



UNIVERSIDAD NACIONAL AUTÓNOMA DE
MÉXICO

**PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS
QUÍMICAS**

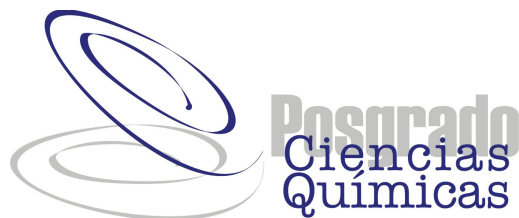
**DISEÑO DE FÁRMACOS ASISTIDO POR COMPUTADORA.
INHIBIDORES DE LA TRANSCRIPTASA REVERSA DEL VIRUS DEL VIH,
HIPOCOLESTEROLEMIANTES Y ANTIPARASITARIOS**

TESIS
PARA OPTAR POR EL GRADO DE

DOCTOR EN CIENCIAS

PRESENTA

M. en C. JOSÉ LUIS MEDINA FRANCO



TUTOR: Dr. Rafael Castillo Bocanegra

2005



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

A Karina, mi amada esposa

A mis padres y hermanos

Agradecimientos

Mi más profundo agradecimiento al Dr. Rafael Castillo por la dirección y apoyo para la realización de esta tesis.

Al Dr. Jesús Jiménez, Dr. Alexander Tropsha, Dr. Gerald M. Maggiora y a los miembros de sus respectivos grupos de investigación que enriquecieron enormemente mi formación. De igual manera, agradezco al M. en C. Sergio Rodríguez por su invaluable guía para adentrarme en el campo de diseño de fármacos asistido por computadora y para la realización de este trabajo.

Muchas de las observaciones plasmadas en esta tesis se deben a discusiones muy enriquecedoras con el Q. Fabián López. Agradezco también las valiosas discusiones con el Dr. Alexander Golbraikh, Dr. Scott Oloff, Dra. Min Shen y el Dr. Thomas Scior.

Agradezco a la M. en C. Cecilia Juárez cuyas ideas y trabajo inicial fueron muy importantes para desarrollar el estudio de los inhibidores de la transcriptasa reversa de VIH-1. De igual forma, agradezco al Dr. Francisco Hernández y a la M. en C. Alicia Hernández por su constante apoyo para la realización de esta tesis, especialmente en el desarrollo del estudio de los compuestos antiparasitarios. Mi más sincera gratitud al Dr. Joaquín Tamariz y al Dr. Germán Chamorro por su apoyo para realizar el estudio de los compuestos hipocolesterolemiantes.

A la Dra. Lilián Yépez y Biol. Amparo Tapia por la dirección y realización, respectivamente, de las pruebas biológicas reportadas en este trabajo. También agradezco a todos aquellos colegas, aquí citados, que sintetizaron los compuestos antiparasitarios.

Al Dr. Benjamín Ortiz, Dra. Carmen Soler, miembros del Comité Tutelar y al Dr. Enrique Ángeles y Dr. Antonio Cogordán, miembros del Comité Tutelar Ampliado, por sus valiosas observaciones.

Al Dr. Joaquín Tamariz, Dr. Vladimir Basiuk, Dr. Benjamín Ortiz, Dra. Lilián Yépez, Dr. Enrique Ángeles, Dra. Patricia Gudarrama y Dr. Fernando Cortés, miembros del Jurado, quienes revisaron y enriquecieron este manuscrito.

Agradezco al Dr. Artur Olson y su grupo de investigación por proporcionar el programa AutoDock y al Dr. Roman Laskowski por proporcionar el programa LigPlot. De igual forma, al Dr. Alexander Tropsha y a su grupo de investigación por proporcionar los programas para realizar los estudios *k*NN-QSAR y búsqueda automatizada en bases de datos.

Mi más profunda gratitud a la Dra. Karina Martínez que escuchó día a día los resultados de este trabajo y los enriqueció con sus comentarios. A las familias Medina Franco, Medina Gaona y Martínez Mayorga por su incansable apoyo.

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACyT) y a la Dirección General de Estudios de Posgrado por el financiamiento para realizar los estudios de doctorado. Agradezco a la Coordinación del Posgrado en Ciencias Químicas y al Programa CONACyT de Becas Mixtas por el apoyo económico para realizar estancias de investigación. Agradezco al Programa de Apoyo a los Estudios de Posgrado, proyecto número 208309, a los proyectos CONACyT G-34851-M y DGAPA IN 202101, por financiar parte de esta investigación.

Esta tesis se desarrolló en el laboratorio 122 del Conjunto E de la Facultad de Química de la Universidad Nacional Autónoma de México, bajo la dirección del Dr. Rafael Castillo Bocanegra.

Durante los estudios de doctorado se realizaron estancias de investigación, en orden cronológico, en el Centro de Investigaciones Biológicas del Consejo Superior de Investigaciones Científicas, en Madrid, España, bajo la dirección del Dr. Jesús Jiménez Barbero; en el Laboratorio de Modelado Molecular de la Escuela de Farmacia de la Universidad de Carolina del Norte, en Chapel Hill, bajo la dirección del Dr. Alexander Tropsha y en la Escuela de Farmacia e Instituto BIO5 de la Universidad de Arizona, bajo la dirección del Dr. Gerald M. Maggiora.

Índice

Resumen	i
Abstract	iii
Abreviaturas	v
Introducción	1
1. Antecedentes	3
1.1. Diseño de fármacos asistido por computadora	3
1.2. Diseño basado en la estructura del receptor	6
Acoplamiento (<i>docking</i>) molecular	7
1.3. Diseño basado en la estructura del ligando	9
1.3.1. Relaciones estructura-actividad cuantitativas (QSAR)	10
1.3.2. Búsqueda en bases de datos	13
2. Objetivos	14
3. Metodología computacional	16
3.1. Acoplamiento molecular	16
3.2. QSAR-3D	18
3.3. QSAR-2D	20
3.4. Búsqueda automática en bases de datos	22
Resultados y discusión	
Parte I. Compuestos antivirales	24
4. Acoplamiento molecular con la transcriptasa reversa de VIH-1 y diseño de nuevos compuestos	25
4.1. Inhibidores de la transcriptasa reversa de VIH-1	25
4.2. Estudio de la interacción de análogos de piridinona con la transcriptasa reversa de VIH-1	30
4.3. Diseño de análogos de piridinona con potencial actividad contra cepas resistentes de VIH-1	45

5. Estudios QSAR-3D	49
5.1. Modelos QSAR-3D	54
5.2. Análisis comparativo de campos moleculares (CoMFA)	57
5.3. Análisis comparativo de índices de similitud molecular (CoMSIA)	60
6. Análisis QSAR-2D y búsqueda de nuevos líderes potenciales en la base de datos del Instituto Nacional de Cáncer	64
6.1. Estudio QSAR de análogos de piridinona usando el método <i>k-Nearest Neighbors</i>	65
6.2. Búsqueda automatizada de nuevos líderes en la base de datos del Instituto Nacional de Cáncer	69
Parte II. Compuestos hipocolesterolemiantes	75
7. Acoplamiento molecular con la 3-hidroxi-3-metilglutaril-CoA reductasa y diseño de nuevos compuestos	76
7.1. Inhibidores de la 3-hidroxi-3-metilglutaril-CoA reductasa	76
7.2. Estudio de la interacción de α -asarona con la 3-hidroxi-3-metilglutaril-CoA reductasa	79
7.3. Diseño basado en la estructura de compuestos con potencial actividad inhibitoria de la 3-hidroxi-3-metilglutaril-CoA reductasa	86
Parte III. Compuestos antiparasitarios	89
8. Estudio QSAR-2D y predicción de la actividad de diversos compuestos contra <i>Giardia intestinalis</i>	90
8.1. Compuestos antiparasitarios contra <i>G. intestinalis</i>	90
8.2. Estudio QSAR-2D de análogos de bencimidazol y benzoxazol usando el método <i>k-Nearest Neighbors</i>	92
8.3. Búsqueda automatizada de nuevos líderes en la base de datos del Instituto Nacional de Cáncer	102
9. Conclusiones	108
Referencias bibliográficas	111
Anexo	
Publicaciones	

Lista de figuras

Figura 1.1	Métodos generales del DiFAC y su integración en el diseño racional de fármacos.	6
Figura 3.1	Etapas principales para hacer acoplamiento molecular con AutoDock 3.0.	16
Figura 3.2	Diagrama de flujo de la búsqueda en la base de datos empleando modelos QSAR.	23
Figura 4.1	Estructura de la TR de VIH (código PDB: 1vrt). Residuos de la cadena A están en amarillo y de la cadena B en púrpura. La nevirapina (esferas y barras) muestra el sitio de unión de los INNTR.	27
Figura 4.2	Nevirapina en el sitio de unión de los INNTR. Residuos representativos del sitio alostérico están señalados en verde. Residuos clave del sitio catalítico están en amarillo.	27
Figura 4.3	Inhibidores no-nucleosídicos de la TR en uso clínico. Se indica el año de aprobación o comercialización.	27
Figura 4.4	Conformación tipo mariposa que adopta la nevirapina (verde), efavirenz (rojo), 9-CI-TIBO (naranja), Cl- α -APA (amarillo) y MKC-442 (azul) en el sitio de unión de los INNTR.	27
Figura 4.5	Estructuras de cuatro análogos de piridinona con potente actividad inhibitoria de la TR.	28
Figura 4.6	Estrategia para el diseño de compuestos con potencial actividad inhibitoria de la TR de VIH-1.	30
Figura 4.7	Estructuras de los INNTR encontrados en las estructuras cristalográficas usadas en este estudio.	31
Figura 4.8	Comparación entre el modo de unión de la nevirapina encontrada en la estructura cristalográfica (blanco) y la posición de unión predicha por AutoDock (negro).	32
Figura 4.9	Modo de unión predicho para 1 (rojo), 2 (verde), 3 (azul) y 4 (amarillo) en el sitio de unión de (a) nevirapina y (b) Cl- α -APA. Nevirapina y Cl- α -APA (color por tipo de átomo) se muestran para comparación. Para mayor claridad, se omiten los hidrógenos.	36
Figura 4.10	Modo de unión predicho para 2 (color por tipo de átomo) en el sitio de unión de (a) MKC-442 (barras color púrpura) y (b) TNK-651 (barras color naranja). En ambos casos se compara con el modo de unión de la nevirapina (barras color verde). Se muestra la posición de la Pro236 cuando el inhibidor es MKC-442 (en líneas color púrpura), TNK-651 (en líneas color naranja) y nevirapina (en líneas color verde). Para mayor claridad, se omiten los hidrógenos.	36
Figura 4.11	Modo de unión predicho para 1 (barras) en el sitio de unión de (a) 1051U91; (b) 9-CI-TIBO y (c) S-1153. Los INNTR cristalográficos correspondientes (esferas y barras) se muestran para comparación. Para mayor claridad, se omiten los hidrógenos.	37

Figura 4.12	Modelo de unión de 2 con la TR. Se muestra la geometría del sitio de unión de la nevirapina. El puente de hidrógeno se muestra con una línea discontinua amarilla. Los residuos que tienen nombre en la figura se encuentran a una distancia de 3.6 Å del anillo de benzoxazol o piridinona. Se incluye el nombre para Trp229 como referencia. Para mayor claridad, se omiten los hidrógenos excepto el hidrógeno polar del anillo de la piridinona.	39
Figura 4.13	Representación 2D de las principales interacciones del compuesto 2 con la TR.	39
Figura 4.14	Relación entre el fragmento amida del anillo de piridinona con el fragmento amida y tioamida de otros INNTR. Todos ocupan la misma región de unión con la TR.	40
Figura 4.15	Estructuras de híbrido piridinona-efavirenz (5) e híbrido piridinona-MKC-442 (6).	41
Figura 4.16	Modo de unión predicho para 2 (esferas y barras) en el sitio de unión de efavirenz. Se muestra la posición del efavirenz (barras) para comparación. Para mayor claridad, se omiten los hidrógenos.	42
Figura 4.17	Modo de unión predicho para 5 (esferas y barras) en el sitio de unión de efavirenz. Se muestra la posición del efavirenz (barras) para comparación. Para mayor claridad, se omiten los hidrógenos.	42
Figura 4.18	Modo de unión predicho para 2 (esferas y barras) en el sitio de unión de nevirapina. Se muestra la posición de MKC-442 (barras) para comparación. Para mayor claridad, se omiten los hidrógenos.	44
Figura 4.19	Modo de unión predicho para 6 (esferas y barras) en el sitio de unión de MKC-442. Se muestra la posición de MKC-442 (barras) para comparación. Para mayor claridad, se omiten los hidrógenos.	44
Figura 4.20	Modo de unión predicho para 2 (esferas y barras) en el sitio de unión de UC-781. Se muestra la conformación con la mejor evaluación de energía del grupo más poblado. Se da la posición de UC-781 (barras) para comparación. Para mayor claridad, se omiten los hidrógenos.	45
Figura 4.21	Ejemplos de nuevos derivados de piridinona con potencial actividad inhibitoria de la TR, especialmente contra cepas con la mutación Tyr181Cys.	46
Figura 4.22	Modo de unión predicho para (a) 9 y (b) 10 (esferas y barras) en el sitio de unión de nevirapina. Se muestra la conformación predicha para 2 (barras) para comparación. Para mayor claridad, se omiten los hidrógenos.	47
Figura 5.1	Actividad biológica experimental versus energías libres de unión calculadas para compuestos estructuralmente similares (15-21 , 28 , 29 , 30 , 34 , 35).	52
Figura 5.2	Alineación de los análogos de piridinona.	53

Figura 5.3	Comparación del modo de unión de los compuestos 18 , 19 , 28 y 29 (naranja) con el modo de unión de 11 , 12 , 23 y 39 (verde). La posición de los residuos Tyr181, Trp229, Tyr318 y Lys101 se muestra como referencia. El puente de hidrógeno se muestra con una línea discontinua magenta. Para mayor claridad, se omiten los hidrógenos excepto el hidrógeno polar del anillo de la piridinona.	54
Figura 5.4	Estrategia del estudio QSAR-3D de los análogos de piridinona.	54
Figura 5.5	Predicciones de los modelos con 28 compuestos para los grupos de entrenamiento y prueba. Las predicciones de los modelos a 2 Å están representadas con círculos llenos y líneas continuas mientras que los modelos a 1 Å están representados con círculos vacíos y líneas punteadas. Las predicciones con los modelos de CoMFA para los grupos de entrenamiento y prueba están mostrados en a y b, respectivamente. Las correspondientes predicciones con los modelos de CoMSIA están mostradas en c y d, respectivamente.	56
Figura 5.6	Mapas de contorno para el mejor modelo de CoMFA. Se muestra al compuesto 29 y a los residuos que se encuentran a 5 Å de distancia de este compuesto. (a) Mapas de contorno estéricos. Las regiones favorecidas están en verde y las desfavorecidas en amarillo. (b) Mapas de contorno electrostáticos. Las regiones favorecidas por cargas negativas están en rojo y las desfavorecidas en azul. El puente de hidrógeno se muestra con una línea discontinua amarilla. Para mayor claridad, se omiten los hidrógenos excepto el hidrógeno polar del anillo de la piridinona.	59
Figura 5.7	Mapas de contorno para el mejor modelo de CoMSIA. Se muestra al compuesto 29 y a los residuos que se encuentran a 5 Å de distancia de este compuesto. (a) Mapas de contorno estéricos. (b) Mapas de contorno electrostáticos. Los colores de los mapas tienen la misma interpretación que el de las Figuras 5.6a y 5.6b, respectivamente. (c) Mapas de contorno hidrófobos. Las regiones favorecidas están en amarillo y las desfavorecidas en blanco. El puente de hidrógeno está representado como en la Figura 5.6. Para mayor claridad, se omiten los hidrógenos excepto el hidrógeno polar del anillo de la piridinona.	62
Figura 6.1	Gráficos de q^2 versus el número de descriptores seleccionados por los mejores modelos k NN-QSAR. Se muestran los resultados para el grupo de datos con actividades reales (triángulos) y aleatorias (cuadrados). Cada valor de q^2 es el promedio de 10 cálculos independientes.	67
Figura 6.2	Valores de actividad reales y calculados para el grupo de entrenamiento (cuadrados) y prueba (triángulos) empleando el modelo k NN-QSAR número 1 ($q^2 = 0.62$, $R^2 = 0.89$) (Tabla 6.2).	68
Figura 6.3	Estructuras químicas de compuestos encontrados en la búsqueda en la base de datos del NCI. Se muestra la estructura de la nevirapina (cf., Capítulo 4) como referencia.	69
Figura 6.4	Modo de unión predicho para NSC 11635 (esferas y barras) en el sitio de unión de la TR. Se muestra a la nevirapina (barras) para comparación. Para mayor claridad, se omiten los hidrógenos.	71

Figura 6.5	Comparación de las conformaciones de unión de las pirazo[3,4-d]pirimidinas (esferas y barras) con 27 (barras) (Capítulo 4). (a) NSC 11635. (b) NSC 11646. (c) NSC 1456. (d) NSC 11648. Las posiciones de Val179, Tyr181, Tyr188, Gly190, Trp229 y Tyr318 se muestran como referencia. Para mayor claridad, se omiten los hidrógenos.	72
Figura 6.6	Modo de unión más favorable (grupo 2) de NSC 127 (esferas y barras) en el sitio de unión de la TR. Se muestra a la nevirapina (barras) para comparación. Para mayor claridad, se omiten los hidrógenos.	73
Figura 7.1	Estructura de (a) HMG-CoA, (b) α -asarona y las estatinas (c) simvastatina y (d) rosuvastatina.	77
Figura 7.2	Estructura del homotetrámero de la HMGR humana (código PDB: 1dqa). Los residuos de los cuatro monómeros están en amarillo, púrpura, rojo y azul, respectivamente.	78
Figura 7.3	Representación 2D de las principales interacciones de HMG-CoA con la HMGR.	78
Figura 7.4	Comparación entre el modo de unión de la rosuvastatina que se encuentra en la estructura cristalográfica (color por tipo de átomo) y la conformación y posición predicha por AutoDock. (a) Acoplamiento flexible (rojo) y (b) acoplamiento rígido (amarillo). Para mayor claridad, se omiten los hidrógenos.	81
Figura 7.5	Modelo de acoplamiento derivado para la α -asarona con la porción catalítica de la HMGR. Se muestra la geometría de unión de HMG-CoA. Los puentes de hidrógeno están en líneas discontinuas amarillas. Se etiquetan residuos representativos a 3.6 Å de distancia de la α -asarona. Residuos de un monómero están en naranja y los del otro monómero en púrpura. Para mayor claridad, se omiten los hidrógenos excepto los hidrógenos involucrados en puentes de hidrógeno.	83
Figura 7.6	Comparación del modo de unión predicho para la α -asarona con el modo de unión cristalográfico de (a) el sustrato; (b) la simvastatina y (c) la rosuvastatina. Se muestran residuos representativos a 3.6 Å de distancia de la α -asarona. El color de los residuos es como en la Figura 7.5. Para mayor claridad, se omiten los hidrógenos.	84
Figura 7.7	(a) Subestructura tipo HMG del sustrato (en negritas). (b) Fragmento de la estructura de α -asarona (en negritas) que juega un papel similar que la subestructura tipo HMG en la unión con la HMGR.	85
Figura 7.8	Conformación de unión predicha para el grupo metoxilo C-1 de la α -asarona en el modelo de acoplamiento con el sitio de unión del sustrato (amarillo) y en el sitio de unión de la rosuvastatina (azul). Se muestran residuos representativos. Para mayor claridad, se omiten los hidrógenos.	85
Figura 7.9	Estructuras de un análogo fenoxicético de la α -asarona (1) y ejemplo de un nuevo análogo de α -asarona (2) con potencial actividad inhibitoria de la HMGR.	86

Figura 7.10	Acoplamiento de 1 con la HMGR. (a) Residuos del cis-loop que forman la cavidad ocupada por el carboxilato y (b) comparación del modo de unión con rosuvastatina (barras). Para mayor claridad, se omiten los hidrógenos.	87
Figura 7.11	Modelo de unión de 2 con la HMGR. Se muestra la geometría del sitio de unión de la rosuvastatina. Los puentes de hidrógeno se muestran con líneas discontinuas verdes. Se muestran residuos representativos a 3.6 Å de distancia de la α -asarona. El color de los residuos es como en la Figura 7.5. Para mayor claridad, se omiten los hidrógenos excepto los involucrados en puentes de hidrógeno.	88
Figura 8.1	Estructura de (a) metronidazol, (b) mebendazol y (c) albendazol.	90
Figura 8.2	Estrategia para el diseño de compuestos con potencial actividad contra <i>G. intestinalis</i> .	91
Figura 8.3	Diagrama de flujo de los modelos QSAR.	99
Figura 8.4	Gráficos de q^2 versus el número de descriptores seleccionados por los mejores modelos <i>k</i> NN-QSAR. Se muestran los resultados para el grupo de datos con actividades reales (círculos llenos) y aleatorias (círculos vacíos). Cada valor de q^2 es el promedio de 10 cálculos independientes.	99
Figura 8.5	Diagrama de flujo de la búsqueda en la base de datos del NCI a partir de los modelos QSAR.	103
Figura 8.6	Número de prototipos computacionales predichos por diferente número de modelos QSAR.	104

Lista de tablas

Tabla 1.1	Ejemplos de fármacos comerciales desarrollados con la ayuda de métodos computacionales.	5
Tabla 3.1	Valores frecuentemente usados de α y los valores críticos de Z_c para la prueba de una cola.	21
Tabla 4.1	Estructuras cristalográficas empleadas en este estudio.	31
Tabla 4.2	Acoplamiento de los INNTR encontrados en los cristales.	32
Tabla 4.3	Acoplamiento de 1-4 en las diferentes geometrías de unión de la TR.	34
Tabla 4.4	Comparación entre la energía de unión y la actividad biológica para 1-4 .	36
Tabla 4.5	Acoplamiento de los híbridos de piridinona en diferentes geometrías de unión de la TR.	43
Tabla 4.6	Acoplamiento de nuevos derivados de piridinona en el sitio de unión de la nevirapina.	47
Tabla 5.1	Estructuras, actividad inhibitoria de la TR de VIH-1 y energías libres de unión (en kcal/mol) calculadas por AutoDock para los compuestos del grupo de entrenamiento.	50
Tabla 5.2	Estructuras, actividad inhibitoria de la TR de VIH-1 y energías libres de unión (en kcal/mol) calculadas por AutoDock para los compuestos del grupo de prueba.	51
Tabla 5.3	Estadística de los modelos CoMFA y CoMSIA.	55
Tabla 5.4	Residuales de las predicciones para el grupo de prueba con los modelos de CoMFA y CoMSIA.	56
Tabla 5.5	Promedio de los valores de q^2 después de randomizar la actividad biológica.	57
Tabla 6.1	Estructuras, actividad inhibitoria de la TR, experimental y calculada por los modelos kNN -QSAR, de 44 análogos de piridinona usados en este estudio.	65
Tabla 6.2	Los diez mejores modelos kNN -QSAR.	68
Tabla 6.3	Acoplamiento molecular de compuestos seleccionados de la búsqueda de la base de datos del NCI en el sitio de unión de la nevirapina.	70
Tabla 6.4	Actividad predicha por los modelos kNN -QSAR, energía de acoplamiento y energía libre de unión calculada por AutoDock para las pirazo[3,4-d]pirimidinas.	73
Tabla 7.1	Acoplamiento flexible de HMG y estatinas representativas.	79
Tabla 7.2	Acoplamiento rígido de estatinas representativas.	80

Tabla 7.3	Resultados del acoplamiento de α -asarona con la porción catalítica de la HMGR humana.	81
Tabla 7.4	Resultados del acoplamiento de 1 y 2 en el sitio de unión de rosuvastatina.	87
Tabla 8.1	Estructuras y actividades biológicas de los compuestos empleados para el estudio <i>k</i> NN-QSAR.	93
Tabla 8.2	Estructuras y actividades biológicas predichas para los compuestos del grupo externo.	96
Tabla 8.3	Prueba de hipótesis de una cola para varios modelos QSAR.	100
Tabla 8.4	Los diez mejores modelos <i>k</i> NN-QSAR.	101
Tabla 8.5	Actividades experimentales y predichas para los compuestos en el grupo externo.	102
Tabla 8.6	Prototipos computacionales seleccionados de la búsqueda en la base de datos del NCI.	104
Tabla 8.7	Predicción de la actividad giardicida de prototipos computacionales adicionales obtenidos de la búsqueda en la base de datos del NCI.	105

Resumen

Se aplicaron diversas técnicas computacionales para el análisis de la interacción fármaco-receptor y diseño de compuestos antivirales, hipocolesterolemiantes y antiparasitarios. Los métodos computacionales fueron seleccionados de acuerdo a la información experimental disponible para los diferentes tipos de compuestos estudiados.

Los compuestos antivirales estudiados fueron derivados de piridinona. Estos compuestos son inhibidores no-nucleosídicos de la transcriptasa reversa (TR) del virus de la inmunodeficiencia humana tipo 1 (VIH-1). Los compuestos hipocolesterolemiantes analizados y diseñados fueron la α -asarona y análogos de α -asarona. Los compuestos antiparasitarios estudiados fueron derivados de bencimidazol y benzoxazol que muestran actividad contra el protozoario *Giardia intestinalis*.

La TR del VIH-1 es un blanco atractivo para el tratamiento del síndrome de la inmunodeficiencia adquirida (SIDA). Se conoce una base de datos amplia de análogos de piridinona con actividad inhibitoria de la TR. Esto, aunado al hecho de que se conoce la estructura tridimensional de esta enzima, permite la aplicación de métodos computacionales, tanto de diseño basado en la estructura del receptor como métodos basados en la estructura del ligando, para el análisis de las interacciones fármaco-receptor y diseño de nuevos compuestos antivirales.

Utilizando acoplamiento molecular automatizado (*docking* molecular), se generó un modelo de unión para cuatro análogos de piridinona con la TR. Con el modelo, se analizaron las interacciones fármaco-receptor identificando la función que pueden ejercer los diversos elementos estructurales de este tipo de compuestos en la unión con la TR. Los resultados permitieron explicar, a nivel molecular, estudios publicados de relación estructura-actividad y establecer las posibles causas de resistencia del VIH-1 a este tipo de inhibidores. El análisis está de acuerdo con la actividad biológica de híbridos de piridinona reportados recientemente. A partir del modelo de acoplamiento desarrollado para los análogos de piridinona, se diseñaron compuestos con potencial actividad inhibitoria de la TR, especialmente contra cepas resistentes.

Los estudios de acoplamiento molecular se extendieron a cuarenta análogos de piridinona. Las conformaciones y posiciones de unión se correlacionaron con las actividades biológicas reportadas en la literatura por medio del análisis de relaciones estructura-actividad cuantitativas (QSAR) en tres dimensiones (QSAR-3D). Los modelos de CoMFA y CoMSIA obtenidos permitieron profundizar sobre el tipo de interacciones que ocurren entre los análogos de piridinona y la TR.

Con los análogos de piridinona se llevaron a cabo estudios QSAR en dos dimensiones (QSAR-2D) utilizando el método *k-Nearest Neighbors* (*k*NN). Los modelos generados se validaron rigurosamente y se emplearon para efectuar una búsqueda automática de nuevos líderes en la base de datos del Instituto Nacional de Cáncer (NCI). Combinando los resultados de la búsqueda con acoplamiento molecular, se encontraron compuestos con potencial actividad inhibitoria de la TR.

La α -asarona es un compuesto con actividad hipocolesterolemiante que inhibe a la enzima 3-hidroxi-3-metilglutaril-coenzimaA reductasa (HMGR). Recientemente se ha publicado la estructura tridimensional de esta enzima lo que permite la aplicación de métodos basados en la estructura del receptor para analizar las interacciones fármaco-receptor correspondientes.

Empleando acoplamiento molecular se generó un modelo de unión para la α -asarona con la HMGR. El modelo de acoplamiento desarrollado permitió identificar elementos farmacofóricos de la α -asarona explicando, a nivel molecular, estudios de relación estructura-actividad para análogos de este compuesto. El análisis de las interacciones fármaco-receptor dio origen al diseño de nuevos análogos de α -asarona con potencial actividad inhibitoria de la HMGR. Este estudio computacional representa el primer estudio de acoplamiento molecular automatizado con la HMGR que tenemos conocimiento.

Se conocen una gran cantidad de derivados de bencimidazol y benzoxazol con actividad antiparasitaria contra *Giardia intestinalis*. Sin embargo, no se conoce el receptor de muchas de estas estructuras. Aunque no se pueden aplicar métodos de diseño basado en la estructura del receptor, la gran cantidad de estructuras con actividad biológica conocida permite la aplicación de métodos basados en la estructura del ligando.

Para una base de datos de más de cien derivados de bencimidazol y benzoxazol con actividad contra *G. intestinalis*, se desarrollaron modelos QSAR utilizando el método *k*NN. Los modelos se validaron y emplearon para predecir la actividad de ocho compuestos antes de que se efectuaran las pruebas biológicas. Las predicciones fueron confirmadas experimentalmente.

Con los modelos *k*NN-QSAR generados, se realizó una búsqueda de nuevos líderes potenciales contra *G. intestinalis* en la base de datos del NCI. De los compuestos identificados como posibles moléculas activas contra este parásito, se evaluaron a cinco compuestos. La actividad experimental no correspondió con la actividad predicha por los modelos QSAR. Se discuten las posibles causas de esta diferencia entre valores experimentales y calculados.

Abstract

Several computational techniques were applied to analyze drug-receptor interactions and design antiviral, hypocholesterolaemic and antiparasitic compounds. The computational methods were selected based on the experimental information available for each type of compounds.

The antiviral compounds studied were pyridinone derivatives. These compounds are non-nucleoside human immunodeficiency virus type 1 (HIV-1) reverse transcriptase (RT) inhibitors. The hypocholesterolemic compounds studied and designed were α -asarone and α -asarone analogues. The antiparasitic compounds studied were benzoxazol and benzimidazol derivatives active against the protozoa *Giardia intestinalis*.

HIV-1 RT is an attractive target in the treatment of the acquired immune deficiency syndrome (AIDS). Several pyridinone derivatives with HIV-1 RT inhibitory activity are reported. Also, the structure of RT has been solved. Therefore, it is feasible to employ structure-based and ligand-based computational methods to analyze drug-receptor interactions and design antiviral compounds.

Automated molecular docking was used to develop a docking model for four pyridinone analogues with RT. Based on this model, drug-receptor interactions were analyzed. As a result, different structural elements of the pyridinone derivatives, that may play a key role in the binding to RT, were identified. Docking results were useful to explain, at the molecular level, structure-activity relationships previously published and to establish the possible causes of resistance of RT to this type of inhibitors. The analysis is in agreement with the biological activity of pyridinone hybrids reported recently. Based on the docking model derived for the pyridinone analogues, novel compounds with possible activity against RT, especially against resistant strains of RT, were designed.

The molecular docking studies were applied to forty pyridinone analogues. The conformations and binding positions were correlated with the biological activities reported in the literature by means of three-dimensional quantitative structure-activity relationships (QSAR). CoMFA and CoMSIA models derived helped to better understand the type of interactions that occur between pyridinone analogues and RT.

Two-dimensional QSAR studies using the *k*-Nearest Neighbors (*k*NN) method were applied to pyridinone analogues. The models generated were rigorously validated and employed to search the National Cancer Institute (NCI) database for novel leads. The results were combined with docking studies. Novel compounds were identified as potential RT inhibitory leads.

α -Asarone is a hypocholesterolemic compound that inhibits the enzyme 3-hydroxy-3-methylglutaryl-coenzymeA reductase (HMGR). The structure of this enzyme has been solved recently. Therefore, it is possible to employ structure-based drug design methods to analyze the corresponding drug-receptor interactions.

Molecular docking was used to develop a binding model of α -asarone with HMGR. The docking model developed helped to identify pharmacophoric elements in the α -asarone structure. The model was also a valuable tool to explain, at the molecular level, structure-activity relationships studies of α -asarone analogues. Novel compounds, with potential HMGR inhibitory activity, were designed based on the drug-receptor interactions analysis conducted for α -asarone. To the best of our knowledge, this is the first automated docking study conducted with HMGR.

Several benzimidazol and benzoxazol compounds have been reported with antiparasitic activity against *G. intestinalis*. However, the molecular target of most of these structures remains unknown. Despite structure-based drug design methods cannot be employed, the great amount of structures with known biological activity enables the application of ligand-based design methods.

QSAR models using the *k*NN method were developed for a database containing more than one hundred benzimidazol and benzoxazol analogues active against *G. intestinalis*. The models were thoroughly validated and used to predict the activity of eight compounds before the biological testing was conducted. The predictions were experimentally confirmed.

The *k*NN-QSAR models developed were also used to search the NCI database for novel potential leads against *G. intestinalis*. The biological testing was conducted for five compounds. The experimental activity did not match the activity predicted by the QSAR models. Possible causes of the differences between the experimental and predicted activities are discussed.

Abreviaturas

ADT	AutoDock Tools
CoMFA	Del inglés: <i>Comparative molecular field analysis</i>
CoMSIA	Del inglés: <i>Comparative molecular similarity indeces analysis</i>
DiFAC	Diseño de fármacos asistido por computadora
HEPT	1-[(2-Hidroxietoxi)-metil]-6-(feniltio)timina
HMG-CoA	3-Hidroxi-3-metilglutaril-CoA
HMGR	3-Hidroxi-3-metilglutaril-CoA reductasa
IC ₅₀	Concentración inhibitoria media
INNTR	Inhibidores no-nucleosídicos de la transcriptasa reversa
kNN	Del inglés: <i>k-Nearest Neighbors</i>
NCI	Del inglés: <i>National Cancer Institute</i>
Noc	Del inglés: <i>Number of components</i>
PLS	Del inglés: <i>Partial least squares</i>
QSAR	Del inglés: <i>Quantitative structure-activity relationships</i>
QSAR-2D	QSAR en dos dimensiones
QSAR-3D	QSAR en tres dimensiones
RMSD	Del inglés: <i>Root mean square deviation</i>
SAR	Del inglés: <i>Structure-activity relationships</i>
SIDA	Síndrome de la inmunodeficiencia adquirida
TR	Transcriptasa reversa
VIH	Virus de la inmunodeficiencia humana

Introducción

El diseño de fármacos es un proceso complejo que requiere la participación de muchas disciplinas para encontrar a un nuevo líder, optimizarlo y desarrollarlo hasta su aprobación para el tratamiento de una enfermedad. Una de las muchas estrategias del diseño racional de fármacos es el diseño asistido por computadora. Esta estrategia es actualmente una herramienta muy poderosa para el descubrimiento de moléculas biológicamente activas.

Los métodos modernos del diseño de fármacos asistido por computadora pueden agruparse en dos categorías principales: métodos basados en el ligando y métodos basados en el receptor. Los primeros, que incluyen los estudios de relación estructura-actividad cuantitativos, métodos de asignación del farmacóforo y búsqueda automática en bases de datos, están fundamentados completamente en estudios de relación estructura-actividad experimentales de ligandos o inhibidores enzimáticos. La aplicación de estos métodos en las últimas tres o cuatro décadas ha dado origen a varios fármacos que actualmente están en el mercado. Los métodos basados en el receptor, que incluyen experimentos de acoplamiento (*docking*) molecular automatizado, requieren información sobre la estructura del receptor o de la enzima. Esta estrategia, que está cobrando cada vez mayor importancia gracias a los avances en los métodos de elucidación estructural, ha dado origen a varias moléculas que son líderes prometedores. Más aún, el acoplamiento molecular ha conducido al desarrollo de fármacos aprobados para su uso en la terapia de varias enfermedades.

En nuestro país se hacen amplios esfuerzos para el descubrimiento de compuestos antiparasitarios contra el protozooario *Giardia lamblia*, de compuestos antivirales contra el virus causante del SIDA y de compuestos hipocolesterolemiantes del tipo de análogos de α -asarona. En todos los casos se trata de atender problemas de salud pública mundial. La amplia información experimental que se ha generado en cada caso ofrece la posibilidad de aplicar diversos métodos computacionales a la búsqueda de nuevas alternativas quimioterapéuticas.

La transcriptasa reversa (TR) del virus de la inmunodeficiencia humana (VIH) es uno de los blancos más atractivos para el tratamiento del síndrome de la inmunodeficiencia adquirida (SIDA). A pesar de que se cuenta con información sobre la estructura tridimensional de esta enzima y se tienen datos experimentales sobre la actividad biológica de una gran cantidad de inhibidores, uno de los principales retos del diseño de nuevos antivirales es la alta capacidad de mutación que tiene la TR. Experimentos de acoplamiento molecular automático, que

pueden ser complementados con estudios de relaciones-estructura actividad cuantitativos para inhibidores conocidos, ayudan a estudiar, a nivel molecular, la forma de unión de los inhibidores de la TR con esta enzima. Los resultados computacionales, aunados a los datos experimentales, permiten racionalizar los mecanismos de resistencia y, como consecuencia, diseñar nuevas moléculas para ser activas contra cepas donde la TR presenta mutaciones. La información disponible sobre los estudios de relación estructura-actividad para los inhibidores de la TR también permite la posibilidad de aplicar un método basado en el ligando: la búsqueda automatizada en bases de datos públicas de nuevos líderes potenciales con actividad anti-VIH.

En cuanto a los compuestos hipocolesterolemiantes, se cuenta con información sobre la actividad biológica en estudios *in vivo* de un producto natural, la α -asarona, y su actividad sobre la enzima 3-hidroxi-3-metilglutaril-CoA reductasa (HMGR). Recientemente se ha resuelto la estructura tridimensional de esta enzima por cristalografía de rayos X pero no se conoce, a nivel molecular, la forma de unión de la α -asarona con la HMGR. Estudios de acoplamiento molecular automático constituyen una herramienta muy útil para generar un modelo de unión α -asarona-HMGR y, de esta manera, estudiar las interacciones fármaco-receptor correspondientes. Estos estudios, aunados a la información experimental conocida, representan una herramienta poderosa para asistir en el diseño de nuevos agentes hipocolesterolemiantes.

En el caso de compuestos antiparasitarios, se han sintetizado y evaluado biológicamente a más de cien derivados de bencimidazol y benzoxazol contra *Giardia intestinalis*. A pesar de que actualmente no se conoce el blanco de acción de estos compuestos, la basta información experimental con la que se cuenta ofrece la oportunidad de aplicar métodos computacionales basados en el ligando. De esta manera, se pueden realizar estudios de relación estructura-actividad cuantitativos y búsqueda automatizada en bases de datos que predigan la actividad de nuevos compuestos y sugieran nuevas estructuras con potencial actividad giardicida.

En este trabajo se aplicaron algunos métodos usados actualmente en el diseño de fármacos asistido por computadora como herramientas que participen en la búsqueda de nuevos compuestos antivirales, hipocolesterolemiantes y antiparasitarios. En los casos donde se conoce la estructura tridimensional del receptor, los estudios de acoplamiento molecular ayudarán a explicar, a nivel molecular, la actividad biológica de compuestos activos. El trabajo realizado con la α -asarona forma parte de una colaboración entre los grupos del Dr. Rafael Castillo, el Dr. Joaquín Tamariz y el Dr. Germán Chamorro, reflejando la importancia del trabajo interdisciplinario para el desarrollo de fármacos.

Capítulo 1

Antecedentes

1.1. Diseño de fármacos asistido por computadora

El diseño de fármacos asistido por computadora (DiFAC) es una área del diseño racional de fármacos (Boyd, 1999) y consiste en aplicar algún procedimiento realizado por una computadora para relacionar la actividad de un compuesto con su estructura (Hopfinger, 1985). Los objetivos principales del DiFAC son tres: descubrir nuevos líderes, optimizar a los líderes ya existentes y seleccionar, de un grupo dado de moléculas, a los candidatos que tienen mayor probabilidad (o menor probabilidad) de convertirse en fármacos exitosos. Estos objetivos pueden llevarse a cabo si, antes de cualquier prueba de laboratorio, se puede predecir adecuadamente la propiedad biológica o farmacológica de interés. De esta manera, un equipo de cómputo puede ayudar a descubrir y diseñar estructuras químicas que tengan las propiedades que las puedan calificar para entrar en el proceso de desarrollo de un fármaco (Boyd, 1998).

Los sistemas biológicos son tan complejos que en el proceso de diseñar fármacos con la ayuda de un equipo de cómputo se asumen varias cosas (Balbes, 1994):

- a) Un aspecto que asume comúnmente el DiFAC es que compuestos similares, con una actividad biológica en común, se unen en el mismo sitio receptor. Si ocurre que compuestos con la misma actividad farmacológica actúan en sitios diferentes, los resultados del cómputo serán erróneos.
- b) Para simplificar los modelos, se asume en general que tanto el ligando como el receptor interactúan en una sola conformación. En realidad, tanto el ligando como la macromolécula están experimentando continuamente cambios conformacionales.
- c) Usualmente se aproxima o, en algunos casos se ignora, el medio: disolvente, iones, impurezas. Para que los cálculos sean más reales, se deberían considerar moléculas de agua explícitas rodeando tanto al ligando como a la macromolécula. Desgraciadamente, esto complica los cálculos. Diferentes métodos se han desarrollado para manejar el problema de la solvatación. Una forma común de manejar este problema es concentrarse en las diferencias

entre sistemas similares. Asumiendo que el efecto del disolvente afectará a cada sistema por igual, se esperaría que los errores se cancelen.

d) Los resultados de los estudios biológicos *in vivo* e *in vitro*, con los que se evalúa a un grupo de compuestos, pueden tener un alto grado de incertidumbre. Sin embargo, se asume que la calidad de los datos biológicos es lo suficientemente adecuada para hacer los estudios computacionales.

Desde luego, las predicciones obtenidas con el DiFAC, como en muchas otras disciplinas científicas, deben someterse a pruebas experimentales. Concretamente, las moléculas diseñadas deben ser sometidas a pruebas de actividad biológica, toxicidad, especificidad, etc., para que, pasando éstas y otras consideraciones de manufactura y costos de producción, puedan llegar a convertirse en un fármaco aprobado para su uso clínico.

Los orígenes del DiFAC pueden remontarse a los estudios de Corwin Hansch quien, a principios de la década de los 60's, inició el desarrollo del método de relaciones estructura-actividad cuantitativas (QSAR) (Hansch, 1963; Hansch, 1964). Los primeros estudios QSAR hicieron uso de varias propiedades fisicoquímicas y descriptores moleculares, basados en la topología química, para derivar modelos que relacionaban la estructura con la actividad biológica. En los primeros años de la década de los 70's, el DiFAC comenzó a hacer uso de las gráficas moleculares generadas por computadoras. Esto condujo al desarrollo del modelado molecular en tres dimensiones, que puede definirse como la generación, manipulación, cálculo y predicción de estructuras químicas reales y sus propiedades fisicoquímicas y biológicas asociadas (Leach, 2001). La aplicación del modelado molecular en el diseño de fármacos dio origen al primer método tridimensional de asignación del farmacóforo: la estrategia del análogo activo (Martin, 1991). Posteriormente surgió el desarrollo de métodos de búsqueda en bases de datos basados en el farmacóforo y el establecimiento de estudios QSAR en tres dimensiones (QSAR-3D). Un ejemplo de este último es el análisis comparativo de campos moleculares (*Comparative Molecular Field Analysis*, CoMFA) (Cramer, 1988). En la década de los 80's comenzó a crecer el número de aplicaciones de diseño de fármacos que utilizan la información estructural de las macromoléculas biológicas blanco de los fármacos (Tropsha, 2001a). Actualmente, hay un marcado interés en desarrollar métodos computacionales no solamente para descubrir líderes y optimizar a los fármacos ya existentes sino también para la predicción de las propiedades farmacocinéticas o propiedades ADME/Tox (absorción, distribución, metabolismo, eliminación y toxicidad) de moléculas bioactivas (Waterbeemd, 2003).

Desde sus primeras aplicaciones al diseño de fármacos, las herramientas computacionales han contribuido significativamente al desarrollo de fármacos que se han aprobado para su uso

clínico. Uno de los primeros casos de éxito lo constituyó el antibacteriano norfloxacin. Estudios QSAR ayudaron al desarrollo de este fármaco que se comercializó por primera vez en 1983 con el nombre de Noroxin® (Boyd, 1999). Ejemplos de otros fármacos que se han comercializado con la asistencia de métodos computacionales se encuentran resumidos en la Tabla 1.1.

Tabla 1.1. Ejemplos de fármacos comerciales desarrollados con la ayuda de métodos computacionales

Compuesto	Nombre comercial (compañía farmacéutica)	Uso	Método computacional	Ref.
Norfloxacin	NOROXIN® (Kyorin Pharmaceutical)	Antibacteriano	QSAR	Koga, 1980
Losartan	COZAAR® (DuPont (BMS) Merck)	Tratamiento de hipertensión	Modelado molecular y QSAR	Duncia, 1990
Dorzolamida	TRUSPOT® (Merck)	Tratamiento del glaucoma	Diseño basado en el receptor y cálculos <i>ab initio</i>	Baldwin, 1989 Greer, 1994
Nelfinavir	VIRACEPT® (Pfizer)	Tratamiento del SIDA	Diseño basado en el receptor	Kaldor, 1997
Eprosartan	TEVETEN® (Smithkline Beecham)	Tratamiento de la hipertensión	Modelado molecular	Weinstock, 1991 Keenan, 1993
Zolmitriptan	ZOMING® (Wellcome, Zeneca)	Tratamiento de la migraña	Modelado molecular, desarrollo del farmacóforo	Buckingham, 1995
Indinavir	CRIVAN® (Merck)	Tratamiento del SIDA	Diseño basado en el receptor	Dorsey, 1994

Además de los fármacos que han llegado al mercado con la asistencia de computadoras, se reportan constantemente casos, por ejemplo en la revista *Journal of Medicinal Chemistry*, en donde el modelado molecular, estudios QSAR o algún otro método del DiFAC, han tenido impacto en el proyecto sugiriendo nuevos líderes potenciales (Charifson, 1997; Wlodawer, 2002).

El empleo de técnicas computacionales para el desarrollo de fármacos se ha visto favorecido por la disponibilidad de equipos de cómputo cada vez más rápidos y más baratos. Así mismo, el número de programas computacionales va creciendo en forma acelerada habiendo muchos programas disponibles gratuitamente. Hoy en día, la industria farmacéutica

y muchos grupos de investigación alrededor del mundo, incorporan herramientas computacionales al descubrimiento y desarrollo de fármacos (Balbes, 1994).

Los métodos del DiFAC pueden agruparse en dos áreas principales: métodos basados en el ligando y métodos basados en el receptor (Tropsha, 2001a). A los primeros también se les conoce como métodos basados en el farmacóforo (Reddy, 1999). El papel que juega el DiFAC en el descubrimiento y optimización de fármacos, y algunas de sus técnicas más comunes que se discutirán más adelante, están esquematizadas en la Figura 1.1. La aplicación de algún método basado en la estructura del ligando o basado en la estructura del receptor depende, en primera instancia, si se conoce o no la estructura tridimensional de la macromolécula blanco (Ooms, 2000). Ya sea que se utilice un tipo de método u otro, la finalidad es la predicción de la actividad biológica, o de alguna otra propiedad farmacológica, de una serie de moléculas. La siguiente etapa consiste en sintetizar, si es que no están disponibles, las moléculas seleccionadas como mejores candidatos y realizar las evaluaciones biológicas correspondientes. Los resultados de las evaluaciones biológicas pueden servir para nuevos diseños (Figura 1.1). Esta serie de etapas tiene como objetivo central la generación o selección de nuevos líderes o la optimización de líderes ya existentes. Si, además de conocer la estructura del receptor, se cuenta con información experimental sobre relaciones estructura-actividad para un grupo de moléculas que se unen al mismo receptor, se pueden combinar técnicas de diseño basado en el receptor y diseño basado en el ligando (por ejemplo: Lozano, 2000; Sippl, 2001; Sippl, 2002).

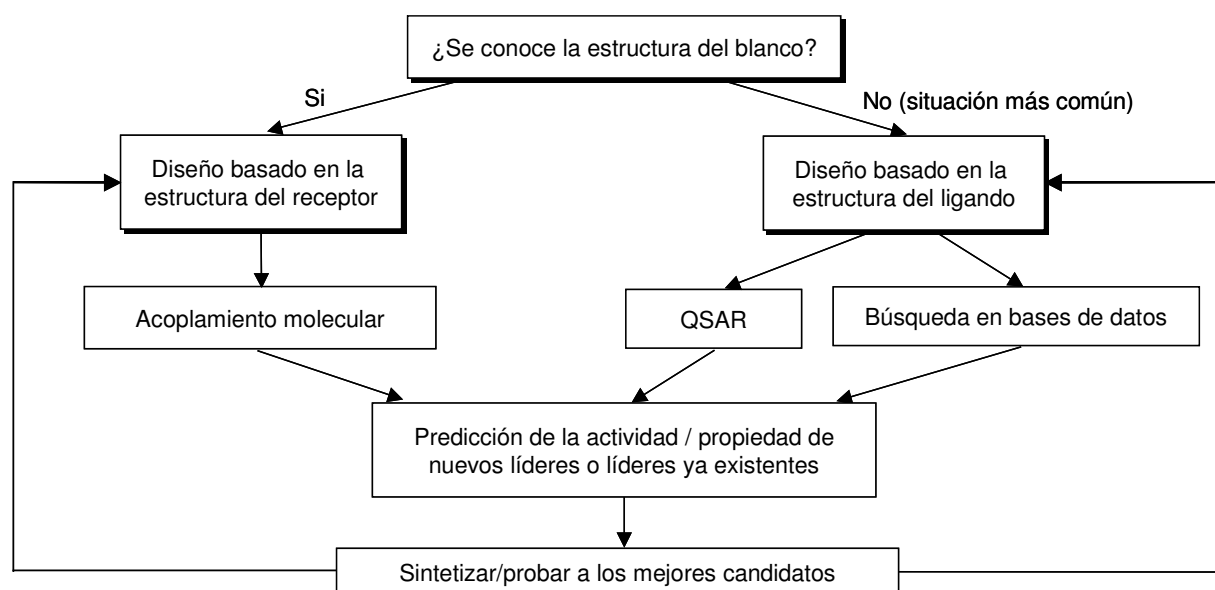


Figura 1.1. Métodos generales del DiFAC y su integración en el diseño racional de fármacos. Adaptado de (Reddy, 1999).

1.2. Diseño basado en la estructura del receptor

Recientemente, se han identificado a macromoléculas que juegan un papel importante en procesos metabólicos y rutas biosintéticas. La estructura de muchas de estas macromoléculas se conocen experimentalmente. De esta manera, se pueden aplicar procedimientos computacionales para identificar o diseñar ligandos que tengan una buena complementariedad estructural y energética en algún sitio de la biomolécula (Gohlke, 2002). A estos procesos computacionales se les da el nombre de diseño basado en la estructura del receptor (Oshiro, 1998).

Hoy en día hay tres fuentes de obtención de estructuras tridimensionales de biomoléculas. Dos de ellas son experimentales y son la cristalografía de rayos X y experimentos de resonancia magnética nuclear. El *Protein Data Bank* (RCSB) es una base de datos pública que tiene las coordenadas de miles de proteínas obtenidas por cristalografía de rayos X y resonancia magnética nuclear. Si no hay datos experimentales para una biomolécula, el modelado por homología (Hillisch, 2004) puede proporcionar datos estructurales aproximados si se conoce la estructura de una biomolécula similar.

Una vez que se cuenta con la estructura del receptor, ya sea en forma experimental o en un modelo, hay que buscar una zona en particular de la biomolécula con la cual interaccionará el ligando. Esto reduce significativamente el tamaño de la búsqueda que habrá que hacer posteriormente. Esta zona puede ser, por ejemplo, el sitio activo de una enzima o un sitio alostérico. En algunos casos se conoce este sitio a través de datos de mutaciones u otros experimentos. En otras ocasiones, la estructura de rayos X de una biomolécula con el ligando indica directamente el sitio de unión de la molécula. En ciertos casos, los sitios de unión pueden identificarse por un análisis visual de las estructuras. Cuando no se conocen los sitios de unión, pueden emplearse métodos computacionales que sugieren lugares probables de unión (Hetényi, 2002). Uno de los programas ampliamente usados para este propósito es Grid (Goodford, 1985).

Uno de los métodos de diseño basado en la estructura del receptor es el llamado diseño *de novo* que consiste en construir moléculas directamente en el receptor (Oshiro, 1998). Otro método, y el más empleado, es el acoplamiento molecular automatizado, comúnmente conocido en la literatura como *docking* molecular. Uno de los casos de éxito más notable del acoplamiento molecular es su participación en el desarrollo del indinavir (Tabla 1.1).

Acoplamiento (*docking*) molecular

El entendimiento detallado de los mecanismos de reconocimiento molecular es importante en el descubrimiento de fármacos, porque muchos de ellos interactúan con biomoléculas como blancos (Gohlke, 2002; Muegge, 2001). De esta manera, el diseño racional de fármacos requiere de herramientas que estén fundamentadas en el conocimiento del proceso de reconocimiento mutuo entre las estructuras de un fármaco y su receptor (Sotriffer, 2003). Los métodos computacionales que predicen la estructura tridimensional de un complejo ligando-receptor reciben comúnmente el nombre de acoplamiento molecular (Brooijmans, 2003, Muegge, 2001).

El objetivo en el acoplamiento molecular es encontrar el modo de unión de un ligando en un complejo ligando-receptor (Kitchen, 2004; Oshiro, 1998). El acoplamiento computacional requiere siempre de dos componentes que pueden ser caracterizados, en forma general, como la etapa de “búsqueda” y la etapa de “evaluación” (Brooijmans, 2003). La búsqueda se refiere a la exploración del espacio configuracional accesible para el ligando dentro del receptor. El objetivo de esta exploración es encontrar la orientación y conformación del ligando que corresponda al mínimo global de la energía libre de unión (Sotriffer, 2003). Por su parte, la etapa de la evaluación se refiere a la asignación de un valor numérico a cada una de las configuraciones generadas durante la etapa de la búsqueda (Sotriffer, 2003). Esto permite establecer un orden entre las diferentes posiciones y configuraciones encontradas. Generalmente, aquella posición con la mejor evaluación es la que representa el modo de unión más probable.

Desde el primer programa que surgió para hacer estudios de acoplamiento molecular por computadora (Kuntz, 1982), se han desarrollado muchos programas computacionales para este propósito. Algunos de ellos están resumidos en Oshiro, 1998; Sotriffer, 2003 y Kitchen, 2004. Actualmente, es de particular interés la comparación de los diferentes programas y funciones de evaluación para hacer acoplamiento molecular automatizado (Bursulaya, 2003; Kontoyianni, 2004).

Los métodos para hacer acoplamiento molecular pueden clasificarse de acuerdo a la forma en la que representan a la proteína, a la forma en la que representan al ligando y a la manera de efectuar la evaluación de las configuraciones generadas durante la etapa de la búsqueda (Sotriffer, 2003). La mayoría de los programas actuales consideran a la proteína como una macromolécula rígida. En cuando al ligando, algunos lo consideran rígido pero otros, como el programa AutoDock (Morris, 1998), consideran la flexibilidad.

Las funciones de evaluación pueden dividirse en dos tipos básicos: funciones relacionadas con la geometría y funciones relacionadas con la energía (Oshiro, 1998). Las primeras se basan en la complementariedad de las formas o superficies entre el ligando y el receptor. Este tipo de funciones se usa con mayor frecuencia para estudiar el acoplamiento entre dos macromoléculas.

Las funciones de evaluación relacionadas con la energía aproximan la energía libre de unión ligando-receptor como la energía de interacción calculada con mecánica molecular o bien consideran a la energía libre de unión como una suma ponderada de diferentes tipos de energía de interacción. Muchas de estas funciones se basan en la comparación de complejos ligando-receptor con datos experimentales de unión (Oshiro, 1998).

En otra clasificación de las funciones de evaluación se considera que estas pueden ser de tres tipos: las basadas en el conocimiento de estructuras de complejos ligando-proteína, las empíricas y los métodos basados en campos de fuerza (Gohlke, 2001; Sotriffer, 2003).

Hay tres aplicaciones importantes de los programas de acoplamiento: la predicción del modo de unión, la optimización de una molécula líder y la evaluación virtual (*in silico*) de bases de datos (Brooijmans, 2003). La predicción del modo de unión de un ligando puede ser aplicada *a posteriori* para racionalizar y entender la unión y las relaciones estructura-actividad de una serie de inhibidores y asistir en el diseño manual de nuevos inhibidores. Para evaluar la capacidad que tiene un programa para hacer predicciones, se compara la configuración predicha para el ligando con la estructura obtenida por cristalografía de rayos X.

Para propósitos de optimización de una molécula líder, el programa de acoplamiento tiene que tener la habilidad de evaluar correctamente, por orden de afinidad con el receptor, a compuestos químicamente similares. Por esto, se necesitan conocer valores de afinidad experimentales cuando se está evaluando la capacidad de un programa para la optimización de líderes.

Una de las mayores tendencias actuales del diseño basado en la estructura del receptor es el evaluar a una gran cantidad de moléculas organizadas en lo que se conocen como bibliotecas virtuales (Leach, 2000). Para estos fines, se necesitan programas de acoplamiento que sean a la vez rápidos y eficientes. Esto hace que el desarrollo de nuevos programas computacionales para hacer experimentos de acoplamiento, y la comparación entre ellos, sea una área de investigación muy importante en la actualidad (Kitchen, 2004).

1.3. Diseño basado en la estructura del ligando

Cuando no se conoce la estructura del receptor se utilizan los métodos basados en el ligando. Éstos están fundamentados en la información experimental disponible para una serie de estructuras químicas con actividad biológica conocida. Algunos métodos comunes dentro de esta área son los estudios QSAR y la búsqueda en bases de datos.

1.3.1. Relaciones estructura-actividad cuantitativas (QSAR)

La metodología involucrada en los estudios QSAR fue introducida por Hansch a principios de la década de los 60's (Hansch, 1963; Hansch, 1964). Esta estrategia está fundamentada en que la diferencia estructural de los compuestos es responsable de la diferencia de la actividad biológica de los mismos. Los factores estructurales que tienen mayor impacto en la actividad biológica son principalmente de tres tipos: electrónicos, estéricos e hidrófobos. La relación entre la actividad biológica (o la propiedad molecular) y los parámetros estructurales se obtienen a través de un análisis de regresión múltiple.

