *Trends Cell Biol.* **1996**, *6*, 393-397.
- [8] Gibson, A. E.; Arris, C. E.; Bentley, J.; Boyle, F. T.; Curtin, N. J.; Davies, T. G.; Endicott, J. A.; Golding, B. T.; Grant, S.; Griffin, R. J.; Jewsbury, P.; Johnson, L. N.; Mesguiche, V.; Newell, D. R.; Noble, M.E.M.; Tucker, J. A.; Whitfield, H. J. Probing the ATP Ribose-binding Domain of Cyclin-Dependent Kinases 1 and 2 with O<sup>6</sup>-Substituted Guanine Derivatives. *J. Med. Chem.* **2002**, *45*, 3381-3393.
- [9] Hanks, S. K.; Hunter, T. The eukaryotic protein kinase super family: kinase (catalytic) domain structure and classification. *FASEB J.* **1995**, *9*, 576-596.
- [10] Gray, N.; Detivaud, L.; Doering, C.; Meijer, L. ATP-site directed inhibitors of cyclin dependent kinases. *Curr. Med. Chem.* **1999**, *6*, 850-875.
- [11] Schulze-Gahmen, U.; Brandsen, J.; Jones, H. D.; Morgan, D. O.; Meijer, L.; Vesely, J.; Kim, S. H. Multiple modes of ligand recognition: crystal structures of cyclin dependent protein kinase 2 in complex with ATP and two inhibitors, olomoucine and isopentenyladenine. *Proteins: Struct, Funct, Genet* **1995**, *22*, 378-391.
- [12] De Azevedo, W. F.; Leclerc, S.; Meijer, L.; Havlicek, L.; Strnad, M.; Kim, S. H. Inhibition of cyclin- dependent kinases by purine analogues. Crystal structure of human cdk2 complexed with roscovitine. *Eur. J. Biochem.* **1997**, *243*, 518-526.
- [13] Davies, T. G.; Bentley, J.; Arris, C. E.; Boyle, F. T.; Curtin, N. J.; Endicott, J. A.; Gibson, A. E.; Golding, B. T.; Griffin, R. J.; Hardcastle, I. R.; Jewsbury, P.; Johnson, L. N.; Mesguiche, V.; Newell, D. R.; Noble, M. E. M.; Tucker, J. A.; Wang, L.; Whitfield, H. J. Structure-based design of a potent purine-based cyclin-dependent kinase inhibitor. *Nat. Struct. Biol.* **2002**, *9*, 745-749.

- [14] Hardcastle, I. R.; Arris, C. E.; Bentley, J.; Boyle, F. T.; Chen, Y.; Curtin, N. J.; Endicott, J. A.; Gibson, A. E.; Golding, B. T.; Griffin, R. J.; Jewsbury, P.; Menyerol, J.; Mesguiche, V.; Newell, D. R.; Noble, M. E. M.; Pratt, D. J.; Wang, L. Z.; Whitfield, H. J. N<sup>2</sup>-Substituted O<sup>6</sup>-Cyclohexylmethylguanine Derivatives: Potent Inhibitors of Cyclin-Dependent Kinases 1 and 2. *J. Med. Chem.* **2004**, *47*, 3710-3722.
- [15] Cavalli, A.; Dezi, C.; Folkers, G.; Scapozza, L.; Recanatini, M. Three-dimensional model of the cyclin-dependent kinase 1 (CDK1): Ab initio active site parameters for molecular dynamics studies of CDKs. *Proteins:Structure Function and Genetics* 2001, *45* (4), 478-485.
- [16] Sims, P.A.; Wong, C. F.; McCammon, J. A. A Computational Model of Binding Thermodynamics: The Design of Cyclin-dependent Kinase 2 inhibitors. *J. Med. Chem.* **2003**, *46*, 3314-3325.
- [17] Sims, P.A.; Wong, C. F.; McCammon, J. A. Charge Optimization of the Interface between Protein Kinases and Their Ligands. *J. Comput. Chem.* **2004**, *25*(11), 1416-1429.
- [18] Gould, C.; Wong, C. F. Designing specific protein kinase inhibitors: insights from computer simulations and comparative sequence/structure analysis. *Pharmacol. Ther.* **2002**, *93*, 169-178.
- [19] Wong, C. F.; Hünenberger, P. H.; Akamine, P.; Narayana, N.; Diller, T.; McCammon, J. A.; Taylor, S.; Xuong, N. H. Computational Analysis of PKA-Balanol Interactions. *J. Med. Chem.* **2001**, *44*, 1530-1539.
- [20] Teague, S. J. Implications of protein flexibility for drug discovery. *Nature Rev. Drug Discov.* **2003**, *2*, 527-541.

- [21] Lin, J. H.; Perryman, A. L.; Schames, J. R.; McCammon, J. A. Computational drug design accommodating receptor flexibility: The relaxed complex scheme. *J. Am. Chem. Soc.* **2002**, *124*, 5632-5633.
- [22] Carlson, H. A. Protein flexibility and drug design: how to hit a moving target. *Curr. Opin. Chem. Biol.* **2002**, *6*, 447-452.
- [23] Carlson, H. A.; McCammon, J. A. Accommodating protein flexibility in computational drug design. *Mol. Pharmacol.* **2000**, *57*, 213-218.
- [24] Huse, M.; Kuriyan, J. The conformational plasticity of protein kinases. *Cell* **2002**, *109*, 275-282.
- [25] Gohlke, H.; Klebe, G. Statistical potentials and scoring functions applied to protein-ligand binding. *Curr. Opin. Struct. Biol.* **2001**, *11*(2), 231-235.
- [26] Aqvist, J.; Luzhkov, V. B.; Brandsdal, B. O. Ligand binding affinities from MD simulations. *Acc. Chem. Res.* **2002**, *35*, 358-365.
- [27] Simonson, T.; Archontis, G.; Karplus, M. Free energy simulations come of age: Protein-ligand recognition. *Acc. Chem. Res.* **2002**, *35*, 430-437.
- [28] Soriano, A.; Silla, E.; Tuñon, I.; Ruiz-López, M. F. Dynamic and Electrostatic Effects in Enzymatic Processes. An Analysis of the Nucleophilic Substitution Reaction in Haloalkane Dehalogenase. *J. Am. Chem. Soc.* **2005**, *127*, 1946-1957.
- [29] Roca, M.; Andrés, J.; Moliner, V.; Tuñon, I.; Bertrán, J. On the Nature of the Transition State in Catechol O-Methyltransferase. A Complementary Study Based on Molecular Dynamics and Potential Energy Surface Explorations. *J. Am. Chem. Soc.* **2005**, *127*, 10648-10655.

- [30] Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J. Preorganization and Reorganization as Related Factors in Enzyme Catalysis: The Chorismate Mutase Case. *Chem. Eur. J.* **2003**, *9*, 984-991.
- [31] Gao, J.; Xia, X. A Priori Evaluation of Aqueous Polarization Effects Through Monte Carlo QM-MM Simulations. *Science* **1992**, *258*, 631-635.
- [32] Gao, J. Methods and Applications of Combined Quantum Mechanical and Molecular Mechanical Potentials. *Rev. Comput. Chem.* **1995**, *7*, 119-185.
- [33] Field, M. J.; Bash, P. A.; Karplus, M. A. A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations. *J. Comput. Chem.* **1990**, *11*, 700-733.
- [34] Parr R. G. and Yang W. *Density Functional Theory of Atoms and Molecules* (Oxford Univ. Press, Oxford) **1989**.
- [35] Hensen, C.; Hermann, J.; Nam, K.; Ma, S.; Gao, J.; Höltje, H. D. A Combined QM/MM Approach to Protein-Ligand Interactions: Polarization Effects of the HIV-1 Protease on Selected High Affinity Inhibitors. *J. Med. Chem.* **2004**, *47*, 6673-6680.
- [36] Gräter, F.; Schwarzl, S. M.; Dejaegere, A.; Fischer, S.; Smith, J. C. Protein/Ligand Binding Free Energies Calculated with Quantum Mechanics/Molecular Mechanics. *J. Phys. Chem. B* **2005**, *109*, 10474-10483.
- [37] Field, M. J.; Albe, M.; Bret, C.; Proust-De Martin, F.; Thomas, A. The Dynamo library for molecular simulations using hybrid quantum mechanical and molecular mechanical potentials. *J. Comput. Chem.* **2000**, *21*, 1088-1100.

- [38] Schaftenaar, G.; Noordik, J. H. Molden: a pre- and post-processing program for molecular and electronic structures. *J. Comput.-Aided Mol. Design.* **2000**, *14*, 123-134.
- [39] Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. Development and use of Quantum Mechanical Molecular Models. 76. AM1: a new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* **1985**, *107*, 3902-3909.
- [40] Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225-11236.
- [41] Jorgensen, W. L.; Chandrasekar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926-935.
- [42] Gao, J.; Truhlar, D. G. Quantum Mechanical Methods for Enzyme Kinetics. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467-505.
- [43] Hestenes, M. R.; Stiefel, E. Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bureau of Standards* **1952**, *49*, 409-436.
- [44] Gray, N. S.; Wodicka, L.; Thunnissen, A. M. W. H.; Norman, T. C.; Kwon, S.; Espinoza, F. N.; Morgan, D. O.; Barnes, G.; LeClerc, S.; Meijer, L.; Kim, S. H.; Lockhart, D. J.; Schultz, P. G. Exploiting Chemical Libraries, Structure, and Genomics in the Search for Kinase Inhibitors. *Science* **1998**, *281*, 533-538.
- [45] Davies, T. G.; Tunnah, T.; Meijer, L.; Marko, D.; Eisenbrand, G.; Endicott, J. A.; Noble, M. E. M. Inhibitor Binding to Active and Inactive CDK2: The Crystal Structure of CDK2-Cyclin A/Indirubin-5-Sulphonate. *Structure* **2001**, *9*, 389-397.

## CAPÍTULO IV

### Estudio teórico en inhibidores de CDK2 utilizando una blandura global obtenida a partir de la densidad de estados

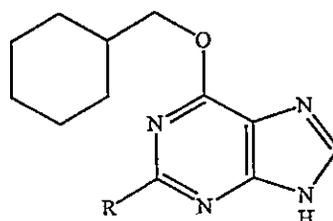
#### IV.1 Introducción

En este último capítulo complementaremos los estudios realizados en el Capítulo II, incorporando los índices globales y locales de reactividad en el estudio de la interacción de los inhibidores de CDK2 con el sitio activo de la enzima. Primero presentaremos los resultados del estudio en el que utilizamos la aproximación de orbitales moleculares de frontera (Frontier Molecular Orbitals (FMO))<sup>1, 2, 3</sup> para determinar los índices globales y locales de reactividad en un grupo de 19 moléculas pertenecientes a la familia de las O<sup>6</sup>-ciclo-hexil-metoxi purinas N<sup>2</sup>-sustituídas. También se determinaron índices regionales de reactividad en los dos fragmentos que interactúan (proteína-ligando), que permite definir un índice de similaridad entre ellos. Finalmente, se obtuvieron los índices globales de reactividad para el mismo grupo de moléculas, pero utilizando la aproximación de la densidad de estados (Density of States (DOS)).<sup>4</sup> Esta última aproximación nos dio mejores resultados que la primera y nos permitió obtener coeficientes de correlación ( $R^2$ ) razonablemente significativos, con respecto a la actividad biológica medida como IC<sub>50</sub>. El análisis de la DOS fue complementado con un estudio del potencial electrostático molecular (Molecular Electrostatic Potential (MEP)),<sup>5, 6</sup> el cual nos permitió predecir los sitios nucleofílicos y electrofílicos en el sitio activo de la proteína y los ligandos, y por tanto predecir la orientación que toman estos en el sitio activo de la enzima CDK2. Actualmente están disponibles nuevas estrategias para encontrar inhibidores o moléculas bio-activas, que hacen uso de relaciones estructura-actividad obtenidas a partir de cálculos computacionales.<sup>7, 8, 9</sup> En esta parte de la tesis proponemos que el análisis con los índices de reactividad global y

local puede ser utilizado para estimar la afinidad de una serie de inhibidores de CDK2, que pertenecen a la familia de las O<sup>6</sup>-ciclo-hexil-metoxi purinas N<sup>2</sup>-sustituidas, por la región bisagra del sitio activo de la misma enzima (residuos 81-83) (Ver Esquema 4.1).<sup>10</sup> Los descriptores globales de reactividad definidos en el contexto de la Teoría de Funcionales de la Densidad (DFT), incluyendo el potencial químico electrónico ( $\mu$ ), la blandura global (S) y la electrofilia global ( $\omega$ ), han demostrado ser herramientas útiles para entender los aspectos mecanísticos<sup>11</sup> y termodinámicos de los sistemas en reacción<sup>12</sup> en sus estados fundamentales. Por ejemplo, en un artículo pionero, Maynard *et al.*<sup>13</sup> utilizaron la relación electronegatividad-dureza,  $\chi^2/\eta$ , para explicar la capacidad de una serie de electrófilos para promover una reacción blanda (covalente) con la proteína p7 de la nucleocápside (NCp7) del virus de la inmunodeficiencia humana tipo 1 (VIH-1). Ellos encontraron una correlación estadísticamente significativa entre las velocidades de reacción y la habilidad de esos agentes para reaccionar como electrófilos blandos. En la segunda parte de este estudio reportamos una correlación significativa entre la actividad biológica de los inhibidores de CDK2 y la blandura global definida en términos de la Densidad de Estados (DOS), lo cual permite describir cualitativamente la afinidad de una serie de inhibidores de CDK2 por la región bisagra, localizada en el sitio activo de la CDK2. Este método fue propuesto por Santos *et al.*,<sup>14</sup> y fue aplicado exitosamente en la evaluación de la basicidad de zeolitas alcalinas intercambiadas. Por otra parte, los aspectos locales de reactividad (selectividad) se pueden evaluar exitosamente por medio de descriptores locales de reactividad como la blandura local  $s(r)$ ,<sup>15</sup> el índice de electrofilia local  $\omega(r)$ <sup>16</sup> y la función de Fukui  $f(r)$ .<sup>17</sup> Estos índices locales pueden, por integración, ser convertidos en índices regionales condensados en átomos o fragmentos moleculares. Todos ellos brindan información útil relacionada con los productos de control cinético esperados en un proceso químico.<sup>18, 19</sup> Otra cantidad pertinente para evaluar la reactividad local son los valores críticos (mínimos en el presente caso) del potencial electrostático.

Esta última aproximación será utilizada aquí para evaluar las interacciones de largo alcance que guían el acercamiento de los ligandos a la región bisagra de la CDK2, antes

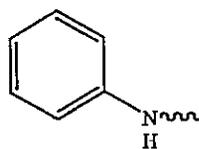
de la interacción intermolecular específica. Hemos utilizado el potencial electrostático,<sup>20</sup> trazado sobre una isosuperficie de potencial, para localizar los sitios de



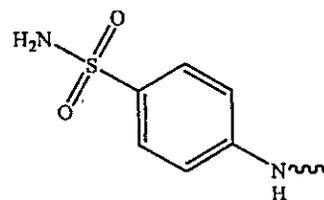
Núcleo Básico

NH<sub>2</sub><sup>-</sup>

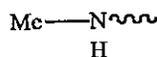
1



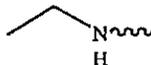
2



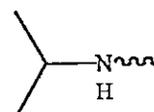
3



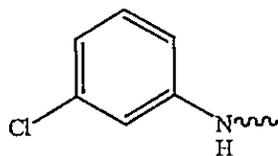
7



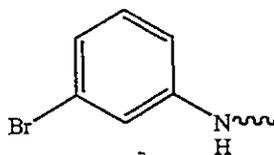
8



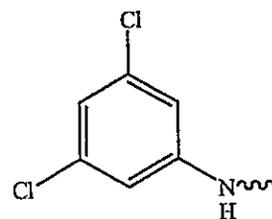
9



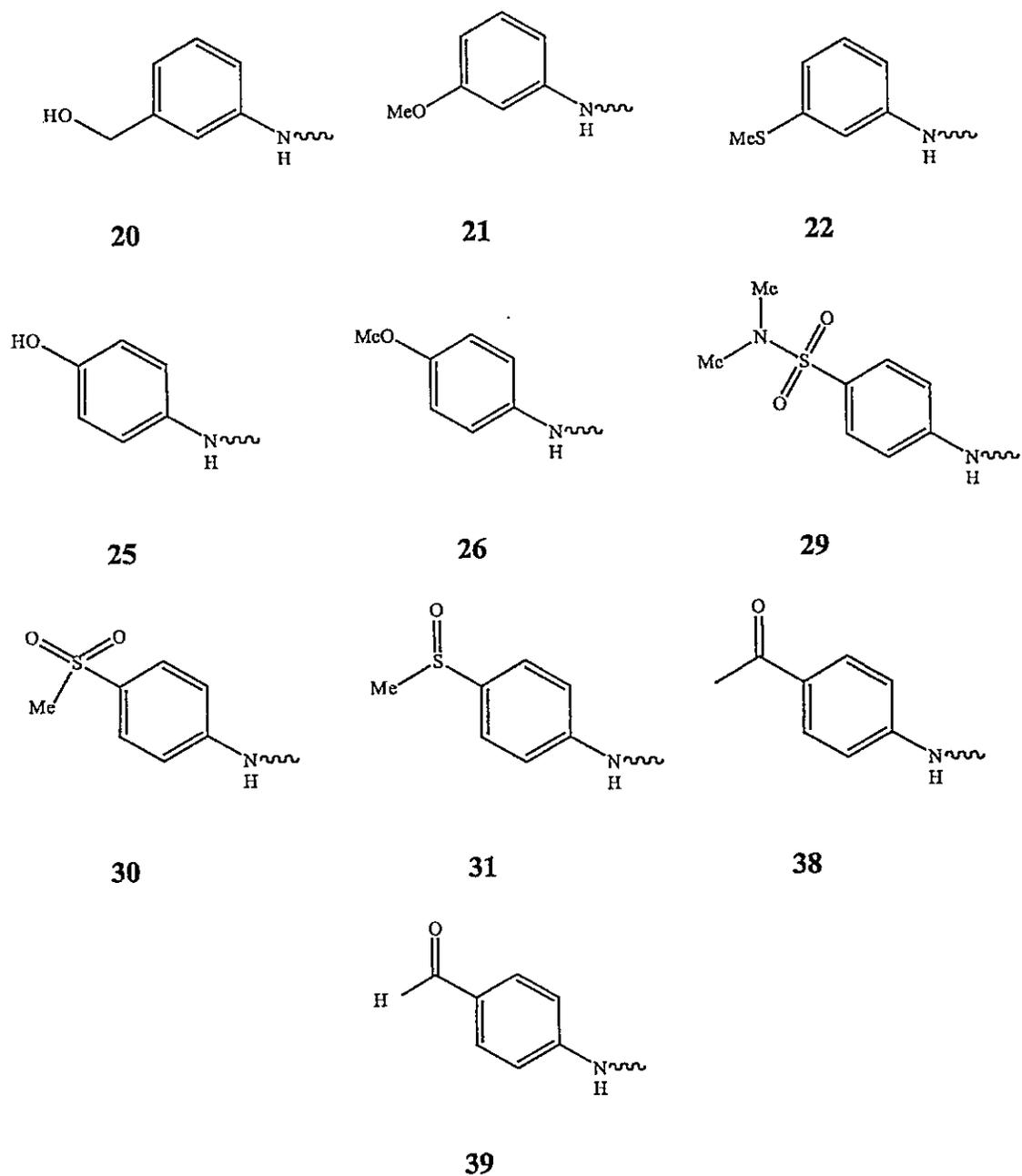
15



17



18



Esquema 4.1. Estructuras moleculares para el conjunto de inhibidores de CDK2 estudiados.

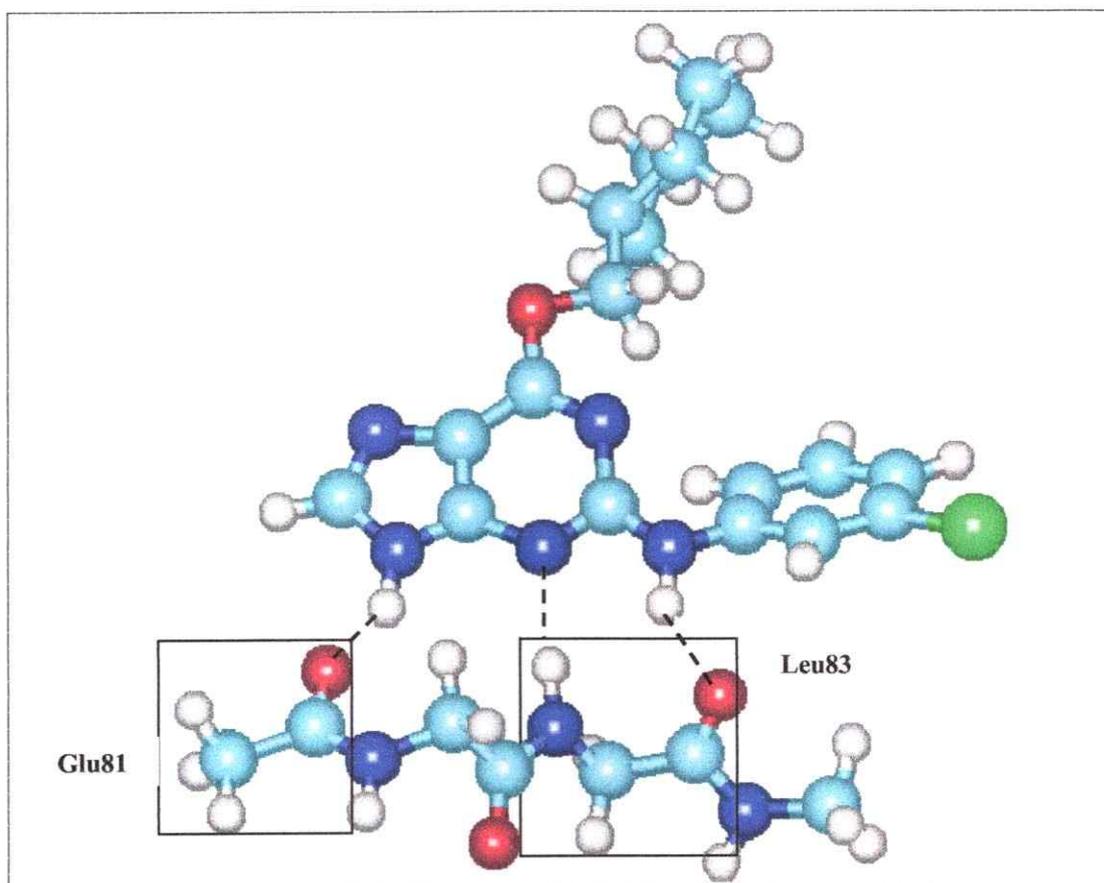
reconocimiento más importantes en la interacción proteína-ligando entre los inhibidores de CDK2 y la región bisagra de la misma. Los resultados reproducen cualitativamente el patrón de reconocimiento típico, definido en el modelo clásico de "llave y cerradura",

entre los sitios nucleofílicos y electrofílicos localizados en ambos sistemas (ligando y proteína). Todos los índices teóricos de reactividad están definidos en el Apéndice 2.

## **IV.2. Ecuaciones modelo y detalles computacionales.**

### **IV.2.1. Índices de reactividad obtenidos a partir de la aproximación de orbitales moleculares de frontera.**

Los modelos utilizados para realizar esta primera parte del estudio de índices globales de reactividad se obtuvieron a partir de las estructuras cristalográficas de la CDK2 en interacción con los inhibidores NU2058, NU6094, NU6086 y NU6102, reportadas previamente por Davies *et al.*<sup>21</sup> (Códigos en PDB (Protein Data Bank) 1H1P, 1H1Q, 1H1R y 1H1S, respectivamente). En la Figura 4.1 se muestra el modelo simplificado estudiado de la región bisagra del sitio activo de la CDK2 que consta del esqueleto de los aminoácidos (a.a.) Leu83 y Phe82, y además del grupo carbonilo y el carbono-alfa del a.a. Glu81 y el grupo amido y el carbono-alfa del a.a. His84. También se incluyen los diferentes inhibidores en interacción con este modelo de la región bisagra. Se asumió que todos los aminoácidos carecían de sus cadenas laterales, con el propósito de tener un modelo computacionalmente simplificado. Se adicionaron los átomos de hidrógeno faltantes para obtener un modelo en el cual todas las valencias estuvieran satisfechas, y luego se realizó la optimización de los átomos de hidrógeno del modelo de la región bisagra-inhibidor con el método B3LYP/6-31G\*<sup>22,23</sup>. Es importante mencionar que los átomos pesados del modelo se mantuvieron en sus posiciones cristalográficas y por lo tanto no se consideró en ningún momento la flexibilidad, tanto de la región bisagra como de los inhibidores. Los modelos moleculares para los demás sistemas fueron obtenidos a partir de las estructuras cristalográficas mencionadas, y se construyeron reemplazando algunos átomos de los modelos iniciales con el fin de obtener el inhibidor deseado (Ver Esquema 4.1.). Luego de esto se realizó una minimización de energía incluyendo los átomos añadidos al modelo con la funcional híbrida B3LYP/6-31G\*.



**Figura 4.1.** Modelo de la región bisagra de la CDK2 en interacción con el inhibidor Cp15. Las interacciones de puentes de hidrógeno son representadas por líneas negras discontinuas y el esqueleto de los aminoácidos implicados en dichas interacciones está encerrado en los cuadrados delineados y corresponden a los residuos Glu81 y Leu83.

Todos los cálculos se realizaron con el paquete de programas GAUSSIAN98.<sup>24</sup> Los análisis de las energías de interacción ( $\Delta E^g$ ) entre los diferentes inhibidores y el modelo de la región bisagra estudiados, se calcularon a partir de la siguiente ecuación:

$$\Delta E^g = E_{AB} - E_A - E_B \quad (4.1)$$

Donde  $\Delta E^g$  representa la energía de interacción calculada en fase gas,  $E_{AB}$  es la energía obtenida para el sistema en interacción,  $E_A$  es la energía calculada para el modelo de la región bisagra y  $E_B$  es la energía calculada para el inhibidor. Esta aproximación nos da

una estimación gruesa de la energía de interacción de cada inhibidor con la región bisagra en fase gas. Por otro lado los valores de los índices globales, locales y regionales de reactividad se obtuvieron usando las ecuaciones presentadas en el Apéndice 2.

Específicamente el análisis de la blandura grupal para la interacción inhibidor-modelo del sitio activo se calculó a partir de las siguientes ecuaciones:

$$s_G^+ \text{ inhibidor} = s^+(\text{H9}) + s^+(\text{C8}) + s^+(\text{N9}) + s^+(\text{C4}) + s^+(\text{C2}) + s^+(\text{N-2}) + s^+(\text{HN-2}) + s^+(\text{H2 ó s}^+(\text{átomo que este ubicado en esta misma posición en los demás inhibidores})) \quad (4.2)$$

$$s_G^- \text{ modelo-sitioactivo} = s^-(\text{C}\alpha[\text{E}]) + s^-(\text{C}[\text{E}]) + s^-(\text{O}[\text{E}]) + s^-(\text{N}[\text{F}]) + s^-(\text{C}\alpha[\text{L}]) + s^-(\text{C}[\text{L}]) + s^-(\text{O}[\text{L}]) + s^-(\text{N}[\text{H}]) \quad (4.3)$$

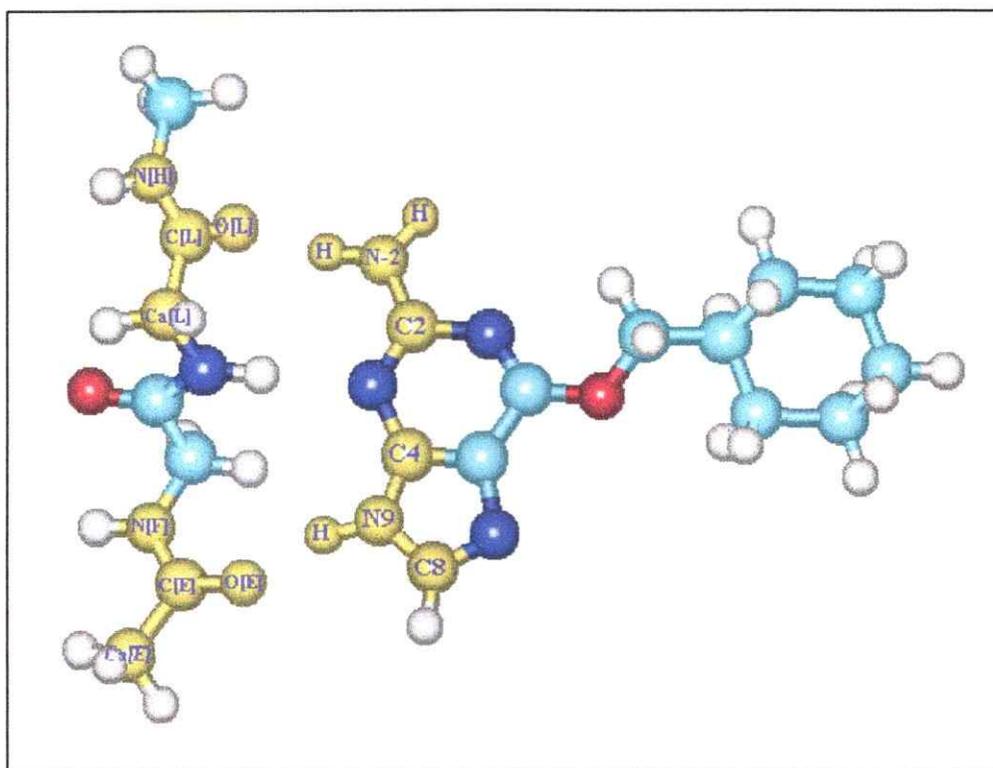
Donde  $s_G^+$  **inhibidor** es la blandura regional del inhibidor (el electrofilo) y  $s_G^-$  **modelo-sitioactivo** es la blandura regional del modelo del sitio activo (el nucleofilo). En este último, los nombres de los residuos, a los que pertenecen los átomos involucrados en el análisis, están asignados según la notación corta adoptada para nombrar los aminoácidos, así: [E], [F], [L] y [H] corresponden a los residuos Glutámico, Fenilalanina, Leucina e Histidina, respectivamente. Utilizando estas ecuaciones encontramos las contribuciones de cada átomo a la blandura regional de grupo, y en los inhibidores el único átomo que cambia es el que está unido al N2 y que forma parte de los diferentes grupos sustituyentes que se incluyen en esta posición del inhibidor. En la Figura 4.2 podemos visualizar de una manera más clara el modelo escogido para realizar el análisis de la blandura regional, el cual nos va a permitir luego calcular los porcentajes de similitud entre los dos fragmentos e interpretar los resultados en el marco del principio HSAB.

### IV.2.2. Índices de reactividad obtenidos a partir de la aproximación de la densidad de estados (DOS).

De acuerdo con la aproximación de Parr y Yang,<sup>19</sup> la blandura global,  $S$ , de un sistema extendido esta dada por la densidad de estados total en el nivel de Fermi (Ver Apéndice 2). Si extendemos cada estado electrónico como una distribución Gaussiana, podemos escribir:

$$S = g(E_F) = \frac{1}{\Delta\sqrt{\pi}} \sum_i g_i \exp\left[-\left(\frac{E_F - \epsilon_i}{\Delta}\right)^2\right] \quad (4.4)$$

donde  $\epsilon_i$  son los valores propios de los orbitales,  $g_i$  es la degeneración del estado  $i$ ,  $E_F$  es el nivel de Fermi, y  $\Delta$  es la amplitud de la función Gaussiana. Al parámetro  $\Delta$  se le asigna típicamente un valor de 0.5 eV.



**Figura 4.2.** Modelo adoptado para evaluar las blanduras regionales en cada fragmento y el porcentaje de similaridad entre ellos. Los átomos en verde y debidamente etiquetados son los implicados en el análisis.

Otra herramienta útil para el análisis local de la reactividad en sistemas con muchos electrones es la densidad de estados condensada a átomos. Esta cantidad nos indica cuantos estados electrónicos proporciona un átomo en un rango dado de energía, y está definida como:

$$g(E)_k = \frac{1}{\Delta\sqrt{\pi}} \sum_i g_i P_{ki} \exp\left[-\left(\frac{E - \epsilon_i}{\Delta}\right)^2\right] \quad (4.5)$$

donde  $P_{ki}$  es la fracción de densidad electrónica del orbital molecular  $i$ , el cual está localizado alrededor del átomo  $k$ . Esta fracción se obtiene a partir de un método alternativo para evaluar funciones de Fukui regionales condensadas a átomos propuesta por Contreras *et al.*<sup>25</sup>

El índice de electrofilia  $\omega = \mu^2/2\eta$ , fue evaluado asumiendo que  $\mu_{DOS} = E_{HOMO} \approx E_{Fermi}$  con  $\eta = 1/g(E_F)$ .

Las estructuras modelo utilizadas para evaluar los índices globales de reactividad a través de la DOS, para cada inhibidor y su correspondiente sitio activo, fueron construidas partiendo de estructuras finales minimizadas que se obtuvieron de estudios con métodos híbridos QM/MM (Quantum Mechanics/Molecular Mechanics) previos que se realizaron con el sistema Thr160-CDK2-CiclinaA complejado con 5 inhibidores diferentes.<sup>26</sup>

El modelo escogido esta vez corresponde a la secuencia de aminoácidos Phe80-Glu81-Phe82-Leu83-His84-Gln85-Asp86 y el inhibidor. Los enlaces no saturados en los extremos del polipéptido fueron saturados con átomos de hidrógeno. La geometría de los grupos de átomos adicionados para obtener un nuevo inhibidor y los átomos de hidrógeno adicionados, fueron optimizados manteniendo fija la posición de los átomos del polipéptido y del resto de la estructura del inhibidor.

El potencial electrostático molecular (PEM) representado por  $V(\mathbf{r})$ , en el punto  $\mathbf{r}$  debido a un sistema molecular con cargas nucleares  $\{Z_A\}$  localizadas en  $\{R_A\}$  y con densidad electrónica  $\rho(\mathbf{r})$  se puede expresar como<sup>27</sup>:

$$V(\vec{r}) = \sum_A^N \frac{Z_A}{|\vec{r} - \vec{R}_A|} - \int \frac{\rho(\vec{r}') d^3 r'}{|\vec{r} - \vec{r}'|} \quad (4.6)$$

donde  $N$  es el número total de núcleos localizados en  $\mathbf{R}_A$  en la molécula y  $Z_A$  es la carga nuclear del átomo A.

El PEM fue evaluado a través de la Ecuación (4.6) utilizando las coordenadas Cartesianas obtenidas a partir de los cálculos de los índices de reactividad derivados de la DOS. Todos los cálculos fueron realizados con el paquete de programas GAUSSIAN98<sup>24</sup> a un nivel DFT, con la funcional de intercambio y correlación B3LYP y con la base 6-31G\*. Primero, se calculó la densidad total a partir de la función de onda, para ello se utilizó una malla de 80 puntos por lado. Luego, el PEM fue trazado sobre la iso-superficie de densidad total a un valor de 0.02 unidades atómica (u.a.).

### IV.3. Resultados y discusión.

#### IV.3.1. Índices globales, locales y regionales de reactividad calculados a partir de la aproximación de orbitales moleculares de frontera.

Los valores de los índices globales de reactividad calculados a partir de la aproximación de orbitales de frontera están consignados en la Tabla 4.1.

A partir de estos resultados se realizaron algunas correlaciones de cada índice global de reactividad contra la actividad biológica expresada como IC<sub>50</sub>. Los resultados de las gráficas de correlación (Ver Anexo 1) nos sugieren la clasificación del conjunto de inhibidores estudiados en diferentes subgrupos de acuerdo al grupo sustituyente presente en la posición C2 del anillo purina del inhibidor. En general podemos ver que hay correspondencia entre las actividades inhibitorias y la blandura y el potencial químico electrónico para cada molécula; esto es a medida que las moléculas se hacen más blandas y tienen una mayor capacidad para intercambiar electrones con el entorno se hacen consistentemente más activas biológicamente. En la Figura 4.3 podemos ver la gráfica de correlación de la blandura global contra la IC<sub>50</sub>, obtenida para el primer grupo

de moléculas descritas en la Tabla 4.1. Aunque el índice de correlación,  $R^2$ , es bastante aceptable, la correlación carece de valor estadístico debido a la baja cantidad de puntos analizados. Al realizar las gráficas de correlación incluyendo todos los compuestos, se pudo observar que no existía ninguna correlación significativa entre los valores de los índices de reactividad calculados y la actividad media inhibitoria,  $IC_{50}$ . Por lo tanto podemos concluir que la deficiencia del método para clasificar los inhibidores de acuerdo a su  $IC_{50}$  se puede deber a dos causas principales:

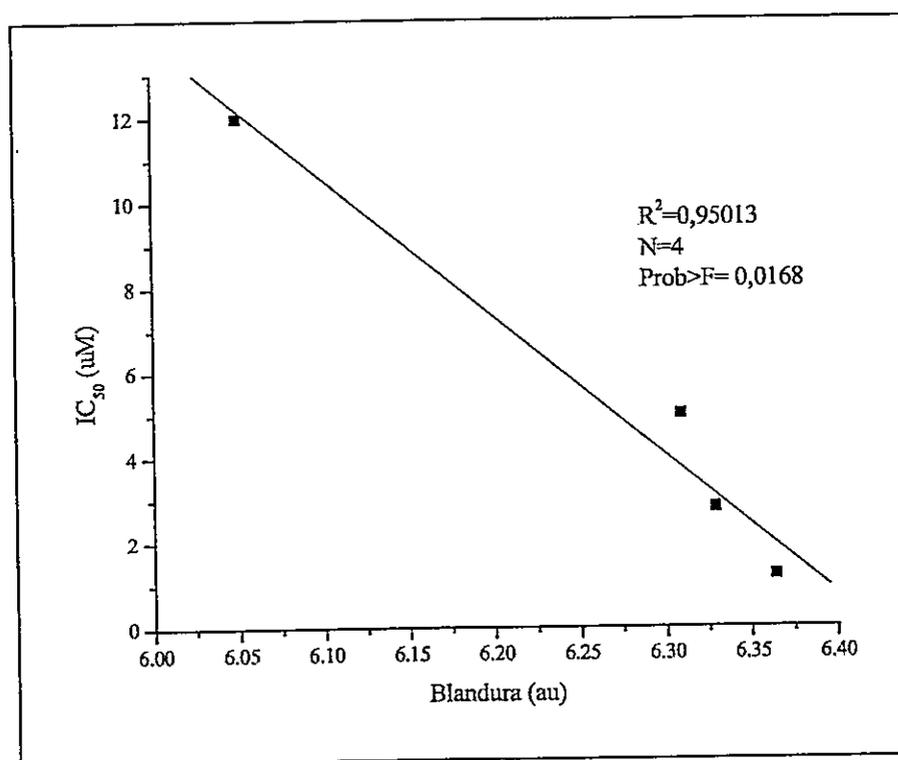


Figura 4.3. Correlación entre la actividad media inhibitoria ( $IC_{50}$ ) y la blandura global de los compuestos del grupo I.

1. La rigidez y el tamaño del modelo escogido; lo cual no permite describir de una manera adecuada las interacciones de cada inhibidor con la región bisagra de la enzima y a su vez no permite incluir los efectos de los demás aminoácidos y el solvente.
2. La utilización del modelo de orbitales de frontera, en el cual sólo se toma en cuenta para el cálculo de los índices el último orbital ocupado y el primer orbital virtual, con lo cual se ignora por completo las contribuciones de los demás orbitales. Esto último

es bien sabido que tiene suma importancia en sistemas condensados y biológicos como el estudiado en la presente tesis.

Más adelante veremos como la aproximación de la densidad de estados nos permitirá mejorar considerablemente las correlaciones encontradas, debido a la resolución de uno de los puntos problemáticos anteriormente expuestos.

En cuanto al análisis de las energías de interacción podemos decir que, aunque se encontraron correlaciones de las mismas características que las encontradas para los índices de reactividad, estas no tuvieron un significado correcto en la interpretación de las afinidades de los inhibidores por la región bisagra de la CDK2, ya que las energías de interacción son más favorables para los compuestos menos activos y viceversa. Esto se puede deber a la incapacidad del modelo para representar de una manera adecuada todos los componentes implicados en la interacción proteína-ligando y está relacionado, obviamente, con los dos puntos débiles del modelo explicados anteriormente.

Finalmente, nos referiremos a los índices regionales de blandura calculados en cada fragmento del sistema en estudio. En la Tabla 4.2 se muestran las principales cantidades calculadas, como son los valores de blandura grupal en el inhibidor y en el modelo de la región bisagra y el porcentaje de complementariedad (Match) obtenido. Este tipo de análisis ha sido aplicado exitosamente por Rivas *et al.*<sup>15</sup> en el estudio de la transferencia de hidruro entre un modelo del grupo isoaloxazina de la flavina y el grupo nicotinamida del NAD(P)H.

En los resultados presentados, las moléculas están organizadas de acuerdo a su actividad biológica  $IC_{50}$  en orden descendente, y podemos ver que hay correspondencia entre los porcentajes de complementariedad obtenidos y esta actividad medida experimentalmente. Sin embargo no en todos los grupos de moléculas se cumple la correspondencia mencionada, en particular en los compuestos de los grupos IIIa y IIIb. Este análisis realizado en el grupo de moléculas es interesante desde el punto de vista de las interacciones duras y blandas consignadas en la teoría de ácidos y bases duros y blandos (Hard and Soft Acids and Bases (HSAB)). Como podemos observar en los

CÁPITULO IV. ÍNDICES DE REACTIVIDAD A PARTIR DE LA DOS

Compuesto	$\mu_{\text{OFM}}$ (eV)	$\omega_{\text{OFM}}$ (eV)	$S_{\text{OFM}}$ (a.u.)	IC <sub>50</sub> ( $\mu\text{M}$ )
Grupo I				
Nu2058	-3.581	1.426	6.05	12.0
Comp7	-3.447	1.377	6.309	5.0
Comp8	-3.434	1.372	6.329	2.8
Comp9	-3.411	1.361	6.364	1.2
Grupo IIa				
Comp18	-3.923	1.84	6.508	12.0
Comp15	-3.757	1.714	6.607	2.3
Comp20	-3.666	1.667	6.749	0.4
Comp25	-3.492	1.595	7.118	0.069
Grupo IIb				
Comp18	-3.923	1.84	6.508	12.0
Comp17	-3.76	1.707	6.573	6.8
Comp21	-3.566	1.593	6.819	1.8
Comp22	-3.5	1.584	7.039	1.7
Nu6094	-3.6	1.621	6.808	1.0
Comp26	-3.438	1.567	7.215	0.65
Grupo IIIa				
Comp39	-3.933	1.927	6.778	0.8
Comp38	-3.919	1.932	6.844	0.3
Comp31	-3.832	1.861	6.894	0.1
Grupo IIIb				
Comp30	-4.135	2.055	6.541	0.063
Comp29	-4.085	2.021	6.589	0.056
NU6102	-4.072	2.001	6.568	0.0054

Tabla 4.1. Valores de los índices globales de reactividad calculados a partir del OFM y clasificación sugerida de los compuestos, de acuerdo al grupo sustituyente en la posición C2 del anillo purina.

resultados presentados en la Tabla 4.2, las cantidades calculadas corresponden a las blanduras regionales obtenidas en cada fragmento. Además estamos suponiendo que todos los átomos del inhibidor, incluidos en el análisis, se comportan como nucleófilos y a su vez todos los átomos presentes en la región bisagra se comportan como electrofilos. El principio HSAB sugiere que las interacciones blandas corresponden a las interacciones mediadas por orbitales y en nuestro modelo las interacciones predominantes son mediadas por carga, ya que corresponden a los enlaces de hidrógeno característicos que se forman entre los ligandos y los residuos Glu81 y Leu83 en la región bisagra de la CDK2. No se ha analizado la "dureza regional" en nuestro modelo, la cual debería representar bien las interacciones de enlace de hidrógeno, ya que no existe en el momento un modelo teórico adecuado para hacerlo. Por lo tanto, en lugar de analizar todas las interacciones en el sistema, solo estamos analizando la polarizabilidad (que también es una de las propiedades con las cuales se puede relacionar a la blandura) de un determinado grupo presente en los inhibidores. Con esto podemos decir que el modelo a pesar de estar incompleto describe razonablemente bien, y de una manera cualitativa, el comportamiento de los inhibidores en interacción con la región bisagra de la CDK2. En resumen, podemos decir que en esta primera aproximación del análisis de la afinidad de los inhibidores por la región bisagra, encontramos algunas dificultades en la aplicación del principio HSAB y que los problemas mayores se encuentran en el tamaño del modelo estudiado y en la rigidez del mismo (el análisis se realizó sobre las configuraciones atómicas de la estructura cristalográfica). Estas dificultades pueden ser corregidas utilizando un modelo que no haga diferencia acerca de las características de la interacción (nucleófilo ó electrofilo) y especialmente el último problema, referido a la rigidez del modelo, puede ser resuelto utilizando estructuras de mínima energía que provengan de estudios de dinámica molecular. Como veremos en la sección siguiente, vamos a utilizar las estructuras obtenidas de los estudios con métodos híbridos QM/MM reportados en el Capítulo III para construir nuevos modelos y vamos a analizar los índices globales de reactividad a partir de la aproximación de la densidad de estados, con el fin de mejorar nuestro modelo.

CÁPITULO IV. ÍNDICES DE REACTIVIDAD A PARTIR DE LA DOS

Compuesto	$s_G^+$ (Inhibidor)	$s_G^-$ (Modelo sitio activo)	Match (%)
Grupo I			
Nu2058	5.752	6.323	90.97
Comp7	11.861	12.762	92.94
Comp8	11.957	12.759	93.71
Comp9	11.853	12.767	92.84
Grupo IIa			
Comp18	10.755	12.169	88.38
Comp15	10.992	12.167	90.34
Comp20	11.29	12.166	92.80
Comp25	12.016	12.167	98.76
Grupo IIb			
Comp18	10.755	12.169	88.38
Comp17	10.913	12.168	89.69
Comp21	11.412	12.160	93.85
Nu6094	5.735	6.083	94.28
Comp22	11.772	12.162	96.79
Comp26	12.158	12.165	99.94
Grupo IIIa			
Comp39	9.735	10.979	88.67
Comp38	9.561	10.975	87.12
Comp31	11.377	10.994	96.63
Grupo IIIb			
Comp30	10.377	10.985	94.47
Comp29	10.448	10.983	95.13
NU6102	5.215	5.491	94.97

Tabla 4.2. Valores de blandura grupal y complementariedad (Match) para los inhibidores y el modelo de sitio activo.

### IV.3.2. Índices globales de reactividad calculados a partir de la DOS.

Los valores de los índices globales de reactividad calculados a partir de la densidad de estados (DOS) están consignados en la Tabla 4.3. Es de conocimiento general que el potencial químico ( $\mu$ ) es el descriptor natural de la dirección de transferencia de carga (TC) durante una interacción química.<sup>28</sup> Hemos encontrado que los valores de  $\mu$  obtenidos para cada inhibidor son más pequeños que los valores correspondientes obtenidos para el modelo del sitio activo. Este resultado puede ser un indicio de que la transferencia de carga ocurre desde el sitio activo hacia el inhibidor, principalmente en los sitios donde el ligando puede establecer un enlace de hidrógeno. Con los valores de  $\mu$  y  $\eta$  para los ligandos y el modelo de la región bisagra (RB), se puede tener una estimación de la transferencia de carga (TC) del sistema en interacción, ligando (I)-modelo de la RB, utilizando la ecuación de Pearson  $\Delta N = (\mu_I - \mu_{HR}) / (2 * (\eta_I + \eta_{HR}))$ .<sup>29</sup> Se pudo observar que en general, la transferencia de carga entre ambos fragmentos es despreciable (Ver Tabla 4.3). Luego graficamos los valores de  $\mu$  contra el  $\text{LogIC}_{50}$  para el conjunto de moléculas consignadas en la Tabla 4.3. La comparación (no se muestra aquí) mostró un índice de correlación muy pobre:  $R^2 = 0.19$ . Por lo tanto, podríamos concluir que no existe una correlación significativa entre los valores de  $\mu$  y la actividad biológica de la serie de compuestos considerados, un resultado que indica que la TC no puede ser uno de los principales factores de los cuales depende la interacción ligando-RB. Luego realizamos dos análisis estadísticos adicionales, uno incorporando la blandura global,  $S$ , y el otro con el índice global de electrofilia,  $\omega$ . Los análisis por separado, incorporando la serie completa de compuestos de la Tabla 4.3, sugieren una clasificación de la serie completa en dos grupos de ligandos que están en concordancia cualitativa con la clasificación realizada por Hardcastle *et al.*<sup>10</sup>, la cual está basada en los sitios explorados en el bolsillo activo de la CDK2. El primero, grupo 1, está formado por los compuestos 17, 18, 21 y 22 que pertenecen a las  $C^2$ -anilino- $O^6$ -ciclo-hexil-metil-purinas 3'-sustituídas que se estudiaron. Este grupo fue aumentado con el compuesto 31 porque en nuestro modelo este compuesto está explorando la misma región en el sitio activo. También se adicionó el compuesto líder 1. Los primeros 4 compuestos

mencionados fueron diseñados para simular una interacción favorable, reportada anteriormente, entre el grupo 3'-cloro en el purvanalol y la enzima.<sup>30</sup> El segundo, grupo 2, incluye los compuestos 2, 3, 15, 20, 25, 26, 29, 30, 38, y 39, más el compuesto líder 1. Todos fueron diseñados para explorar un grupo aceptor de enlace de hidrógeno en una región de la proteína que es accedida por el grupo 2-anilino, y que es presumiblemente el residuo Asp86.<sup>21</sup> Los valores de la blandura global ( $S$ ) basados en la DOS para los grupos 1 y 2 que se presentan en la Tabla 4.3, muestran una correlación más significativa con el  $\text{LogIC}_{50}$  cuando se comparan con la obtenida con el potencial químico electrónico. El resultado de esta comparación da coeficientes de correlación con valores de 0.82 y 0.80 para los grupos 1 y 2, respectivamente (Ver Figuras 4.4 y 4.5). Puesto que la blandura global está relacionada con la polarizabilidad electrónica del sistema,<sup>31</sup> podemos concluir que los efectos de polarización son más importantes que los efectos de TC para estas interacciones.

Motivados por los resultados reportados por Maynard *et al.*,<sup>13</sup> decidimos realizar una última comparación entre un índice de reactividad similar, el cual es definido nuevamente en términos de la DOS, y que llamaremos de aquí en adelante  $\omega_{\text{DOS}}$ , y el  $\text{LogIC}_{50}$ .

En principio, este índice debería dar mejores resultados en comparación con la correlación obtenida para el  $\text{LogIC}_{50}$ -Blandura global, ya que contiene más información acerca de la estructura electrónica de los ligandos, a saber, el cuadrado de la electronegatividad y la blandura global  $S=1/\eta$ . Los valores calculados para el índice global de electrofilia también se muestran en la Tabla 4.3. Los coeficientes de correlación son 0.85 y 0.65 para los grupos 1 y 2, respectivamente. A pesar que para el grupo 2 la correlación  $\text{LogIC}_{50}$ -electrofilia es ligeramente mejor, es evidente que cuando se consideran los dos grupos, la correlación  $\text{LogIC}_{50}$ -Blandura es más estable. El fracaso de la electrofilia global en la correlación cualitativa con los valores de  $\text{LogIC}_{50}$  puede ser probablemente explicada por el hecho que, en forma contraria a los resultados de Maynard *et al.*, la interacción ligando-RB tiene un componente covalente despreciable, el cual es un resultado que ya se anticipaba a partir de los resultados de la correlación del  $\text{LogIC}_{50}$  con el potencial químico electrónico.

CÁPITULO IV. ÍNDICES DE REACTIVIDAD A PARTIR DE LA DOS

Compd #	$\mu_{\text{DOS-SITIO}}$ (eV)	$\mu_{\text{DOS}}$ (eV)	$S_{\text{DOS}}$ (eV)	$\omega_{\text{DOS}}$ (eV)	$\Delta N$ (electrones)	IC <sub>50</sub> (μM)
1	-0.60	-5.46	1.7105	25.48	-0.095	17.0
2	-0.60	-5.30	1.8156	25.46	-0.094	0.97
3	-0.60	-5.77	2.4889	41.44	-0.126	0.0054
7	-0.60	-5.33	1.6680	23.68	-0.091	5.0
8	-0.64	-5.30	1.6597	23.34	-0.090	2.8
9	-0.64	-5.31	1.6632	23.46	-0.090	1.2
15	-1.18	-5.44	1.8856	27.95	-0.094	2.3
17	-0.66	-5.51	2.0624	31.31	-0.106	6.8
18	-1.17	-5.59	1.9259	30.11	-0.099	12.0
20	-0.60	-5.27	1.8777	26.03	-0.096	0.4
21	-1.14	-5.30	2.5008	35.12	-0.105	1.8
22	-1.14	-5.33	3.2105	45.53	-0.121	1.7
25	-0.64	-5.18	2.0223	27.10	-0.098	0.069
26	-0.68	-5.11	1.9523	25.50	-0.093	0.65
29	-1.12	-5.70	2.0085	32.65	-0.103	0.056
30	-0.82	-5.76	2.0753	34.37	-0.111	0.063
31	-0.84	-5.46	3.3637	50.17	-0.135	0.10
38	-1.11	-5.50	2.0363	30.75	-0.102	0.30
39	-1.21	-5.57	1.8084	28.03	-0.090	0.8

Tabla 4.3. Descriptores globales de reactividad calculados a partir de la aproximación de la DOS para el conjunto de inhibidores de CDK2 estudiados y sus respectivos modelos del sitio activo.

Una ventaja adicional de la aproximación DOS es la posibilidad de analizar la DOS condensada en átomos individuales. En la Figura 4.6 se muestra la DOS global y condensada a átomos para los compuestos más representativos. Las gráficas de la densidad de estados local y condensada para los demás compuestos se presentan en el

Anexo 1. Analizamos las contribuciones individuales de los átomos a la DOS global para cada ligando.

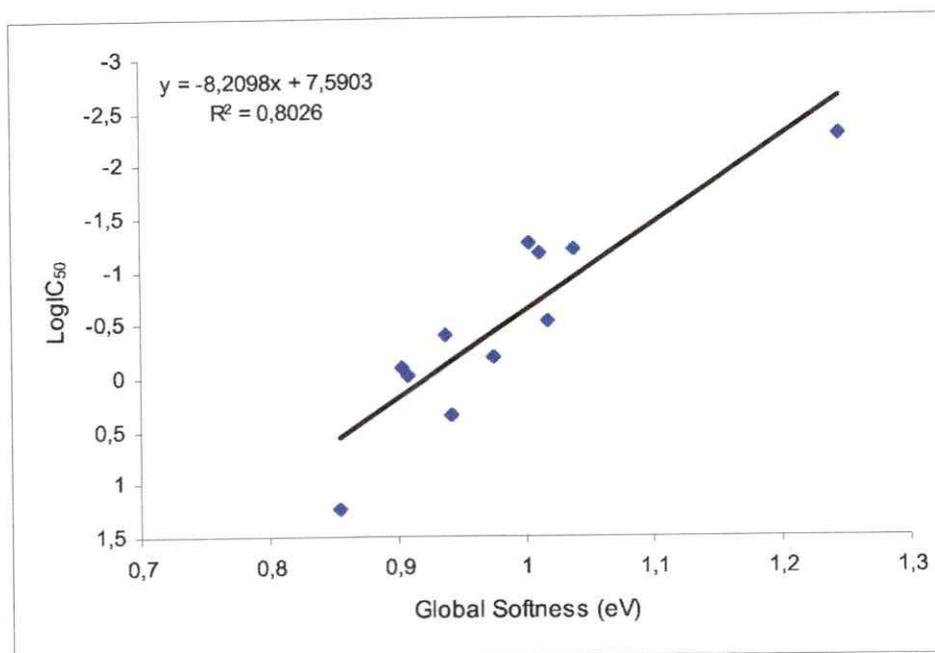


Figura 4.4. Gráfica de LogIC<sub>50</sub> vs Blandura Global ( $S_{DOS}$ ) (Grupo II).

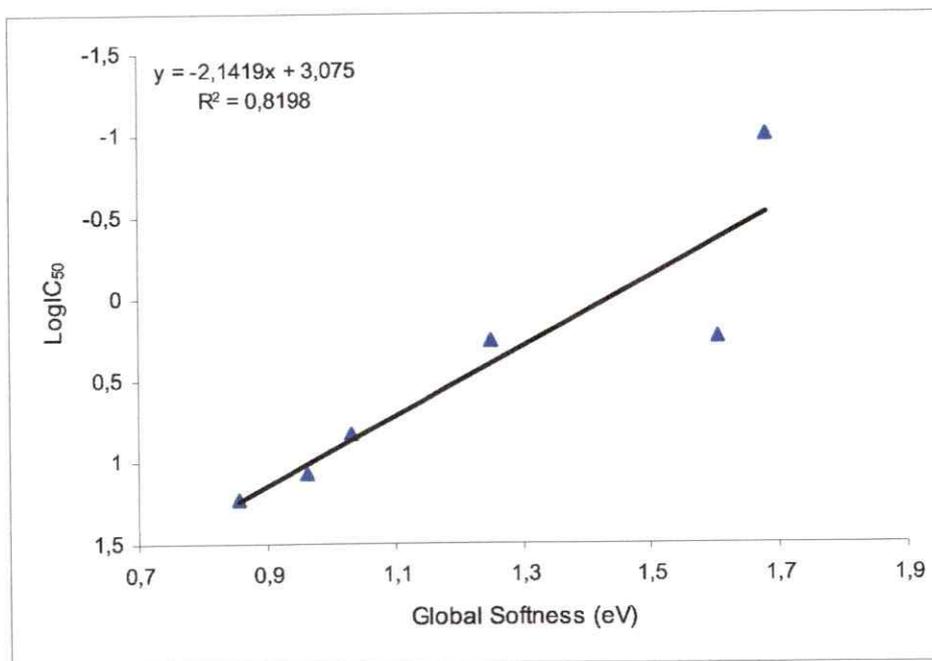


Figura 4.5. Gráfica de LogIC<sub>50</sub> vs Blandura Global ( $S_{DOS}$ ) (Grupo I).

Los átomos seleccionados fueron N3, N9 y N<sup>2</sup>, que pertenecen al anillo purínico, debido al conocido rol que ellos cumplen en la interacción con la RB.

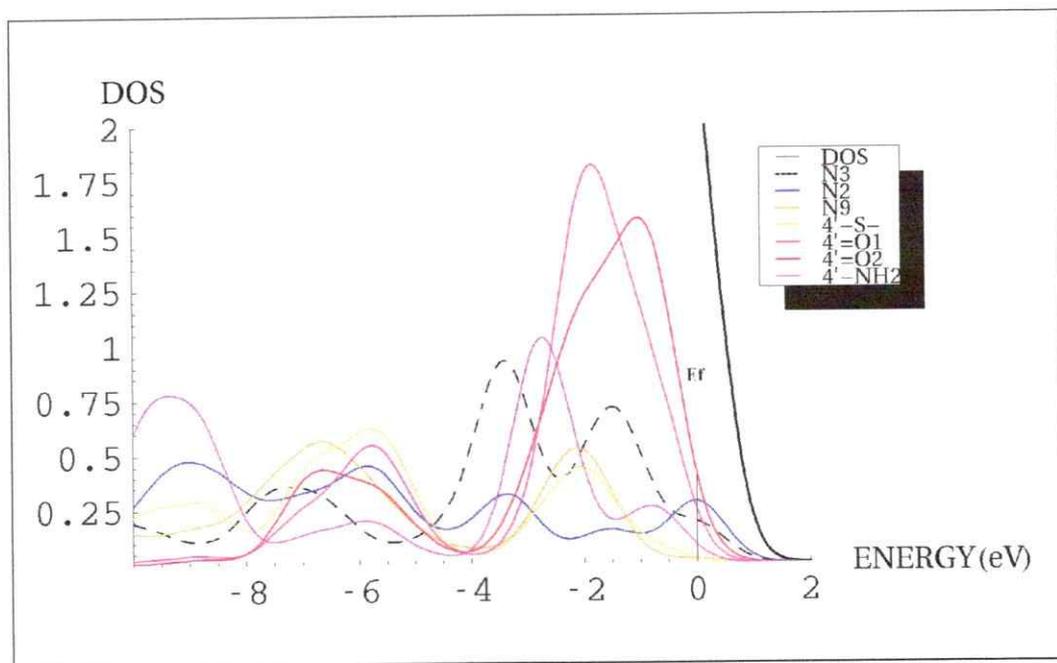
Además hemos analizado las contribuciones debidas a los nuevos átomos, adicionados en la posición N<sup>2</sup> para obtener un inhibidor diferente, a la DOS global de los ligandos. Hemos encontrado que las contribuciones de los átomos de carbono en los grupos metil, etil e isopropil de los compuestos 7, 8 y 9 son casi las mismas.

La ganancia en potencia de esos compuestos con respecto al compuesto líder 1 puede ser explicada por una polarización aumentada del ligando debida a la contribución de cada grupo sustituyente a la DOS cerca del nivel de Fermi. Santos *et al.* han demostrado que el cuadrado de la DOS condensada a átomos puede ser interpretada como una blandura local resuelta en un rango de energía significativo.<sup>14</sup> Por ejemplo, cuando se adiciona un grupo anilino en la posición 2 (compuesto 2) del anillo purina, las contribuciones más importantes a la DOS global provienen de los átomos de carbono pertenecientes al anillo. Además, hay un ligero aumento en la contribución del átomo N3 del anillo purina. El patrón de sustitución con halógenos en las posiciones 3' y 5' (compuestos 15, 17 y 18) del anillo anilino da una importante contribución a la DOS global en el rango de energía E<sub>F</sub>-2 y E<sub>F</sub>-3 eV. Las contribuciones de los átomos de carbono pertenecientes al anillo anilino permanecen iguales. Cuando el sustituyente en la posición 3' es un grupo hidroximetilo (compuesto 20) la contribución del átomo de oxígeno es pequeña, pero debido a la posición de este grupo en el anillo, éste podría formar las interacciones esperadas con el residuo Asp86. Las contribuciones de los grupos 3'-metoxi y 3'-metilmercapto (compuestos 21 y 22, respectivamente) son importantes, siendo más significativa la del grupo metilmercapto. Sin embargo, esos grupos no pudieron formar las interacciones esperadas de enlace de hidrógeno con el residuo Asp86. Cuando el patrón de sustitución se cambia a la posición 4' en el anillo anilino, el cuadro de reactividad es muy diferente. Los grupos hidroxilo y metoxi en los compuestos 25 y 26 muestran contribuciones similares a la DOS global, pero en el compuesto 26 la contribución del átomo de oxígeno, del grupo metoxi, se reduce por la presencia del grupo -CH<sub>3</sub> unido a él. Esto puede ser observado experimentalmente por la imposibilidad del grupo metoxi para formar una interacción de enlace de hidrógeno con

el residuo Asp86 y consecuentemente la baja actividad biológica del compuesto 26, cuando se compara con la del compuesto 25. Cuando se adiciona un grupo sulfonamido en la posición 4' (compuesto 3) las principales contribuciones a la DOS global provienen de los átomos de oxígeno y nitrógeno pertenecientes al grupo sulfonamido, los cuales se espera realicen interacciones de enlace de hidrógeno fuertes con el residuo Asp86 ubicado en la superficie específica de la CDK2.

Es digno de mencionar que los valores de energía para las contribuciones atómicas mencionadas anteriormente están en el rango de  $E_F - 4$  y  $E_F$  eV, lo cual significa que esas contribuciones están más cercanas al nivel de Fermi cuando se comparan con las contribuciones realizadas por los demás átomos en los compuestos restantes. Como consecuencia, esos átomos son más blandos que cualquier otro presente en los compuestos estudiados. En el compuesto 29 las contribuciones de los grupos metilo en el sustituyente *N,N*-dimetil-sulfonamida sobrepasan las correspondientes al átomo de nitrógeno. Además hay una ligera disminución de las contribuciones de los átomos de oxígeno del grupo sulfonamida causada por los grupos metilo. Esto claramente indica una actividad biológica disminuida debido a la imposibilidad para formar las interacciones de enlace de hidrógeno con el residuo Asp86. Lo mismo es aplicable al derivado 4'-metil-sulfona (compuesto 30), donde las contribuciones más importantes del átomo de nitrógeno en el grupo sulfonamida se reemplazan por las contribuciones de un grupo metilo. En el derivado 4'-metil sulfoxido (compuesto 31) la ausencia de uno de los átomos de oxígeno en el fragmento reactivo provoca que el compuesto sea menos activo, aunque la DOS global es la más grande de toda la serie estudiada. Esto puede ser explicado por un gran incremento en las contribuciones del grupo metil-sulfoxido y los átomos N3, N9 y N<sup>2</sup>, cuando estas se comparan con los diferentes derivados sulfona (compuestos 3, 29 y 30). Sin embargo, como se pudo verificar en nuestro modelo final para este compuesto, el fragmento S=O en el grupo sulfoxido no puede lograr una geometría efectiva para la interacción con el residuo Asp86, lo que está en correcto acuerdo con los resultados cristalográficos reportados por Hardcastle *et al.*<sup>10</sup> Finalmente, en los compuestos 38 y 39, los grupos en la posición 4' fueron colocados para formar los correspondientes derivados acetofenona y ácido benzoico,

a)



b)

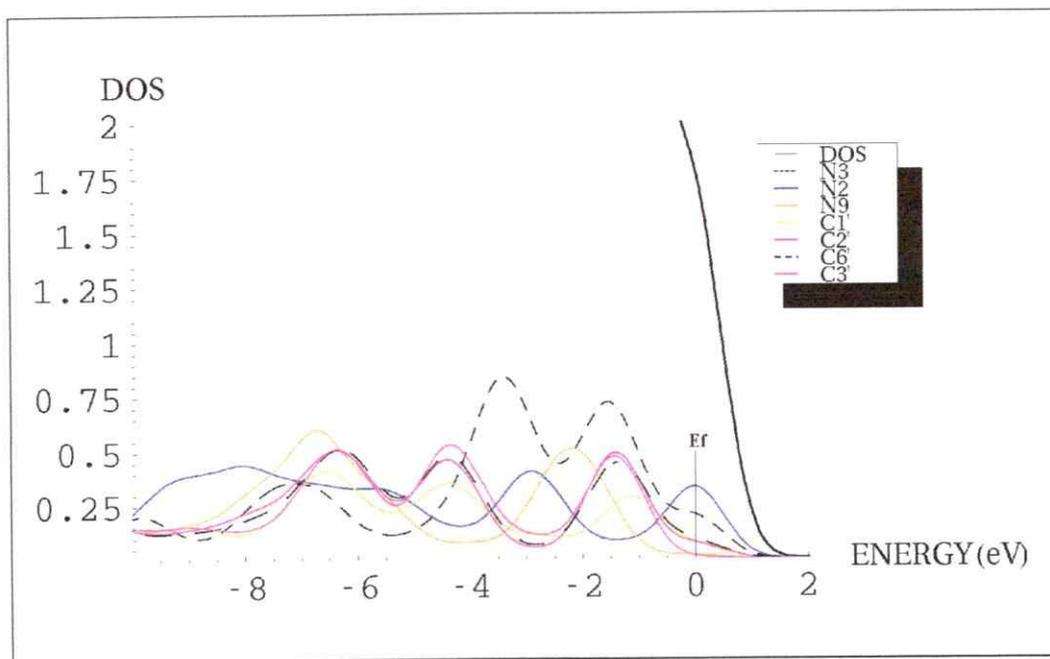


Figura 4.6. Densidad de estados global y condensada para los compuestos: a) 3 y b) 2.

respectivamente. Las contribuciones de los grupos  $-C=O$  tienen casi la mitad del valor que las correspondientes para los grupos  $-S=O$  en los derivados sulfona y sulfoxido. Estos compuestos muestran una actividad biológica disminuida debido a la carencia de una interacción efectiva con el residuo Asp86.

En resumen, este esquema simple parece ofrecer mejores resultados que los que fueron obtenidos con el análisis de los índices de reactividad basado en los modelos moleculares formulados en base a la teoría de orbitales moleculares de frontera. Este resultado puede ser debido a la inclusión de estados electrónicos adicionales que están muy cerca en energía al nivel HOMO, los cuales pueden realizar contribuciones relevantes al patrón de reactividad desplegado en estos sistemas. El patrón de reactividad para cada compuesto puede ser explicado cualitativamente en términos de las contribuciones atómicas o de grupos de átomos a la DOS global.

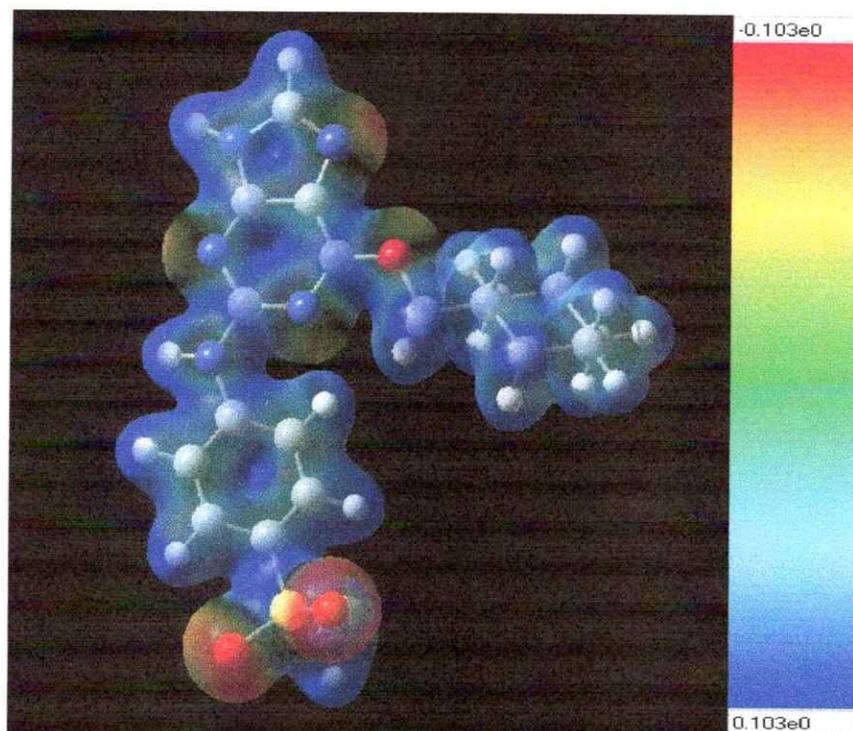
#### IV.3.2. Análisis del Potencial Electrostático Molecular.