Cualquier método QSAR puede definirse, en general, como la aplicación de métodos matemáticos y estadísticos al problema de encontrar relaciones empíricas (modelos QSAR) de la forma $P_i = k(D_1, D_2, \dots, D_n)$, donde P_i son las actividades biológicas (o alguna otra propiedad de interés) de las moléculas. D_1, D_2, \dots, D_n son propiedades estructurales calculadas (o algunas veces medidas experimentalmente) de los compuestos y se les conoce comúnmente con el nombre de descriptores. En la expresión, k , representa alguna transformación matemática establecida empíricamente que habría que aplicar a los descriptores para calcular el valor de la propiedad P_i de todas las moléculas. La relación entre los descriptores D y la propiedad P puede ser lineal, donde la propiedad puede calcularse directamente a partir del valor de los descriptores, o no lineal, donde el valor de los descriptores se utilizan para caracterizar la similitud química entre las moléculas, lo cual se usa para predecir la propiedad P de los compuestos (Tropsha, 2003a).

Desde los primeros trabajos de Hansch se han desarrollado diferentes estrategias para hacer estudios QSAR (Debnath, 2001; Kubini, 1998b). Los diferentes métodos que se han desarrollado pueden analizarse desde dos puntos de vista: 1) los tipos de parámetros estructurales que se utilizan para caracterizar a las moléculas, desde simples fórmulas químicas a estructuras tridimensionales; y 2) los procedimientos matemáticos que se emplean para obtener las relaciones cuantitativas entre los descriptores moleculares y la actividad biológica.

De acuerdo al origen de los descriptores moleculares usados en los cálculos, los métodos QSAR pueden dividirse en tres grupos. Un grupo usa un número relativamente pequeño de propiedades fisicoquímicas y parámetros que describen, por ejemplo, efectos hidrófobos, estéricos y electrostáticos. Estos métodos son conocidos en la literatura con el nombre de análisis de Hansch. Métodos más recientes se basan en características cuantitativas de gráficas moleculares. Debido a que las gráficas moleculares o fórmulas estructurales son bidimensionales, a estos métodos se les refiere como estudios QSAR en dos dimensiones o QSAR-2D (Tropsha, 2003a).

Un tercer grupo de métodos se basa en los descriptores obtenidos de la representación tridimensional de las estructuras de las moléculas. A estos estudios se les conoce como QSAR en tres dimensiones o QSAR-3D (Greco, 1997; Tropsha, 2003a). Estos métodos se han hecho cada vez más populares con el desarrollo rápido y preciso de métodos computacionales para generar conformaciones tridimensionales y alineaciones de las estructuras químicas. Probablemente el ejemplo más conocido de QSAR-3D es CoMFA. Otro ejemplo representativo es el análisis comparativo de índices de similitud molecular (*Comparative Molecular Similarity Indices Analysis*, CoMSIA) (Klebe, 1994).

Los estudios QSAR pueden también ser clasificados por el tipo de métodos de correlación usados en el desarrollo de los modelos. Estos pueden ser métodos lineales, como es el caso de la regresión lineal o regresión lineal múltiple, y no lineales. Ejemplo de este último es el análisis de k -vecinos más cercanos (*k-Nearest Neighbors*, k NN) (Zheng, 2000).

Los diferentes modelos QSAR tienen sus propias ventajas y desventajas. Por ejemplo, los métodos QSAR-3D tienen la ventaja de mostrar sus resultados en forma gráfica que pueden interpretarse fácilmente en términos de las interacciones estéricas y electrostáticas importantes para que el ligando se una al receptor. Sin embargo, estos métodos requieren de la alineación de las estructuras en tres dimensiones, lo cual es subjetivo, consume tiempo y dificulta el análisis de bases de datos grandes. Por otra parte, los métodos QSAR-2D son mucho más rápidos y más fácilmente adaptables a la automatización. Esto se debe a que no requieren búsqueda conformacional y alineación de las estructuras. De esta forma, los métodos QSAR-2D son mejores para analizar a una gran cantidad de compuestos y hacer búsquedas computacionales en bases de datos. Sin embargo, la interpretación de los modelos en términos químicos comunes es, con frecuencia, difícil o prácticamente imposible (Tropsha, 2003a).

Después de desarrollar un modelo que correlacione la actividad biológica con los descriptores, la siguiente etapa de un estudio QSAR es la validación de los modelos. Esta etapa es necesaria debido a la posibilidad de que las correlaciones encontradas sean

producto de la casualidad (Topliss, 1972). La validación consiste en evaluar la capacidad que tiene el modelo QSAR para predecir con exactitud la propiedad de interés (por ejemplo, la actividad biológica) de los compuestos que no fueron utilizados en el desarrollo del modelo. La predicción es una de las características más importantes de los modelos QSAR (Golbraikh, 2002b; Oprea, 1997). Al grupo de compuestos con los que se desarrolla el modelo se le denomina grupo de entrenamiento mientras que, al grupo con que se valida, se le llama grupo de prueba. Se ha sugerido que el grupo de prueba debe contener al menos cinco compuestos (Golbraikh, 2002a). En el caso de que, al momento de desarrollar los modelos QSAR, no se cuente con un grupo de prueba, se puede dividir el grupo de datos original en los grupos de entrenamiento y prueba. Una alternativa a la selección de un grupo de entrenamiento y un grupo de prueba, aunque menos eficiente, es la llamada validación cruzada (Kubini, 1998b). En esta estrategia se desarrollan varios modelos QSAR en donde se han eliminado uno o varios compuestos del grupo de datos que después son predichos por los correspondientes modelos. La validación cruzada más común es la llamada dejar-uno-afuera en donde cada compuesto se excluye una vez. Si se tienen n compuestos se derivan entonces n modelos y las n predicciones se comparan con los valores experimentales. Como resultado se obtiene un coeficiente de correlación cruzada q^2 (Kubini, 1998b). La capacidad de predicción de un método QSAR puede considerarse como el criterio más importante para evaluar la eficiencia de ese método (Tropsha, 2003b).

Para evaluar qué tan robusto es un modelo QSAR se puede usar la estrategia de asignar valores aleatorios a la actividad biológica y repetir las predicciones (Golbraikh, 2002a). Si el modelo es robusto, las predicciones con las actividades reales deben ser estadísticamente superiores que cuando se derivaron los modelos con valores aleatorios (Tropsha, 2003b).

Las diferentes metodologías QSAR han tenido aplicaciones exitosas al diseño de fármacos. Numerosos ejemplos de estas aplicaciones pueden encontrarse en la literatura (Debnath, 2001; Fujita, 1990; Hansch, 1995). Hoy en día, una de las aplicaciones más importantes que tienen los estudios QSAR es la predicción de una gran cantidad de compuestos en forma rápida. Esto se debe al creciente desarrollo de la química combinatoria y a las pruebas biológicas de alto rendimiento. De esta manera, se puede predecir en forma virtual la actividad de compuestos que podrán ser sintetizados posteriormente, de manera que los estudios QSAR sirven como guía para la creación de bases de datos. Todo esto hace que en la era moderna de la química farmacéutica, los estudios QSAR constituyan una de las herramientas más importantes del diseño de fármacos asistido por computadora (Tropsha, 2003a).

1.3.2. Búsqueda en bases de datos

En el diseño de fármacos, la búsqueda en bases de datos o evaluación virtual se entiende como el uso de un procedimiento computacional para evaluar una serie de compuestos químicos contenidos en una base de datos, y sugerir qué compuestos son los más prometedores para su posterior síntesis y/o evaluación biológica (Tropsha, 2003a; Walters, 1998). Esta estrategia es cada vez más popular en el diseño de fármacos y su uso se ha visto promovido por la química combinatoria y la evaluación biológica de alta eficiencia.

Para hacer una búsqueda en una base de datos es necesario contar con un patrón de búsqueda. Usualmente este patrón está relacionado con la estructura de una molécula líder. En muchos casos, los patrones de búsqueda se derivan de farmacóforos tridimensionales (Martin, 1992) que contienen información sobre los requerimientos estructurales esenciales para la actividad biológica. Sin embargo, se ha demostrado recientemente que modelos QSAR-2D pueden ser empleados para hacer búsquedas en bases de datos (Shen, 2004).

Hay dos estrategias principales para realizar una búsqueda en base de datos. Una de ellas es la búsqueda de subestructuras. Se utiliza como patrón de la búsqueda a la subestructura del compuesto líder que se cree es importante para la actividad. El resultado será todas aquellas estructuras en la base de datos que contengan esa subestructura. Un problema que se puede presentar en este tipo de búsquedas es que resulte una gran cantidad de compuestos para poder analizarse en un tiempo razonable.

Otra estrategia es la búsqueda por similitud. En este caso, se emplean una serie de descriptores que, en forma colectiva, describen a la estructura completa del compuesto líder. Esta serie de descriptores se compara con el correspondiente grupo de descriptores para cada uno de los compuestos en la base de datos y posteriormente se calcula la similitud entre ellos. Se han desarrollado una variedad de descriptores aplicados a este tipo de búsqueda así como una serie de métodos para evaluar a la similitud (Tropsha, 2003a).

Aunque muchas bases de datos son propiedad de compañías farmacéuticas, hay bases de datos públicas importantes. Una de las bases de datos públicas más grandes es la base de datos del Instituto Nacional de Cáncer (NCI) (NCI Database) que contiene *ca.* 250,000 compuestos.

Recientemente, se han reportado varias aplicaciones exitosas de la búsqueda en la base de datos del NCI en donde han surgido nuevos líderes potenciales contra diferentes blancos, por ejemplo, la transcriptasa reversa (Keller, 2003), la proteasa de VIH (Wang, 1996) y agentes anticonvulsivos (Shen, 2004).

Capítulo 2

Objetivos

General

Analizar interacciones ligando-receptor y diseñar compuestos antivirales, hipocolesterolemiantes y antiparasitarios usando métodos empleados en el diseño de fármacos asistido por computadora.

Particulares

1. Desarrollar un modelo de unión para derivados de piridinona con la transcriptasa reversa (TR) de VIH-1, mediante experimentos de acoplamiento molecular automatizado, y analizar las interacciones ligando-receptor correspondientes.
2. Con el modelo de unión generado para derivados de piridinona con la TR explicar, a nivel molecular, estudios de relación estructura-actividad cualitativos reportados para análogos de piridinona y explicar observaciones experimentales de la respuesta biológica de este tipo de inhibidores a mutaciones en la enzima.
3. Diseñar, a partir de los estudios de la interacción piridinona-TR, nuevos análogos de piridinona con potencial actividad anti-VIH, especialmente activos contra cepas resistentes.
4. Generar y validar modelos de relaciones estructura-actividad cuantitativas en tres dimensiones, aunadas a experimentos de acoplamiento molecular automático, para profundizar en el conocimiento de las interacciones fármaco-receptor de los análogos de piridinona con la TR.
5. Derivar y validar modelos de relaciones estructura-actividad cuantitativas en dos dimensiones (QSAR-2D) para análogos de piridinona.
6. Con los modelos QSAR-2D para los análogos de piridinona, identificar nuevos líderes potenciales contra la TR mediante una búsqueda automática en una base de datos pública.

7. Desarrollar un modelo de unión para la α -asarona con la 3-hidroxi-3-metilglutaril-CoA reductasa (HMGR), utilizando acoplamiento molecular automatizado y analizar las interacciones ligando-receptor correspondientes.
8. Diseñar nuevos inhibidores potenciales de la HMGR.
9. Generar y validar modelos QSAR-2D para análogos de bencimidazol y benzoxazol con actividad contra *G. intestinalis*.
10. Predecir la actividad giardicida de derivados de bencimidazol y comprobar las predicciones experimentalmente.
11. Identificar, en una base de datos pública, compuestos potencialmente antiparasitarios utilizando los modelos QSAR-2D. Comprobar experimentalmente la actividad antiparasitaria de compuestos seleccionados.

Capítulo 3

Metodología computacional

Las moléculas se construyeron en todos los casos con Sybyl 6.8 o Sybyl 6.9 (Sybyl). La visualización de estructuras se realizó con Sybyl 6.8, Spartan'02 (Kong, 2000), ViewerPro 5.0 (ViewerPro) y AutoDock Tools (ADT).

3.1. Acoplamiento molecular

Las estructuras cristalográficas se obtuvieron del *Protein Data Bank* (RCSB). La preparación de los ligandos y de las estructuras de las proteínas se realizó con Sybyl 6.8. Los cálculos de acoplamiento molecular se realizaron con AutoDock 3.0 (Morris, 1998). En la Figura 3.1 se esquematizan las etapas principales para realizar el acoplamiento con este programa.

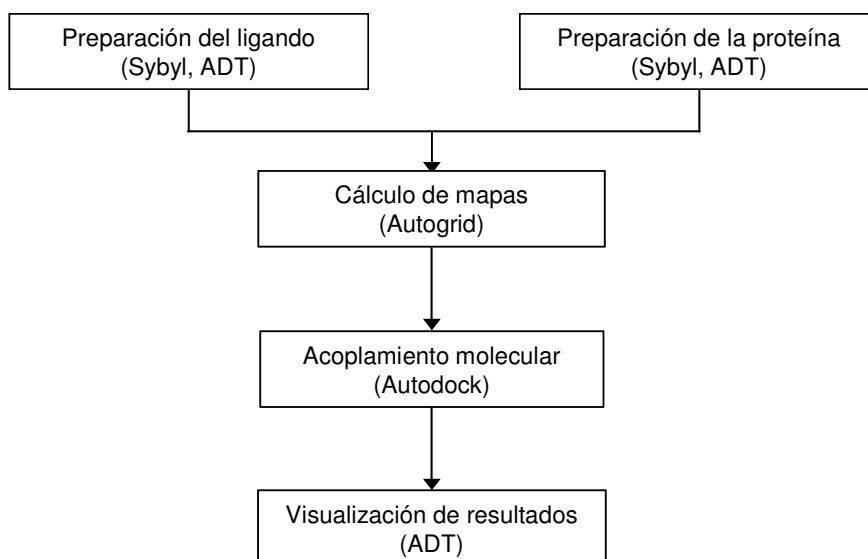


Figura 3.1. Etapas principales para hacer acoplamiento molecular con AutoDock 3.0. Entre paréntesis se indican los programas con los que se realizó cada etapa.

El programa AutoDock fue desarrollado para predecir la interacción de moléculas pequeñas con algún blanco (Morris, 2001). En resumen, AutoDock efectúa el acoplamiento automático de un ligando, con una flexibilidad previamente determinada, dentro de un sitio de unión rígido de una proteína. El programa realiza diversas búsquedas en cada experimento de acoplamiento. Cada búsqueda da como resultado la predicción de un modo de unión.

Todas las moléculas de agua, ion magnesio, adenosina-5'-difosfato y 2,3-dihidroxi-1,4-ditiobutano, en el caso de estar presentes, se removieron con Sybyl de los archivos originales obtenidos del *Protein Data Bank*.

Antes de realizar el acoplamiento de los inhibidores de la TR, los sitios de unión de los inhibidores no-nucleosídicos fueron superpuestos en forma independiente sobre la estructura del complejo TR-nevirapina, que es la estructura con la mayor resolución disponible. La superposición se llevó a cabo con Sybyl utilizando el comando FIT MONOMERS de Biopolymer. Se consideraron los átomos de carbono alfa de los residuos 94-118, 156-215, 225-243 y 317-319 de la subunidad p66 y a los residuos 137-139 de la subunidad p51 (Hopkins, 1996).

Utilizando ADT, se adicionaron los átomos de hidrógeno polares a las proteínas, las cargas de Kollman (Weiner, 1984), los parámetros de solvatación atómicos y los volúmenes fraccionarios. A todos los ligandos se les asignaron las cargas parciales de Gasteiger-Marsili (Gasteiger, 1980) y posteriormente los hidrógenos no-polares fueron eliminados sumando su carga a los carbonos vecinos. Se permitieron rotar a todos los enlaces posibles, excepto en los casos que se indica.

Para la validación del protocolo de acoplamiento, se removieron las coordenadas de los ligandos presentes en las estructuras cristalográficas. Con Sybyl, se corrigió la estructura del ligando 1051U91 de la estructura cristalográfica con código PDB: 1rth.

En AutoDock, para hacer las evaluaciones de la energía de las diferentes conformaciones del ligando dentro de la proteína en forma rápida, se calculan previamente a las evaluaciones energéticas, mapas de afinidad para cada tipo de átomo presente en el ligando y un mapa de potencial electrostático (Morris, 2001). Estos mapas fueron calculados por el programa auxiliar AutoGrid (Figura 3.1). En el procedimiento de AutoGrid, la proteína se coloca dentro de una rejilla. Esta rejilla consiste en un arreglo tridimensional de puntos regularmente espaciados. En este trabajo, cada rejilla se centró en las coordenadas del ligando cristalográfico correspondiente. En todos los casos, las dimensiones fueron $23 \times 23 \times 23 \text{ \AA}^3$ con puntos separados por 0.375 \AA . Para cada punto se calcula un valor de energía potencial entre un átomo de prueba que se sitúa en esa posición y la proteína. La energía del átomo de prueba en cada punto de la rejilla queda determinada por el grupo de parámetros

suministrados por AutoDock para este tipo particular de átomo y todos los átomos de la macromolécula que se encuentran a una distancia de no-enlace. Para el cálculo del mapa de potencial electrostático, AutoDock calcula interacciones de tipo coulómbicas entre la proteína y una carga de prueba e , $+1.60219 \times 10^{-19}$ C. AutoDock utiliza una constante dieléctrica dependiente de la distancia sigmoidal para modelar al disolvente, basándose en el trabajo de Mehler y Solmajer (Mehler, 1991). En el presente trabajo, el cálculo de los mapas de energía se realizó utilizando los parámetros especificados en AutoDock 3.0.

Para todos los ligandos se emplearon posiciones iniciales, orientaciones y torsiones aleatorias. Para la búsqueda de las posiciones de menor energía se utilizó el algoritmo genético de AutoDock 3.0. Los parámetros asociados a este método de búsqueda fueron los suministrados por defecto en el programa. El número total de búsquedas fue de cien.

Después de la etapa de búsqueda y evaluación de la energía de las diferentes posiciones y conformaciones del ligando, las cien soluciones fueron organizadas en grupos con desviaciones (RMSD) inferiores a 1.0 Å. En cada grupo, las diferentes soluciones se acomodaron por orden creciente de energía.

Después de cada experimento de acoplamiento, la energía del complejo ligando-receptor se optimizó con el campo de fuerza de Tripos (Clark, 1989). Durante la minimización, se permitió el movimiento de los átomos que se encuentran en una esfera de 6 Å de ligando mientras que se restringió el movimiento de los átomos localizados a una distancia mayor a esta distancia.

3.2. QSAR-3D

El análisis QSAR-3D se efectuó utilizando Sybyl. En un estudio QSAR-3D se relaciona la actividad biológica de moléculas con sus propiedades calculadas en un espacio tridimensional (Greco, 1997). El estudio comprende cuatro etapas generales (Kubini, 1998a). La primera etapa consiste en la alineación y cálculo de cargas atómicas parciales de las moléculas. En este estudio la alineación de los análogos de piridinona se hizo a partir de la conformación y posición de unión predicha por AutoDock. A las moléculas se les calcularon cargas parciales de Gasteiger-Hückel y cargas atómicas parciales obtenidas con el método semiempírico AM1 (Dewar, 1985).

En la segunda etapa, se coloca una caja alrededor de las moléculas y se define una rejilla dentro de esta caja con un espacio determinado entre sus puntos. En el presente trabajo se emplearon rejillas con un espacio equidistante entre sus puntos de 1 y 2 Å, de acuerdo a la

metodología original de CoMFA (Cramer, 1988). Cada rejilla se extiende 4 Å más allá de las coordenadas de cada átomo en cualquiera de las tres direcciones.

En la tercera etapa, se colocan diferentes átomos de prueba en cada punto de la rejilla, por ejemplo un átomo de carbono o un átomo con carga positiva o negativa. En los estudios de CoMFA se calculan valores de campo en cada punto, es decir, valores de energía que tendría cada átomo de prueba en cada posición correspondiente de la rejilla. En los estudios de CoMSIA se calculan valores de similitud entre cada átomo de prueba y las moléculas en cada punto (Klebe, 1994). Para los estudios de CoMFA se calcularon campos estéricos (van der Waals 6-12) y electrostáticos (coulómbica) con el campo de fuerza de Tripos usando como átomo de prueba a un carbono con hibridación sp^3 , con carga +1.0, y con una constante dieléctrica dependiente de la distancia (Cramer, 1988). Los campos de CoMFA también se calcularon con un oxígeno con hibridación sp^3 con carga -1.0, y un átomo de hidrógeno con carga +1.0, como átomos de prueba. Se empleó un filtro de 2.0 kcal/mol para eliminar aquellos puntos de la red con poca influencia sobre la actividad biológica. Para CoMSIA, se emplearon las mismas rejillas usadas en el estudio de CoMFA calculando campos estéricos, electrostáticos, hidrófobos, donadores y aceptores de hidrógenos utilizando como átomo de prueba a un carbono sp^3 con carga +1.0. Al igual que en CoMFA, se empleó un filtro de 2.0 kcal/mol. Los índices de CoMSIA se calcularon de acuerdo a lo que está descrito en la literatura (Klebe, 1994).

En la cuarta etapa de un estudio QSAR-3D, los campos calculados, que corresponden a tablas conteniendo usualmente miles de columnas, deben ser correlacionados con los valores de actividad biológica. En este estudio, los campos calculados para CoMFA y CoMSIA se utilizaron como variables independientes mientras que la actividad biológica, expresada como $-\log I_{C_{50}}$, fue utilizada como variable dependiente. Para obtener los modelos QSAR-3D de CoMFA y CoMSIA, se utilizó el análisis de regresión de mínimos cuadrados parciales (PLS) (Wold, 1998). La predictibilidad general de los modelos se evaluó por la validación cruzada de dejar-uno-afuera. El coeficiente de validación cruzada q^2 se calculó de acuerdo a la siguiente ecuación:

$$q^2 = (SSY - PRESS)/SSY$$

donde SSY es la varianza de la actividad biológica alrededor del valor promedio y PRESS es la suma de los cuadrados de los errores de predicción, derivados del método de dejar-uno-afuera. La incertidumbre de la predicción está definida por:

$$SPRESS = [PRESS/(n-k-1)]^{1/2}$$

donde k es el número de las variables en el modelo y n es el número de los compuestos empleados en el estudio. El número de componentes (noc) que correspondió al menor valor

de PRESS se empleó para derivar los modelos finales de regresión. También se calculó el coeficiente de correlación convencional r^2 así como el coeficiente de correlación convencional para la predicción de los compuestos en el grupo de prueba r^2_{pred} .

El resultado de un análisis QSAR-3D corresponde a una ecuación de regresión con miles de coeficientes, que se presenta en este estudio como un grupo de mapas de contorno. Estos grupos de contorno muestran las regiones alrededor de las moléculas en las cuales una sustitución tiene mayor impacto en la actividad biológica (Oprea, 1997).

3.3. QSAR-2D

Se emplearon índices topológicos moleculares (Leach, 2003; Randic, 1975) calculados con el programa MolConnZ 4.05 (MolConnZ). En general, MolConnZ genera más de 400 descriptores dentro de los que se encuentran los descriptores chi, kappa y electrotopológicos de Kier y Hall (Kier, 1976; Kier, 1999). La mayoría de estos descriptores caracterizan a la estructura química pero varios dependen de la numeración de los átomos de las moléculas y se introducen únicamente con fines de registro. En el estudio de los inhibidores de la TR se utilizaron solamente 193 descriptores químicos (después de eliminar a aquellos con un valor de cero o con una varianza igual a cero). En el estudio de los compuestos antiparasitarios se emplearon 247 descriptores químicos. Debido a que las escalas absolutas de los descriptores MolConnZ pueden variar por ordenes de magnitud, los descriptores fueron escalados antes de hacer otro cálculo (los detalles de este escalamiento están dados en Zheng, 2000). El escalamiento evita dar a los descriptores originalmente calculados con un valor alto, una contribución desproporcionada en los cálculos de la distancia, que forman parte del procedimiento de k NN en el espacio multidimensional de los descriptores MolConnZ.

El método QSAR empleado fue el de k NN-QSAR (Zheng, 2000), que utiliza el principio de clasificación de k vecinos más cercanos, en donde un compuesto es clasificado de acuerdo con los valores correspondientes de los k vecinos más próximos en el grupo de entrenamiento en un espacio n -dimensional (Sharaf, 1986), y un procedimiento de selección de variables. En resumen, un grupo de descriptores es seleccionado en forma aleatoria como un farmacóforo de descriptores hipotético. Se modifican entonces los descriptores hasta obtener el mejor valor posible de q^2 . El farmacóforo de descriptores hipotético es validado por el método de validación cruzada de dejar-uno-afuera, donde cada uno de los compuestos se elimina del grupo de entrenamiento y su actividad biológica es predicha como el promedio de actividades del número k de moléculas más similares (se emplearon valores de k entre 1 y 5). La similitud es caracterizada por la distancia euclidiana entre los compuestos en un espacio

multidimensional. Para optimizar la selección de variables se utiliza un método de anelación simulada con un criterio de aceptación tipo Metrópolis. Mayores detalles de la implementación del método k NN, incluyendo el procedimiento de anelación simulada para el muestreo de los descriptores, están descritos en la literatura (Zheng, 2000).

Para evaluar qué tan robustos fueron los modelos QSAR, los valores de q^2 obtenidos con los grupos de entrenamiento con actividades biológicas experimentales fueron comparados con los valores obtenidos de grupos de datos con actividades biológicas aleatorias. La significancia estadística de los modelos QSAR para los grupos de entrenamiento fue evaluada con la prueba de hipótesis estándar (Gilbert, 1976). En este método se postulan dos hipótesis alternas: (1) para H_0 , $h = \mu$; (2) para H_1 , $h > \mu$, donde μ es el valor promedio de q^2 para los grupos de datos aleatorios y h es el valor de q^2 para el grupo de datos con actividades experimentales. De esta forma, la hipótesis nula, H_0 , establece que el modelo QSAR para el grupo de datos con actividades experimentales no es significativamente mejor que los modelos obtenidos con actividades aleatorias. Por su parte, la hipótesis alterna, H_1 , asume lo contrario (que el modelo obtenido con datos experimentales es significativamente mejor que los modelos obtenidos con datos aleatorios). La toma de decisión está basada en la prueba estándar de una cola, la cual involucra el siguiente procedimiento.

(1) Determinar el valor promedio de q^2 (μ) y su desviación estándar (σ) para grupos de datos con actividades aleatorias.

(2) Calcular el valor de Z que corresponda al valor de q^2 para el grupo de datos con actividades experimentales: $Z = (h - \mu) / \sigma$

(3) Comparar el valor de Z con los valores críticos tabulados de Z_c a diferentes niveles de significancia (α) (Gilbert, 1976) para determinar el nivel al cual H_0 debe ser rechazado. Si el valor de Z es más alto que el valor tabulado de Z_c (Tabla 3.1) se concluye que al valor de significancia al que corresponde el valor de Z_c , H_0 debe ser rechazado y, por tanto, H_1 debe ser aceptado. Si este es el caso, se concluye que el resultado obtenido para el grupo de datos con actividades experimentales es significativamente mejor que los resultados obtenidos para grupos de datos con actividades aleatorias a ese nivel dado de significancia.

Tabla 3.1. Valores frecuentemente usados de α y los valores críticos de Z_c para la prueba de una cola (Gilbert, 1976)

α	Z_c
0.10	1.28
0.05	1.64
0.01	2.33

Para la selección de los grupos de entrenamiento y grupos de prueba se utilizó el método de exclusión de esferas descrito en la literatura (Golbraikh, 2002). Este algoritmo permite la construcción de grupos de entrenamiento y prueba de varios tamaños que cubren todo el espacio de los descriptores de todos los compuestos. Para estimar el poder predictivo de un modelo QSAR se utilizaron los siguientes parámetros (Sachs, 1984). (i) Coeficiente de correlación R^2 entre las actividades predichas y observadas para el grupo de entrenamiento; (ii) coeficientes de determinación (actividades predichas versus observadas R_0^2 , y actividades observadas versus predichas $R_0'^2$); (iii) pendientes k y k' de las líneas de regresión (actividades predichas versus observadas y actividades observadas versus predichas) a través del origen. En este trabajo se consideró que un modelo QSAR tiene un poder predictivo aceptable si se cumplen las siguientes condiciones (Golbraikh, 2002b):

$$q^2 > 0.5$$

$$R^2 > 0.6$$

$$|R_0^2 - R_0'^2| < 0.3$$

$$0.85 \leq k \leq 1.15 \text{ o } 0.85 \leq k' \leq 1.15$$

3.4. Búsqueda automática en bases de datos

La búsqueda se efectuó con programas desarrollados en el Laboratorio de Modelado Molecular de la Universidad de Carolina del Norte en Chapel Hill dirigido por el profesor Alexander Tropsha.

La estrategia de la búsqueda en la base de datos realizada en este trabajo hace uso de modelos QSAR validados para series de compuestos con actividad biológica. Las series de compuestos que se estudiaron fueron los análogos de piridinona con actividad inhibitoria de la TR de VIH-1 y los derivados de bencimidazol y benzoxazol con actividad antiparasitaria. El diagrama de flujo del procedimiento de la búsqueda está descrito en la Figura 3.2. Este procedimiento involucra las siguientes etapas principales:

1. Para un grupo de datos de compuestos con estructuras y actividades conocidas, desarrollar modelos QSAR utilizando selección de variables y validar a los modelos; definir farmacóforos de descriptores (Tropsha, 2001b) y dominios de aplicabilidad (Tropsha, 2003b) para todos los modelos.
2. Calcular los descriptores químicos usados en el desarrollo de los modelos QSAR para todos los compuestos en la base de datos disponible. Para estos estudios se usaron los índices topológicos moleculares calculados con el programa MolConnZ (MolConnZ). Los

descriptores de la base de datos fueron calculados en el Laboratorio de Modelado Molecular de la Universidad de Carolina del Norte en Chapel Hill. La base de datos empleada fue la del Instituto Nacional de Cáncer (NCI Database).

3. Calcular los valores de similitud química (se usaron las distancias Euclidianas en el espacio de los descriptores farmacofóricos) entre todas las moléculas activas (estructuras usadas para el desarrollo de los modelos QSAR) y cada una de las estructuras de la base de datos.
4. Ordenar a todas las estructuras de la base de datos de acuerdo a su similitud a una molécula o moléculas activas y seleccionar a un número M de estructuras dentro de ciertos límites establecidos de similitud.
5. Predecir los valores de la actividad biológica para las M estructuras a partir de los modelos QSAR previamente desarrollados usando un dominio de aplicabilidad.
6. Seleccionar a las estructuras con predicciones altas de actividad biológica como prototipos computacionales.

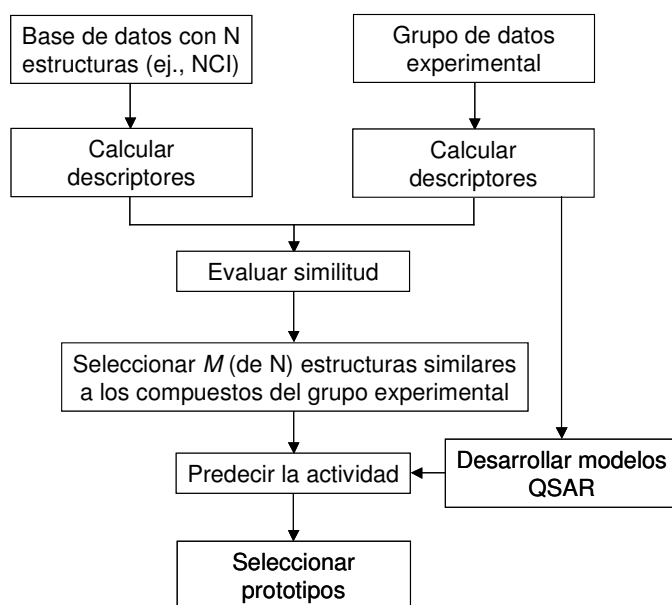


Figura 3.2. Diagrama de flujo de la búsqueda en la base de datos empleando modelos QSAR.

Un procedimiento similar se utilizó recientemente para el descubrimiento de nuevos agentes anticonvulsivos (Shen, 2004).

Resultados y discusión

Parte I

Compuestos antivirales

Capítulo 4

Acoplamiento molecular con la transcriptasa reversa de VIH-1 y diseño de nuevos compuestos

4.1. Inhibidores de la transcriptasa reversa de VIH-1

El síndrome de la inmunodeficiencia adquirida (SIDA) es la etapa terminal de la infección por el virus de la inmunodeficiencia humana (VIH), principalmente por el virus de la familia VIH-1 (Siliciano, 2001). Se estima que en el mundo hay aproximadamente cuarenta millones de personas infectadas por el VIH (ONUSIDA, 2004).

Con el afán de desarrollar terapias efectivas contra esta epidemia mundial se han logrado muchos avances en el conocimiento de la estructura y proceso replicativo del VIH (Parniak, 2004). Hoy en día se han identificado diversos blancos claves para detener el proceso de infección (De Clercq, 2004). Hasta la fecha se han aprobado veinte medicamentos para el tratamiento del SIDA (Balzarini, 2004). Con la llamada terapia anti-retroviral altamente activa se han visto avances significativos en el tratamiento de esta enfermedad. Sin embargo, el alto costo de esta terapia la hace inaccesible a la mayoría de la población infectada alrededor del mundo. Otro problema que se presenta en pacientes que tienen acceso a los medicamentos es el surgimiento de cepas resistentes a la mayoría, si no es que a todos, los fármacos disponibles (Parniak, 2004).

Uno de los blancos más atractivos para el tratamiento del SIDA es la enzima transcriptasa reversa (TR). Esta enzima es un heterodímero asimétrico (Figura 4.1) constituido por las subunidades p66, que contiene 560 amino ácidos, y p51 con 440 amino ácidos (Ding, 1997). Los inhibidores de esta enzima se clasifican en nucleosídicos y no-nucleosídicos dependiendo de su modo de acción. Los nucleosídicos se unen en el sitio catalítico mientras que los no-nucleosídicos, por ejemplo la nevirapina (cf. Figura 4.3), se unen en un sitio alostérico (Figura 4.2) (De Clercq, 2002). Una de las ventajas del segundo grupo es su alta selectividad lo que los hace relativamente poco tóxicos. Sin embargo, los inhibidores no-nucleosídicos generan fácilmente mutaciones en el virus con diferentes niveles de resistencia a los fármacos (Balzarini, 2004). La primera generación de inhibidores no-nucleosídicos de la transcriptasa

reversa (INNTR), tales como la nevirapina y delavirdina (cf., Figura 4.3), pierden su capacidad inhibitoria contra cepas que contienen al menos una mutación en la TR. La segunda generación de INNTR requiere al menos de dos o más mutaciones en la TR para que pierdan significativamente su potencia antiviral. Evidentemente, se requieren de tiempos más prolongados para tener una resistencia significativa contra los INNTR de segunda generación (Balzarini, 2004).

Actualmente se han aprobado a tres inhibidores no-nucleosídicos para su uso clínico: nevirapina (VIRAMUNE®), delavirdina (RESCRIPTOR®) y efavirenz (SUSTIVA®) (Figura 4.3). Sin embargo, todos estos fármacos generan resistencia. Los INNTR han probado ser especialmente benéficos cuando se utilizan en combinación con otros fármacos, especialmente con inhibidores de la proteasa de VIH e inhibidores nucleosídicos de la TR (Balzarini, 2004).

Además de los INNTR aprobados para su uso clínico, se han descrito a muchos otros compuestos como inhibidores de la TR que corresponden a diversos tipos estructurales (De Clercq, 1998). Algunos de ellos se encuentran en fase preclínica o clínica. Ejemplos de estos compuestos son la etravirina y TMC-125 (De Clercq, 2004).

Se han obtenido diversas estructuras cristalográficas de la TR tanto en forma libre (Esnouf, 1995) como formando un complejo con diferentes inhibidores (por ejemplo Chan, 2001; Das, 2004; Esnouf, 1997; Hogberg, 2000; Hopkins, 1996; Hsiou, 1998; Ren, 1995; Ren, 1995b; Ren, 1998; Ren, 1999; Ren, 2000; Ren, 2000b). Esto ha permitido conocer con detalle el sitio de unión de los INNTR y las interacciones ligando-receptor con la TR. El sitio de unión de los inhibidores no-nucleosídicos está formado por los residuos Pro95, Leu100, Lys101, Lys103, Val106, Val179, Tyr181, Tyr188, Gly190, Phe227, Trp229, Leu234, His235, Pro236 y Tyr318 de la subunidad p66 y Glu138 de la subunidad p51 (Ding, 1997). Los INNTR están a una distancia aproximada de 10 a 15 Å de los residuos catalíticos Asp110, Asp185 y Asp186 (Figura 4.2).

Las estructuras cristalográficas de los complejos indican que algunos de los residuos del sitio de unión adoptan diferentes conformaciones cuando están unidos a diferentes inhibidores (Esnouf, 1997; Hopkins, 1996; Hsiou, 1998). Esto da origen a variaciones en la geometría del sitio de unión dependiendo del inhibidor que esté unido. Una modificación marcada en la geometría ocurre cuando la TR está formando un complejo con los derivados de la 1-[(2-hidroxietoxi)-metil]-6-(feniltio)timina (derivados HEPT) (Hopkins, 1996). A pesar de estas variaciones, muchos INNTR adoptan un modo de unión semejante que comúnmente se refiere en la literatura como una conformación tipo mariposa (Figura 4.4) (Chan, 2001; Hopkins, 1996; Ren, 1995; Ren, 2000).

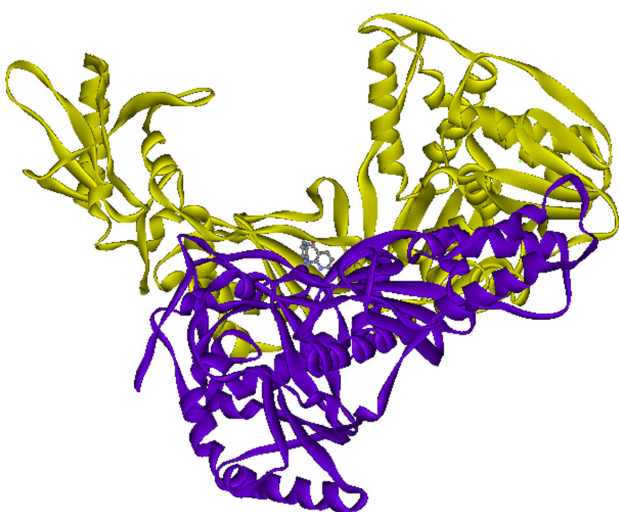


Figura 4.1. Estructura de la TR de VIH (código PDB: 1vrt). Residuos de la cadena A están en amarillo y de la cadena B en púrpura. La nevirapina (esferas y barras) muestra el sitio de unión de los INNTR.

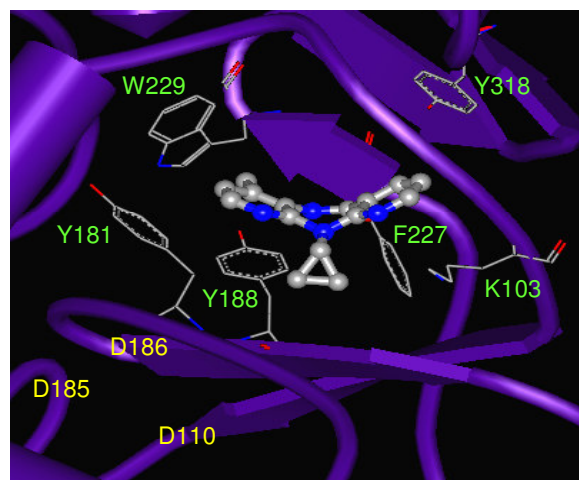


Figura 4.2. Nevirapina en el sitio de unión de los INNTR. Residuos representativos del sitio alostérico están señalados en verde. Residuos clave del sitio catalítico están en amarillo.

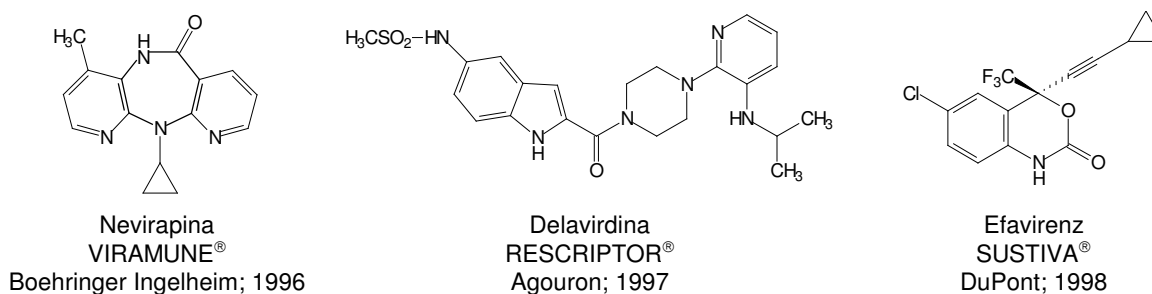


Figura 4.3. Inhibidores no-nucleosídicos de la TR en uso clínico. Se indica el año de aprobación o comercialización.

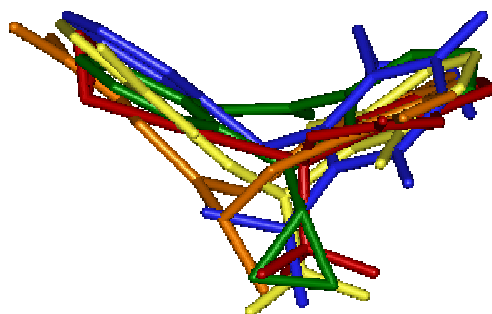


Figura 4.4. Conformación tipo mariposa que adopta la nevirapina (verde), efavirenz (rojo), 9-Cl-TIBO (naranja), Cl-α-APA (amarillo) y MKC-442 (azul) en el sitio de unión de los INNTR.

Los derivados de piridinona (Goldman, 1991) son una clase de INNTR (Figura 4.5). Mediante estudios cinéticos (Goldman, 1991) y análisis de mutaciones que inducen resistencia, se concluyó que los análogos de piridinona ocupan el mismo sitio de unión de otros INNTR (Nunberg, 1991). Sin embargo, no hay una estructura cristalográfica que indique la orientación y conformación que adoptan estos compuestos dentro del sitio de unión.

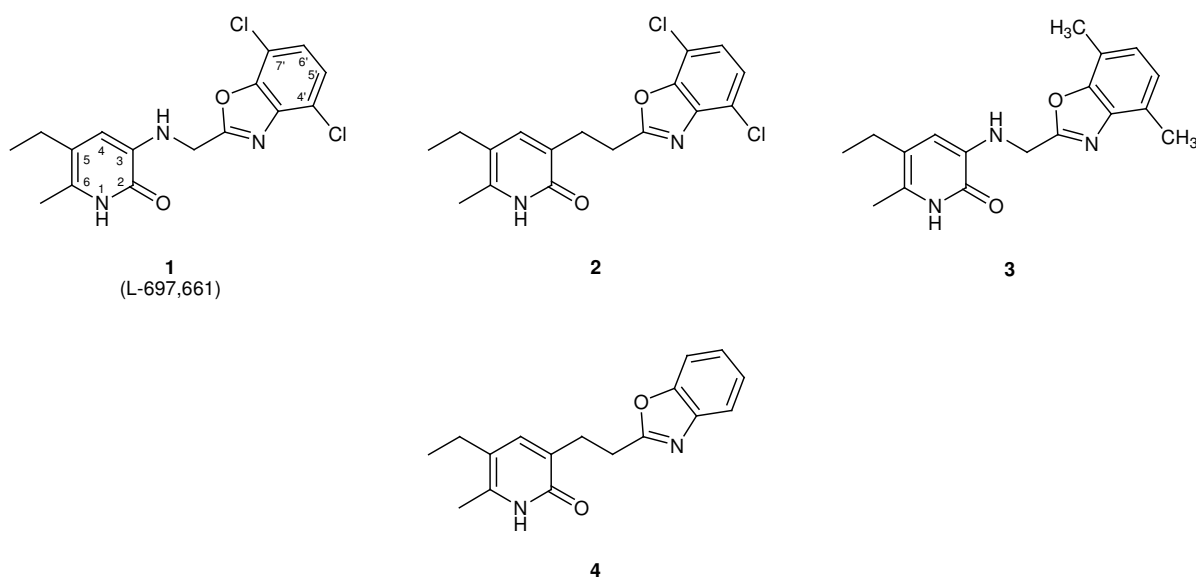


Figura 4.5. Estructuras de cuatro análogos de piridinona con potente actividad inhibitoria de la TR.

Se han realizado numerosos estudios de relación estructura-actividad (SAR) con los análogos de piridinona logrando el desarrollo de inhibidores potentes (Hoffman, 1992; Hoffman, 1993; Saari, 1991; Saari, 1992). Uno de los compuestos desarrollados, L-697,661 (Figura 4.5), fue sometido a estudios clínicos mostrando buena actividad. Sin embargo, emergieron cepas resistentes del virus en poco tiempo (6-12 semanas) (Davey, 1993; Saag, 1993). Las mutaciones, principalmente de los residuos en las posiciones 103 y 181 de la TR, son los responsables de la generación de cepas resistentes. Esto ocurre no solamente para L-697,661, sino también para otros derivados de piridinona (Byrnes, 1993; Saag, 1993; Sardana, 1992). De los estudios SAR se concluyó que el patrón de sustitución en el anillo de la piridinona fue importante para la potencia inhibitoria de la TR (Hoffman, 1992; Hoffman, 1993; Saari, 1991). De los sustituyentes evaluados se encontró que un etilo en la posición cinco y un metilo en la posición seis dan a las moléculas la mayor potencia inhibitoria (Hoffman, 1992; Hoffman, 1993; Saari, 1992). También se estableció que se necesita una unión flexible de dos átomos, concretamente un grupo aminometileno o etileno, entre los anillos de la piridinona y el benzoxazol (Hoffman, 1992; Hoffman, 1993).

Se han reportado diversos análisis QSAR-2D con los derivados de piridinona (Gaarg, 1999; Gupta, 2002). Como resultado de estos estudios se puede proponer que la hidrofobicidad de los cloros o metilos de **1** y **3** (Figura 4.5) están relacionados con la actividad de estos compuestos.

El diseño de nuevos INNTR es especialmente atractivo por sus características de alta selectividad y baja toxicidad (Balzarini, 2004). Los nuevos inhibidores de este tipo representan potenciales alternativas a los INNTR actuales que continúan presentando problemas de resistencia y algunos de ellos de toxicidad (De Clercq, 2004). Por este motivo, los compuestos diseñados no solamente deben ser activos contra cepas nativas sino también contra cepas resistentes.

Estrategia para el estudio computacional de inhibidores de la TR de VIH-1

El conocimiento de la estructura tridimensional del blanco de acción de los INNTR permite utilizar métodos computacionales para el estudio de la interacción fármaco-receptor de derivados de piridinona y hacer diseño basado en la estructura del receptor de nuevos inhibidores. Por su parte, los estudios SAR, reportados para un número considerable de piridinonas, permiten aplicar métodos de diseño basado en la estructura del ligando.

La estrategia del estudio y diseño de inhibidores de la TR está esquematizada en la Figura 4.6. En este capítulo se presentan los resultados de los experimentos de acoplamiento molecular para analizar las interacciones fármaco-receptor de los derivados de piridinona con la TR. Los experimentos de acoplamiento se efectuaron utilizando diversas estructuras cristalográficas de la TR para considerar la flexibilidad del sitio de unión. Esta estrategia para modelar la flexibilidad el sitio de unión se ha reportado previamente (Mao, 2000; Ragno, 2005; Titmuss, 1999; Yadav, 1998). Los estudios de interacción fármaco-receptor constituyen la base para el diseño de nuevos derivados de piridinona con potencial actividad inhibitoria de la TR, especialmente contra cepas resistentes. Cabe mencionar que el acoplamiento molecular automatizado se ha empleado para entender el modo de unión de otras clases estructurales de INNTR (Barreca, 1999; Ragno, 2005; Titmuss, 1999; Zhou, 2002). De manera semejante, se ha reportado recientemente la aplicación de métodos computacionales al diseño exitoso de inhibidores de la TR (Artico, 2000; Mai 2001, Ragno, 2004; Smith, 2003).

En los capítulos 5 y 6 se presentan los resultados de los estudios de diseño basado en la estructura del ligando y la combinación de técnicas de diseño basado en la estructura del ligando con acoplamiento molecular.

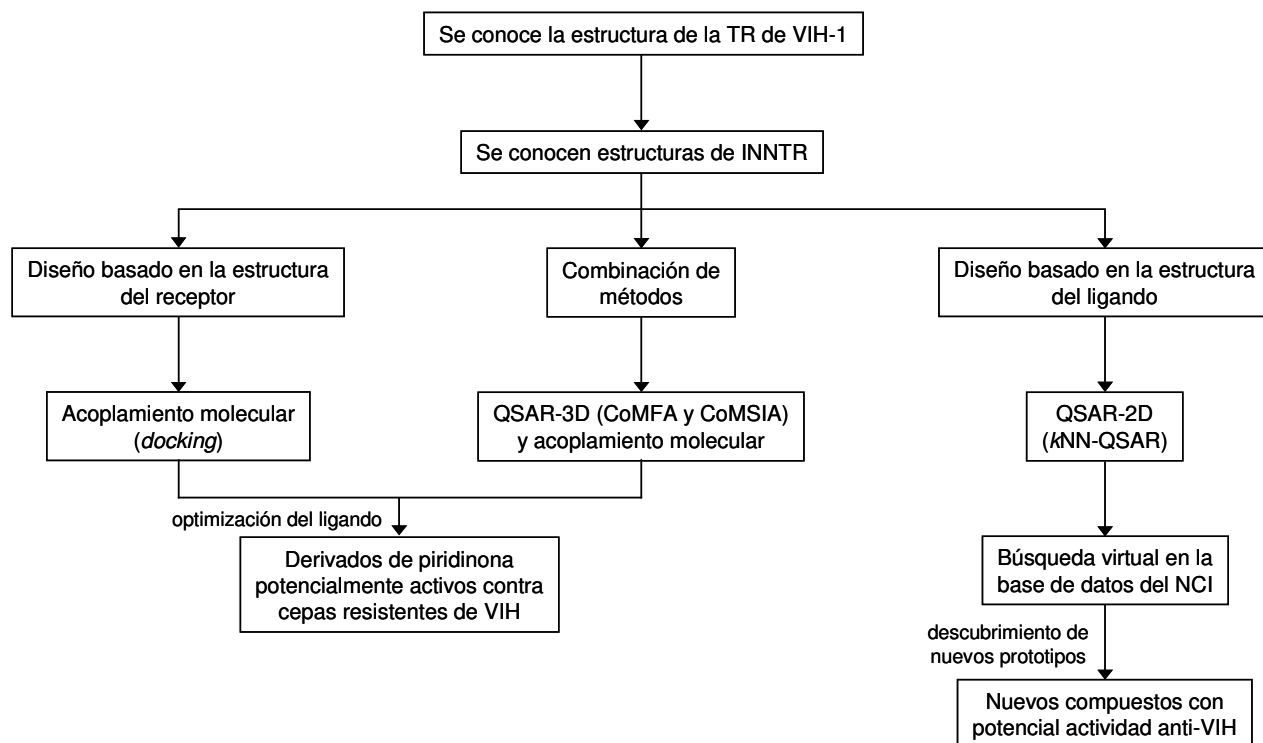


Figura 4.6. Estrategia para el diseño de compuestos con potencial actividad inhibitoria de la TR de VIH-1.

4.2. Estudio de la interacción de análogos de piridinona con la transcriptasa reversa de VIH-1

Se utilizaron las estructuras cristalográficas de la TR de VIH-1 formando complejos con los inhibidores no-nucleosídicos nevirapina, efavirenz (Figura 4.3), 1051U91, 9-CI-TIBO, Cl- α -APA, MKC-442, TNK-651, UC-781 y S-1153 (Figura 4.7). De las estructuras cristalográficas disponibles en el *Protein Data Bank* (RCSB) estos complejos fueron seleccionados de acuerdo a criterios de resolución, diversidad estructural de los ligandos y diversidad de la geometría del sitio de unión. Dentro de los ligandos seleccionados se encuentran los inhibidores de segunda generación efavirenz, UC-781 y S-1153. Como se discute más adelante, los complejos con los inhibidores MKC-442 y TNK-651 representan geometrías del sitio de unión particulares. En la Tabla 4.1 están resumidos los datos y referencias de estas estructuras.

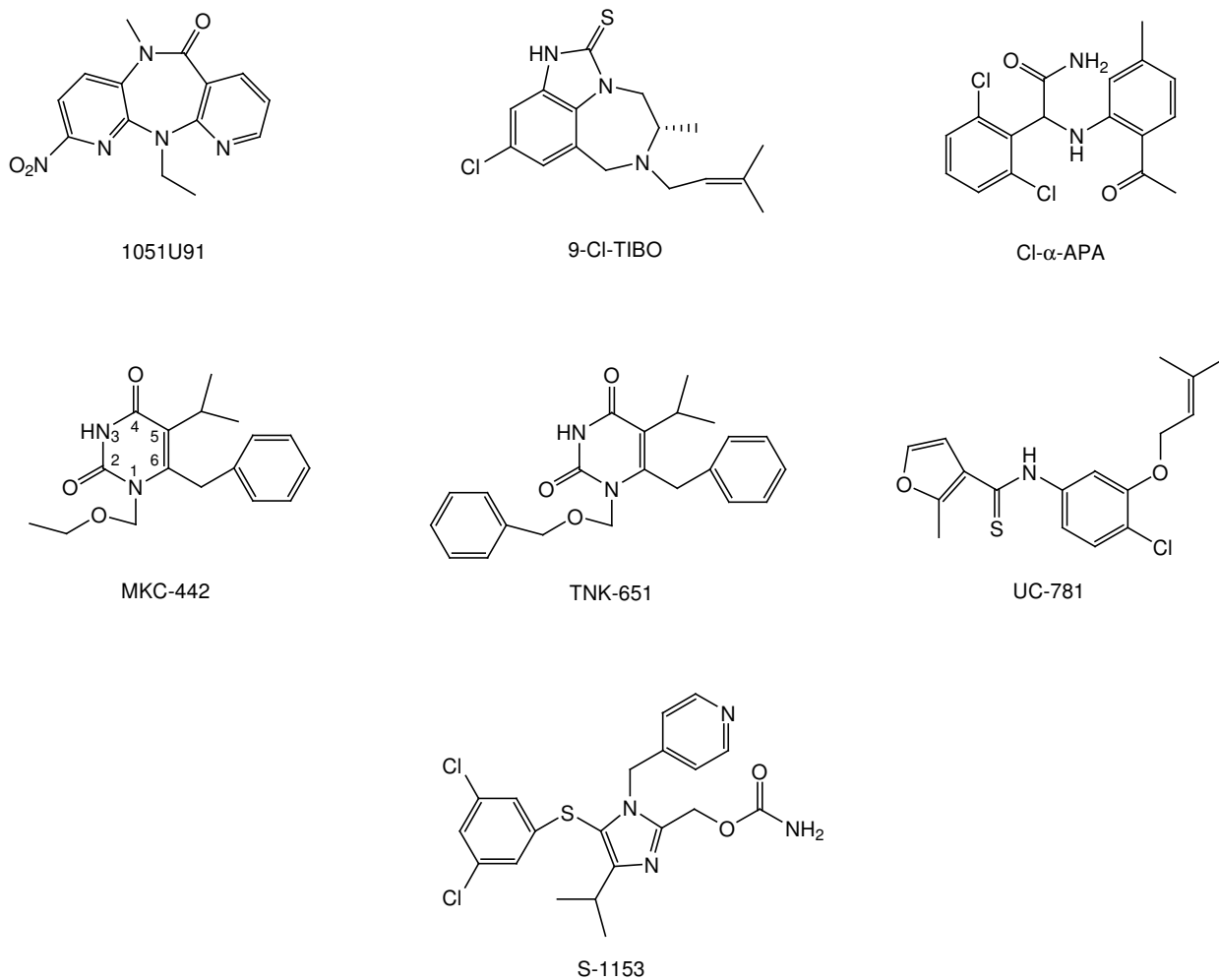


Figura 4.7. Estructuras de los INNTR encontrados en las estructuras cristalográficas usadas en este estudio.

Tabla 4.1. Estructuras cristalográficas empleadas en este estudio

INNTR	Código PDB	Resolución	Referencia
Nevirapina	1VRT	2.2 Å	Ren, 1995
1051U91	1RTH	2.2 Å	Ren, 1995
9-Cl-TIBO	1REV	2.6 Å	Ren, 1995b
Cl- α -APA	1VRU	2.4 Å	Ren, 1995
Efavirenz	1FK9	2.5 Å	Ren, 2000
MKC-442	1RT1	2.55 Å	Hopkins, 1996
TNK-651	1RT2	2.55 Å	Hopkins, 1996
UC-781	1RT4	2.90 Å	Ren, 1998
S-1153	1EP4	2.50 Å	Ren, 2000b

Validación del protocolo de acoplamiento

Antes de realizar el acoplamiento de los análogos de piridinona con las estructuras cristalográficas de la TR, se realizó la validación para cada una de estas estructuras. La validación consistió en predecir la orientación y conformación de unión de cada INNTR en su respectiva estructura cristalográfica. En la Tabla 4.2 se resumen los resultados de la validación.

Tabla 4.2. Acoplamiento de los INNTR encontrados en los cristales

INNTR	Miembros en el grupo ^a	Energía más baja (kcal/mol) ^b	Energía promedio (kcal/mol) ^c	RMSD (Å)
Nevirapina	100	-9.68	-9.67	0.70
1051U91	100	-9.97	-9.95	1.25
9-CI-TIBO	100	-11.83	-11.82	0.89
CI- α -APA	68	-10.49	-10.31	0.96
Efavirenz	99	-10.97	-10.96	0.24
MKC-442	97	-11.15	-11.05	0.75
TNK-651	88	-14.01	-13.82	0.46
UC-781	93	-12.38	-12.03	0.92
S-1153	59	-15.63	-15.30	0.54

^aSe muestra el primer grupo de un total de 100 repeticiones.

^bValor para la estructura óptima del grupo. Las energías de acoplamiento son las determinadas por AutoDock.

^cEnergía de acoplamiento promedio del grupo.

Para cada INNTR cristalográfico, AutoDock predijo satisfactoriamente las posiciones de unión. Esto puede observarse en los valores de las desviaciones (RMSD) de las coordenadas de los átomos predichas por AutoDock respecto a las coordenadas de los átomos encontradas en las estructuras cristalográficas (Tabla 4.2) (Figura 4.8). En todos los casos, estos valores fueron inferiores o cercanos a 1.0 Å. Las altas poblaciones obtenidas en cada grupo indican que AutoDock encontró, en la mayoría de los casos, una forma de unión principalmente.