Hemos llevado a cabo un análisis cualitativo basado en el potencial electrostático molecular (PEM) en el conjunto de moléculas estudiadas. Los análisis fueron realizados sobre la configuración geométrica final obtenida a partir del análisis de la DOS. En la Figura 4.7 se muestra el PEM para los compuestos más representativos de la serie estudiada. Las gráficas del PEM para los demás compuestos se presentan en el Anexo 1. Como se puede observar en la Figura 4.7, los sitios más nucleofílicos en cada inhibidor son los centros N3, N7, N1 y O<sup>6</sup> que pertenecen al anillo purina. Cuando se adicionan el anillo anilino, y sus diferentes grupos sustituyentes, en posición 3' y 4', en la posición 2 del anillo purina para conseguir los nuevos inhibidores; la reactividad es transferida a esos nuevos grupos de acuerdo al patrón de interacción esperado para cada compuesto. Por ejemplo, en los compuestos 15, 17, y 18, los átomos tipo halógeno son los más nucleofílicos del grupo sustituyente en la posición 2 del anillo purina.

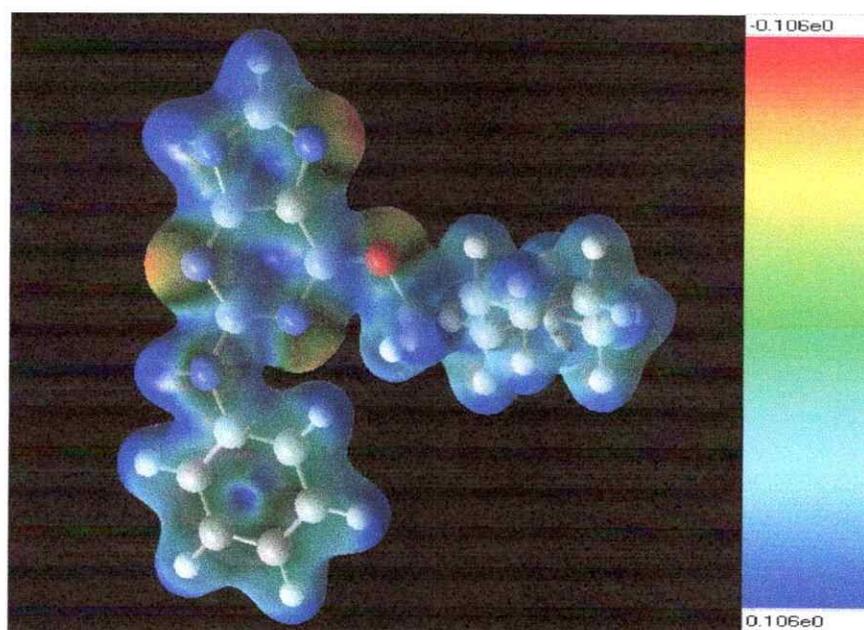
En general, en los compuestos restantes, los átomos más nucleofílicos son los de oxígeno y los halógenos que están localizados en los grupos sustituyentes en las

posiciones 3' y 4' del anillo anilino. Es importante mencionar que el centro del anillo anilino es uno de los sitios más nucleofílicos en cada ligando que posee este tipo de sustituyente. Este hallazgo está en perfecta concordancia con una interacción débil del tipo  $-C-H\cdots\pi$  entre el hidrógeno del carbono  $\alpha$  en el residuo Gln85 y el centro del anillo anilino, reportada previamente.<sup>21, 26</sup> Por el contrario, los sitios más electrofílicos en el anillo purina son los átomos de hidrógeno unidos a los átomos N9 y N<sup>2</sup> del mismo anillo. Obviamente, existen otros sitios electrofílicos en cada molécula de inhibidor, pero los anteriormente mencionados son especialmente importantes en la interacción complementaria con la RB de la CDK2, como se discutirá a continuación. En el modelo del sitio activo (residuos 80-86) existen algunos sitios nucleofílicos ubicados en los átomos de oxígeno carbonílicos pertenecientes al esqueleto (backbone) de la proteína y en los átomos de oxígeno carboxílicos de los residuos Glutamato y Aspártico. Los sitios nucleofílicos más importantes están situados sobre los átomos de oxígeno carbonílicos en el backbone de los residuos Glu81 y Leu83. Los sitios más electrofílicos están situados sobre los átomos de hidrógeno que pertenecen al grupo amido en el backbone del modelo del sitio activo. Habiendo definido este conjunto de sitios nucleofílicos y electrofílicos en los ligandos y en el modelo del sitio activo de la CDK2, podemos determinar la interacción complementaria entre los ligandos y la RB de la CDK2, la cual se da principalmente entre los átomos N3, H-N9 y H-N<sup>2</sup> en los ligandos y los átomos Glu81-C=O, Leu83-NH y Leu83-C=O (Ver Figura 4.8a) en la RB. Si analizamos más detenidamente el sistema en la búsqueda de otros posibles sitios en los ligandos y en el modelo del sitio activo para lograr esta interacción específica, no encontraremos esa complementariedad perfecta en otros sitios nucleofílicos y electrofílicos en cada parte que interacciona. Otra interacción complementaria importante que ocurre en este sistema se da entre un átomo de hidrógeno ubicado en un grupo sustituyente en la posición 4' del anillo anilino y uno de los átomos de oxígeno carboxílicos de la cadena lateral del residuo Asp86. Ésta última interacción entre los sitios electrofílico y nucleofílico mencionados solo puede ser llevada a cabo por el compuesto 3, el cual posee la geometría correcta y el grupo reactivo adecuado en la posición 4' del anillo anilino (Ver Figura 4.8b).

a)



b)



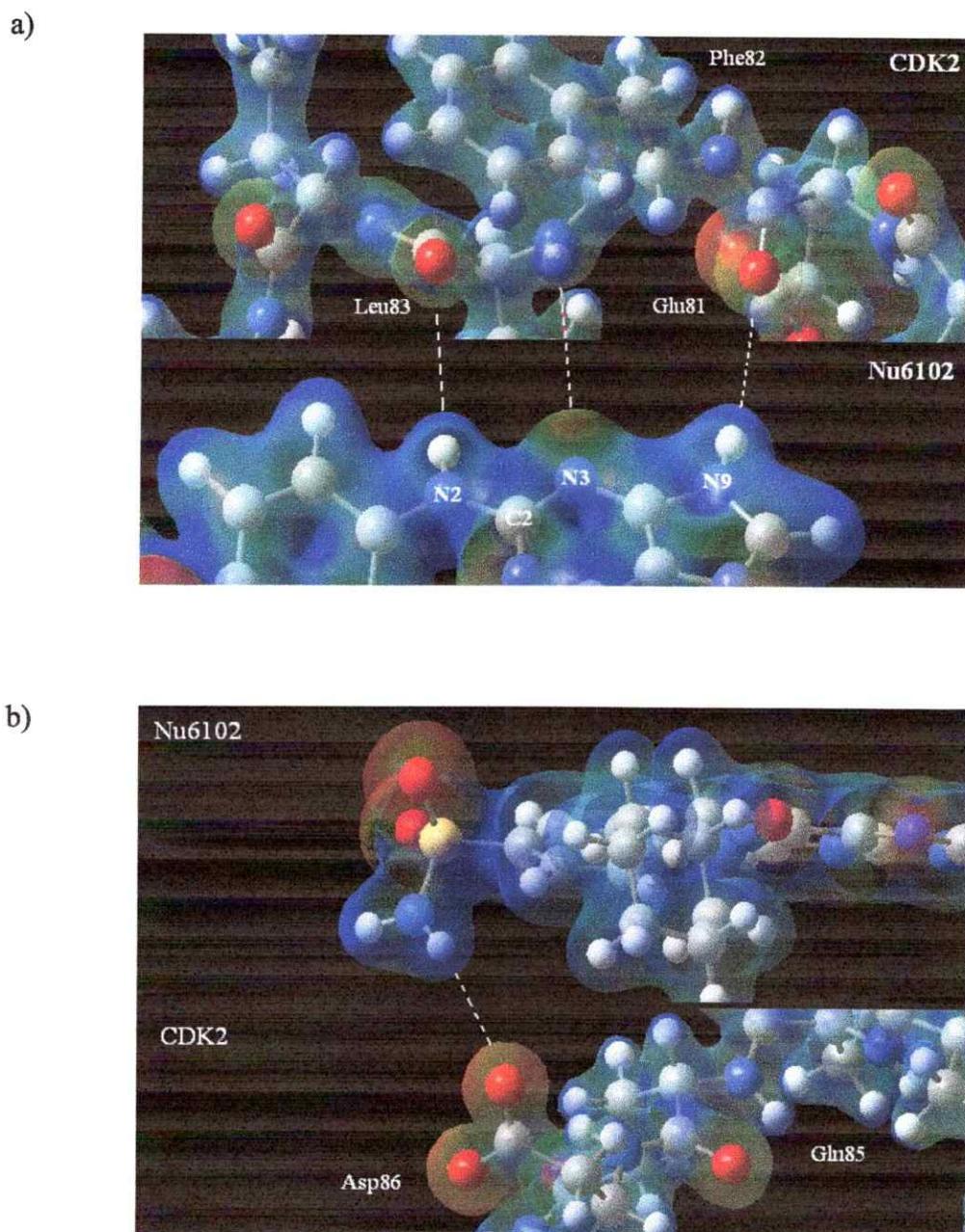
**Figura 4.7.** Potencial Electrostático Molecular (PEM) para los inhibidores de CDK2 más representativos: a) 3 y b) 2.

Finalmente, podemos decir que los resultados presentados sugieren que existe un paso de reconocimiento inicial entre el ligando y la enzima. Este paso inicial está gobernado por los sitios nucleofílicos y electrofílicos localizados en ambas partes que interactúan y esos sitios son descritos de una manera adecuada por el PEM.

#### IV.4. Conclusiones

En esta última parte de la tesis hemos realizado un estudio teórico para describir la afinidad de un conjunto de moléculas inhibitoras de la enzima CDK2 utilizando un modelo del sitio activo de la misma enzima. Hemos utilizado algunos descriptores globales de reactividad obtenidos a partir de las aproximaciones de orbitales moleculares de frontera y de la densidad de estados, para tratar de explicar el patrón de reactividad de una serie de compuestos estudiados. El trabajo con los índices de reactividad es complementado con un estudio del potencial electrostático molecular, el cual es muy útil en el contexto del modelo llave y cerradura propuesto por Fisher. Hemos encontrado que la descripción de la afinidad de los ligandos por la RB utilizando la aproximación del OFM tiene serias desventajas. Estas desventajas son claramente la utilización de un modelo muy pequeño que no involucra todos los aminoácidos participantes en la interacción proteína-ligando y la obtención de los índices de reactividad a partir de los orbitales moleculares frontera, que para estos sistemas condensados prácticamente no existe debido a la gran cantidad de electrones y al comportamiento típico de bandas en el nivel de Fermi. Cuando el mismo análisis se lleva a cabo utilizando la DOS, en la cual no se utilizan las energías de los orbitales HOMO y LUMO para la obtención de los índices de reactividad, sino una banda de energía en el nivel de Fermi; es posible obtener una mejor descripción de la afinidad de cada uno de los inhibidores estudiados por el sitio activo de la CDK2. Hemos demostrado que los índices de reactividad obtenidos a partir de la aproximación





**Figura 4.8.** Fotografía del PEM para: a) La interacción de enlace de hidrógeno entre la región bisagra de la CDK2 y el compuesto 3 (Nu6102) y b) la interacción de enlace de hidrógeno entre el residuo Asp86 y el grupo sulfonamida de Nu6102.

de la DOS, especialmente la blandura global y la electrofilia global, describen de una manera correcta la afinidad de los inhibidores de CDK2 por la región bisagra. Esta mejora en la descripción de la afinidad de los ligandos puede ser explicada por la

inclusión de estados electrónicos que están muy cerca de los orbitales frontera. Esos estados electrónicos pueden contener información más completa, y útil, acerca del patrón de reactividad global y local que los orbitales HOMO y LUMO por sí solos. Esto es especialmente útil para los sistemas biológicos y condensados donde la representación en términos de los orbitales de frontera es reemplazada por la representación de bandas de valencia y de conducción, ampliando la cantidad de estados electrónicos que pueden participar en un cuadro de reactividad. Finalmente, hemos confirmado la importancia y utilidad del PEM en el estudio y la predicción de los sitios reactivos en la interacción de reconocimiento proteína-ligando. Hemos demostrado que existe solo una posibilidad de interacción, en el sentido de la orientación inicial dentro del sitio activo, entre cada ligando y la región bisagra de la CDK2; y ésta corresponde con la ubicación de los átomos de nitrógeno en el anillo purina y a los grupos carbonilo y amido en los backbone de los residuos Glu81 y Leu83. Las demás interacciones se supone que ocurren después de esta interacción de reconocimiento inicial, y ellas dan la especificidad y potencia finales esperadas para cada inhibidor.

### IV.5. Bibliografía

- [1] Fukui, K. Recognition of stereochemical paths by orbital interaction *Acc. Chem. Res.* **1971**, *4*, 57-64.
- [2] Dewar, M.J.S. *The Molecular Orbital Theory of Organic Chemistry*. p. 47, McGraw-Hill: New York. **1969**.
- [3] Fleming, I. *Frontier Orbitals and Organic Chemical Reactions*. p. 24, Wiley: New York and London. **1976**.
- [4] Hoffmann, R. *Solids and Surfaces: A Chemist's View of Bonding in Extended Structures*, VCH, New York, **1988**.
- [5] Tomasi, J.; Mennucci, B.; Cammi, R. 'MEP: a tool for interpretation and prediction. From molecular structure to solvation effects', in: J.S. Murray, K. Sen (eds.), *Molecular Electrostatic Potentials. Concepts and Applications*, Elsevier, Amsterdam, pp. 1-103, **1996**.
- [6] Tomasi, J. 'Models and modeling in theoretical chemistry', *Journal of Molecular Structure (Theochem)*, **1988**, *179*, 273-292.
- [7] Sims, P. A.; Wong, C. F.; McCammon, J. A. *J. Med. Chem.* **2003**, *46*, 3314-3325.
- [8] Tominaga, Y.; Jorgensen, W. L. *J. Med. Chem.* **2004**, *47*, 2534-2549.
- [9] Gräter, F.; Schwarzl, S. M.; Dejaegere, A.; Fischer, S.; Smith, J. C. *J. Phys. Chem. B* **2005**, *109*, 10474-10483.
- [10] Hardcastle, I. R.; Arris, C. E.; Bentley, J.; Boyle, F. T.; Chen, Y.; Curtin, N. J.; Endicott, J. A.; Gibson, A. E.; Golding, B. T.; Griffin, R. J.; Jewsbury, P.; Menyerol,

- J.; Mesguiche, V.; Newell, D. R.; Noble, M. E. M.; Pratt, D. J.; Wang, L. Z.; Whitfield, H. J. *J. Med. Chem.* **2004**, *47*, 3710-3722.
- [11] Domingo, L. R.; Aurell, M. J.; Pérez, P.; Contreras, R. *Tetrahedron* **2002**, *58*, 4417-4423.
- [12] Pérez, P.; Simón-Manso, Y.; Aizman, A.; Fuentealba, P.; Contreras, R. *J. Am. Chem. Soc.* **2000**, *122*, 4756-4762.
- [13] Maynard, A. T.; Huang, M.; Rice, W. G.; Covell, D. G. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 11578-11583.
- [14] Santos, J. C.; Contreras, R.; Chamorro, E.; Fuentealba, P. *J. Chem. Phys.* **2002**, *116*, 4311-4316.
- [15] Rivas, P.; Zapata-Torres, G.; Melin, J.; Contreras, R. *Tetrahedron* **2004**, *60*, 4189-4196.
- [16] Domingo, L. R.; Aurell, M. J.; Pérez, P.; Contreras, R. *J. Phys. Chem. A* **2002**, *106*, 6871-6875.
- [17] Chattaraj, P. K.; Poddar, A. *J. Phys. Chem. A* **1999**, *103*, 8691-8699.
- [18] Parr, R. G.; Yang, W. *J. Am. Chem. Soc.* **1984**, *106*, 4591.
- [19] Yang, W.; Parr, R. G. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 6723.
- [20] Murray, J. S.; Sen, K. *Theoretical and Computational Chemistry*, vol. 3, Elsevier Science B. V., Amsterdam, 1996.
- [21] Davies, T. G.; Bentley, J.; Arris, C. E.; Boyle, F. T.; Curtin, N. J.; Endicott, J. A.; Gibson, A. E.; Golding, B. T.; Griffin, R. J.; Hardcastle, I. R.; Jewsbury, P.; Johnson, L. N.; Mesguiche, V.; Newell, D. R.; Noble, M. E. M.; Tucker, J. A.; Wang, L.; Whitfield, H. J. *Nature. Struct. Biol.* **2002**, *9*, 745-749.

[22] Lee, C.; Yang, W. and Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

[23] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

[24] Gaussian 98 (Revision A.11), M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, P. Salvador, J. J. Dannenberg, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, **2001**.

[25] Contreras, R.; Fuentealba, P.; Galván, M.; Pérez, P. *Chem. Phys. Lett.* **1999**, *304*, 405-413.

[26] Alzate-Morales, J. H.; Contreras, R.; Soriano, A.; Tuñón, I.; Silla, E. *Biophys. J.*, To appear in Vol. 92 (2), January, **2007**, pp. 1-10.

[27] Bonaccorsi, R.; Scrocco, E.; Tomasi, J. *J. Chem. Phys.* **1970**, *52*, 5270.

[28] Parr, R. G.; Yang, W. *Density Functional Theory of atoms and molecules*. Oxford University Press: New York, **1989**.

[29] Pearson R. G. *J. Chem. Ed.* **1987**, *64*, 561-567.

[30] Gray, N. S.; Wodicka, L.; Thunnissen, A. -M. W. H.; Norman, T. C.; Kwon, S.; Espinoza, F. H.; Morgan, D. O.; Barnes, G.; LeClerc, S.; Meijer, L.; Kim, S. -H.; Lockhart, D. J.; Schultz, P. G. *Science* **1998**, *281*, 533-538.

[31] Simón-Manso, Y.; Fuentealba, P. *J. Phys. Chem. A* **1998**, *102*, 2029-2032.

## CONCLUSIONES GENERALES Y PROYECCIONES FUTURAS

Las energías de interacción de puente de hidrógeno determinadas con el método ONIOM nos permiten concluir que; las interacciones específicas realizadas entre el anillo purina de cada inhibidor y el aminoácido Leu83 son de gran importancia en la afinidad global de cada ligando por el sitio activo de la enzima CDK2. Estas interacciones aportan la mayor parte de la energía de interacción entre los ligandos estudiados y el modelo de la región bisagra de la CDK2.

De acuerdo a los resultados obtenidos con el método QM/MM, aunque la contribución mayoritaria a la energía de interacción entre los diferentes ligandos y el sitio activo de la CDK2 sea la energía de vdW, hemos observado que la formación de enlaces de hidrógeno es la responsable de un incremento en la actividad biológica. Este último sería el aspecto más importante a tener en cuenta al momento de diseñar un nuevo inhibidor de CDK2 con la estructura O<sup>6</sup>-ciclohexil-metilpurina como núcleo básico.

Las interacciones de puente de hidrógeno del compuesto Cp3 con los aminoácidos Lys89 e Ile10, obtenidas a partir de los métodos híbridos QM/MM, no han sido reportadas antes. Estas nuevas interacciones sugieren que la "superficie específica" situada en el sitio activo de la CDK2 debe ser explorada con más detalle para conseguir inhibidores más selectivos con respecto a los otros subtipos de CDK's.

Los resultados obtenidos a partir de los índices globales y regionales de reactividad sugieren que en general el sitio activo de la CDK2 es blando debido a una respuesta electrónica concertada a la formación de múltiples enlaces de hidrógeno.

Los índices globales de reactividad obtenidos con la DOS, especialmente la blandura y la electrofilia, se pueden utilizar para predecir la afinidad, por el sitio activo de la CDK2, de nuevos inhibidores tipo purínicos.

Los índices globales de reactividad obtenidos a partir de la aproximación del orbital molecular frontera (FMO) no son la elección correcta para describir el patrón de reactividad (en nuestro caso la afinidad por un modelo del sitio activo de la CDK2) de moléculas orgánicas de gran tamaño, debido a que se desprecian las contribuciones de estados energéticos cercanos al último orbital ocupado y que podrían ser importantes en la reactividad del sistema.

Con las estructuras del modelo de la región bisagra y los inhibidores, obtenidos a través de los estudios QM/MM, es posible calcular descriptores basados en la DFT que predicen correctamente la afinidad de una serie de moléculas por el sitio activo de la CDK2 y a un costo computacional razonable.

Un paso siguiente en el estudio de la afinidad de inhibidores de CDK por el sitio activo del ATP, sería validar los descriptores obtenidos con las diferentes aproximaciones y modelos utilizados. Esto se haría empleando otros grupos de inhibidores y sistemas enzimáticos de la familia de las CDK's, con el fin de evaluar la capacidad de los descriptores y de los modelos utilizados en predecir la afinidad en otro grupo de moléculas y sistemas diferentes.

Actualmente estamos aplicando la metodología QM/MM (la energía de interacción, la descomposición de energía y la contribución energética de cada aminoácido) al estudio de la selectividad de una serie de inhibidores de fosfodiesterasas, utilizados desde hace algunos años en el tratamiento de la disfunción eréctil, por los subtipos PDE5 y PDE6. Estos estudios nos permitirán predecir cuales son las determinantes de la selectividad de los diferentes inhibidores por los subtipos de la enzima.

Recientemente hemos abordado el estudio de la catálisis de la transferencia del fosfato gamma del ATP a un sustrato específico, mediada por la CDK2. Para ello utilizamos los métodos QM/MM para trazar los diferentes caminos de reacción que pueden existir a lo largo de la superficie de energía potencial (PES) y determinar cual es el más viable energéticamente. Con ello contribuiremos a esclarecer cual de los dos mecanismos propuestos actualmente en la literatura es el más favorecido.

## APÉNDICE 1

### A.1.1. Introducción

Muchos de los sistemas de interés en medicina y biología, que los químicos medicinales y los bioquímicos les interesa abordar en sus investigaciones en modelación molecular, son generalmente demasiado grandes y complejos para ser estudiados por medio de la química cuántica tradicional. Los métodos mecánico cuánticos resuelven una ecuación electrónica de Schrödinger en un campo clásico de núcleos fijos, lo cual alivia significativamente el problema computacional, y el problema pasa a depender únicamente del número de electrones en el sistema. Pero aún cuando se podrían ignorar algunos electrones (como por ejemplo en los métodos semi-empíricos que incluyen únicamente los electrones de valencia) para estudiar nuestros sistemas enzimáticos, el problema sigue considerando un gran número de electrones y los cálculos computacionales se hacen muy costosos. Para solucionar en gran medida este problema se recurre generalmente a la utilización de campos de fuerza moleculares (más conocidos como mecánica molecular (MM)) que ignoran los movimientos electrónicos y calculan la energía de un sistema solamente en función de la posición de los núcleos. Así, la mecánica molecular es invariablemente utilizada para realizar cálculos en sistemas que contienen un número significativo de átomos. En algunos casos los campos de fuerza pueden ofrecer respuestas que son tan exactas como las que ofrecen los cálculos mecánico-cuánticos del más alto nivel, y solo en una fracción de tiempo computacional. Sin embargo, la mecánica molecular no puede ofrecer, por supuesto, respuestas claras sobre las propiedades que dependen de la distribución electrónica en una molécula.

La MM funciona asumiendo la validez de varias aproximaciones. La primera es la aproximación de Born-Oppenheimer, sin la cual sería imposible escribir la energía en función de las coordenadas nucleares. La segunda es el hecho de que la MM está basada

en modelos simples de las interacciones que ocurren en un sistema, con contribuciones de procesos tales como el estiramiento de los enlaces, la abertura o cierre de los ángulos y las rotaciones sobre enlaces simples (ángulos diedros). Por último, la transferibilidad de las propiedades es un atributo clave de los campos de fuerza, ya que ésta permite que un juego de parámetros que han sido desarrollado y probados en un número relativamente pequeño de casos, pueda ser aplicado a un rango más amplio de problemas. Así por ejemplo, los parámetros desarrollados a partir de datos en moléculas pequeñas pueden ser utilizados para estudiar moléculas muchos más grandes como por ejemplo polímeros.

En los siguientes apartados de este apéndice trataremos de definir simplídicamente cada uno de los componentes de la energía total de un campo de fuerza y más adelante introduciremos brevemente los métodos de cálculo híbridos QM/MM (Quantum Mechanics/Molecular Mechanics), los cuales son bastante útiles actualmente en el estudio de reacciones en medio enzimático y/o acuoso, ya que permiten tratar una pequeña parte del sistema de interés con algunos de los métodos cuánticos disponibles y el resto del sistema con la MM.

### **A.1.2. La energía como función de parámetros de geometría internos.**<sup>1,2</sup>

Muchos de los campos de fuerza utilizados actualmente pueden ser interpretados en términos de un cuadro relativamente simple, que consta de 4 componentes que describen las fuerzas intra e intermoleculares dentro del sistema. Así se pueden asignar "penalties" de energía que están asociados con las desviaciones de los valores de referencia o equilibrio de los enlaces y de los ángulos. También hay una función que describe como cambia la energía cuando se rotan los enlaces, y finalmente el campo de fuerza contiene términos que describen la interacción entre partes no enlazadas del sistema. En la Figura A.1.1 se presentan en forma esquemática las fuerzas que forman un campo de fuerza típico. En el desarrollo de esta tesis hemos empleado la versión del campo de fuerza OPLS-all-atom (OPLS-AA) que ha sido desarrollado por W. Jorgensen

y colaboradores.<sup>3</sup> Este es un campo de fuerza del tipo más simple que ha sido aplicado a un amplio rango de sistemas, incluyendo biomacromoléculas, tales como proteínas y ácidos nucleicos, líquidos orgánicos y solutos en solución. Existen campos de fuerza más sofisticados que incluyen términos adicionales en la función de energía, pero todos ellos contienen básicamente los cuatro componentes descritos anteriormente. La función de energía se puede dividir entonces en dos categorías de términos que describen las

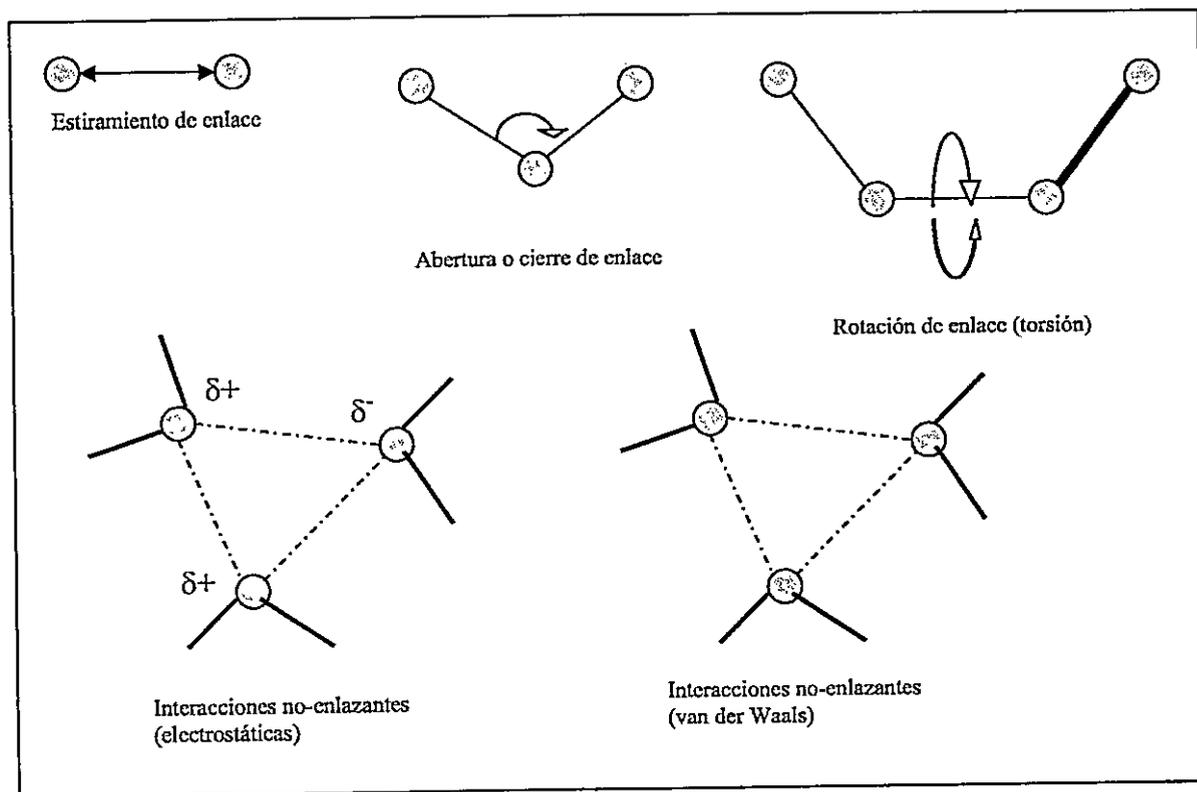


Figura A.1.1. Representación esquemática de las cuatro interacciones clave en un campo de fuerza molecular.

interacciones enlazantes y no-enlazantes entre los átomos y que se explican a continuación.

### A.1.2.1. Términos enlazantes de la función de energía.

Los términos de energía enlazante son aquellos que ayudan a definir la estructura enlazante o covalente de una molécula, por ejemplo, su forma local. En un campo de

fuerza típico, y simple, la energía covalente o de enlace,  $V_{cov}$ , consistirá de una suma para el enlace, el ángulo de enlace, el ángulo diedro (o ángulo de torsión) y la distorsión fuera del plano (o diedro impropio), así:

$$V_{cov} = V_{enlace} + V_{ángulo} + V_{diedro} + V_{impropio} \quad (\text{A.1.1})$$

#### A.1.2.1.1. Energía de enlace

La energía de enlace toma frecuentemente la forma de un armónico:

$$V_{enlace} = \sum_{enlaces} \frac{1}{2} k_b (b - b_0)^2 \quad (\text{A.1.2})$$

donde  $k_b$  es la constante de fuerza para el enlace,  $b$  es la distancia del enlace en la estructura entre los dos átomos que definen el enlace y  $b_0$  es la distancia de equilibrio para el enlace. La suma se realiza sobre todos los enlaces que han sido definidos en el sistema. Debido a que este tipo de energía tiene forma de armónico, veremos que la energía del enlace aumentará continuamente, sin límite, a medida que este es distorsionado de su valor de equilibrio,  $b_0$ .

#### A.1.2.1.2. Energía de ángulo de enlace.

El término de energía de ángulo está diseñado para imitar como cambia la energía de un ángulo de enlace cuando este es distorsionado lejos de su posición de equilibrio. Al igual que el término de energía de enlace, este se toma frecuentemente como una forma de armónico:

$$V_{ángulo} = \sum_{ángulos} \frac{1}{2} k_\theta (\theta - \theta_0)^2 \quad (\text{A.1.3})$$

Los parámetros extra son similares a aquellos de la energía de enlace, así  $k_\theta$  es la constante de fuerza para el ángulo y  $\theta_0$  es su valor de equilibrio. La suma corre sobre todos los ángulos definidos para el sistema y cada ángulo es definido en la misma forma como se muestra en la Figura A.1.1.

#### A.1.2.1.3. Energía asociada a ángulos diedros.

El tercer tipo de término enlazante es el que describe como cambia la energía de una molécula a medida que esta experimenta una rotación sobre uno de sus enlaces definidos en un plano específico, o sea, la energía dependiente de un ángulo diedro o de torsión para el sistema. Al contrario que los términos de energía de enlace y de ángulo de enlace, la forma del armónico para definir una energía asociada al ángulo diedro no es usualmente la más apropiada. Lo anterior se debe a que, para muchos ángulos diedros en moléculas, el rango entero de ángulos desde  $0^\circ$  a  $360^\circ$  puede ser accesible sin causar grandes diferencias en su energía. Tales efectos se pueden reproducir con una función periódica que sea continua en todo el rango completo de ángulos posibles. La energía de un ángulo diedro puede ser entonces escrita como:

$$V_{\text{diedro}} = \sum_{\text{diedros}} \frac{1}{2} V_n [1 + \cos(n\phi - \delta)] \quad (\text{A.1.4})$$

La suma es sobre todos los ángulos diedros definidos en el sistema. En la expresión (A.1.4),  $n$  es la periodicidad del ángulo (la cuál determina cuantos picos y pozos hay en el potencial),  $\delta$  es la "fase" del ángulo y  $V_n$  es la constante de fuerza. Frecuentemente el valor de  $\delta$  es restringido a tomar valores de  $0^\circ$  o  $180^\circ$ , en cuyo caso lo único que cambiará será el signo del término coseno en la expresión anterior. Se debe tener en cuenta que la periodicidad de cada término en la suma cambia dependiendo de tipo de ángulo diedro. Así en el campo de fuerza OPLS-AA para cada ángulo diedro se

incluyen tres términos, en la expansión de Fourier, con periodicidades 1, 2 y 3. La fase para cada uno de los términos es  $0^\circ$  para periodicidades de 1 y 3 y  $180^\circ$  para periodicidades de valor 2. Esto conduce a la siguiente expresión en la cual ninguno de los coeficientes de Fourier,  $V_n$ , es cero:

$$V_{\text{diedro}} = \sum_{\text{diedros}} \frac{V_1}{2}(1 + \cos \phi) + \frac{V_2}{2}(1 - \cos 2\phi) + \frac{V_3}{2}(1 + \cos 3\phi) \quad (\text{A.1.5})$$

#### A.1.2.1.4. Energía asociada a un ángulo impropio

El cuarto término en la suma de la Ecuación (A.1.1) es un poco más complicado en su definición que los que ya han sido discutidos, y describe la energía de los movimientos fuera del plano. Este término es necesario frecuentemente para grupos planares, tales como los carbonos híbridos  $sp^2$  en los grupos carbonilos y en los sistemas aromáticos, ya que se ha encontrado que el uso de los términos de energía asociada al ángulo diedro, no es suficiente para mantener la planaridad de estos grupos durante los cálculos computacionales. Una manera común de evitar este tipo de problema es definir un *ángulo diedro impropio*, el cual difiere del ángulo diedro propio en que los átomos que definen el ángulo diedro,  $i-j-k-l$ , no están directamente unidos uno con otro (Ver Figura A.1.2). Con esta definición de un ángulo diedro impropio, el cual denotaremos como  $\omega$ , algunos campos de fuerza utilizan la misma forma para este tipo de energía tal cual se ha definido en la Ecuación (A.1.4), así:

$$V_{\text{impropio}} = \sum_{\text{impropios}} \frac{1}{2} V_n [1 + \cos(n\omega - \delta)] \quad (\text{A.1.6})$$

En el campo de fuerza OPLS-AA, la energía de ángulo impropio tiene la misma forma que la del ángulo diedro (Ecuación (A.1.5)), excepto que los ángulos diedros,  $\phi$ , son reemplazados por los impropios,  $\omega$ .

Los cuatro términos definidos anteriormente son solo los términos de energía enlazante y que consideraremos en los estudios realizados en esta tesis, aunque se pueden encontrar otro tipo de términos en algunos de los tantos campos de fuerza disponibles actualmente. En general, los términos extras son adicionados al campo de fuerza con el fin de obtener una mejor concordancia con los datos experimentales (especialmente espectros vibracionales), pero también es cierto que estos aumentan el grado de complejidad del campo de fuerza y el número de parámetros que se necesitan obtener. Los términos extra que algunas veces se incluyen son, por ejemplo, términos de energía de enlace y ángulo de enlace de la misma forma de los presentados en las ecuaciones 2 y 3, excepto que los términos ya no son armónicos sino que son reemplazados por términos lineales, cúbicos y de cuarto grado.