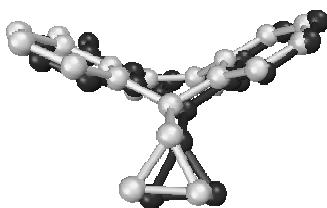


Figura 4.8. Comparación entre el modo de unión de la nevirapina encontrada en la estructura cristalográfica (blanco) y la posición de unión predicha por AutoDock (negro).

Acoplamiento de los análogos de piridinona

Para explorar el modo de unión de los análogos de piridinona con la TR se efectuó el acoplamiento de cuatro moléculas con potente actividad inhibitoria (Figura 4.5) (cf., Tabla 4.4).

Se efectuó primero el acoplamiento de **1-4** en la geometría de unión de nevirapina, estructura con la mejor resolución disponible. Los resultados están en la Tabla 4.3. A partir de la conformación de menor energía encontrada para cada uno de los cuatro compuestos, se efectuó el acoplamiento en las correspondientes geometrías de los ocho INNTR restantes (Tabla 4.3). Los valores de RMSD en la Tabla 4.3 son una aproximación de la similitud entre la conformación encontrada para cada compuesto en un sitio de unión determinado y el sitio de unión de la nevirapina. Esto se debe a que, antes de los experimentos de acoplamiento, las coordenadas de las diferentes geometrías fueron superpuestas con la estructura de la TR en complejo con la nevirapina (ver la sección de Metodología computacional para detalles). De acuerdo a los grupos más poblados, los compuestos **1-4** ocupan el mismo sitio de unión del correspondiente INNTR y tienen la misma orientación que la nevirapina, 1051U91, 9-CI-TIBO, CI- α -APA, efavirenz, UC-781 y S-1153. En los sitios de unión de la nevirapina y CI- α -APA, el grupo más poblado de cada compuesto también mostró la menor energía de acoplamiento promedio (Figura 4.9).

En los sitios de unión de MKC-442 y TNK-651 (derivados HEPT) se calculó que los análogos de piridinona tendrían un modo de unión diferente que el del inhibidor encontrado en la estructura cristalográfica. En estos sitios de unión, el anillo de benzoxazol de los análogos de piridinona se localiza en la misma región que el grupo etoximetilo de MKC-442 y que el grupo benziloximetilo de TNK-651 (Figura 4.10). En todos los sitios de unión se calculó que los análogos de piridinona tendrían una conformación tipo mariposa.

Cuando se emplearon los sitios de unión de nevirapina, 1051U91, 9-CI-TIBO, CI- α -APA, efavirenz, UC-781 y S-1153, se encontró una buena relación entre las energías de acoplamiento calculadas y la actividad biológica. En la Tabla 4.4 se resumen los valores de IC₅₀ reportados para **1-4** (Goldman, 1991) y el promedio de las energías. Para **2**, la piridinona más activa, se estima la menor energía. Para **4**, la piridinona menos activa de esta serie, la energía de acoplamiento es la más alta. Para **1** y **3**, con actividades casi iguales, la energía también es muy similar (Tabla 4.4). La buena relación entre la actividad biológica y las energías calculadas indican que los cálculos de acoplamiento dieron modos de unión razonables para las piridinonas en estas geometrías de unión.

Tabla 4.3. Acoplamiento de 1-4 en las diferentes geometrías de unión de la TR

Sitio de unión	Compuesto	Grupo ^a	Miembros en el grupo	Energía más baja (kcal/mol) ^b	Energía promedio (kcal/mol) ^c	RMSD (Å) ^d
Nevirapina	1	1	92	-11.13	-11.08	0.00
	2	1	93	-11.69	-11.64	0.00
	3	1	94	-11.18	-11.14	0.00
	4	1	96	-10.63	-10.58	0.00
1051U91	1	1	91	-10.90	-10.81	0.33
	2	1	99	-11.73	-11.70	0.55
	3	1	93	-10.97	-10.89	0.34
	4	1	3	-10.85	-10.74	2.65
	4	2	89	-10.58	-10.45	1.08
9-Cl-TIBO	1	1	3	-11.50	-11.40	3.01
	1	4	49	-11.08	-10.90	0.75
	2	1	93	-12.12	-11.75	1.07
	3	1	3	-11.47	-11.37	2.95
	3	6	65	-11.13	-10.98	0.74
	4	1	97	-11.23	-11.13	0.78
Cl- α -APA	1	1	92	-10.96	-10.82	0.68
	2	1	94	-11.65	-11.53	0.62
	3	1	96	-11.02	-10.91	0.67
	4	1	95	-10.68	-10.56	0.63
Efavirenz	1	1	10	-11.49	-11.25	1.47
	1	2	74	-11.47	-11.26	0.56
	2	1	84	-11.95	-11.84	0.63
	3	1	75	-11.57	-11.39	0.60
	4	1	80	-11.12	-10.97	0.53
MKC-442	1	1	24	-11.76	-11.66	3.18
	2	1	4	-12.09	-12.05	3.06
	2	2	35	-12.04	-11.99	2.89
	3	1	25	-11.80	-11.72	3.15
	4	1	5	-11.21	-11.00	3.29
	4	9	59	-10.66	-10.60	0.59

^aSe muestran los mejores grupos de un total de 100 repeticiones.

^bValor para la estructura óptima del grupo. Las energías de acoplamiento son las determinadas por AutoDock.

^cEnergía de acoplamiento promedio del grupo.

^dLas coordenadas encontradas en el acoplamiento con el sitio de unión de la nevirapina se emplearon como referencia.

Tabla 4.3. (continuación)

Sitio de unión	Compuesto	Grupo ^a	Miembros en el grupo	Energía más baja (kcal/mol) ^b	Energía promedio (kcal/mol) ^c	RMSD (Å) ^d
TNK-651	1	1	1	-11.84	-11.84	3.88
	1	2	32	-11.71	-11.66	3.03
	2	1	4	-12.49	-12.29	3.52
	2	2	34	-12.19	-12.13	2.87
	3	1	3	-12.28	-12.19	3.63
	3	2	34	-11.76	-11.71	2.97
	4	1	3	-11.74	-11.53	3.39
	4	8	29	-10.89	-10.82	2.76
UC-781	1	1	16	-11.39	-11.21	3.15
	1	4	67	-10.60	-10.54	0.68
	2	1	15	-12.05	-11.88	2.78
	2	2	75	-11.46	-11.38	0.74
	3	1	13	-11.41	-11.14	3.15
	3	3	78	-10.82	-10.77	0.79
	4	1	2	-10.76	-10.49	2.87
	4	2	93	-10.73	-10.67	0.62
S-1153	1	1	2	-10.93	-10.91	3.08
	1	3	71	-10.76	-10.62	1.22
	2	1	73	-11.30	-11.10	1.34
	3	1	3	-10.90	-10.83	3.65
	3	2	88	-10.77	-10.70	1.37
	4	1	18	-10.41	-10.27	3.69
	4	9	53	-10.03	-9.98	1.14

^aSe muestran los mejores grupos de un total de 100 repeticiones.

^bValor para la estructura óptima del grupo. Las energías de acoplamiento son las determinadas por AutoDock.

^cEnergía de acoplamiento promedio del grupo.

^dLas coordenadas encontradas en el acoplamiento con el sitio de unión de la nevirapina se emplearon como referencia.

Después del acoplamiento, la geometría del sitio de unión de la nevirapina en complejo con la conformación mejor evaluada energéticamente de **1-4** se optimizó con mecánica molecular (ver la sección de Metodología computacional para detalles). Se observaron muy pocas variaciones entre las conformaciones predichas por AutoDock y las conformaciones de **1-4** después de la optimización geométrica.

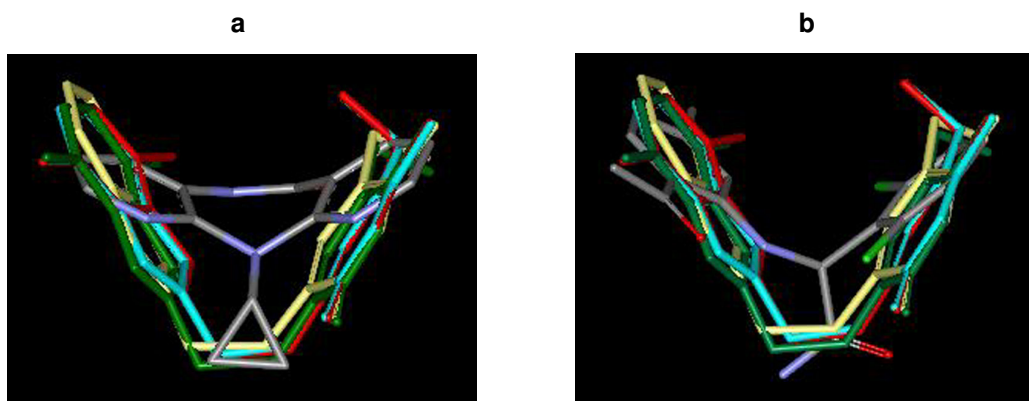


Figura 4.9. Modo de unión predicho para **1** (rojo), **2** (verde), **3** (azul) y **4** (amarillo) en el sitio de unión de (a) nevirapina y (b) Cl- α -APA. Nevirapina y Cl- α -APA (color por tipo de átomo) se muestran para comparación. Para mayor claridad, se omiten los hidrógenos.

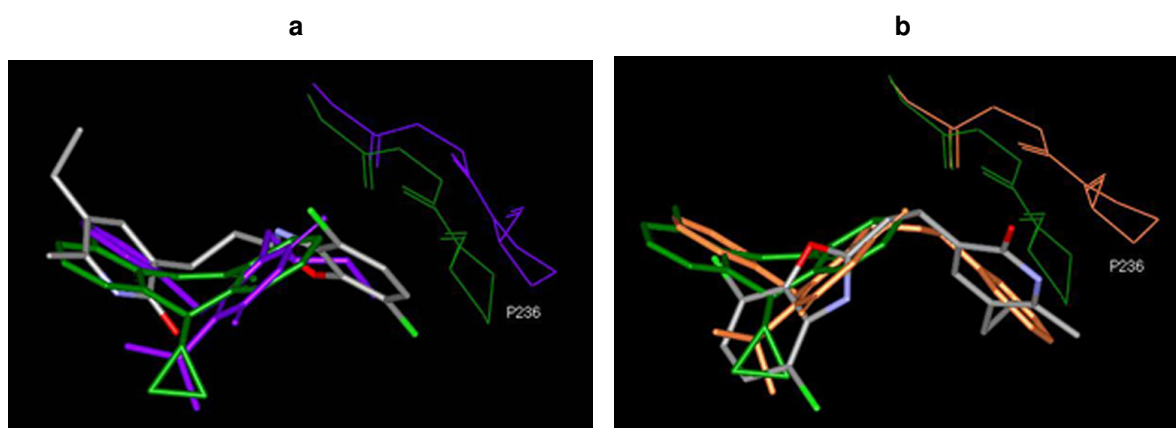


Figura 4.10. Modo de unión predicho para **2** (color por tipo de átomo) en el sitio de unión de (a) MKC-442 (barras color púrpura) y (b) TNK-651 (barras color naranja). En ambos casos se compara con el modo de unión de la nevirapina (barras color verde). Se muestra la posición de la Pro236 cuando el inhibidor es MKC-442 (en líneas color púrpura), TNK-651 (en líneas color naranja) y nevirapina (en líneas color verde). Para mayor claridad, se omiten los hidrógenos.

Tabla 4.4. Comparación entre la energía de unión y la actividad biológica para **1-4**

Compuesto	IC ₅₀ (nM) ^a	Energía de acoplamiento (kcal/mol) ^b
1	19	-10.99
2	9.6	-11.70
3	20	-11.07
4	22	-10.71

^a(Goldman, 1991).

^bPromedio de las energías de acoplamiento de la estructura óptima en los grupos más poblados obtenidos en los sitios de unión de nevirapina, 1051U91, 9-CI-TIBO, Cl- α -APA, efavirenz, UC-781 y S-1153.

Los resultados de acoplamiento indican que los derivados de piridinona tienen un modo de unión común en las geometrías de los sitios de unión de la nevirapina, 1051U91, 9-Cl-TIBO, Cl- α -APA, efavirenz, UC-781 y S-1153 (Figura 4.9). El carácter hidrófobo del sitio de unión está de acuerdo con los datos de QSAR disponibles para los análogos de piridinona (Garg, 1999; Gupta, 2002). Los análogos de piridinona adoptan una conformación tipo mariposa y pueden superponerse aproximadamente sobre las estructuras cristalográficas de otros INNTR (Figuras 4.9, 4.11). Resultados similares se han obtenido para otros INNTR para los cuales no están disponibles estructuras cristalográficas de los complejos (Zhou, 2002).

Utilizando métodos *ab-initio* se ha calculado que la conformación de menor energía para muchos análogos de piridinona, cuando se encuentran en forma libre, es también tipo mariposa (Parreira, 2001). Conclusiones similares se obtuvieron cuando se realizó un análisis conformacional sistemático para el compuesto **1**. Con diferentes conformaciones de mínima energía como conformaciones iniciales, se realizaron estudios de acoplamiento con la TR. Fue muy interesante concluir que, independientemente de la conformación de inicio, AutoDock encontró el mismo resultado. Esto demostró la capacidad que tiene el protocolo empleado de acoplamiento flexible para convergir en un mismo resultado.

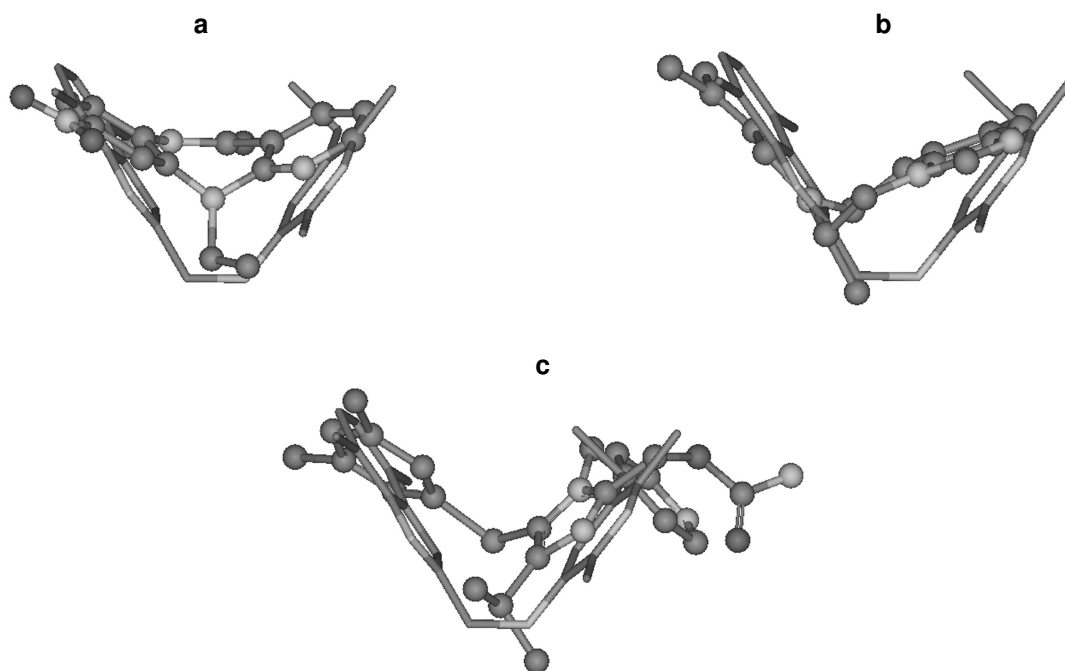


Figura 4.11. Modo de unión predicho para **1** (barras) en el sitio de unión de (a) 1051U91; (b) 9-Cl-TIBO y (c) S-1153. Los INNTR cristalográficos correspondientes (esferas y barras) se muestran para comparación. Para mayor claridad, se omiten los hidrógenos.

El modelo de unión desarrollado para los análogos de piridinona está representado en la Figura 4.12. Un modelo en dos dimensiones de este modelo, generado con el programa Ligplot (Wallace, 1995), se muestra en la Figura 4.13. De acuerdo a los complejos optimizados de los análogos de piridinona con la TR, los residuos que se localizan próximos al anillo de benzoxazol son Pro95, Leu100, Try181 y Tyr188. Los residuos Lys101, Lys103, His235 y Tyr318 forman el sitio de unión del anillo de la piridinona.

Se puede formar un puente de hidrógeno entre el protón polar del anillo de la piridinona y el oxígeno de la cadena principal de la Lys101 (Figuras 4.12 y 4.13). De acuerdo a las estructuras cristalográficas de complejos con otros INNTR, por ejemplo 9-CI-TIBO, efavirenz, MKC-442, TNK-561, UC-781, se forma un puente de hidrógeno similar entre un protón polar del inhibidor con el mismo oxígeno de Lys101 (Chan, 2001; Hogberg, 2000; Hopkins, 1996; Hsiou, 1998; Ren, 1995b; Ren 1998; Ren, 2000a; Ren, 2000b). Este puente de hidrógeno puede anclar al anillo de piridinona en el sitio de unión, tal como se ha propuesto para otros inhibidores (Ren, 1998).

Comparando la posición de unión de los protones de los INNTR que forman el puente de hidrógeno con la Lys101, se encontró una relación estructural entre diversos fragmentos de estos compuestos. El fragmento amida del anillo de piridinona ocupa una región de unión similar que el fragmento amida de efavirenz, MKC-442 y TNK-561 y que el fragmento tioamida de 9-CI-TIBO y UC-781 (Figura 4.14). Esto ayuda a explicar la observación experimental de que no hay una diferencia muy grande en la actividad de las piridinonas comparada con la actividad de las piridintionas (Hoffman, 1993). Estudios QSAR-2D han confirmado esta observación (Garg, 1999).

Los contactos entre el anillo aromático de la Tyr181 y el anillo de benzoxazol de 1-4 (Figuras 4.12, 4.13) pueden estabilizar a los complejos y ser también responsables del decremento de la actividad biológica cuando muta este residuo (Byrnes, 1993; Sardana, 1992;). Interacciones similares se observan entre los anillos aromáticos de otros INNTR y la Tyr181 (Tantillo, 1994). Los contactos hidrófobos con la Lys103 también pueden explicar la menor actividad biológica de los análogos de piridinona cuando sucede la mutación de este residuo (Byrnes, 1993).

Otra característica sobresaliente del modelo de unión de los derivados de piridinona es la región que rodea al grupo aminometileno o etileno que funcionan como puente entre los anillos aromáticos (Figura 4.5). Este puente se localiza en la misma región que el grupo ciclopropilo de la nevirapina, el grupo amida de Cl- α -APA, el grupo trifluorometilo de efavirenz, el metilfuran-3-carbotioamida de UC-781, el isopropilo de S-1153, el etilo de 1051U91 y el grupo 5(S)-metilo de 9-CI-TIBO (Figuras 4.9, 4.11. Cf., Figuras 4.16, 4.18).

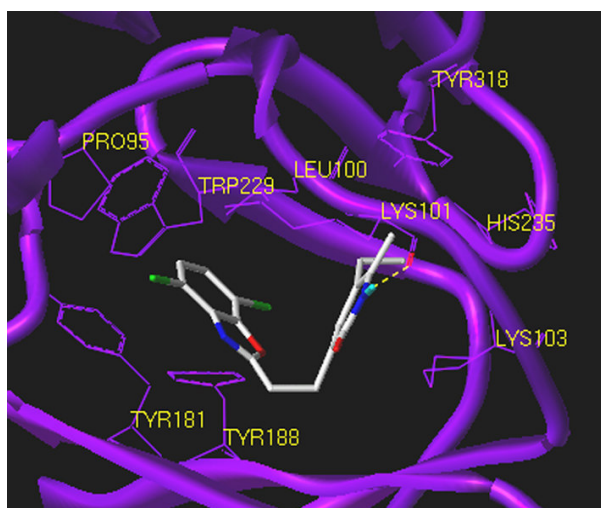


Figura 4.12. Modelo de unión de **2** con la TR. Se muestra la geometría del sitio de unión de la nevirapina. El puente de hidrógeno se muestra con una línea discontinua amarilla. Los residuos que tienen nombre en la figura se encuentran a una distancia de 3.6 Å del anillo de benzoxazol o piridinona. Se incluye el nombre para Trp229 como referencia. Para mayor claridad, se omiten los hidrógenos excepto el hidrógeno polar del anillo de la piridinona.

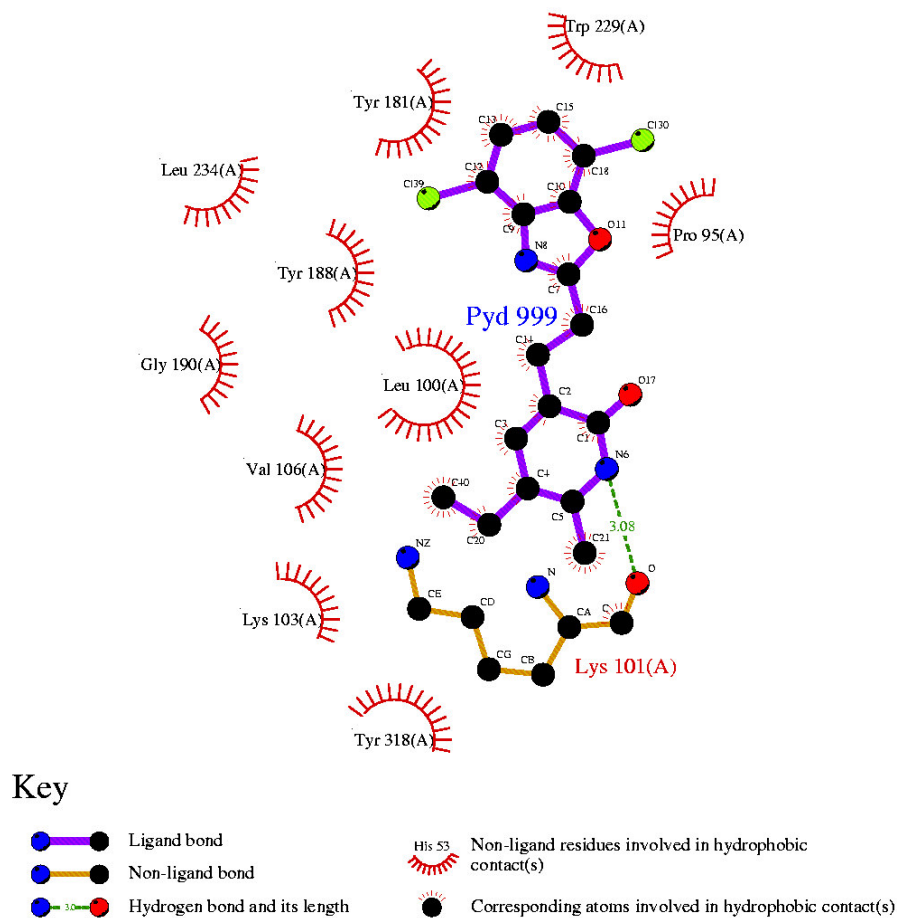


Figura 4.13. Representación 2D de las principales interacciones del compuesto **2** con la TR.

Es interesante observar que aunque las estructuras químicas de estos grupos son diferentes, todas parecen tener una función similar para interactuar con la TR (Hopkins, 1996). Esto está de acuerdo con los estudios SAR disponibles para derivados de piridinona. Más aún, en estudios QSAR-2D para análogos de piridinona, no se observó una diferencia significativa del efecto del grupo aminometileno o etileno sobre la actividad biológica (Garg, 1999).

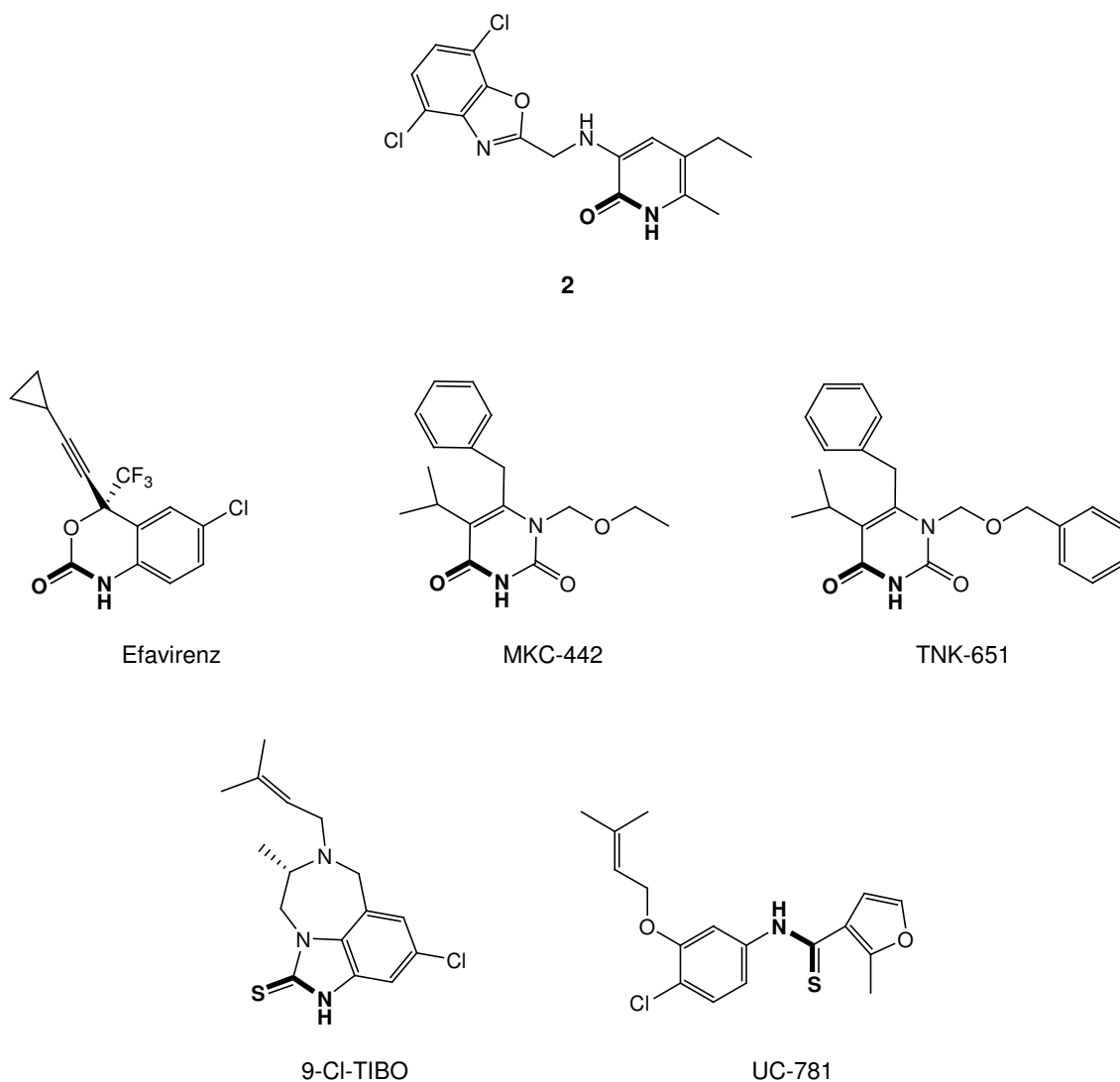


Figura 4.14. Relación entre el fragmento amida del anillo de piridinona con el fragmento amida y tioamida de otros INNTR. Todos ocupan la misma región de unión con la TR.

AutoDock predijo un modo de unión diferente para los análogos de piridinona en las geometrías de unión de MKC-442 y TNK-651. En estas geometrías, uno de los anillos aromáticos de los análogos de piridinona ocupa aproximadamente la misma región de los

sustituyentes *N*-1 de los análogos HEPT (Figura 4.10). En estos casos, la diferente orientación predicha para los análogos de piridinona puede ser explicada por las diferentes geometrías de los sitios de unión. Está descrito que en las estructuras cristalográficas de los derivados HEPT, el sustituyente en *N*-1 causa que la Pro236, y residuos adyacentes, adquieran una conformación parecida a la que se observa en la estructura de la TR cuando no está unida a ningún inhibidor (Hopkins, 1996). AutoDock encontró mejores interacciones si uno de los anillos aromáticos de la piridinonas ocupan el espacio que se forma cuando ocurre el movimiento de la Pro236 y los amino ácidos cercanos a este residuo (Figura 4.10). Debido a que los análogos de piridinona carecen de una estructura equivalente al sustituyente *N*-1 de los derivados HEPT para mantener a la Pro236 en esa conformación característica, la geometría del sitio de unión de los análogos de piridinona debe ser más parecida a la geometría de los sitios de unión de los otros INNTR estudiados.

Acoplamiento de híbridos de piridinona

Se han reportado numerosos híbridos de piridinona con otros INNTR. Los híbridos consisten en una combinación del anillo aromático de piridinona con un fragmento estructural de otro INNTR, específicamente de efavirenz o derivados HEPT (Corbett, 2001; Dollé, 1995; Dollé, 2000). Ejemplo de estas moléculas son los compuestos **5** y **6** (Figura 4.15). Aunque esta estrategia continua empleándose para tratar de disminuir los problemas de resistencia asociados a los análogos de piridinona (Benjahad, 2004), no se han reportado cálculos o estudios a nivel molecular que ayuden a explicar el comportamiento de estos híbridos o que guíen el diseño de nuevas moléculas. Esto está de acuerdo con el hecho de que se continúen sintetizando y evaluando biológicamente a una gran cantidad de compuestos para encontrar a un nuevo híbrido de piridinona líder (Benjahad, 2004).

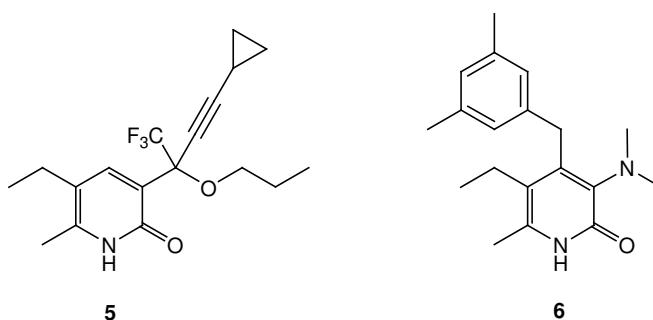


Figura 4.15. Estructuras de híbrido piridinona-efavirenz (**5**) e híbrido piridinona-MKC-442 (**6**).

Para racionalizar el diseño de **5**, considerando la estructura del receptor, se comparó el modo de unión de los análogos de piridinona, en la geometría de unión de efavirenz, con la posición cristalográfica de este inhibidor. Como resultado se observó que el anillo de benzoxazol de los análogos de piridinona ocupa una posición similar a la que ocupa el grupo ciclopropil-propinil de efavirenz (Figura 4.16). De esta misma comparación se puede entender, a nivel molecular, la estrategia de sustituir el anillo de piridinona en la posición 4 con un grupo ciclopropil-propinil para obtener un compuesto activo (**5**) (Corbett, 2001).

Con la finalidad de evaluar esta hipótesis, se realizó el acoplamiento de **5** en diferentes geometrías de unión. Los resultados están mostrados en la Tabla 4.5. En la geometría del sitio de unión de efavirenz, y en la mayoría de los sitios de unión, la conformación con la mejor evaluación de energía mostró la posición y conformación de mariposa esperadas: el grupo ciclopropil-propinil de la molécula híbrida y efavirenz, así como el anillo de piridinona y el anillo de benzoxazin-2-ona, ocupan la misma región de unión en la TR (Figura 4.17).

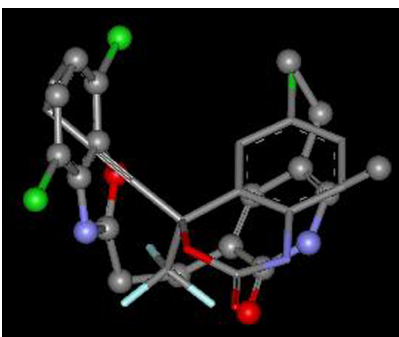


Figura 4.16. Modo de unión predicho para **2** (esferas y barras) en el sitio de unión de efavirenz. Se muestra la posición del efavirenz (barras) para comparación. Para mayor claridad, se omiten los hidrógenos.

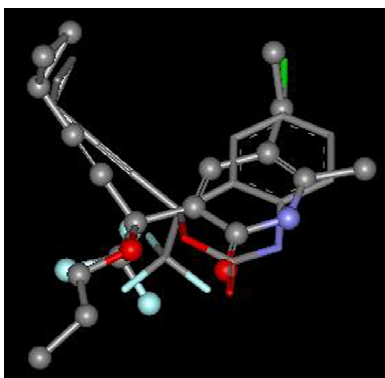


Figura 4.17. Modo de unión predicho para **5** (esferas y barras) en el sitio de unión de efavirenz. Se muestra la posición del efavirenz (barras) para comparación. Para mayor claridad, se omiten los hidrógenos.

Tabla 4.5. Acoplamiento de los híbridos de piridinona en diferentes geometrías de unión de la TR

Sitio de unión	Compuesto	Grupo ^a	Miembros en el grupo	Energía más baja (kcal/mol) ^b	Energía promedio (kcal/mol) ^c	RMDS (Å) ^d
Nevirapina	5	1	100	-9.67	-9.53	0.40
	6	1	3	-10.62	-10.60	3.10
	6	2	92	-10.51	-10.49	0.15
1051U91	5	1	96	-9.42	-9.22	0.70
	6	1	96	-10.42	-10.39	0.59
9-CI-TIBO	5	1	100	-10.15	-9.96	0.62
	6	1	5	-10.37	-10.32	1.88
	6	2	73	-10.35	-10.27	1.04
CI- α -APA	5	1	91	-9.32	-9.12	0.55
	6	1	4	-10.07	-10.03	1.68
	6	2	94	-9.90	-9.88	0.35
Efavirenz	5	1	83	-10.42	-10.18	0.67
	6	1	15	-10.75	-10.54	2.91
	6	3	40	-10.43	-10.28	2.43
MKC-442	5	1	79	-9.56	-9.34	1.41
	6	1	97	-11.01	-10.99	0.72
TNK-651	5	1	98	-9.95	-9.61	0.96
	6	1	96	-11.00	-10.98	0.66
UC-781	5	1	56	-9.84	-9.57	1.20
	5	2	38	-9.67	-9.41	1.24
	6	1	96	-10.65	-10.61	1.02
S-1153	5	1	37	-8.99	-8.76	1.28
	6	1	99	-10.07	-10.05	0.56

^{a-d} Ver pie de Tabla 4.3.

De esta forma, la actividad que tiene **5** frente a cepas con la mutación Lys103Asn (Corbett, 2001) puede deberse a los mismos factores que explican la actividad de efavirenz frente a cepas que presentan esta mutación. A partir de estructuras cristalográficas de TR en complejo con efavirenz, se ha propuesto que este inhibidor presenta interacciones con la cadena lateral de asparagina (Lindberg, 2002). Se ha sugerido también que cuando ocurre la mutación Lys103Asn, ocurren acomodos estructurales tanto en el sitio de unión como en la molécula de efavirenz que le permiten adaptarse en el sitio mutado (Ren, 2000a).

Un factor estructural que puede contribuir a la actividad de **5** frente a cepas resistentes es la presencia del grupo ciclopropil-propinil. A partir de la estructura cristalográfica de efavirenz (Ren, 2000a), se ha planteado que el grupo ciclopropil-propinil, al ser más pequeño que un anillo aromático, tiene interacciones más limitadas con Tyr181 y Tyr188, que son residuos que

mutan fácilmente. De esta manera, cuando las tirosinas mutan a residuos no aromáticos como cisteína, esto tiene un menor impacto en la afinidad de efavirenz (Ren, 2000a).

Comparando el modo de unión predicho de los análogos de piridinona en las geometrías de unión de nevirapina, 1051U91, 9-Cl-TIBO, Cl- α -APA, efavirenz, UC-781 y S-1153, con el modo de unión de MKC-442 y TNK-651, se puede observar que el anillo de piridinona ocupa aproximadamente la misma zona de unión que el anillo de uracilo de los derivados HEPT (Figura 4.18). El grupo aminometileno o etileno de las piridinonas se localiza en la misma región que el grupo isopropilo de MKC-442 y TNK-651. Así mismo, el anillo de benzoxazol de los análogos de piridinona ocupa una región similar que el grupo bencilo de los derivados HEPT. Estas relaciones estructurales ayudan a entender, a nivel molecular, diversos estudios SAR obtenidos para híbridos piridinona-HEPT (Benjahad, 2004; Dollé, 1995; Dollé, 2000). El modelo de acoplamiento está totalmente de acuerdo con la estrategia de sustituir la posición 4 del anillo de la piridinona con un grupo bencilo o feniltio (Figura 4.18).

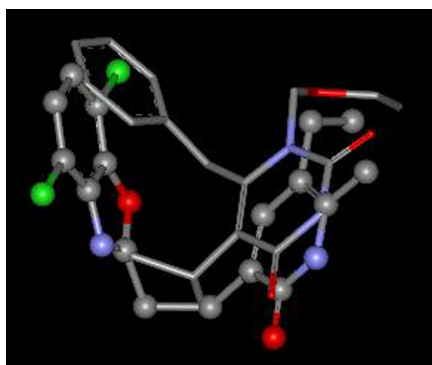


Figura 4.18. Modo de unión predicho para **2** (esferas y barras) en el sitio de unión de nevirapina. Se muestra la posición de MKC-442 (barras) para comparación. Para mayor claridad, se omiten los hidrógenos.

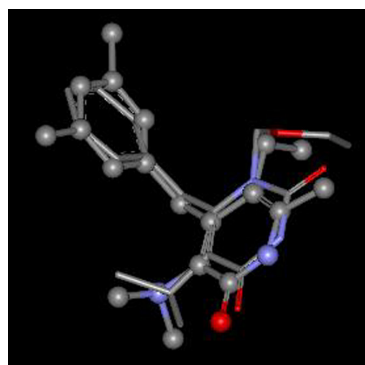


Figura 4.19. Modo de unión predicho para **6** (esferas y barras) en el sitio de unión de MKC-442. Se muestra la posición de MKC-442 (barras) para comparación. Para mayor claridad, se omiten los hidrógenos.

Experimentos de acoplamiento del híbrido piridinona-MKC-442 (**6**) (Figura 4.15) en diferentes geometrías de unión, muestra que la molécula híbrida adopta la conformación esperada de mariposa (Figura 4.19). Los resultados del acoplamiento se encuentran en la Tabla 4.5.

La actividad que presentan los híbridos de piridinona-MKC-442 contra cepas resistentes puede deberse, al menos en parte, a la presencia de la disustitución por dos grupos metilo en posiciones 3 y 5 del anillo de bencilo de moléculas como **6**. Estos dos sustituyentes pueden presentar múltiples interacciones con la TR, especialmente con Trp229, que es un residuo que no muta (Hopkins, 1999; Ren, 2000b).

4.3. Diseño de análogos de piridinona con potencial actividad contra cepas resistentes de VIH-1

Uno de los principales objetivos del diseño actual de INNTR es desarrollar compuestos que puedan inhibir no sólo a cepas nativas de la TR, sino también cepas que contengan mutaciones. Este es el caso de los INNTR de segunda generación (*vide supra*).

Uno de los inhibidores de segunda generación más atractivos a la fecha es el compuesto UC-781 (Figura 4.7) (De Clercq, 2004) mostrando cierta actividad contra la mutación Tyr181Cys (Ren, 1998). En la estructura cristalográfica de este compuesto se observa que el grupo (3-metilbut-2-enil)oxi interacciona con el residuo Tyr181 (Ren, 1998). De esta manera, a semejanza de efavirenz, la conformación de mariposa que adopta UC-781 en el sitio de unión no está formada por dos sistemas aromáticos como el caso, por ejemplo, de nevirapina, Cl- α -APA y las piridinonas **1-4**.

A partir de la estructura cristalográfica de la TR formando complejo con UC-781, se ha propuesto que el grupo (3-metilbut-2-enil)oxi, al ser más pequeño que un anillo aromático, tiene interacciones más limitadas con Tyr181. De esta manera, cuando las tirosina muta a cisteína, esto tiene un menor impacto en la afinidad de UC-781 (Ren, 1998). Una propuesta análoga se ha hecho con el grupo ciclopropil-propinil de efavirenz (*vide supra*).

Con el propósito de diseñar análogos de piridinona activos contra cepas resistentes, se comparó el modo de unión predicho por AutoDock para **1-4** con la posición de unión cristalográfica de UC-781. La comparación está mostrada en la Figura 4.20.

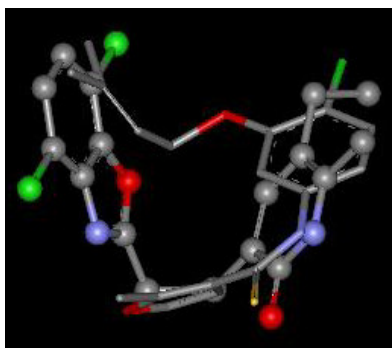
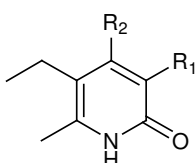


Figura 4.20. Modo de unión predicho para **2** (esferas y barras) en el sitio de unión de UC-781. Se muestra la conformación con la mejor evaluación de energía del grupo más poblado. Se da la posición de UC-781 (barras) para comparación. Para mayor claridad, se omiten los hidrógenos.

En la Figura 4.20 se observa que el anillo de benzoxazol de las piridinonas ocupa una región de unión similar que el grupo (3-metilbut-2-enil)oxi de UC-781. Como se discutió anteriormente en el modelo de unión de las piridinonas, el anillo de benzoxazol interacciona

con Tyr181 (Figuras 4.12, 4.13) y puede ser el responsable de la resistencia a estos compuestos. Por tanto, puede esperarse que en las piridinonas, al sustituir el anillo de benzoxazol por un sistema no aromático, por ejemplo el grupo (3-metilbut-2-enil)oxi, las interacciones con la Tyr181 sean limitadas como en el caso de UC-781. Este tipo de sustitución puede dar origen a análogos de piridinona activos contra la TR que contiene la mutación Tyr181Cys. El grupo (3-metilbut-2-enil)oxi, o un grupo análogo no aromático, puede estar en la posición 4 del anillo de piridinona (Figura 4.20). Ejemplos de nuevos derivados de piridinona que se proponen siguiendo esta estrategia de diseño, están mostrados en la Figura 4.21.



- 7, $R_1 = \text{NH}_2$; $R_2 = \text{OCH}_2\text{CH}=\text{C}(\text{CH}_3)_2$
 8, $R_1 = \text{N}(\text{CH}_3)_2$; $R_2 = \text{OCH}_2\text{CH}=\text{C}(\text{CH}_3)_2$
 9, $R_1 = \text{COOCH}_2\text{CH}_3$; $R_2 = \text{OCH}_2\text{CH}=\text{C}(\text{CH}_3)_2$
 10, $R_1 = \text{N}(\text{CH}_3)_2$; $R_2 = \text{OCOCH}_2\text{CH}(\text{CH}_3)_2$

Figura 4.21. Ejemplos de nuevos derivados de piridinona con potencial actividad inhibitoria de la TR, especialmente contra cepas con la mutación Tyr181Cys.

Se realizó el acoplamiento de las moléculas **7-10** en la geometría de unión de nevirapina. Los resultados están en la Tabla 4.6. La posición y conformación de unión predicha para estos compuestos es la esperada, es decir, una conformación de mariposa en donde el grupo (3-metilbut-2-enil)oxi ocupa una región similar al benzoxazol de las piridinonas en el sitio de unión. De esta forma, este sustituyente haría contactos con la Tyr181. El modo de unión predicho para estos compuestos está representado en la Figura 4.22.

Comparando las energías de acoplamiento calculadas para **7**, **8** y **9** puede deducirse que es conveniente la presencia de un grupo lo suficientemente voluminoso en la posición 3 (R_1) del anillo de piridinona, como el grupo dimetilamino (**8**) o etoxicarbonilo (**9**) (Figura 4.21). Este grupo puede ejercer el efecto estérico del puente aminometileno o etileno de los análogos **1-4**. Las sustituciones propuestas para la posición 3 del anillo de piridinona están de acuerdo con los estudios SAR de los híbridos de piridinona-MKC-442. Estos estudios de relación-estructura actividad muestran que la actividad biológica se incrementa cuando en la posición 3 del anillo de piridinona está un grupo dimetilamino (Dollé, 2000) o un grupo etoxicarbonilo (Dollé, 1995).

Tabla 4.6. Acoplamiento de nuevos derivados de piridinona en el sitio de unión de la nevirapina

Compuesto	Miembros en el grupo ^a	Energía más baja (kcal/mol) ^b	Energía promedio (kcal/mol) ^c
7	39	-8.80	-8.75
8	76	-9.33	-9.19
9	45	-10.73	-10.53
10	95	-9.68	-9.60

^{a-c} Ver pie de Tabla 4.2.

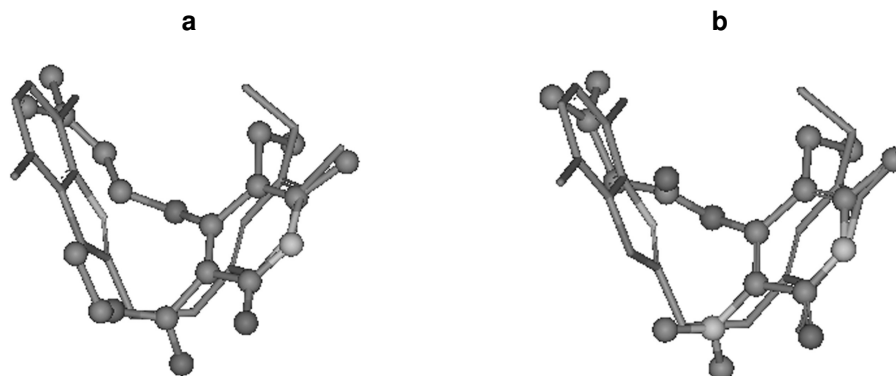


Figura 4.22. Modo de unión predicho para (a) **9** y (b) **10** (esferas y barras) en el sitio de unión de nevirapina. Se muestra la conformación predicha para **2** (barras) para comparación. Para mayor claridad, se omiten los hidrógenos.

Las energías de acoplamiento calculadas para la serie de los nuevos híbridos de piridinona de la Figura 4.21 sugieren que el compuesto **9** tendría la mayor actividad contra la cepa nativa. De acuerdo a la discusión anterior, se espera que este compuesto también muestre actividad contra cepas con la mutación Tyr181Cys. La mayor energía de acoplamiento calculada para **7** sugiere que este compuesto mostraría la menor actividad.

Conclusiones

Se empleó acoplamiento molecular automatizado para estudiar el modo de unión de los análogos de piridinona en el sitio de unión de los INNTR. Se derivó un solo modelo de acoplamiento considerando las geometrías de unión de nevirapina, 1051U91, 9-Cl-TIBO, Cl- α -APA, efavirenz, UC-781 y S-1153. De esta manera, la estructura cristalográfica del complejo con nevirapina, la estructura con la mejor resolución disponible a la fecha, puede seleccionarse como modelo de la conformación de unión de los análogos de piridinona, mientras no esté disponible una estructura cristalográfica para estos derivados unidos a la TR.

De acuerdo al modelo de unión propuesto, que está de acuerdo con diversos estudios SAR para estos compuestos, los análogos de piridinona adoptan un modo de unión similar a la forma de unión de los ligandos cristalográficos. Es característica la conformación tipo mariposa que tienen en común.

El modelo de acoplamiento ayudó a entender, a nivel molecular, el diseño y la actividad biológica reportada para moléculas híbridas de piridinona.

AutoDock predijo un modelo de unión diferente para los análogos de piridinona dentro de las geometrías de unión de MKC-442 y TNK-651. Se propone que la conformación del giro de la Pro236 en estas geometrías, que es una consecuencia particular de la estructura de los derivados HEPT, es la responsable de modificar las predicciones del acoplamiento con AutoDock.

Se diseñaron nuevos análogos de piridinona con potencial actividad inhibitoria de la TR, especialmente activas contra cepas resistentes. Los resultados del acoplamiento de las moléculas diseñadas indican que pueden adoptar el modo de unión esperado. Las evaluaciones de la energía de acoplamiento permiten establecer un orden de las moléculas en que se esperaba fueran más activas contra la cepa nativa.

La estrategia del diseño de los nuevos análogos de piridinona consistió en comparar el modo de unión predicho para las piridinonas con las posiciones cristalográficas de un inhibidor de segunda generación. Esta estrategia, que es una forma de hacer diseño de fármacos asistido por computadora basado en el receptor, puede aplicarse para diseñar otros análogos de piridinona además de los ejemplificados en este trabajo. Más aún, la estrategia puede aplicarse a otras clases estructurales de INNTR para proponer, considerando la estructura de la TR, compuestos activos contra cepas resistentes. La perspectiva de este trabajo es confirmar experimentalmente las predicciones realizadas para los análogos de piridinona diseñados.

Capítulo 5

Estudios QSAR-3D

La gran cantidad de estudios SAR disponibles para los derivados de piridinona (Hoffman, 1992; Hoffman, 1993; Saari, 1991; Saari, 1992) permite realizar estudios QSAR con estas moléculas. Para otros INNTR se han realizado varios estudios QSAR utilizando diferentes métodos. Ejemplos de trabajos recientes pueden encontrarse en Douali, 2003 y Pungpo, 2003.

La estrategia general del estudio computacional de inhibidores de la TR de VIH-1 está descrita en el capítulo anterior (cf., Figura 4.6). En este capítulo se presentan los resultados de los estudios QSAR-3D con derivados de piridinona. Se han reportado estudios QSAR-3D que han contribuido a identificar el tipo de interacciones que ocurren entre diversas clases estructurales de INNTR y la TR (Barreca, 1999; Chen, 2003; Hannongbua, 1999; Hannongbua, 2001; Pungpo, 2000).

Se realizaron dos estudios QSAR-3D, empleando los métodos de CoMFA y CoMSIA. Los estudios se efectuaron con 40 análogos de piridinona reportados en la literatura (Hoffman 1993, Saari, 1992). Esta base de datos se dividió en dos grupos, uno de entrenamiento con 29 compuestos y otro de prueba con 11 compuestos. Como criterio de selección se consideró la variabilidad estructural y la diversidad de la potencia biológica de cada compuesto. Las pruebas biológicas de inhibición de la TR fueron determinadas *in vitro* bajo las mismas condiciones experimentales (Hoffman 1993, Saari, 1992). Los valores de las actividades se expresaron en los correspondientes valores de $-\log IC_{50}$, donde IC_{50} es la concentración efectiva del compuesto para inhibir el 50% de la actividad de la TR. Las actividades biológicas están reportadas en las Tablas 5.1 y 5.2.

Una etapa crucial en un análisis QSAR-3D es la alineación de las moléculas (Cramer, 1988). El objetivo de la alineación es generar una superposición espacial de los ligandos que reproduzca modos de unión experimentales (Klebe, 1998). Se han propuesto diferentes métodos para la alineación (Greco, 1997). Sin embargo, hay que considerar que la conformación propuesta para las moléculas, que con frecuencia se obtiene de una optimización de la geometría, puede no corresponder a la conformación bioactiva (Greco, 1997; Oprea, 1997).

Tabla 5.1. Estructuras, actividad inhibitoria de la TR de VIH-1 y energías libres de unión (en kcal/mol) calculadas por AutoDock para los compuestos del grupo de entrenamiento

Comp.	L	Ar	-LogIC ₅₀	ΔG	Comp.	L	Ar	-LogIC ₅₀	ΔG
1	NH		6.55	-9.34	16	NH		5.90	-9.02
2	NH		6.46	-9.25	17	NH		6.59	-9.65
3	NH		6.34	-9.53	18	NH		7.7	-10.01
4	NH		6.30	-9.24	19	NH		7.72	-9.83
5	NH		6.28	-9.56	20	NH		6.74	-9.31
6	NH		5.96	-9.29	21	NH		6.36	-9.09
7	NH		5.72	-8.70	22	NH		7.24	-9.75
8	NH		5.60	-9.51	23	CH ₂		4.30	-7.74
9	NH		5.57	-10.27	24	CH ₂		5.61	-9.94
10	NH		5.36	-9.10	25	CH ₂		6.43	-9.68
11	NH		5.27	-7.59	26	CH ₂		7.24	-9.72
12	NH		4.82	-7.55	27	CH ₂		6.41	-9.32
13	NH		4.65	-8.55	28	CH ₂		7.55	-10.20
14	NH		3.98	-8.85	29	CH ₂		7.85	-10.18
15	NH		6.92	-9.55					

Tabla 5.2. Estructuras, actividad inhibitoria de la TR de VIH-1 y energías libres de unión (en kcal/mol) calculadas por AutoDock para los compuestos del grupo de prueba

Comp.	L	Ar	-LogIC ₅₀	ΔG	Comp.	L	Ar	-LogIC ₅₀	ΔG
30	NH		6.68	-8.83	36	NH		6.47	-9.40
31	NH		5.63	-9.05	37	NH		5.71	-9.47
32	NH		5.12	-8.97	38	CH ₂		5.08	-10.00
33	NH		4.49	-8.88	39	CH ₂		3.59	-7.90
34	NH		5.78	-9.06	40	CH ₂		4.87	-9.22
35	NH		7.26	-9.50					

Sin embargo, cuando se conoce la estructura tridimensional del receptor, como es el caso de la TR (Balzarini, 2004), se puede hacer un alineamiento considerando a esta estructura, por ejemplo, utilizando acoplamiento molecular automatizado. Esta estrategia se ha aplicado exitosamente en estudios QSAR-3D en donde el sitio de unión de los ligandos es conocido (Lozano, 2000; Sippl, 2001; Sippl, 2002).

Alineación de los análogos de piridinona

La alineación de los 40 análogos de piridinona se hizo a partir de la posición y conformación que predice AutoDock para cada compuesto dentro del sitio de unión de los INNTR. La estructura cristalográfica de la TR en complejo con la nevirapina se utilizó como modelo de la geometría del sitio de unión de las piridinonas. En el capítulo anterior se demostró que esta geometría es un modelo apropiado para estudiar el acoplamiento de los derivados de piridinona (cf., Capítulo 4).

AutoDock encontró dos formas de unión posibles para algunos análogos de piridinona. Se seleccionaron los complejos análogo de piridinona-TR de acuerdo a los grupos más poblados y a características geométricas de los complejos. En muchos casos, la conformación

seleccionada correspondió a la conformación y posición de menor energía. De acuerdo a estos criterios, todos los compuestos estudiados adoptan una conformación de mariposa y tienen la misma orientación que la nevirapina. Los dos anillos aromáticos de los análogos de piridinona (que pueden compararse con las “alas” de la conformación tipo mariposa), ocupan aproximadamente el mismo sitio de unión de los anillos aromáticos de la nevirapina. Los residuos que interactúan con la mayoría de los análogos de piridinona estudiados son Pro95, Leu100, Lys101, Lys103, Tyr181, Tyr188 y Tyr318. Estos resultados son análogos a los descritos en el capítulo anterior (cf., Figuras 4.12 y 4.13).

La energía libre de unión calculada por AutoDock está dada en las Tablas 5.1 y 5.2. En general, la energía mostró una buena correlación con la actividad biológica. De esta manera, compuestos con la mayor actividad biológica mostraron las energías más bajas. Esto sugiere que los modelos de unión para los análogos de piridinona con la TR son razonables. La correlación cuantitativa entre la energía libre de unión y la actividad biológica fue mejor para compuestos estructuralmente similares, por ejemplo, aquellos compuestos con un anillo de benzoxazol (Figura 5.1). Todos estos compuestos adoptan un modo de unión muy parecido (cf., Figura 5.3).

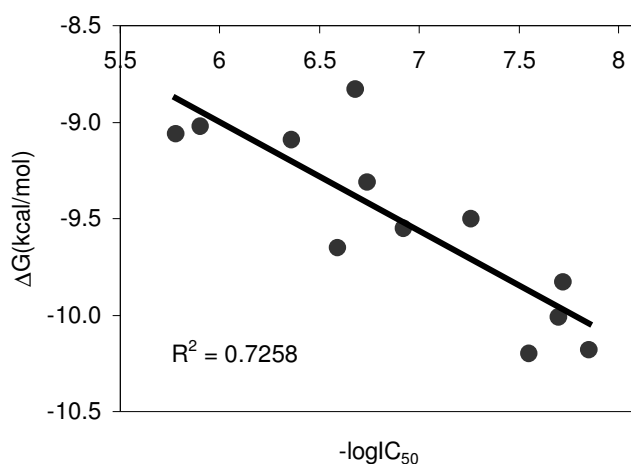


Figura 5.1. Actividad biológica experimental *versus* energías libres de unión calculadas para compuestos estructuralmente similares (15-21, 28, 29, 30, 34, 35).

Después de los cálculos de acoplamiento, se optimizó la geometría de los complejos con mecánica molecular (ver sección de Metodología computacional para detalles). Se observaron muy pocas variaciones entre las conformaciones predichas por AutoDock y las conformaciones de las piridinonas después de la optimización. En la Figura 5.2 se muestra la

superposición de las conformaciones de unión calculadas para los análogos de piridinona después de la optimización geométrica de los complejos con la TR. Esta alineación fue empleada en los estudios QSAR-3D.

En la superposición de las piridinonas mostrada en la Figura 5.2 se observa que todas las moléculas tienen un modo de unión común en donde los anillos de piridinona (“ala 1” de la conformación tipo mariposa) ocupan aproximadamente la misma zona de unión en la TR. Todas las piridinonas que se acoplaron pueden formar un puente de hidrógeno entre el protón polar del anillo de piridinona y el oxígeno del grupo carbonilo de la Lys101 (cf., Figura 5.3). Sin embargo, el “ala 2” de la conformación tipo mariposa (el sustituyente “Ar” en las Tablas 5.1 y 5.2) puede adoptar una orientación diferente dependiendo de la estructura química del sistema aromático (Figura 5.3).

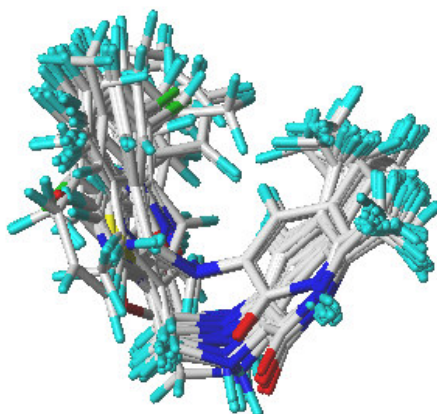


Figura 5.2. Alineación de los análogos de piridinona.