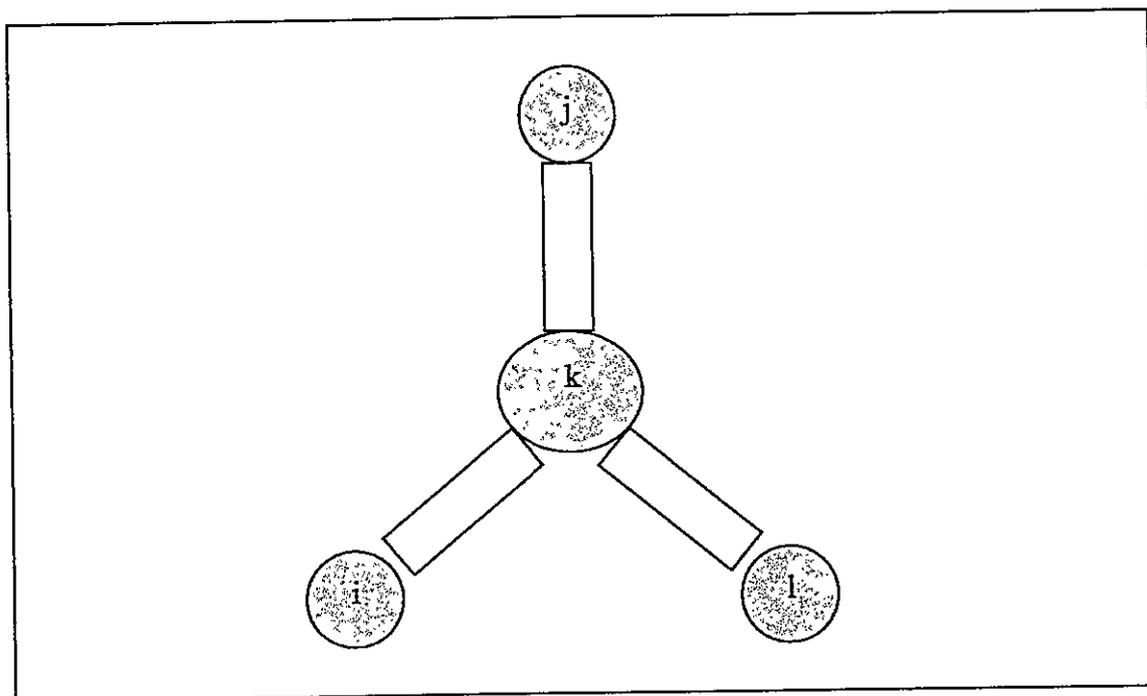


Figura A.1.2. Arreglo de átomos en la definición de un ángulo diedro impropio.

Otro tipo de términos que también se pueden observar son los *términos cruzados* que unen distorsiones en coordenadas internas diferentes. Por ejemplo, un término cruzado enlace/ángulo podría ser proporcional a  $(b-b_0)(\theta - \theta_0)$ .

### A.1.2.2. Términos no-enlazantes de la función de energía.

Los términos de energía no-enlazante describen las interacciones entre los átomos de diferentes moléculas o entre los átomos que no están directamente enlazados en una misma molécula. Estas interacciones ayudan a determinar la conformación global de un sistema molecular.

Las interacciones no-enlazantes son utilizadas para imitar las interacciones que surgen entre las distribuciones electrónicas que rodean diferentes átomos. Así, a corto alcance, las interacciones son principalmente repulsivas debido a la interacción entre las nubes electrónicas y a un efecto puramente mecánico cuántico llamado *repulsión de intercambio*, el cual resulta cuando dos nubes de electrones son empujadas una contra otra. A larga distancia existen diferentes clases de interacción. La primera corresponde a las interacciones *electrostáticas* que se presentan a partir de la interacción de las distribuciones de carga (incluyendo el núcleo) alrededor de cada molécula o parte de una molécula. La segunda son las interacciones de *dispersión*, las cuales se producen por la distorsión de las distribuciones de carga de dos grupos. Finalmente, existen interacciones *inducidas* o de *polarización*, que son causadas por las distorsiones de la distribución de carga de una molécula a medida que ésta interactúa con grupos vecinos. Los términos no-enlazantes en un campo de fuerza intentan reproducir todo este tipo de interacciones descritas anteriormente. Aquí consideramos una energía no-enlazante compuesta por tres términos:

$$V_{ne} = V_{elect} + V_{LJ} + V_{polar} \quad (\text{A.1.7})$$

#### A.1.2.2.1. Energía electrostática

La energía electrostática,  $v_{elect}$ , imita la energía que proviene de las interacciones electrostáticas entre dos distribuciones de carga. En una molécula, el núcleo cargado positivamente estará rodeado por una nube de electrones cargados negativamente. Así, en los campos de fuerza, se deben elegir modelos que puedan representar la distribución

de carga alrededor de una molécula. La representación más simple es asignar una carga fraccionada para cada átomo, la cual representa su carga neta total. La energía electrostática se calcula como:

$$V_{elect} = \frac{1}{4\pi \epsilon_0 \epsilon} \sum_{\text{pares } ij} \frac{q_i q_j}{r_{ij}} \quad (\text{A.1.8})$$

donde  $q_i$  y  $q_j$  son las cargas fraccionadas en los átomos  $i$  y  $j$ , y  $r_{ij}$  es la distancia entre las dos partículas. Los términos en el prefactor  $1/(4\pi \epsilon_0)$ , es el término estándar cuando se calculan interacciones electrostáticas en el sistema de unidades MKSA (metro, kilogramo, segundo, amperio), y  $\epsilon$ , es la constante dieléctrica que tendrá un valor igual a 1 cuando el sistema está en el vacío. La suma en la Ecuación (A.1.8) corre sobre todos los pares de átomos para los cuales se quiera calcular las interacciones electrostáticas. Hay que tener en cuenta que las cargas fraccionadas sobre los átomos son constantes y por lo tanto no cambian durante el cálculo.

#### A.1.2.2.2. Energía de Lennard-Jones

El segundo término en la Ecuación (A.1.7) es la energía de Lennard-Jones, la cual describe las interacciones de dispersión a largo alcance y las interacciones repulsivas a corto alcance. Esta tiene la siguiente forma:

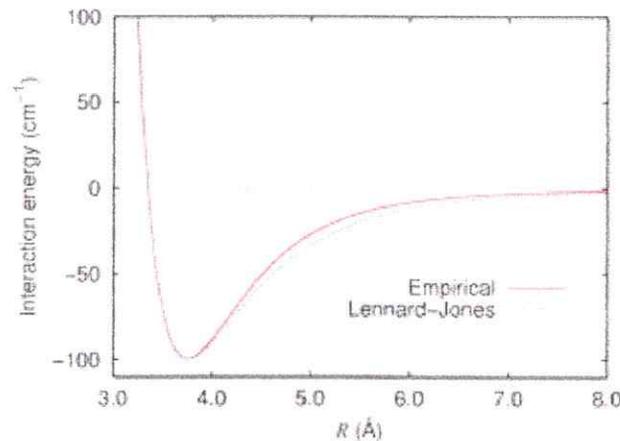
$$V_{LJ} = \sum_{\text{pares } ij} \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \quad (\text{A.1.9})$$

donde  $A_{ij}$  y  $B_{ij}$  son constantes positivas cuyos valores dependen de los tipos de átomos,  $i$  y  $j$ , y la suma se realiza sobre todos los pares de átomos para los cuales se quiere calcular este tipo de interacción. La forma del potencial de Lennard-Jones se muestra en

la Figura A.1.3. La parte repulsiva de la curva es representada por el término  $1/r_{ij}^{12}$  y la parte atractiva por el término  $1/r_{ij}^6$ . Para completar la forma de la energía de Lennard-Jones es necesario tener a mano una expresión empírica para determinar los parámetros  $A_{ij}$  y  $B_{ij}$ . Estos son usualmente definidos en términos de la profundidad del pozo de la curva de Lennard-Jones,  $\epsilon_{ij}$ , o también la distancia a la cual la energía de interacción es cero,  $s_{ij}$ , o la posición de la parte inferior de el pozo,  $\sigma_{ij}$ . Con esos parámetros la interacción Lennard-Jones entre los átomos  $i$  y  $j$ ,  $v_{LJ}^{ij}$ , toma la siguiente forma:

$$v_{LJ}^{ij} = 4\epsilon_{ij} \left[ \left( \frac{s_{ij}}{r_{ij}} \right)^{12} - \left( \frac{s_{ij}}{r_{ij}} \right)^6 \right] = \epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - 2 \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \quad (\text{A.1.10})$$

donde  $\sigma_{ij}^6 = 2s_{ij}^6$ .



**Figura A.1.3.** Energía potencial de Lennard-Jones para un par de átomos de argón (Tomado de [http://en.wikipedia.org/wiki/Lennard-Jones\\_potential](http://en.wikipedia.org/wiki/Lennard-Jones_potential)).

Aunque cada uno de los parámetros,  $\epsilon_{ij}$ ,  $s_{ij}$ , o  $\sigma_{ij}$ , depende formalmente de dos átomos, es normal especificar un juego de *combinación de reglas* con el fin de que estos puedan ser definidos a partir de parámetros para un solo átomo. Es usual utilizar la media geométrica como una regla de combinación para la profundidad del pozo:

$$\varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}} \quad (\text{A.1.11})$$

Para los parámetros de distancia son comunes tanto la combinación de reglas de la media geométrica como la aritmética. Para el parámetro  $\sigma_{ij}$ , por ejemplo, podemos tener bien sea,

$$\sigma_{ij} = \sqrt{\sigma_{ii}\sigma_{jj}} \quad (\text{A.1.12})$$

ó

$$\sigma_{ij} = \frac{1}{2}(\sigma_{ii} + \sigma_{jj}) \quad (\text{A.1.13})$$

El tercer término de energía no-enlazante que se considera frecuentemente es la energía de polarización. En contraste con los dos términos previos, las energías electrostática y de Lennard-Jones, este término de energía no es un término estándar en muchos campos de fuerza y por lo tanto no será usado en los cálculos realizados en esta tesis y también su explicación en este apéndice será obviada.

Por lo tanto, en el campo de fuerza OPLS-AA utilizado en esta tesis, la energía no-enlazante,  $v_{ne}$ , esta formado sólo por la suma de los términos electrostático y de Lennard-Jones descritos anteriormente, así:

$$v_{ne} = \sum_{\text{pares } ij} \left\{ \frac{q_i q_j}{4\pi \epsilon_0 \epsilon r_{ij}} + 4\varepsilon_{ij} \left[ \left( \frac{s_{ij}}{r_{ij}} \right)^{12} - \left( \frac{s_{ij}}{r_{ij}} \right)^6 \right] \right\} f_{ij} \quad (\text{A.1.14})$$

La variable  $f_{ij}$  es un factor de peso para las interacciones. Para las interacciones 1-2 y 1-3 (se explican en el siguiente apartado) tiene un valor de 0 y así estas interacciones son excluidas de la suma. Para las interacciones 1-4 el valor es  $\frac{1}{2}$  y para todas las otras interacciones su valor es 1.

### A.1.2.2.3. Otros aspectos a considerar

Existe un punto adicional que se necesita resolver cuando se quiere calcular la energía electrostática, de Lennard-Jones y de polarización. Este tiene que ver con las interacciones entre partículas que se deben incluir en la suma para las energías electrostática y de Lennard-Jones (Ecuaciones (A.1.8) y (A.1.10), respectivamente) o para el cálculo de los campos en el caso de la energía de polarización.

Existen dos problemas para las partículas (átomos) que están unidas o que están separadas sólo por unos pocos enlaces. Primero, las interacciones no-enlazantes entre ellas son grandes debido a que su separación interpartícula es pequeña y, segundo, también existirán términos de enlace (enlace, ángulo, diedros, etc.) entre tales átomos.

El dilema se puede resolver introduciendo el concepto de *exclusiones no-enlazantes* (Ver Figura A.1.4). Las interacciones no-enlazantes se calculan solo para las partículas que no están involucradas en interacciones enlazantes directas.

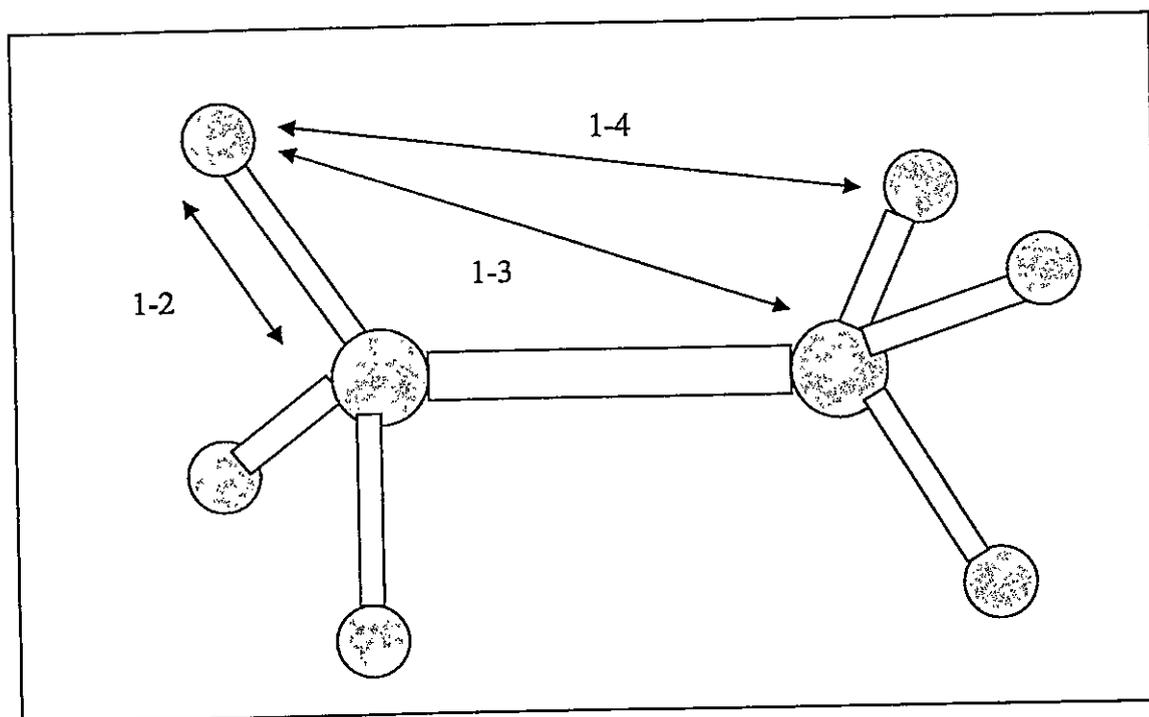


Figura A.1.4. Ejemplos de exclusiones no-enlazantes 1-2, 1-3 y 1-4 para un átomo de hidrógeno en una molécula de etano.

Para las partículas que estén enlazadas o separadas por sólo unos pocos enlaces, las interacciones no-enlazantes entre ellas no son calculadas y son los términos enlazantes los que determinarán su energía de interacción. Esto evita el problema de obtener energías de interacción muy grandes y también evita sobrecontar las interacciones, lo cual llevaría a incluir en el cálculo ambos tipos de interacción.

El número y tipo de exclusiones no-enlazantes utilizadas depende del campo de fuerza. Es típico excluir las interacciones entre átomos que están enlazados directamente (lo que es llamado interacciones 1-2) y también aquellas entre átomos que están separados por dos enlaces (interacciones 1-3). El tratamiento de las interacciones entre átomos separados por tres enlaces (interacciones 1-4) es el más variable. En algunos campos de fuerza estas son excluidas, y en otros son incluidas pero además las interacciones son escaladas utilizando, un factor especial, un juego especial de cargas 1-4 o parámetros de Lennard-Jones. El motivo para tratar de forma diferente las interacciones 1-4 es que es la combinación de los términos enlazantes de ángulo diedro y los términos 1-4 electrostáticos y de Lennard-Jones la que determina las barreras de rotación sobre los enlaces.

Para cerrar este apartado es importante mencionar que, como es el caso para los términos de energía enlazante, se pueden utilizar otro tipo de términos no-enlazantes según el campo de fuerza. Un ejemplo importante incluido en algunos campos de fuerza es la energía introducida para modelar el *enlace de hidrógeno*. Se han utilizado varias aproximaciones para este tipo de energía, pero este término es frecuentemente omitido en los campos de fuerza recientes ya que la experiencia ha demostrado que la combinación de las interacciones electrostáticas y de repulsión/dispersión es suficiente para reproducir las interacciones de enlace de hidrógeno con una precisión razonable.

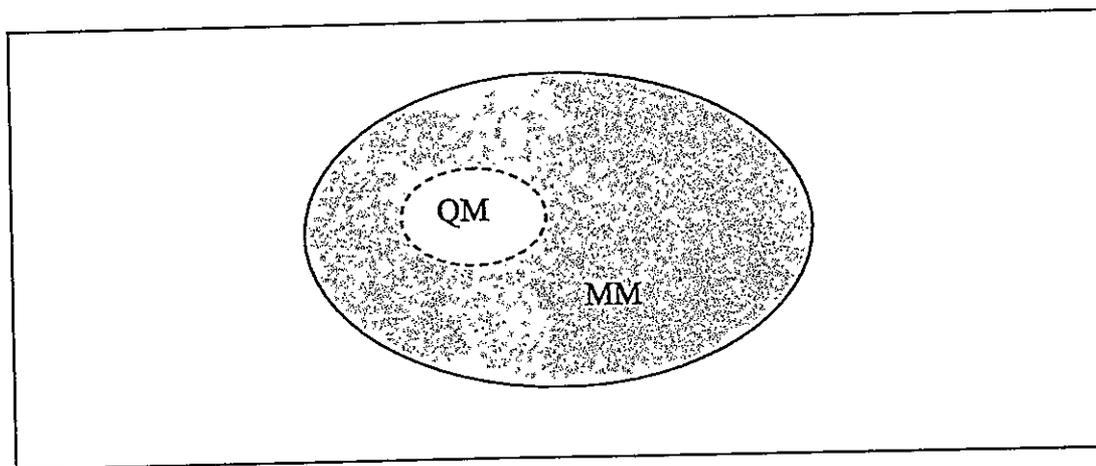
### A.1.3. Métodos híbridos de cálculo QM/MM <sup>1</sup>

La descripción mecano-cuántica de una molécula pasa por la resolución de la ecuación de Schrödinger. Sin embargo, hasta el momento únicamente se ha conseguido resolver exactamente esta ecuación para un problema de dos partículas no relativista. En el resto

de sistemas debemos utilizar métodos que proporcionen soluciones aproximadas. Generalmente, cuanto mayor sea el tamaño del sistema los métodos que se aplican son menos rigurosos, ya que al requerir mayores recursos computacionales se introducen más aproximaciones.<sup>4, 5</sup>

Para sistemas con un gran número de átomos, como pueden ser las enzimas, se encuentra el problema de no poder tratar todo el sistema mediante la mecánica cuántica (QM). Por otro lado la mecánica molecular (MM) no permite estudiar procesos que implican una reacción química, ya que no describe electrones. La metodología híbrida que combina la mecánica cuántica y la mecánica molecular (QM/MM) utiliza la mecánica cuántica (QM), para describir la parte del sistema donde se producen los procesos de ruptura y formación de enlaces para un número limitado de átomos, mientras que el resto del sistema se describe con potenciales clásicos (descritos en el apartado anterior) de la mecánica molecular, incluyendo así el efecto del medio.<sup>6, 7</sup>

En consecuencia, el sistema a estudiar se divide en dos regiones como se representa en la Figura A.1.5.



**Figura A.1.5.** Representación esquemática de la división de un sistema molecular en sus zonas QM y MM, es decir, los dos subsistemas presentes en los métodos QM/MM.

La energía potencial total del sistema será la suma de la energía de la parte cuántica, la energía de la parte clásica y la energía de interacción entre ambos subsistemas, así:

$$E_{TOT} = E_{QM} + E_{MM} + E_{QM/MM} \quad (\text{A.1.15})$$

Los átomos de la región QM se representan como electrones y núcleos. El Hamiltoniano asociado a esta región, en la aproximación de Born-Oppenheimer, se escribe:

$$H_{QM} = -\frac{1}{2} \sum_i \nabla_i^2 + \sum_{ij} \frac{1}{r_{ij}} - \sum_{i\alpha} \frac{Z_\alpha}{r_{i\alpha}} + \sum_{\alpha\beta} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}} \quad (\text{A.1.16})$$

donde  $i$  y  $j$  representan las coordenadas electrónicas,  $\alpha$  y  $\beta$  las coordenadas nucleares,  $r$  la distancia electrón-electrón o electrón-núcleo,  $R$  las distancias núcleo-núcleo,  $-\nabla_i^2/2$  es el operador de energía cinética electrónica y  $Z$  es la carga nuclear.

La zona MM contiene el resto de los átomos del sistema. Éstos se describen clásicamente, y sus interacciones se determinan mediante funciones de energía potencial empíricas, de modo que dichos átomos no pueden estar implicados en procesos de formación o ruptura de enlaces, ni de transferencia de carga.

Por último, el término  $E_{QM/MM}$  describe como interaccionan los átomos QM con los centros MM. El Hamiltoniano, en unidades atómicas, viene dado por la siguiente ecuación:

$$H_{QM/MM} = -\sum_i \sum_M \frac{q_M}{r_{iM}} + \sum_\alpha \sum_M \frac{Z_\alpha q_M}{R_{\alpha M}} + \sum_\alpha \sum_M \left( \frac{A_{\alpha M}}{R_{\alpha M}^{12}} - \frac{C_{\alpha M}}{R_{\alpha M}^6} \right) \quad (\text{A.1.17})$$

donde los subíndices  $i$ , hacen referencia a los electrones de los átomos QM, los subíndices  $\alpha$  a los núcleos de los átomos QM y los subíndices  $M$  a los átomos MM. El primer término de la Ecuación (A.1.17) representa las interacciones electrostáticas entre los átomos MM y los electrones de los átomos QM. El segundo término representa las interacciones electrostáticas entre los átomos MM y los núcleos de los átomos QM, y el tercer término describe las interacciones de van der Waals entre los centros MM y QM.

Los dos últimos términos de esta ecuación no incluyen coordenadas electrónicas, de modo que pueden ser calculados y sumados directamente a la energía total. Sin embargo, el primer término sí que incluye coordenadas electrónicas por lo que debe tratarse en el procedimiento autoconsistente.

Los métodos híbridos QM/MM no están, por supuesto, restringidos al estudio de reacciones, sino que además se pueden utilizar para estudiar procesos de asociación y transiciones conformacionales. La mayoría de las implementaciones utilizan un modelo de dos subsistemas como se describió anteriormente, pero Morokuma y colaboradores han descrito una aproximación que utiliza varios subsistemas o capas y que han llamado ONIOM.<sup>8</sup> ONIOM es particularmente un nombre apropiado dado que el cálculo típico se construye a partir de una serie de "capas". Una característica particular de ONIOM y sus métodos relacionados es que puede proporcionar gradientes rigurosos y segundas derivadas, posibilitando de esta forma que se puedan calcular propiedades como las frecuencias vibracionales.<sup>9</sup>

#### A.1.4. Bibliografía

- [1] Leach, Andrew R. *Empirical Force Fields Models : Molecular Mechanics*. In: *Molecular Modelling: Principles and Applications*. Second Edition. Great Britain, Pearson Education Limited, 2001. pp. 165-252.
- [2] Field, Martin J. *The energy function*. In: *A Practical Introduction to the Simulation of Molecular Systems*. United Kingdom, Cambridge University Press, 1999. pp. 43-66.
- [3] Jorgensen, W.L., Maxwell, D.S., Tirado-Rives, J. Development and testing of the PLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* 1996, 118, 11225-11236.
- [4] Minkin, V. I., Simkin, B. Y. y Minyaev, R. M. *Quantum Chemistry of Organic Compounds*, New York, 1990.
- [5] Levine, R. D. *Quantum Mechanics of Molecular Rate Processes*, Oxford, 1969.
- [6] Kollman, P. A.; Kuhn, B., Donini, O., Peräkylä, M., Stanton, R. V. y Bakowies, D. *Acc. Chem. Res.* 2001, 34, 72-79.
- [7] Hammes-Schiffer, S. *J. Chem. Phys.* 1996, 105, 2236-2246.
- [8] Svensson, M., Humbel, S., Froese, R. D. J., Matsubara, T., Sieber, S., Morokuma, K. *J. Phys. Chem.* 1996, 100, 19357-19363.
- [9] Dapprich, S., Komiroi, I., Byun, K. S., Morokuma, K., Frisch, M. J. *THEOCHEM* 1999, 461-462, 1-21.

## APÉNDICE 2

### A.2.1. Introducción

La Teoría de los Funcionales de la Densidad (Density Functional Theory (DFT)) se ha convertido, desde su formulación en el principio de los años 80 por Parr y Yang<sup>1</sup>, en una de las herramientas teóricas más útiles y explotadas por químicos teóricos, químicos orgánicos e inorgánicos, e incluso bioquímicos. Esto ha sido posible gracias a que esta teoría proporciona lo que se ha llamado “precisión química” cuando se quiere obtener información acerca de la energía, estructura y propiedades de sistemas atómicos y moleculares. Además es importante resaltar que los métodos derivados de la DFT tienen un costo computacional menor que los métodos basados en la resolución de la función de onda convencional, los cuales dependen de las coordenadas espaciales y de espín de todos los  $N$  electrones del sistema bajo estudio. También esta teoría representa una forma alternativa, a la de los métodos *ab initio* convencionales, de introducir los efectos de la correlación electrónica en la resolución de la ecuación de Schrödinger electrónica. Luego de describir brevemente las bondades de los métodos DFT, vamos a introducir rápidamente las definiciones más relevantes de la teoría, y entre las cuales las más importantes para el desarrollo de esta tesis tienen que ver principalmente con la “DFT conceptual” desarrollada por Parr y Yang. Esta trata de extraer los conceptos y principios químicos más relevantes a partir de la DFT.

### A.2.2. Las bases de la DFT

#### A.2.2.1. Los teoremas de Hohenberg y Kohn.

Basada en los conocidos teoremas de Hohenberg-Kohn<sup>2</sup> (HK), la DFT se centra en la densidad electrónica,  $\rho(r)$  como la portadora de toda la información de un sistema molecular o atómico en su estado fundamental. Así la energía del estado fundamental de

un sistema polielectrónico puede expresarse a partir de la densidad electrónica, y de hecho el uso de la densidad electrónica en lugar de la función de onda para el cálculo de la energía constituye la base fundamental de la DFT.

El primer teorema de HK sostiene que la densidad electrónica,  $\rho(\mathbf{r})$ , determina el potencial externo (debido a los núcleos),  $v(\mathbf{r})$ . También  $\rho(\mathbf{r})$  determina  $N$ , el número total de electrones del sistema, por medio de la normalización al número total de partículas,

$$\int \rho(\mathbf{r}) d(\mathbf{r}) = N \quad (\text{A.2.1})$$

y  $N$  y  $v(\mathbf{r})$  determinan el Hamiltoniano molecular,  $H_{op}$ , que escrito en términos de la aproximación de Born-Oppenheimer e ignorando los efectos relativistas, es

$$H_{op} = -\sum_i^N \frac{1}{2} \nabla_i^2 - \sum_A^n \sum_i^N \frac{Z_A}{r_{iA}} + \sum_{i<j}^N \sum_j^N \frac{1}{r_{ij}} + \sum_{B<A}^n \sum_A^n \frac{Z_A Z_B}{R_{AB}} \quad (\text{A.2.2})$$

Aquí las sumas en  $i$  y  $j$  corren sobre los electrones, y las sumas en  $A$  y  $B$  sobre los núcleos;  $r_{ij}$ ,  $r_{iA}$ , y  $R_{AB}$  significan las distancias electrón-electrón, electrón-núcleo, y núcleo-núcleo, respectivamente. Así, puesto que  $H_{op}$  determina la energía del sistema por medio de la ecuación de Schrödinger,

$$H_{op} \psi = E \psi \quad (\text{A.2.3})$$

siendo  $\psi$  la función de onda electrónica, se puede concluir que finalmente  $\rho(\mathbf{r})$  determina la energía del sistema y todas las demás propiedades electrónicas del estado fundamental. Por consiguiente se puede afirmar que la energía del sistema,  $E$ , es una funcional de la densidad electrónica,  $\rho$ , así:

$$E = E_v[\rho] \quad (\text{A.2.4})$$

El índice “v” se escribe para hacer énfasis en la dependencia en  $v$ .

El segundo teorema de HK proporciona un planteamiento variacional para obtener  $\rho$ , ósea buscar la  $\rho(\mathbf{r})$  minimizando la energía,  $E$ . Para el valor óptimo de  $\rho(\mathbf{r})$ , la energía  $E$  no cambia cuando hay una variación en  $\rho(\mathbf{r})$ , siempre y cuando que  $\rho(\mathbf{r})$  integre todo el tiempo al número total de electrones,  $N$  (Ecuación (A.2.1)):

$$\delta(E - \mu\rho(\mathbf{r})) = 0 \quad (\text{A.2.5})$$

donde  $\mu$  es el multiplicador Lagrangiano correspondiente.

Finalmente se obtiene la ecuación de Euler-Lagrange, que en los métodos DFT, es la análoga a la ecuación de Schrödinger independiente del tiempo (Ecuación (A.2.3)).

$$\mu = v(\mathbf{r}) + \frac{\delta F_{HK}}{\delta \rho(\mathbf{r})} \quad (\text{A.2.6})$$

donde  $F_{HK}$  es la funcional de Hohenberg-Kohn que contiene la funcional de la energía cinética electrónica,  $T[\rho]$ , y la funcional de interacción electrón-electrón,  $V_{ee}[\rho]$ :

$$E_v[\rho] = \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_{HK}[\rho] \quad (\text{A.2.7})$$

siendo,

$$F_{HK}[\rho] = T[\rho] + V_{ee}[\rho] \quad (\text{A.2.8})$$

La Ecuación (A.2.6) nos proporciona una fórmula para minimizar la energía y determinar así la densidad del estado fundamental. El problema que presenta es que no se conoce con exactitud la expresión que relaciona  $F_{HK}$  con la densidad. En particular la forma exacta de  $T[\rho]$  no se conoce con suficiente precisión. Por el contrario, la energía cinética se calcula fácilmente si se conoce  $\psi$ .

### A.2.2.2. Las ecuaciones de Kohn-Sham.

El tratamiento práctico de la Ecuación (A.2.6) fue proporcionado por Kohn y Sham<sup>3</sup> quienes ingeniosamente la transformaron a una forma que muestra una cercana analogía con las ecuaciones de Hartree.<sup>4</sup> Este último aspecto facilitó su implementación en los paquetes de software basados en la función de onda, como por ejemplo Gaussian.<sup>5</sup> Lo anterior se logró introduciendo orbitales, de tal forma que la energía cinética se pudiera calcular de forma simple y con buena exactitud. Finalmente se obtiene la siguiente ecuación:

$$\left( -\frac{1}{2} \nabla^2 + v_{eff}(\mathbf{r}) \right) \Psi_i = \epsilon_i \Psi_i \quad (\text{A.2.9})$$

que está sujeta a la condición de ortonormalidad,

$$\int \Psi_i^*(\mathbf{x}) \Psi_j(\mathbf{x}) d\mathbf{x} = \delta_{ij} \quad (\text{A.2.10})$$

y que corresponde a la ecuación de Kohn-Sham de un solo electrón y puede ser resuelta iterativamente, en forma análoga a las ecuaciones de Hartree o Hartree-Fock. El precio que se tiene que pagar por la incorporación de la correlación electrónica es la aparición del potencial de intercambio y correlación,

$$v_{XC} = \frac{\delta E_{XC}}{\delta \rho(\mathbf{r})} \quad (\text{A.2.11})$$

el cual contiene los componentes de energía restantes en la funcional de HK,  $F_{HK}$ , y que hasta ahora no tiene una forma conocida y por lo tanto no se dispone de una estrategia de cálculo sistemática para mejorar sus resultados. Sin embargo la situación ha cambiado dramáticamente en los últimos años, como han demostrado las amplias pruebas disponibles actualmente para la funcional de intercambio y correlación,  $v_{XC}$  probablemente más popular, llamada B3LYP.<sup>6, 7</sup> Para una comprensión más profunda y actual de los fundamentos y orígenes de la DFT se recomienda ver las revisiones realizadas por Geerlings *et al.*<sup>8</sup> y por Caballol y Solá.<sup>9</sup>

### A.2.3. Conceptos químicos y su definición.

Usando el lenguaje de la DFT es posible definir de forma cuantitativa algunos conceptos clásicos muy ligados a la estructura y reactividad química como son el potencial químico, la dureza, la blandura y las funciones de Fukui, en lo que se denomina la DFT conceptual. Esto ha sido posible gracias al papel central que juegan la densidad electrónica y el número de electrones en esta teoría.

#### A.2.3.1. La electronegatividad y el potencial químico electrónico.

La identificación del multiplicador Lagrangiano,  $\mu$ , en la Ecuación (A.2.6) con el negativo de la electronegatividad,  $\chi$ ,<sup>10</sup>

$$\mu = -\chi = -\left(\frac{\partial E}{\partial N}\right)_v \quad (\text{A.2.12})$$

ofrece una forma de calcular los valores de electronegatividad para átomos, grupos funcionales, clusters, y moléculas. Partiendo de la ecuación (A.2.6) y haciendo la derivada a potencial externo constante por diferencias finitas se encuentra que:

$$\mu = \left( \frac{\partial E}{\partial N} \right)_v \cong \frac{\Delta E}{\Delta N} = \frac{E(N+1) - E(N-1)}{(N+1) - (N-1)} = \frac{(E(N) - A) - (E(N) + I)}{2} \quad (\text{A.2.13})$$

$$\mu = -\frac{(A + I)}{2} = -\chi \quad (\text{A.2.14})$$

donde  $A$  es la afinidad electrónica e  $I$  es el potencial de ionización del sistema bajo estudio. Este potencial químico, cambiado de signo, no es más que la definición original de Mulliken<sup>11</sup> para la electronegatividad. Como una aproximación a la ecuación (A.2.14), el potencial de ionización y la afinidad electrónica se pueden reemplazar por las energías del HOMO y LUMO, respectivamente, utilizando el teorema de Koopmans<sup>12</sup> dentro del esquema Hartree-Fock, obteniendo así

$$\chi = \frac{\varepsilon_{HOMO} + \varepsilon_{LUMO}}{2} \quad (\text{A.2.15})$$

En definitiva, desde un punto de vista energético, se favorece que haya una transferencia de carga desde el sistema de mayor potencial químico o menor electronegatividad, al de menor potencial químico o mayor electronegatividad. El potencial químico es pues una medida de la tendencia que tienen los electrones a escapar del sistema. Una diferencia de potencial químico grande entre dos sistemas favorece la transferencia electrónica.

### A.2.3.2. La dureza y la blandura globales

Los conceptos de dureza y blandura química fueron introducidos en los principios de los años 60 por Pearson con respecto al estudio generalizado de las reacciones ácido-base de Lewis,



donde **A** es un ácido de Lewis o aceptor de pares de electrones y **B** es una base de Lewis o dador de pares de electrones.<sup>13</sup> En el artículo publicado por Parr y Pearson<sup>14</sup> se identifica la dureza como la segunda derivada de la energía con respecto al número de electrones a potencial externo constante,

$$\eta = \frac{1}{2} \left( \frac{\partial^2 E}{\partial N^2} \right)_v = \frac{1}{2} \left( \frac{\partial \mu}{\partial N} \right)_v \quad (\text{A.2.16})$$

De una manera similar a la identificación de  $\chi$  como  $-(\delta E / \delta N)_v$ , esta última ecuación ofrece una definición clara que permite el cálculo de esta cantidad y su posterior comparación con datos experimentales. Es importante aclarar que en algunos textos el factor arbitrario  $\frac{1}{2}$  se omite. La dureza puede ser considerada pues como una medida de la resistencia a la transferencia de carga. Variar el número de electrones en un sistema duro va a ser costoso energéticamente y por tanto este sistema tendrá poca tendencia a dar o aceptar electrones. Como sucede con la electronegatividad, se han llevado a cabo muchos cálculos utilizando el método de las diferencias finitas<sup>15</sup> o con aproximaciones a él, indicando que la dureza no es más que la diferencia energética entre HOMO y LUMO.

$$\eta = \frac{I - A}{2} = \frac{\epsilon_{LUMO} - \epsilon_{HOMO}}{2} \quad (\text{A.2.17})$$

Es previsible que, en general, una diferencia energética grande corresponda a sistemas estables y poco reactivos, mientras que cuando la diferencia de energía HOMO-LUMO sea pequeña es de esperar que el sistema sea poco estable y altamente reactivo.

Ligados a la dureza hay dos principios de reactividad de gran importancia. El primero es el principio de máxima dureza que afirma que un sistema a potencial químico,

temperatura y potencial externo constantes tiende a un estado de máxima dureza.<sup>16, 17, 18</sup>

El segundo es el principio HSAB (Hard-Soft-Acids-Bases) que dice que ácidos duros prefieren bases duras y ácidos blandos prefieren bases blandas.<sup>19</sup>

Junto a este último principio, y estrechamente ligada, se encuentra la blandura global,  $S$ , que ha sido definida como el recíproco de la dureza,

$$S = \frac{1}{2\eta} = \left( \frac{\partial N}{\partial \mu} \right)_{v(r)} \quad (\text{A.2.18})$$

Siguiendo la relación dureza-polarizabilidad introducida en la aproximación original del principio HSAB de Pearson, no es sorprendente que la blandura sea una medida de la polarizabilidad de un sistema. Varios estudios que relacionan la polarizabilidad atómica con la blandura, confirman esta afirmación. Los dos índices globales de reactividad descritos anteriormente se pueden utilizar como herramientas útiles en la comprensión de las propiedades globales relacionadas con los aspectos termodinámicos de los sistemas reactantes en sus estados fundamentales.<sup>1, 14</sup>

### A.2.3.3. Electrofilia Global

Existe un índice de reactividad que combina la electronegatividad y la dureza: el índice de electrofilia, que ha sido introducido recientemente por Parr, Von Szentpaly y Liu.<sup>20,</sup>

<sup>21</sup>

Los autores comienzan su estudio refiriéndose a un trabajo previo realizado por Maynard *et al.*, en el cual se estudia la transferencia parcial de carga en la interacción ligando-proteína en un sistema de interés bioquímico,<sup>22</sup> y del cual surge la cantidad  $\chi^2_A/\eta_A$ , que se sugirió inicialmente como la capacidad de un electrofilo para estabilizar una interacción covalente (blanda). Luego ellos se formularon la siguiente pregunta: ¿hasta que medida la transferencia parcial electrónica entre un electro-dador y un electro-aceptor contribuye a la disminución de la energía total de unión en el caso de un

flujo máximo de electrones (nótese la diferencia con la afinidad electrónica, que mide la capacidad de un electro-aceptor de aceptar, precisamente, solo un electrón)? Utilizando un modelo de un ligando electrofilico sumergido en un mar de electrones libres ideal, a temperatura cero y potencial químico cero, el punto de saturación del ligando para la entrada de electrones fue definido como,

$$\Delta E / \Delta N = 0 \quad (\text{A.2.19})$$

Al realizar para  $\Delta E$ , el cambio de energía hasta segundo orden a potencial externo constante, tenemos que,

$$\Delta E = \mu \Delta N + \frac{1}{2} \eta \Delta N^2 \quad (\text{A.2.20})$$

donde  $\mu$  y  $\eta$  son el potencial químico y la dureza del ligando, respectivamente.

Si el mar de electrones provee suficientes electrones, el ligando está saturado cuando (combinando Ecuaciones (A.2.19) y (A.2.20)),

$$\Delta N_{\max} = -\frac{\mu}{\eta} \quad (\text{A.2.21})$$

Lo cual produce una estabilización de energía igual a,

$$\Delta E = -\frac{\mu^2}{2\eta} \quad (\text{A.2.22})$$

la cual es siempre positiva. La cantidad  $\mu^2/2\eta$ , abreviada como  $\omega$ , ha sido considerada una medida de la electrofilia de un ligando:

$$\omega = \frac{\mu^2}{2\eta} \quad (\text{A.2.23})$$

Así pues, la electrofilia es una cantidad que mide la estabilización de energía cuando el sistema adquiere una cantidad de carga electrónica desde el medio que la rodea.

#### A.2.3.4. La función de Fukui y la blandura local y grupal.

Además de los índices globales de reactividad, existe un conjunto de descriptores locales de reactividad que han sido relacionados con la selectividad de las moléculas hacia reacciones específicas en algunos sistemas orgánicos.<sup>23, 24</sup>

La función de Fukui ( $f(\mathbf{r})$ ) es uno de los descriptores más utilizados en reactividad local y fue introducida por Parr y Yang<sup>25, 26</sup> como una generalización del concepto de orbital molecular frontera (OMF) de Fukui,<sup>27, 28, 29</sup> y juega un papel clave en la teoría del OMF enlazante y el principio HSAB.<sup>19</sup>

La función de Fukui se puede interpretar o bien como una medida de la sensibilidad del potencial químico de un sistema ante un cambio en el potencial externo, o bien como una medida de cómo varía la densidad electrónica al aumentar el número de electrones del sistema. Un valor alto de este índice local está asociado con una alta reactividad en ese punto de la región molecular. La fórmula operacional para evaluar este índice condensado a átomos ya ha sido descrita<sup>30, 31, 32</sup>,

$$f(\mathbf{r}) = \left( \frac{\partial \mu}{\partial \nu(\mathbf{r})} \right)_N = \left( \frac{\partial \rho(\mathbf{r})}{\partial \mathbf{N}} \right)_\nu \quad (\text{A.2.24})$$

Esta última definición permite demostrar fácilmente que la función de Fukui está normalizada a la unidad:

$$\int f(\vec{r}) d\vec{r} = 1 \quad (\text{A.2.25})$$

A partir de la Ecuación (A.2.24) y utilizando una vez más la aproximación de diferencias finitas se pueden definir tres tipos de funciones de Fukui. La  $f^+(r)$  se define según:

$$f^+(\vec{r}) \cong \frac{\rho_{N+1}(\vec{r}) - \rho_N(\vec{r})}{N+1 - N} = \rho_{N+1}(\vec{r}) - \rho_N(\vec{r}) \cong \rho_{LUMO}(\vec{r}) \quad (\text{A.2.26})$$

y nos da la región del espacio en la que se localiza la carga cuando el sistema adquiere un electrón. Es una buena medida de la reactividad de nuestro sistema frente a un nucleófilo y en primera aproximación es la densidad electrónica del LUMO. Por su parte, la  $f^-(r)$  es una medida de la reactividad frente a un electrófilo:

$$f^-(\vec{r}) \cong \frac{\rho_N(\vec{r}) - \rho_{N-1}(\vec{r})}{N - N + 1} = \rho_N(\vec{r}) - \rho_{N-1}(\vec{r}) \cong \rho_{HOMO}(\vec{r}) \quad (\text{A.2.27})$$

Finalmente  $f^0(r)$  es la media de las dos anteriores y nos da la reactividad frente a un ataque radicalario:

$$f^0(\vec{r}) = \frac{1}{2} [f^+(\vec{r}) + f^-(\vec{r})] \cong \frac{1}{2} [\rho_{N+1}(\vec{r}) - \rho_{N-1}(\vec{r})] \cong \frac{1}{2} [\rho_{LUMO}(\vec{r}) + \rho_{HOMO}(\vec{r})] \quad (\text{A.2.28})$$

donde en notación convencional HOMO representa el orbital molecular ocupado más alto (*highest occupied molecular orbital*) de la especie en cuestión, y LUMO es el orbital molecular desocupado más bajo (*lowest unoccupied molecular orbital*).

Este índice es quizás una forma muy simplificada de ver la reactividad química, pero es muy útil. En cualquier caso  $f(r)$  se ha establecido como un índice de considerable

importancia para entender el comportamiento molecular, o sea es el índice natural de selectividad posicional en la teoría de los funcionales de la densidad.

La función de Fukui claramente contiene información relativa sobre diferentes regiones en una molécula dada. Pero cuando se quieren comparar diferentes regiones en moléculas diferentes, la blandura local resultar ser más interesante (ver referencia 33). La cantidad  $s(r)$  fue introducida en 1985 por Yang y Parr como<sup>34</sup>,

$$s(\mathbf{r}) = \left( \frac{\partial \rho(\mathbf{r})}{\partial \mu} \right)_{v(\mathbf{r})} \quad (\text{A.2.29})$$

y, esta a su vez, es un análogo de la blandura total,  $S$ , que se puede escribir como:

$$S = \left( \frac{\partial N}{\partial \mu} \right)_{v(\mathbf{r})} \quad (\text{A.2.30})$$

Aplicando la regla de la cadena,  $s(r)$  se puede escribir como un producto de la blandura total y la función de Fukui,

$$s(\mathbf{r}) = \left( \frac{\partial \rho(\mathbf{r})}{\partial \mu} \right)_{v} = \left( \frac{\partial \rho(\mathbf{r})}{\partial N} \right)_{v} \left( \frac{\partial N}{\partial \mu} \right)_{v} = f(\mathbf{r})S \quad (\text{A.2.31})$$

indicando esto que  $f(r)$  distribuye la blandura global entre las diferentes regiones de la molécula. Además, que  $s(r)$  integra a  $S$ :

$$\int s(\mathbf{r})d\mathbf{r} = \int Sf(\mathbf{r})d\mathbf{r} = S \int f(\mathbf{r})d\mathbf{r} = S \quad (\text{A.2.32})$$

El poder de predicción de la reactividad intermolecular de la blandura local emerge claramente a partir de la Ecuación (A.2.29), mostrando que  $f(r)$  y  $s(r)$  contienen la misma información sobre los sitios reactivos relativos dentro de una misma molécula, pero que  $s(r)$ , en vista de la información de la blandura molecular total que contiene, es más apropiado para evaluar la reactividad intermolecular.

Las cantidades  $s(r)$ ,  $f(r)$  y  $S$  son bien conocidas en la teoría de metales. De allí surgen varias ecuaciones que son análogas a estas y que se pueden aplicar en sistemas moleculares grandes y donde casi no existe una diferencia clara entre los orbitales HOMO y LUMO debido a la degeneración de los estados de energía. Así, si dejamos que  $g(\epsilon_F, r)$  sea la densidad de estados local en el nivel de Fermi y que  $g(\epsilon_F)$  sea la densidad de estados total en el nivel de Fermi, a 0K, entonces tenemos:

$$s(\mathbf{r}) = g(\epsilon_F, \mathbf{r}) \quad (\text{A.2.33})$$

$$S = g(\epsilon_F) \quad (\text{A.2.34})$$

y por tanto,

$$f(\mathbf{r}) = \frac{g(\epsilon_F, \mathbf{r})}{g(\epsilon_F)} \quad (\text{A.2.35})$$

Por último podemos decir que existen también dos índices de reactividad regional, la blandura regional,  $s_k$ , y la electrofilia regional,  $\omega_k$ , condensados a un átomo  $k$ , que han sido derivados a través de la función de Fukui:<sup>1</sup>

$$s_k^\pm = f_k^\pm S \quad (\text{A.2.36})$$

$$\omega_k = f_k^{\pm} \omega \quad (\text{A.2.37})$$

Donde  $f_k$  es la función de Fukui en el sitio  $k$  y esta asociado en la dirección de aumentar (+) o disminuir (-) el número de electrones. Estos índices están asociados con ataques nucleofílicos y electrofílicos en el sitio  $k$ , respectivamente. Es importante resaltar que de acuerdo con las Ecuaciones (A.2.36) y (A.2.37), la blandura y la electrofilia local también están distribuidas siguiendo la función de Fukui.

Con el objetivo de caracterizar una región molecular más extendida, que incluya más de un centro atómico, se puede utilizar el concepto de blandura grupal o regional,

$$S_G = \sum_{k \in G} S_k^{\pm} \quad (\text{A.2.38})$$

Donde  $S_k^{\pm}$  está definido como en la Ecuación (A.2.36). El concepto de blandura grupal ya fue introducido y aplicado por Gázquez y Méndez<sup>35</sup> y otros<sup>36, 37</sup>

---

#### A.2.4. Bibliografía

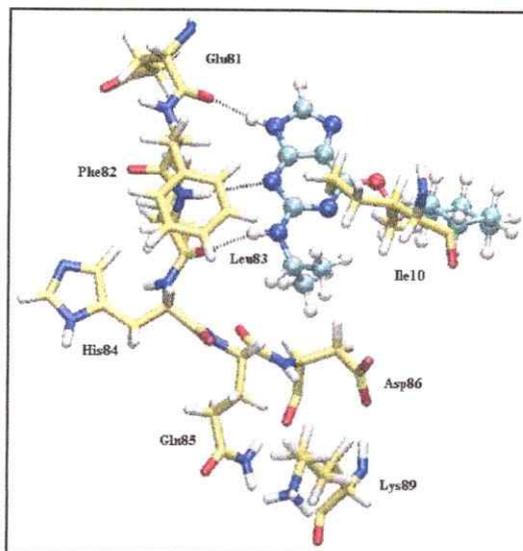
- [1] Parr RG, Yang W. *Density Functional Theory of Atoms and Molecules*. Oxford University Press, New York, 1989.
- [2] Hohenberg, P.; Kohn, W. *Phys Rev* 1964, A136, 864.
- [3] Kohn, W.; Sham, L. J. *Phys. Rev. A* 1965, 140, 1133.
- [4] Hohenberg, P.; Kohn, W.; Sham, L. *Adv. Quantum Chem.* 1990, 21, 7.
- [5] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*; Gaussian Inc.: Pittsburgh PA, 1998, and previous versions.
- [6] Becke, A. D. *J. Chem. Phys.* 1993, 98, 5648.
- [7] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* 1988, 37, 785.
- [8] Geerlings, P.; De Proft, F. and Langenaeker, W. *Chem. Rev.* 2003, 103, 1793-1873.

- 
- [9] Caballol, R.; Solá, M. En *Métodos de la Química Cuántica*, Andrés, J.; Beltrán, J., Ed.; Química Teórica y Computacional, Colección Ciencias Experimentales 2; Universitat Jaume I: Castelló de la Plana, 2000, p 107.
- [10] Parr, R. G.; Donnelly, R. A.; Levy, M.; Palke, W. E. *J. Chem. Phys.* **1978**, *68*, 3801.
- [11] Mulliken, R. S. *J. Chem. Phys.* **1934**, *2*, 782.
- [12] Koopmans, T. A. *Physica* **1933**, *1*, 104.
- [13] Pearson, R. G. *J. Am. Chem. Soc.* **1963**, *85*, 3533.
- [14] Parr, R. G.; Pearson, R. G. *J. Am. Chem. Soc.* **1983**, *105*, 7512.
- [15] Geerlings, P.; De Proft, F.; Martin, J. M. L. In *Recent Developments and Applications of Modern Density Functional Theory*; Seminario, J. M., Ed.; Theoretical and Computational Chemistry 4; Elsevier: Amsterdam, **1996**; p 773.
- [16] Pearson, R. G. *J. Chem. Educ.* **1968**, *45*, 981.
- [17] Parr, R. G.; Chattaraj, P. K. *J. Am. Chem. Soc.* **1991**, *113*, 1854.
- [18] Pearson, R. G. *Chemical Hardness*; Wiley-VCH: Weinheim, 1997.
- [19] Li, Y.; Evans, J. N. S. *J. Am. Chem. Soc.* **1995**, *117*, 7756.
- [20] Parr, R. G.; Von Szentpaly, L.; Liu, S. *J. Am. Chem. Soc.* **1999**, *121*, 1922.

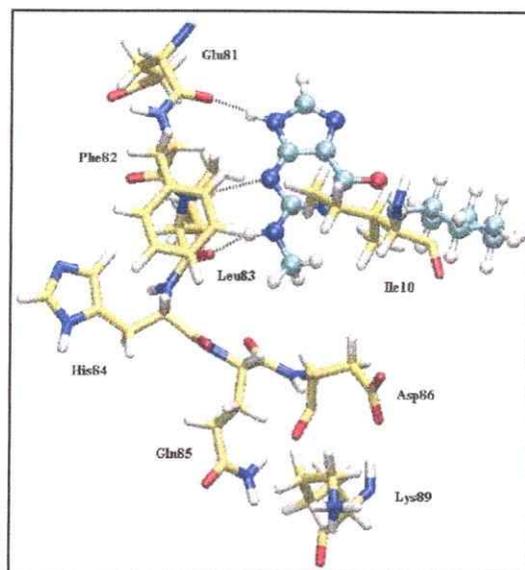
- 
- [21] Von Szentpaly, L. *Int. J. Quantum Chem.* **2000**, *76*, 222.
- [22] Maynard, A. T.; Huang, M.; Rice, W. G.; Covell, D. G. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 11578.
- [23] Contreras, R.; Domingo, L. R.; Andrés, L.; Pérez, P.; Tapia, O. *J. Phys. Chem. A* **1999**, *103*, 1367-1375.
- [24] Pérez, P.; Toro-Labbé, A.; Aizman, A.; Contreras, R. *J. Org. Chem.* **2002**, *67*, 4747-4752.
- [25] Parr, R. G.; Yang, W. *J. Am. Chem. Soc.* **1984**, *106*, 4049.
- [26] Para una perspectiva, ver: Ayers, P. W.; Levy, M. *Theor. Chem. Acc.* **2000**, *103*, 353.
- [27] Fukui, K.; Yonezawa, Y.; Shingu, H. *J. Chem. Phys.* **1952**, *20*, 722.
- [28] Parr, R. G.; Parr, J. B. *Theor. Chem. Acc.* **1999**, *102*, 5.
- [29] Kato, S. *Theor. Chem. Acc.* **2000**, *103*, 219.
- [30] Yang, W.; Mortier, W. *J. Am. Chem. Soc.* **1986**, *108*, 5708-5711.
- [31] Contreras, R.; Fuentealba, P.; Galván, M.; Pérez, P. *Chem. Phys. Lett.* **1999**, *304*, 239-413.
- [32] Fuentealba, P.; Pérez, P.; Contreras, R. *J. Chem. Phys.* **2000**, *113*, 2544-2551.

- 
- [33] Geerlings, P.; De Proft, F.; Langenaeker, W. *Adv. Quantum Chem.* **1999**, *33*, 303.
- [34] Yang, W.; Parr, R. G. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 6723.
- [35] Gázquez, J. L.; Méndez, F. *J. Phys. Chem. A* **1994**, *98*, 4591–4593.
- [36] Krishnamurty, S.; Pal, S. *J. Phys. Chem. A* **2000**, *104*, 7639–7645.
- [37] Rivas, P.; Zapata-Torres, G.; Melin, J.; Contreras, R. *Tetrahedron*, **2004**, *60*, 4189-4196.

## ANEXO 1



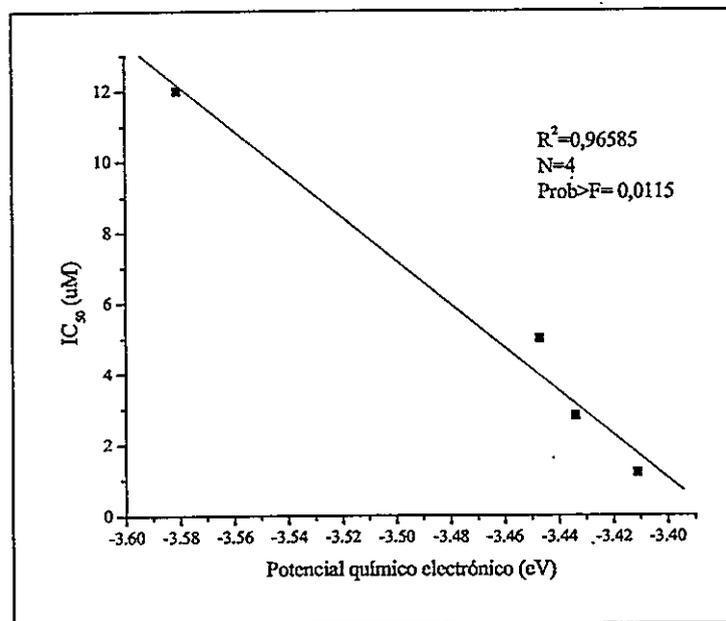
**Figura A.3.1.** Fotografía de la estructura de CDK2-Cp9. La representación es realizada como en la Figura 3.5.



**Figura A.3.2.** Fotografía de la estructura de CDK2-Cp7. La representación es realizada como en la Figura 3.5.

Compuesto	IC <sub>50</sub> (μM) Experimentales	LogIC <sub>50</sub> Experimentales	Valores de IC <sub>50</sub> calculados (μM)	Valores de LogIC <sub>50</sub> calculados
3	0.0054 (±0.0010)	-2.268	0.0043	-2.368
25	0.069 (±0.001)	-1.161	0.158	-0.800
2	0.970 (±0.030)	-0.013	0.497	-0.303
9	1.200 (±0.200)	0.079	1.680	0.225
7	5.000 (±0.300)	0.699	3.820	0.582

**TABLA A.3.1.** Resultados del análisis de regresión múltiple (Ver Ecuación (3.5)) para la serie de compuestos que fueron estudiados y pertenecen a la familia de las 6-ciclohexil-metoxi-purinas N<sup>2</sup>-sustituidas.



**Figura A.4.1.** Correlación entre la actividad media inhibitoria (IC<sub>50</sub>) y el Potencial químico electrónico de los compuestos del grupo I.

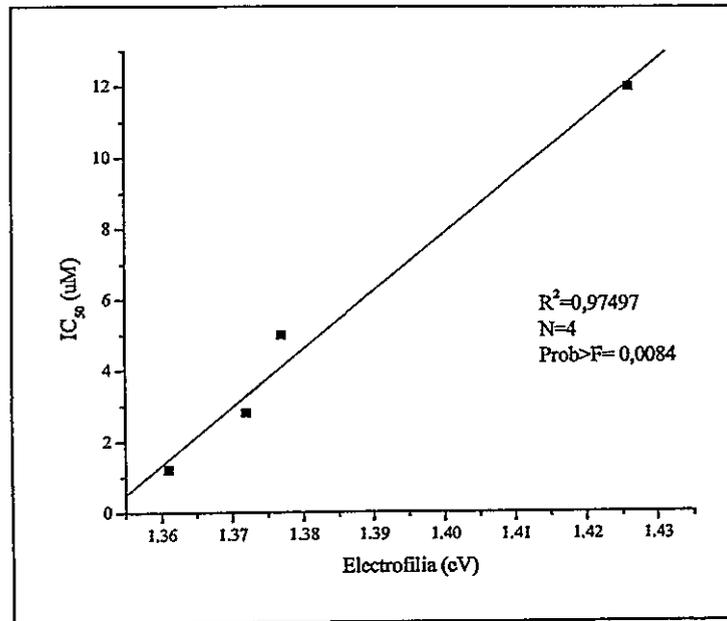


Figura A.4.2. Correlación entre la actividad media inhibitoria ( $IC_{50}$ ) y la electrofilia de los compuestos del grupo I.

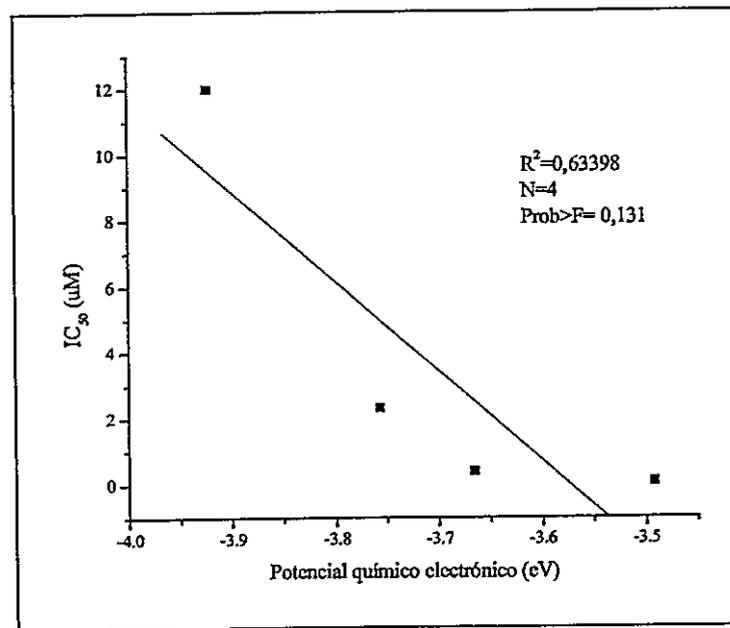


Figura A.4.3. Correlación entre la actividad media inhibitoria ( $IC_{50}$ ) y el Potencial químico electrónico de los compuestos del grupo IIa.

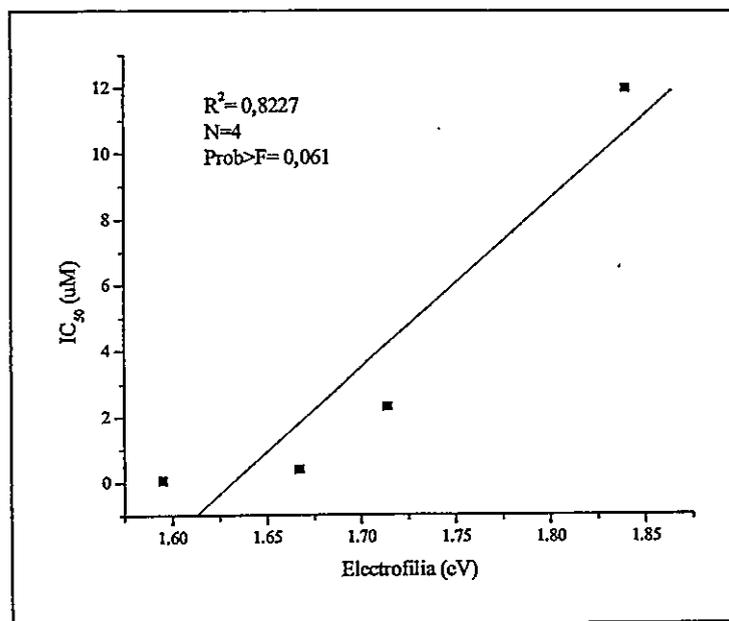


Figura A.4.4. Correlación entre la actividad media inhibitoria ( $IC_{50}$ ) y la electrofilia de los compuestos del grupo IIa.

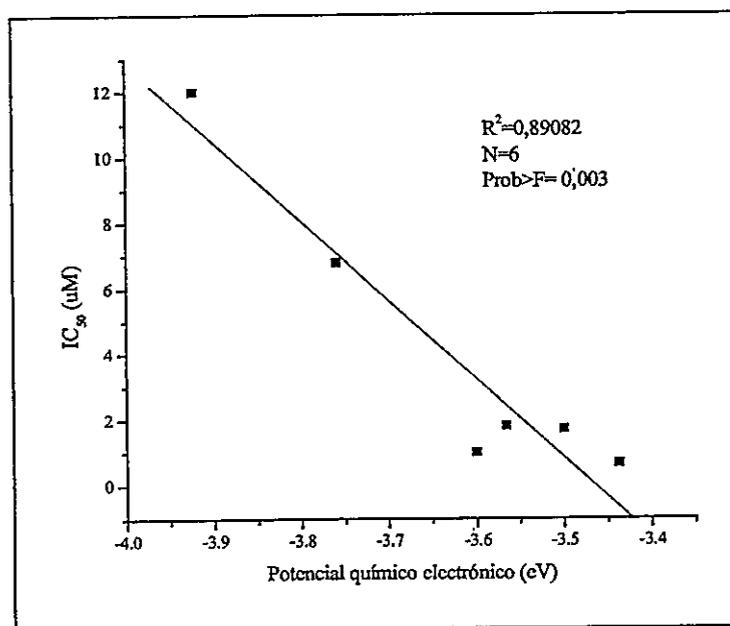


Figura A.4.5. Correlación entre la actividad media inhibitoria ( $IC_{50}$ ) y el Potencial químico electrónico de los compuestos del grupo IIb.

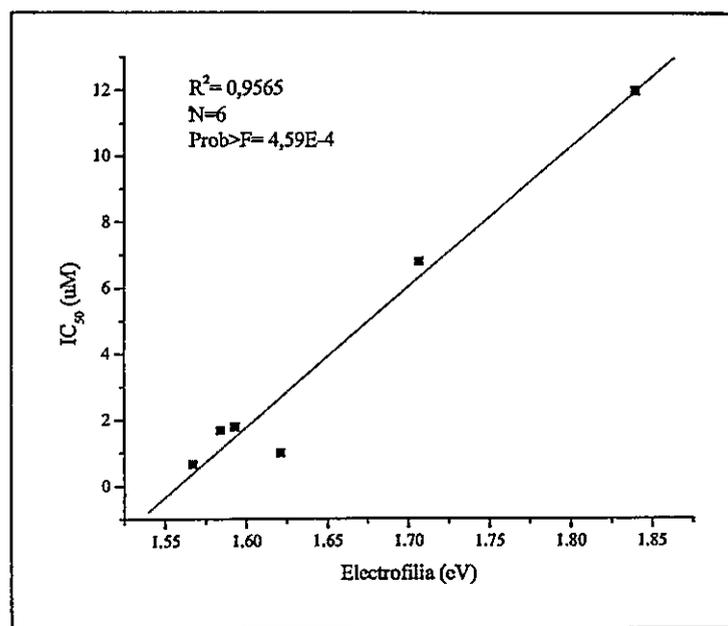


Figura A.4.6. Correlación entre la actividad media inhibitoria (IC<sub>50</sub>) y la electrofilia de los compuestos del grupo IIb.

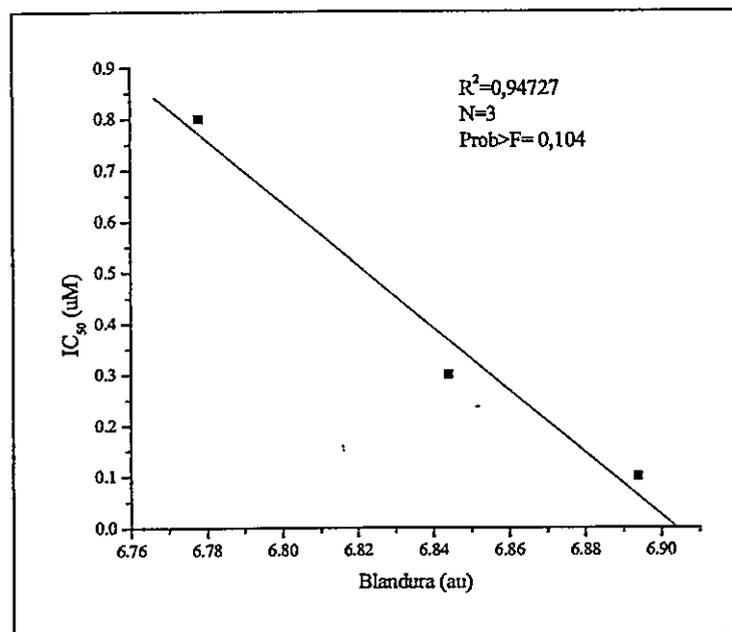
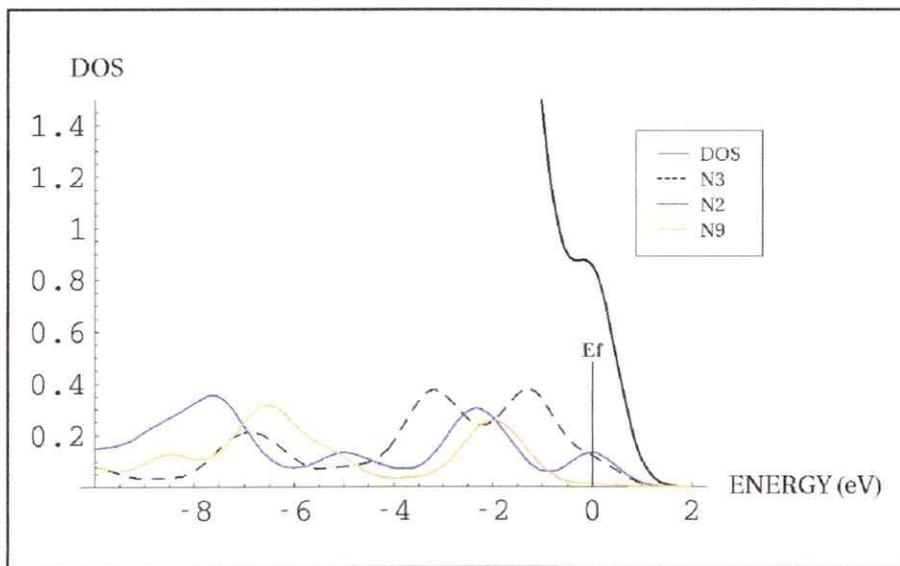
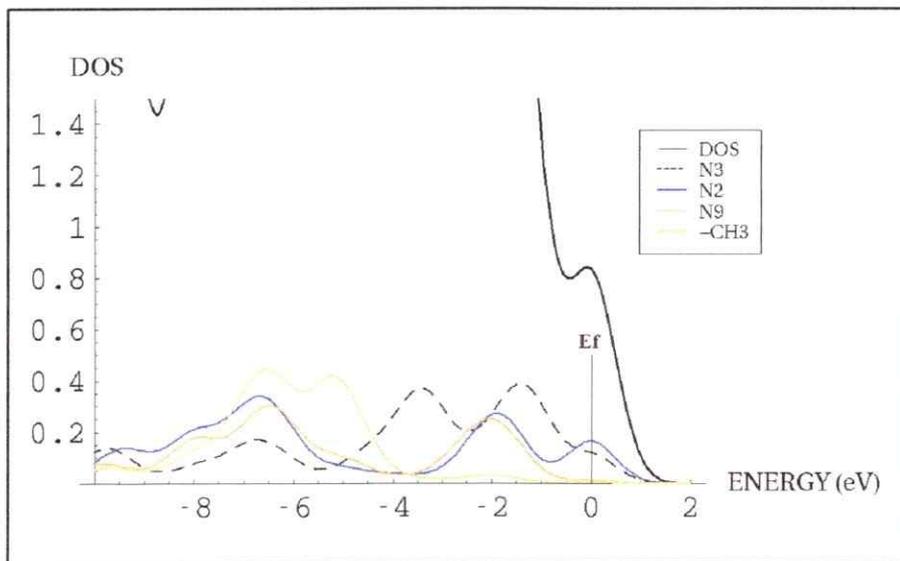


Figura A.4.7. Correlación entre la actividad media inhibitoria (IC<sub>50</sub>) y la blandura global de los compuestos del grupo IIIa.