Resulta interesante que los compuestos más activos contienen un anillo de benzoxazol. El benzoxazol, que a su vez está constituido por otros dos anillos, está próximo a la Tyr181 (Figura 5.3). Las posiciones paralelas del anillo de benzoxazol y el anillo aromático de Tyr181 sugieren que pueden estar formando interacciones tipo empaquetamiento- π (McGaughey, 1998; Meyer, 2003) y de esta forma contribuir en la estabilización del complejo. Interacciones similares pueden ocurrir para otros análogos de piridinona con un sistema de dos anillos aromáticos en la “ala 2”.

Los compuestos **11**, **12**, **23** y **39**, que contienen un solo anillo en la “ala 2”, adoptan una conformación de acoplamiento que es ligeramente diferente a la de los otros análogos de piridinona. En **11**, **12**, **23** y **39** el grupo aminometileno o etileno que unen a las dos “alas”, está desplazado hacia arriba respecto al aminometileno o etileno de los otros análogos de piridinona (Figura 5.3). A pesar de que la “ala 2” está cercana a la Tyr181, las posiciones no

paralelas del grupo fenilo de Tyr181 y la “ala 2” sugieren interacciones más débiles entre estos sistemas (Figura 5.3) (McGaughey, 1998; Meyer, 2003). Esto ayuda a explicar, al menos en parte, la baja actividad biológica para estos compuestos (Tablas 5.1 y 5.2). En forma similar, la baja actividad para **13** puede ser explicada por la falta de contactos con Tyr181.

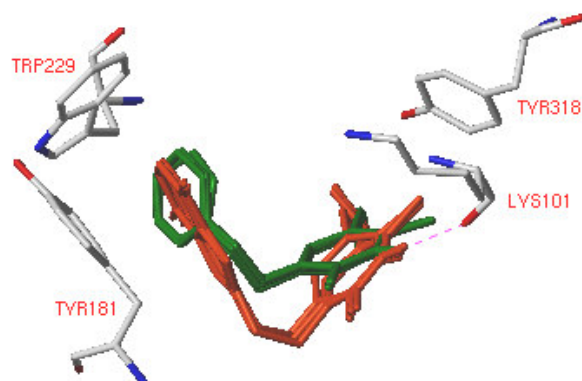


Figura 5.3. Comparación del modo de unión de los compuestos **18**, **19**, **28** y **29** (naranja) con el modo de unión de **11**, **12**, **23** y **39** (verde). La posición de los residuos Tyr181, Trp229, Tyr318 y Lys101 se muestra como referencia. El puente de hidrógeno se muestra con una línea discontinua magenta. Para mayor claridad, se omiten los hidrógenos excepto el hidrógeno polar del anillo de la piridinona.

5.1. Modelos QSAR-3D

En la Figura 5.4 se esquematiza la estrategia general con la cual se obtuvieron los modelos de CoMFA y CoMSIA. Se hace notar la combinación de técnicas computacionales de diseño basado en la estructura (acoplamiento molecular automático) con técnicas de diseño basado en el ligando (QSAR). Esta combinación de técnicas es muy atractiva para aplicar a los análogos de piridinona para los cuales se conoce una gran cantidad de estudios SAR y además se conoce la estructura tridimensional de su sitio de unión.

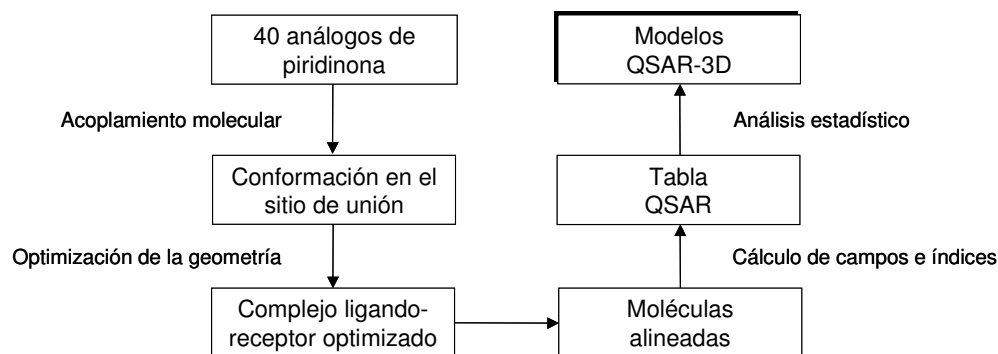


Figura 5.4. Estrategia del estudio QSAR-3D de los análogos de piridinona.

Los resultados del análisis PLS de los estudios CoMFA y CoMFA están resumidos en la Tabla 5.3. En la Figura 5.5 se muestran las predicciones de los modelos QSAR-3D. En la Tabla 5.4 se presentan los residuales para las predicciones de las actividades del grupo de prueba.

Tabla 5.3. Estadística de los modelos CoMFA y CoMSIA

Cargas	Espacio en la rejilla	Estadística PLS	Modelos con 29 compuestos		Modelos con 28 compuestos	
			CoMFA	CoMSIA	CoMFA	CoMSIA
AM1	2 Å	q^2	0.656	0.643	0.706	0.722
		PRESS	0.639	0.651	0.603	0.586
		Noc	4	4	4	4
		r^2	0.943	0.926	0.954	0.943
		S	0.259	0.295	0.239	0.266
		F	100.055	75.566	119.076	94.725
		r^2_{pred}	0.688	0.785	0.720	0.759
		Contribución de los campos				
		Estérico	0.608	0.121	0.600	0.122
		Electrostático	0.392	0.421	0.400	0.434
	Hidrófobo		0.457		0.445	
	1 Å	q^2	0.613	0.643	0.650	0.723
		PRESS	0.664	0.651	0.644	0.585
		Noc	3	4	3	4
		r^2	0.934	0.932	0.947	0.946
		S	0.275	0.284	0.251	0.259
		F	117.468	82.484	142.534	100.113
		r^2_{pred}	0.693	0.778	0.731	0.750
		Contribución de los campos				
		Estérico	0.563	0.121	0.559	0.121
Electrostático		0.437	0.426	0.441	0.434	
Hidrófobo		0.456		0.445		
Gasteiger-Hückel	2 Å	q^2	0.558	0.569	0.627	0.621
		PRESS	0.710	0.715	0.694	0.685
		Noc	3	4	5	4
		r^2	0.920	0.934	0.955	0.941
		S	0.303	0.280	0.241	0.269
		F	95.341	84.812	93.199	92.289
		r^2_{pred}	0.771	0.729	0.715	0.780
		Contribución de los campos				
		Estérico	0.582	0.099	0.590	0.102
		Electrostático	0.418	0.498	0.410	0.505
Hidrófobo		0.403		0.393		

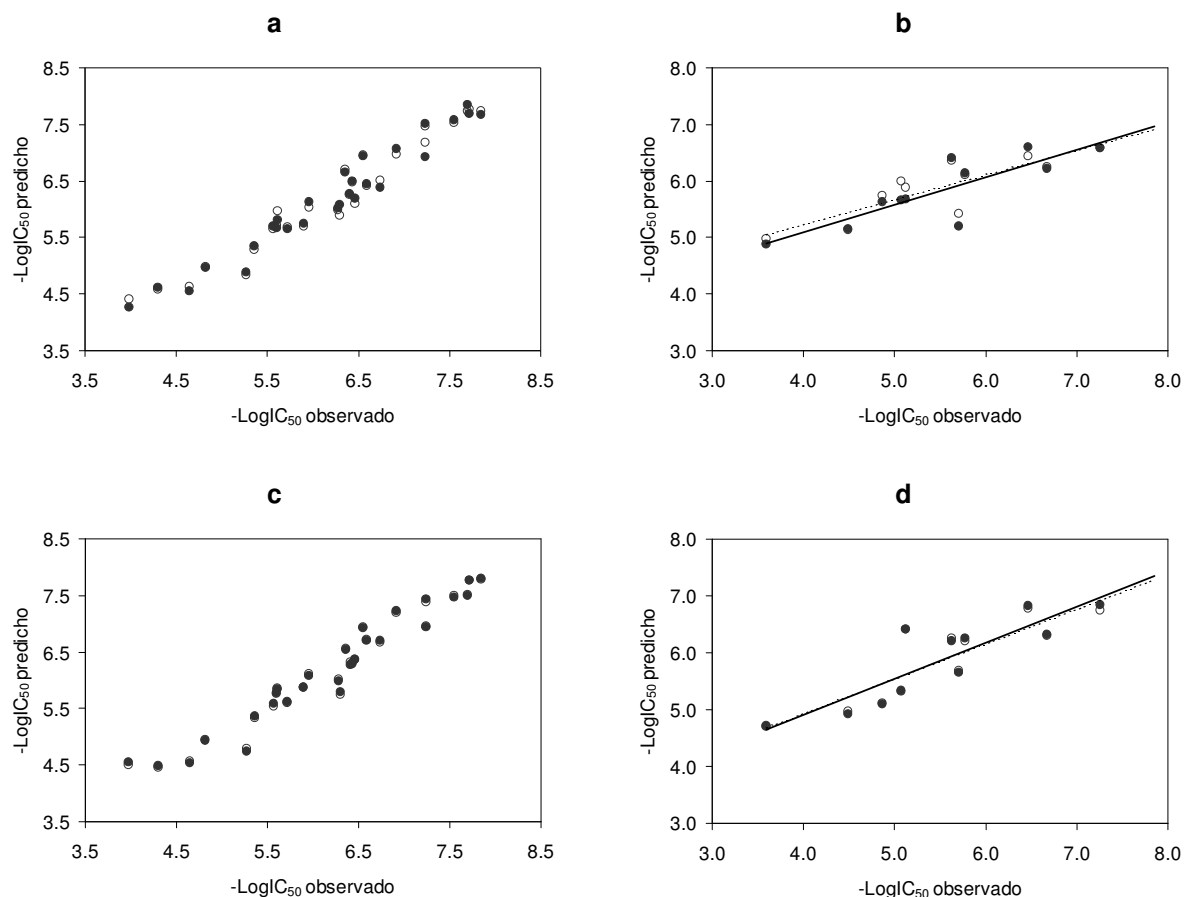


Figura 5.5. Predicciones de los modelos con 28 compuestos para los grupos de entrenamiento y prueba. Las predicciones de los modelos a 2 Å están representadas con círculos llenos y líneas continuas mientras que los modelos a 1 Å están representados con círculos vacíos y líneas punteadas. Las predicciones con los modelos de CoMFA para los grupos de entrenamiento y prueba están mostrados en a y b, respectivamente. Las correspondientes predicciones con los modelos de CoMSIA están mostradas en c y d, respectivamente.

Tabla 5.4. Residuales de las predicciones para el grupo de prueba con los modelos de CoMFA y CoMSIA^a

Compuesto	$-\text{LogIC}_{50}$	Residuales			
		CoMFA 2 Å	CoMFA 1 Å	CoMSIA 2 Å	CoMSIA 1 Å
30	6.68	0.462	0.429	0.360	0.376
32	5.63	-0.772	-0.736	-0.583	-0.628
33	5.12	-0.563	-0.758	-1.300	-1.300
34	4.49	-0.667	-0.646	-0.429	-0.474
35	5.78	-0.362	-0.318	-0.467	-0.424
36	7.26	0.671	0.674	0.424	0.511
37	6.47	-0.136	0.036	-0.348	-0.302
38	5.71	0.514	0.287	0.061	0.035
39	5.08	-0.576	-0.910	-0.258	-0.233
40	3.59	-1.283	-1.390	-1.128	-1.101
41	4.87	-0.752	-0.867	-0.228	-0.235

^a Modelos derivados con 28 compuestos.

Para evaluar qué tan robustos son los modelos CoMFA y CoMSIA desarrollados, se asignaron valores aleatorios a las actividades biológicas de los compuestos del grupo de entrenamiento. Se calcularon los valores de q^2 para cada grupo de datos con las actividades aleatorias y se compararon con los valores de q^2 obtenidos con las actividades biológicas experimentales. En estudios QSAR, esta estrategia ha mostrado probar la estabilidad de los modelos generados (Kuo, 2004; Tropsha, 2003b).

El promedio de los valores de q^2 , después de asignar valores aleatorios a las actividades biológicas en diez ocasiones independientes, están presentados en la Tabla 5.5. En todos los casos, la predictibilidad de los modelos con las actividades biológicas aleatorias fue considerablemente inferior a los modelos derivados con las actividades biológicas reales. Esto indica que los modelos CoMFA y CoMSIA desarrollados son robustos.

Tabla 5.5. Promedio de los valores de q^2 después de asignar valores aleatorios a la actividad biológica

Modelo	29 Compuestos		28 Compuestos	
	CoMFA	CoMSIA	CoMFA	CoMSIA
2 Å	-0.144	-0.288	-0.157	-0.295
1 Å	-0.211	-0.286	-0.212	-0.293

5.2. Análisis comparativo de campos moleculares (CoMFA)

El análisis PLS, usando las cargas de AM1 y los parámetros de defecto, dio un modelo de CoMFA con un valor de q^2 igual a 0.656 y un valor de r^2_{pred} igual a 0.688. En este análisis, el compuesto **3** estuvo fuera de la correlación. Al omitir al compuesto **3** del análisis, se incrementó el valor de q^2 a 0.706 y el valor de r^2_{pred} a 0.720 para el resto de los 28 compuestos del grupo de entrenamiento (Tabla 5.3).

Se investigó el efecto del espacio de los puntos de la rejilla sobre la predictibilidad de los modelos. Con la reducción del espacio de 2 a 1 Å, disminuyó el valor de q^2 de 0.656 a 0.613 para el modelo con los 29 compuestos. Para el modelo con 28 compuestos, después de omitir a **3**, el valor de q^2 disminuyó de 0.706 a 0.650.

La capacidad predictiva de los modelos con 28 compuestos se muestra en la Figura 5.5. El mejor modelo CoMFA que se derivó fue a 2 Å con 28 compuestos. Este modelo mostró una alta capacidad predictiva para el grupo de entrenamiento ($q^2 = 0.706$) y el grupo de prueba ($r^2_{pred} = 0.720$). Sin embargo, el modelo a 1 Å con 28 compuestos mostró la predictibilidad más alta para el grupo de prueba con un valor de r^2_{pred} de 0.731. Esto, a pesar del menor valor

de q^2 que mostró respecto al modelo correspondiente a 2 Å. En forma similar, el modelo a 1 Å con todos los compuestos mostró un valor de r^2_{pred} superior que el modelo correspondiente a 2 Å a pesar de su valor de q^2 más bajo. Esto significa que un mayor valor de q^2 no significa que el modelo sea necesariamente más predictivo respecto a un grupo externo de prueba (Golbraikh, 2002a). Esta misma observación se ha hecho en otros estudios QSAR-3D (por ejemplo, Buolamwini, 2002).

En los modelos a 2 Å con 29 y 28 compuestos, la contribución de las interacciones estéricas fue del 60.8 y 60.0 %, respectivamente. En los modelos a 1 Å, la contribución estérica fue cercana al 56 %. Esto coincide con estudios CoMFA reportados para otros INNTR en donde la contribución de las interacciones estéricas también predomina sobre las interacciones de tipo electrostático (Hannongbua, 1999; Hannongbua, 2001; Pungpo, 2000). Más aún, en algunos de estos estudios CoMFA, la contribución de las interacciones tipo estéricas son también alrededor del 60 % (Hannongbua, 2001; Pungpo, 2000).

En todos los casos estudiados, los modelos generados con las cargas de AM1 mostraron mayores valores de q^2 que los modelos obtenidos con las cargas de Gasteiger-Hückel. Sin embargo, los modelos obtenidos con las cargas de Gasteiger-Hückel mostraron una predictibilidad satisfactoria para el grupo de prueba (Tabla 5.3).

Respecto a la variación del átomo de prueba, todos los modelos derivados con un carbono sp^3 fueron mejores que los modelos obtenidos con un átomo de oxígeno sp^3 o un átomo de hidrógeno.

Los mapas de contorno estéricos y electrostáticos correspondientes al mejor modelo de CoMFA están mostrados en la Figura 5.6. Para ayudar en la visualización de los mapas, se muestra al compuesto **29** y a los residuos que se encuentran a 5 Å de distancia. Los mapas en color verde indican regiones en donde sustituyentes voluminosos incrementan la actividad biológica (Norinder, 2002).

Hay un mapa de contorno verde cercano a la “ala 2” de los análogos de piridinona y a la posición del amino ácido Tyr181. Esto, que sugiere una interacción favorable entre estos dos sistemas, está de acuerdo con la discusión anterior de que esta interacción puede estar estabilizando al complejo. Resultados experimentales muestran que cuando ocurre la mutación Tyr181Cys disminuye la capacidad inhibitoria de **29** y otros análogos de piridinona (Byrnes, 1993; Nunberg, 1991). De esta manera, esta mutación parece eliminar las interacciones entre el anillo aromático de la tirosina y el inhibidor, reduciendo así la afinidad de las piridinonas por la TR.

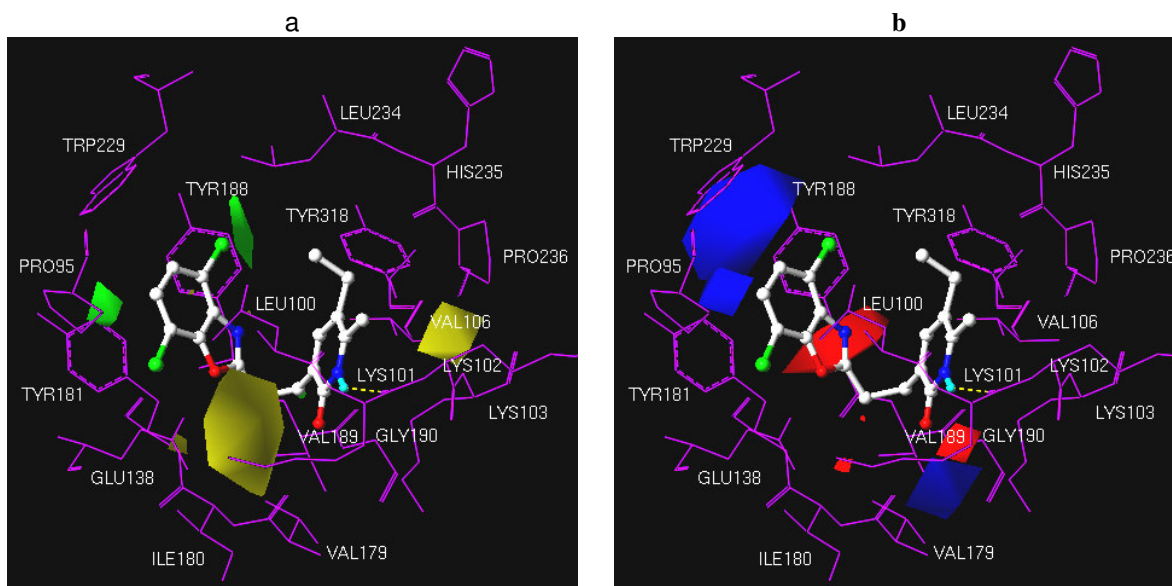


Figura 5.6. Mapas de contorno para el mejor modelo de CoMFA. Se muestra al compuesto **29** y a los residuos que se encuentran a 5 Å de distancia de este compuesto. (a) Mapas de contorno estéricos. Las regiones favorecidas están en verde y las desfavorecidas en amarillo. (b) Mapas de contorno electrostáticos. Las regiones favorecidas por cargas negativas están en rojo y las desfavorecidas en azul. El puente de hidrógeno se muestra con una línea discontinua amarilla. Para mayor claridad, se omiten los hidrógenos excepto el hidrógeno polar del anillo de la piridinona.

Otra interacción estérica favorable está cercana a la posición de la “ala 2” y a la posición de los residuos Tyr188 y Leu234 (Figura 5.6a). Este contorno está de acuerdo con las posiciones de acoplamiento de las moléculas **17**, **20**, **21**, **35**, y **36** en donde el sustituyente en el anillo de benzoxazol o benzofurano hace contactos con estos residuos. El sustituyente en la posición 4 del anillo de benzoxazol de los compuestos **18**, **19**, **28** y **29**, que es importante para la actividad de los análogos de piridinona (Saari, 1992), también hace contactos con Tyr188 y Leu234. Este mismo contorno color verde está muy próximo al anillo de benceno de los compuestos **27** y **31**. Estos compuestos, con una actividad biológica de media o baja, no hacen contactos con Tyr181 pero sí con Tyr188. Puede ser que las interacciones con Tyr188 no los hagan tan inactivos como se esperaba por carecer de interacciones con Tyr181.

Estas observaciones están de acuerdo con los resultados experimentales de que la mutación de la Tyr188 también es responsable del decremento de actividad de los análogos de piridinona (Byrnes, 1993; Sardana, 1992), aunque en menor grado que cuando ocurre la mutación de la Tyr181. De hecho, los residuos Tyr 181 y Tyr188 son muy importantes para la inhibición de la TR por los INNTR. Se ha propuesto que la pérdida de interacciones aromáticas con estos residuos es la responsable del decremento de la actividad biológica de los INNTR de primera generación cuando estos residuos mutan (Ren, 1998; Ren, 2000a).

En la Figura 5.6a hay un contorno color amarillo cercano a la “ala 2” y al grupo etileno (o aminometileno en las moléculas correspondientes) que une a los dos anillos aromáticos de los análogos de la piridinona. Esto indica una restricción estérica en esta zona. Las cadenas laterales de la Lys101 y Val179 están próximas a esta región.

En la Figura 5.6b están representados los mapas de contorno electrostáticos. Los mapas en color rojo muestran zonas en donde se predice que grupos electronegativos favorecen la actividad biológica. Por su parte, los mapas en color azul muestran las áreas en donde grupos electronegativos disminuyen la actividad (Norinder, 2002). El mapa de color rojo cercano a la “ala 2” de los análogos de piridinona indica qué grupos con densidad de carga en esta zona incrementan la actividad. Estos resultados están de acuerdo con la posición de la “ala 2” predicha por AutoDock para las moléculas **2, 4, 5, 10, 18, 19, 22, 24, 25, 28, 29 y 36**. En todos estos casos, un átomo con alta densidad de carga está dirigido hacia esta región. Este contorno explica la contribución electrostática a la alta actividad biológica de los compuestos **18, 28 y 29**.

Hay un contorno azul cercano a la posición de Trp229 (Figura 5.6b). La baja actividad biológica de los compuestos **14, 33 y 40** puede ser explicada por la presencia de un átomo de nitrógeno con carga negativa cerca de este contorno. En forma similar, un contorno azul próximo a Pro95 explica la baja actividad de **7** que tiene un átomo de nitrógeno con carga negativa cercano a esta región.

5.3. Análisis comparativo de índices de similitud molecular (CoMSIA)

Se obtuvieron ligeramente mejores datos estadísticos para los modelos de CoMSIA que para los modelos de CoMFA (Tabla 5.3). Cuando se emplearon las cargas de AM1, a los parámetros de defecto de Sybyl y se consideraron a todos los compuestos en el grupo de entrenamiento, se obtuvo un modelo con una q^2 de 0.643 y una r^2_{pred} de 0.785 (comparar con una q^2 de 0.656 y una r^2_{pred} de 0.688 para CoMFA). Sin embargo, cuando se omitió del análisis al mismo compuesto **3**, los valores de q^2 y r^2_{pred} se incrementaron a 0.722 y 0.759, respectivamente (comparar con los valores obtenidos para CoMFA de 0.706 y 0.720, para q^2 y r^2_{pred} , respectivamente).

Cuando se disminuyó el tamaño del espacio de la rejilla de 2 a 1 Å se obtuvieron modelos CoMSIA casi con los mismos valores de q^2 (Tabla 5.3). Sin embargo, el poder predictivo para estimar las actividades del grupo de prueba fue ligeramente menor (aunque mayor a la capacidad predictiva de los modelos de CoMFA correspondientes) (Tabla 5.3).

El modelo a 1 Å con 28 compuestos fue el mejor modelo de CoMSIA que se generó, con una q^2 de 0.723 y una r^2_{pred} de 0.750 (comparar con el mejor modelo de CoMFA con una q^2 de 0.706 y una r^2_{pred} de 0.720). Para este grupo de compuestos, el mejor desempeño de CoMSIA comparado con CoMFA puede ser atribuido a los campos hidrófobos presentes en los modelos de CoMSIA. Este tipo de descriptores no está disponible en la versión estándar de CoMFA.

En CoMSIA se obtuvieron contribuciones similares de los campos estéricos y electrostáticos a las contribuciones correspondientes obtenidas con CoMFA (Tabla 5.3).

En forma similar a CoMFA, los modelos obtenidos con cargas de AM1 también mostraron valores de q^2 superiores que aquellos modelos obtenidos con cargas de Gasteiger-Hückel. Los modelos derivados con cargas de Gasteiger-Hückel mostraron una predictibilidad aceptable para el grupo de prueba (Tabla 5.3).

Se investigó el efecto de los campos donadores y aceptores de hidrógenos de CoMSIA. Sin embargo, cuando se consideraron a estos campos, no se incrementó el poder predictivo de los modelos. Esto puede explicarse con los resultados de acoplamiento. Para los 29 compuestos del grupo de entrenamiento se predice que se puede formar un puente de hidrógeno entre el protón polar del anillo de la piridinona y el oxígeno de la cadena principal de la Lys101 (Figura 5.3). De esta manera, no puede esperarse una diferencia en la afinidad de estos compuestos debida a la formación de este puente de hidrógeno.

Los mapas de contorno correspondientes al mejor modelo de CoMSIA están mostrados en la Figura 5.7. Para ayudar en la visualización de los mapas, se muestra al compuesto **29** y a los residuos que se encuentran a 5 Å de distancia. El color e interpretación de los campos estéricos y electrostáticos son los mismos que los correspondientes mapas de contorno de CoMFA.

Los campos estéricos están colocados en regiones similares que los obtenidos en los modelos de CoMFA (Figura 5.7a). Sin embargo, los dos contornos de CoMSIA color verde, cercanos a las posiciones de los residuos Tyr181 y Tyr188, son más prominentes que los obtenidos en CoMFA. Esta observación remarca la importancia de las interacciones entre los análogos de piridinona con Tyr181 y Tyr188 discutidas anteriormente. Por su parte, los campos electrostáticos de CoMSIA están, en general, de acuerdo con la distribución de los correspondientes campos de CoMFA (Figura 5.7b).

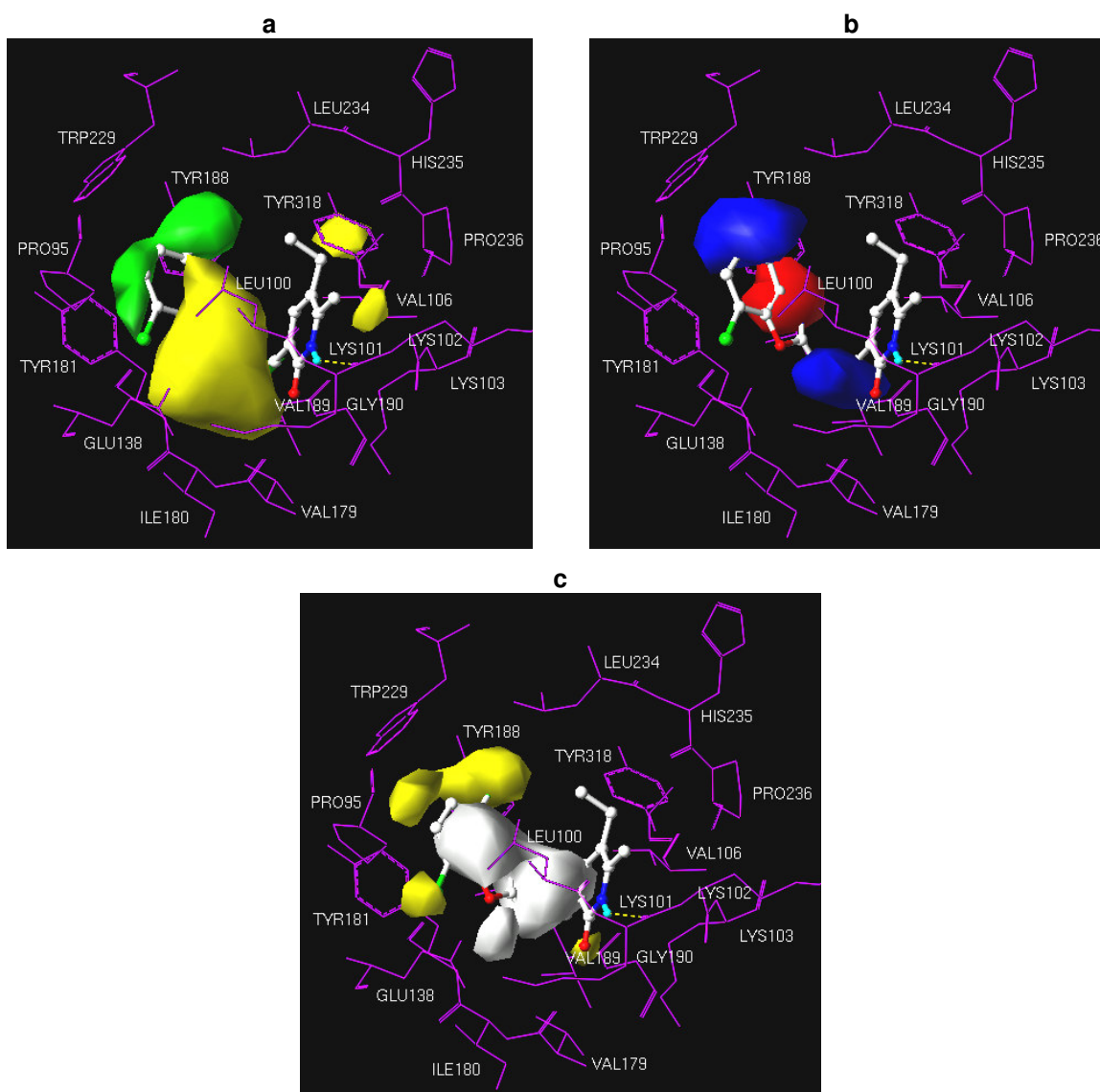


Figura 5.7. Mapas de contorno para el mejor modelo de CoMSIA. Se muestra al compuesto **29** y a los residuos que se encuentran a 5 Å de distancia de este compuesto. (a) Mapas de contorno estéricos. (b) Mapas de contorno electrostáticos. Los colores de los mapas tienen la misma interpretación que el de las Figuras 5.6a y 5.6b, respectivamente. (c) Mapas de contorno hidrófobos. Las regiones favorecidas están en amarillo y las desfavorecidas en blanco. El puente de hidrógeno está representado como en la Figura 5.6. Para mayor claridad, se omiten los hidrógenos excepto el hidrógeno polar del anillo de la piridinona.

Para los mapas de contorno hidrófobos de CoMSIA, mapas de color amarillo indican regiones en donde interacciones de tipo hidrófobas están asociadas con un incremento en la actividad biológica. Mapas de color blanco indican zonas donde un incremento en las interacciones hidrofílicas están relacionadas con un incremento en la actividad. Los contornos de color amarillo alrededor de los sustituyentes en las posiciones 4 y 7 del anillo de benzoxazol de **29** (Figura 5.7c) indican que estos grupos interactúan con los residuos

próximos en el sitio de unión a través de interacciones hidrófobas. Los residuos cercanos a estos contornos son la Tyr188 y las cadenas laterales de los residuos Pro95, Leu100 y Tyr181. Esta observación está de acuerdo con los estudios SAR para análogos de piridinona que tienen un anillo de benzoxazol. En estos estudios se concluyó que compuestos disustituidos con cloros o grupos metilo en las posiciones 4 y 7 del anillo de benzoxazol son más activos que los correspondientes compuestos no sustituidos (Garg, 1999; Hoffman, 1993; Saari, 1992). Esto también está de acuerdo con estudios QSAR-2D que se han reportado para análogos de piridinona. En estos estudios se concluyó que suceden interacciones hidrófobas entre los sustituyentes en el anillo de benzoxazol y el receptor (Gaarg, 1999; Gupta, 2002).

Conclusiones

Se hizo la predicción de la conformación de unión de 40 análogos de piridinona en el sitio de unión de los INNTR. De acuerdo a los modelos de acoplamiento, las moléculas adoptan una conformación tipo mariposa y la mayoría de ellas tienen un modo de unión similar a la nevirapina. Las energías libres de unión calculadas mostraron una correlación aceptable con la actividad biológica reportada. La correlación fue mejor para compuestos estructuralmente similares.

A partir de las posiciones de acoplamiento predichas, se efectuaron estudios QSAR-3D. De esta manera, se combinaron técnicas computacionales de diseño basado en el receptor (acoplamiento molecular) con técnicas de diseño basado en el ligando (QSAR).

Se obtuvieron modelos de CoMFA y CoMSIA robustos con alta predictibilidad para los grupos de entrenamiento y prueba. Los modelos de CoMSIA mostraron ligeramente mejor predictibilidad que los modelos de CoMFA. Los modelos QSAR-3D desarrollados también sirvieron para dar mayor validación a las predicciones de acoplamiento hechas por AutoDock. De acuerdo a estos modelos, suceden interacciones de tipo estérico, electrostático e hidrófobo entre los análogos de piridinona y el sitio de unión de los INNTR. Se predicen interacciones importantes con los residuos Tyr181 y Tyr188, las cuales ayudan a explicar la pérdida de actividad de las piridinonas contra cepas de TR que contienen mutaciones.

El análisis de las interacciones fármaco-receptor son valiosas para racionalizar, a nivel molecular, el diseño de nuevos análogos de piridinona. Este análisis apoya la estrategia de modificar el anillo de benzoxazol de los análogos de piridinona para superar el problema de las mutaciones Tyr181Cys y Tyr188Cys (cf., Capítulo 4).

Capítulo 6

Análisis QSAR-2D y búsqueda de nuevos líderes potenciales en la base de datos del Instituto Nacional de Cáncer

La basta información de estudios SAR disponible para los derivados de piridinona discutidos en el Capítulo 4 presenta la oportunidad de aplicar métodos basados en la estructura del ligando para el diseño de nuevos compuestos antivirales. En el Capítulo 5 se presentó la aplicación de dos métodos QSAR-3D en tres dimensiones, CoMFA y CoMSIA, para análogos de piridinona. Debido al inconveniente que tienen los métodos QSAR-3D, que es conocer la conformación bioactiva de los ligandos, se utilizó acoplamiento molecular para realizar la alineación. La estrategia anterior implicó la conveniencia de utilizar un método computacional basado en la estructura (acoplamiento molecular) para poder realizar un estudio (QSAR) que, por su propia naturaleza, no necesita conocer la estructura tridimensional del receptor.

Una de las ventajas que tienen los métodos QSAR-2D sobre los métodos QSAR-3D es que no necesitan conocer o calcular la conformación bioactiva de las moléculas. Otra ventaja de los métodos QSAR-2D es que pueden aplicarse fácilmente a la tarea de hacer búsquedas automatizadas en bases de datos (Shen, 2002).

En este capítulo se presentan los resultados de la aplicación de uno de los métodos QSAR-2D, *k*NN-QSAR, a una base de datos de derivados de piridinonas con actividad inhibitoria de la TR. Los modelos QSAR son empleados para la búsqueda de nuevos compuestos antivirales en una base de datos pública. Este estudio computacional forma parte de una estrategia general para el diseño de inhibidores de la TR de VIH-1 (cf., Figura 4.6).

Este estudio es un QSAR típico en cuanto a que se puede obtener información valiosa sin necesidad de conocer la estructura del receptor. En el caso de los INNTR, esto es especialmente atractivo porque la flexibilidad del sitio de unión de los INNTR (cf., Capítulo 4) representa un reto adicional al aplicar métodos de diseño basado en la estructura del receptor.

Además de la importancia que tiene este estudio QSAR en la investigación de nuevos compuestos antivirales, es atractivo desde el punto de vista computacional. El método QSAR-

2D empleado en el presente trabajo y su aplicación a la búsqueda de nuevos compuestos en una base de datos pública es de reciente aparición (Shen, 2004). El presente estudio es la primera búsqueda automatizada en una base de datos de nuevos compuestos con actividad inhibitoria de la TR de VIH-1, a partir de modelos *k*NN-QSAR. Esto contribuye a extender la aplicabilidad de esta metodología a otros sistemas biológicos además de los ya estudiados (Shen, 2004).


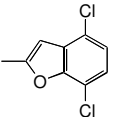
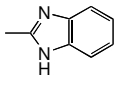
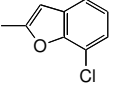

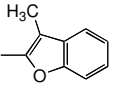


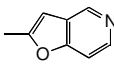
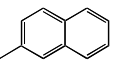
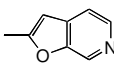

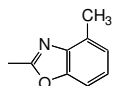
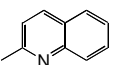
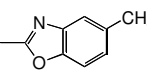
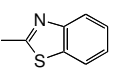
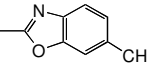
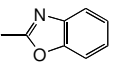
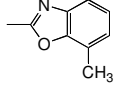
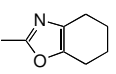
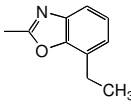
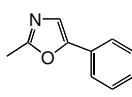
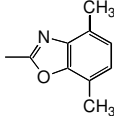
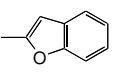
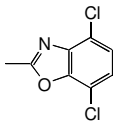
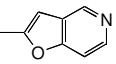
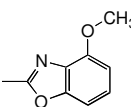
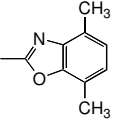
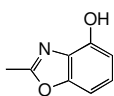
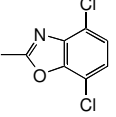
6.1. Estudio QSAR de análogos de piridinona usando el método *k*-Nearest Neighbors

Las estructuras y actividad *in vitro* de los cuarenta y cuatro derivados de piridinona empleados en este estudio están listados en la Tabla 6.1. Las actividades fueron transformadas en los correspondientes valores $-\log IC_{50}$, donde IC_{50} es la concentración efectiva de compuesto requerida para alcanzar el 50 % de inhibición de la TR. La actividad biológica fue tomada de la literatura (Hoffman, 1993; Saari, 1992).

Tabla 6.1. Estructuras, actividad inhibitoria de la TR, experimental y calculada por los modelos *k*NN-QSAR, de 44 análogos de piridinona usados en este estudio

Comp.	L	Ar	Exp. $-\log IC_{50}$	Calc. $-\log IC_{50}$	Comp.	L	Ar	Exp. $-\log IC_{50}$	Calc. $-\log IC_{50}$
1	NH		6.68	6.49	8	NH		6.28	6.10
2	NH		6.48	6.39	9	NH		5.96	5.78
3	NH		6.55	6.70	10	NH		5.72	5.51
4	NH		6.47	6.11	11	NH		5.63	5.79
5	NH		6.46	6.54	12	NH		5.60	5.68
6	NH		6.34	6.04	13	NH		5.57	5.57
7	NH		6.30	6.31	14	NH		5.36	5.33

Tabla 6.1. (continuación)

Comp.	L	Ar	Exp. -LogIC ₅₀	Calc. -LogIC ₅₀	Comp.	L	Ar	Exp. -LogIC ₅₀	Calc. -LogIC ₅₀
15	NH		5.27	4.91	30	NH		7.24	7.22
16	NH		5.12	5.54	31	NH		6.47	6.51
17	NH		4.82	4.44	32	NH		5.71	6.51
18	NH		4.65	5.07	33	CH ₂		4.30	4.64
19	NH		4.49	4.54	34	CH ₂		5.08	5.43
20	NH		3.98	4.54	35	CH ₂		3.59	4.10
21	NH		6.92	6.80	36	CH ₂		5.61	5.82
22	NH		5.90	5.83	37	CH ₂		6.43	6.43
23	NH		5.78	5.81	38	CH ₂		7.64	7.18
24	NH		7.26	7.21	39	CH ₂		7.24	7.02
25	NH		6.59	6.43	40	CH ₂		6.41	6.12
26	NH		7.70	7.65	41	CH ₂		6.89	6.72
27	NH		7.72	7.61	42	CH ₂		4.87	5.19
28	NH		6.74	6.56	43	CH ₂		7.55	7.65
29	NH		6.36	6.43	44	CH ₂		7.85	7.67

Estabilidad de los modelos QSAR

La estabilidad de un modelo QSAR se establece comparando los resultados obtenidos para el grupo original de datos con los resultados obtenidos para un grupo de datos con actividades aleatorias (Tropsha, 2003b). La Figura 6.1 muestra un gráfico de los valores de q^2 versus el número de descriptores para grupos de datos con actividades biológicas reales y aleatorias. Cada valor de q^2 es el promedio de diez cálculos independientes. En general, se obtuvieron valores de q^2 significativamente mayores cuando se emplearon valores de actividad biológica reales que cuando se emplearon actividades aleatorias (Figura 6.1).

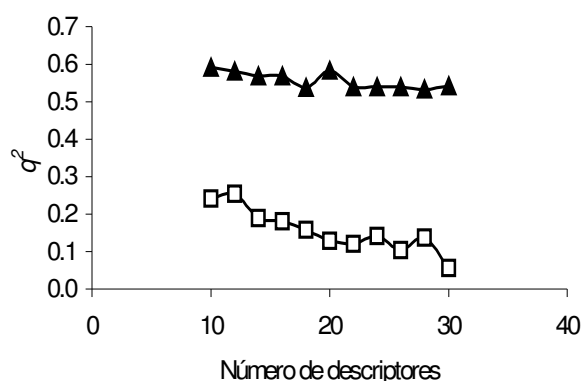


Figura 6.1. Gráficos de q^2 versus el número de descriptores seleccionados por los mejores modelos k NN-QSAR. Se muestran los resultados para el grupo de datos con actividades reales (triángulos) y aleatorias (cuadrados). Cada valor de q^2 es el promedio de 10 cálculos independientes.

El análisis estadístico de los resultados se realizó con la prueba de hipótesis de una-cola, tal como está descrito en la metodología computacional. El valor de q^2 promedio para los modelos con 30 descriptores con actividades aleatorias fue de 0.06. La desviación estándar fue de 0.18. El valor Z para el modelo más significativo con 30 descriptores, para el grupo con actividades biológicas reales, fue de 2.71. Esto indica que la probabilidad de que el modelo k NN-QSAR desarrollado para el grupo con actividades reales, sea aleatorio, es de menos de 0.01.

Validación de los modelos QSAR

Para obtener modelos QSAR predictivos y confiables es necesario demostrar que los modelos generados para el grupo de entrenamiento pueden predecir con precisión las actividades de los compuestos en grupos de prueba (Tropsha, 2003b). Se aceptaron modelos con valores de q^2 , para el grupo de entrenamiento, superiores a 0.5 y con valores de R^2 , para

los grupos de prueba, mayores a 0.6 (Golbraikh, 2002a; Golbraikh, 2002b). La Tabla 6.2 presenta a los 10 mejores modelos QSAR obtenidos de múltiples análisis k NN. En la Figura 6.2 se muestran valores de actividades reales *versus* calculados para los grupos de entrenamiento y prueba, basados en el mejor modelo. El número de descriptores para los 10 mejores modelos varió entre 10 y 26. Aún para grupos de prueba tan grandes como de 19 compuestos, se obtuvo un modelo con $q^2 = 0.52$ y $R^2 = 0.75$ (Tabla 6.2).

Tabla 6.2. Los diez mejores modelos k NN-QSAR

Número de modelo	Tamaño del grupo de entrenamiento (núm. moléculas)	Tamaño del grupo de prueba (núm. moléculas)	Número de descriptores	q^2 (grupo de entrenamiento)	R^2 (grupo de prueba)	R_0^2	$R_{0'}^2$
1	36	8	14	0.62	0.89	0.88	0.89
2	36	8	14	0.60	0.87	0.85	0.87
3	36	8	12	0.61	0.75	0.70	0.75
4	36	8	26	0.55	0.74	0.48	0.72
5	36	8	10	0.58	0.72	0.68	0.71
6	31	13	10	0.59	0.72	0.72	0.68
7	31	13	20	0.63	0.71	0.60	0.71
8	28	16	16	0.63	0.75	0.46	0.72
9	27	17	24	0.53	0.71	0.48	0.70
10	25	19	10	0.52	0.75	0.72	0.42

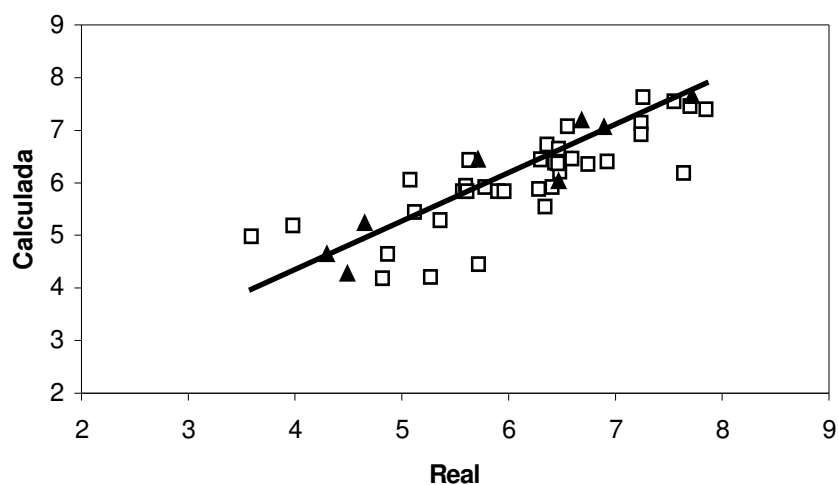


Figura 6.2. Valores de actividad reales y calculados para el grupo de entrenamiento (cuadrados) y prueba (triángulos) empleando el modelo k NN-QSAR número 1 ($q^2 = 0.62$, $R^2 = 0.89$) (Tabla 6.2).

6.2. Búsqueda automatizada de nuevos líderes en la base de datos del Instituto Nacional de Cáncer

Se encontraron varios compuestos en la base de datos del NCI para los cuales no se ha documentado ampliamente su actividad, ya sea como inhibidores de la TR o como compuestos anti-VIH. La estructura de algunos de estos compuestos está en la Figura 6.3. Fue especialmente interesante observar la similitud estructural entre el compuesto NCI 127 y nevirapina: ambas estructuras consisten en un sistema de tres anillos que involucran a dos anillos de seis miembros en los extremos. Ambos compuestos tienen un sustituyente pequeño (grupo metilo o ciclopropilo) unido a un nitrógeno localizado en el anillo central (Figura 6.3).

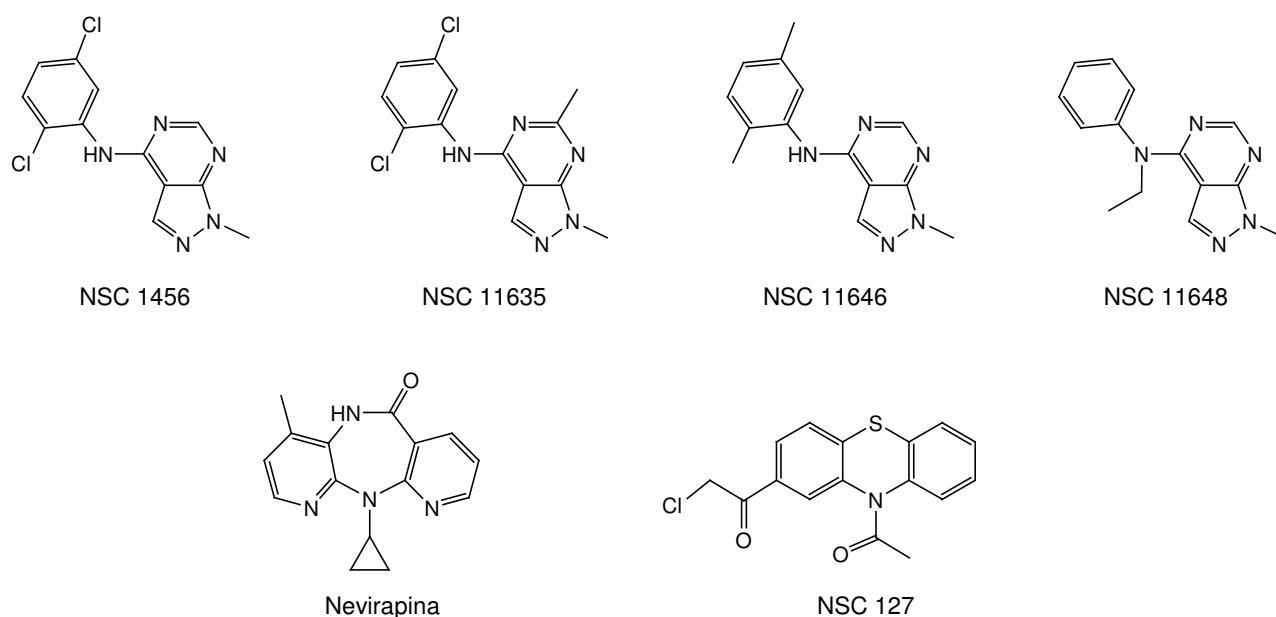


Figura 6.3. Estructuras químicas de compuestos encontrados en la búsqueda en la base de datos del NCI. Se muestra la estructura de la nevirapina (cf., Capítulo 4) como referencia.

Acoplamiento de moléculas seleccionadas de la búsqueda en la base de datos del NCI

Se inspeccionaron visualmente los compuestos que fueron predichos más activos en la búsqueda de la base de datos. De acuerdo con algunas características estructurales comunes a los INNTR, por ejemplo, la posibilidad de adoptar una conformación de mariposa dentro del sitio de unión (Ren, 1995a) (cf., Capítulo 4), se seleccionaron algunos compuestos que fueron acoplados en la geometría del sitio de unión de la nevirapina. Los resultados se muestran en la Tabla 6.3.

Tabla 6.3. Acoplamiento molecular de compuestos seleccionados de la búsqueda de la base de datos del NCI en el sitio de unión de la nevirapina

Compuesto	Grupo ^a	Miembros en el grupo	Energía más baja (kcal/mol) ^b	Energía promedio (kcal/mol) ^c
NSC 11635	1	93	-10.03	-10.02
NSC 11646	1	100	-9.37	-9.36
NSC 1456	1	62	-9.34	-9.29
NSC 11648	1	9	-8.61	-8.57
NSC 127	2	43	-10.28	-10.16
	4	40	-9.74	-9.71

^aSe muestran los mejores grupos de un total de 100 repeticiones.

^bValor para la estructura óptima del grupo. Las energías de acoplamiento son las determinadas por AutoDock.

^cEnergía de acoplamiento promedio del grupo.

Se encontró que las pirazo[3,4-*d*]pirimidinas NSC 1456, 11635, 11646 y 11648 (Cheng, 1956) pueden adoptar una conformación de mariposa en el sitio de unión de la nevirapina. De hecho, la conformación de unión predicha es similar a la conformación de la nevirapina. El sitio de unión de la mayoría de estos compuestos está formado por los residuos Pro95, Leu100, Lys101, Lys102, Lys103, Val106, Tyr181, Tyr188, Phe227, Trp229, Leu234, His235, Pro236, Asp237, Tyr318 de la subunidad p66 y Glu138 de la subunidad p51 (Ren, 1995a). La mayoría de los residuos son comunes al sitio de unión de otros INNTR (por ejemplo, Das, 2004; Hopkins, 1996; Ren, 1995a; Ren, 1995b). El modo de unión del compuesto para el cual se predice la mayor actividad, NSC 11635, está representado en la Figura 6.4. Para comparación, se muestra el modo de unión de la nevirapina. El grupo metilo en la posición 6 del anillo de pirazolo[3,4-*d*]pirimidina de NSC 11635 mostró contactos adicionales con la Pro225 de la subunidad p66. Estos contactos adicionales pueden explicar la energía de interacción más favorable calculada por AutoDock para NSC 11635 comparada con la energía de interacción calculada para NSC 11646 y NSC 1456 (Tabla 6.3). Este grupo metilo, que no está presente en las estructuras de NSC 11646 y NSC 1456 (Figura 6.3), pudiera ser una característica estructural importante en la posible actividad de NSC 11635. Por su parte, NSC 11648 presentó contactos adicionales entre el grupo etilo y los residuos Val179 y Gly190 de la subunidad p66 (cf., Figura 6.5d). Sin embargo, para este compuesto no se observaron contactos con los residuos Asp237 y Glu138.

Para comparar el modo de unión predicho para las pirazo[3,4-*d*]pirimidinas con el de los derivados de piridinona (cf., Capítulo 4), el modo de unión de las pirazo[3,4-*d*]pirimidinas fue superpuesto con el modo de unión calculado para **27**. Anteriormente se mostró que el modo

de unión de **27** es representativo de los análogos de piridinona más potentes (cf., Capítulo 4). Las comparaciones se muestran en la Figura 6.5.

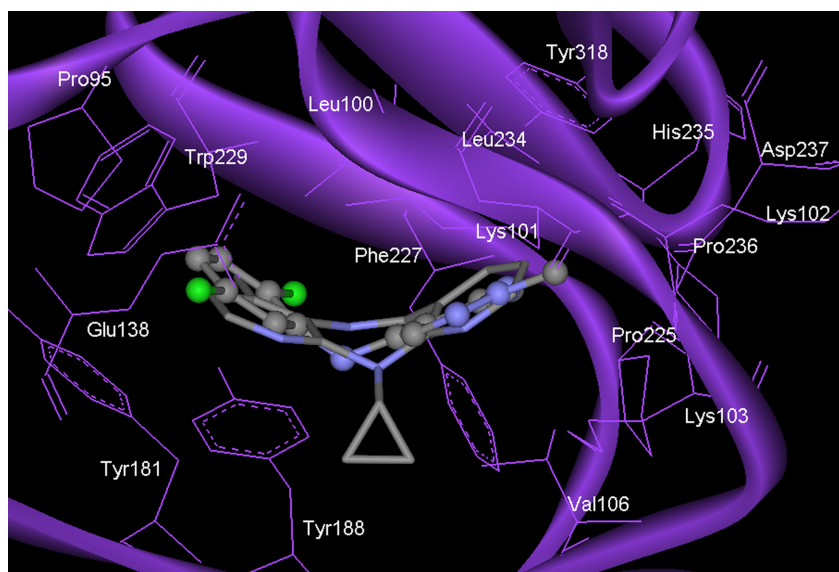


Figura 6.4. Modo de unión predicho para NSC 11635 (esferas y barras) en el sitio de unión de la TR. Se muestra a la nevirapina (barras) para comparación. Para mayor claridad, se omiten los hidrógenos.

El modo de unión predicho para todos los compuestos es similar al modo de unión del derivado de piridinona. El grupo fenilo de NSC 11635, NSC 11646 y NSC 1456, todos con una disustitución 2,5-dicloro o 2,5-dimetilo, pueden superponerse, aproximadamente, con el anillo de benzoxazol de **27**. Es notable el hecho de que ambos grupos cloro o metilos sobre el anillo de fenilo de las pirazo[3,4-*d*]pirimidinas ocupan casi la misma posición de unión que los átomos de cloro de **27** (Figura 6.5a-c) interaccionando principalmente con los residuos Pro95, Tyr181, Tyr188 y Trp229. Es interesante notar que los análogos de piridinona disustituídos con cloros o metilos en las posiciones 4 y 7 del anillo de benzoxazol son más activos que los derivados correspondientes sin sustituir (Hoffman, 1993; Saari, 1992). Estas observaciones sugieren la posible importancia de estos grupos cloro o metilo de los compuestos NSC 11635, NSC 11646 y NSC 1456, en la actividad biológica predicha para las pirazo[3,4-*d*]pirimidinas.

Para NSC 11648, que carece de sustituyentes en el fenilo, se predijo una conformación alterna respecto a la orientación de unión de las otras tres pirazo[3,4-*d*]pirimidinas. El anillo de pirazo[3,4-*d*]pirimidina ocupa una región similar que el anillo de benzoxazol de **27** (Figura 6.5d). Se predice que el grupo etilo ocupa la misma región de unión que el grupo aminometileno del análogo de piridinona.

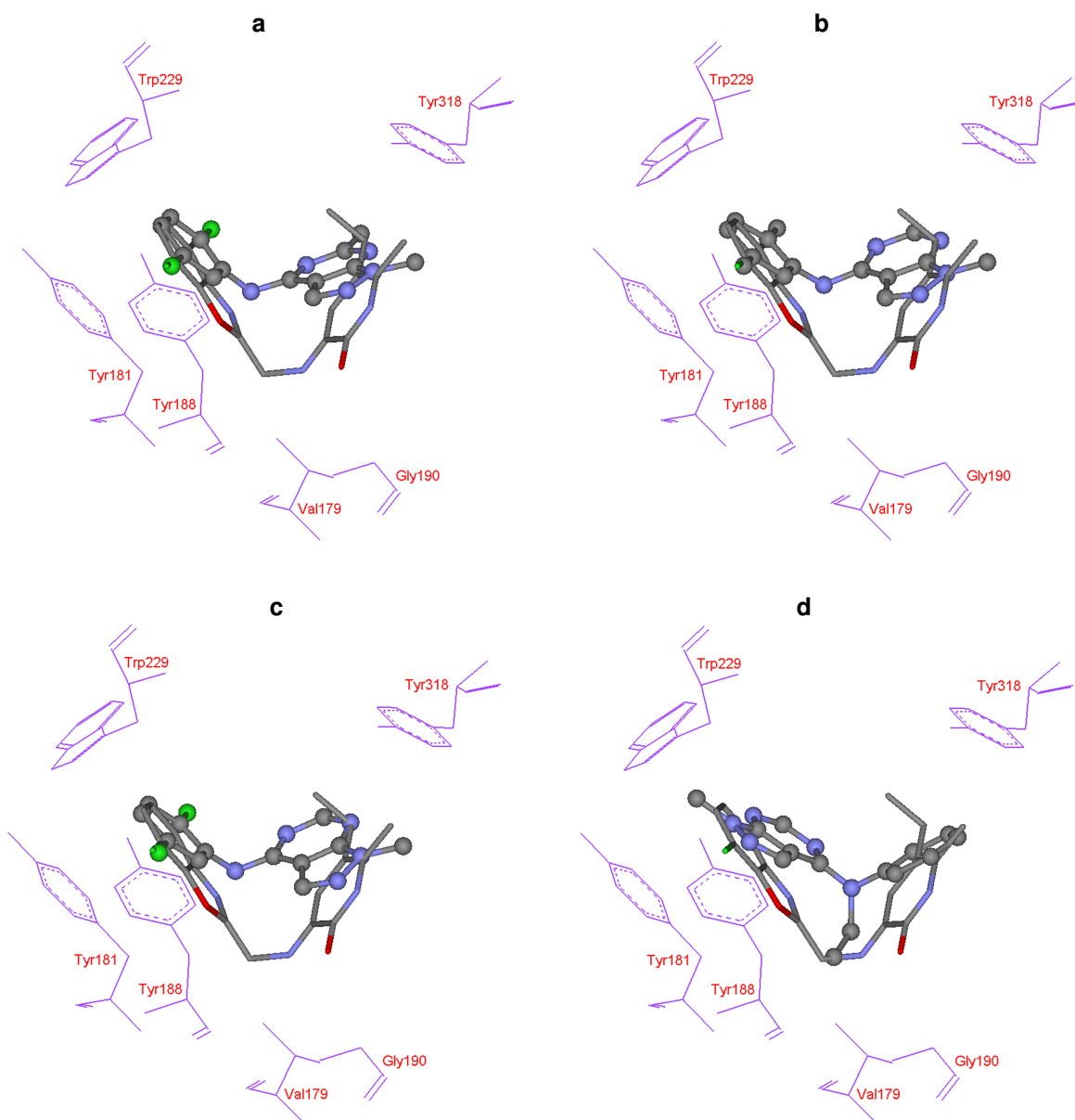


Figura 6.5. Comparación de las conformaciones de unión de las pirazo[3,4-*d*]pirimidinas (esferas y barras) con **27** (barras) (Capítulo 4). (a) NSC 11635. (b) NSC 11646. (c) NSC 1456. (d) NSC 11648. Las posiciones de Val179, Tyr181, Tyr188, Gly190, Trp229 y Tyr318 se muestran como referencia. Para mayor claridad, se omiten los hidrógenos.

En la Tabla 6.4 se muestra una comparación entre la actividad predicha por los modelos *k*NN-QSAR, la energía de acoplamiento y la energía libre de unión calculada por AutoDock para las cuatro pirazo[3,4-*d*]pirimidinas. Se observó una buena relación entre las predicciones de QSAR y las evaluaciones de AutoDock. También se observó una muy buena correlación entre las actividades predichas por los modelos *k*NN-QSAR para las 44 piridinonas con las actividades biológicas experimentales (Tabla 6.1). Por otra parte, en el Capítulo 5, se observó

una correlación aceptable entre las energías de acoplamiento calculadas por AutoDock para análogos de piridinona, especialmente aquellos que son estructuralmente similares, por ejemplo, los análogos que contienen un anillo de benzoxazol. Todas estas observaciones sugieren que las pirazo[3,4-*d*]pirimidinas representan nuevos líderes potenciales del tipo de INNTR.

Para el compuesto NSC 127 (Chan, 1998) también se encontró que adopta una conformación de mariposa en el sitio de unión de la nevirapina. Los dos modos de unión predichos, representados por los grupos más poblados 2 y 4 (Tabla 6.3), son similares al modo de unión de la nevirapina. El sitio de unión de NSC 127 está formado por los residuos Pro95, Leu100, Lys101, Lys102, Lys103, Val106, Val179, Tyr181, Tyr183, Tyr188, Val189, Gly190, Phe227, Trp229, Leu234, His235, Pro236, Asp237 y Tyr318 (Figura 6.6). El grupo metilcetona de NSC 127 puede estar ocupando la misma región de unión que el grupo ciclopropilo de la nevirapina haciendo contactos con Val106, Val179, Tyr181, Tyr188, Val189 y Gly190.

Tabla 6.4. Actividad predicha por los modelos *k*NN-QSAR, energía de acoplamiento y energía libre de unión calculada por AutoDock para las pirazo[3,4-*d*]pirimidinas

Compuesto	Actividad predicha por los modelos <i>k</i> NN (-LogIC ₅₀) ^a	Energía de acoplamiento (kcal/mol) ^b	Energía libre de unión (kcal/mol) ^b
NSC 11635	6.44	-10.03	-9.40
NSC 11646	6.34	-9.37	-8.72
NSC 1456	6.12	-9.34	-8.65
NSC 11648	5.89	-8.61	-7.98

^aPromedio de las predicciones. ^bValor para la estructura óptima.

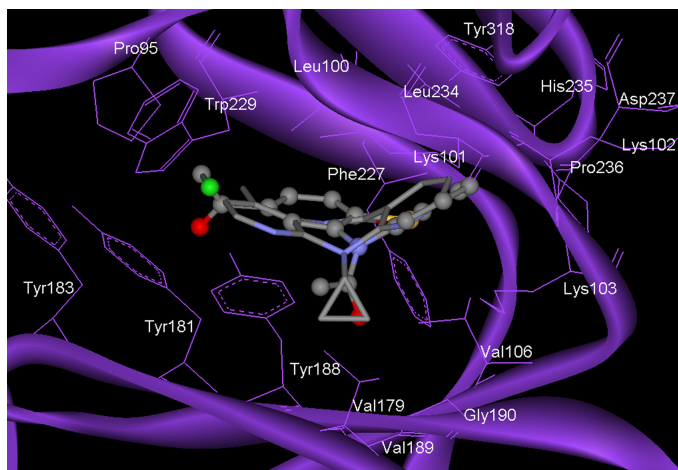


Figura 6.6. Modo de unión más favorable (grupo 2) de NSC 127 (esferas y barras) en el sitio de unión de la TR. Se muestra a la nevirapina (barras) para comparación. Para mayor claridad, se omiten los hidrógenos.

Conclusiones

La estrategia aplicada en este capítulo para el diseño de nuevos líderes con potencial actividad inhibitoria de la TR involucró el uso de tres técnicas computacionales. A partir de modelos QSAR-2D (modelos *k*NN-QSAR), que fueron validados por diversas técnicas, se hizo una búsqueda en una base de datos pública. Para aplicar estos dos métodos, QSAR y búsqueda virtual en bases de datos, no fue necesario considerar la estructura del receptor. Los prototipos computacionales encontrados en la búsqueda en la base de datos representan en sí nuevos líderes antivirales potenciales.

Gracias a que se conoce la estructura de la TR, fue posible utilizar acoplamiento molecular para refinar los resultados de la búsqueda. Esto representa una de las prácticas más recomendadas en diseño de fármacos asistido por computadora. Esto es, el uso de la mayor cantidad de información disponible para hacer predicciones lo más confiables posibles. Fue muy interesante observar en este estudio que, tanto los métodos de diseño basado en el ligando como los basados en el receptor, permitieron llegar a conclusiones similares.

Se predijo que diversas clases de compuestos tienen actividad. Los líderes más prometedores son los análogos de pirazolo[3,4-*d*]pirimidina y fenotiazina. Experimentos de acoplamiento flexible de este tipo de compuestos con la TR de VIH-1 sugieren que estos compuestos pueden unirse en el sitio de unión de los INNTR con un modo de unión muy similar que otros INNTR, especialmente la nevirapina.

La perspectiva de este trabajo es realizar las pruebas biológicas de los análogos que pueden ser potentes tales como NSC 11635 y NSC 127. Aunque las moléculas propuestas pudieran generar resistencia, representan estructuras que pueden optimizarse posteriormente con la asistencia de métodos computacionales. Por ejemplo, empleando los modelos de acoplamiento y *k*NN-QSAR desarrollados en este trabajo.

La estrategia empleada en este capítulo para el diseño racional de nuevos inhibidores de la TR de VIH-1 puede aplicarse para estudios similares de otras series de compuestos con otro tipo de actividad biológica.

Parte II

Compuestos hipocolesterolemiantes

Capítulo 7

Acoplamiento molecular con la 3-hidroxi-3-metilglutaril-CoA reductasa y diseño de nuevos compuestos

7.1. Inhibidores de la 3-hidroxi-3-metilglutaril-CoA reductasa

La cardiopatía isquémica, que es la principal manifestación de la artereosclerosis, es la causa de muerte primordial en países industrializados. También es un problema en crecimiento en países en vías de desarrollo (Packard, 2002).

La artereosclerosis está fuertemente asociada a la hipercolesterolemia. De hecho, cambios en la dieta y el uso de fármacos hipocolesterolemiantes (Packard, 2002; Wild, 2001) han mostrado reducir el riesgo de infarto y los niveles de mortalidad (Istvan, 2001).

Uno de los blancos más atractivos para desarrollar fármacos que reduzcan los niveles de colesterol en el plasma es la enzima 3-hidroxi-3-metilglutaril-CoA reductasa (HMGR) (Corsini, 1995). Esta enzima cataliza en el hígado la reducción de 3-hidroxi-3-metilglutaril-CoA (HMG-CoA) a mevalonato (Istvan 2000b) y regula principalmente la biosíntesis del colesterol (Corsini, 1995). En estudios clínicos se ha observado que la inhibición de la HMGR disminuye significativamente los niveles de colesterol, reduce el riesgo de infarto en un 29 % y la mortalidad en un 22 % (Istvan, 2000a).

Inhibidores de la HMGR, comúnmente referidos en la literatura como estatinas (Tobert, 2003), son fármacos ampliamente prescritos en la terapia de reducción del colesterol (Cheng, 2004; Link, 2004). Las estatinas son inhibidores competitivos de la HMGR respecto a la unión del sustrato (HMG-CoA) (Istvan, 2001). Se caracterizan porque en su estructura tienen un fragmento que es similar a HMG-CoA (Figura 7.1). Este fragmento puede estar presente en forma de lactona inactiva como en la simvastatina. *In vivo*, estos profármacos son hidrolizados enzimáticamente a la forma activa de ácido (Istvan, 2001). Ejemplos de estatinas que actualmente están en el mercado son simvastatina (ZOCOR[®], Merck) y rosuvastatina (CRESTOR[®], AstraZeneca) (Figura 7.1). Sin embargo, muchos pacientes no muestran los efectos esperados después de una terapia con estatinas (Cheng, 2004).

La HMGR es una proteína transmembranal con 888 amino ácidos. La región amino-terminal consta de 339 residuos y están en la membrana del retículo endoplásmico. La porción catalítica de la enzima está formada por los residuos 460-888 en la región C-terminal. La porción catalítica está en el citoplasma. Los residuos 340-459 conectan a las dos porciones de la proteína (Istvan, 2000b).

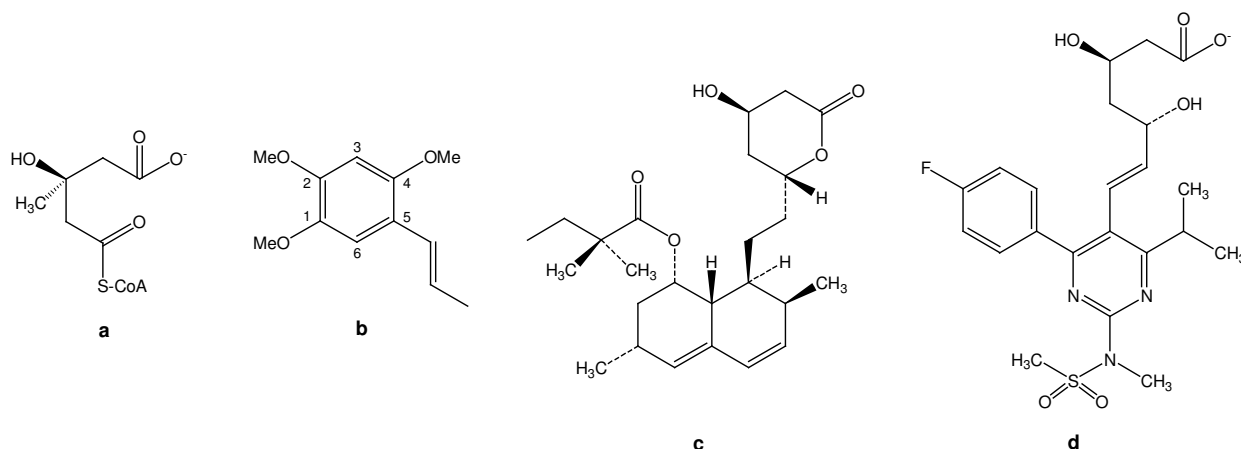


Figura 7.1. Estructura de (a) HMG-CoA, (b) α -asarona y las estatinas (c) simvastatina y (d) rosuvastatina.

Recientemente, se ha determinado la estructura cristalográfica de la porción catalítica de la HMGR humana en complejo con el sustrato (Istvan, 2000), y con seis diferentes estatinas (Istvan, 2001a). La enzima forma tetrámeros y tiene cuatro sitios activos formados por los residuos de dos monómeros (Figura 7.2). El sitio de unión está caracterizado por el llamado *cis-loop* (residuos 682-694). En la Figura 7.3 se muestra un modelo en dos dimensiones de las interacciones del sustrato con el sitio activo de la enzima. Un mecanismo catalítico propuesto sugiere que los residuos Lys691 y Glu559 participan directamente en la reducción del sustrato (Istvan, 2000b).

La α -asarona (Figura 7.1) es el principal componente activo del extracto de la corteza de *Gutteria gaumeri* (Enríquez, 1980), una planta medicinal utilizada en México para el tratamiento de la hipercolesterolemia (Martínez, 1992). α -Asarona ha mostrado actividad hipolipidémica (Chamorro, 1993; Garduño, 1997) y recientemente ha mostrado la capacidad que tiene para inhibir a la HMGR (Rodríguez-Páez, 2003). Esto ha sugerido la posibilidad de que la actividad hipolipidémica de la α -asarona se deba a un mecanismo similar al de las estatinas. Sin embargo, no está disponible una estructura cristalográfica de la HMGR en complejo con la α -asarona que permita saber con precisión el tipo de interacciones que ocurren entre ambas moléculas.



Figura 7.2. Estructura del homotetrámero de la HMGR humana (código PDB: 1dqa). Los residuos de los cuatro monómeros están en amarillo, púrpura, rojo y azul, respectivamente.

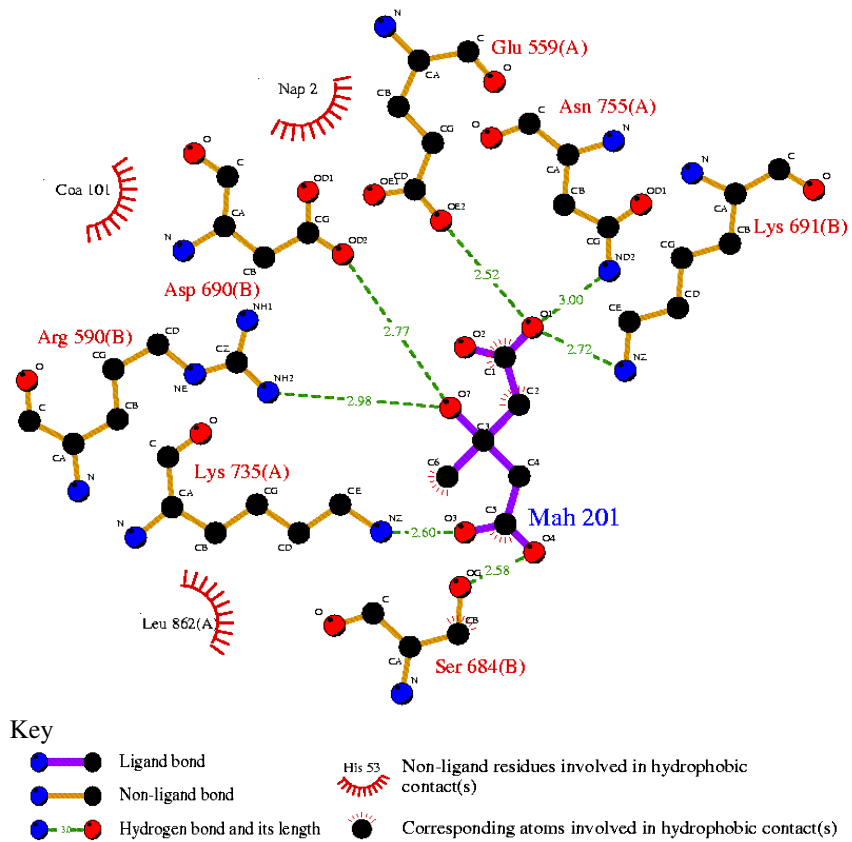


Figura 7.3. Representación 2D de las principales interacciones de HMG-CoA con la HMGR.

Gracias al conocimiento de la estructura tridimensional de la HMGR, se pueden aplicar técnicas de diseño basado en la estructura del receptor para generar un modelo de unión α -asarona-HMGR. En este capítulo se muestran los resultados de acoplamiento molecular para desarrollar este modelo. A partir del modelo de unión, se pueden identificar posibles partes farmacofóricas de la α -asarona y diseñar nuevos inhibidores potenciales de la HMGR.

7.2. Estudio de la interacción de α -asarona con la 3-hidroxi-3-metilglutaril-CoA reductasa

Para efectuar los experimentos de acoplamiento molecular se empleó la estructura de la HMGR formando complejo con HMG-CoA (Istvan, 2000a) y con las estatinas (Istvan, 2001). Las estructuras de la HMG se obtuvieron del *Protein Data Bank* (RCSB) (códigos PDB: 1dq8, 1hw8, 1hw9, 1hwi, 1hwj, 1hwk y 1hwl).

Validación del protocolo de acoplamiento

Antes de efectuar el acoplamiento de la α -asarona, se validó el protocolo para cada una de las estructuras cristalográficas. La validación consistió en predecir la orientación y conformación de unión de HMG-CoA y cada estatina en su respectiva estructura cristalográfica.

En todos los casos, se obtuvieron desviaciones (RMSD), entre las coordenadas cristalográficas y las calculadas por AutoDock, inferiores o cercanos a 1 Å. Simvastatina y rosuvastatina (Figura 7.1) fueron las estatinas para las cuales se obtuvieron las mejores predicciones. En la Tabla 7.1 se resumen los resultados de la validación para el substrato, la simvastatina y la rosuvastatina. En la tabla se incluyen los valores de energía de acoplamiento, la energía libre de unión (ΔG) y la constante de inhibición (K_i) calculados por AutoDock. Se incluye también la actividad biológica experimental (Istvan, 2001, McTaggart, 2001).

Tabla 7.1. Acoplamiento flexible de HMG y estatinas representativas

Compuesto	Energía de acoplamiento (kcal/mol)	ΔG (kcal/mol)	K_i (a 298.15 K)	RMSD	IC ₅₀ (M) ^a experimental
HMG	-11.27	-10.74	23.5 x 10 ⁻⁹	1.19	4000 x 10 ^{-9b}
Simvastatina	-15.29	-11.70	2.6 x 10 ⁻⁹	0.89	11.2 x 10 ⁻⁹
Rosuvastatina	-15.81	-12.34	0.9 x 10 ⁻⁹	0.81	5.4 x 10 ⁻⁹

^aIC₅₀ promedio de inhibición de HMGR (McTaggart, 2001). ^bConstante de Michaelis para HMG-CoA (Istvan, 2001).

Las energías de interacción calculadas para simvastatina y rosuvastatina (Tabla 7.1) están de acuerdo con resultados experimentales recientemente publicados que muestran que la rosuvastatina es un inhibidor más potente que la simvastatina (Holdgate, 2003; McTaggart, 2001). La energía de acoplamiento y la energía libre de unión calculadas para la rosuvastatina son más favorables que las calculadas para simvastatina.

Los resultados obtenidos indican que los parámetros usados por AutoDock reprodujeron satisfactoriamente las estructuras cristalográficas. Es de hacer notar la habilidad de AutoDock para reproducir la conformación de unión de las estatinas las cuales tienen varios enlaces flexibles.

Para analizar la influencia que tiene la flexibilidad de las estatinas sobre las predicciones de acoplamiento, se realizó un acoplamiento rígido. Para esto, se restringió, durante la etapa de la búsqueda, el movimiento de los enlaces de las estructuras cristalográficas de la estatinas. Los resultados se muestran en la Tabla 7.2. En la Figura 7.4 se muestra, para rosuvastatina, una comparación entre el modo de unión cristalográfico y el predicho por AutoDock en el acoplamiento flexible y rígido, respectivamente.

Tabla 7.2. Acoplamiento rígido de estatinas representativas

Compuesto	Energía de acoplamiento (kcal/mol)	ΔG (kcal/mol)	K_i (a 298.15 K)	RMSD
Simvastatina	-13.97	-10.54	1.88×10^{-8}	0.67
Rosuvastatina	-15.08	-11.97	1.69×10^{-9}	0.55

Como es de esperarse, en el acoplamiento rígido se obtuvieron valores de desviación más pequeños porque se está partiendo de la conformación de los ligandos que se encuentran en el cristal. De esta manera, AutoDock realiza únicamente una búsqueda translacional y no translacional más conformacional. Estas observaciones llevarían a concluir que el acoplamiento rígido es más eficiente. Sin embargo, hay que tomar en cuenta que en muchos casos, como sucede con la α -asarona, no se conoce *a priori* la conformación bioactiva. En estas circunstancias es más conveniente emplear un protocolo de acoplamiento flexible. De acuerdo a los resultados de la Tabla 7.1 se demostró que AutoDock es capaz de reproducir, con valores aceptables de desviación, las estructuras cristalográficas de las estatinas empleando acoplamiento flexible.

Las energías de interacción fueron menos favorables para las dos moléculas en el acoplamiento rígido. Esto se debe a que en un acoplamiento rígido se restringen las

posibilidades de encontrar mínimos de energía. Sin embargo, aún en el protocolo de acoplamiento rígido, AutoDock calcula que la rosuvastatina tiene mayor afinidad por la HMGR que la simvastatina (Tabla 7.2).

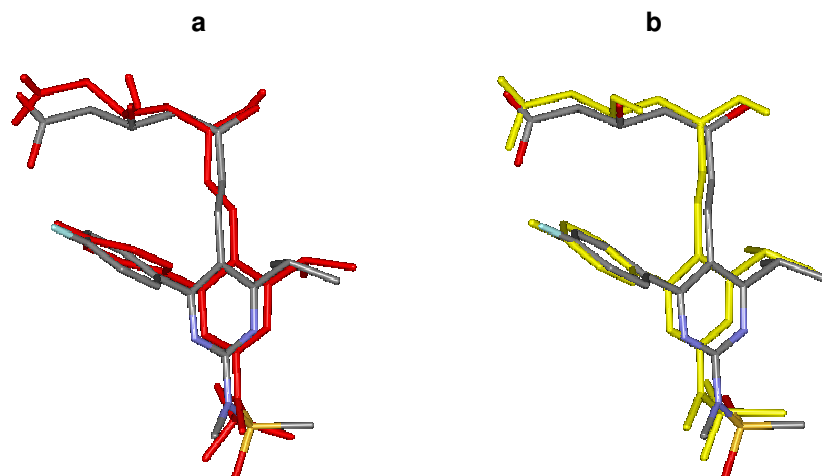


Figura 7.4. Comparación entre el modo de unión de la rosuvastatina que se encuentra en la estructura cristalográfica (color por tipo de átomo) y la conformación y posición predicha por AutoDock. (a) Acoplamiento flexible (rojo) y (b) acoplamiento rígido (amarillo). Para mayor claridad, se omiten los hidrógenos.

Acoplamiento de α -asarona

La α -asarona se acopló primero con la estructura cristalográfica de HMGR en complejo con el sustrato y en complejo con la simvastatina y rosuvastatina. Se seleccionaron las estructuras de estas dos estatinas porque con ellas se obtuvieron los menores valores de desviación en la validación. Los resultados del acoplamiento de la α -asarona están resumidos en la Tabla 7.3.

Tabla 7.3. Resultados del acoplamiento de α -asarona con la porción catalítica de la HMGR humana

Sitio de unión	Energía de acoplamiento (kcal/mol)	ΔG (kcal/mol)	K_i (a 298.15 K)
HMG-CoA	-6.60	-6.02	3.9×10^{-5}
Simvastatina	-6.29	-5.68	6.8×10^{-5}
Rosuvastatina	-6.24	-5.66	7.1×10^{-5}

La energía de acoplamiento y la energía libre de unión estimada para la α -asarona, cuando se acopló en el sitio de unión de la simvastatina y rosuvastatina, son casi idénticas (Tabla 7.3). Esto no es de sorprender ya que las geometrías de estos sitios de unión son muy parecidas (Istvan, 2001). La energía ligeramente inferior calculada con el sitio de unión del sustrato puede explicarse por las interacciones favorables con el residuo Leu862. Este residuo no está presente en el sitio de unión de la simvastatina y rosuvastatina (cf., Figura 7.6) (Istvan, 2001). Las diferencias en la conformación de las cadenas laterales de algunos residuos del sitio de unión tales como Asp690, Arg590, Leu857 y Leu853 (cf., Figura 7.8) también pueden contribuir a esta diferencia.

Es importante hacer notar que la energía libre de unión calculada para α -asarona es mucho menos favorable que la calculada para la simvastatina y rosuvastatina (comparar las Tablas 7.1 y 7.3). Esto está de acuerdo con la tendencia experimental de la inhibición de HMGR determinada para la α -asarona y las estatinas. Para la rosuvastatina y simvastatina se han determinado experimentalmente constantes de inhibición (IC_{50}) a nivel nanomolar (McTaggart, 2001), mientras que para la α -asarona el IC_{50} experimental está a nivel milimolar (Rodríguez-Páez, 2003). Estas observaciones sugieren que el modo de unión predicho para α -asarona es razonable.

AutoDock encontró un modo de unión principal para la α -asarona en todas las geometrías de unión empleadas. α -Asarona se une en el sitio activo de la HMGR ocupando la región donde se une la porción 3-hidroximetilglutaril del sustrato (Figura 7.5. Cf. Figura 7.6a).

De acuerdo con el modelo de acoplamiento desarrollado, la α -asarona interacciona con dos monómeros del tetrámero de la HMGR. Se calculan interacciones polares con los residuos Ser684, Asp690, Lys691 y Lys692, del *cis-loop*, y con Arg590 de un monómero (Figura 7.5). El grupo C-4 metoxilo de la α -asarona hace contactos polares con Glu559 del segundo monómero y forma puentes de hidrógeno con las cadenas laterales de Lys691 y Asn755. Se forman puentes de hidrógeno adicionales entre los grupos metoxilo C-1 y C-2 de la α -asarona con Ser684 y Arg590 (Figura 7.5). La cadena propenilica de la α -asarona está próxima a Glu559, Leu562, His752, Asn755 y Leu853, todos los residuos de un monómero. El anillo aromático está cercano a Leu853 y Leu862. Estas interacciones parecen ser las responsables de la afinidad de la α -asarona con HMGR representando un impedimento estérico para la unión del sustrato con la enzima. Se obtuvieron conclusiones similares cuando los experimentos de acoplamiento se efectuaron utilizando las estructuras de la HMGR en complejo con compactina, fluvastatina, cerivastatina y atorvastatina.

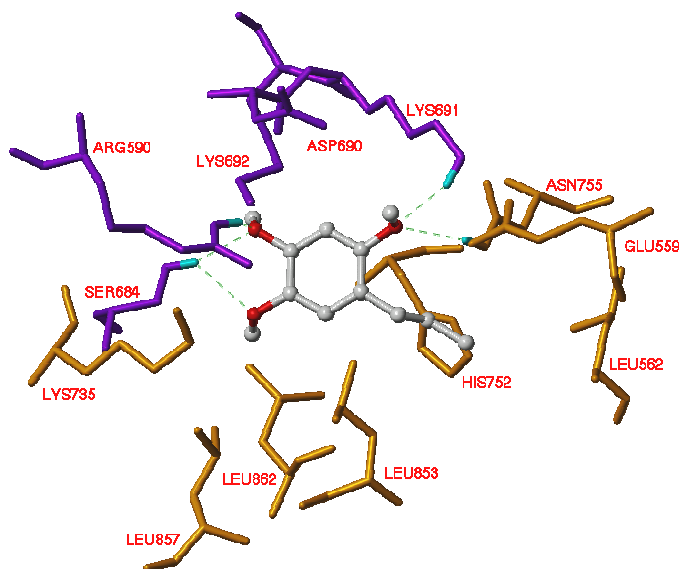


Figura 7.5. Modelo de acoplamiento derivado para la α -asarona con la porción catalítica de la HMGR. Se muestra la geometría de unión de HMG-CoA. Los puentes de hidrógeno están en líneas discontinuas amarillas. Se etiquetan residuos representativos a 3.6 Å de distancia de la α -asarona. Residuos de un monómero están en naranja y los del otro monómero en púrpura. Para mayor claridad, se omiten los hidrógenos excepto los hidrógenos involucrados en puentes de hidrógeno.

Las interacciones de la α -asarona con el sitio catalítico de HMGR tienen relación con las interacciones observadas en los complejos de la HMGR con el sustrato y las estatinas. El grupo metoxilo C-4 de la α -asarona ocupa una región similar que el oxígeno del carbonilo C-5 de HMG-CoA y que el grupo hidroxilo C-5 de las estatinas. Este último interacciona con las cadenas laterales de Lys691 y Glu559, los cuales parecen participar directamente en la catálisis (Figura 7.6) (Istvan, 2000b). De acuerdo con el mecanismo de catálisis propuesto para la reducción de HMG-CoA (Istvan, 2000b), el cual sugiere que el residuo Glu559 está protonado en condiciones fisiológicas, el oxígeno del grupo metoxilo C-4 de la α -asarona también puede formar un enlace de hidrógeno con el grupo carboxilo de Glu559.

Los grupos metoxilos C-1 y C-2 de la α -asarona también tienen interacciones similares que el grupo carboxilato C-1, tanto del sustrato como de las estatinas. Estas interacciones involucran enlaces de hidrógeno con Ser684 e interacciones con las cadenas laterales polares de Lys692 y Lys735 (Istvan, 2000a; Istvan, 2001). El grupo metoxilo C-1 está localizado en la región de unión del grupo butirilo de la simvastatina y el grupo fluorofenilo de la rosuvastatina (Figuras 7.6b y 7.6c, respectivamente). El metilo del grupo metoxi C-1 está próximo a la cadena lateral de Leu857. La cadena propenilica de la α -asarona se localiza cerca del grupo hidrófobo de las estatinas.

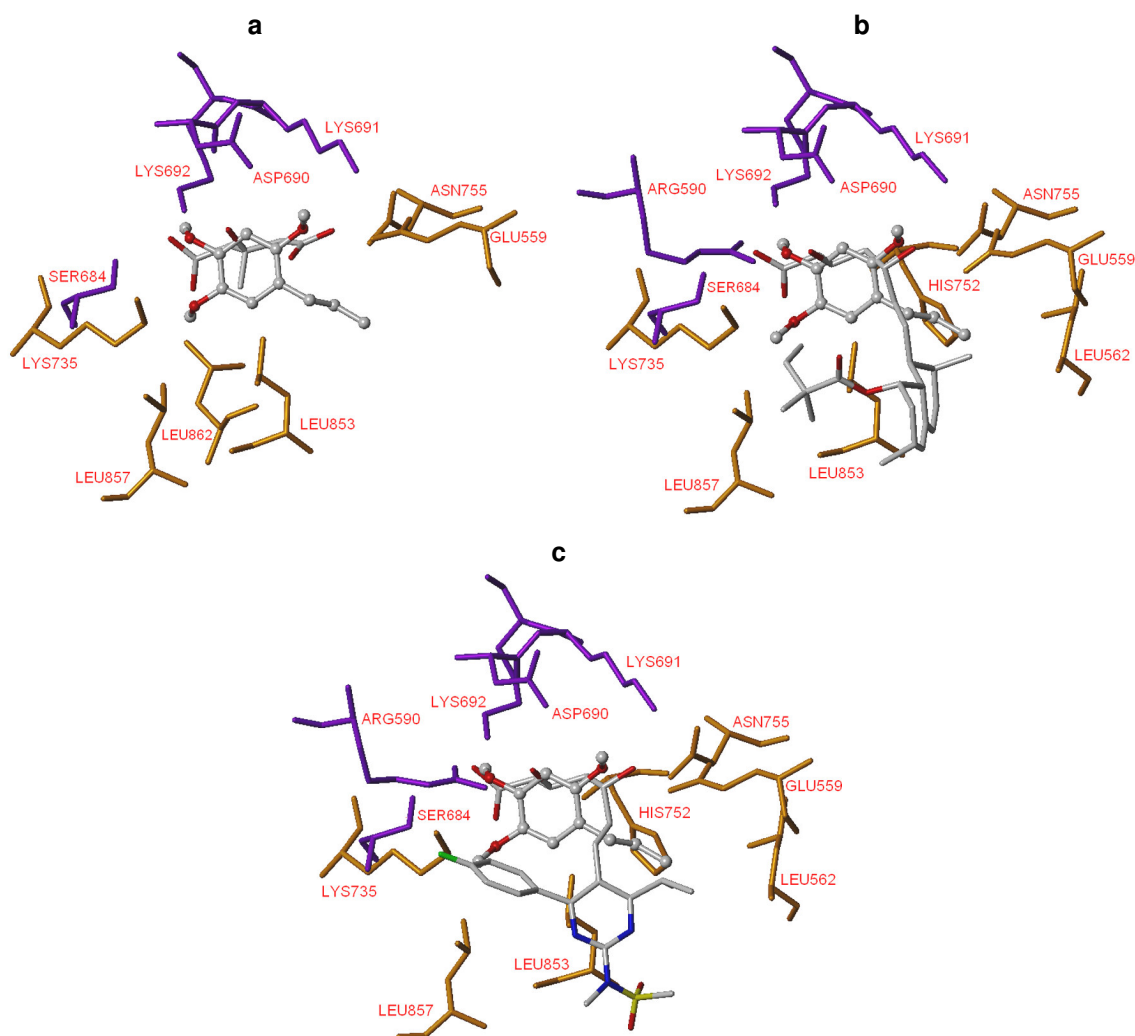


Figura 7.6. Comparación del modo de unión predicho para la α -asarona con el modo de unión cristalográfico de (a) el sustrato; (b) la simvastatina y (c) la rosuvastatina. Se muestran residuos representativos a 3.6 Å de distancia de la α -asarona. El color de los residuos es como en la Figura 7.5. Para mayor claridad, se omiten los hidrógenos.

Las observaciones anteriores sugieren que las interacciones de los átomos de oxígeno en C-1, C-2 y C-4 de la α -asarona son muy importantes en la unión de este compuesto con la HMGR. Estas interacciones tienen una función similar en la unión con la enzima que la subestructura tipo HMG presente en el sustrato y en las estatinas (Figura 7.7). Esto está de acuerdo con la observación de que el grupo trimetoxibenceno parece ser fundamental para la actividad hipocolesterolemia de análogos de α -asarona (Cruz, 2001). El modelo de acoplamiento también ayuda a visualizar, a nivel molecular, los resultados experimentales que indican que la cadena lateral de la α -asarona debería estar en posición *meta* y *para* de los grupos metoxilos C-1 y C-2, respectivamente (Cruz, 2003).

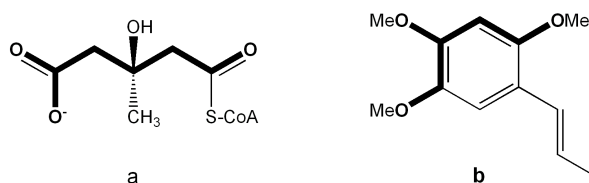


Figura 7.7. (a) Subestructura tipo HMG del sustrato (en negritas). (b) Fragmento de la estructura de α -asarona (en negritas) que juega un papel similar que la subestructura tipo HMG en la unión con la HMGR.

AutoDock encontró dos posibles conformaciones de unión para el grupo metoxilo C-1, dependiendo del sitio de unión. En el sitio de unión del sustrato, el grupo metoxilo está en posición relativa *cis* al grupo metoxilo vecino C-2 (ángulo diedro de 281.6°). Sin embargo, en el sitio de unión de las estatinas, la posición relativa es *trans* (ángulo diedro de 72.6°). La diferente conformación predicha para la conformación del grupo metoxilo en C-1 puede ser explicada por la presencia del residuo Leu862 en el sitio de unión del sustrato. Este residuo está ausente en la geometría de unión de las estatinas (Figura 7.8) (Istvan, 2001). De esta forma, en el sitio de unión del sustrato, el metoxilo en C-1 posiblemente tiene impedimento estérico con la cadena lateral de Leu862 y prefiere la posición *cis*. La ligera diferencia de la conformación de las cadenas laterales de los residuos Leu857 y Arg590 también puede contribuir a la diferente orientación del grupo metoxilo. A pesar de las diferentes conformaciones predichas para el grupo metoxilo en C-1, la predicción de la orientación del anillo de benceno de la α -asarona, las conformaciones de ambos grupos metoxilos C-2 y C-4, y la cadena propenilica, fueron las mismas en todas las geometrías de los sitios de unión (Figura 7.8).

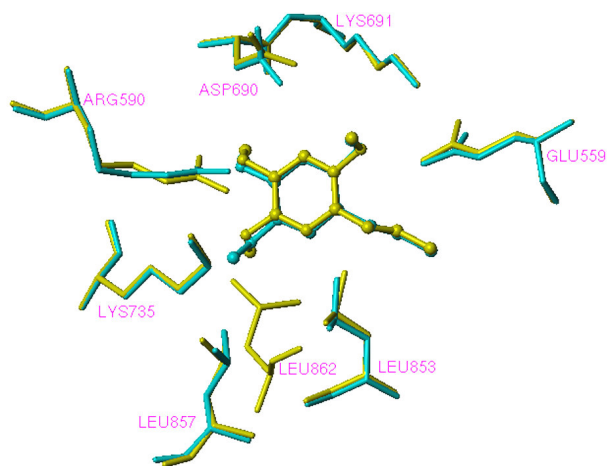


Figura 7.8. Conformación de unión predicha para el grupo metoxilo C-1 de la α -asarona en el modelo de acoplamiento con el sitio de unión del sustrato (amarillo) y en el sitio de unión de la rosuvastatina (azul). Se muestran residuos representativos. Para mayor claridad, se omiten los hidrógenos.

7.3. Diseño basado en la estructura de compuestos con potencial actividad inhibitoria de la 3-hidroxi-3-metilglutaril-CoA reductasa

El modelo de unión propuesto para la α -asarona sugiere modificaciones que podrían incrementar su afinidad por la HMGR y, posiblemente, su actividad hipocolesterolémica.

Una estrategia es reemplazar el grupo metoxilo C-4 por un grupo hidroxilo o amino para incrementar las interacciones polares con las cadenas laterales de residuos importantes en la catálisis (Lys691 y Glu559). Este tipo de grupos puede formar puentes de hidrógeno con estos residuos a semejanza del oxígeno en la posición 3 de HMG y las estatinas.

En el modelo de acoplamiento de la α -asarona se observa una cavidad, formada por los residuos del *cis-loop*, que tiene la posibilidad de ser ocupada por un grupo voluminoso. Esta cavidad puede ser ocupada por un grupo de mayor tamaño unido al oxígeno C-2 de la α -asarona que pudiera hacer interacciones polares con estos residuos. Resulta interesante observar que análogos fenoxiacéticos de α -asarona, por ejemplo el ácido 4-(2-propenil)-5-hidroxifenoxiacético (**1**) (Figura 7.9), tienen mayor actividad hipocolesterolémica *in vivo* que la α -asarona (Labarrios, 1999). Acoplamiento de análogos fenoxiacéticos de α -asarona con la HMGR indica que el sustituyente en C-2 ocupa la cavidad formada por el *cis-loop* (Figura 7.10). De hecho, la energía de interacción calculada para **1** es más favorable que para α -asarona (Tablas 7.3 y 7.4).

Ejemplo de un compuesto que considera las estrategias de modificar los sustituyentes en las posiciones C-2 y C-4 de α -asarona es el ácido 2-metoxi-4-(2-propenil)-5-hidroxifenoxiacético (**2**) (Figura 7.9). Se realizó el acoplamiento de este compuesto con la HMGR en el sitio de unión de la rosuvastatina. Se seleccionó a este sitio de unión porque, en la validación del protocolo de acoplamiento, se obtuvo el menor valor de desviación para la rosuvastatina (Tabla 7.1).

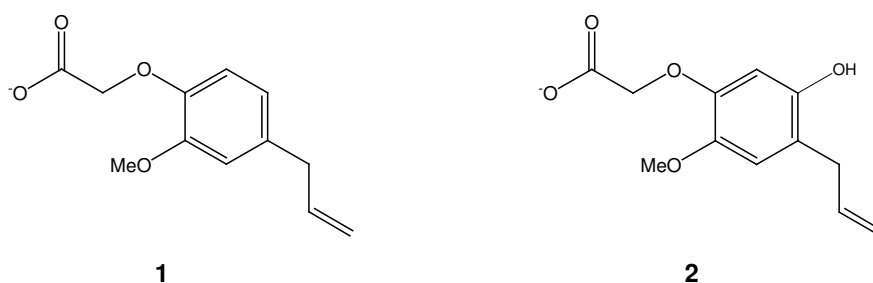


Figura 7.9. Estructuras de un análogo fenoxiacético de la α -asarona (**1**) y ejemplo de un nuevo análogo de α -asarona (**2**) con potencial actividad inhibitoria de la HMGR.

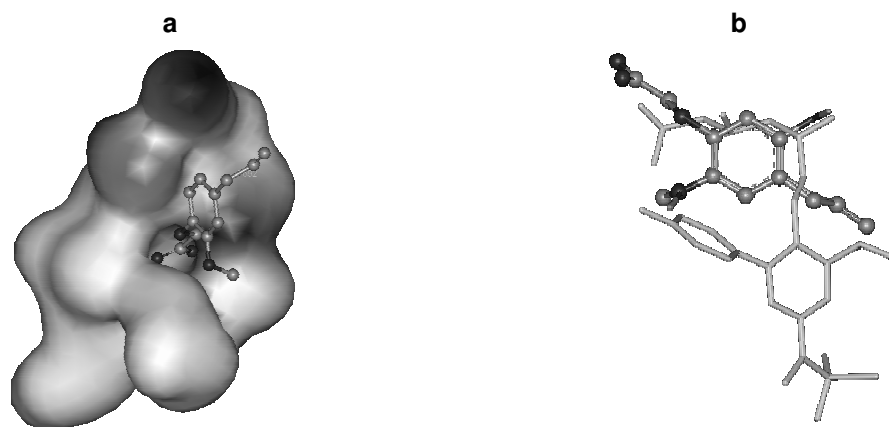


Figura 7.10. Acoplamiento de **1** con la HMGR. (a) Residuos del *cis-loop* que forman la cavidad ocupada por el carboxilato y (b) comparación del modo de unión con rosuvastatina (barras). Para mayor claridad, se omiten los hidrógenos.

La energía de acoplamiento calculada para **2** es más favorable que la energía calculada para α -asarona y que para **1** (Tablas 7.3 y 7.4). Esto sugiere que **2** puede tener mayor actividad inhibitoria de la HMGR que la α -asarona y, probablemente, mayor actividad hipocolesterolémica.

Tabla 7.4. Resultados del acoplamiento de **1** y **2** en el sitio de unión de rosuvastatina

Compuesto	Energía de acoplamiento (kcal/mol)	ΔG kcal/mol	K_i (a 298.15 K)
1	-10.24	-8.45	6.39×10^{-7}
2	-11.16	-9.45	1.17×10^{-7}

El modelo de unión generado con AutoDock para **2** está representado en la Figura 7.11. De acuerdo a este modelo, el grupo hidroxilo tiene interacciones polares con los residuos Lys691 y Glu559, y puede formar puentes de hidrógeno con Lys691 y Asn755. Por su parte, el grupo carboxilato puede formar una red extensa de puentes de hidrógeno con residuos del *cis-loop* (Figura 7.11).

Conclusiones

Se empleó acoplamiento molecular automatizado para desarrollar un modelo de unión de α -asarona con la HMGR y estudiar las interacciones fármaco-receptor correspondientes. Este trabajo constituye el primer estudio de acoplamiento molecular automático reportado con la HMGR que se tenga conocimiento. El protocolo de acoplamiento usado con AutoDock predijo

en forma aceptable el modo de unión de los ligandos cristalográficos, especialmente de simvastatina y rosuvastatina, a pesar de la gran cantidad de enlaces rotables de estas moléculas. Esto es notable ya que se demuestra que se puede aplicar AutoDock al estudio computacional de las estatinas y de la HMGR.

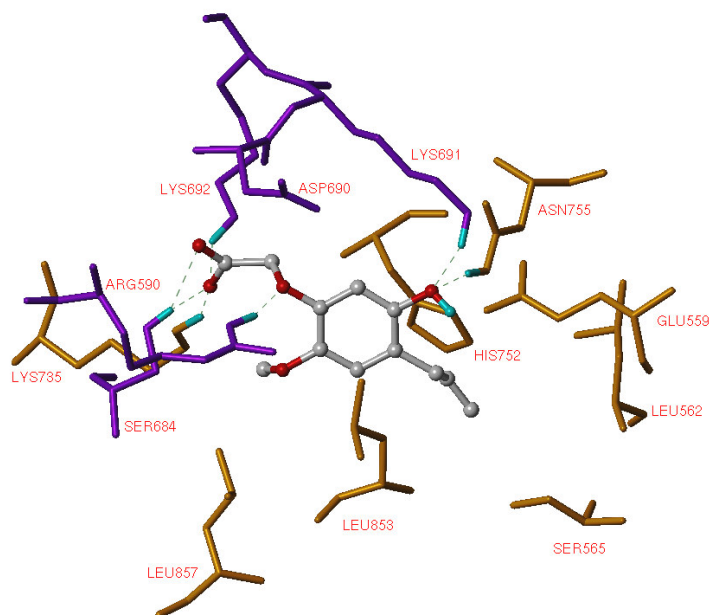


Figura 7.11. Modelo de unión de **2** con la HMGR. Se muestra la geometría del sitio de unión de la rosuvastatina. Los puentes de hidrógeno se muestran con líneas discontinuas verdes. Se muestran residuos representativos a 3.6 Å de distancia de la α -asarona. El color de los residuos es como en la Figura 7.5. Para mayor claridad, se omiten los hidrógenos excepto los involucrados en puentes de hidrógeno.

El estudio de acoplamiento de la α -asarona indica que esta molécula ocupa el mismo sitio de unión que HMG-CoA y las estatinas. Las energías de unión calculadas por AutoDock están de acuerdo con la actividad biológica relativa de α -asarona, simvastatina y rosuvastatina.

El modelo de acoplamiento, que está de acuerdo con estudios SAR reportados para análogos de α -asarona, indica que los átomos de oxígeno en C-1, C-2 y C-4 de la α -asarona participan en forma importante en la unión con la HGMR. Su función está relacionada con los oxígenos en posiciones 1 y 5 de HMG-CoA y de las estatinas. Es particularmente interesante la interacción del oxígeno en C-4 con los residuos catalíticos Lys691 y Glu559.

Siguiendo la estrategia general de diseño de fármacos basado en la estructura, se plantean estrategias para el diseño de análogos de α -asarona con potencial actividad inhibitoria de la HMGR. El diseño parte del modelo de unión desarrollado para la α -asarona.

Parte III

Compuestos antiparasitarios

Capítulo 8

Estudio QSAR-2D y predicción de la actividad de diversos compuestos contra *Giardia intestinalis*

8.1. Compuestos antiparasitarios contra *G. intestinalis*

La giardiasis, enfermedad intestinal causada por el protozooario *G. intestinalis*, causa la muerte de aproximadamente dos y medio millones de personas anualmente. Alrededor de doscientos ochenta millones de personas se infectan con este protozooario cada año a nivel mundial (IMSS, 2002).

El fármaco de elección contra este parásito es el metronidazol (Figura 8.1) (Mineno, 2003). Sin embargo, el metronidazol tiene asociados diversos efectos secundarios que pueden ser severos. A pesar de los avances en el estudio de la biología de *G. intestinalis*, el desarrollo de nuevos medicamentos es escaso. Adicionalmente, todos los fármacos generan resistencia (Mineno, 2003). Por este motivo, la búsqueda de nuevos compuestos activos contra *G. intestinalis* es primordial.

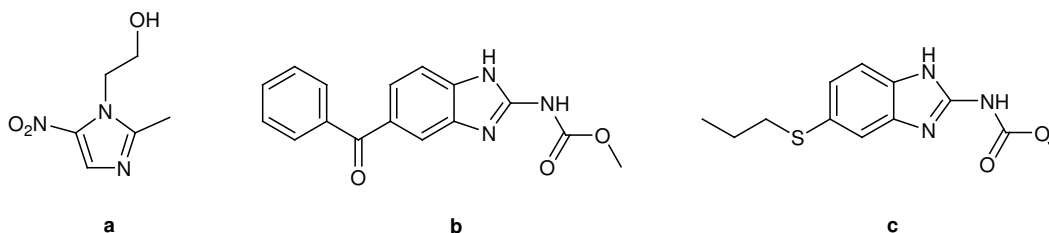


Figura 8.1. Estructura de (a) metronidazol, (b) mebendazol y (c) albendazol.

Uno de los compuestos de mayor interés con actividad giardicida corresponden a los derivados de bencimidazol. Se ha reportado que el mebendazol y albendazol (Figura 8.1) inhiben el crecimiento de *G. intestinalis* (Cedillo-Rivera, 1992; Chávez, 1992). El mecanismo de acción del albendazol está relacionado con la inhibición de la polimerización de la tubulina del citoesqueleto de *G. intestinalis* (Chávez, 1992). Se ha sugerido que un requerimiento estructural para este efecto es la presencia de un grupo metilcarbamato en la posición 2 y un

átomo de hidrógeno en la posición 1 en el anillo de bencimidazol (Sharma, 1994). Sin embargo, se ha encontrado que núcleos bencimidazólicos que no poseen estas características estructurales pueden ser más activos que el metronidazol y albendazol pero no interferir con la polimerización de la tubulina (Navarrete-Vázquez, 2001; Navarrete-Vázquez, 2003; Valdez, 2002). Para este segundo tipo de análogos de bencimidazol no se conoce su mecanismo de acción.

Con el propósito de encontrar los requerimientos básicos para la actividad giardicida de los derivados de bencimidazol, se han sintetizado y evaluado biológicamente a más de cien compuestos (por ejemplo, Andrzejewska, 2002; Andrzejewska, 2004; Navarrete-Vázquez, 2001; Navarrete-Vázquez, 2003; Navarrete-Vázquez, 2004; Sánchez-Díaz, 1999; Valdez, 2002). Recientemente, se han sintetizado y evaluado derivados del benzoxazol (Arroyo-Sánchez, 2003; Luna-González, 2004). Es especialmente interesante para la aplicación de estudios QSAR, que todas las pruebas biológicas de estos compuestos se han determinado en el mismo laboratorio utilizando la misma metodología.

Estrategia para el estudio computacional de compuestos contra *G. intestinalis*

La basta información de actividades biológicas determinadas para más de cien derivados de bencimidazol y benzoxazol contra *G. intestinalis* presenta la oportunidad de aplicar métodos computacionales, basados en la estructura del ligando, para el diseño de nuevos compuestos. La estrategia del presente trabajo consistió en realizar estudios QSAR-2D, utilizando el método *k*NN-QSAR, para generar modelos predictivos. Con estos modelos se efectuó una búsqueda automatizada en una base de datos pública de nuevos compuestos contra *G. intestinalis*. En la Figura 8.2 está esquematizada la estrategia.

La investigación en el diseño de derivados de bencimidazol como compuestos antiparasitarios continúa siendo muy activa. Recientemente se han sintetizado y evaluado biológicamente diversos derivados de bencimidazol (López-Balbiaux, 2005; Sandoval-Rivera). Los modelos QSAR desarrollados fueron utilizados para predecir la actividad biológica de estos compuestos antes de que fueran evaluados biológicamente. Así mismo, se predijo la actividad de compuestos que actualmente se están sintetizando (Méndez-Cuesta; Reza-Ramírez; Rivera-Pérez; Sandoval-Rivera; Vargas-Benítez), o que ya están sintetizados pero no se han evaluado (Arroyo-Sánchez, 2005; Luna-González, 2004) para identificar, de acuerdo a los modelos QSAR, qué compuestos tienen las mayores probabilidades de ser activos (Figura 8.2).

Debido a que no se conoce el blanco de acción de bencimidazoles que no posean un grupo carbamato en la posición 2 y un hidrógeno en la posición 1 del anillo de bencimidazol, no es posible aplicar acoplamiento molecular para el diseño de nuevos compuestos.

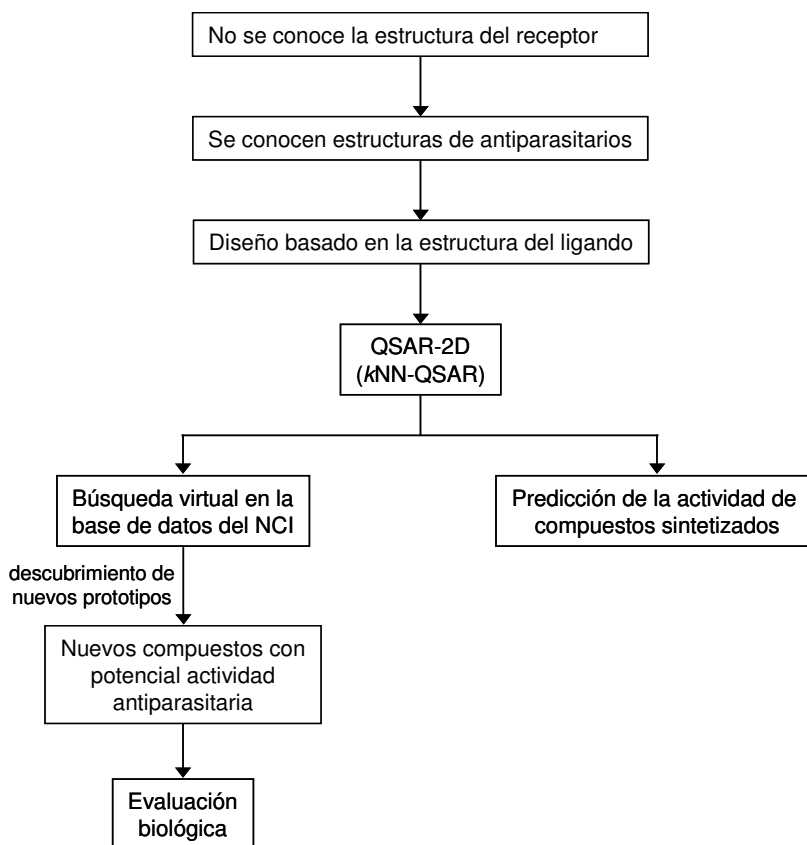


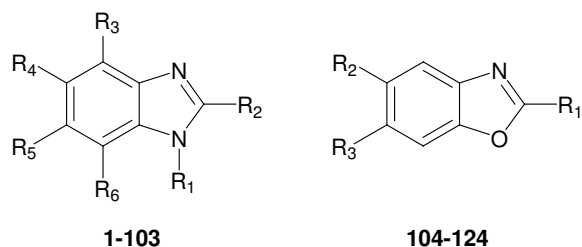
Figura 8.2. Estrategia para el diseño de compuestos con potencial actividad contra *G. intestinalis*.

8.2. Estudio QSAR-2D de análogos de bencimidazol y benzoxazol usando el método *k-Nearest Neighbors*

Las estructuras y actividad *in vitro* de 124 derivados de bencimidazol y benzoxazol usados en este estudio están descritas en la Tabla 8.1. Los derivados de bencimidazol no contienen un grupo metilcarbamato en la posición 2 y un átomo de hidrógeno en la posición 1. Las actividades fueron transformadas en los correspondientes valores $-\log IC_{50}$, donde IC_{50} es la concentración efectiva de compuesto requerida para alcanzar el 50% de inhibición del crecimiento de *G. intestinalis*. La actividad biológica fue tomada de la literatura (Arroyo-Sánchez, 2003; Andrzejewska, 2002; Andrzejewska, 2004; González-Aguilar, 2004; Luna-González, 2004; Navarrete-Vázquez, 2001; Navarrete-Vázquez, 2003; Sánchez-Díaz, 1999; Soriano-Agaton, 2002; Valdez, 2002; Vilchis-Reyes, 2004). En la Tabla 8.2 se listan las

estructuras de los compuestos sintetizados cuyas actividades biológicas fueron predichas por los modelos *k*NN-QSAR. Este grupo de estructuras sintetizadas es referido en este capítulo como un “grupo externo”.