**Figura A.4.8.** Densidad de estados global y condensada (DOS) para el compuesto 1.



**Figura A.4.9.** Densidad de estados global y condensada (DOS) para el compuesto 7.

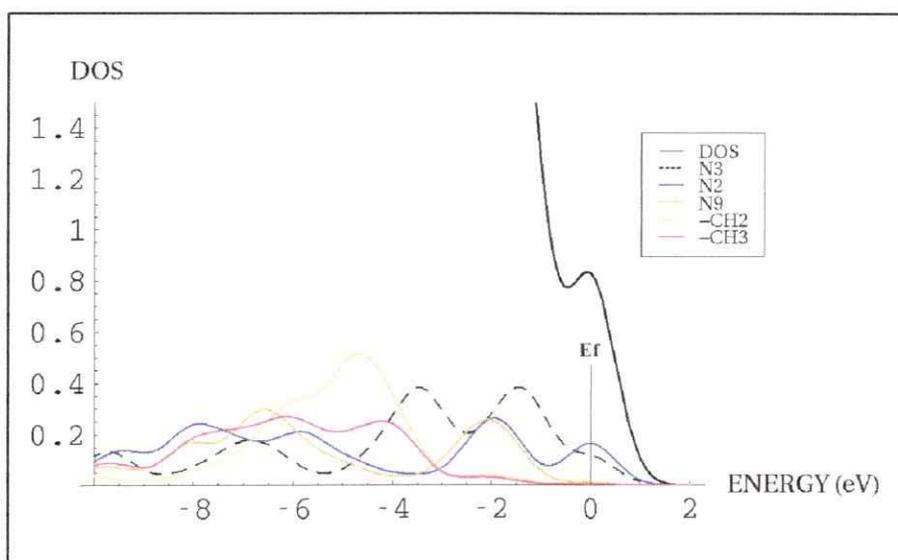


Figura A.4.10. Densidad de estados global y condensada (DOS) para el compuesto 8.

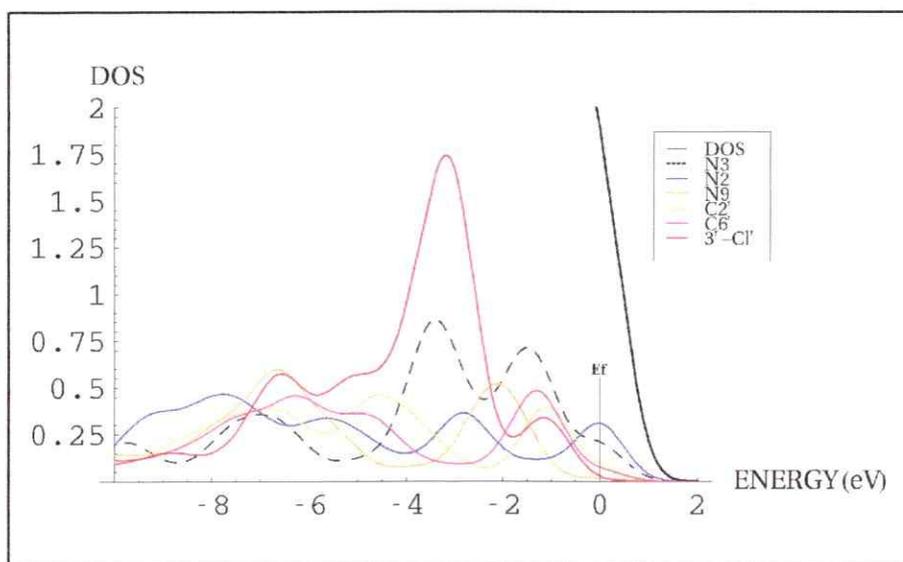


Figura A.4.11. Densidad de estados global y condensada (DOS) para el compuesto 15.

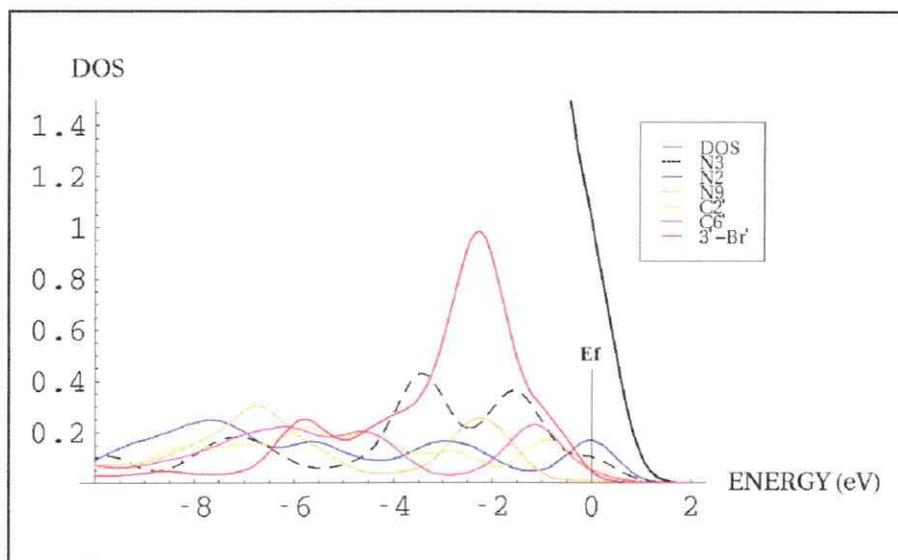


Figura A.4.12. Densidad de estados global y condensada (DOS) para el compuesto 17.

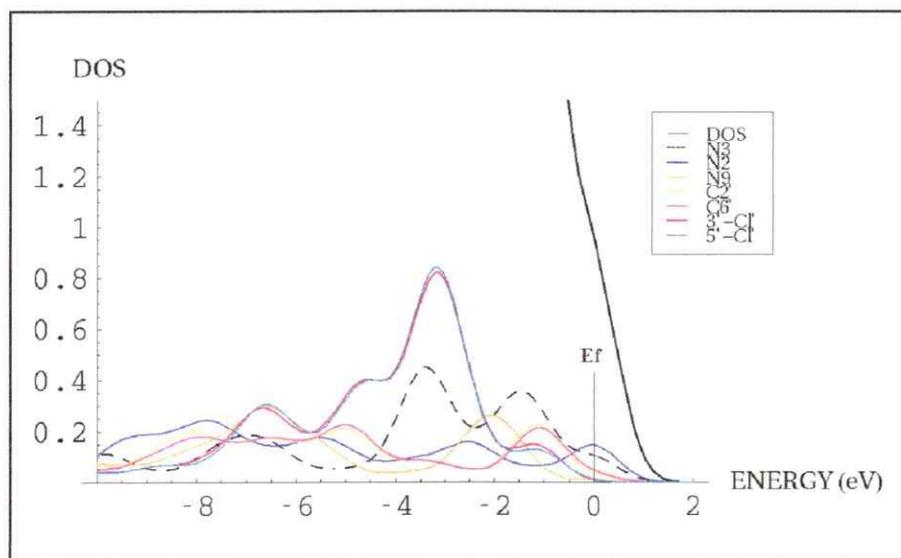


Figura A.4.13. Densidad de estados global y condensada (DOS) para el compuesto 18.

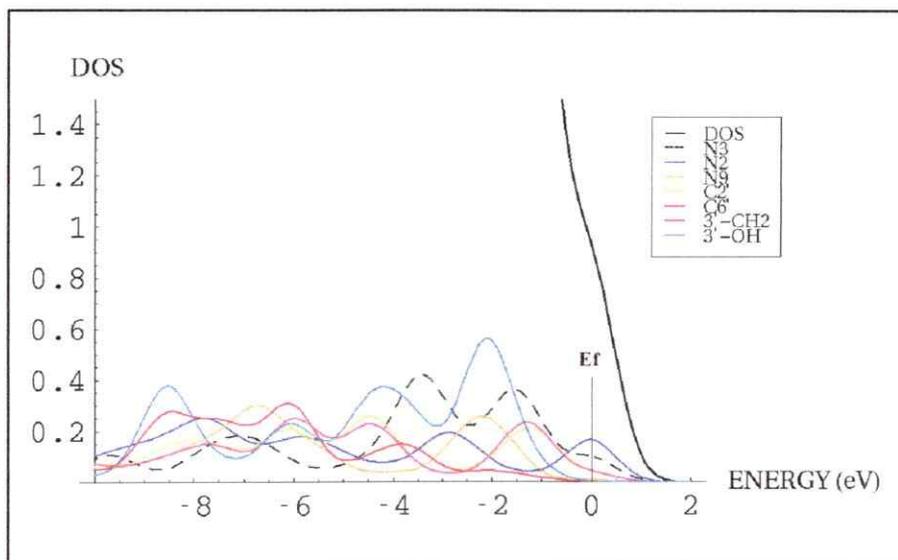


Figura A.4.14. Densidad de estados global y condensada (DOS) para el compuesto 20.

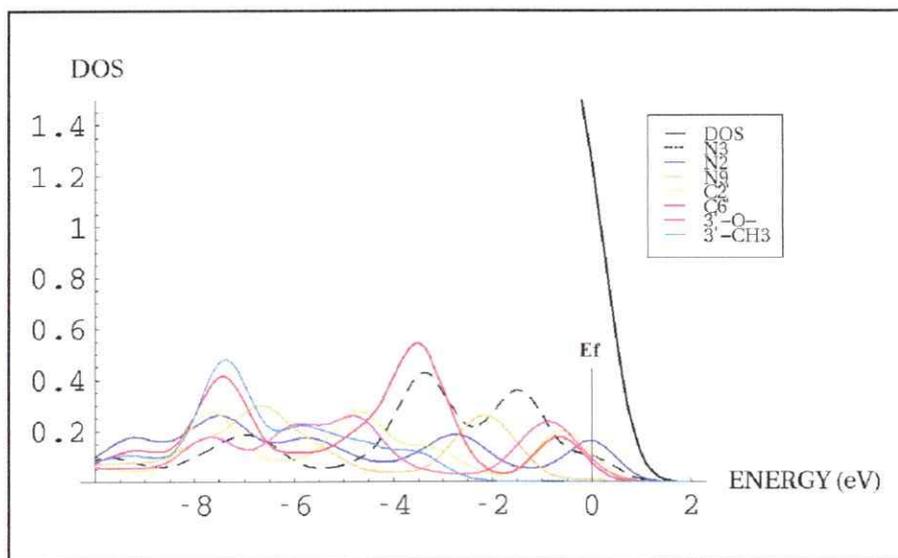


Figura A.4.15. Densidad de estados global y condensada (DOS) para el compuesto 21.

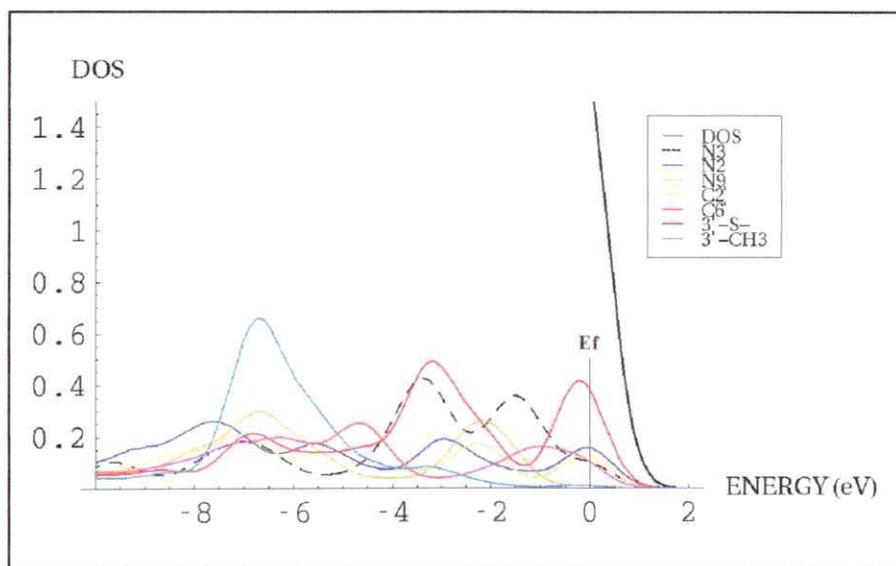


Figura A.4.16. Densidad de estados global y condensada (DOS) para el compuesto 22.

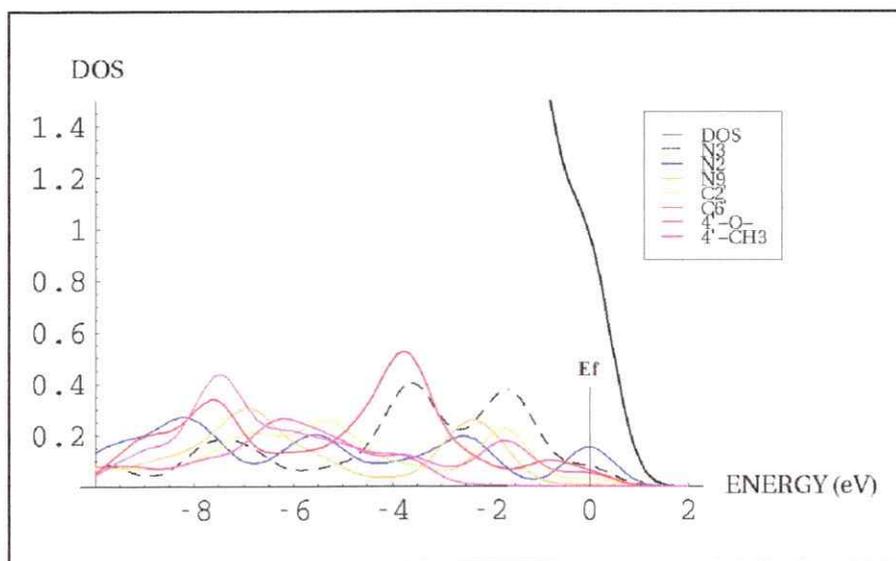
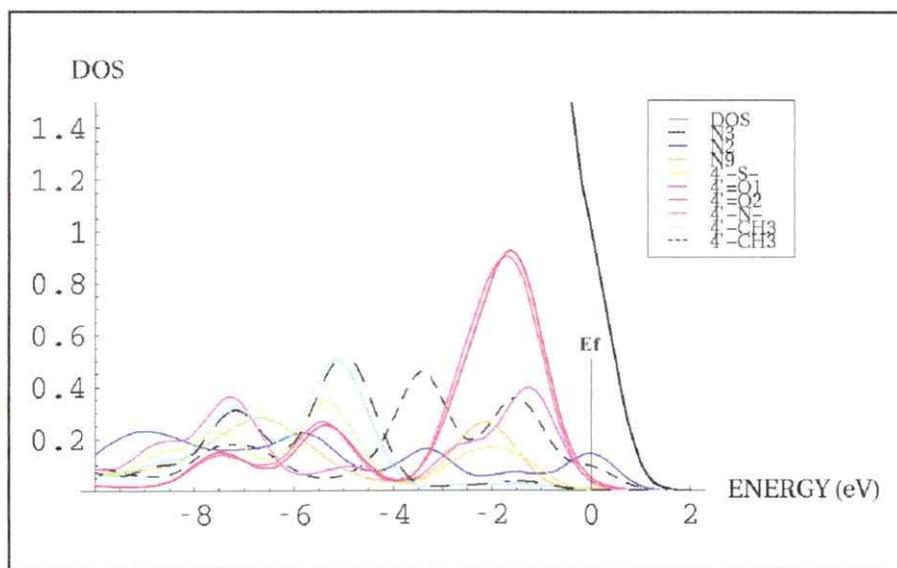
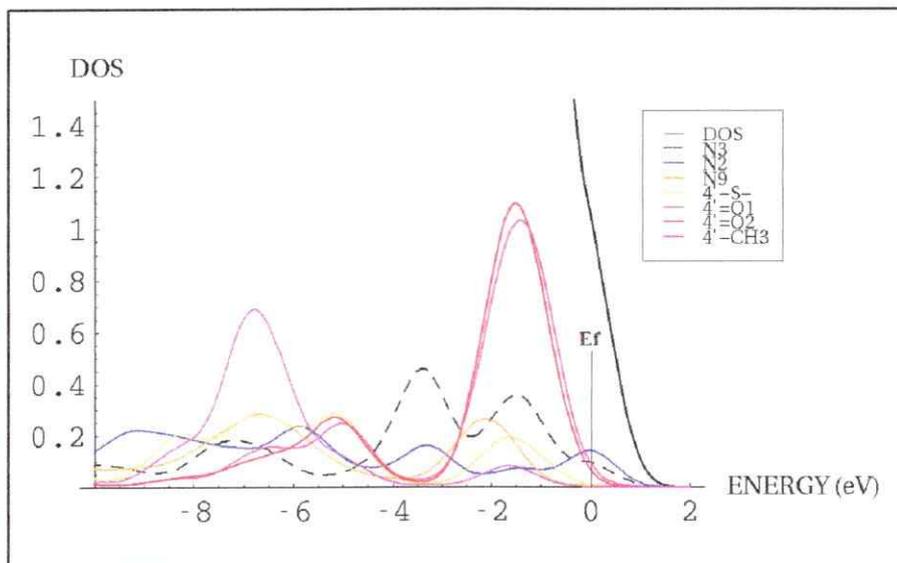


Figura A.4.17. Densidad de estados global y condensada (DOS) para el compuesto 26.



**Figura A.4.18.** Densidad de estados global y condensada (DOS) para el compuesto 29.



**Figura A.4.19.** Densidad de estados global y condensada (DOS) para el compuesto 30.

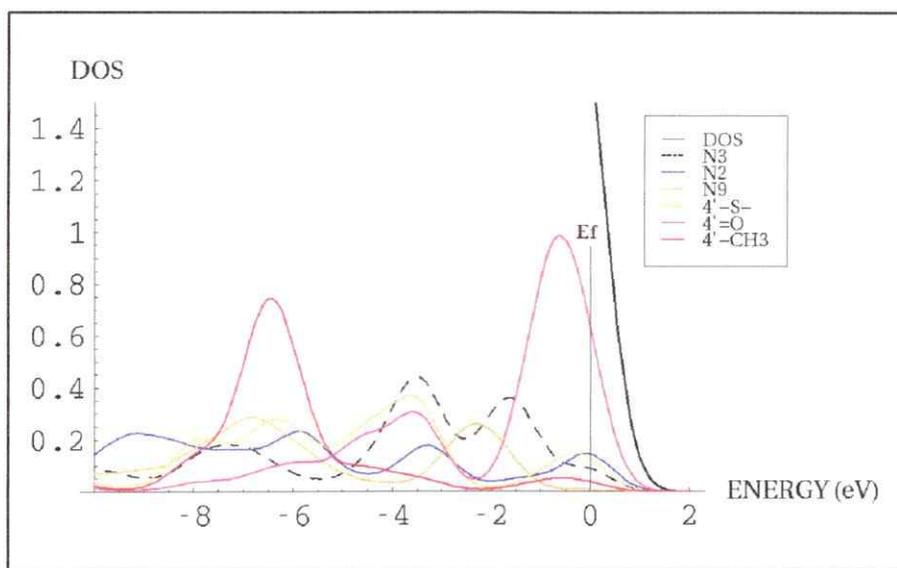


Figura A.4.20. Densidad de estados global y condensada (DOS) para el compuesto 31.

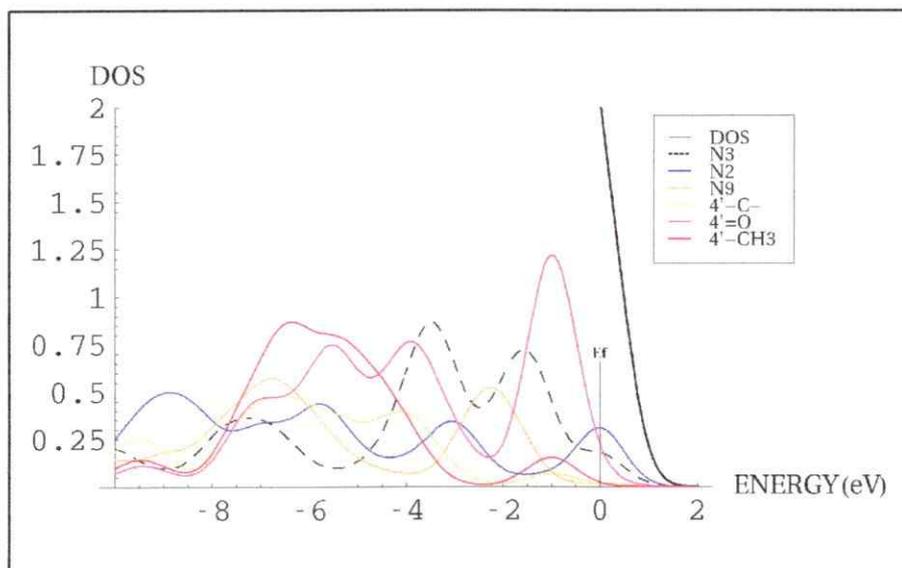
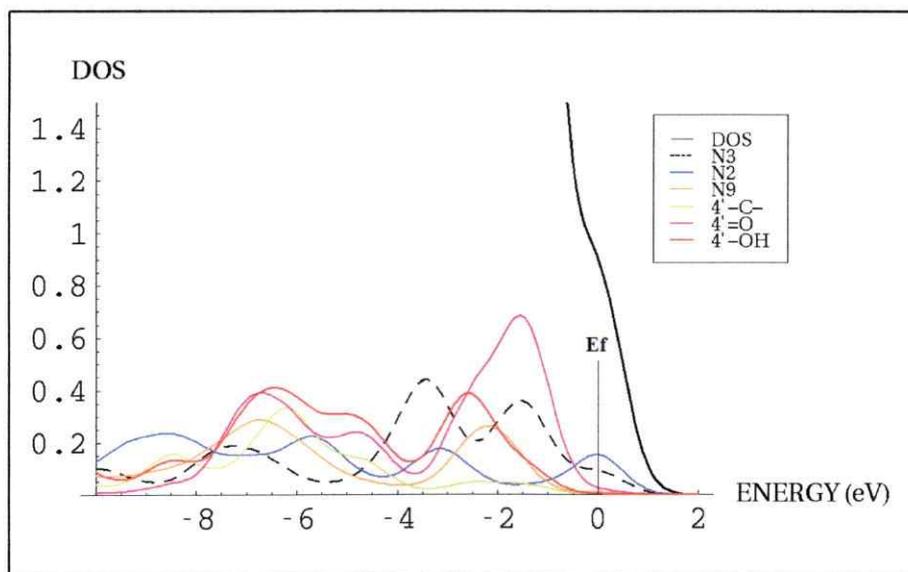
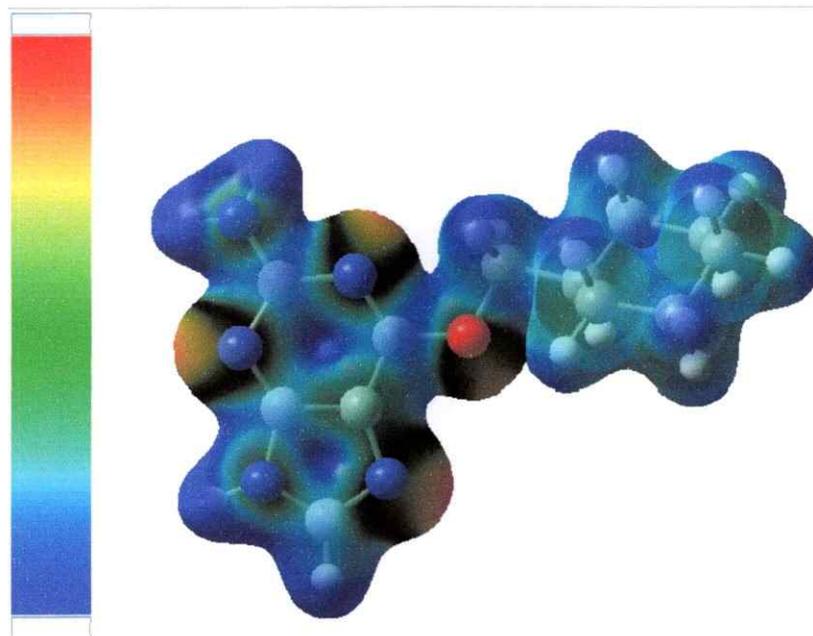


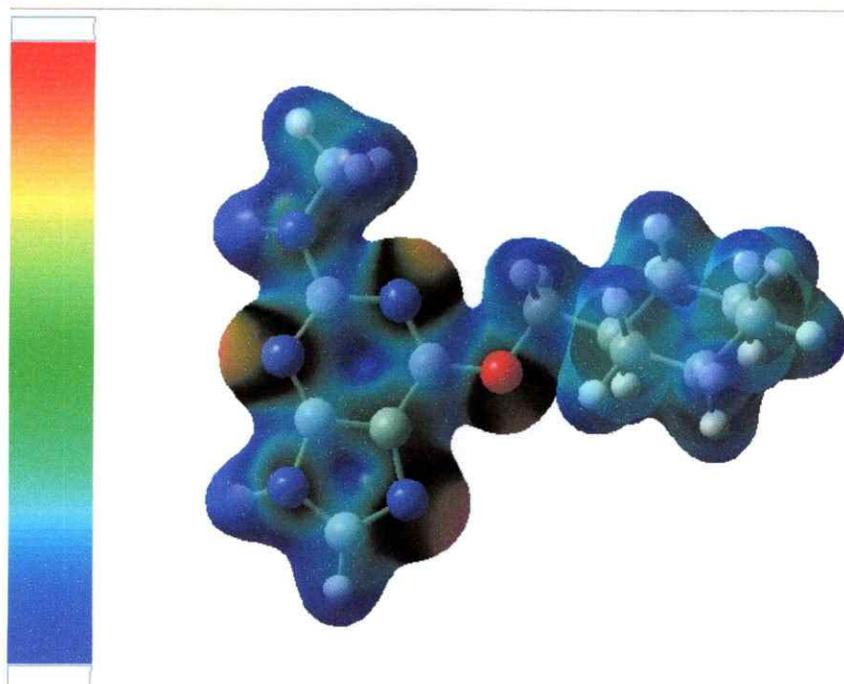
Figura A.4.21. Densidad de estados global y condensada (DOS) para el compuesto 38.



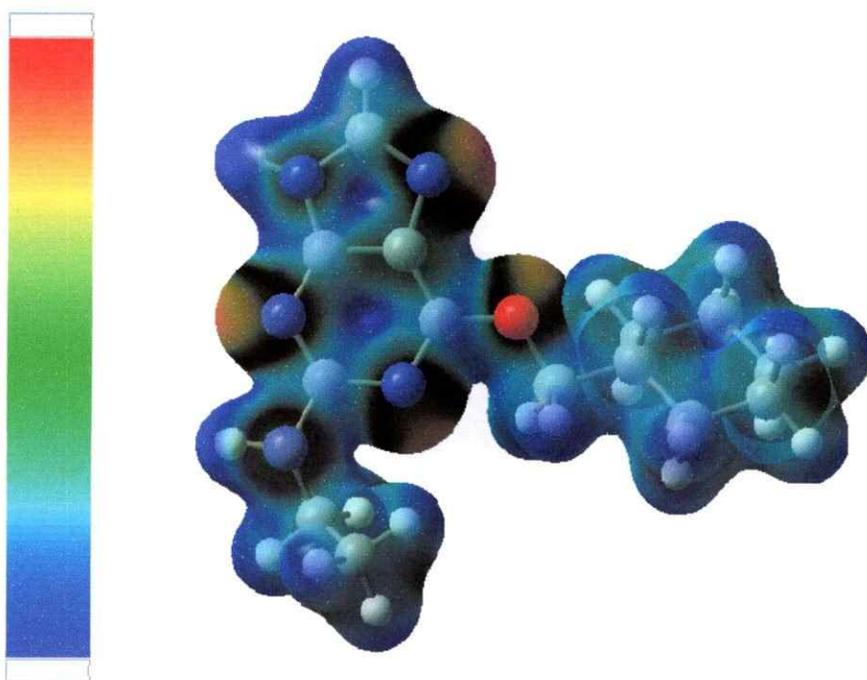
**Figura A.4.22.** Densidad de estados global y condensada (DOS) para el compuesto 39.



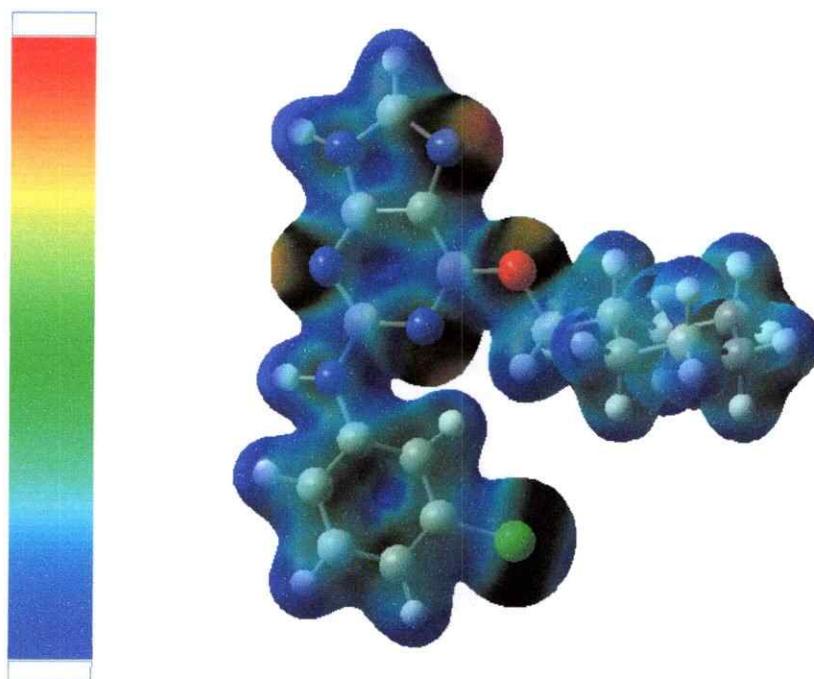
**Figura A.4.23.** Potencial Electrostático Molecular para el compuesto 1.



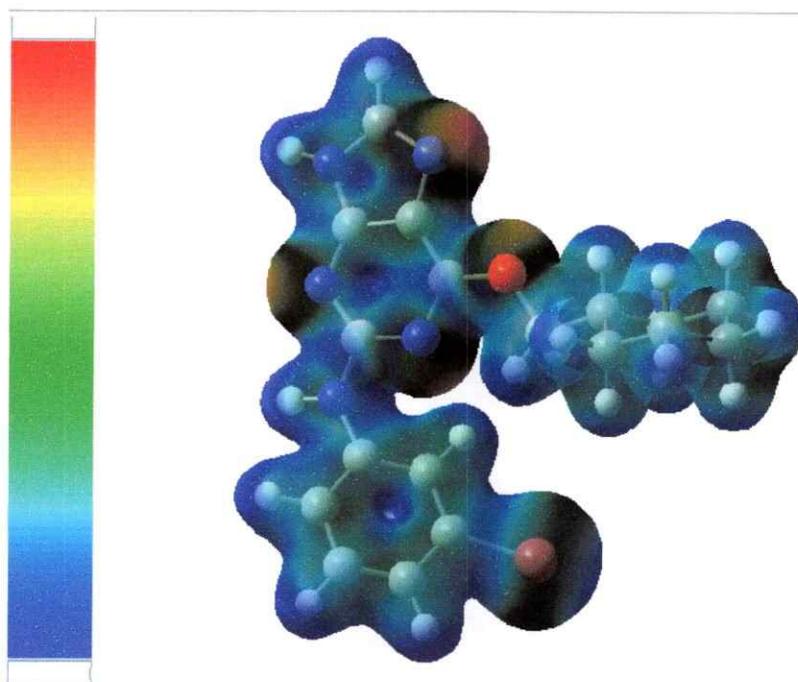
**Figura A.4.24.** Potencial Electrostático Molecular para el compuesto 7.



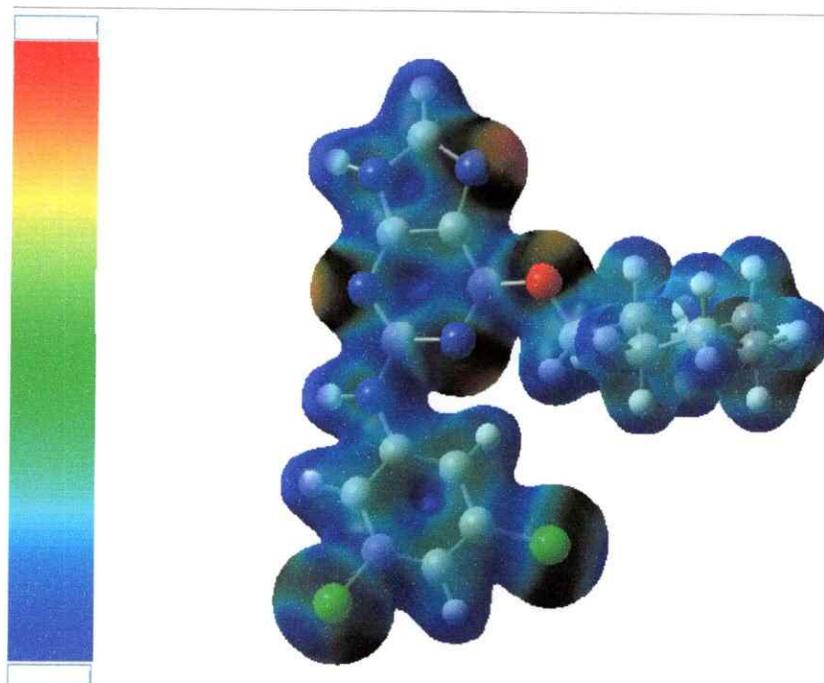
**Figura A.4.25.** Potencial Electrostático Molecular para el compuesto 8



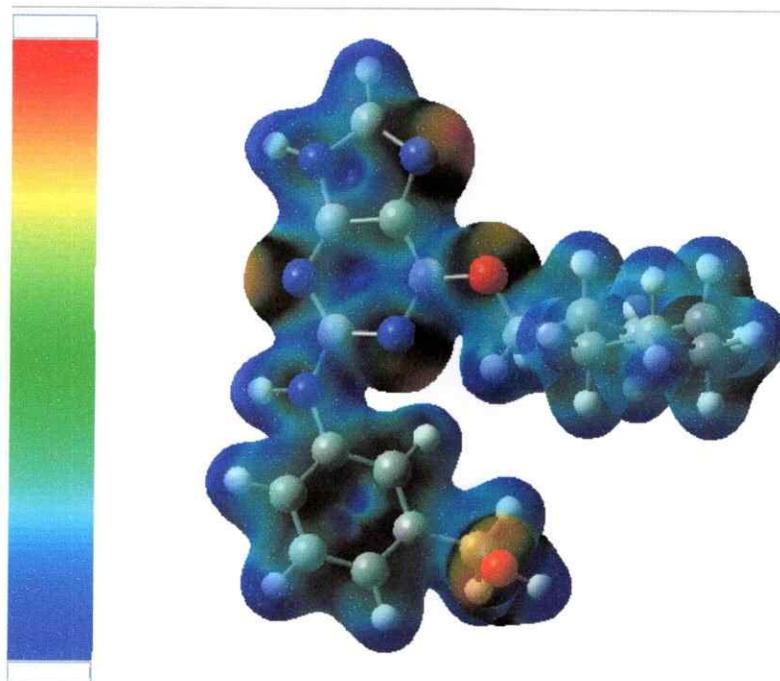
**Figura A.4.26.** Potencial Electrostático Molecular para el compuesto 15.