Tabla 8.1. Estructuras y actividades biológicas de los compuestos empleados para el estudio *k*NN-QSAR



Comp.	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆	Exp. -LogIC ₅₀	Pred. -LogIC ₅₀
1	H	CH ₃	H	H	H	H	6.50	6.72
2	H	NH ₂	H	H	H	H	5.72	6.26
3	H	SH	H	H	H	H	7.40	7.12
4	H	SCH ₃	H	H	H	H	7.25	7.25
5	H	H	H	H	H	H	8.10	6.96
6	H	CH ₃	H	Cl	H	H	6.81	7.05
7	H	NH ₂	H	Cl	H	H	7.52	6.78
8	H	SH	H	Cl	H	H	7.09	7.25
9	H	SCH ₃	H	Cl	H	H	8.30	7.42
10	H	H	H	Cl	H	H	6.55	6.86
11	H	CH ₃	H	Cl	Cl	H	7.19	6.94
12	H	NH ₂	H	Cl	Cl	H	6.66	6.75
13	H	SH	H	Cl	Cl	H	7.11	7.14
14	H	SCH ₃	H	Cl	Cl	H	6.64	7.15
15	H	H	H	Cl	Cl	H	6.45	6.76
16	H	CF ₃	H	H	H	H	6.97	6.58
17	H	CF ₃	H	Cl	H	H	5.89	6.56
18	H	CF ₃	H	Cl	Cl	H	7.11	6.69
19	CH ₃	CF ₃	H	H	H	H	7.19	6.72
20	CH ₃	CF ₃	H	Cl	H	H	7.38	6.77
21	CH ₃	CF ₃	H	H	Cl	H	6.90	6.75
22	CH ₃	CF ₃	H	Cl	Cl	H	6.59	6.64
23	H	CF ₃	H	Br	H	H	6.19	6.55
24	H	CF ₃	H	Br	Br	H	6.92	6.76
25	H	CF ₃	Br	H	Br	H	6.60	6.71
26	H	CF ₃	Br	Br	Br	H	7.85	7.06
27	H	CF ₃	Br	Br	Br	Br	7.25	7.29
28	H	C ₂ F ₅	H	Cl	Cl	H	6.25	6.44

Tabla 8.1. (continuación)

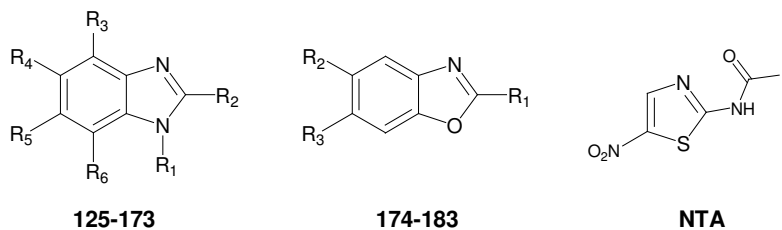
Comp.	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆	Exp. -LogIC ₅₀	Pred. -LogIC ₅₀
29	H	CF ₃	H	NO ₂	NO ₂	H	6.62	6.49
30	H	C ₂ F ₅	Br	Br	Br	Br	7.64	7.29
31	H	SC ₂ H ₄ N(CH ₃) ₂	Br	H	Br	H	7.60	7.49
32	CH ₃	CH ₃	H	H	H	H	7.05	7.03
33	CH ₃	NH ₂	H	H	H	H	5.87	6.52
34	CH ₃	SH	H	H	H	H	7.74	7.42
35	CH ₃	SCH ₃	H	H	H	H	7.48	7.40
36	CH ₃	CH ₃	H	Cl	H	H	6.84	7.06
37	CH ₃	NH ₂	H	Cl	H	H	6.62	6.84
38	CH ₃	SH	H	Cl	H	H	7.35	7.40
39	CH ₃	SCH ₃	H	Cl	H	H	7.55	7.37
40	CH ₃	CH ₃	H	H	Cl	H	7.36	7.10
41	CH ₃	NH ₂	H	H	Cl	H	6.88	6.89
42	CH ₃	SH	H	H	Cl	H	7.70	7.40
43	CH ₃	SCH ₃	H	H	Cl	H	6.91	7.35
44	CH ₃	CH ₃	H	Cl	Cl	H	7.26	7.23
45	CH ₃	NH ₂	H	Cl	Cl	H	7.13	6.96
46	CH ₃	SH	H	Cl	Cl	H	8.10	7.48
47	CH ₃	SCH ₃	H	Cl	Cl	H	7.62	7.33
48	CH ₃	SCH ₃	H	Cl	piperidil	H	6.91	6.22
49	CH ₃	SCH ₃	H	Cl	metilpiperidilcarbonil	H	7.00	6.24
50	CH ₃	SCH ₃	H	piperidil	Cl	H	5.15	6.21
51	CH ₃	SCH ₃	H	metilpiperidilcarbonil	Cl	H	5.37	6.19
52	CH ₃	SCH ₃	H	H	piperidil	H	6.59	6.40
53	CH ₃	SCH ₃	H	H	metilpiperidilcarbonil	H	7.27	6.43
54	CH ₃	CF ₃	H	H	propiltio	H	5.85	5.37
55	CH ₃	CF ₃	H	H	propilsulfonil	H	5.37	5.35
56	CH ₃	CF ₃	H	propiltio	H	H	4.98	5.35
57	CH ₃	CF ₃	H	propilsulfonil	H	H	4.68	5.34
58	CH ₃	CF ₃	H	benzoil	H	H	5.96	5.95
59	CH ₃	CF ₃	H	H	benzoil	H	5.89	5.93
60	H	SCH ₃	H	H	CONH ₂	H	7.15	7.13
61	H	SCH ₃	H	H	CONHCH ₃	H	7.74	7.35
62	H	SCH ₃	H	H	CON(CH ₃) ₂	H	7.14	7.13
63	H	SCH ₃	H	H	CONHC ₂ H ₅	H	7.07	7.08
64	H	SCH ₃	H	H	CON(C ₂ H ₅) ₂	H	6.46	6.96
65	H	SC ₂ H ₄ OH	Cl	H	Cl	H	7.42	7.32
66	H	SC ₂ H ₄ OH	Br	H	Br	H	6.98	7.33

Tabla 8.1. (continuación)

Comp.	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆	Exp. -LogIC ₅₀	Pred. -LogIC ₅₀
67	H	SC ₂ H ₄ N(CH ₃) ₂	Cl	H	Cl	H	7.74	7.44
68	H	SC ₂ H ₄ N(C ₂ H ₅) ₂	Cl	H	Cl	H	7.77	7.46
69	H	SC ₂ H ₄ N(C ₂ H ₅) ₂	Br	H	Br	H	7.50	7.47
70	H	SCH ₂ CH ₂ CH ₂ N(CH ₃) ₂	Cl	H	Cl	H	7.22	7.49
71	H	SCH ₂ CH ₂ CH ₂ N(CH ₃) ₂	Br	H	Br	H	7.76	7.52
72	H	SCH ₂ CH ₂ -(N-piperidil)	Cl	H	Cl	H	7.31	7.25
73	H	SCH ₂ CH ₂ -(N-piperidil)	Br	H	Br	H	6.93	7.25
74	H	SCH ₂ CH ₂ -(N-morfolinil)	Cl	H	Cl	H	7.45	7.29
75	H	SCH ₂ CH ₂ -(N-morfolinil)	Br	H	Br	H	7.26	7.26
76	H	SCH ₂ CH ₂ -(p-nitrofenil)	Cl	H	Cl	H	7.13	7.20
77	H	SCH ₂ CH ₂ -(p-nitrofenil)	Br	H	Br	H	7.46	7.26
78	H	CF ₃	H	CO ₂ H	H	H	5.52	5.88
79	H	CF ₃	H	CO ₂ CH ₃	H	H	5.92	6.24
80	H	CF ₃	H	CO ₂ CH ₂ CH ₃	H	H	5.39	5.85
81	H	CF ₃	H	COCH ₂ CON(C ₂ H ₅) ₂	H	H	5.51	5.71
82	H	CF ₃	H	CONC ₅ H ₁₀	H	H	5.22	5.72
83	H	CF ₃	H	propiltio	H	H	5.82	5.60
84	CH ₃	CF ₃	H	feniltio	H	H	6.03	5.33
85	CH ₃	CF ₃	H	fenilsulfinil	H	H	7.02	5.69
86	CH ₃	CF ₃	H	CF ₃	H	H	5.94	5.73
87	CH ₃	CF ₃	H	2,3-diclorofenoxi	H	H	7.52	6.85
88	CH ₃	CF ₃	H	H	feniltio	H	4.08	5.26
89	CH ₃	CF ₃	H	H	fenilsulfinil	H	4.72	5.59
90	H	CF ₃	H	propilsulfinil	H	H	5.20	5.42
91	CH ₃	CF ₃	H	H	CF ₃	H	5.05	5.71
92	CH ₃	CF ₃	H	H	2,3-diclorofenoxi	H	7.20	6.81
93	H	CF ₃	H	feniltio	H	H	5.80	5.48
94	H	CF ₃	H	fenilsulfinil	H	H	5.64	5.58
95	H	CF ₃	H	benzoil	H	H	5.64	5.85
96	H	CF ₃	H	CF ₃	H	H	6.17	5.78
97	H	CF ₃	H	2,3-diclorofenoxi	H	H	7.27	6.71
98	CH ₃	CF ₃	H	CO ₂ H	H	H	5.56	6.00
99	CH ₃	CF ₃	H	COCH ₂ CON(C ₂ H ₅) ₂	H	H	6.89	6.61
100	CH ₃	CF ₃	H	H	CO ₂ H	H	6.48	6.03
101	CH ₃	CF ₃	H	H	CO ₂ CH ₂ CON(C ₂ H ₅) ₂	H	6.33	6.57
102	CH ₃	CF ₃	H	CONHCH ₃	H	H	6.38	6.40
103	CH ₃	CF ₃	H	H	CONHCH ₃	H	6.68	6.42
104	H	H	H				5.12	5.30
105	CH ₃	H	H				5.58	5.25
106	NH ₂	H	H				4.53	5.17

Tabla 8.1. (continuación)

Comp.	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆	Exp. -LogIC ₅₀	Pred. -LogIC ₅₀
107	SH	H	H				5.11	5.29
108	CF ₃	H	H				5.43	5.98
109	SCH ₃	H	H				6.09	5.54
110	NHCO ₂ CH ₃	H	H				6.17	5.64
111	SH	Cl	H				4.99	5.41
112	SCH ₃	Cl	H				5.08	5.46
113	NH ₂	Cl	H				5.14	5.35
114	NHCO ₂ CH ₃	Cl	H				5.44	5.68
115	OCO ₂ CH ₃	Cl	H				4.44	5.12
116	H	Cl	Cl				4.94	5.16
117	CH ₃	Cl	Cl				5.17	5.41
118	CF ₃	Cl	Cl				6.02	6.19
119	SH	Cl	Cl				5.60	5.47
120	SCH ₃	Cl	Cl				5.44	5.48
121	NH ₂	Cl	Cl				5.74	5.45
122	NHCO ₂ CH ₃	Cl	Cl				6.45	6.04
123	CF ₃	CO ₂ CH ₃	H				6.74	6.60
124	CF ₃	H	CO ₂ CH ₃				7.22	6.62

Tabla 8.2. Estructuras y actividades biológicas predichas para los compuestos del grupo externo

Comp.	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆	Predicción -LogIC ₅₀
125	CH ₃	COCH ₃	H	Cl	H	H	7.22
126	CH ₃	CO ₂ H	H	Cl	H	H	7.20
127	CH ₃	CONH ₂	H	Cl	H	H	7.00
128	CH ₃	CONHCH ₃	H	Cl	H	H	7.15
129	CH ₃	CON(CH ₃) ₂	H	Cl	H	H	7.09
130	CH ₃	CO ₂ C ₂ H ₅	H	Cl	H	H	7.10
131	CH ₃	COCH ₃	H	H	H	H	7.13
132	CH ₃	CO ₂ H	H	H	H	H	7.12
133	CH ₃	CONH ₂	H	H	H	H	6.76
134	CH ₃	CONHCH ₃	H	H	H	H	7.27

Tabla 8.2. (continuación)

Comp.	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆	Predicción -LogIC ₅₀
135	CH ₃	CON(CH ₃) ₂	H	H	H	H	7.19
136	CH ₃	CO ₂ C ₂ H ₅	H	H	H	H	7.04
137	CH ₃	SCH ₃	H	CONH ₂	H	H	7.09
138	CH ₃	SCH ₃	H	CONHCH ₃	H	H	7.12
139	CH ₃	SCH ₃	H	CON(CH ₃) ₂	H	H	6.92
140	CH ₃	SCH ₃	H	CONHC ₂ H ₅	H	H	6.92
141	CH ₃	SCH ₃	H	CON(C ₂ H ₅) ₂	H	H	6.58
142	CH ₃	SCH ₃	H	H	CONH ₂	H	7.09
143	CH ₃	SCH ₃	H	H	CONHCH ₃	H	7.13
144	CH ₃	SCH ₃	H	H	CON(CH ₃) ₂	H	6.59
145	CH ₃	SCH ₃	H	H	CONHC ₂ H ₅	H	6.94
146	CH ₃	SCH ₃	H	H	CON(C ₂ H ₅) ₂	H	6.57
147	CH ₃	NHCO ₂ CH ₃	H	H	CONH ₂	H	NR
148	CH ₃	NHCO ₂ CH ₃	H	H	CONHCH ₃	H	NR
149	CH ₃	NHCO ₂ CH ₃	H	H	CON(CH ₃) ₂	H	NR
150	CH ₃	NHCO ₂ CH ₃	H	H	CONHC ₂ H ₅	H	NR
151	CH ₃	NHCO ₂ CH ₃	H	H	CON(C ₂ H ₅) ₂	H	NR
152	CH ₃	NHCO ₂ CH ₃	H	CONH ₂	H	H	NR
153	CH ₃	NHCO ₂ CH ₃	H	CONHCH ₃	H	H	NR
154	CH ₃	NHCO ₂ CH ₃	H	CON(CH ₃) ₂	H	H	NR
155	CH ₃	NHCO ₂ CH ₃	H	CONHC ₂ H ₅	H	H	NR
156	CH ₃	NHCO ₂ CH ₃	H	CON(C ₂ H ₅) ₂	H	H	NR
157	H	CF ₃	H	CONH ₂	H	H	5.93
158	H	CF ₃	H	CON(C ₂ H ₅) ₂	H	H	5.66
159	H	CF ₃	H	CONHC ₂ H ₅	H	H	5.86
160	H	CF ₃	H	CON(CH ₃) ₂	H	H	6.16
161	H	CF ₃	H	CONHCH ₃	H	H	6.23
162	H	CH ₃	H	NTA	H	H	NR
163	CH ₃	CH ₃	H	NTA	H	H	NR
164	CH ₃	CH ₃	H	H	NTA	H	NR
165	H	CF ₃	H	NTA	H	H	NR
166	H	SH	H	NTA	H	H	NR
167	H	SCH ₃	H	NTA	H	H	NR
168	CH ₃	CF ₃	H	NTA	H	H	NR
169	CH ₃	SH	H	NTA	H	H	NR
170	CH ₃	SCH ₃	H	NTA	H	H	NR
171	CH ₃	CF ₃	H	H	NTA	H	NR
172	CH ₃	SH	H	H	NTA	H	NR

Tabla 8.2. (continuación)

Comp.	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆	Predicción -LogIC ₅₀
173	CH ₃	SCH ₃	H	H	NTA	H	NR
174	OCO ₂ CH ₃	H	H				5.11
175	H	Cl	H				5.18
176	CH ₃	Cl	H				5.31
177	CF ₃	Cl	H				6.06
178	H	H	Cl				5.18
179	CH ₃	H	Cl				5.30
180	CF ₃	H	Cl				6.05
181	SH	H	Cl				5.39
182	SCH ₃	H	Cl				5.46
183	NHCOCH ₃	H	Cl				5.68

NR = predicción no confiable (ver texto para detalles)

En la Figura 8.3 se resumen los resultados de los modelos obtenidos. De un total de 4070 modelos QSAR generados, solamente 59 modelos cumplieron con el criterio de aceptación (Golbraikh, 2002b) (ver sección de Metodología computacional para detalles).

Estabilidad de los modelos QSAR

La estabilidad de los modelos QSAR se estableció con el procedimiento descrito en la sección de Metodología computacional. La Figura 8.4 muestra el gráfico de los valores de q^2 versus el número de descriptores para grupos de datos con actividades biológicas reales y aleatorias. Cada valor de q^2 es el promedio de diez cálculos independientes. En general, se obtuvieron valores de q^2 significativamente mayores cuando se emplearon valores de actividad biológica reales que cuando se emplearon actividades aleatorias (Figura 8.4).

El análisis estadístico de los resultados se realizó con la prueba de hipótesis de una-cola tal como se describe en la sección de Metodología computacional. En la Tabla 8.3 se muestran los valores de q^2 para varios modelos k NN-QSAR, con diferente número de descriptores. Estos valores se obtuvieron a partir de diez diferentes grupos con actividades biológicas aleatorias. En la misma tabla se muestran los valores de q^2 promedio, la desviación estándar de los valores q^2 , el valor Z para los modelos más significativos para el grupo con actividades reales. En todos los casos, el valor Z fue igual o superior a 4.84, indicando que la probabilidad que el modelo k NN-QSAR derivado con actividades biológicas reales sea aleatorio es de aproximadamente 10^{-5} .

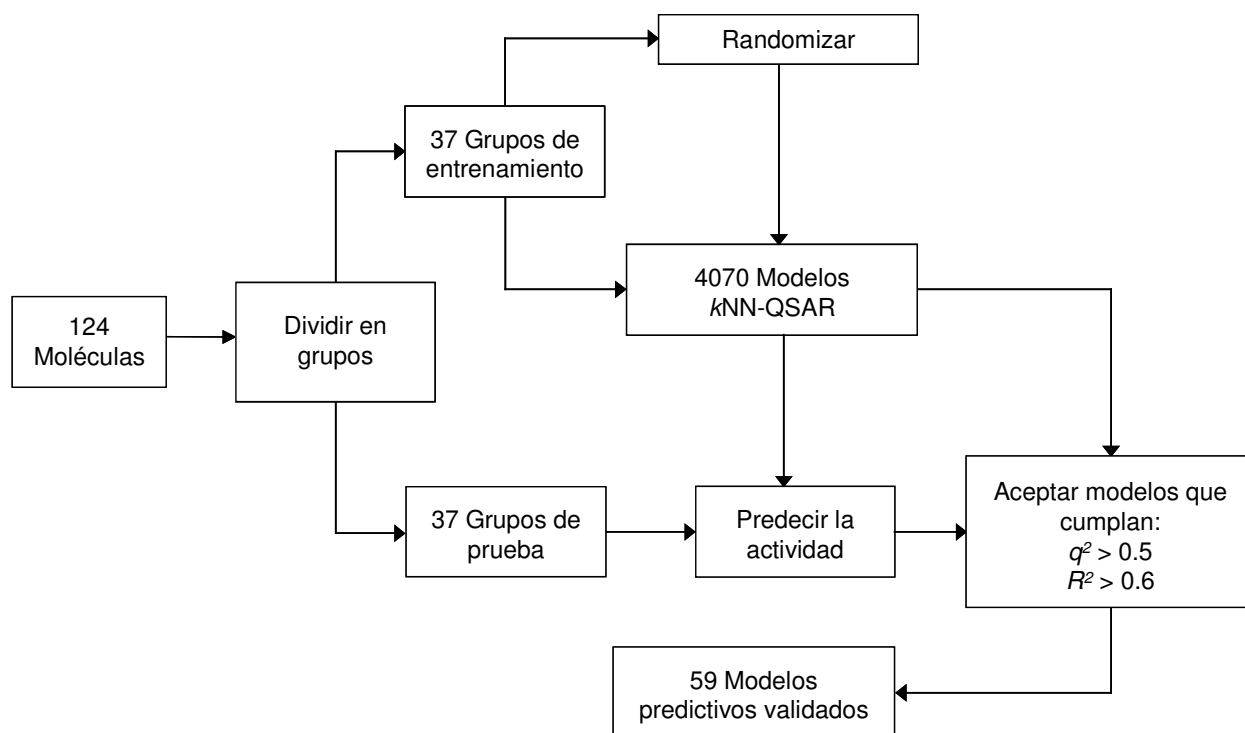


Figura 8.3. Diagrama de flujo de los modelos QSAR.

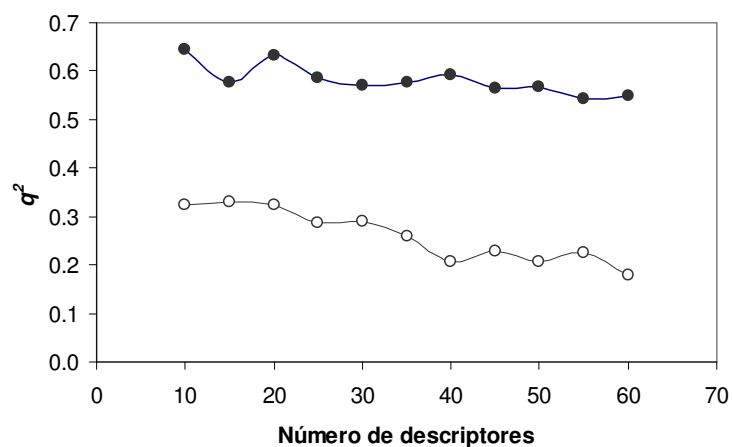


Figura 8.4. Gráficos de q^2 versus el número de descriptores seleccionados por los mejores modelos k NN-QSAR. Se muestran los resultados para el grupo de datos con actividades reales (círculos llenos) y aleatorias (círculos vacíos). Cada valor de q^2 es el promedio de diez cálculos independientes.

Tabla 8.3. Prueba de hipótesis de una cola para varios modelos QSAR

Número de descriptores	h (q^2 con actividades reales)	m (promedio de q^2 de 10 modelos aleatorios)	σ	Z
10	0.64	0.32	0.05	6.15
15	0.58	0.33	0.05	4.84
20	0.63	0.32	0.06	5.10
25	0.59	0.29	0.06	4.94
30	0.57	0.29	0.05	5.72
35	0.58	0.26	0.06	5.22
40	0.59	0.21	0.05	8.50
45	0.56	0.23	0.07	4.94
50	0.57	0.21	0.07	5.16
55	0.54	0.22	0.05	5.97
60	0.55	0.18	0.05	6.95

Validación de los modelos QSAR

Para obtener modelos QSAR predictivos y confiables, es necesario demostrar que los modelos generados para el grupo de entrenamiento pueden predecir con precisión las actividades de los compuestos en grupos de prueba (Tropsha, 2003b). Se aceptaron modelos con valores de q^2 , para el grupo de entrenamiento, superiores a 0.5 y con valores de R^2 , para los grupos de prueba, mayores a 0.6 (Golbraikh, 2002a; Golbraikh, 2002b). La Tabla 8.4 presenta a los 10 mejores modelos QSAR obtenidos de múltiples análisis k NN. El número de descriptores para los 10 mejores modelos varió entre 15 y 40. Aún para grupos de prueba tan grandes como de 54 compuestos, se obtuvo un modelo con $q^2 = 0.573$ y $R^2 = 0.666$.

Predicción de la actividad anti-giardia para los compuestos en el grupo de datos inicial

Con los 59 modelos k NN-QSAR se hizo la predicción de la actividad biológica de los 124 compuestos del grupo de datos inicial. El resultado del promedio de las predicciones para cada molécula está en la Tabla 8.1. En general, el valor de la predicción fue semejante al valor experimental. En casi todos los casos, la diferencia entre el valor predicho y el experimental fue menor a uno. Esto indica una capacidad predictiva de los modelos QSAR muy aceptable.

Predicción de la actividad anti-giardia para los compuestos en el grupo externo

Empleando a todos los modelos *k*NN-QSAR aceptables, se hizo la predicción de la actividad biológica de 49 derivados de bencimidazol y de 10 derivados de benzoxazol. Los resultados de las predicciones están listados en la Tabla 8.2. Los compuestos **125-136** tienen, en general, las actividades predichas más altas de todos los compuestos. Para el derivado de bencimidazol **125** se predice la actividad más alta de esta serie. Es interesante notar que los compuestos **131-136** fueron diseñados por un estudio QSAR-3D (Rodríguez-Morales).

Para los derivados de bencimidazol **137-145** se predican actividades biológicas elevadas. La alta actividad predicha para **142-145** puede explicarse por su similitud estructural con los correspondientes compuestos **60-63**, que también tienen una actividad elevada (Tabla 8.1). La diferencia entre estas estructuras es un grupo metilo en la posición 1 del anillo de bencimidazol.

Para los análogos de benzoxazol se predican, en general, actividades biológicas bajas. Esto puede explicarse por la baja actividad de los benzoxazoles **104-123** de la base de datos inicial (Tabla 8.1). Sin embargo, es interesante observar que para los benzoxazoles **177** y **180**, con un grupo trifluorometilo en la posición 2 del anillo de benzoxazol, se predican las actividades más favorables respecto a sus series análogas **175-177** y **178-183**, respectivamente (Tabla 8.2).

Tabla 8.4. Los diez mejores modelos *k*NN-QSAR

Modelo no.	Tamaño del grupo de entrenamiento	Tamaño del grupo de prueba	No. de descriptores	q^2	R^2
1	103	21	35	0.618	0.725
2	103	21	15	0.623	0.713
3	101	23	40	0.570	0.711
4	101	23	25	0.580	0.678
5	70	54	30	0.573	0.666
6	103	21	25	0.654	0.664
7	72	52	20	0.553	0.657
8	70	54	20	0.552	0.652
9	74	50	15	0.595	0.651
10	101	23	25	0.653	0.643

No fue posible predecir en forma confiable la actividad de 22 derivados de bencimidazol (Tabla 8.2). Esto se debe a que estos compuestos se encuentran afuera del dominio de

aplicación de los modelos *k*NN-QSAR, es decir, que no son similares a los compuestos en la base de datos inicial con 124 moléculas (de acuerdo al farmacóforo de descriptores). Ejemplos de moléculas para las cuales no se logró predecir su actividad son **162-173**. Estos compuestos tienen como sustituyente a un grupo 1,3-tiazol que no está presente en los compuestos de la base de datos inicial.

La comparación entre la actividad experimental y predicha contra *G. intestinalis* para ocho compuestos del grupo externo está en la Tabla 8.5. La actividad biológica se determinó siguiendo el mismo protocolo que para los compuestos del grupo de datos inicial (Cedillo-Rivera, 1992). Los detalles de la determinación de la evaluación biológica están descritos en la literatura, de donde se tomaron estos valores (López-Balbiaux, 2005; Sandoval-Rivera).

Tabla 8.5. Actividades experimentales y predichas para los compuestos en el grupo externo

Compuesto	Experimental -LogIC ₅₀ ^a	Predicción -LogIC ₅₀	Residuales ^b
131	6.76	7.13 ± 0.25	-0.37
137	6.60	7.09 ± 0.23	-0.49
138	6.65	7.12 ± 0.25	-0.47
139	6.97	6.92 ± 0.40	0.05
140	6.62	6.92 ± 0.27	-0.30
141	7.17	6.58 ± 0.38	0.59
142	6.96	7.09 ± 0.23	-0.13
143	6.60	7.13 ± 0.25	-0.53

^a(López-Balbiaux, 2005; Sandoval-Rivera). ^bPara el cálculo del residual se considera el valor promedio de la predicción de la actividad.

Para todos los compuestos en la Tabla 8.5, las predicciones son aceptables con residuales menores a 1 en unidades de potencia. Si se considera la desviación asociada a las predicciones, el valor de los residuales es aún inferior a los reportados en la Tabla 8.5 (que corresponden al promedio de la predicción de la actividad). Para el compuesto **139**, el promedio de la predicción fue casi igual a la determinada experimentalmente.

8.3. Búsqueda virtual de nuevos líderes en la base de datos del Instituto Nacional de Cáncer

Se emplearon **59** modelos QSAR para buscar nuevos líderes antiparasitarios contra *G. intestinalis* en la base de datos del NCI (Figura 8.5). El procedimiento de la búsqueda está descrito en la sección de Metodología computacional.

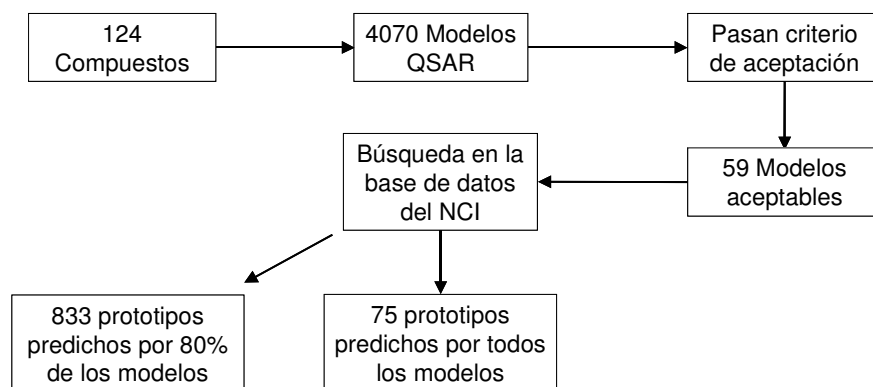


Figura 8.5. Diagrama de flujo de la búsqueda en la base de datos del NCI a partir de los modelos QSAR.

Se seleccionaron a todos los compuestos en la base de datos que se encuentran a un valor determinado de similitud (estancia Euclidiana de 0.5) de cualquiera de las moléculas activas. Debido a las diferencias en los farmacóforo de descriptores, las 59 listas resultantes no fueron iguales. De esta forma, las listas iniciales se refinaron seleccionando a aquellos compuestos que se encuentran en las 59 listas o en la mayoría de ellas. Esto condujo a 833 prototipos computacionales presentes en el 80% de las listas y a 75 prototipos presentes en todas las listas. La Figura 8.6 muestra el número completo de prototipos contenidos en diferentes números de listas.

Cinco compuestos, con alta actividad biológica esperada, fueron adquiridos comercialmente de ALDRICH. Los compuestos se evaluaron contra *G. intestinalis* por el grupo de la Dra. Lilián Yépez, en el Centro Médico Nacional Siglo XXI. Se utilizó el mismo protocolo de prueba con la que se determinó la actividad de los 124 compuestos de la base de datos inicial (Cedillo-Rivera, 1992). Los resultados están resumidos en la Tabla 8.6.

Para los compuestos **184-187** se obtuvo una desviación significativa entre los valores predichos y los experimentales (residuales mayores a uno en unidades de potencia). Solamente para el compuesto **188** se obtuvo un residual aceptable (menor a uno), aunque alto. Para los cinco compuestos se obtuvieron actividades bajas.

Es interesante observar que, dos de las actividades más bajas, se obtuvieron para compuestos que no son derivados de bencimidazol (**184** y **185**). Los valores altos de los residuales se pueden explicar, al menos en parte, porque los modelos QSAR se derivaron de un grupo de datos donde hay solamente derivados de bencimidazol y benzoxazol.

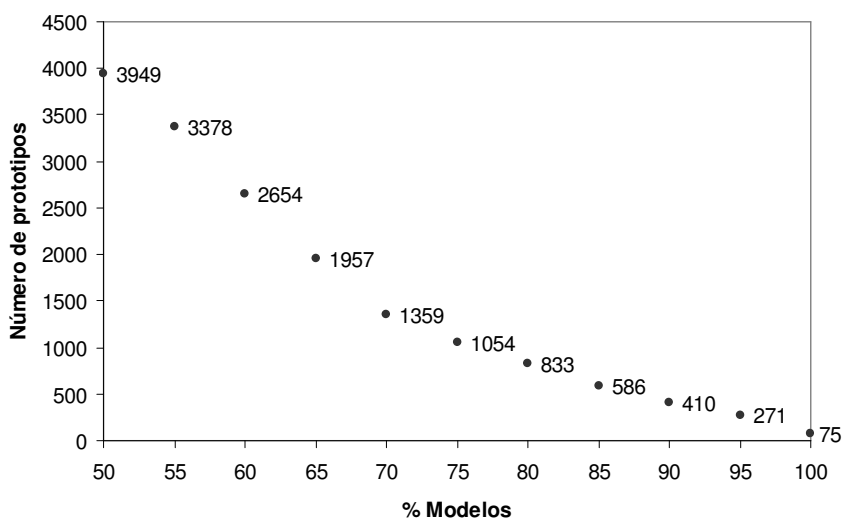
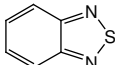
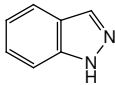
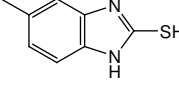
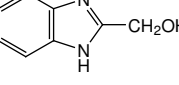
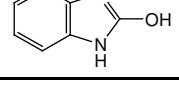


Figura 8.6. Número de prototipos computacionales predichos por diferente número de modelos QSAR.

Tabla 8.6. Prototipos computacionales seleccionados de la búsqueda en la base de datos del NCI

Compuesto	NSC	Estructura	Predicción -LogIC ₅₀	Experimental -LogIC ₅₀ ^a	Residuales ^b
184	43636 679		7.50-6.86	5.05	-2.13
185	26336 90357		7.36-6.55	2.94	-4.00
186	231400		7.49-6.98	3.49	-3.79
187	18284		7.23-6.47	5.29	-1.62
188	10383 178108		7.01-6.42	5.75	-0.98

^a Promedio de dos ensayos. ^b Se considera el valor promedio de la predicción.

La baja actividad biológica observada para **186**, que es un bencimidazol, no puede explicarse fácilmente con los presentes resultados. De hecho, **186** es parecido al compuesto **8** que tiene una buena actividad (Tabla 8.1). El valor alto del residual puede ser atribuido a la falta de compuestos en el grupo de datos inicial con un grupo metilo en las posiciones 5 y 6 del anillo de bencimidazol.

Las moléculas más activas de esta serie fueron **187** y **188**, ambos derivados de bencimidazol. Para estos compuestos se obtuvieron los residuales más bajos. También es difícil explicar la baja actividad de **188** ya que el compuesto **3**, que tiene un átomo de oxígeno en lugar de un átomo de azufre, tiene una buena actividad (Tabla 8.1).

La baja predictibilidad para las actividades de los compuestos de la base de datos del NCI sugiere que se necesitan modelos *k*NN-QSAR con valores de q^2 y R^2 superiores a los obtenidos con el grupo de datos de la Tabla 8.1. Es interesante observar que en el estudio *k*NN-QSAR aplicado a búsqueda en la base de datos del NCI para compuestos antivirales (capítulo 6) y anticonvulsantes (Shen, 2004), los valores de q^2 y R^2 son superiores a los valores de predictibilidad obtenidos en este estudio (Tabla 8.4).

Además de los prototipos computacionales reportados en la Tabla 8.6, se obtuvieron otras estructuras con alta actividad biológica esperada. Ejemplos de estas estructuras están listados en la Tabla 8.7. En la misma tabla se incluye la predicción de la actividad giardicida obtenida con los modelos *k*NN-QSAR. Estas estructuras representan compuestos antiparasitarios potenciales contra *G. intestinalis*.

Tabla 8.7. Predicción de la actividad giardicida de prototipos computacionales adicionales obtenidos de la búsqueda en la base de datos del NCI

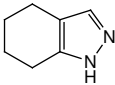
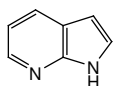
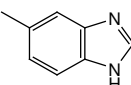
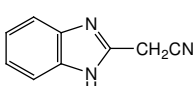
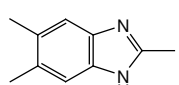
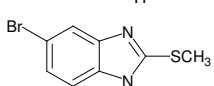
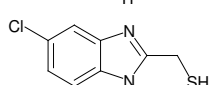
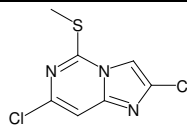
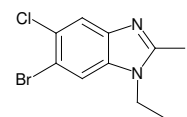
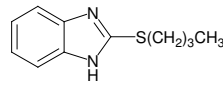
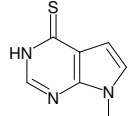
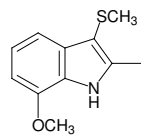
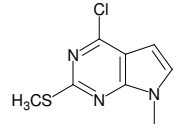
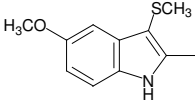
Compuesto	NSC	Estructura	Predicción -LogIC ₅₀
189	195333		7.35-6.61
190	77951		7.40-6.55
191	3826		7.22-6.56
192	525203		7.19-6.41
193	401569		7.24-6.76
194	523987		7.67-7.21
195	92672		7.58-7.10

Tabla 8.7. (continuación)

Compuesto	NSC	Estructura	Predicción -LogIC ₅₀
196	70905		7.54-7.06
197	141832		7.39-7.05
198	191914		7.59-7.15
199	102339		7.63-7.09
200	298289		7.70-7.04
201	126311		7.58-7.10
202	298287		7.62-7.04

Conclusiones

Se realizó un estudio QSAR-2D para compuestos antiparasitarios utilizando la técnica *k*NN-QSAR. La aplicación de esta técnica computacional basada en el ligando fue posible gracias al conocimiento de una basta serie de estructuras con actividad biológica conocida.

Los modelos QSAR fueron validados y se emplearon para predecir la actividad biológica de 59 compuestos ya sintetizados o que están en proceso de síntesis en el grupo del Dr. Rafael Castillo. Se validaron experimentalmente las actividades de ocho de estos compuestos. En todos los casos, la diferencia entre las actividades experimentales y predichas fue inferior a 0.6 en unidades de potencia. Resultados similares se obtuvieron cuando se hizo la predicción de la actividad biológica de los 124 derivados de bencimidazol y benzoxazol del grupo de datos inicial. Esto demuestra un grado aceptable de predictibilidad de los modelos *k*NN-QSAR generados.

Los modelos *k*NN-QSAR se emplearon para buscar, en aproximadamente 220,000 compuestos contenidos en la base de datos del NCI, compuestos activos contra *G. intestinalis*. Cinco compuestos fueron obtenidos comercialmente y evaluados contra el parásito por el grupo de la Dra. Lilián Yépez. Las predicciones no fueron precisas para estas moléculas. La falta de predictibilidad para algunos de estos compuestos puede atribuirse, entre otros factores, a la falta de información en el grupo inicial de datos. Por ejemplo, la ausencia de otros núcleos además del bencimidazol y benzoxazol, y a la falta de grupos metilo como sustituyentes en el anillo de bencimidazol.

Una perspectiva de este trabajo es evaluar la actividad biológica de más prototipos computacionales obtenidos de la búsqueda en la base de datos del NCI.

Otra perspectiva es extender el análisis *k*NN-QSAR a una base de datos más grande que incluya las actividades biológicas de las moléculas en el grupo externo y los prototipos computacionales obtenidos en la búsqueda en la base de datos del NCI. Esto aumentará el dominio de aplicabilidad de los modelos QSAR y se podrá predecir la actividad biológica de más moléculas. Por ejemplo, puede hacerse nuevamente la búsqueda en la base de datos del NCI.

En el caso de derivados de bencimidazol que no tienen un grupo metilcarbamato en la posición 2 y a un átomo de hidrógeno en la posición 1, no se conoce el blanco específico de acción. Por tanto, no pueden aplicarse métodos de diseño basado en la estructura del receptor para este grupo de compuestos.

Capítulo 9

Conclusiones

Las conclusiones y perspectivas específicas para cada tipo de compuestos están descritas al final de los capítulos 4-8. Conclusiones generales se presentan a continuación:

- Se aplicaron métodos computacionales de diseño de fármacos basados en la estructura del receptor y de diseño basado en la estructura de ligando, para el estudio y diseño de compuestos antivirales, hipocolesterolemiantes y antiparasitarios.
- El método de diseño basado en la estructura de receptor fue el acoplamiento molecular automático. Los métodos de diseño basado en la estructura de ligando fueron cuatro: CoMFA, CoMSIA (ambos QSAR-3D), *k*NN-QSAR (QSAR-2D) y búsqueda automatizada en bases de datos a partir de modelos QSAR-2D.
- Para los compuestos antivirales (derivados de piridinona, inhibidores de la TR de VIH-1) e hipocolesterolemiantes (α -asarona, inhibidor de la HMGR), el conocimiento de la estructura tridimensional de la macromolécula blanco permitió la aplicación de acoplamiento molecular. Hasta donde se sabe, el presente trabajo es el primer estudio de acoplamiento molecular automatizado que se realiza con la HMGR.
- En el caso de los inhibidores de la TR y compuestos antiparasitarios (derivados de bencimidazol y benzoxazol, contra el protozooario *G. intestinalis*), para los cuales hay una gran cantidad de información disponible de estudios SAR, se realizaron estudios QSAR-2D. En ambos casos, los modelos fueron utilizados para hacer una búsqueda de nuevos compuestos con potencial actividad antiviral y antiparasitaria, respectivamente, en una base de datos pública. Para los compuestos contra *G. intestinalis*, los modelos QSAR-2D se emplearon también para predecir la actividad de compuestos sintetizados y que fueron evaluados biológicamente después de las predicciones.
- Para los inhibidores de la TR fue posible hacer la combinación de técnicas de diseño basado en la estructura del ligando con una técnica de diseño basado en la estructura del receptor. De esta manera, se realizaron dos estudios QSAR-3D, CoMFA y CoMSIA,

aunados a estudios de acoplamiento molecular. La combinación de métodos permitió obtener mayor información sobre las interacciones ligando-receptor.

- Los estudios de acoplamiento molecular permitieron analizar las interacciones ligando-receptor de los derivados de piridinona con la TR y de la α -asarona con la HMGR. Este análisis ayudó a entender, a nivel molecular, la actividad antiviral e hipocolesterolemia, respectivamente, de este tipo de compuestos.
- El análisis de las interacciones ligando-receptor de los derivados de piridinona fue la base para el diseño de inhibidores potenciales de la TR, especialmente activos contra cepas resistentes. De esta manera, se hizo un diseño de nuevos análogos de piridinona a partir de la estructura de la TR con la finalidad de optimizar las propiedades biológicas de los análogos de piridinona como compuestos antivirales.
- El análisis de las interacciones ligando-receptor para la α -asarona constituyó la base para el diseño de análogos de este compuesto con potencial actividad inhibitoria de la HMGR. Este diseño, basado en la estructura de la HMGR, tiene como objetivo final optimizar la actividad de α -asarona como compuesto hipocolesterolemia.
- Los estudios QSAR-3D para los análogos de piridinona fueron especialmente valiosos para conocer el tipo de interacciones que ocurren entre estos compuestos y la TR. Las conclusiones derivadas de este estudio apoyaron la estrategia de diseño basado en la estructura para la optimización de las propiedades de los derivados de piridinona como compuestos antivirales.
- El estudio QSAR-2D para los derivados de piridinona, acoplada a la búsqueda automatizada en bases de datos, permitió proponer estructuras diferentes a los análogos de piridinona como inhibidores potenciales de la TR. De esta forma, a partir de la estructura del ligando, se lograron proponer nuevos líderes. La propuesta estuvo apoyada por cálculos de acoplamiento molecular.
- Los modelos QSAR-2D para los derivados de bencimidazol y benzoxazol, además de predecir satisfactoriamente la actividad giardicida de compuestos sintetizados, permitió identificar a estructuras contenidas en la base de datos del NCI, como posibles compuestos giardicidas. Para los compuestos de la base de datos las predicciones no correspondieron a los valores experimentales. Sin embargo, esta información aparentemente negativa, será valiosa para robustecer al método k NN-QSAR aplicado a compuestos antiparasitarios.

- Los métodos computacionales de diseño de fármacos empleados en este trabajo pueden ser aplicados a compuestos con otro tipo de actividades biológicas. La elección del método computacional dependerá de la información experimental disponible.
- El acoplamiento molecular con el programa AutoDock ayudó a justificar, a nivel molecular, el diseño de híbridos de piridinona que se ha hecho durante más de cinco años siguiendo una estrategia de sintetizar una gran cantidad de compuestos y efectuar las pruebas biológicas correspondientes. El caso presentado en este trabajo, muestra cómo el uso de un programa computacional, que es gratuito, puede simplificar el tiempo y costos del diseño de fármacos.
- La búsqueda en la base de datos del NCI, a partir de modelos *k*NN-QSAR, es un ejemplo de cómo puede realizarse en una institución académica, haciendo uso de una PC, estrategias que se utilizan en la industria farmacéutica. Esta técnica puede aplicarse a diversos sistemas en los que se cuente con una serie de estructuras químicas con actividad biológica conocida.
- Los métodos empleados en este trabajo son herramientas poderosas para analizar las interacciones ligando-receptor, ayudar a entender el modo de acción de moléculas activas y diseñar nuevos compuestos. Sin embargo, los sistemas biológicos son tan complejos que resulta prácticamente imposible modelarlos con precisión. De esta manera, la perspectiva general de este estudio computacional es que se realice la validación experimental de las predicciones que aún no han sido evaluadas. Los resultados experimentales, aunados a los cálculos computacionales, serán muy valiosos para la búsqueda de fármacos antivirales, hipocolesterolemiantes y antiparasitarios.

Referencias bibliográficas

- Andrzejewska, M.; Yepez-Mulia, L.; Cedillo-Rivera, R.; Tapia, A.; Vilpo, L.; Vilpo, J.; Kazimierczuk, Z. Synthesis, antiprotozoal and anticancer activity of substituted 2-trifluoromethyl and 2-pentafluoroethylbenzimidazoles. *Eur. J. Med. Chem.* **2002**, *37*, 973-978.
- Andrzejewska, M.; Yepez-Mulia, L.; Tapia, A.; Cedillo-Rivera, R.; Laudy, A. E.; Staroścaciak, B. J.; Kazimierczuk, Z. Synthesis, and antiprotozoal and antibacterial activities of S-substituted 4,6-dibromo- and 4,6-dichloro-2-mercaptobenzimidazoles. *Eur. J. Pharm. Sci.* **2004**, *21*, 323-329.
- Arroyo-Sánchez, V. Síntesis de benzoxazoles sustituidos en posición 2 con potencial actividad antiparasitaria. Tesis de Licenciatura. Facultad de Química, UNAM. México D.F. 2003.
- Arroyo-Sánchez, V. Síntesis de nitrocarboxamidas aromáticas. Tesis de Maestría. Facultad de Química, UNAM. México D.F. 2005.
- Artico, M.; Silvestri, R.; Pagnozzi, E.; Bruno, B.; Novellino, E.; Greco, G.; Massa, S.; Ettore, A.; Loi, A. G.; Scintu, F.; La Colla, P. Structure-based design, synthesis, and biological evaluation of novel pyrrolyl aryl sulfones: HIV-1 non-nucleoside reverse transcriptase inhibitors active at nanomolar concentrations. *J. Med. Chem.* **2000**, *43*, 1886-1891.
- Balbes, L. M.; Mascarella, S. W.; Boyd, D. B. A perspective of modern methods in computer-aided drug design. En *Reviews in Computational Chemistry*. Lipkowitz, K.; Boyd, D., Eds.; VCH Publishers, New York, 1994; Vol. 5, pp. 337-379.
- Baldwin, J. J.; Ponticello, G. S.; Anderson, P. S.; Christy, M. E.; Murcko, M. A.; Randall, W. C.; Schwam, H.; Sugrue, M. F.; Gautheron, P.; Grove, J.; Mallorga, P.; Viader, M-P.; McKeever, B. M.; Navia, M. A. Thienothiopyran-2-sulfonamides: novel topically active carbonic anhydrase inhibitors for the treatment of glaucoma. *J. Med. Chem.* **1989**, *32*, 2510-2513.
- Balzarini, J. Current status of the non-nucleoside reverse transcriptase inhibitors of human immunodeficiency virus type 1. *Curr. Top. Med. Chem.* **2004**, *4*, 921-944.
- Barreca, M. L.; Carotti, A.; Carrieri, A.; Chimiri, A.; Monforte, A. M.; Calace, M. P.; Rao, A. Comparative molecular field analysis (CoMFA) and docking studies of non-nucleoside HIV-1 RT inhibitors (NNIs). *Bioorg. Med. Chem.* **1999**, *7*, 2283-2292.
- Benjahad, A.; Courté, K.; Guillemont, J.; Mabire, D.; Coupa, S.; Poncelet, A.; Csoka, I.; Andries, K.; Pauwels, R.; De Béthune, M-P.; Monneret, C.; Bisagni, E.; Nguyen, C. H.; Grierson, D. S. 4-Benzyl- and 4-benzoyl-3-dimethylaminopyridin-2(1H)-ones, a new family of potent anti-HIV agents: optimization and in vitro evaluation against clinically important HIV mutant strains. *J. Med. Chem.* **2004**, *47*, 5501-5514.
- Boyd, D. B. Drug design. En *Encyclopedia of Computational Chemistry*. Schleyer, P. V. R., Ed.; John Wiley and Sons, West Sussex, UK, 1998; Vol. 1, pp 795-804.

- Boyd, D. B. Is rational design good for anything? En *Rational Drug Design: Novel Methodology and Practical Applications*. Parril, A. L., Reddy, M. R., Eds.; ACS Symposium Series 719; American Chemical Society, Washington, DC, 1999; pp 347-356.
- Brooijmans, N.; Kuntz, I. D. Molecular recognition and docking algorithms. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 335-373.
- Buckingham, J.; Glen, R. C.; Hill, A. P.; Hyde, R. M.; Martin, G. R.; Robertson, A. D.; Salmon, J. A.; Woollard, P. M. Computer-aided design and synthesis of 5-substituted tryptamines and their pharmacology at the 5-HT_{1D} receptor: discovery of compounds with potential anti-migraine properties. *J. Med. Chem.* **1995**, *38*, 3566-3580.
- Buolamwini, J. K.; Assefa, H. CoMFA and CoMSIA 3D QSAR and docking studies on conformationally-restrained cinnamoyl HIV-1 integrase inhibitors: Exploration of a binding mode at the active site. *J. Med. Chem.* **2002**, *45*, 841-852.
- Bursulaya, B. D.; Totrov, M.; Abagyan, R.; Brooks III, C. L. Comparative study of several algorithms for flexible ligand docking. *J. Comput.-Aided. Mol. Des.* **2003**, *17*, 755-763.
- Byrnes, V. W.; Sardana, V. V.; Schleif, W. A.; Condra, J. H.; Waterbury, J. A.; Wolfgang, J. A.; Long, W. J.; Schneider, C. L.; Schlabach, A. J.; Wolanski, B. S.; Graham, D. J.; Gotlib, L.; Rhodes, A.; Titus, D. L.; Roth, E.; Blahy, O. M.; Quintero, J. C.; Staszewski, S.; Emini, E. A. Comprehensive mutant enzyme and viral variant assessment of human immunodeficiency virus type 1 reverse transcriptase resistance to nonnucleoside inhibitors. *Antimicrob. Agents Chemother.* **1993**, *37*, 1576-1579.
- Cedillo-Rivera, R.; Muñoz, O. In-vitro susceptibility of *Giardia lamblia* to albendazole, mebendazole and other chemotherapeutic agents. *J. Med. Microbiol.* **1992**, *37*, 221-224.
- Chamorro, G.; Salazar, M.; Salazar, S.; Mendoza, T. Farmacología y toxicología de *Gutteria gaumeri*. *Rev. Inv. Clin.* **1993**, *45*, 597-604.
- Chan, C., Yin, H., Garforth, J., McKie, J. H., Jaouhari, R., Speers, P., Douglas, K. T., Rock, P. J., Yardley, V., Croft, S. L. and Fairlamb, A. H. Phenothiazine inhibitors of trypanothione reductase as potential antitrypanosomal and antileishmanial drugs. *J. Med. Chem.* **1998**, *41*, 4910-4910.
- Chan, J. H.; Hong, J. S.; Hunter III, R. N.; Orr, G. F.; Cowan, J. R.; Sherman, D. B.; Sparks, S. M.; Reitter, B. E.; Andrews III, C. W.; Hazen, R. J.; St Clair, M.; Boone, L. R.; Ferris, R. G.; Creech, K. L.; Roberts, G. B.; Short, S. A.; Weaver, K.; Ott, R. J.; Ren, J.; Hopkins, A.; Stuart, D. I.; Stammers, D. K. 2-Amino-6-arylsulfonylbenzonitriles as non-nucleoside reverse transcriptase inhibitors of HIV-1. *J. Med. Chem.* **2001**, *44*, 1866-1882.
- Charifson, P. S.; Kuntz, I. Recent successes and continuing limitations in computer-aided drug design. En *Application of Computer-Aided Drug Design*. Charifson, P. S. Ed.; Marcel Dekker, New York, NY, 1997; pp 1-37.
- Chávez, B.; Cedillo-Rivera, R.; Martínez-Palomo, A. *Giardia lamblia*: Ultrastructural study of the *in vitro* effect of benzimidazoles. *J. Protozool.* **1992**, *39*, 510-515.
- Chen, H. F.; Yao, X. J.; Li, Q.; Yuan, S. G.; Panaye, A.; Doucet, J. P.; Fan, B. T. Comparative study of non nucleoside inhibitors with HIV-1 reverse transcriptase based on 3D-QSAR and docking. *SAR and QSAR in Environmental Research* **2003**, *14*, 455-474.

- Cheng, C. C.; Robins, R. K. Potential purine antagonists. VI. Synthesis of 1-alkyl- and 1-aryl-4-substituted pyrazolo[3,4-d]pyrimidines. *J. Org. Chem.* **1956**, *21*, 1240-1256.
- Cheng, J. W. M. Rosuvastatin in the management of hyperlipidemia. *Clinical therapeutics* **2004**, *26*, 1368-1387.
- Corbett, J. W.; Kresge, K. J.; Pan, S.; Cordova, B. C.; Klabe, R. M.; Rodgers, J. D.; Erickson-Viitanen, S. K. Trifluoromethyl containing 3-alkoxymethyl- and 3-aryloxymethyl-2-pyridinones are potent inhibitors of HIV-1 non-nucleoside reverse transcriptase. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 309-312.
- Corsini, A.; Maggi, F. M.; Catapano, A. L. Pharmacology of competitive inhibitors of HMG-CoA reductase. *Pharmacol. Res.* **1995**, *31*, 9-27.
- Cramer, R. D. III.; Patterson, D. E.; Bunce, J. D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959-5967.
- Cruz, A.; Garduño, L.; Salazar, M.; Martínez, E.; Jiménez-Vázquez, H. A.; Díaz, F.; Chamorro, G.; Tamariz, J. Synthesis and hypolipidemic activity of modified side chain α -asarone homologues. *Arzneimittel-Forsch.* **2001**, *51*, 535-544.
- Cruz, MC.; Salazar, M.; Garciafigueroa, Y.; Hernández, D.; Díaz, F.; Chamorro, G.; Tamariz, J. Hypolipidemic activity of new phenoxyacetic derivatives related to α -asarone with minimal pharmacophore features. *Drug Dev. Res.* **2003**, *60*, 186-195.
- Das, K.; Clark, Jr. A. D.; Lewi, P. J.; Heeres, J.; De Jonge, M. R.; Koymans, L. M. H.; Vinkers, H. M.; Daeyaert, F.; Ludovici, D. W.; Kukla, M. J.; De Corte, B.; Kavash, R. W.; Ho, C. Y.; Ye, H.; Lichtenstein, M. A.; Andries, K.; Pauwels, R.; De Béthune M.-P.; Boyer, P. L.; Clark, P.; Hughes, S. H.; Janssen, P. A. J.; Arnold, E. Roles of conformational and positional adaptability in structure-based design of TMC125-R165335 (etravirine) and related non-nucleoside reverse transcriptase inhibitors that are highly potent and effective against wild-type and drug-resistant HIV-1 variants. *J. Med. Chem.* **2004**, *47*, 2550-2560.
- Davey, R. T.; Dewar, R. L.; Reed, G. F.; Vasudevachari, M. B.; Polis, M. A.; Kovacs, J. A.; Fallon, J.; Walker, R. E.; Masur, H.; Haneiwich, S. E.; O'Neill, D. G.; Decker, M. R.; Metcalf, J. A.; Deloria, M. A.; Laskin, O. L.; Salzman, N.; Lane, H. C. Plasma viremia as a sensitive indicator of the antiretroviral activity of L-697,661. *Proc. Natl. Acad. Sci. USA.* **1993**, *90*, 5608-5612.
- Debnath, A. K. Quantitative structure-activity relationship (QSAR): a versatile tool in drug design. En *Combinatorial Library Design and Evaluation for Drug Discovery: Principles, Methods, Software Tools and Applications*. Ghose, A. K.; Viswanadhan, V. N., Eds.; Marcel Dekker, New York, 2001; Chapter 3, pp 73-129.
- De Clercq, E. The role of non-nucleoside reverse transcriptase inhibitors (NNRTIs) in the therapy of HIV-1 infection. *Antivir. Res.* **1998**, *38*, 153-179.
- De Clercq, E. Strategies in the design of antiviral drugs. *Nat. Rev. Drug Discov.* **2002**, *1*, 13-25.
- De Clercq, E. HIV-chemotherapy and -prophylaxis: new drugs, leads and approaches. *Int. J. Biochem. Cell Biol.* **2004**, *36*, 1800-1822.

- Dewar, M. J. S.; Zoebisch, E. G., Healy, E. F. and 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.
- Ding, J.; Das, K.; Hsiou, Y.; Zhang, W.; Arnold, E. Structural studies of HIV-1 reverse transcriptase and implications for drug design. En *Structure-based Drug Design*. Veerapandian, P., Ed.; Marcel Dekker, New York, NY, 1997; pp 41-82.
- Dollé, V.; Fan, E.; Nguyen, C. H.; Aubertin, A. M.; Kirn, A.; Andreola, M. L.; Jamieson, G.; Tarrago-Litvak, L.; Bisagni, E. A New series of pyridinone derivatives as potent non-nucleoside human immunodeficiency virus type 1 specific reverse transcriptase inhibitors. *J. Med. Chem.* **1995**, *38*, 4679-4686.
- Dollé, V.; Nguyen, C. H.; Legraverend, M.; Aubertin, A. M.; Kirn, A.; Andreola, M. L.; Ventura, M.; Tarrago-Litvak, L.; Bisagni, E. Synthesis and antiviral activity of 4-benzyl pyridinone derivatives as potent and selective non-nucleoside human immunodeficiency virus type 1 reverse transcriptase inhibitors. *J. Med. Chem.* **2000**, *43*, 3949-3962.
- Dorsey, B. D.; Levin, R. B.; McDaniel, S. L.; Vacca, J. P.; Guare, J. P.; Darke, P. L.; Zugay, J. A.; Emini, E. A.; Schleif, W. A.; Quintero, J. C.; Lin, J. H.; Chen, I.W.; Holloway, M. K.; Fitzgerald, P. M. D.; Axel, m. G.; Ostovic, D.; Anderson, P. S.; Huff, J. R. L-735,524: The design of a potent and orally bioavailable HIV protease inhibitor. *J. Med. Chem.* **1994**, *37*, 3443-3451.
- Douali, L. ; Villemin, D.; Cherqaoui, D. Neural networks: Accurate nonlinear QSAR model for HEPT derivatives. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1200-1207.
- Duncia, J. V.; Chiu, A. T.; Carini, D. J.; Gregory, G. B.; Jonson, A. L.; Price, W. A.; Wells, G. J.; Wong, P. C.; Calabrese, J. C.; Timmermans, P. B. M. W. M. The discovery of potent nonpeptide angiotensin II receptor antagonists: a new class of potent antihypertensives. *J. Med. Chem.* **1990**, *33*, 1312-1329.
- Enríquez, R. G.; Chávez, M. A.; Jaúregui, F. Propenyl benzenes from *Guatteria gaumeri*. *Phytochemistry* **1980**, *19*, 2024-2025.
- Esnouf, R.; Ren, J.; Ross, C.; Jones, Y.; Stammers, D.; Stuart, D. Mechanism of inhibition of HIV-1 reverse transcriptase by non-nucleoside inhibitors. *Nat. Struct. Biol.* **1995**, *2*, 303-308.
- Esnouf, R. M.; Ren, J.; Hopkins, A. L.; Ross, C. K.; Jones, E. Y.; Stammers, D. K.; Stuart, D. I. Unique features in the structure of the complex between HIV-1 reverse transcriptase and the bis(heteroaryl)piperazine (BHAP) U-90152 explain resistance mutations for this nonnucleoside inhibitor. *Proc. Natl. Acad. Sci. USA.* **1997**, *94*, 3984-3989.
- Fujita, T.; Iwasa, J.; Hansch, C. A new substituent constant, π , derived from partition coefficients. *J. Am. Chem. Soc.* **1964**, *86*, 5175-5180.
- Fujita, T. The extrathermodynamic approach to drug design. En *Comprehensive Medicinal Chemistry*. Hansch, C.; Sammes, P. G.; Taylor, J. B.; Ramsden, C. A., Eds.; Pergamon Press, Oxford, 1990; Vol. 4, pp 497-560.
- Garduño, L.; Salazar, M.; Salazar, S.; Morelos, M. E.; Labarrios, F.; Tamariz, J.; Chamorro, G. A. Hypolipidemic activity of α -asarone in mice. *J. Ethnopharmacol.* **1997**, *55*, 161-163.