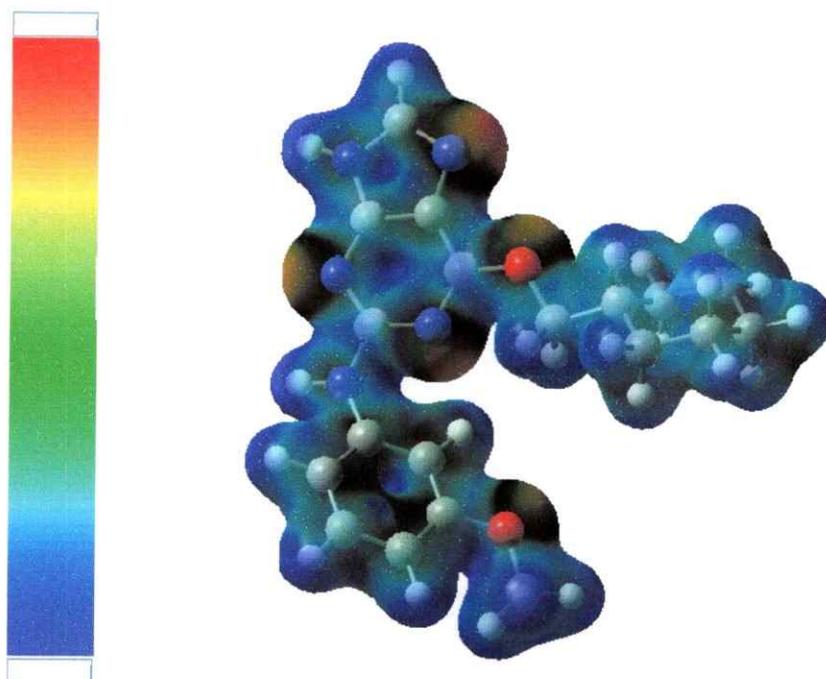
**Figura A.4.27.** Potencial Electrostático Molecular para el compuesto 17.



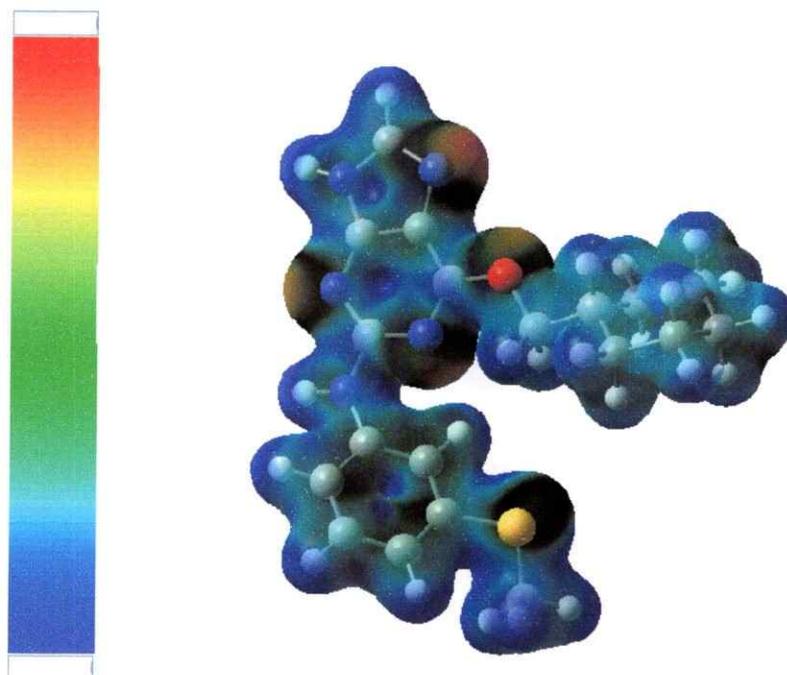
**Figura A.4.28.** Potencial Electrostático Molecular para el compuesto 18.



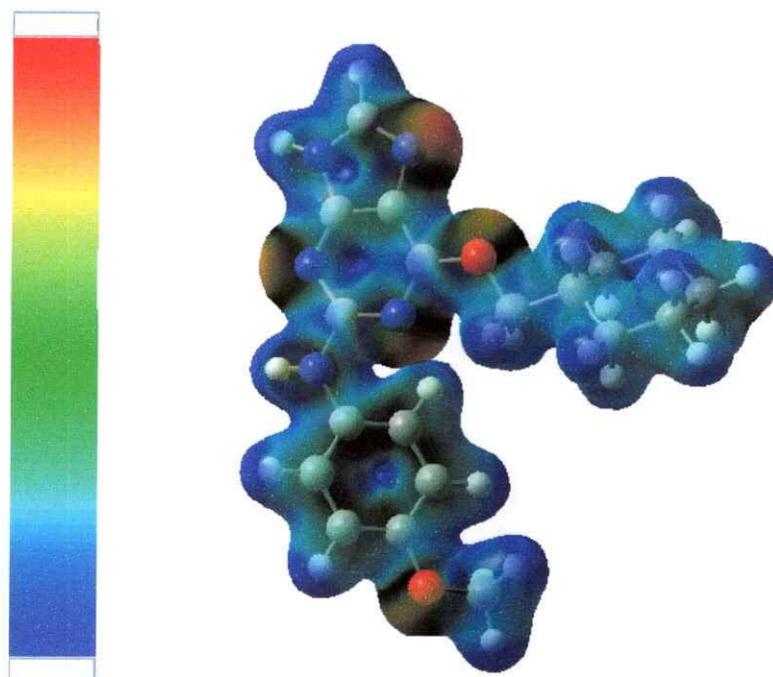
**Figura A.4.29.** Potencial Electrostático Molecular para el compuesto 20.



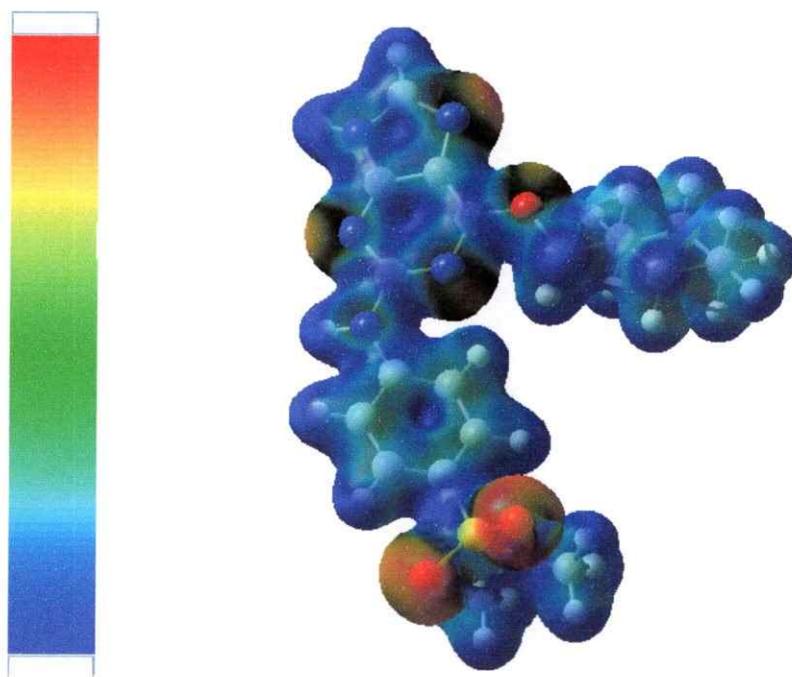
**Figura A.4.30.** Potencial Electrostático Molecular para el compuesto 21.



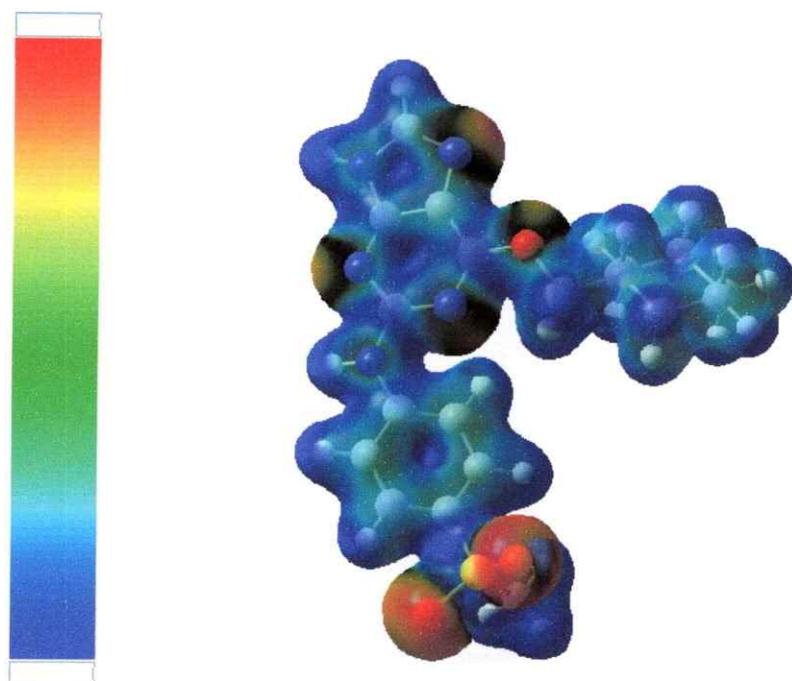
**Figura A.4.31.** Potencial Electrostático Molecular para el compuesto 22.



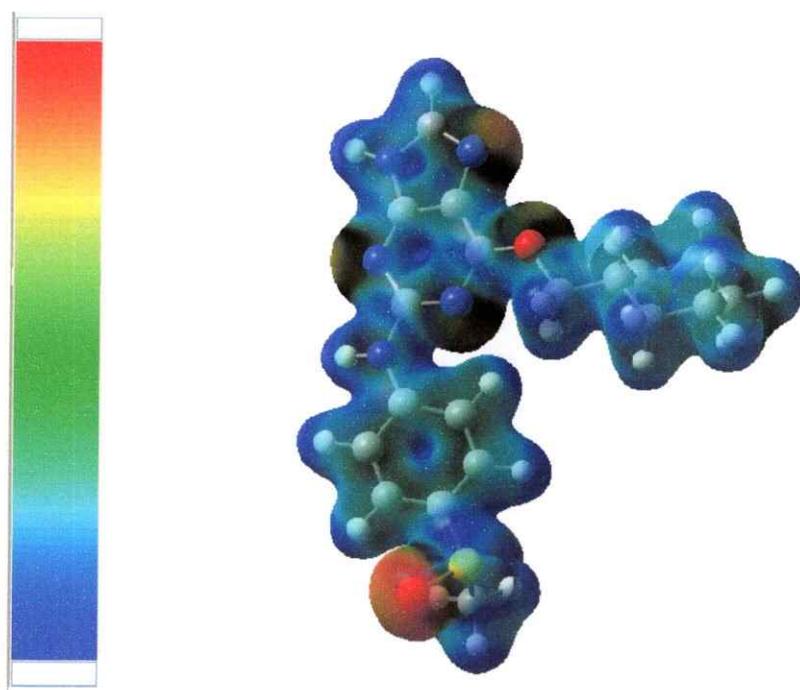
**Figura A.4.32.** Potencial Electrostático Molecular para el compuesto 26.



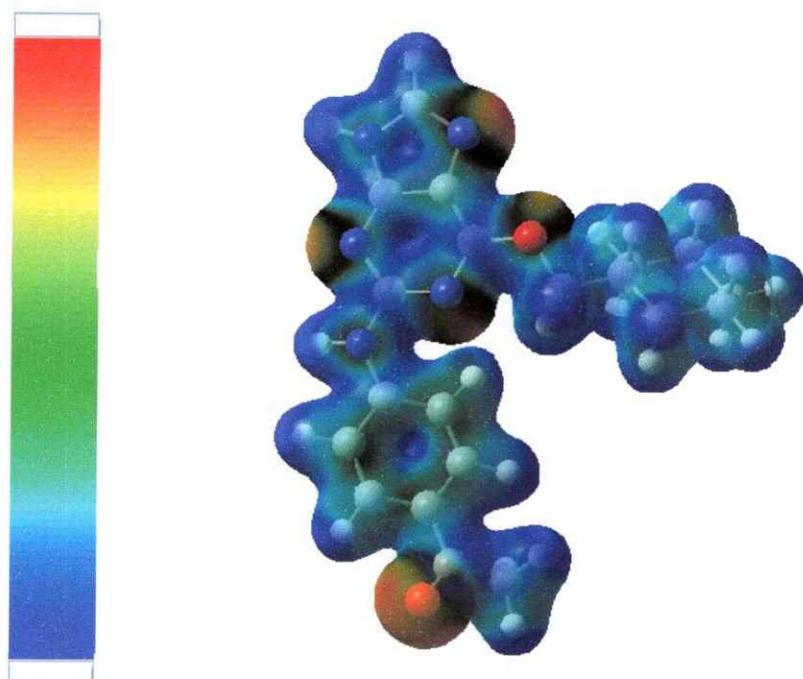
**Figura A.4.33.** Potencial Electrostático Molecular para el compuesto 29.



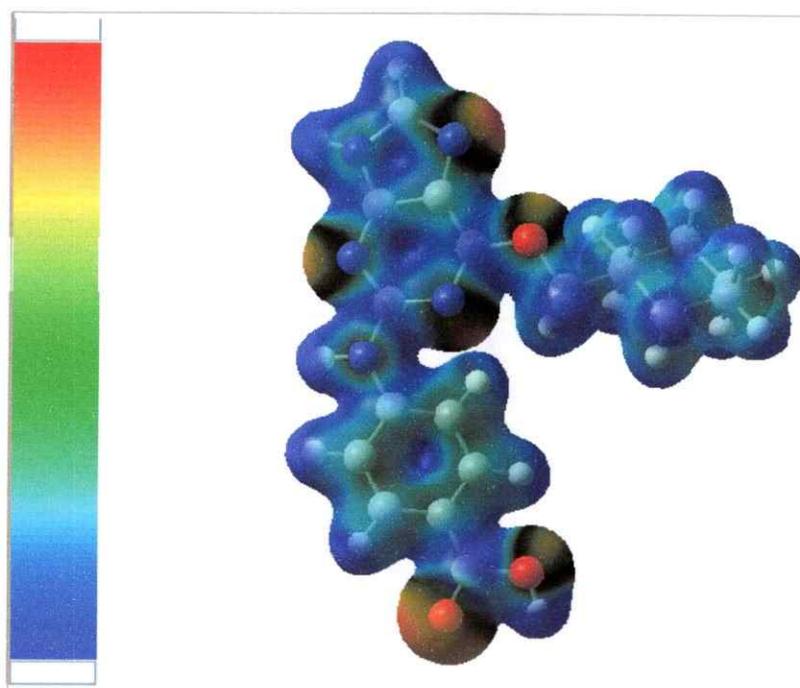
**Figura A.4.34.** Potencial Electrostático Molecular para el compuesto 30.



**Figura A.4.35.** Potencial Electrostático Molecular para el compuesto 31.



**Figura A.4.36.** Potencial Electrostático Molecular para el compuesto 38.



**Figura A.4.37.** Potencial Electrostático Molecular para el compuesto 39.

## ANEXO 2

### PUBLICACIONES GENERADAS DURANTE LA TESIS

Alzate-Morales, J. H., Contreras, R., Soriano, A., Tuñón, I. and Silla E. "A Computational Study of the Protein-Ligand Interactions in CDK2 Inhibitors: Using QM/MM Interaction Energy as a Predictor of the Biological Activity" *Biophysical Journal*. To appear in Vol. 92, Issue 2, January 2007, pp. 1-10.

Alzate-Morales, J. H., Tiznado, W., Santos, J. C., Cárdenas, C. and Contreras, R. "Theoretical study on CDK2 inhibitors using a global softness obtained from the density of states" *Journal of Physical Chemistry B*. In first revision.

# A Computational Study of the Protein-Ligand Interactions in CDK2 Inhibitors: Using Quantum Mechanics/Molecular Mechanics Interaction Energy as a Predictor of the Biological Activity

Jans H. Alzate-Morales,\* Renato Contreras,\* Alejandro Soriano,<sup>†</sup> Iñaki Tuñon,<sup>†</sup> and Estanislao Silla<sup>†</sup>

Q1] \*Departamento de Química, Facultad de Ciencias, Universidad de Chile, Santiago, Chile; and <sup>†</sup>Departamento de Química Física/IcMol, Universidad de Valencia, Valencia, Spain

Q2] **ABSTRACT** We report a combined quantum mechanics/molecular mechanics (QM/MM) study to determine the protein-ligand  
Q3] interaction energy between CDK2 (cyclin-dependent kinase 2) and five inhibitors with the N<sup>2</sup>-substituted 6-cyclohexylmethoxypurine scaffold. The computational results in this work show that the QM/MM interaction energy is strongly correlated to the biological activity and can be used as a predictor, at least within a family of substrates. A detailed analysis of the protein-ligand structures obtained from molecular dynamics simulations shows specific interactions within the active site that, in some cases,  
Q4] have not been reported before to our knowledge. The computed interaction energy gauges the strength of protein-ligand interactions. Finally, energy decomposition and multiple regression analyses were performed to check the contribution of the electrostatic and van der Waals energies to the total interaction energy and to show the capabilities of the computational model to identify new potent inhibitors.

## INTRODUCTION

The cyclin-dependent kinases (CDKs) play an essential role in regulating eukaryotic cell-cycle progression (1). These protein kinases are generally categorized into G1, S, and G2 phase regulators because they are present at various checkpoints in the cell cycle (2). As their name suggests, the CDKs are dependent on larger proteins known as cyclins for activation. Only as a complex can these proteins regulate cell growth and DNA synthesis properly. Partial activation occurs upon binding of these positive regulatory subunits; complete activation requires phosphorylation of the CDK subunit by the CDK-activating kinase at the conserved threonine residue. The CDK considered in this study is CDK2, which combines with cyclin E at an S-phase checkpoint known as the restriction point. In the same way, the completion of the S-phase depends on a complex of CDK2 and cyclin A (2).

The activity of the CDK-cyclin complex can be reduced by at least two major mechanisms: the phosphorylation of the CDK subunit at inhibitory sites and the binding of the specialized protein inhibitors known as CKIs or CDK inhibitors. These inhibitors compete with ATP (adenosine 5'-triphosphate) for binding to the CDK active site. However, in some cancer cells it has been shown that the CKIs are underexpressed, and medicinal chemists have made numerous efforts to replace the CKIs with synthetic inhibitors (3). Considerable progress has been made in the identification of pharmacologic agents targeting the CDKs (4). A large number of ATP-competitive inhibitors from a variety of chemical classes have been identified (2,5–7). Among noteworthy attempts to

produce such inhibitors are a series of compounds based on O<sup>6</sup>-cyclohexylmethylpurine or NU2058 (8) (see Fig. 1), which are competitive inhibitors of both CDK1 and CDK2 with respect to ATP. They also display good selectivity over CDK4 (9). Several authors, with the aid of iterative structure-based drug design, have carefully explored this scaffold. In this way, it has been possible to identify three kinds of characteristic interactions for this class of compounds within the active site of some CDKs. The first is the presence of the triplet of hydrogen bonds formed between the different tested compounds and the hinge region in CDK1 and CDK2. This feature induces a different orientation of these compounds inside the active site of the enzyme with respect to other inhibitors such as flavopiridol and olomoucine, and of course it has direct consequences for enzyme and cell growth inhibition (8). A second characteristic for this family of compounds is that optimum binding occurs with a moderately sized aliphatic O<sup>6</sup> substituent that tightly packs against the hydrophobic patch presented by the glycine loop, centered on Val-18—an interaction promoted by the conformational constraints imposed in a cyclohexylmethyl or cyclohexenylmethyl ring.

Thus, the parent compound O<sup>6</sup>-cyclohexylmethylguanine (NU2058) is the preferred starting point for exploring other areas of the kinase active site (10). The third characteristic kind of interactions are those established with the so-called 'specificity surface'; this is with residues that lie outside the highly conserved ATP binding site cleft (11,12). Sequence differences between the different members of the CDK family exist in this region, and targeting them may afford selectivity. For example, CDK2 residues His-84, Gln-85, and Lys-89 are, respectively, an aspartate, a glutamine, and a threonine in CDK4. Studies on olomoucine (13) and roscovitine (14) have

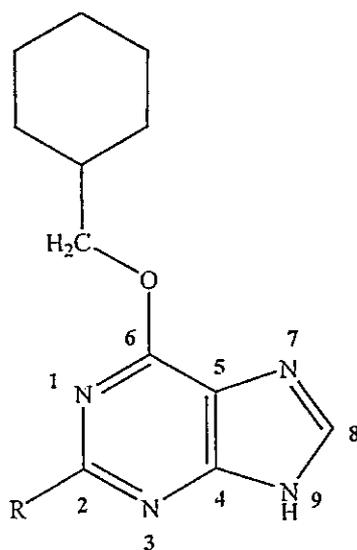
Submitted June 20, 2006, and accepted for publication October 6, 2006.

Address reprint requests to Jans H. Alzate-Morales, Tel.: 56-2-978-7272; Fax: 56-2-271-3888; E-mail: jalzate@ciq.uchile.cl; or Iñaki Tuñon, Fax: 34-96-386-4564; E-mail: ignacio.tunon@uv.es.

© 2007 by the Biophysical Society

0006-3495/07/01/1/10 \$2.00

doi: 10.1529/biophysj.106.091512



COMPOUND 2 R = 2-anilino

COMPOUND 3 R = 2-(4'-sulfamoylanilino)

COMPOUND 7 R = 2-methylamino

COMPOUND 9 R = 2-(1-methylethylamino)

COMPOUND 25 R = 2-(4'-hydroxyanilino)

FIGURE 1 Structures of the  $N^2$ -substituted  $O^6$ -cyclohexylmethylguanine derivatives of Hardcastle et al. (15). Numbering of atoms in the purine moiety is also displayed.

shown that large gains in potency, in addition to specificity, are possible by targeting this 'specificity surface'. Structural analysis had indicated that an aromatic ring at the  $N^2$  position of NU2058 would improve inhibitory activity against CDK1 and CDK2 (9). This was found to be the case with the resulting compound produced by maintaining the specificity for the inhibition of CDK1/2 over CDK4. Additional potency is conferred by the presence of a group capable of donating a hydrogen bond at the 4'-position in the aromatic ring. The resulting compound of all these previous structural-based investigations is NU6102 (compound 3 (Cp3) in Fig. 1). This inhibitor is selective and one of the most active CDK2 inhibitors described so far (9,15). Regardless of this significant progress in the structural-activity relationships for this kind of anticancer compounds, there remains the need to find a broader spectrum inhibitor that can, for example, selectively inhibit CDK1, CDK2, CDK4, and CDK6 at low nanomolar  $IC_{50}$  concentrations. The investigation of CDK2-ligand interactions can provide further insight for developing new compounds that are still required to achieve selective cytotoxicity and valuable anticancer activity.

Supplementary to a tremendous amount of experimental work, a few theoretical studies on CDKs, based on classical and quantum mechanical calculations, have also been reported in the literature. For instance, Cavalli et al. have developed force field parameters based on quantum chemical calculations on a relevant model system for simulating the binding of ATP to CDKs (16). Molecular dynamics (MD) simulation with the newly calculated parameters showed that CDK1 should be dynamically more stable than CDK2, implying that differences in protein structure flexibilities among CDKs should be considered in designing their selective inhibitors. More recent computational studies on CDK2 by Sims et al. have established a binding free energy model for the solvent (17). This approach proved to be successful in reproducing the relative inhibitory activities of flavopiridol analogs. On the basis of the validated utility of the computational approach, they suggested new putative inhibitors that were predicted to be more potent than the lead compound. In a later work, these same authors implemented the previously developed, continuum solvent-based charge optimization model with a simple, quadratic programming algorithm and the UHBD Poisson-Boltzmann solver (18). This method allowed them to compute the best set of point charges for a ligand or ligand region based on the ligand and receptor shape and the receptor partial charges by optimizing the binding free energy obtained from a continuum-solvent model (18). However, there are several limitations to such a model of binding free energy (17,19,20). In their particular case, they applied a fixed conformational model that assumes that the ligand and receptor undergo no conformational change upon binding.

Consideration of protein flexibility is indispensable for a critical evaluation of ligand-binding affinity (21–24), especially here because the conformational plasticity of the catalytic domain is a hallmark of protein kinases (25). As ever, a tradeoff between accuracy and computational expense is unavoidable. Empirical scoring functions designed for the fast estimation of binding affinities for high-throughput virtual screening often neglect conformational flexibility of the protein, entropic effects, and desolvation during the binding process (26). At the other end of the accuracy range are techniques based on MD simulations in explicit solvent that make use of free energy perturbation, thermodynamic integration (TI), and linear interaction energy methods (27,28). Given sufficient sampling, MD methods can be used to calculate the free energy of protein/ligand binding with high accuracy, including all entropic and solvent effects as well as receptor flexibility. The main problem associated with this methodology is the computational effort required to adequately sample the full configurational space and also to parametrize each new ligand assayed. Here we investigate a strategy to overcome these two computational bottlenecks based on the use of combined quantum mechanical/molecular mechanical (QM/MM) methods. These methods are now widely used for the analysis of enzymatic reactions (29–31). In combined QM/MM

methods, the ligand/substrate species is treated explicitly by a QM model (32–35), avoiding the derivation of parameters for each new ligand.

The protein and solvent environment is represented by MM force fields, which are computationally efficient. Treating the ligand quantum mechanically and the protein molecular mechanically has the additional advantage of the inclusion of ligand polarization upon binding. Besides these two important computational aspects, there exists an additional point that is worth mentioning. With a QM definition of the ligand, a better treatment of its accessible configurations could be expected, no matter what kind of functional group it may contain. It is important to mention that some functional groups are very difficult to parametrize, in the MM approach, when effects like electronic delocalization are present. There are some recent computational studies that successfully applied the combined QM/MM method to the study of protein-ligand interactions in the HIV-1 protease system (36) and in the trypsin system (37). In this work we propose an additional step in the use of QM/MM method for inhibitors analysis. We show below that the QM/MM interaction energy can be safely used to predict the biological activity of a particular inhibitor, at least when comparing within a given family of compounds. As mentioned before, binding free energies are computationally expensive, whereas interaction energies can be much more easily obtained. In this way two of the largest inconveniences in the use of MD simulations, the need of reparametrization for each new compound and the extensive sampling of different states required to evaluate free energy changes, can be conveniently overcome. In particular, we will show the effectiveness of the proposed strategy for the set of CDK2 inhibitors presented in Fig. 1.

## METHODS

### Simulations

All simulations were performed using the DYNAMO software package (38). The initial coordinates for the QM/MM calculations were taken from the x-ray crystal structure of the Nu6102 inhibitor bound to the fully activated Thr-160-phosphorylated CDK2-cyclin A complex (T160pCDK2-cyclinA) (Protein Data Bank code 1H1S) (9). The amino acids Arg-297 and Leu-298 were added to the CDK2 subunit of the initial crystal structure with the aid of the program Molden (39) to have a complete protein system. To mimic the aqueous environment, an equilibrated water box with sides of 79.5 Å, centered on the mass center of each inhibitor, was used to solvate the T160pCDK2-cyclinA-ligand system. Amino acid residues and water molecules placed more than 25 Å away from the center of the box were fixed during the simulations. The inhibitors inside the ATP active site were chosen to be the QM subsystem, described at the AM1 level (40), whereas the rest of the T160pCDK2-cyclinA complex and the water molecules of crystallization and those belonging to the solvent box constitute the MM subsystem were described using the OPLS force field (41) and the flexible TIP3P potential (42,43), respectively.

There are no bonds between the QM and the MM subsystems; so we do not need to consider any special treatment to complete the valence of the frontier quantum atoms (44). The remaining systems studied were obtained from this initial model by simply deleting and replacing a few atoms of the inhibitor Nu6102 (see Fig. 1). Initially, the hydrogen atoms of the system were relaxed using the conjugate gradient subroutine implemented in the

DYNAMO program (45). Then the full system was minimized up to a gradient tolerance of 1.0 kJ/mol. Subsequently the system was heated up to 300 K by a sequence of MD simulations. Afterwards the system was further equilibrated during additional 100 ps using the NVT ensemble at 300 K. The production run consisted of an MD simulation of 250 ps. In all cases we used a time step of 1 fs to solve the equation of motion and a switched cutoff distance of 13.5 Å. One protein-ligand configuration was saved each 15 time steps for a posteriori energetic and structural analysis. This procedure was applied to the five inhibitors presented in Fig. 1.

[AQ6]

### Energy decomposition

The energy for a QM/MM system can be obtained as

$$E = \langle \Psi | \hat{H}_0 | \Psi \rangle + \langle \Psi | \hat{V}_{\text{QM/MM}} | \Psi \rangle + E_{\text{MM}}, \quad (1)$$

where  $\hat{H}_0$  is the in vacuo Hamiltonian for the selected QM method,  $\Psi$  is the polarized wave function (that is, the wave function obtained in the presence of the MM field),  $E_{\text{MM}}$  represents the force field energy, and  $\hat{V}_{\text{QM/MM}}$  is the coupling operator between the QM and MM subsystems and includes an electrostatic and a van der Waals term. Eq. 1 can be written in a different form to define the QM/MM interaction energy. In our partition scheme the total energy is written as the sum of the in vacuo energy of the QM subsystem, the QM/MM interaction energy, and the MM energy (46):

$$E = \langle \Psi^0 | \hat{H}_0 | \Psi^0 \rangle + [\langle \Psi | \hat{H}_0 | \Psi \rangle - \langle \Psi^0 | \hat{H}_0 | \Psi^0 \rangle] + \langle \Psi | \hat{V}_{\text{QM/MM}} | \Psi \rangle + E_{\text{MM}}, \quad (2)$$

where  $\Psi^0$  stands for the unpolarized (gas phase) wave function of the QM subsystem. The first term on the right side represents the gas phase energy of the quantum subsystem (usually the ligand), the second term (the one inside the brackets) the polarization energy of the quantum subsystem wave function, and the third one the QM/MM interaction energy with the polarized wave function. The sum of these two last terms is the total QM/MM interaction energy. This energy can be further decomposed considering that the coupling operator between the QM and MM subsystems ( $\hat{V}_{\text{QM/MM}}$ ) includes an electrostatic and a van der Waals term. This last contribution is usually evaluated using the Lennard-Jones expression

$$E_{\text{vdw,QM/MM}} = \sum_{i \in \text{QM}} \sum_{j \in \text{MM}} 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{R_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{R_{ij}} \right)^6 \right], \quad (3)$$

where  $R_{ij}$  is the distance between the QM and MM interaction centers and  $\epsilon_{ij}$  and  $\sigma_{ij}$  are the Lennard-Jones parameters. As this expression does not involve electronic coordinates, this energy term does not need to be included in the SCF evaluation of the QM subsystem wave function.

The interaction energy can, then, be calculated as the energy difference between the full QM/MM system and the separated QM and MM subsystems. Taking into account that in nonpolarizable force fields the MM energy exactly cancels out, we can write

$$E_{\text{QM/MM}} = E - \langle \Psi^0 | \hat{H}_0 | \Psi^0 \rangle - E_{\text{MM}}, \quad (4)$$

which means that in vacuo single point energy calculation must be carried out for given nuclear configurations of the quantum atoms along the trajectory and then subtract this energy and the MM energy to the full QM/MM potential energy. Note that according to our definition of the polarization energy (term inside the brackets in Eq. 2) as the gas phase energy difference between the polarized and unpolarized wave functions, this is a positive contribution. A different energy decomposition scheme was used in Hensen et al. (36). Finally, it is also worth noting that the set of configurations over which the average is carried out depends on the temperature at which the simulation is performed. So, the average interaction energy is in fact temperature dependent.

[AQ5]

## RESULTS AND DISCUSSION

In this section we first present the structural features of each of the CDK2-inhibitor complexes. Then we analyze the interaction energies and the relationship to the biological activity (measured as  $IC_{50}$ ). Both geometrical data and energy terms have been averaged during the production run, and the standard deviation is provided as a number in parenthesis accompanying the average value.

### Ligand-active site interactions: compounds 2 (Nu6094) and 25

It is well known that several compounds such as purvalanols (47) and indirubins (48,49) pack aromatic moieties with the specificity surface of CDK2. Compound 2 (**Cp2**) was initially synthesized taking this property into account. Then, according to Davies et al., the aniline group of **Cp2** projects out of the adenine site through a largely hydrophobic tunnel constituted of the side chains of Phe-82 and Ile-10 and packs against the kinase surface, forming a  $\pi$ - $\pi$  stacking interaction with the peptide backbone between Gln-85 and Asp-86 (9). The final structure obtained from the MD simulations is shown in Fig. 2. As a common feature for the five inhibitors analyzed here, a triplet of hydrogen bonds is formed between the purine ring and the hinge region of CDK2: NH-9 acts as a hydrogen-bond donor to the backbone carbonyl group of Glu-81, and N-3 and 2-NH<sub>2</sub> accept and donate a hydrogen bond to the backbone carbonyl and amide groups of Leu-83, respectively. The averaged distances for the triplet of hydrogen bonds for each ligand are given in Table 1, being the values that are quite similar to the five compounds. From the analysis of geometrical data from the MD simulation, we have found other weak interactions between the **Cp2** and the protein, the C-H $\cdots$  $\pi$  weak interaction between the  $\alpha$ -carbon hydrogen of Gln-85 and the center of the aniline ring of ligand being the most significant. The averaged distance from the hydrogen atom to the centroid of the ring is  $\sim 2.90$  Å. There exists another weak interaction of the type C-H $\cdots$ O=C between one of the aniline ring hydrogens and the carboxylate group of Asp-86 (**Cp2** Ph-H to Asp-86 OD2 = 3.35 Å ( $\pm 0.49$  Å)). It is worth mentioning that the side chain of Lys-89 is tilted away from the aniline ring of the ligand. The NH<sub>3</sub><sup>+</sup> group of this residue makes a stabilizing hydrogen bond with the carbonyl side-chain group of Gln-85 (Fig. 2).

The potencies of the 4'-substituted C<sup>2</sup>-anilino-O<sup>6</sup>-cyclohexylmethyl-purine series were evaluated by Hardcastle et al.

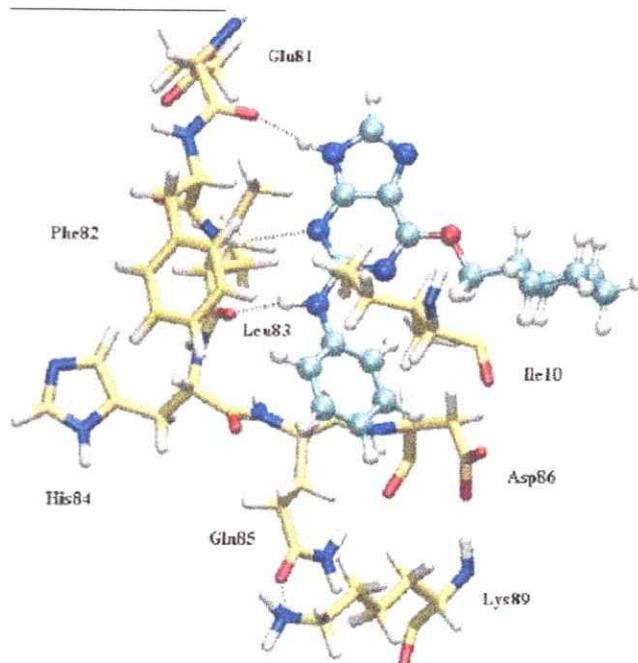


FIGURE 2 Snapshot of the CDK2-Cp2 structure. Selected residues of the CDK2 active site are shown in licorice representation, and Cp2 is presented in ball-and-stick representation, with carbon atoms colored in yellow (CDK2) and cyan (inhibitor). The hydrogen bonds are depicted as black dashed lines.

(15). They found that 4'-hydroxy substitution in this series generated compound 25 (**Cp25**) which was substantially more potent than the parent aniline **Cp2**. This result suggested that there was a hydrogen-bond acceptor in the region of the protein accessed by the 2-anilino group. An obvious candidate was the Asp-86 residue. The data obtained from our MD simulations (see Fig. 3) show the hydroxyl group of **Cp25** interacting with the carboxylate moiety of Asp-86 via a water-mediated hydrogen bond (**Cp25** O-TIP3-Asp-86 OD2 = 3.29 Å/2.62 Å ( $\pm 0.42$  Å/ $\pm 0.10$  Å)). This result is in agreement with the crystallographic data reported by Davies et al., although the hydrogen-bond distances are different from those quoted by them (9). We observed two additional hydrogen bonds between the hydroxyl group and the Lys-89 residue. One of them is formed between the oxygen atom of the hydroxyl group and the NH<sub>3</sub><sup>+</sup> group of Lys-89 (**Cp25** O-Lys-89 NZ = 3.79 Å ( $\pm 0.66$  Å)). The hydroxyl group also interacts with Lys-89 via a water-mediated hydrogen bond (**Cp25** O-TIP21-Lys-89 NZ = 3.31 Å/2.94 Å ( $\pm 0.37$  Å/ $\pm 0.51$  Å)).