- Garg, R.; Gupta, S. P.; Gao, H.; Babu, M. S.; Debnath, A. K.; Hansch, C. Comparative quantitative structure-activity relationship studies on anti-HIV drugs. *Chem. Rev.* **1999**, *99*, 3525-3602.
- Gilbert, N. *Statistics*, W. B. Saunders, Co., Philadelphia, PA, 1976.
- Gohlke, H.; Klebe, G. Statistical potentials and scoring functions applied to protein-ligand binding. *Curr. Opin. Struc. Biol.* **2001**, *11*, 231-235.
- Gohlke, H.; Klebe, G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angew. Chem. Int. Ed.* **2002**, *41*, 2644-2676.
- (a) Golbraikh, A.; Tropsha, A. Beware of q^2 ! *J. Mol. Graphics Model.* **2002**, *20*, 269-276.
- (b) Golbraikh, A.; Tropsha, A. Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. *J. Comput.-Aided. Mol. Des.* **2002**, *16*, 357-369.
- Goldman, M. E.; Nunberg, J. H.; O'Brien, J. A.; Quintero, J. C.; Schleif, W. A.; Freund, K. F.; Gaul, S. L.; Saari, W. S.; Wai, J. S.; Hoffman, J. M.; Anderson, P. S.; Hupe, D. J.; Emini, E. A.; Stern, A. M. Pyridinone derivatives: Specific human immunodeficiency virus type 1 reverse transcriptase inhibitors with antiviral activity. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 6863-6867.
- González-Aguilar, J. Síntesis y determinación de la actividad antiparasitaria de derivados diclorados del benzoxazol. Tesis de Maestría. Facultad de Química, UNAM. México D.F. 2004.
- Goodford, P. J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* **1985**, *28*, 849-857.
- Greco, G.; Novellino, E.; Martin, I. C. Approaches to three-dimensional quantitative structure-activity relationships. En *Reviews in Computational Chemistry*. Lipkowitz, K.; Boyd, D., Eds.; VCH Publishers, New York, 1997; Vol. 11, pp. 183-240.
- Greer, J.; Erickson, J. W.; Baldwin, J. L.; Varney, M. D. Application of the three-dimensional structures of protein target molecules in structure-based drug design. *J. Med. Chem.* **1994**, *37*, 1035-1054.
- Gupta, S. P. In *Progress in Drug Research*; Jucker, E., Ed.; Birkhäuser Verlag: Basel, Switzerland, 2002; Vol. 58, pp 252-253.
- Hannongbua, S., Pungpo, P., Limtrakul, J.; Wolschann, P. Quantitative structure-activity relationships and comparative molecular field analysis of TIBO derivatised HIV-1 reverse transcriptase inhibitors. *J. Comput.-Aided. Mol. Des.* **1999**, *13*, 563-577.
- Hannongbua, S.; Nivesanond, K.; Lawtrakul, L.; Pungpo, P.; Wolschann, P. 3D-Quantitative structure-activity relationships of HEPT derivatives as HIV-1 reverse transcriptase inhibitors, based on ab initio calculations. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 848-855.
- Hansch, C.; Muir, R. M.; Fujita, T.; Maloney, P. P.; Geiger, E.; Streich, M. The correlation of biological activity of plant growth regulators and chloromycetin derivatives with Hammett constants and partition coefficients. *J. Am. Chem. Soc.* **1963**, *85*, 2817-2824.

- Hansch, C.; Fujita, T. ρ - σ - π Analysis; method for the correlation of biological activity and chemical structure. *J. Am. Chem. Soc.* **1964**, *85*, 1616-26.
- Hansch, C.; Leo, A. Exploring QSAR-fundamentals and applications in chemistry and biology. American Chemical Society, Washington, DC, 1995.
- Hetényi, C.; Van der Spoel, D. Efficient docking of peptides to proteins without prior knowledge of the binding site. *Protein Science* **2002**, *11*, 1729-1737.
- Hillisch, A.; Pineda, L. F.; Hilgenfeld, R. Utility of homology models in the drug discovery process. *Drug Discov. Today* **2004**, *9*, 659-669.
- Hoffman, J. M.; Wai, J. S.; Thomas, C. M.; Levin, R. B.; O'Brien, J. A.; Goldman, M. E. Synthesis and evaluation of 2-pyridinone derivatives as HIV-1 specific reverse transcriptase inhibitors. 1. Phthalimidoalkyl and -alkylamino analogs. *J. Med. Chem.* **1992**, *35*, 3784-3791.
- Hoffman, J. M.; Smith, A. M.; Rooney, C. S.; Fisher, T. E.; Wai, J. S.; Thomas, C. M.; Bamberger, D.; Barnes, J. L.; Williams, T. M.; Jones, J. H.; Olson, B. D.; O'Brien, J. A.; Goldman, M. E.; Nunberg, J. H.; Quintero, J. C.; Schleif, W. A.; Emini, E. A.; Anderson, P. S. Synthesis and evaluation of 2-pyridinone derivatives as HIV-1-specific reverse transcriptase inhibitors. 4. 3-[2-(Benzoxazol-2-yl)ethyl]-5-ethyl-6-methylpyridin-2(1*H*)-one and analogs. *J. Med. Chem.* **1993**, *36*, 953-966.
- Högberg, M.; Sahlberg, C.; Engelhardt, P.; Noreen, R.; Kangasmetsa, J.; Johansson, N. G.; Öberg, B.; Vrang, L.; Zhang, H.; Sahlberg, B. L.; Unge, T.; Lovgren, S.; Fridborg, K.; Backbro, K. Synthesis and further structure-activity relationship studies of PETT analogues. *J. Med. Chem.* **2000**, *43*, 304-304.
- Holdgate, G. A.; Ward, W. H. J.; McTaggart, F. Molecular mechanism for inhibition of 3-hydroxy-3-methylglutaryl CoA (HMG-CoA) reductase by rosuvastatin. *Biochem. Soc. Trans.* **2003**, *31*, 528-531.
- Hopfinger, A. J. Computer-assisted drug design. *J. Med. Chem.* **1985**, *28*, 1133-1139.
- Hopkins, A. L.; Ren, J.; Esnouf, R. M.; Willcox, B. E.; Jones, E. Y.; Ross, C.; Miyasaka, T.; Walker, R. T.; Tanaka, H.; Stammers, D. K.; Stuart, D. I. Complexes of HIV-1 reverse transcriptase with inhibitors of the HEPT series reveal conformational changes relevant to the design of potent non-nucleoside inhibitors. *J. Med. Chem.* **1996**, *39*, 1589-1600.
- Hopkins, A. L.; Ren, J.; Tanaka, H.; Baba, M.; Okamoto, M.; Stuart, D. I.; Stammers, D. K. Design of MKC-442 (emivirine) analogues with improved activity against drug-resistant HIV mutants. *J. Med. Chem.* **1999**, *42*, 4500-4505.
- Hsiou, Y.; Das, K.; Ding, J.; Clark Jr., A. D.; Kleim, J. P.; Rosner, M.; Winkler, I.; Riess, G.; Hughes, S. H.; Arnold, E. Structures of Tyr188Leu mutant and wild-type HIV-1 reverse transcriptase complexed with the non-nucleoside inhibitor HBV 097: inhibitor flexibility is a useful design feature for reducing drug resistance. *J. Mol. Biol.* **1998**, *284*, 313-323.
- IMSS. Boletín de vigilancia epidemiológica, 2002. <http://www.ssa.gob.mx>
- (a) Istvan, E. S.; Palnitkar, M.; Buchanan, S. K.; Deisenhofer, J. Crystal structure of the catalytic portion of human HMG-CoA reductase: insights into regulation of activity and catalysis. *EMBO J.* **2000**, *19*, 819-830.

- (b) Istvan, E. S.; Deisenhofer, J. The structure of the catalytic portion of human HMG-CoA reductase. *Biochim. Biophys. Acta* **2000**, *1529*, 9-18.
- Istvan, E. S.; Deisenhofer, J. Structural mechanism for statin inhibition of HMG-CoA reductase. *Science* **2001**, *292*, 1160-1164.
- Kaldor, S.W.; Kalish, V. J.; Davies, J. F.; Shetty, B. V.; Fritz, J. E.; Appelt, K.; Burgess, J. A.; Campanale, K. M.; Chirgadze, N. Y.; Clawson, D. K.; Dressman, B. A.; Hatch, S. D.; Khalil, D. A.; Kosa, M. B.; Lubbehusen, P. P.; Muesing, M. A.; Patick, A. K.; Reich, S. H.; Su, K. S.; Tatlock, J. H. Viracept (nelfinavir mesylate, AG1343): A potent, orally bioavailable inhibitor of HIV-1 protease. *J. Med. Chem.* **1997**, *40*, 3979-3985.
- Keenan, R. M.; Weinstock, J.; Finkelstein, J. A.; Franz, R. G.; Gaitanopoulos, D. E.; Girard, G. R.; Hill, D. T.; Morgan, T. M.; Samanen, J. M. Peishoff, C. E.; Tucker, L. M.; Aiyar, N.; Griffin, e.; Ohlstein, E. H.; Stack, E. J.; Weidley, E. F.; Edwards, R. M. Potent nonpeptide angiotensin II receptor antagonists. 2. 1-(Carboxybenzyl)imidazole-5-acrylic acids. *J. Med. Chem.* **1993**, *36*, 1880-1892.
- Keller, P. A.; Birch, C.; Leach, S. P.; Tyseen, D.; Griffith, R. Novel pharmacophore-based methods reveal gossypol as a reverse transcriptase inhibitor. *J. Mol. Graphics Model.* **2003**, *21*, 365-373.
- Kier, L. B.; Hall, L. H. Molecular connectivity in chemistry and drug research. Academic Press, New York, 1976.
- Kier, L. B.; Hall, L. H. Molecular structure description: The electrotopological state. Academic Press, San Diego, 1999.
- Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discov.* **2004**, *3*, 935-949.
- Klebe, G.; Abraham, U.; Mietzner, T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. *J. Med. Chem.* **1994**, *37*, 4130-4146.
- Klebe, G. Comparative molecular similarity indices analysis: COMSIA. En *3D QSAR in Drug Design: Recent Advances*. Kubini, H.; Folkers, G.; Martin, Y. C., Eds.; Kluwer/Escom, Great Britain, 1998; pp 87-104.
- Koga, H.; Itoh, A.; Murayama, S.; Suzue, S.; Irikura, T. Structure-activity relationships of antibacterial 6,7- and 7,8-disubstituted 1-alkyl-1,4-dihydro-4-oxoquinoline-3-carboxylic acids. *J. Med. Chem.* **1980**, *23*, 1358-1363.
- Kong, J.; White, C. A.; Krylov, A. I.; Sherrill, C. D.; Adamson, R. D.; Furlani, T. R.; Lee, M. S.; Lee, A. M.; Gwaltney, S. R.; Adams, T. R.; Ochsenfeld, C.; Gilbert, A. T. B.; Kedziora, G. S.; Rassolov, V. A.; Maurice, D. R.; Nair, N.; Shao, Y.; Besley, N. A.; Maslen, P. E.; Dombroski, J. P.; Daschel, H.; Zhang, W.; Korambath, P. P.; Baker, J.; Byrd, E. F. C.; Van Vooris, T.; Oumi, M.; Hirata, S.; Hsu, C.-P.; Ishikawa, N.; Florian, J. Warshel, A.; Johnson, B. G.; Gill, P. M. W., Head-Gordon, M.; Pople, J. A. Q-Chem 2.0: a high-performance ab initio electronic structure program package. *J. Comput. Chem.* **2000**, *21*, 1532-1548.
- Kontoyianni, M.; McClellan, L. M.; Sokol, G. S. Evaluation of docking performance: comparative data on docking algorithms. *J. Med. Chem.* **2004**, *47*, 558-565.

- (a) Kubini, H. Comparative molecular field analysis (CoMFA). En *Encyclopedia of Computational Chemistry*. Schleyer, P. V. R., Ed.; John Wiley and Sons, West Sussex, UK, 1998; Vol. 1, pp 448-460.
- (b) Kubini, H. Quantitative structure-activity relationships in drug design. En *Encyclopedia of Computational Chemistry*. Schleyer, P. V. R., Ed.; John Wiley and Sons, West Sussex, UK, 1998; Vol. 4, pp 2309-2320.
- Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.* **1982**, *161*, 269-288.
- Kuo, CL.; Assefa, H.; Kamath, S.; Brzozowski, Z.; Slawinski, J.; Saczewski, F.; Buolamwini, J. K.; Neamati, N. Application of CoMFA and CoMSIA 3D-QSAR and docking studies in optimization of mercaptobenzenesulfonamides as HIV-1 integrase inhibitors. *J. Med. Chem.* **2004**, *47*, 385-399.
- Labarrios, F.; Garduño, L.; Vidal, MDR.; Garcia, R.; Salazar, M.; Martinez, E.; Diaz, F.; Chamorro, G.; Tamariz, J. Synthesis and hypolipidaemic evaluation of a series of α -asarone analogues related to clorofibrate in mice. *J. Pharm. Pharmacol.* **1999**, *51*, 1-7.
- Leach, A. R; Hann, M. M. The *in silico* world of virtual libraries. *Drug Discov. Today* **2000**, *5*, 326-336.
- Leach, A. R. Molecular modeling. Principles and Applications. Pearson Education Limited, Essex, England, 2001; pp 2.
- Leach, A. R.; Gillet, V. An Introduction to Chemoinformatics. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2003; pp 2, 57.
- Link, M. Statin switch to OTC. *Nat. Rev. Drug Discov.* **2004**, *3*, 641-642.
- Lindberg, J.; Sigurosson, S.; Löwgren, S.; Andersson, H. O.; Sahlberg, C.; Noréen, R.; Fridborg, K.; Zhang, H.; Unge, T. Structural basis for the inhibitory efficacy of efavirenz (DMP-266) MSC194 and PNU142721 towards the HIV-1 RT K103N mutant. *Eur. J. Biochem.* **2002**, *269*, 1670-1677.
- López-Balbiaux, N. Diseño, síntesis y actividad antiparasitaria de nuevas carboxamidas 1-metilbencimidazólicas. Tesis de Maestría. Facultad de Química, UNAM. México D.F. 2005.
- Lozano, J. J.; Pastor, M.; Cruciani, G.; Gaedt, K.; Centeno N. B.; Gago, F; Sanz, F. 3D-QSAR methods on the basis of ligand-receptor complexes. Application of COMBINE and GRID/GOLPE methodologies to a series of CYP1A2 ligands. *J. Comput.-Aided. Mol. Des.* **2000**, *14*, 341-353.
- Luna-González, A. Síntesis y actividad antiprotozoaria contra *G. lamblia* de 5-clorobenzoxazoles 2-sustituidos. Tesis de Maestría. Facultad de Química, UNAM. México D.F. 2004.
- Mai, A.; Sbardella, G.; Artico, M.; Ragno, R.; Massa, S.; Novellino, E.; Greco, G.; Lavecchia, A.; Musiu, C.; La Colla, M.; Murgioni, C.; La Colla, P.; Loddo, R. Structure-based design, synthesis, and biological evaluation of conformationally restricted novel 2-alkylthio-6-[1-(2,6-difluorophenyl)alkyl]-3,4-dihydro-5-alkylpyrimidin-4(3H)-ones as non-nucleoside inhibitors of HIV-1 reverse transcriptase. *J. Med. Chem.* **2001**, *44*, 2544-2554.

- Mao, C.; Subeck, E. A.; Venkatachalam, T. K.; Uckun, F. M. Structure-based design of non-nucleoside inhibitors for wild-type and drug resistant HIV reverse transcriptase. *Biochem. Pharmacol.* **2000**, *60*, 1251-1265.
- Martin, Y. C. Overview of concepts and methods in computer-assisted drug design. *Methods Enzymol.* **1991**, *203*, 587-613.
- Martin, Y. C. 3D Database searching in drug design. *J. Med. Chem.* **1992**, *35*, 2145-2154.
- Martínez, M. *Las Plantas Medicinales de México*. Botas, México City, 1992; pp 125-127.
- McGaughey, G. B.; Gagné, M.; Rappé, A. π -Stacking interactions. Alive and well in proteins. *J. Biol. Chem.* **1998**, *273*, 15458-15463.
- McTaggart, F.; Buckett, L.; Davidson, R.; Holdgate, G.; McCormick, A.; Scheneck, D.; Smith, G.; Warwick, M. Preclinical and clinical pharmacology of rosuvastatin, a new 3-hydroxy-3-methylglutaryl coenzyme A reductase inhibitor. *Am. J. Cardiol.* **2001**, *87* (suppl.), 28B-32B.
- Mehler, E. L.; Solmajer, T. Electrostatic effects in proteins: comparison of dielectric and charge models. *Protein Engineering* **1991**, *4*, 903-910.
- Méndez-Cuesta, C. A. Tesis de Licenciatura. Facultad de Química, UNAM. México D.F. En desarrollo.
- Meyer, E. A.; Castellano, R. K.; Diederich, F. Interactions with aromatic rings in chemical and biochemical recognition. *Angew. Chem. Int. Ed.* **2003**, *42*, 1210-1250.
- Mineno, T.; Avery, M. A. Giardiasis: Recent progress in chemotherapy and drug development. *Curr. Pharm. Design* **2003**, *9*, 841-855.
- Molconn-Z, version 4.05, Hall Associates Consulting: Quincy, MA.
- Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comp. Chem.* **1998**, *19*, 1639-1662.
- Morris, G. M.; Goodsell, D. S.; Huey, R.; Olson, A. J. *AutoDock 3.0.5 user guide*; The Scripps Research Institute, Department of Molecular Biology: La Jolla, CA, 2001.
- Muegge, I.; Rarey, M. Small molecule docking and scoring. En *Reviews in Computational Chemistry*. Lipkowitz, K.; Boyd, D., Eds.; VCH Publishers, New York, 2001; Vol. 17, pp. 1-60.
- Navarrete-Vazquez, G.; Cedillo, R.; Hernandez-Campos, A.; Yopez, L.; Hernandez-Luis, F.; Valdez, J.; Morales, R.; Cortes, R.; Hernandez, M.; Castillo, R. Synthesis and antiparasitic activity of 2-(trifluoromethyl)-benzimidazole derivatives. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 187-190.
- Navarrete-Vazquez, G.; Yopez, L.; Hernandez-Campos, A.; Tapia, A.; Hernandez-Luis, F.; Cedillo, R.; González, J.; Martínez-Fernández, A.; Martínez-Grueiro, M.; Castillo, R. Synthesis and antiparasitic activity of albendazole and mebendazole analogues. *Bioorg. Med. Chem.* **2003**, *11*, 4615-4622.

- Navarrete-Vazquez, G. Diseño, síntesis y actividad antiparasitaria de derivados del 2-(trifluorometil) bencimidazol. Tesis de Doctorado. Facultad de Química, UNAM. México D.F. 2004.
- NCI DB. http://dtp.nci.nih.gov/docs/3d_database/structural_information/smiles_strings.html.
- Norinder, U.; Högberg, T. Quantitative structure-activity relationships and experimental design. En *Textbook of drug design and discovery*. Krogsgaard-Larsen, P.; Liljefors, T.; Madsen, U., Eds.; Taylor & Francis, London, 2002; pp. 117-155.
- Nunberg, J. H.; Schleif, W. A.; Boots, E. J.; O'Brien, J. A.; Quintero, J. C.; Hoffman, J. M.; Emini, E. A.; Goldman, M. E. Viral resistance to human immunodeficiency virus type 1-specific pyridinone reverse transcriptase inhibitors. *J. Virol.* **1991**, *65*, 4887-4892.
- ONUSIDA: Programa Conjunto de las Naciones Unidas sobre el VIH/SIDA y Organización Mundial de la Salud 2004 (<http://www.unaids.org>).
- Ooms, F. Molecular modeling and computer aided drug design. Examples of their applications in medicinal chemistry. *Curr. Med. Chem.* **2000**, *7*, 141-158.
- Oprea, T.; Waller, C. L. Theoretical and practical aspects of three-dimensional quantitative structure-activity relationships. En *Reviews in Computational Chemistry*. Lipkowitz, K.; Boyd, D., Eds.; VCH Publishers, New York, 1997; Vol. 11, pp. 127-182.
- Oshiro, C. M.; Kuntz, I. D.; Knegt, R. M. A. Molecular docking and structure-based design. En *Encyclopedia of Computational Chemistry*. Schleyer, P. V. R., Ed.; John Wiley and Sons, West Sussex, UK, 1998; Vol. 3, pp 1606-1612.
- Packard, C. J.; Gaw, A.; Shepherd, J. Cholesterol-lowering agents and their use. *Encyclopedia of Life Sciences*. Nature Publishing Group. Macmillan Publishers, Inglaterra, 2002 (<http://www.els.net>).
- Parniak, M. A. HIV/AIDS after twenty-five years. *Int. J. Biochem. Cell Biol.* **2004**, *36*, 1666-1667.
- Parreira, R. L. T.; Abrahão, O.; Galembeck, S. E. Conformational preferences of non-nucleoside HIV-1 reverse transcriptase inhibitors. *Tetrahedron* **2001**, *57*, 3243-3253.
- Pungpo, P. and Hannongbua, S., Three-dimensional quantitative structure-activity relationships study on HIV-1 reverse transcriptase inhibitors in the class of dipyridodiazepinone derivatives, using comparative molecular field analysis. *J. Mol. Graph. Mod.* **2000**, *18*, 581-590.
- Pungpo, P.; Hannongbua, S.; Wolschann, P. Hologram quantitative structure-activity relationships investigations of non-nucleoside reverse transcriptase inhibitors. *Curr. Med. Chem.* **2003**, *10*, 1661-1677.
- Ragno, R.; Mai, A.; Sbardella, G.; Artico, M.; Massa, S.; Musiu, C.; Mura, M.; Marturana, F.; Cadeddu, A.; La Colla, P. Computer-aided design, synthesis, and anti-HIV-1 activity in vitro of 2-alkylamino-6-[1-(2,6-difluorophenyl)alkyl]-3,4-dihydro-5-alkylpyrimidin-4(3H)-ones as novel potent non-nucleoside reverse transcriptase inhibitors, also active against the Y181C variant. *J. Med. Chem.* **2004**, *47*, 928-934.

- Ragno R.; Frasca, S.; Manetti, F.; Brizzi, A.; Massa, S. HIV-reverse transcriptase inhibition: inclusion of ligand-induced fit by cross-docking studies. *J. Med. Chem.* **2005**, *48*, 200-212.
- Randic, M. Characterization of molecular branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609-6615.
- RCSB. Brookhaven Protein Data Bank. <http://www.rcsb.org>
- Reddy, M. R.; Parrill, A. L. Overview of rational drug design. En *Rational Drug Design: Novel Methodology and Practical Applications*, Parril, A. L., Reddy, M. R., Eds.; ACS Symposium Series 719; American Chemical Society, Washington, DC, 1999; pp 1-11.
- (a) Ren, J.; Esnouf, R.; Garman, E.; Somers, D.; Ross, C.; Kirby, I.; Keeling, J.; Darby, G.; Jones, Y.; Stuart, D.; Stammers, D. High resolution structures of HIV-1 RT from four RT-inhibitor complexes. *Nat. Struct. Biol.* **1995**, *2*, 293-302.
- (b) Ren, J.; Esnouf, R.; Hopkins, A.; Ross, C.; Jones, Y.; Stammers, D.; Stuart, D. The structure of HIV-1 reverse transcriptase complexed with 9-chloro-TIBO: lessons for inhibitor design. *Structure* **1995**, *3*, 915-926.
- Ren, J.; Esnouf, R. M.; Hopkins, A. L.; Warren, J.; Balzarini, J.; Stuart, D. I.; Stammers, D. K. Crystal structures of HIV-1 reverse transcriptase in complex with carboxanilide derivatives. *Biochemistry* **1998**, *37*, 14394-14403.
- Ren, J.; Esnouf, R. M.; Hopkins, A. L.; Stuart, D. I.; Stammers, D. K. Crystallographic analysis of the binding modes of thiazoloisoindolinone non-nucleoside inhibitors to HIV-1 reverse transcriptase and comparison with modeling studies. *J. Med. Chem.* **1999**, *42*, 3845-3851.
- (a) Ren, J.; Milton, J.; Weaver, K. L.; Short, S. A.; Stuart, D. I.; Stammers, D. K. Structural basis for the resilience of efavirenz (DMP-266) to drug resistance mutations in HIV-1 reverse transcriptase. *Structure* **2000**, *8*, 1089-1094.
- (b) Ren, J.; Nichols, C.; Bird, L. E.; Fujiwara, T.; Sugimoto, H.; Stuart, D. I.; Stammers, D. K. Binding of the second generation non-nucleoside inhibitor S-1153 to HIV-1 reverse transcriptase involves extensive main chain hydrogen bonding. *J. Biol. Chem.* **2000**, *275*, 14316-14320.
- Reza-Ramírez, B. E. Tesis de Licenciatura. Facultad de Química, UNAM. México D.F. En desarrollo.
- Rivera-Pérez, J. A. Tesis de Licenciatura. Facultad de Química, UNAM. México D.F. En desarrollo.
- Rodríguez-Morales, S. Estudio CoMFA de compuestos antiparasitarios. En preparación.
- Rodríguez-Páez, L.; Juárez-Sánchez, M.; Antúnez-Solís, J.; Baeza, I.; Wong, C. α -Asarone inhibits HMG-CoA reductase, lowers serum LDL-cholesterol levels and reduces biliary CSI in hypercholesterolemic rats. *Phytomedicine* **2003**, *10*, 397-404.
- Saag, M. S.; Emini, E. A.; Laskin, O. L.; Douglas, J.; Lapidus, W. I.; Schleif, W. A.; Whitley, R. J.; Hildebrand, C.; Byrnes, V. W.; Kappes, J. C.; Anderson, K. W.; Massari, F. E.; Shaw, G. M.; L-697,661 Working Group. A short-term clinical evaluation of L-697,661, a non-nucleoside inhibitor of HIV-1 reverse transcriptase. *N. Engl. J. Med.* **1993**, *329*, 1065-1072.

- Saari, W. S.; Hoffman, J. M.; Wai, J. S.; Fisher, T. E.; Rooney, C. S.; Smith, A. M.; Thomas, C. M.; Goldman, M. E.; O'Brien, J. A.; Nunberg, J. H.; Quintero, J. C.; Schleif, W. A.; Emini, E. A.; Stern, A. M.; Anderson, P. S. 2-Pyridinone derivatives: a new class of nonnucleoside, HIV-1-specific reverse transcriptase inhibitors. *J. Med. Chem.* **1991**, *34*, 2922-2925.
- Saari, W. S.; Wai, J. S.; Fisher, T. E.; Thomas, C. M.; Hoffman, J. M.; Rooney, C. S.; Smith, A. M.; Jones, J. H.; Bamberger, D. L.; Goldman, M. E.; O'Brien, J. A.; Nunberg, J. H.; Quintero, J. C.; Schleif, W. A.; Emini, E. A.; Anderson, P. S. Synthesis and evaluation of 2-pyridinone derivatives as HIV-1-specific reverse transcriptase inhibitors. 2. Analogues of 3-aminopyridin-2(1H)-one. *J. Med. Chem.* **1992**, *35*, 3792-3802.
- Sachs, L. Applied Statistics. A Handbook of Techniques. Springer-Verlag, 1984, p. 349.
- Sánchez-Díaz, R. Síntesis de derivados del 1-metilbencimidazol con actividad antihelmíntica potencial. Tesis de Maestría. Facultad de Química, UNAM. México D.F. 1999.
- Sandoval-Rivera E. Tesis de Maestría. Facultad de Química, UNAM. México D.F. En desarrollo.
- Sardana, V. V.; Emini, E. A.; Gotlib, L.; Graham, D. J.; Lineberger, D. W.; Long, W. J.; Schlabach, A. J.; Wolfgang, J. A.; Condra, J. H. Functional analysis of HIV-1 reverse transcriptase amino acids involved in resistance to multiple nonnucleoside inhibitors. *J. Biol. Chem.* **1992**, *267*, 17526-17530.
- Sharaf, M. A., Illman, D. L. and Kowalski, B. R. Chemometrics, John Wiley & Sons, New York, 1986. pp 234-239.
- Sharma, S. Design new drugs for Helminth diseases: Lead optimization in benzimidazoles. *Adv. Drug Res.* **1994**, *25*, 105-163.
- Shen, M.; LeTiran, A.; Xiao, T.; Golbraikh, A.; Kohn, H.; Tropsha, A. Quantitative structure-activity relationship analysis of functionalized amino acid anticonvulsant agents using *k* nearest neighbor and simulated annealing PLS methods. *J. Med. Chem.* **2002**, *45*, 2811-2823.
- Shen, M.; Béguin, C.; Golbraikh, A.; Stables, J. P.; Kohn, H.; Tropsha, A. Application of predictive QSAR models to database mining: identification and experimental validation of novel anticonvulsant compounds. *J. Med. Chem.* **2004**, *47*, 2356-2364.
- Siliciano, R. F. Acquired Immune Deficiency Syndrome (AIDS). *Encyclopedia of Life Sciences*. Nature Publishing Group. Macmillan Publishers, Inglaterra, 2001 (<http://www.els.net>).
- Sipl, W.; Contreras, J-M.; Parrot, I.; Rival, Y. M.; Wermuth, C. G. Structure-based 3D QSAR and design of novel acetylcholinesterase inhibitors. *J. Comput.-Aid. Mol. Des.* **2001**, *15*, 395-410.
- Sipl, W. Development of biologically active compounds by combining 3D QSAR and structure-based design methods. *J. Comput.-Aid. Mol. Des.* **2002**, *16*, 825-830.
- Smith, M. B. K.; Hose, B. M.; Hawkins, A.; Lipchock, J.; Farnsworth, D. W.; Rizzo, R. C.; Tirado-Rives, J.; Arnold, E.; Zhang, W.; Hughes, S. H.; Jorgensen, W. L.; Michejda, C. J.; Smith, Jr. R. H. Molecular modeling calculations of HIV-1 reverse transcriptase nonnucleoside inhibitors: correlation of binding energy with biological activity for novel 2-aryl-substituted benzimidazole analogues. *J. Med. Chem.* **2003**, *46*, 1940-1947.

- Soriano-Agaton, F. N. Síntesis del 5-carboxi-2-(trifluorometil)-1 h-bencimidazol, un bioprecursor y un análogo de tipo carboxamida. Tesis de Licenciatura. Facultad de Química, UNAM. México D.F. 2002.
- Sotriffer, C.; Klebe, G.; Stahl, M.; Böhm, H-J. Docking and scoring functions / virtual screening. En *Burger's Medicinal Chemistry and Drug Discovery*. Abraham, D. J., Ed.; John Wiley and Sons, Hoboken, New Jersey, 2003; Vol. 1, pp 281-331.
- Sybyl version 6.8; Tripos Associates, Inc., St. Louis, MO, USA.
- Tantillo, C.; Ding, J.; Jacobo-Molina, A.; Nanni, R. G.; Boyer, P. L.; Hughes, S. H.; Pauwels, R.; Andries, K.; Janseen, P. A. J.; Arnold, E. Locations of anti-AIDS drug binding sites and resistance mutations in the three-dimensional structure of HIV-1 reverse transcriptase: implications for mechanisms of drug inhibition and resistance. *J. Mol. Biol.* **1994**, *243*, 369-387.
- Titmuss, S. J.; Keller, P. A.; Griffith, R. Docking experiments in the flexible non-nucleoside inhibitor binding pocket of HIV-1 reverse transcriptase. *Bioorg. Med. Chem.* **1999**, *7*, 1163-1170.
- Tobert, J. A. Lovastatin and beyond: The history of the HMG-CoA reductase inhibitors. *Nat. Rev. Drug Discov.* **2003**, *2*, 517-526.
- Topliss, J. G.; Costello, R. J. Chance correlations in structure-activity studies using multiple regression analysis. *J. Med. Chem.* **1972**, *15*, 1066-1068.
- Tropsha, A., Cho, S. J. and Zheng, W. "New Tricks For an Old Dog": Development and application of novel QSAR methods for rational design of combinatorial chemical libraries and database mining. En Parrill, A. L. and Reddy, M. R. (Eds.), ACS Symposium Series 719, American Chemical Society, Washington, DC, 1999, pp 198-211.
- (a) Tropsha, A.; Zheng, W. Computer aided drug design. En *Computational Biochemistry and Biophysics*. Becker, O. M.; Mackerell, Jr. A. D.; Roux, B.; Watanabe, M., Eds.; Marcel Dekker, New York, NY, 2001; pp 351-369.
- (b) Tropsha, A.; Zheng, W. Identification of the descriptor pharmacophores using variable selection QSAR: Applications to database mining. *Curr. Pharm. Design* **2001**, *7*, 599-612.
- (a) Tropsha, A. Recent trends in quantitative structure-activity relationships. En *Burger's Medicinal Chemistry and Drug Discovery*. Abraham, D. J., Ed.; John Wiley and Sons, Hoboken, New Jersey, 2003; Vol. 1, pp 49-76.
- (b) Tropsha, A.; Gramatica, P.; Gombar, V. K. The importance of being earnest: Validation is the absolute essential for successful application and interpretation of QSPR models. *Quant. Struct.-Act. Relat. Comb. Sci.* **2003**, *22*, 69-77.
- Valdez, J.; Cedillo, R.; Hernández-Campos, A.; Yépez-Mulia, L.; Hernández-Luis, F.; Navarrete-Vázquez, G.; Morales, R.; Cortés, R.; Hernández, M.; Castillo, R. Synthesis and antiparasitic activity of 1H-benzimidazole derivatives. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2221-2224.
- Vargas-Benítez, P. Tesis de Maestría. Facultad de Química, UNAM. México D.F. En desarrollo.

- Vilchis-Reyes, M. A. Síntesis, actividad giardicida y permeabilidad aparente de derivados del 2-(trifluorometil) bencimidazol. Tesis de Maestría. Facultad de Química, UNAM. México D.F. 2004.
- ViewerPro version 5.0; Accelrys, Inc., San Diego, CA, USA.
- Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Prot. Eng.* **1995**, *8*, 127-134.
- Walters, W. P.; Stahl, M. T.; Murcko, M. A. Virtual screening - an overview. *Drug Discov. Today* **1998**, *3*, 160-178.
- Wang, S.; Milne, G. W. A.; Yan, X.; Posey, I. J.; Nicklaus, M. C.; Graham, L.; Rice, W. G. Discovery of novel, non-peptide HIV-1 protease inhibitors by pharmacophore searching. *J. Med. Chem.* **1996**, *39*, 2047-2054.
- Waterbeemd, van de H.; Gifford, E. ADME *in silico* modelling: towards prediction paradise? *Nat. Rev. Drug Discov.* **2003**, *2*, 192-204.
- Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **1984**, *106*, 765-784.
- Weinstock, J.; Keenan, R. M.; Samanen, J.; Hempel, J.; Finkelstein, J. A.; Franz, R. G.; Gaitanopoulos, D. E.; Girard, G. R.; Gleason, J. G. Hill, D. T.; Morgan, T. M.; Peishoff, C. E.; Aiyar, n.; Brooks, D. P.; Fredrickson, T. A.; Ohlstein, E. H.; Ruffolo, Jr. R. R.; Stack, E. J.; Sulpizio, A. C.; Wiedley, E. F.; Edwards, R. M. 1-(Carboxybenzyl)imidazole-5-acrylic acids: potent and selective angiotensin II receptor antagonists. *J. Med. Chem.* **1991**, *34*, 1514-1517.
- Wild, SH.; Byrne, CD. Cholesterol and vascular disease. *Encyclopedia of Life Sciences*. Nature Publishing Group. Macmillan Publishers, Inglaterra, 2001 (<http://www.els.net>).
- Wlodawer, A. Rational approach to AIDS drug design through structural biology. *Annu. Rev. Med.* **2002**, *53*, 595-614.
- Wold, S.; Sjöström, M.; Eriksson, L. Partial least squares projections to latent structures (PLS) in chemistry. En *Encyclopedia of Computational Chemistry*. Schleyer, P. V. R., Ed.; John Wiley and Sons, West Sussex, UK, 1998; Vol. 3, pp 2006-2021.
- Yadav, P. N. S.; Das, K.; Ding, J.; Arnold, E.; Yadav, J. S.; Modak, M. J. Structure-based design of inhibitors of HIV-1 reverse transcriptase: a potential class of bidentate nonnucleoside inhibitors. *J. Mol. Struct. (Theochem)* **1998**, *423*, 101-111.
- Zheng, W.; Tropsha, A. Novel variable selection quantitative structure-property relationship approach based on the *k*-Nearest-Neighbor principle. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 185-194.
- Zhou, Z.; Madrid, M.; Madura, J. D. Docking of non-nucleoside inhibitors: Neotripterifordin and its derivatives to HIV-1 reverse transcriptase. *Proteins* **2002**, *49*, 529-542.

Anexo

Flexible docking of pyridinone derivatives into the non-nucleoside inhibitor binding site of HIV-1 reverse transcriptase[☆]

José Luis Medina-Franco,^a Sergio Rodríguez-Morales,^a Cecilia Juárez-Gordiano,^a Alicia Hernández-Campos,^a Jesús Jiménez-Barbero^{b,*} and Rafael Castillo^{a,*}

^aDepartamento de Farmacia, Facultad de Química, UNAM, CU, DF 04510, Mexico

^bCentro de Investigaciones Biológicas, CSIC, Madrid 28040, Spain

Received 26 February 2004; revised 3 September 2004; accepted 8 September 2004

Available online 2 October 2004

Abstract—Potent non-nucleoside reverse transcriptase inhibitors (NNRTIs) of the pyridinone derivative type were docked into nine NNRTIs binding pockets of HIV-1 reverse transcriptase (RT) structures. The docking results indicate that pyridinone analogues adopt a butterfly conformation and share the same binding mode as the crystal inhibitors in the pocket geometries of nevirapine, 1051U91, 9-Cl-TIBO, Cl- α -APA, efavirenz, UC-781, and S-1153. The results are in agreement with the data concerning mutational and structure–activity relationships available for pyridinone analogues and aid in the understanding, at the molecular level, of the biological response of published hybrid pyridinone molecules. Strategies to design further pyridinone derivatives active against RT containing mutations are discussed.

© 2004 Elsevier Ltd. All rights reserved.

1. Introduction

The reverse transcriptase (RT) of the human immunodeficiency virus type 1 (HIV-1) is an attractive target in the treatment of the acquired immune deficiency syndrome (AIDS), for which no completely successful chemotherapy is yet available.^{1,2} RT inhibitors are classified as nucleoside and non-nucleoside, depending on their action mechanism. One advantage of the latter group is that they lack the toxic effects associated with the former. To date, three non-nucleoside RT inhibitors have been approved for clinical use named nevirapine (VIRAMUNE[®]), delavirdine (RESCRIPTOR[®]), and efavirenz (SUSTIVA[®]),¹ however, all these and other approved drugs induce drug resistant variants of HIV-1.

Pyridinone derivatives³ are a class of non-nucleoside reverse transcriptase inhibitors (NNRTIs). From kinetic³ and analysis of resistance mutations⁴ it was concluded that this type of inhibitors should bind in the same pocket as other NNRTIs. Extensive structure–activity relationship (SAR) studies have been conducted for these compounds leading to potent RT inhibitors (Fig. 1).^{5–8}

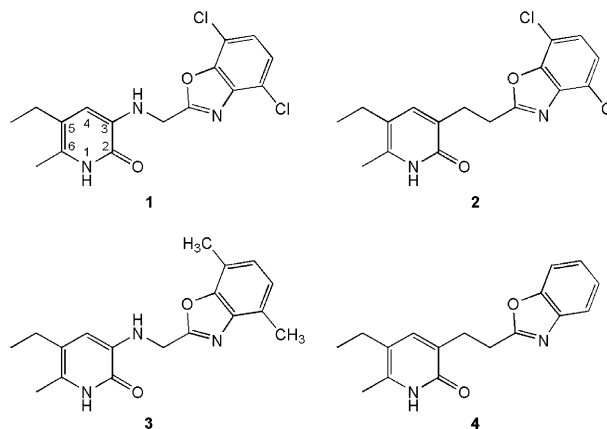


Figure 1. Pyridinone derivatives studied.

Keywords: AutoDock; Binding mode model; Mutation; Non-nucleoside reverse transcriptase inhibitors.

[☆]Taken in part from the Ph.D. dissertation of José Luis Medina-Franco.

* Corresponding authors. Tel./fax: +52 5622 5287 (R.C.); tel.: +34 91534 6623; fax: +34 91564 4853 (J.J.); e-mail addresses: jjbarbero@cib.csic.es; rafaelc@servidor.unam.mx

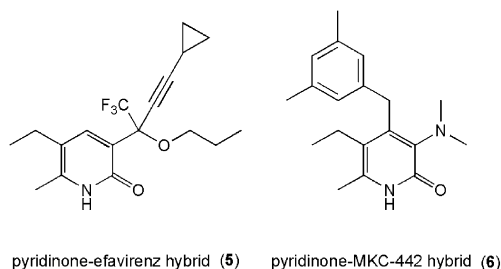


Figure 2. Hybrid pyridinone analogues with anti-HIV-1 RT inhibitory activity.

Compound **1** was subject to clinical studies, which showed good activity in patients, but rapid resistant strains of the virus emerged.^{9,10} Amino acid substitutions mainly at positions 103 and 181 are responsible for the emergence of resistant virus not only for **1** but also for other pyridinone derivatives.^{10–12}

Several quantitative structure–activity relationship analyses (QSAR)^{13,14} have been performed for pyridinone analogues in which a hydrophobic character of the binding pocket appears. Recent hybrid pyridinone–NNRTIs designs have generated molecules with RT inhibitory activity against wild type and mutant strains (Figs. 2 and 3).^{15–17} This suggests the importance of the pyridinone ring for RT inhibitory activity and demonstrates how the pyridinone ring can lead to active compounds against mutant strains; however, in these studies^{15–17} an explanation for the activity of the hybrid pyridinone inhibitors against the mutant strains at the molecular level is not provided.

Studies of several crystal structures of HIV-1 RT in the apo form¹⁸ and complexed with different NNRTIs have been published.^{19–30} The non-nucleoside inhibitor binding pocket of HIV-1 RT is formed by the amino acids Pro95, Leu100, Lys101, Lys103, Val106, Val179, Tyr181, Tyr188, Gly190, Phe227, Trp229, Leu234, His235, Pro236, and Tyr318 of the p66 subunit and Glu138 of the p51 subunit. Crystal structures of RT–NNRTIs complexes reveal that some amino acids of the binding pocket adopt different conformations when bound to different inhibitors leading to distinct pocket volumes.^{20,24,27} A marked variation in the pocket geometry is due to the modified position of the amino acid Pro236 when RT is complexed with 1-[(2-hydroxyethoxy)-methyl]-6-(phenylthio)thymine (HEPT) derivatives.²⁰ Despite these variations, several NNRTIs share a common binding mode, which is a butterfly-like conformation,^{19–23,25,26,28,29} however, no crystal structures of HIV-1 RT in complex with pyridinone analogues that explain the resistance of HIV-1 RT to pyridinone analogues are available.

In order to gain an insight into the binding mode and interactions of pyridinone derivatives with HIV-1 RT and, consequently, to improve the development of more efficient HIV-1 RT inhibitors, especially active against mutant strains, we hereby describe a study of the docking of potent pyridinone derivatives (Fig. 1) with HIV-1 RT. The docking studies will provide information for a better understanding of drug resistant mechanisms. The compounds were flexibly docked with the program AutoDock 3.0³¹ into several crystal structures of RT complexed with structural diverse NNRTIs. Since no crystal structure is available for RT in complex with

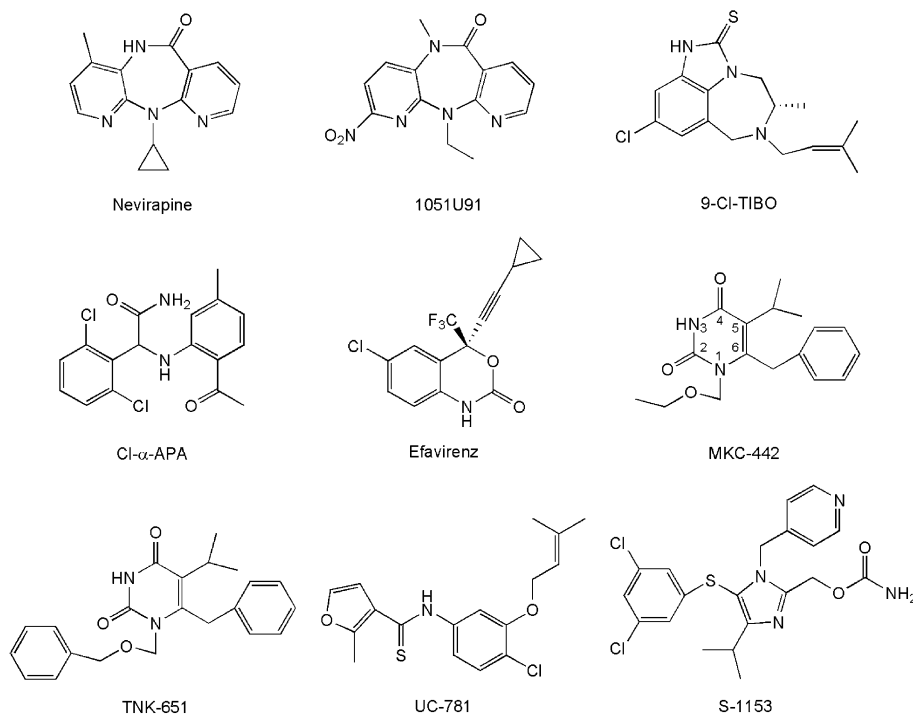


Figure 3. NNRTIs found in the crystal HIV-1 RT complexes.

pyridinones, computational docking is a useful tool to develop models that will assist in the rational design of further inhibitors. Previous docking studies with other NNRTIs have been conducted.^{32–36} Using different pocket geometries is a good approach to model the flexible binding site.^{32,37,38}

2. Results

Crystal HIV-1 RT structures complexed with the NNRTIs nevirapine, 1051U91, 9-Cl-TIBO, Cl- α -APA, efavirenz, MKC-442, TNK-651, UC-781, and S-1153 were used (Fig. 3). Table 1 provides a summary of the data and source of these structures.

2.1. Validation of the docking protocol

The docking protocol was validated for each crystal structure predicting the binding mode of the corresponding inhibitor. Each NNRTI was removed from the active site and docked back into the binding pocket in the conformation found in the crystal structure. The results of the validation are summarized in Table 2.

AutoDock successfully predicted the binding mode of the crystal NNRTIs with RMS deviations below or near 1 Å. The highly-populated clusters obtained in most cases indicate that the docking program found one binding mode mainly.

Table 1. Crystal HIV-1 RT structures used in this study

NNRTI	PDB code	Resolution Å	R-factor	Reference
Nevirapine	1VRT	2.2	0.186	19
1051U91	1RTH	2.2	0.214	19
9-Cl-TIBO	1REV	2.6	0.224	21
Cl- α -APA	1VRU	2.4	0.187	19
Efavirenz	1FK9	2.5	0.218	22
MKC-442	1RT1	2.55	0.197	20
TNK-651	1RT2	2.55	0.207	20
UC-781	1RT4	2.90	0.237	29
S-1153	1EP4	2.50	0.254	23

Table 2. Docking of crystal NNRTIs

NNRTI	Members in cluster ^a	Lowest energy (kcal/mol) ^b	Average energy (kcal/mol)	RMSD (Å)
Nevirapine	100	-9.68	-9.67	0.70
1051U91	100	-9.97	-9.95	1.25
9-Cl-TIBO	100	-11.83	-11.82	0.89
Cl- α -APA	68	-10.49	-10.31	0.96
Efavirenz	99	-10.97	-10.96	0.24
MKC-442	97	-11.15	-11.05	0.75
TNK-651	88	-14.01	-13.82	0.46
UC-781	93	-12.38	-12.03	0.92
S-1153	59	-15.63	-15.30	0.54

^a The first cluster of a total 100 runs is shown.

^b Value for the optimal structure in the cluster. Docking energies are as

2.2. Docking of pyridinone derivatives

The top ranked binding conformation found for each compound in the binding pocket of nevirapine was docked into the binding pockets of the other eight NNRTIs. The results are summarized in Table 3. Since the different binding sites were superimposed onto the structure of the RT–nevirapine complex before docking (see Section 5 for details), the RMSD values in Table 3 are an approximation of the similarity between the conformation found for each compound within a binding pocket and the pocket of nevirapine. According to the most populated clusters, 1–4 occupy the same binding pocket as the corresponding NNRTI and have the same orientation as nevirapine, 1051U91, 9-Cl-TIBO, Cl- α -APA, efavirenz, UC-781, and S-1153. In the pockets of nevirapine and Cl- α -APA, the most populated cluster for each compound also showed the lowest average docked energy (Figs. 4 and 5).

In the binding pockets of MKC-442 and TNK-651 (HEPT derivatives) it was predicted that the pyridinone analogues would have a different binding mode from that of the crystal inhibitor. In these pockets, the benzoxazol ring of the pyridinone derivatives is located in the same sub-pocket volume as the ethoxymethyl group of MKC-442 and the (benzyloxy)methyl group of TNK-651 (Figs. 6 and 7). In all binding pockets it was predicted that the docked pyridinone molecule would have a butterfly-like geometry.

For the binding pockets of nevirapine, 1051U91, 9-Cl-TIBO, Cl- α -APA, efavirenz, UC-781, and S-1153 a very good agreement between the docking energies calculated by AutoDock and the biological activity was observed. Table 4 summarizes the IC₅₀ reported for 1–4³ and the average of docking energies estimated within the seven binding pockets. This good relationship is an indication that the docking calculation produced reasonable binding modes within these binding geometries. For the binding pockets of MKC-442 and TNK-651 no good agreement between the biological activity³ and the docking energies (Table 3) was observed.

After docking, the binding pocket of nevirapine in complex with the top ranked binding conformation of 1–4 was fully optimized.³⁹ Very small variations were observed between the conformations predicted by AutoDock and the conformations of 1–4 after minimization.

3. Discussion

3.1. Docking model

The docking results indicate that the pyridinone derivatives have a common binding mode in the binding pockets of nevirapine, 1051U91, 9-Cl-TIBO, Cl- α -APA, efavirenz, UC-781, and S-1153. The hydrophobic character of the binding pocket is in very good agreement with the QSAR data available for pyridinone derivatives.^{13,14} The pyridinone molecules adopt a

Table 3. Docking of 1–6 into different binding pockets

Binding pocket	Compound	Cluster ^a	Members in cluster	Lowest energy (kcal/mol) ^b	Average energy (kcal/mol)	RMSD (Å) ^c
Nevirapine	1	1	92	−11.13	−11.08	0.00
	2	1	93	−11.69	−11.64	0.00
	3	1	94	−11.18	−11.14	0.00
	4	1	96	−10.63	−10.58	0.00
	5	1	100	−9.67	−9.53	0.40
	6	1	3	−10.62	−10.60	3.10
	6	2	92	−10.51	−10.49	0.15
1051U91	1	1	91	−10.90	−10.81	0.33
	2	1	99	−11.73	−11.70	0.55
	3	1	93	−10.97	−10.89	0.34
	4	1	3	−10.85	−10.74	2.65
	4	2	89	−10.58	−10.45	1.08
	5	1	96	−9.42	−9.22	0.70
	6	1	96	−10.42	−10.39	0.59
9-Cl-TIBO	1	1	3	−11.50	−11.40	3.01
	1	4	49	−11.08	−10.90	0.75
	2	1	93	−12.12	−11.75	1.07
	3	1	3	−11.47	−11.37	2.95
	3	6	65	−11.13	−10.98	0.74
	4	1	97	−11.23	−11.13	0.78
	5	1	100	−10.15	−9.96	0.62
	6	1	5	−10.37	−10.32	1.88
6	2	73	−10.35	−10.27	1.04	
Cl- α -APA	1	1	92	−10.96	−10.82	0.68
	2	1	94	−11.65	−11.53	0.62
	3	1	96	−11.02	−10.91	0.67
	4	1	95	−10.68	−10.56	0.63
	5	1	91	−9.32	−9.12	0.55
	6	1	4	−10.07	−10.03	1.68
	6	2	94	−9.90	−9.88	0.35
Efavirenz	1	1	10	−11.49	−11.25	1.47
	1	2	74	−11.47	−11.26	0.56
	2	1	84	−11.95	−11.84	0.63
	3	1	75	−11.57	−11.39	0.60
	4	1	80	−11.12	−10.97	0.53
	5	1	83	−10.42	−10.18	0.67
	6	1	15	−10.75	−10.54	2.91
	6	3	40	−10.43	−10.28	2.43
MKC-442	1	1	24	−11.76	−11.66	3.18
	2	1	4	−12.09	−12.05	3.06
	2	2	35	−12.04	−11.99	2.89
	3	1	25	−11.80	−11.72	3.15
	4	1	5	−11.21	−11.00	3.29
	4	9	59	−10.66	−10.60	0.59
	5	1	79	−9.56	−9.34	1.41
	6	1	97	−11.01	−10.99	0.72
TNK-651	1	1	1	−11.84	−11.84	3.88
	1	2	32	−11.71	−11.66	3.03
	2	1	4	−12.49	−12.29	3.52
	2	2	34	−12.19	−12.13	2.87
	3	1	3	−12.28	−12.19	3.63
	3	2	34	−11.76	−11.71	2.97
	4	1	3	−11.74	−11.53	3.39
	4	8	29	−10.89	−10.82	2.76
	5	1	98	−9.95	−9.61	0.96
	6	1	96	−11.00	−10.98	0.66
UC-781	1	1	16	−11.39	−11.21	3.15
	1	4	67	−10.60	−10.54	0.68
	2	1	15	−12.05	−11.88	2.78
	2	2	75	−11.46	−11.38	0.74
	3	1	13	−11.41	−11.14	3.15
	3	3	78	−10.82	−10.77	0.79

Table 3 (continued)

Binding pocket	Compound	Cluster ^a	Members in cluster	Lowest energy (kcal/mol) ^b	Average energy (kcal/mol)	RMSD (Å) ^c
	4	1	2	-10.76	-10.49	2.87
	4	2	93	-10.73	-10.67	0.62
	5	1	56	-9.84	-9.57	1.20
	5	2	38	-9.67	-9.41	1.24
	6	1	96	-10.65	-10.61	1.02
S-1153	1	1	2	-10.93	-10.91	3.08
	1	3	71	-10.76	-10.62	1.22
	2	1	73	-11.30	-11.10	1.34
	3	1	3	-10.90	-10.83	3.65
	3	2	88	-10.77	-10.70	1.37
	4	1	18	-10.41	-10.27	3.69
	4	9	53	-10.03	-9.98	1.14
	5	1	37	-8.99	-8.76	1.28
	6	1	99	-10.07	-10.05	0.56

^a The best clusters of a total 100 runs are shown.

^b Value for the optimal structure in the cluster. Docking energies are as determined by AutoDock.

^c Coordinates found with the docking into the nevirapine binding pocket were used as reference.

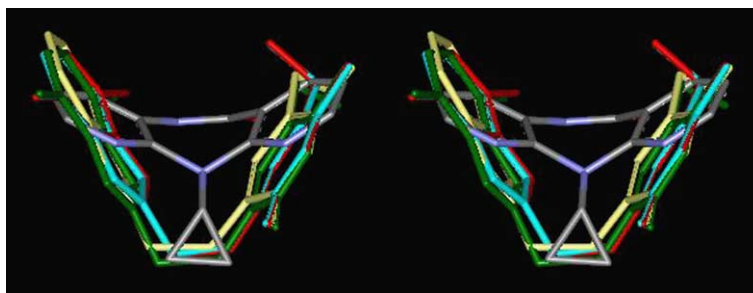


Figure 4. Top ranked binding mode of **1** (red), **2** (green), **3** (blue), and **4** (yellow) into the binding site of nevirapine. Nevirapine (colored by atom type) is displayed for comparison. Hydrogens are omitted for clarity.

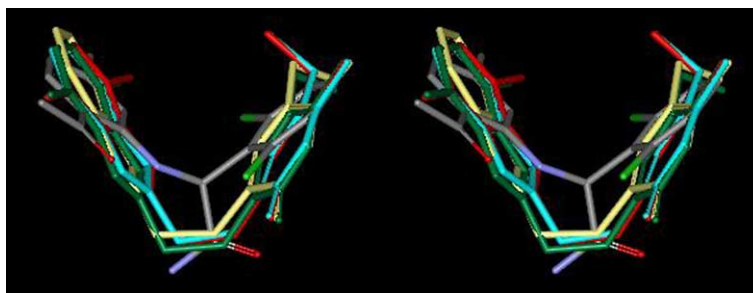


Figure 5. Top ranked binding mode of **1-4** into the binding site of Cl- α -APA (displayed for comparison). The color scheme is that of Figure 4. Hydrogens are omitted for clarity.

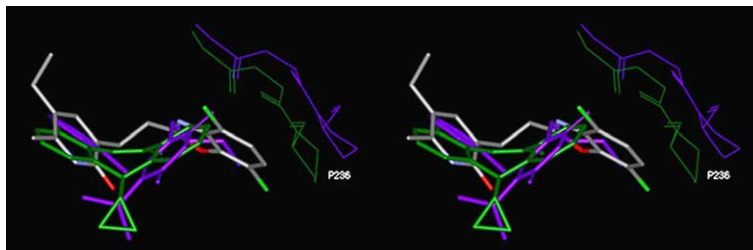


Figure 6. Top ranked binding mode of **2** (colored by atom type) into the binding site of MKC-442 (thick purple sticks). The position of Pro236 is displayed when MKC-442 (thin purple sticks) and nevirapine (thin green sticks) are the crystal inhibitors. Nevirapine is in thick green sticks. Hydrogens are omitted for clarity.

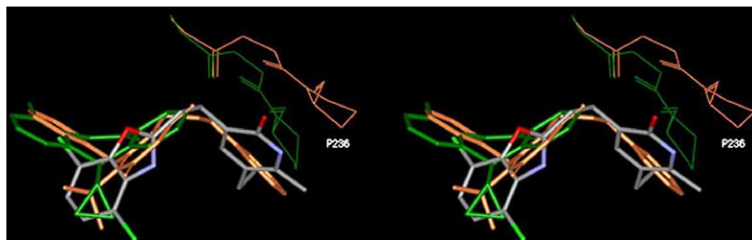


Figure 7. Top ranked binding mode of **2** (colored by atom type) into the binding site of TNK-651 (thick orange sticks). The position of Pro236 is displayed when TNK-651 (thin orange sticks) and nevirapine (thin green sticks) are the crystal inhibitors. Nevirapine is in thick green sticks. Hydrogens are omitted for clarity.