TABLE 1 Hydrogen-bond-averaged distances (Å) between the purine ring of each ligand and the hinge region of the CDK2 active site

Fragment	Cp 2	Cp 25	Cp 3	Cp 9	Cp 7
Glu-81 C=O—H-N9	1.96 ( $\pm 0.16$ )	1.88 ( $\pm 0.14$ )	1.92 ( $\pm 0.15$ )	1.90 ( $\pm 0.15$ )	1.91 ( $\pm 0.15$ )
Leu-83 N-H—N3	2.49 ( $\pm 0.25$ )	2.62 ( $\pm 0.21$ )	2.46 ( $\pm 0.21$ )	2.60 ( $\pm 0.23$ )	2.72 ( $\pm 0.30$ )
Leu-83 C=O—H-N2	1.94 ( $\pm 0.17$ )	1.94 ( $\pm 0.17$ )	1.90 ( $\pm 0.17$ )	1.97 ( $\pm 0.19$ )	2.01 ( $\pm 0.21$ )

The values in parentheses correspond to the standard deviations.

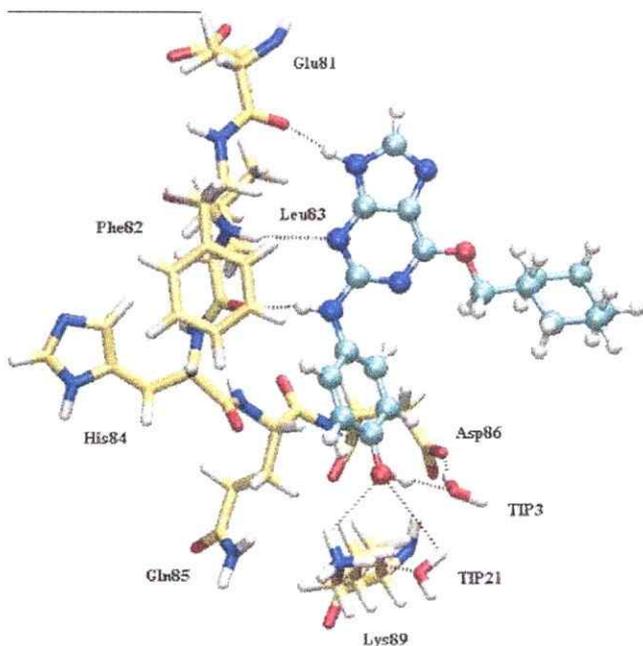


FIGURE 3 Snapshot of the CDK2-Cp25 structure. Representation is made as in Fig. 2.

### Compound 3 (Nu6102)

This compound presents a particular interaction pattern within the active site of CDK2 due to the presence of the sulfonamide group. Hardcastle et al. reported that this compound interacts with Asp-86 through the  $\text{NH}_2$  group of the sulfonamide, which donates a hydrogen bond to the side-chain oxygen of Asp-86 (**Cp3** N26 to Asp-86 OD2 = 2.9 Å) and through one sulfonamide oxygen that accepts a hydrogen bond from the backbone nitrogen of Asp-86 (**Cp3** O24 to Asp-86 NH = 3.1 Å) (15). They have further suggested that the high activity of **Cp3** is due to the two additional H-bonds formed with Asp-86. However, we have found that the interactions formed for this compound inside the active site of CDK2 are different from those reported before. A representative snapshot of this protein-inhibitor complex as obtained in our MD simulation is provided in Fig. 4. In addition to the triplet of hydrogen bonds established with the hinge region, reported in Table 1, we have found other CDK2-Cp3 hydrogen bonds. The  $\text{NH}_2$  group of the sulfonamide donates a hydrogen bond to each of the side-chain oxygen atoms of Asp-86 (**Cp3** O<sub>2</sub>S-N to Asp-86 OD1 = 3.78 Å ( $\pm 0.40$  Å) and **Cp3** O<sub>2</sub>S-N to Asp-86 OD2 = 3.38 Å ( $\pm 0.45$  Å)). One sulfonamide oxygen atom accepts a hydrogen bond from the  $\text{NH}_3^+$  side-chain group of Lys-89 (**Cp3** S=O to Lys-89 NZ = 2.80 Å ( $\pm 0.15$  Å)).

According to Davies et al. (9), although the **Cp3** aniline group packed closely with the specificity surface, the sulfonamide group did not form the proposed hydrogen bond with Lys-89. We have found that this interaction does occur and that the sulfonamide group is positioned in a geometry that

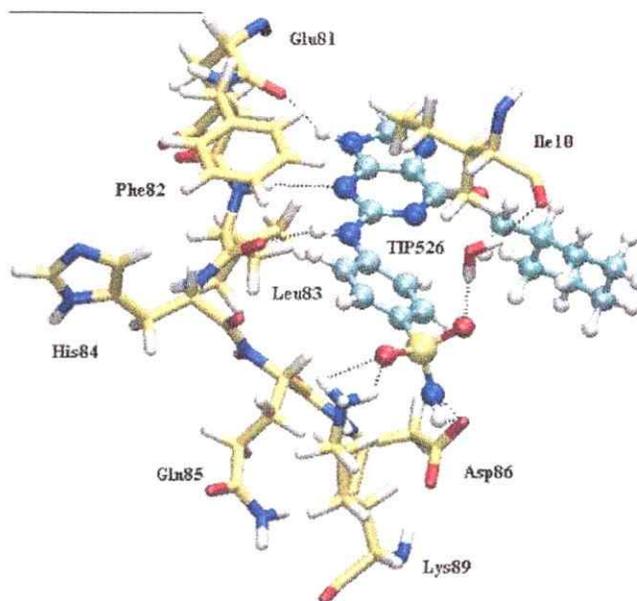


FIGURE 4 Snapshot of the CDK2-Cp3 structure. Representation is made as in Fig. 2.

facilitates this unreported protein-ligand hydrogen bond. Finally, the last hydrogen-bond interaction found for this compound, and not yet reported, occurs between the other sulfonamide oxygen atom and the backbone carbonyl group of residue Ile-10. The hydrogen bond is this time mediated by a water molecule from the solvent water box (**Cp3** S=O-TIP526-Ile-10 C=O = 3.07 Å/3.39 Å ( $\pm 0.66$  Å/ $\pm 0.76$  Å)).

### Compounds 9 and 7

These compounds belong to the nonaromatic series studied by Hardcastle et al. (15). They proved to be more potent CDK1 and CDK2 inhibitors than the parent compound Nu2058 (8), and both form a triplet of hydrogen bonds within the CDK2 ATP binding site (see Table 1). Compound **9** (**Cp9**) is the isopropyl  $\text{N}^2$ -monosubstituted derivative of the series studied and forms a weak C-H $\cdots\pi$  interaction with the peptide backbone between Gln-85 and Asp-86 through one of the isopropyl hydrogen atoms. The protein-ligand complex is shown in Fig. 5. The isopropyl group establishes a nonpolar interaction with the side chain of Ile-10 and a weak hydrogen-bond interaction of the type C-H $\cdots$ O=C, with one of the side-chain oxygens of Asp-86 (**Cp9** isopropyl-H to Asp-86 OD2 = 3.78 Å ( $\pm 0.49$  Å)). On the other hand, compound **7** (**Cp7**) is the methyl  $\text{N}^2$ -monosubstituted derivative in this series. This compound establishes the typical triplet of hydrogen bonds with the hinge region of CDK2 (Table 1) and a nonpolar interaction with the side chain of residue Ile-10 (Fig. 6). The lack of a bulky group in this ligand, like the isopropyl in **Cp9**, causes a diminished interaction with Ile-10.

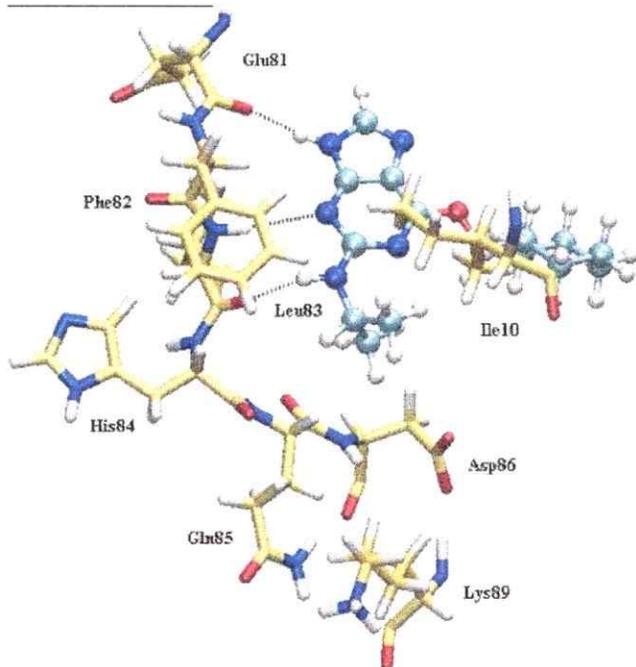


FIGURE 5 Snapshot of the CDK2-Cp9 structure. Representation is made as in Fig. 2.

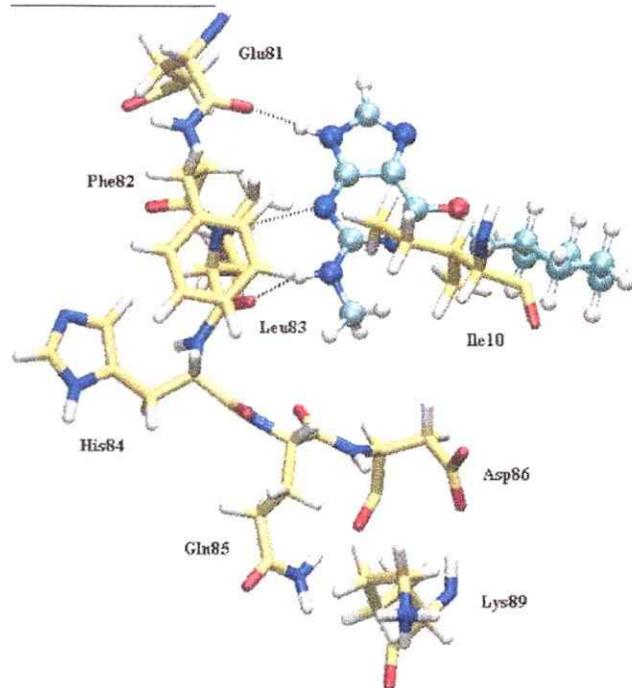


FIGURE 6 Snapshot of the CDK2-Cp7 structure. Representation is made as in Fig. 2.

### Analysis of interaction energies

Averaged interaction energies for each of the inhibitors with CDK2, together with their electrostatic and van der Waals components, are reported in Table 2. The biological activity is also provided as the common logarithm of  $IC_{50}$  (15). The potency of each ligand is obviously related to its binding free energy but assuming that, within a family of compounds, solvation/desolvation energies, enzyme deformation energy, and entropic changes are proportional to the magnitude of the interactions established between the inhibitor and the protein, one may then express the activity as a function of the interaction energy:

$$\text{Log}IC_{50} = \alpha \Delta E_{\text{QM/MM}} + \beta. \quad (5)$$

The result of a least-square fit of Eq. 5 using data of Table 2 is shown in Fig. 7. The high value obtained for the correlation coefficient ( $R = 0.96$ ) makes us confident in our simulation procedure. In addition, it clearly indicates that interaction energies can be safely used as a predictor of the

biological activity of a given inhibitor, at least when comparing them with a series of compounds belonging to the same family. The interest in this kind of correlations is obvious considering the large difference in computational effort needed to evaluate interaction energies or binding free energies. The use of QM/MM procedures has the additional advantage of avoiding extensive reparametrization for each new ligand considered in the study. An interesting practical conclusion can also be drawn from the study of the quality of the fit as a function of the length of the sampling used in the evaluation of the interaction energies. If instead of using the full 250-ps simulation to average the interaction energy, one averages them only over the last 100 ps, the correlation coefficient is slightly reduced up to 0.95, indicating that the sampling, and then the computational effort, could be reduced without a significant loss of quality.

It is also interesting to relate the different interaction energies and their decomposition to the structural features of the different inhibitor-protein complexes. It can be seen from Table 2 that the magnitude of the electrostatic contribution is

TABLE 2 Averaged QM/MM interaction energies and its components (kcal/mol) for the CDK2 inhibitors studied

	Compound 2	Compound 25	Compound 3	Compound 9	Compound 7
Interaction energy	-73.4 ( $\pm 2.9$ )	-86.2 ( $\pm 3.5$ )	-124.5 ( $\pm 8.2$ )	-70.2 ( $\pm 4.0$ )	-65.5 ( $\pm 3.1$ )
Electrostatic energy	-17.5 ( $\pm 2.4$ )	-30.2 ( $\pm 3.4$ )	-65.8 ( $\pm 9.0$ )	-23.4 ( $\pm 3.5$ )	-22.8 ( $\pm 3.3$ )
vdW energy	-55.9 ( $\pm 2.2$ )	-56.1 ( $\pm 2.5$ )	-58.6 ( $\pm 3.0$ )	-46.8 ( $\pm 2.4$ )	-42.7 ( $\pm 2.4$ )
Log $IC_{50}$	-0.013	-1.161	-2.268	0.079	0.699

Biological activities are reported as the common logarithm of  $IC_{50}$  (micromolar) (15).

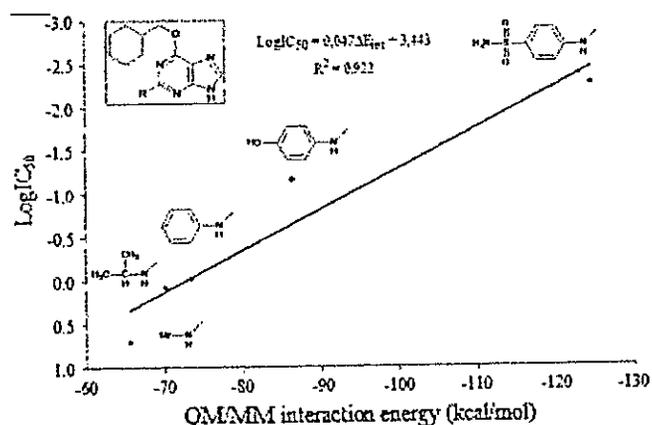


FIGURE 7 Graph of  $\text{LogIC}_{50}$  versus QM/MM interaction energy (kcal/mol) for the five  $\text{N}^2$ -substituted 6-cyclohexylmethylguanine derivatives studied and equation obtained by a least-square fit.

similar for compounds 9 and 7, as result of the very similar hydrogen-bond interactions presented by these ligands with the hinge region in CDK2. The van der Waals interaction energy is different for these two compounds due to the difference in the substitution pattern at the  $\text{N}^2$  position of the purine ring. The main consequence arising from this difference is that Cp9 bears a more polarizable isopropyl group at the  $\text{N}^2$  position, thereby promoting a more favorable interaction with the Ile-10 residue than Cp7, which bears a methyl group at that position. This example clearly shows that the activity can be improved by the optimization of the van der Waals interactions. The opposite is found for Cp2 and Cp25. These compounds show similar van der Waals contributions to the interaction energy, and their electrostatic energy values are different (Table 2). The difference in the electrostatic energy for these two compounds can be attributed to a significant interaction between the 4'-substituent group in the phenyl ring of the aniline group and some amino acids in the specificity surface. For Cp2, the larger van der Waals energy contribution (in absolute value), when compared to the corresponding values for Cp9 and Cp7, is in agreement with the expected interaction pattern with the specificity surface at the active site of CDK2. That means that the presence of an aniline ring in the C-2 position of the purine ring is more favorable than a nonaromatic group like methyl (Cp7) or ethyl (Cp9).

The electrostatic energy component for Cp2 is smaller, in absolute value, than the corresponding values found for Cp7 and Cp9. This can be attributed to a less favorable interaction between the purine ring of this compound and the hinge region of CDK2. However, this low contribution is compensated for by an increase in the van der Waals interaction energy. Cp25, on the other hand, shows an increase in the electrostatic energy component due to the hydrogen-bond interactions formed by the  $-\text{OH}$  group in the 4'-position of the aniline ring with the residues of the specificity surface of the enzyme. It is also interesting to mention that Cp3 is the

only one that has an absolute value of the electrostatic energy component greater than the van der Waals term. This result can be attributed to the multiple hydrogen-bond interactions that this compound can establish with residues Asp-86, Lys-89, and Ile-10, which are present in the CDK2 specificity surface. Comparing the van der Waals energy value for this compound with those evaluated for Cp2 and Cp25, we can conclude that the van der Waals energy component has a similar weight in these compounds, but this energy component is slightly improved due to the interaction formed with the amino acids in the specificity surface and, consequently, to the more effective packing with the hydrophobic surface within the active site of the CDK2. The final result for Cp3 is an improved interaction and then a larger inhibitor activity.

Finally, to evaluate if our interaction energy partition is a reliable model to represent the interaction between each ligand and the active site of CDK2, a two-variable regression analysis was performed. The two independent variables are the electrostatic energy and the vdW energy components, and the dependent variable is the logarithm of the  $\text{IC}_{50}$  inhibitory concentration obtained from the experiments. The resulting regression equation is

$$\begin{aligned} \text{LogIC}_{50} &= \gamma + \alpha * \Delta E_{\text{elec}} + \beta * \Delta E_{\text{vdw}} \\ &= 4.973 + 0.038 * \Delta E_{\text{elec}} + 0.082 * \Delta E_{\text{vdw}}, \quad (6) \end{aligned}$$

where  $\Delta E_{\text{elec}}$  and  $\Delta E_{\text{vdw}}$  are the electrostatic and the van der Waals interaction energy components, respectively. The standard deviation values for the parameters  $\gamma$ ,  $\alpha$ , and  $\beta$  involved in the equation above are 1.46, 0.011, and 0.031, respectively. This equation can be used to predict  $\text{LogIC}_{50}$  values, which are quoted in Table 3. We have plotted these predicted  $\text{LogIC}_{50}$  values against the experimental  $\text{LogIC}_{50}$  values. The comparison yields a very good correlation coefficient  $R = 0.976$  (see Fig. 8). Interestingly, Eq. 6 displays quite different values for the slopes associated with the electrostatic and van der Waals interaction energies, suggesting that an improvement of this last contribution could be more effective, in terms of inhibition, than a similar energetic improvement in the electrostatic component. Obviously, another question is that obtaining differences in the van der Waals interaction energy may be more difficult than for the electrostatic part, as suggested by the ranges covered by these

TABLE 3 Results of multiple regression analysis (see Eq. 5) for the compounds belonging to the  $\text{N}^2$ -substituted-6-cyclohexylmethoxypurine family studied

Compound	Experimental $\text{IC}_{50}$ ( $\mu\text{M}$ )	Experimental $\text{LogIC}_{50}$ values	Calculated $\text{IC}_{50}$ values ( $\mu\text{M}$ )	Calculated $\text{LogIC}_{50}$ values
3	0.0054 ( $\pm 0.0010$ )	-2,268	0.0043	-2,368
25	0.069 ( $\pm 0.001$ )	-1,161	0.158	-0,800
2	0.970 ( $\pm 0.030$ )	-0,013	0.497	-0,303
9	1.200 ( $\pm 0.200$ )	0,079	1.680	0,225
7	5.000 ( $\pm 0.300$ )	0,699	3.820	0,582

[AQ10]

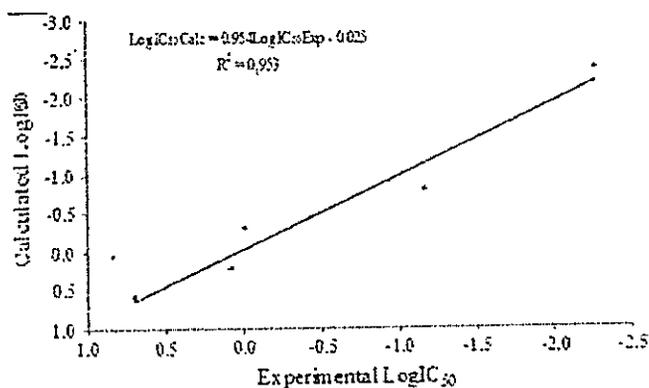


FIGURE 8 Graph of calculated  $\text{LogIC}_{50}$  values (obtained by means of Eq. 5) versus experimental  $\text{LogIC}_{50}$  values for the five compounds studied and equation obtained by least-square fit.

two components of the interaction energies in the five cases analyzed here. As shown in Table 2, the maximum difference in the van der Waals interaction energy among the inhibitors studied in this work is 16.9 kcal/mol, whereas in the case of the electrostatic interaction energy this maximum difference amounts up to 48.3 kcal/mol.

We have also tried to correlate  $\text{LogIC}_{50}$  versus  $\Delta E_{\text{elec}}$  or  $\Delta E_{\text{vdw}}$  separately. The correlation coefficients obtained are significantly worse, 0.89 and 0.82, respectively. This suggests that correlations using only one of those components are not sufficient to describe the behavior of the inhibitors within the CDK2 active site due to the important role that each property has in the global interaction between the inhibitor and the amino acids at the active site. Instead of using those descriptors separately, the whole interaction energy (Eq. 5) or a multiple regression analysis with respect to both terms (Eq. 6) can be used to predict the  $\text{IC}_{50}$  for a new compound structurally related to the scaffold studied here in a quantitative manner.

## CONCLUSIONS

In this study we have demonstrated that the combined QM/MM methods provide a useful approximation to calculate the interactions established at the CDK2 active site by a series of compounds belonging to the  $\text{N}^2$ -substituted- $\text{O}^6$ -cyclohexylmethoxypurine family. This method allows us to describe with some detail the nature of the specific interactions involved in the protein-ligand binding. This fact makes the method extensible to other groups of CDK2 inhibitory molecules and, in fact, to other interesting biological systems. Within the approach used here, it is further possible to obtain an acceptable and representative picture of the real system in which the evaluation of the different contributions from the enzymatic and solvent environments to the protein-ligand binding in the active site of the protein becomes a simpler task. The consideration of flexibility of almost the whole protein and the ligand in the calculations performed with the

combined QM/MM methods allows us to achieve a wider sampling of the configurational space. A huge set of possible structures at a relevant temperature in the MD simulations can be incorporated.

With respect to the specific interactions formed inside the CDK2 active site by each of the compounds studied, we can conclude that the enhanced potency observed for Nu6102 (Cp3) results from the hydrogen-bond interactions formed between the sulfonamide group and residues Asp-86, Lys-89, and Ile-10. There is an additional gain in potency that depends on favorable packing between the aniline aromatic ring and the hydrophobic surface in the CDK2 binding site. It is worth mentioning that this is the first time to our knowledge that the interactions of a compound with Lys-89 are [AQ7] reported.

The statistical validation of the proposed method was performed by means of multiple regression analysis. This analysis permits us to conclude that there exists a high degree of confidence in the data presented for the calculated interaction energy components of each inhibitor using the combined QM/MM described before. Taking into account the values of the two variables proposed and their correlation with the biological activity values reported in the literature, it seems possible to predict the  $\text{IC}_{50}$  value for a new compound. The evaluation of the averaged interaction energies has allowed us to show the strong correlation existing with the biological activity of this series of inhibitors. In fact, the QM/MM interaction energy could be used as a predictor of the biological activity, at least, within this family of compounds. The use of QM/MM interaction energies has a double advantage: i) extensive reparametrization of new ligands is avoided, and ii) the computational effort is considerably smaller than for the evaluation of binding free energies. Energy decomposition shows that in all compounds, except Cp3, the major contribution to the total interaction energy is made by the van der Waals energy component. In some compounds, the hydrogen bonds facilitate the interactions with the hydrophobic surface of the enzyme.

J.H.A.M. thanks DAAD (Deutscher Akademischer Austausch Dienst, Germany) for financial support through a doctoral fellowship and Departamento de Postgrado y Posttulo (U. de Chile) for partial research grant PG/95/2004, and Millennium Nucleus for Applied Quantum Mechanics and Computational Chemistry, grant P02-004-F, Mideplan-Conicyt. We are indebted to DGI for project BQU2003-04168 and Generalitat Valenciana [AQ8] for projects GV04B-021, GV04B-131, and GRUPOS04/08, which supported this research.

## REFERENCES

1. Morgan, D. O. 1995. Principles of CDK regulation. *Nature*. 374:131-134.
2. Sielecki, T. M., J. F. Boylan, P. A. Benfield, and G. L. Trainor. 2000. Cyclin-dependent kinase inhibitors: useful targets in cell cycle regulation. *J. Med. Chem.* 43:1-18.
3. Losiewicz, M. D., B. A. Carlson, G. Kaur, E. A. Sausville, and P. J. Worland. 1994. Potent inhibition of CDC2 kinase activity by the flavonoid L86-8275. *Biochem. Biophys. Res. Commun.* 201:589-595.

4. Senderowicz, A. M., and E. A. Sausville. 2000. Preclinical and clinical development of cyclin-dependent kinase modulators. *J. Natl. Cancer Inst.* 92:376–387.
5. Hardcastle, I. R., B. T. Golding, and R. J. Griffin. 2002. Designing inhibitors of cyclin-dependent kinases. *Annu. Rev. Pharmacol. Toxicol.* 42:25–348.
6. Knockaert, M., P. Greengard, and L. Meijer. 2002. Pharmacological inhibitors of cyclin-dependent kinases. *Trends Pharmacol. Sci.* 23: 417–425.
7. Toogood, P. L. 2002. Progress toward the development of agents to modulate the cell cycle. *Curr. Opin. Chem. Biol.* 6:472–478.
8. Arris, C. E., F. T. Boyle, A. H. Calvert, N. J. Curtin, J. A. Endicott, E. F. Garman, A. E. Gibson, B. T. Golding, S. Grant, R. J. Griffin, P. Jewsbury, L. N. Johnson, et al. 2000. Identification of novel purine and pyrimidine cyclin-dependent kinase inhibitors with distinct molecular interactions and tumor cell growth inhibition profiles. *J. Med. Chem.* 43:2797–2804.
9. Davies, T. G., J. Bentley, C. E. Arris, F. T. Boyle, N. J. Curtin, J. A. Endicott, A. E. Gibson, B. T. Golding, R. J. Griffin, I. R. Hardcastle, P. Jewsbury, L. N. Johnson, et al. 2002. Structure-based design of a potent purine-based cyclin-dependent kinase inhibitor. *Nat. Struct. Biol.* 9:745–749.
10. Gibson, A. E., C. E. Arris, J. Bentley, F. T. Boyle, N. J. Curtin, T. G. Davies, J. A. Endicott, B. T. Golding, S. Grant, R. J. Griffin, P. Jewsbury, L. N. Johnson, et al. 2002. Probing the ATP ribose-binding domain of cyclin-dependent kinases 1 and 2 with O<sup>6</sup>-substituted guanine derivatives. *J. Med. Chem.* 45:3381–3393.
11. Hanks, S. K., and T. Hunter. 1995. The eukaryotic protein kinase super family: kinase (catalytic) domain structure and classification. *FASEB J.* 9:576–596.
12. Gray, N., L. Devivaud, C. Doering, and L. Meijer. 1999. ATP-site directed inhibitors of cyclin dependent kinases. *Curr. Med. Chem.* 6: 850–875.
13. Schulze-Gahmen, U., J. Brandsen, H. D. Jones, D. O. Morgan, L. Meijer, J. Vesely, and S. H. Kim. 1995. Multiple modes of ligand recognition: crystal structures of cyclin dependent protein kinase 2 in complex with ATP and two inhibitors, olomoucine and isopentenyladenine. *Proteins: Struct. Funct. Genet.* 22:378–391.
14. De Azevedo, W. F., S. Leclerc, L. Meijer, L. Havlicek, M. Strnad, and S. H. Kim. 1997. Inhibition of cyclin-dependent kinases by purine analogues. Crystal structure of human CDK2 complexed with roscovitine. *Eur. J. Biochem.* 243:518–526.
15. Hardcastle, I. R., C. E. Arris, J. Bentley, F. T. Boyle, Y. Chen, N. J. Curtin, J. A. Endicott, A. E. Gibson, B. T. Golding, R. J. Griffin, P. Jewsbury, J. Menyerol, et al. 2004. N<sup>2</sup>-substituted O<sup>6</sup>-cyclohexylmethylguanine derivatives: potent inhibitors of cyclin-dependent kinases 1 and 2. *J. Med. Chem.* 47:3710–3722.
16. Cavalli, A., C. Dezi, G. Folkers, L. Scapozza, and M. Recanatini. 2001. Three-dimensional model of the cyclin-dependent kinase 1 (CDK1): ab initio active site parameters for molecular dynamics studies of CDKs. *Proteins: Struct. Funct. Genet.* 45:478–485.
17. Sims, P. A., C. F. Wong, and J. A. McCammon. 2003. A computational model of binding thermodynamics: the design of cyclin-dependent kinase 2 inhibitors. *J. Med. Chem.* 46:3314–3325.
18. Sims, P. A., C. F. Wong, and J. A. McCammon. 2004. Charge optimization of the interface between protein kinases and their ligands. *J. Comput. Chem.* 25:1416–1429.
19. Gould, C., and C. F. Wong. 2002. Designing specific protein kinase inhibitors: insights from computer simulations and comparative sequence/structure analysis. *Pharmacol. Ther.* 93:169–178.
20. Wong, C. F., P. H. Hünenberger, P. Akamine, N. Narayana, T. Diller, J. A. McCammon, S. Taylor, and N. H. Xuong. 2001. Computational analysis of PKA-balanol interactions. *J. Med. Chem.* 44:1530–1539.
21. Teague, S. J. 2003. Implications of protein flexibility for drug discovery. *Nat. Rev. Drug Discov.* 2:527–541.
22. Lin, J. H., A. L. Perryman, J. R. Schames, and J. A. McCammon. 2002. Computational drug design accommodating receptor flexibility: the relaxed complex scheme. *J. Am. Chem. Soc.* 124:5632–5633.
23. Carlson, H. A. 2002. Protein flexibility and drug design: how to hit a moving target. *Curr. Opin. Chem. Biol.* 6:447–452.
24. Carlson, H. A., and J. A. McCammon. 2000. Accommodating protein flexibility in computational drug design. *Mol. Pharmacol.* 57: 213–218.
25. Huse, M., and J. Kuriyan. 2002. The conformational plasticity of protein kinases. *Cell.* 109:275–282.
26. Gohlke, H., and G. Klebe. 2001. Statistical potentials and scoring functions applied to protein-ligand binding. *Curr. Opin. Struct. Biol.* 11: 231–235.
27. Aqvist, J., V. B. Luzhkov, and B. O. Brandsdal. 2002. Ligand binding affinities from MD simulations. *Acc. Chem. Res.* 35:358–365.
28. Simonson, T., G. Archontis, and M. Karplus. 2002. Free energy simulations come of age: protein-ligand recognition. *Acc. Chem. Res.* 35: 430–437.
29. Soriano, A., E. Silla, I. Tuñón, and M. F. Ruiz-López. 2005. Dynamic and electrostatic effects in enzymatic processes. An analysis of the nucleophilic substitution reaction in haloalkane dehalogenase. *J. Am. Chem. Soc.* 127:1946–1957.
30. Roca, M., J. Andrés, V. Moliner, I. Tuñón, and J. Bertrán. 2005. On the nature of the transition state in catechol O-methyltransferase. A complementary study based on molecular dynamics and potential energy surface explorations. *J. Am. Chem. Soc.* 127:10648–10655.
31. Martí, S., J. Andrés, V. Moliner, E. Silla, I. Tuñón, and J. Bertrán. 2003. Preorganization and reorganization as related factors in enzyme catalysis: the chorismate mutase case. *Chem. Eur. J.* 9:984–991.
32. Gao, J., and X. Xia. 1992. A priori evaluation of aqueous polarization effects through Monte Carlo QM-MM simulations. *Science.* 258: 631–635.
33. Gao, J. 1995. Methods and applications of combined quantum mechanical and molecular mechanical potentials. *Rev. Comput. Chem.* 7: 119–185.
34. Field, M. J., P. A. Bash, and M. A. Karplus. 1990. A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J. Comput. Chem.* 11:700–733.
35. Martí, S., M. Roca, J. Andrés, V. Moliner, E. Silla, I. Tuñón, and J. Bertrán. 2004. Theoretical insights in enzyme catalysis. *Chem. Soc. Rev.* 33:98–107.
36. Hensen, C., J. Hermann, K. Nam, S. Ma, J. Gao, and H. D. Höltje. 2004. A combined QM/MM approach to protein-ligand interactions: polarization effects of the HIV-1 protease on selected high affinity inhibitors. *J. Med. Chem.* 47:6673–6680.
37. Gräter, F., S. M. Schwarzl, A. Dejaegere, S. Fischer, and J. C. Smith. 2005. Protein/ligand binding free energies calculated with quantum mechanics/molecular mechanics. *J. Phys. Chem. B.* 109:10474–10483.
38. Field, M. J., M. Albe, C. Bret, F. Proust-De Martin, and A. Thomas. 2000. The dynamo library for molecular simulations using hybrid quantum mechanical and molecular mechanical potentials. *J. Comput. Chem.* 21:1088–1100.
39. Schaftenaar, G., and J. H. Noordik. 2000. Molden: a pre- and post-processing program for molecular and electronic structures. *J. Comput. Aided Mol. Des.* 14:123–134.
40. Dewar, M. J. S., E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart. 1985. Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* 107:3902–3909.
41. Jorgensen, W. L., D. S. Maxwell, and J. Tirado-Rives. 1996. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* 118: 11225–11236.

42. Jorgensen, W. L., J. Chandrasekar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926-935.
43. Steinbach, P. J., and B. R. Brooks. 1994. New spherical cutoff methods for long-range forces in macromolecular simulation. *J. Comput. Chem.* 15:667-683.
44. Gao, J., and D. G. Truhlar. 2002. Quantum mechanical methods for enzyme kinetics. *Annu. Rev. Phys. Chem.* 53:467-505.
45. Hestenes, M. R., and E. Stiefel. 1952. Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bureau of Standards.* 49:409-436.
46. Martí, S., V. Moliner, and I. Tuñón. 2005. Improving the QM/MM description of chemical processes: a dual level strategy to explore the potential energy surface in very large systems. *J. Chem. Theory Comput.* 1:1008-1016.
47. Gray, N. S., L. Wodicka, A. M. W. H. Thunnissen, T. C. Norman, S. Kwon, F. N. Espinoza, D. O. Morgan, G. Barnes, S. LeClerc, L. Meijer, S. H. Kim, D. J. Lockhart, and P. G. Schultz. 1998. Exploiting chemical libraries, structure, and genomics in the search for kinase inhibitors. *Science.* 281:533-538.
48. Hoessel, R., S. Leclerc, J. A. Endicott, M. E. M. Noble, A. Lawrie, P. Tunnah, M. Leost, E. Damiens, D. Marie, D. Marko, E. Niederberger, W. Tang, et al. 1999. Indirubin, the active constituent of a Chinese antileukaemia medicine, inhibits cyclin-dependent kinases. *Nat. Cell Biol.* 1:60-67.
49. Davies, T. G., T. Tunnah, L. Meijer, D. Marko, G. Eisenbrand, J. A. Endicott, and M. E. M. Noble. 2001. Inhibitor binding to active and inactive CDK2: the crystal structure of CDK2-cyclin A/indirubin-5-sulphonate. *Structure.* 9:389-397.