Table 4. Comparison between the docking energy and biological activity

Compound	IC ₅₀ (nM)	Docking energy (kcal/mol) ^a
1	19	−10.99
2	9.6	−11.70
3	20	−11.07
4	22	−10.71

^a Average of docking energies of the optimal structure in the most populated clusters obtained within the binding sites of nevirapine, 1051U91, 9-Cl-TIBO, Cl- α -APA, efavirenz, UC-781, and S-1153.

butterfly-like conformation and may be roughly overlapped with the crystal structures of the NNRTIs (Figs. 4 and 5). Similar results have been obtained for other NNRTIs for which no crystal structures of the complex are available.³⁴ A butterfly-like shape has also been calculated to be the most stable for several pyridinone analogues in the unliganded state based on ab-initio methods.⁴⁰

According to the optimized RT–pyridinone complexes, the residues Pro95, Leu100, Try181, and Tyr188 form the binding pocket of the benzoxazol ring. The amino acids Lys101, Lys103, His235, and Tyr318 form the pocket of the pyridinone ring. A hydrogen bond between the polar proton of the pyridinone ring and the main-chain carbonyl oxygen of Lys101 may be formed (Fig. 8). A similar hydrogen bond to the main-chain carbonyl oxygen of Lys101 is observed for several other NNRTIs.^{20–22,25–27,29,30} The hydrophobic contacts

between the aromatic ring of Tyr181 and the benzoxazol ring of the pyridinone derivative may stabilize the complex and also be responsible for the decrease of activity upon the mutation of this residue.^{11,12} Similar interactions are observed between one of the aromatic rings of other NNRTIs and Tyr181.⁴¹ The hydrophobic contacts with Lys103 may also explain the decrease of potency of pyridinone derivatives when the mutation of this amino acid occurs.¹²

A further interesting feature of the docking model is the region surrounding the linker between the aromatic rings of the pyridinone analogues. The linker occupies the same sub-pocket volume as the cyclopropyl group of nevirapine, the amide group of Cl- α -APA, the trifluoromethyl group of efavirenz, the methylfuran-3-carbothioamide group of UC-781 and the isopropyl group of S-1153 (Figs. 4 and 5. See also Figs. 9, 13, 15), the ethyl group of 1051U91 and the 5(*S*)-methyl group of 9-Cl-TIBO. These structural relationships among several NNRTIs have been previously quoted.²⁰ It is interesting to note that though the chemical structures of these groups are different they all seem to play a steric role,²⁰ which is in full agreement with the SAR data of pyridinone analogues. Moreover, in the QSAR studies of pyridinones no significant difference was found between ethyl or aminomethylene as linkers on the inhibitory potency.¹³

AutoDock predicted a different binding mode for the pyridinone analogues in the pocket geometries of

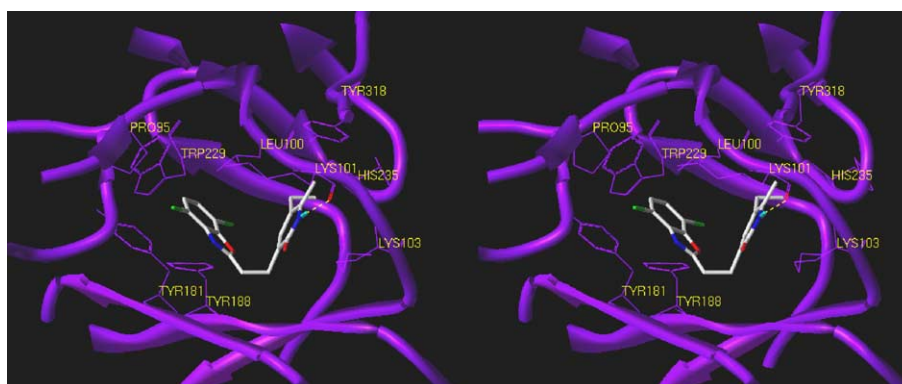


Figure 8. RT–2 complex. Optimized complex of **2** into the binding pocket of nevirapine is shown. The hydrogen bond is displayed as yellow dashes. Amino acids within 3.6 Å of the benzoxazol and pyridinone ring are labeled. Trp229 is also labeled for reference. Hydrogens are omitted.

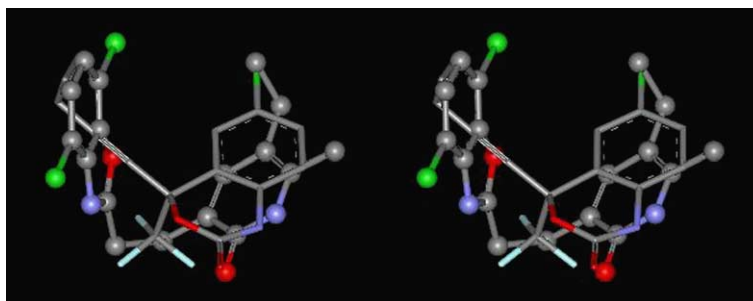


Figure 9. Top ranked binding mode of **2** (ball and sticks) into the binding site of efavirenz. Efavirenz (sticks) is displayed for comparison. Hydrogens are omitted for clarity.

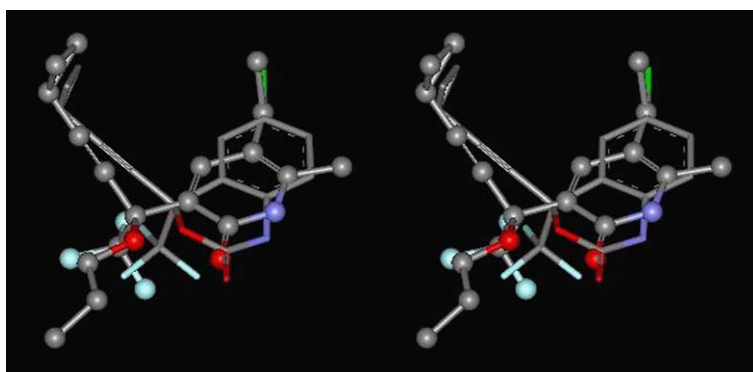


Figure 10. Top ranked binding mode of the hybrid pyridinone–efavirenz (**5**) (ball and sticks) into the binding site of efavirenz. Efavirenz (sticks) is displayed for comparison. Hydrogens are omitted for clarity.

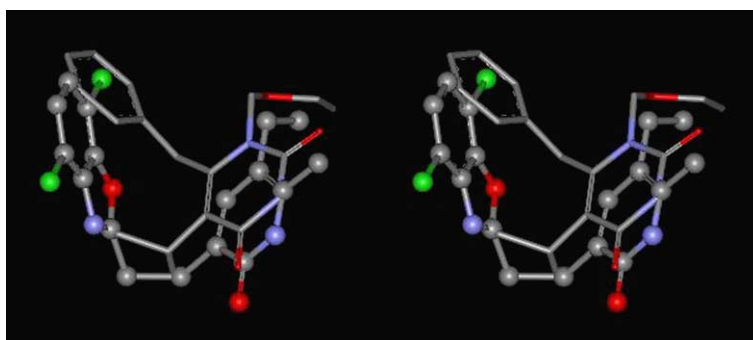


Figure 11. Top ranked binding mode of **2** (ball and sticks) found within the binding site of nevirapine compared with the binding mode of MKC-442 (sticks). Hydrogens are omitted for clarity.

MKC-442 and TNK-651. In these pockets, one of the aromatic rings of the pyridinone analogues roughly occupies the same sub-pocket volume as the N-1 substituent of these HEPT derivatives (Figs. 6 and 7). The different binding orientations found within these pockets may be explained by the different pocket geometries. It has been described that in the crystal structures of HEPT derivatives, the N-1 substituent is a ‘tail’ of the inhibitor that causes the highly flexible Pro236 loop to have a conformation similar to that observed in the unliganded RT structure.²⁰ AutoDock found better interactions if one of the aromatic rings of the pyridinones occupies the sub-pocket volume generated as a consequence of the Pro236 loop position (Figs. 6 and 7).

Since pyridinone derivatives lack the ‘tail’ of HEPT derivatives to maintain the Pro236 loop in that characteristic conformation, the actual binding pocket of pyridinone derivatives might be more related to the pocket geometries of the other NNRTIs studied. Moreover, for the binding pockets of MKC-442 and TNK-651 a good relationship between the docking energies (Table 3) and the biological activity³ was not observed.

3.2. Understanding at the molecular level hybrid pyridinone–NNRTIs designs

Comparing the docked conformations of pyridinone derivatives in the pocket of efavirenz with the position

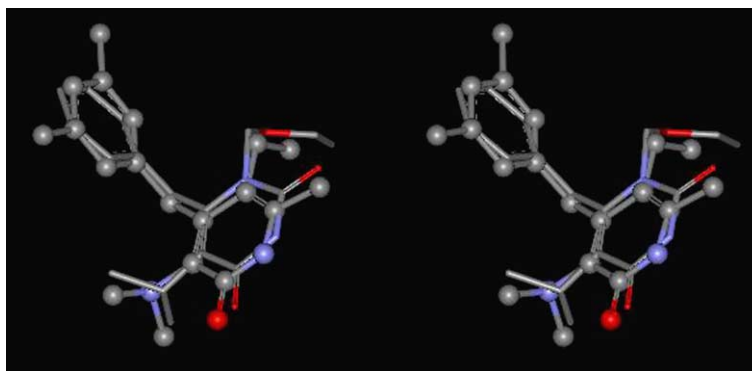


Figure 12. Top ranked binding mode of the hybrid pyridinone–MKC-442 molecule (**6**) (balls and sticks) into the binding site of MKC-442. MKC-442 (sticks) is displayed for comparison. Hydrogens are omitted for clarity.

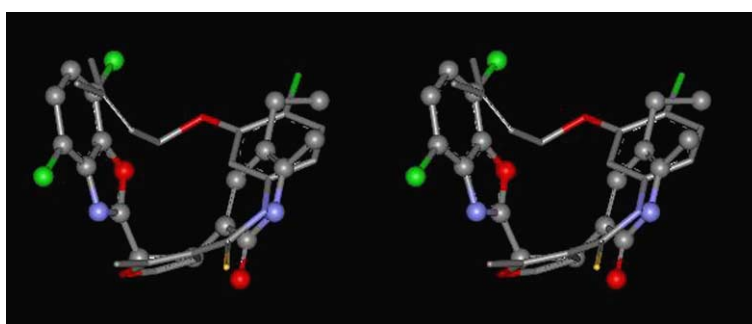


Figure 13. Top ranked binding mode of the most populated cluster of **2** (ball and sticks) into the binding site of UC-781. UC-781 (sticks) is displayed for comparison. Hydrogens are omitted for clarity.

of efavirenz found within the crystal, it is observed that the benzoxazol ring of pyridinones overlaps the cyclopropyl-propynyl group of efavirenz (Fig. 9). A close comparison of the binding modes of pyridinones and efavirenz helps to understand at the molecular level the strategy of substituting the pyridinone ring at position 4 with a cyclopropyl-propynyl group to generate an active compound (**5**).¹⁷ In order to test this hypothesis the pyridinone–efavirenz hybrid molecule (**5**) was flexibly docked into several pocket geometries. Docking results are in Table 3. In the binding pocket of efavirenz, and in most binding geometries, the top ranked binding conformation showed the expected conformation and position with a butterfly-like conformation. The cyclopropyl-propynyl groups of the hybrid molecule and efavirenz overlap, besides, the pyridinone ring of the hybrid overlaps the benzoxazin-2-one ring of efavirenz (Fig. 10).

When the binding mode of pyridinone derivatives found within the pocket geometries of nevirapine, 1051U91, 9-Cl-TIBO, Cl- α -APA, efavirenz, UC-781, and S-1153 was compared with the binding modes of MKC-442 and TNK-651 it was observed that the pyridinone ring overlaps the uracil ring of these HEPT derivatives (Fig. 11). The linker of pyridinones occupies the same sub-pocket volume as the isopropyl group of MKC-442 and TNK-651. Also, the benzoxazol ring of pyridinones is approximately in the same binding region as the benzyl group of the HEPT derivatives. These structural relationships aid in the understanding, at the

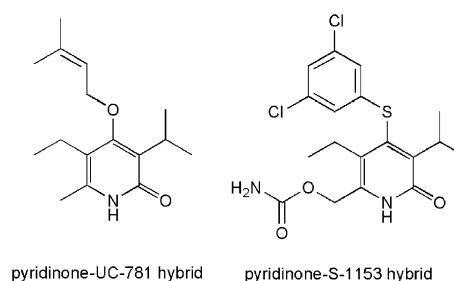


Figure 14. Examples of pyridinone derivatives with potential activity against RT containing the mutations Tyr181Cys (pyridinone–UC-781 hybrid) and Lys103Asn (pyridinone–S-1153 hybrid).

molecular level, of several SAR results obtained for a series of published hybrid pyridinone–HEPT molecules.^{15,16} The docking model totally agrees with the approach of substituting the pyridinone molecule at position 4 with a benzyl (Fig. 2) or phenylthio groups (Fig. 11). The docking of the hybrid pyridinone–MKC-442 molecule (**6**) into different pocket geometries shows that the hybrid molecule adopts the expected butterfly conformation (Fig. 12). The docking results are in Table 3.

3.3. Strategies to design further pyridinone derivatives active against RT containing mutations

One major objective in the current design of NNRTIs is to develop compounds, which can inhibit mutant forms

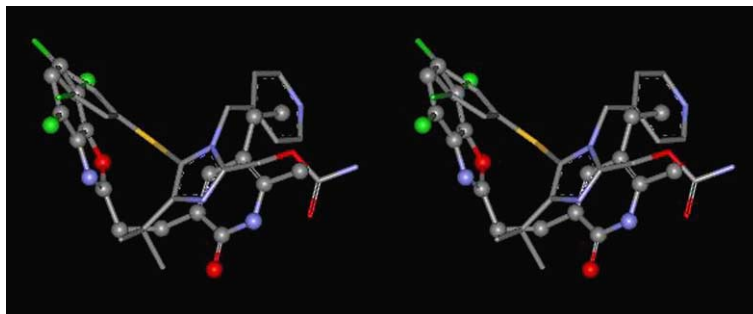


Figure 15. Top ranked binding mode of the most populated cluster of **2** (ball and sticks) into the binding site of S-1153. S-1153 (sticks) is displayed for comparison. Hydrogens are omitted for clarity.

of the enzyme. Some NNRTIs show greater resilience to drug resistance mutations within HIV-1 RT compared with first generation drugs such as nevirapine, 1051U91, and Cl- α -APA. Some of these compounds, so-called second-generation inhibitors, are efavirenz, UC-781, and S-1153.

It has been proposed that the main contribution to drug resistance for the Tyr181Cys mutation is the loss of aromatic ring stacking interactions between Tyr181 and an aromatic system, associated with a ‘butterfly wing’, of several NNRTIs.⁴² A good example of such an aromatic system is the benzoxazol ring of pyridinone derivatives that makes contacts with Tyr181, however, in the case of UC-781, the (3-methylbut-2-enyl)oxy group makes fewer significant interactions with Tyr181 than an aromatic system and it is thought to be responsible for the activity of this compound against RT containing the single point mutation Tyr181Cys.²⁹ Comparing the binding mode of the top ranked conformation of **1–4** of the most populated cluster obtained within the pocket of UC-781 with the binding mode of UC-781 (Fig. 13) it is observed that the benzoxazol ring of the pyridinone derivatives may be replaced by a (3-methylbut-2-enyl)-oxy group. This may lead to pyridinones derivatives active against the RT containing the mutation Tyr181Cys. The (3-methylbut-2-enyl)oxy group can be at position 4 of the pyridinone ring. A dimethylamine or isopropyl group at position-3 can play the steric role of the ethylene or aminomethylene linker of pyridinone derivatives, as previously shown for **6** (Fig. 14).

The carbamate group of S-1153 has the ability to form an extensive net of hydrogen bonds with surrounding amino acids, thus making S-1153 active against RT containing the single point mutation Lys103Asn.²³ Comparing the top ranked conformation of **1–4** of the most populated cluster obtained within the binding pocket of S-1153 with the binding mode of the crystal inhibitor it is observed that the methyl group at the pyridinone ring shares the same sub-pocket volume as the carbamate group of S-1153 (Fig. 15). This observation suggests that substituting the position-6 of the pyridinone ring with a carbamate group or other group with the ability to form hydrogen bonds may lead to pyridinone derivatives active against RT containing the Lys103Asn mutation. The activity against RT containing the Tyr181Cys mutation may also be achieved by replacing

the benzoxazol ring with a phenylthio or benzyl group *m*-substituted with methyl or chlorine groups. This strategy to overcome the Tyr181Cys mutation has already been successful for other NNRTIs.^{15,16,43} Similar to the pyridinone–UC-781 design proposed above, a dimethylamine or isopropyl group at position 3 can play the steric role of the ethylene or aminomethylene linker.

The docking of the pyridinone–UC-781 and pyridinone–S-1153 hybrids in Figure 14 into several NNRTIs binding pockets shows that the hybrid molecule can adopt the expected conformation within these pockets. The proposed molecules represent potential alternatives to RT inhibitors that still present resistance issues in clinical studies. The modified pyridinones could be less toxic than some current second-generation inhibitors such as S-1153, which showed negative results in animal toxicity studies.¹ Also, the novel molecules will provide information for a better understanding of the drug resistant mechanisms and are examples of structure based designed compounds.

4. Conclusions

Automated flexible docking was used to study the binding mode of four pyridinone derivatives in the NNRTIs binding site. A single docking model was derived considering the pocket geometries of nevirapine, 1051U91, 9-Cl-TIBO, Cl- α -APA, efavirenz, UC-781, and S-1153. According to the docking model proposed, that agrees with the mutational and structure–activity relationship data available for pyridinone derivatives, the pyridinone analogues have a very similar binding mode to that of the crystallographic NNRTIs. The docking model helped to understand, at the molecular level, the biological activity reported for hybrid pyridinone–NNRTIs molecules. A different binding mode was predicted for the pyridinone derivatives within the pocket geometries of MKC-442 and TNK-651. The conformation of the Pro236 loop in these binding geometries, which is a consequence of the particular structure of HEPT derivatives, is thought to be the responsible to modify the docking predictions. Comparison of the predicted binding mode of pyridinone analogues with the actual binding mode of second generation NNRTIs suggests substitutions at the pyridinone ring that may lead to compounds active against RT containing mutations.

5. Experimental

Sybyl 6.8⁴⁴ was used for ligand and protein preparation. Molecular structure viewing was carried out with Sybyl 6.8 and Viewer Pro 4.2.⁴⁵ RT crystal structures were obtained from the Protein Data Bank.⁴⁶ Docking calculations were conducted with AutoDock 3.0.³¹ In short, AutoDock performs an automated docking of the ligand with user-specified dihedral flexibility within a protein rigid binding site. The program performs several runs in each docking experiment. Each run provides one predicted binding mode.

All water molecules and magnesium ion, when present, were removed from the original Protein Data Bank files. The binding sites were separately superimposed before docking onto the structure of the RT–nevirapine complex, the structure with the highest resolution available, to compensate for differences in crystal form and domain orientations based on the ‘core’ of the p66 palm domain (residues 94–118, 156–215, and 225–243), residues 317–319 from the p66 connection domain, and residues 137–139 from the p51 fingers domain using Sybyl 6.8.

Polar hydrogen atoms were added and Kollman charges,⁴⁷ atomic solvation parameters and fragmental volumes were assigned to the protein. For validation of the docking protocol, ligand coordinates in the crystal complexes were removed and the bond orders were checked. The structure of 1051U91 in the PDB file 1RTH was corrected with Sybyl 6.8. For all ligands, Gasteiger charges⁴⁸ were assigned and non-polar hydrogen atoms merged. All torsions were allowed to rotate during docking.

The coordinates of the best docked conformation found for **1–4** within the pocket of nevirapine were used as starting coordinates to dock these compounds into the pockets of the other NNRTIs. To conduct docking validations the initial coordinates of NNRTIs were as found within the crystal.

The auxiliary program AutoGrid generated the grid maps. Each grid was centered at the crystal structure of the corresponding NNRTI. The grid dimensions were $23 \times 23 \times 23 \text{ \AA}^3$ with points separated by 0.375 \AA . Lenard-Jones parameters 12–10 and 12–6, supplied with the program, were used for modeling H-bonds and van der Waals interactions, respectively. The distance-dependent dielectric permittivity of Mehler and Solmajer⁴⁹ was used for calculation of the electrostatic grid maps. For all ligands, random starting positions, random orientations, and torsions were used. The translation, quaternion, and torsion steps were taken from default values in AutoDock. The Lamarckian genetic algorithm and the pseudo-Solis and Wets methods were applied for minimization using default parameters. The number of docking runs was 100. The population in the genetic algorithm was 50, the energy evaluations were 250,000 and the maximum number of iterations 27,000.

After docking, the 100 solutions were clustered into groups with RMS deviations lower than 1.0 \AA .

The clusters were ranked by the lowest energy representative of each cluster.

In order to describe pyridinone-binding pocket interactions, the top ranked binding mode found by AutoDock for **1–4** in complex with the binding pocket of nevirapine was subject to full energy minimization utilizing Gasteiger–Hückel partial charges⁵⁰ and the Tripos force field⁵¹ until the gradient of 0.05 kcal/mol was reached. During minimization, atoms within 6 \AA from the ligand were free to move (other atoms were fixed).

Acknowledgements

J.L.M.-F. is grateful to CONACyT and DGAPA, UNAM, for the Ph.D. scholarships. We also would like to acknowledge CONACyT for financing project G-34851-M.

References and notes

- De Clercq, E. *Int. J. Biochem. Cell Biol.* **2004**, *36*, 1800.
- Jacobo-Molina, A.; Arnold, E. *Biochemistry* **1991**, *30*, 6351.
- Goldman, M. E.; Nunberg, J. H.; O'Brien, J. A.; Quintero, J. C.; Schleif, W. A.; Freund, K. F.; Gaul, S. L.; Saari, W. S.; Wai, J. S.; Hoffman, J. M.; Anderson, P. S.; Hupe, D. J.; Emini, E. A.; Stern, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 6863.
- Nunberg, J. H.; Schleif, W. A.; Boots, E. J.; O'Brien, J. A.; Quintero, J. C.; Hoffman, J. M.; Emini, E. A.; Goldman, M. E. *J. Virol.* **1991**, *65*, 4887.
- Saari, W. S.; Hoffman, J. M.; Wai, J. S.; Fisher, T. E.; Rooney, C. S.; Smith, A. M.; Thomas, C. M.; Goldman, M. E.; O'Brien, J. A.; Nunberg, J. H.; Quintero, J. C.; Schleif, W. A.; Emini, E. A.; Stern, A. M.; Anderson, P. S. *J. Med. Chem.* **1991**, *34*, 2922.
- Hoffman, J. M.; Wai, J. S.; Thomas, C. M.; Levin, R. B.; O'Brien, J. A.; Goldman, M. E. *J. Med. Chem.* **1992**, *35*, 3784.
- Saari, W. S.; Wai, J. S.; Fisher, T. E.; Thomas, C. M.; Hoffman, J. M.; Rooney, C. S.; Smith, A. M.; Jones, J. H.; Bamberger, D. L.; Goldman, M. E.; O'Brien, J. A.; Nunberg, J. H.; Quintero, J. C.; Schleif, W. A.; Emini, E. A.; Anderson, P. S. *J. Med. Chem.* **1992**, *35*, 3792.
- Hoffman, J. M.; Smith, A. M.; Rooney, C. S.; Fisher, T. E.; Wai, J. S.; Thomas, C. M.; Bamberger, D.; Barnes, J. L.; Williams, T. M.; Jones, J. H.; Olson, B. D.; O'Brien, J. A.; Goldman, M. E.; Nunberg, J. H.; Quintero, J. C.; Schleif, W. A.; Emini, E. A.; Anderson, P. S. *J. Med. Chem.* **1993**, *36*, 953.
- Davey, R. T.; Dewar, R. L.; Reed, G. F.; Vasudevachari, M. B.; Polis, M. A.; Kovacs, J. A.; Fallon, J.; Walker, R. E.; Masur, H.; Haneiwich, S. E.; O'Neill, D. G.; Decker, M. R.; Metcalf, J. A.; Deloria, M. A.; Laskin, O. L.; Salzman, N.; Lane, H. C. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5608.
- Saag, M. S.; Emini, E. A.; Laskin, O. L.; Douglas, J.; Lapidus, W. I.; Schleif, W. A.; Whitley, R. J.; Hildebrand, C.; Byrnes, V. W.; Kappes, J. C.; Anderson, K. W.; Massari, F. E.; Shaw, G. M.; L-697,661 Working Group. *New Engl. J. Med.* **1993**, *329*, 1065.
- Sardana, V. V.; Emini, E. A.; Gotlib, L.; Graham, D. J.; Lineberger, D. W.; Long, W. J.; Schlabach, A. J.;

- Wolfgang, J. A.; Condra, J. H. *J. Biol. Chem.* **1992**, *267*, 17526.
12. Byrnes, V. W.; Sardana, V. V.; Schleif, W. A.; Condra, J. H.; Waterbury, J. A.; Wolfgang, J. A.; Long, W. J.; Schneider, C. L.; Schlabach, A. J.; Wolanski, B. S.; Graham, D. J.; Gotlib, L.; Rhodes, A.; Titus, D. L.; Roth, E.; Blahy, O. M.; Quintero, J. C.; Staszewski, S.; Emini, E. A. *Antimicrob. Agents Chemother.* **1993**, *37*, 1576.
13. Garg, R.; Gupta, S. P.; Gao, H.; Babu, M. S.; Debnath, A. K.; Hansch, C. *Chem. Rev.* **1999**, *99*, 3525.
14. Gupta, S. P. In *Progress in Drug Research*; Jucker, E., Ed.; Birkhäuser: Basel, Switzerland, 2002; Vol. 58, pp 252–253.
15. Dollé, V.; Fan, E.; Nguyen, C. H.; Aubertin, A. M.; Kirn, A.; Andreola, M. L.; Jamieson, G.; Tarrago-Litvak, L.; Bisagni, E. *J. Med. Chem.* **1995**, *38*, 4679.
16. Dollé, V.; Nguyen, C. H.; Legraverend, M.; Aubertin, A. M.; Kirn, A.; Andreola, M. L.; Ventura, M.; Tarrago-Litvak, L.; Bisagni, E. *J. Med. Chem.* **2000**, *43*, 3949.
17. Corbett, J. W.; Kresge, K. J.; Pan, S.; Cordova, B. C.; Klabe, R. M.; Rodgers, J. D.; Erickson-Viitanen, S. K. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 309.
18. Esnouf, R.; Ren, J.; Ross, C.; Jones, Y.; Stammers, D.; Stuart, D. *Nat. Struct. Biol.* **1995**, *2*, 303.
19. Ren, J.; Esnouf, R.; Garman, E.; Somers, D.; Ross, C.; Kirby, I.; Keeling, J.; Darby, G.; Jones, Y.; Stuart, D.; Stammers, D. *Nat. Struct. Biol.* **1995**, *2*, 293.
20. Hopkins, A. L.; Ren, J.; Esnouf, R. M.; Willcox, B. E.; Jones, E. Y.; Ross, C.; Miyasaka, T.; Walker, R. T.; Tanaka, H.; Stammers, D. K.; Stuart, D. I. *J. Med. Chem.* **1996**, *39*, 1589.
21. Ren, J.; Esnouf, R.; Hopkins, A.; Ross, C.; Jones, Y.; Stammers, D.; Stuart, D. *Structure* **1995**, *3*, 915.
22. Ren, J.; Milton, J.; Weaver, K. L.; Short, S. A.; Stuart, D. I.; Stammers, D. K. *Structure* **2000**, *8*, 1089.
23. Ren, J.; Nichols, C.; Bird, L. E.; Fujiwara, T.; Sugimoto, H.; Stuart, D. I.; Stammers, D. K. *J. Biol. Chem.* **2000**, *275*, 14316.
24. Esnouf, R. M.; Ren, J.; Hopkins, A. L.; Ross, C. K.; Jones, E. Y.; Stammers, D. K.; Stuart, D. I. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 3984.
25. Ren, J.; Diprose, J.; Warren, J.; Esnouf, R. M.; Bird, L. E.; Ikemizu, S.; Slater, M.; Milton, J.; Balzarini, J.; Stuart, D. I.; Stammers, D. K. *J. Biol. Chem.* **2000**, *275*, 5633.
26. Chan, J. H.; Hong, J. S.; Hunter, R. N., III; Orr, G. F.; Cowan, J. R.; Sherman, D. B.; Sparks, S. M.; Reitter, B. E., III; Andrews, C. W., III; Hazen, R. J.; St Clair, M.; Boone, L. R.; Ferris, R. G.; Creech, K. L.; Roberts, G. B.; Short, S. A.; Weaver, K.; Ott, R. J.; Ren, J.; Hopkins, A.; Stuart, D. I.; Stammers, D. K. *J. Med. Chem.* **2001**, *44*, 1866.
27. Hsiou, Y.; Das, K.; Ding, J.; Clark, A. D., Jr.; Kleim, J. P.; Rosner, M.; Winkler, I.; Riess, G.; Hughes, S. H.; Arnold, E. *J. Mol. Biol.* **1998**, *284*, 313.
28. Ren, J.; Esnouf, R. M.; Hopkins, A. L.; Stuart, D. I.; Stammers, D. K. *J. Med. Chem.* **1999**, *42*, 3845.
29. Ren, J.; Esnouf, R. M.; Hopkins, A. L.; Warren, J.; Balzarini, J.; Stuart, D. I.; Stammers, D. K. *Biochemistry* **1998**, *37*, 14394.
30. Hogberg, M.; Sahlberg, C.; Engelhardt, P.; Noreen, R.; Kangasmetsa, J.; Johansson, N. G.; Oberg, B.; Vrang, L.; Zhang, H.; Sahlberg, B. L.; Unge, T.; Lovgren, S.; Fridborg, K.; Backbro, K. *J. Med. Chem.* **2000**, *43*, 304.
31. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998**, *19*, 1639.
32. Titmuss, S. J.; Keller, P. A.; Griffith, R. *Bioorg. Med. Chem.* **1999**, *7*, 1163.
33. Barreca, M. L.; Carotti, A.; Carrieri, A.; Chimiri, A.; Monforte, A. M.; Calace, M. P.; Rao, A. *Bioorg. Med. Chem.* **1999**, *7*, 2283.
34. Zhou, Z.; Madrid, M.; Madura, J. D. *Proteins* **2002**, *49*, 529.
35. Chen, H. F.; Yao, X. J.; Li, Q.; Yuan, S. G.; Panaye, A.; Doucet, J. P.; Fan, B. T. *SAR QSAR Environ. Res.* **2003**, *14*, 455.
36. Kontoyianni, M.; McClellan, L. M.; Sokol, G. S. *J. Med. Chem.* **2004**, *47*, 558.
37. Yadav, P. N. S.; Das, K.; Ding, J.; Arnold, E.; Yadav, J. S.; Modak, M. J. *J. Mol. Struct. (Theochem)* **1998**, *423*, 101.
38. Mao, C.; Subeck, E. A.; Venkatachalam, T. K.; Uckun, F. M. *Biochem. Pharmacol.* **2000**, *60*, 1251.
39. The binding pocket of nevirapine was selected to study pyridinone–RT interactions because it is the structure with the highest resolution available and represent well the binding mode of the pyridinones within the binding geometries of 1051U91, 9-Cl-TIBO, Cl- α -APA, efavirenz, UC-781, and S-1153.
40. Parreira, R. L. T.; Abrahão, O.; Galembeck, S. E. *Tetrahedron* **2001**, *57*, 3243.
41. Tantillo, C.; Ding, J.; Jacobo-Molina, A.; Nanni, R. G.; Boyer, P. L.; Hughes, S. H.; Pauwels, R.; Andries, K.; Janseen, P. A. J.; Arnold, E. *J. Mol. Biol.* **1994**, *243*, 369.
42. Ren, J.; Nichols, C.; Bird, L.; Chamberlain, P.; Weaver, K.; Short, S.; Stuart, D. I.; Stammers, D. K. *J. Mol. Biol.* **2001**, *312*, 795.
43. Hopkins, A. L.; Ren, J.; Tanaka, H.; Baba, M.; Okamoto, M.; Stuart, D. I.; Stammers, D. K. *J. Med. Chem.* **1999**, *42*, 4500.
44. Sybyl version 6.8; Tripos, Inc., St. Louis, MO.
45. ViewerPro version 4.2; Accelrys, Inc, San Diego, CA.
46. Brookhaven Protein Data Bank. <http://www.rcsb.org>.
47. Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765.
48. Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219.
49. Mehler, E. L.; Solmajer, T. *Protein Eng.* **1991**, *4*, 903.
50. Streitwieser, A. *Molecular Orbital Theory for Organic Chemists*; Wiley: New York, 1961.
51. Clark, M.; Cramer, R. D., III; Van Opdenbosch, N. *J. Comput. Chem.* **1989**, *10*, 982.

Docking-based CoMFA and CoMSIA studies of non-nucleoside reverse transcriptase inhibitors of the pyridinone derivative type

J.L. Medina-Franco*, S. Rodríguez-Morales, C. Juárez-Gordiano, A. Hernández-Campos & R. Castillo**

Departamento de Farmacia, Facultad de Química, UNAM, CU, DF 04510, Mexico

Received 19 April 2004; accepted in revised form 26 July 2004

Key words: AutoDock, CoMFA, CoMSIA, 3D-QSAR, docking, HIV-1, non-nucleoside inhibitors, reverse transcriptase

Summary

Comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) were performed on a set of pyridinone derivatives. A molecular alignment obtained by docking of compounds into the non-nucleoside reverse transcriptase inhibitor binding site of HIV-1 was used. Good correlations between the calculated binding free energies and experimental inhibitory activities suggest that the binding conformations of these inhibitors are reasonable. Robust and predictive 3D-QSAR models were obtained with q^2 values of 0.706 and 0.723 for CoMFA and CoMSIA, respectively. The models were validated by an external test set obtaining r^2_{pred} values of 0.720 and 0.750 for CoMFA and CoMSIA, respectively. The CoMFA, CoMSIA and docking results help to understand the type of interactions that occur between pyridinone derivatives with the non-nucleoside reverse transcriptase inhibitor binding pocket, and explain the viral resistance to pyridinone derivatives upon mutation of amino acids Tyr181 and Tyr188. The results obtained provide information for a better understanding of the drug resistant mechanisms. The 3D-QSAR models derived will be used to guide the design of pyridinone derivatives active against mutant strains of reverse transcriptase.

Abbreviations: AIDS, acquired immunodeficiency syndrome; CoMFA, comparative molecular field analysis; CoMSIA, comparative molecular similarity indices analysis; HIV, human immunodeficiency virus; NNRTIs, non-nucleoside reverse transcriptase inhibitors; PLS, partial least squares; QSAR, quantitative structure-activity relationships; RT, reverse transcriptase.

Introduction

The reverse transcriptase (RT) of human immunodeficiency virus type 1 (HIV-1) is an attractive target to treat the acquired immune deficiency syndrome (AIDS) for which no absolute successful chemotherapy is available to date [1]. RT inhibitors are classified in nucleoside and non-nucleoside depending on their mechanism of action. One advantage of the second group is that they lack the toxic effects associated

with the nucleosides. To date, three non-nucleoside RT inhibitors (NNRTIs) have been approved for clinical use named nevirapine (VIRAMUNE®) (Figure 1), delavirdine (RESCRIPTOR®) and efavirenz (SUSTIVA®) [1]. However, all these and other approved drugs induce drug resistant variants of HIV-1 [2].

Several crystal structures of HIV-1 RT in the apo form [3] and complexed with different NNRTIs have been published [4]. The non-nucleoside inhibitor binding pocket of HIV-1 RT is formed by the amino acids Pro95, Leu100, Lys101, Lys103, Val106, Val179, Tyr181, Tyr188, Gly190, Phe227, Trp229, Leu234,

*This work is part of the Ph.D. thesis of J.L. Medina-Franco.

**To whom correspondence should be addressed. Fax: +52-5622-5329; E-mail: rafaelc@servidor.unam.mx

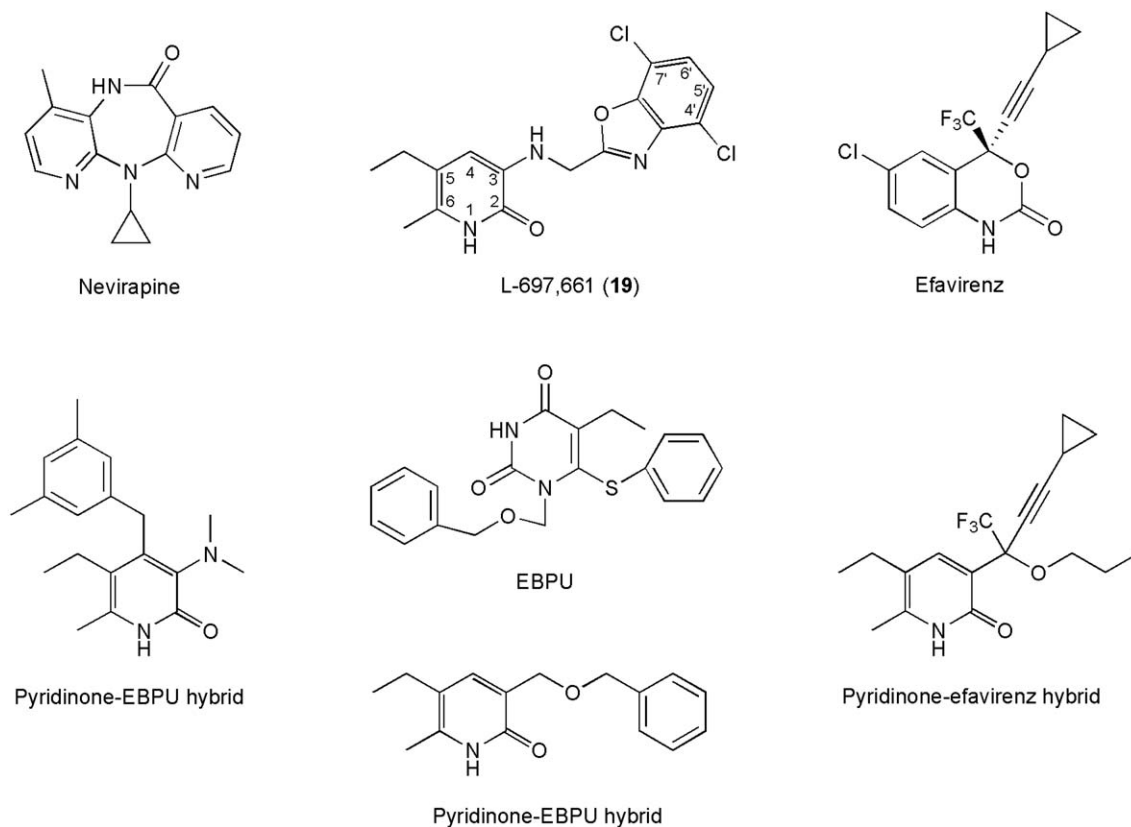


Figure 1. Chemical structures of nevirapine, L-697,661, efavirenz, EBPU, pyridinone-EBPU and pyridinone-efavirenz.

His235, Pro236 and Tyr318 of the p66 subunit and Glu138 of the p51 subunit. Crystal structures of RT-NNRTIs complexes reveal that some amino acids of the binding pocket adopt different conformations when bound to different inhibitors, leading to distinct pocket volumes [4b, 4f, 4i]. Despite these variations, several NNRTIs have a common binding mode with a butterfly-like conformation [4].

Pyridinone derivatives [5] are a class of NNRTIs [2]. No crystal structures of HIV-1 RT in complex with pyridinone derivatives have been published. However, several two-dimensional quantitative structure-activity relationship analyses (2D-QSAR) [6, 7] have been performed for pyridinone analogues suggesting a hydrophobic character of their binding pocket. Furthermore, from kinetic [5] and analysis of resistance mutations [8] it was concluded that this type of inhibitors should be binding in the same pocket as other NNRTIs. Extensive structure-activity relationships (SAR) studies have been conducted for these compounds leading to potent RT inhibitors [9–12]. Compound L-697,661 (**19**) (Figure 1) was subjected to clinical studies

showing good activity in patients, but rapid resistant strains of the virus emerged [13, 14]. The Tyr181Cys, Tyr188Cys and Lys103Asn mutations are responsible for the emergence of resistant virus not only for **19** but also for other pyridinone derivatives [14–16]. The Tyr181Cys mutation is the most important. Interesting to note, hybrid pyridinone-NNRTIs molecules have shown inhibitory activity against wild type and some mutant strains of RT (Figure 1) [17–20].

Several studies employing different QSAR methodologies have been performed on NNRTIs. Examples of recent works can be found in Refs. 21 and 22. Three-dimensional quantitative structure-activity relationship (3D-QSAR) methods such as comparative molecular field analysis (CoMFA) [23] and comparative molecular similarity indices analysis (CoMSIA) [24] have been successfully applied in many instances to guide the design of new active molecules [25]. For several NNRTIs, 3D-QSAR studies have helped to identify the type of interactions between the inhibitors and RT [26–30]. However, these methodologies have not been applied to pyridinone derivatives. Also,

molecular docking has been applied to understand the binding mode of NNRTIs and design novel molecules [26, 30–33]. We recently performed a docking study of potent pyridinone RT inhibitors and pyridinone hybrid molecules into the NNRTI binding pocket, proposing a binding mode for these molecules [34].

We now report in this paper CoMFA and CoMSIA studies of pyridinone derivatives based on the docking positions determined by the automated docking program AutoDock. The combination of ligand-based with structure-based drug design approaches is an attractive strategy for pyridinone derivatives for which the binding site is known but the binding mode has not been determined experimentally. Successful studies applying a combination of ligand-based approaches with structure-based approaches have been reported recently [30, 35–38]. The purposes of this work are (a) to investigate the type of interactions that occur between the pyridinone derivatives and the NNRTI binding pocket; and (b) to identify the structural requirements for HIV-1 RT inhibitory activity of pyridinone derivatives. 3D-QSAR results will further validate the docking models.

Computational methods

Data set and biological data

A set of 40 pyridinone derivatives taken from literature [11, 12] was divided into a training set of 29 compounds and a test set of 11 compounds, sampling from various ranges of logarithmic unit of activities. The chemical structures of the training and test sets are depicted in Tables 1 and 2, respectively. The experimental RT inhibitory activities were estimated under the same experimental conditions. Activities were converted into the corresponding $-\log IC_{50}$ values, where IC_{50} is the effective concentration of compound required to achieve 50% of inhibition of RT. The observed biological activities are given in Tables 1 and 2.

Molecular docking

Automated docking was used to locate the appropriate binding orientations and conformations of pyridinone derivatives into the NNRTI binding pocket. The powerful genetic algorithm method implemented in the program AutoDock 3.0 [39] was employed. RT in complex with nevirapine [4a], the structure with the highest resolution available, was used. Furthermore,

in a previous docking study of pyridinone derivatives into several NNRTI binding pocket geometries it was concluded that the binding pocket of nevirapine is a good model to study the binding mode of pyridinone analogues [34]. The structure was retrieved from the RCSB Protein Data Base (PDB entry 1VRT) [40].

The structures of pyridinone derivatives and RT were prepared using Sybyl 6.8 [41]. Molecular structures of pyridinone derivatives were built using the SKETCH option in Sybyl. Geometry optimizations were performed using the Tripos force field [42] with a distance-dependent dielectric and the Powell conjugate gradient algorithm. Gasteiger-Hückel charges [43] were used.

All water molecules and magnesium ions were removed from the original Protein Data Bank file. Polar hydrogen atoms were added and Kollman charges [44], atomic solvation parameters and fragmental volumes were assigned to the protein using AutoDock Tools (ADT). For validation of the docking protocol, nevirapine coordinates in the crystal complex were removed and the bond orders were checked. For docking calculations, Gasteiger partial charges [45] were assigned to the pyridinone derivatives and nevirapine, and non-polar hydrogen atoms were merged. All torsions were allowed to rotate during docking.

The auxiliary program AutoGrid generated the grid maps. Each grid was centered at the crystal structure of the corresponding NNRTI. The grid dimensions were $23 \times 23 \times 23 \text{ \AA}^3$ with points separated by 0.375 \AA . Lennard-Jones parameters 12-10 and 12-6, supplied with the program, were used for modeling H-bonds and van der Waals interactions, respectively. The distance-dependent dielectric permittivity of Mehler and Solmajer [46] was used for calculation of the electrostatic grid maps. For all ligands, random starting positions, random orientations and torsions were used. The translation, quaternion and torsion steps were taken from default values in AutoDock. The Lamarckian genetic algorithm and the pseudo-Solis and Wets methods were applied for minimization using default parameters. The number of docking runs was 100. The population in the genetic algorithm was 50, the energy evaluations were 250 000 and the maximum number of iterations 27 000.

After docking, the 100 solutions were clustered into groups with RMS deviations lower than 1.0 \AA . The clusters were ranked by the lowest energy representative of each cluster.

The complexes of pyridinone derivatives with HIV-1 RT resulting from molecular docking were further

Table 1. Structures, HIV-1 RT inhibitory activity and AutoDock binding free energies of compounds in the training set.

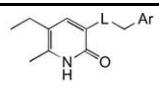

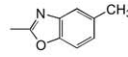

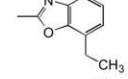

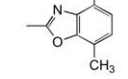

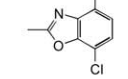

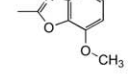

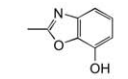

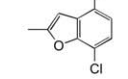











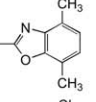

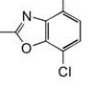

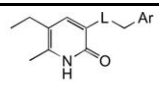
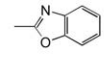

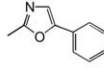

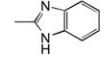

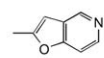

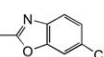

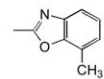
									
Comp.	L	Ar	-LogIC ₅₀	ΔG (kcal/mol)	Comp.	L	Ar	-LogIC ₅₀	ΔG (kcal/mol)
1	NH		6.55	-9.34	16	NH		5.90	-9.02
2	NH		6.46	-9.25	17	NH		6.59	-9.65
3	NH		6.34	-9.53	18	NH		7.7	-10.01
4	NH		6.30	-9.24	19	NH		7.72	-9.83
5	NH		6.28	-9.56	20	NH		6.74	-9.31
6	NH		5.96	-9.29	21	NH		6.36	-9.09
7	NH		5.72	-8.70	22	NH		7.24	-9.75
8	NH		5.60	-9.51	23	CH ₂		4.30	-7.74
9	NH		5.57	-10.27	24	CH ₂		5.61	-9.94
10	NH		5.36	-9.10	25	CH ₂		6.43	-9.68
11	NH		5.27	-7.59	26	CH ₂		7.24	-9.72
12	NH		4.82	-7.55	27	CH ₂		6.41	-9.32
13	NH		4.65	-8.55	28	CH ₂		7.55	-10.20
14	NH		3.98	-8.85	29	CH ₂		7.85	-10.18
15	NH		6.92	-9.55					

Table 2. Structures, HIV-1 RT inhibitory activity and AutoDock binding free energies of compounds in the test set.

									
Comp.	L	Ar	-LogIC ₅₀	ΔG (kcal/mol)	Comp.	L	Ar	-LogIC ₅₀	ΔG (kcal/mol)
30	NH		6.68	-8.83	36	NH		6.47	-9.40
31	NH		5.63	-9.05	37	NH		5.71	-9.47
32	NH		5.12	-8.97	38	CH ₂		5.08	-10.00
33	NH		4.49	-8.88	39	CH ₂		3.59	-7.90
34	NH		5.78	-9.06	40	CH ₂		4.87	-9.22
35	NH		7.26	-9.50					

structurally optimized using Gasteiger-Hückel partial charges and the Tripos force field. During minimization, atoms within 6 Å from the ligand were free to move (other atoms were fixed).

3D-QSAR studies

CoMFA and CoMSIA studies were performed based on the conformational alignment predicted from the molecular docking. Pyridinone derivatives with partial Gasteiger-Hückel charges or partial atomic charges based on the AM1 semiempirical molecular orbital method [47] implemented in Sybyl were used.

CoMFA. A 3D cubic lattice with grid spacing of 2 or 1 Å and extending 4 Å units beyond the aligned molecules in all directions was created. Steric and electrostatic interactions were calculated with the Tripos force field using a sp³ carbon probe atom with a charge of +1.0 with a distance-dependent dielectric at each lattice point. CoMFA descriptor fields were also calculated with a sp³ oxygen with -1.0 charge and a hydrogen atom with +1.0 charge as probe atoms. The minimum-sigma (column filtering) was set to 2.0 kcal/mol to improve the signal-to-noise ratio by

omitting those lattice points whose energy variation was below this threshold. A cutoff of 30 kcal/mol was adopted. The CoMFA steric and electrostatic fields generated were scaled by the CoMFA-STD method in Sybyl.

The CoMFA steric and electrostatic fields were used as independent variables and -logIC₅₀ values were used as dependent variables in the partial least squares (PLS) regression analyses to derive CoMFA models. The overall predictive ability of the analysis was evaluated first by leave-one-out (LOO) cross-validation. The cross-validated coefficient q^2 was calculated according to the following equation:

$$q^2 = (\text{SSY} - \text{PRESS})/\text{SSY}$$

SSY stands for the variance of the biological activities around the mean value, and PRESS is the prediction error sum of the squares, derived from the LOO method. The uncertainty of the prediction is defined as

$$\text{SPRESS} = [\text{PRESS}/(n - k - 1)]^{1/2}$$

where k is the number of variables in the model and n is the number of compounds used in the study. The number of components (noc) corresponding to the

lowest PRESS value was used for deriving the final PLS regression models. The conventional correlation coefficient r^2 was also computed as well as the conventional correlation coefficient for the prediction of compounds in the test set r^2_{pred} .

CoMSIA. The same grids constructed for the CoMFA fields calculation were used for CoMSIA fields calculation. The steric, electrostatic, hydrophobic, H-bond acceptor and H-bond donor similarity indices descriptors were calculated using a sp^3 carbon probe atom with a charge of +1.0 and a radius of 1.0 Å. The minimum-sigma was set to 2.0 kcal/mol. CoMSIA similarity indices (A_F) for a molecule j with atoms i at a grid point q are calculated by the following equation:

$$A_{F,k}^q(j) = -\sum \omega_{probe,k} \omega_{ik} e^{-\alpha r_{iq}^2}$$

where q is the grid point for molecule j ; ω_{ik} the actual value of physicochemical property k of atom i ; $\omega_{probe,k}$ indicates probe atom with charge +1, radius 1.0 Å, hydrophobicity +1, H-bond donor and acceptor property +1; α is the attenuation factor; r_{iq} is the mutual distance between probe atom at grid point q and atom i of the test molecule. Here, steric indices are related to the third power of the atomic radii, electrostatic descriptors are derived from atom-based parameters [48], and H-bond donor and acceptor indices are obtained by a rule-based method based on experimental results [49]. The default value of 0.3 was used as the attenuation factor for the Gaussian-type distance r_{iq} . The statistical evaluation for the CoMSIA analyses was performed in the same way as described for CoMFA.

Results and discussion

Docking

Before docking the pyridinone derivatives into the NNRTI binding site, the docking protocol was validated. Nevirapine was removed from the active site and docked back into the binding pocket. The root mean square deviation (RMSD) between the predicted conformation and the observed X-ray crystallographic conformation of nevirapine was 0.70 Å (Figure 2). Only one cluster was obtained with a mean docking energy of -9.66 kcal/mol. The top ranked binding conformation showed a docking energy of -9.67 kcal/mol and a free energy of binding of

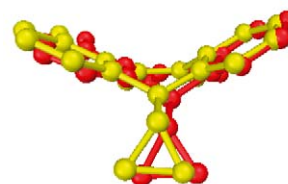


Figure 2. Comparison between the binding position of nevirapine found within the crystal structure (yellow) and the conformation predicted by AutoDock (red).

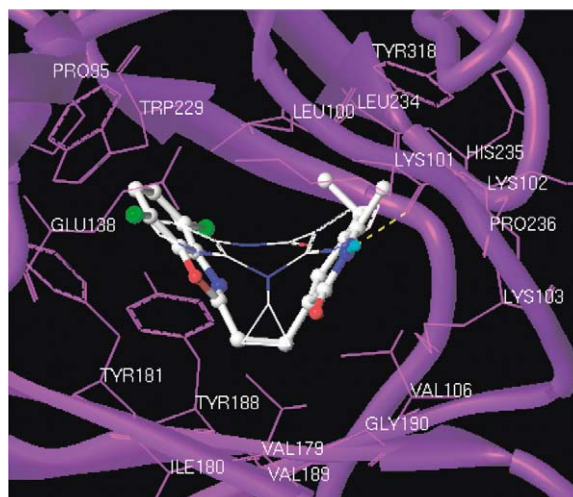


Figure 3. Optimized RT-29 complex. The hydrogen bond is displayed as yellow dashes. Amino acids within 5 Å of the benzoxazol and pyridinone rings are labeled. Nevirapine is displayed for comparison. Hydrogens are omitted for clarity except the polar hydrogen of the pyridinone ring.

-9.49 kcal/mol. These results indicated that the parameters used for AutoDock successfully reproduced the X-ray structure.

AutoDock found two possible binding modes for several pyridinone derivatives. The docked complexes of pyridinones-RT were selected according to the criteria of the most populated cluster combined with geometrical matching quality. In several cases the selected conformation corresponded to the lowest energy binding energy. According to this criterion all compounds adopt a butterfly-like conformation [50] into the NNRTI binding pocket and have the same orientation as nevirapine. The two aromatic rings ('wings') of the pyridinone derivatives may be roughly overlapped with the aromatic rings of nevirapine. The amino acid residues that interact with most of the docked pyridinone derivatives in this study were Pro95, Leu100, Lys101, Lys103, Tyr181, Tyr188 and Tyr318 of the p66 subunit (Figure 3). The corres-

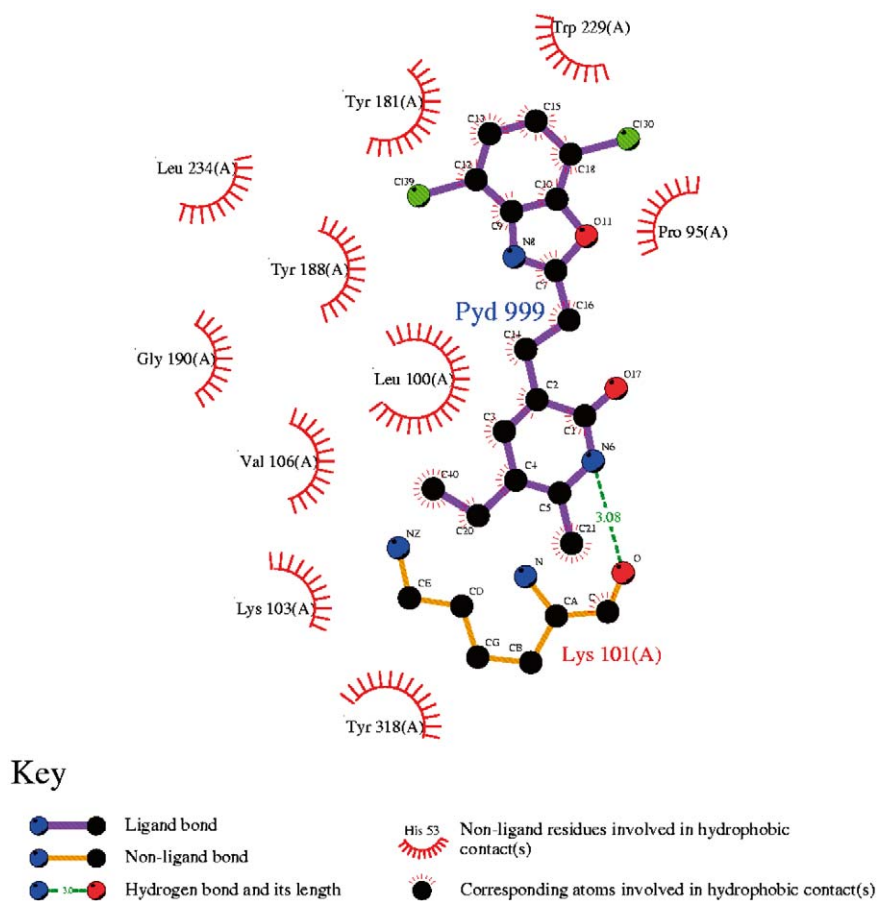


Figure 4. Two-dimensional representation of the main interactions of **29** with RT. The image was generated with the LIGPLOT program [51].

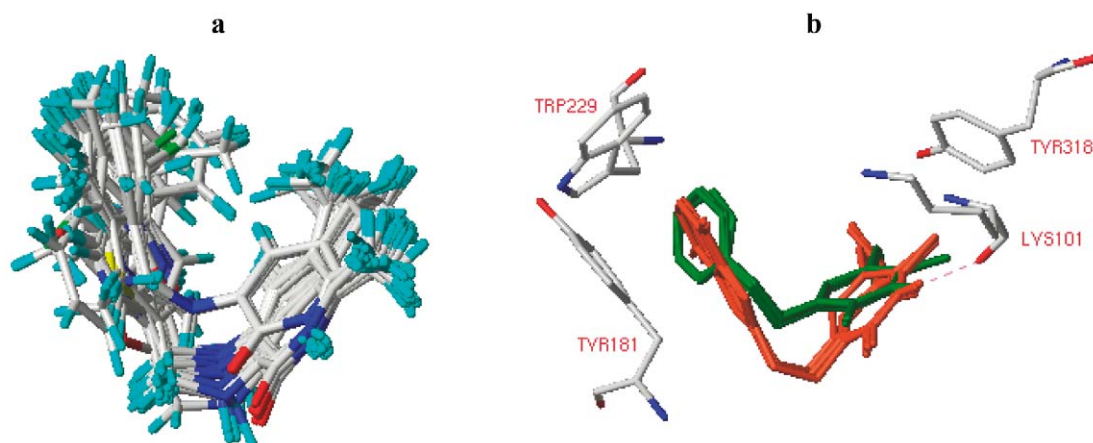


Figure 5. (a) Probable binding conformations of the pyridinone derivatives predicted by AutoDock and their alignment in the NNRTI binding site. (b) Comparison of the binding conformations of compounds **18**, **19**, **28** and **29** (orange) with **11**, **12**, **23** and **39** (green) in the binding site. The positions of Tyr181, Trp229, Tyr318 and Lys101 are displayed for reference. The hydrogen bond is displayed as magenta dashes. Hydrogens are omitted for clarity except for the polar hydrogen of the pyridinone ring.

ponding 2D interactions model for the most active compound is shown in Figure 4.

The molecular superposition of the binding conformations of the pyridinone derivatives extracted from the optimized inhibitor-RT complexes is shown in Figure 5a. All pyridinone derivatives have a common binding mode with a good overlapping of the pyridinone rings ('wing 1'). All docked compounds may form a hydrogen bond between the polar proton of the pyridinone ring and the main-chain carbonyl oxygen of Lys101 (Figures 3 and 4). A similar hydrogen bond to the main-chain carbonyl oxygen of Lys101 is observed for other NNRTIs [4b-d, g-i, k, l]. However, the 'wing 2' (the 'Ar' substituent in Tables 1 and 2) may adopt a different orientation depending on the chemical structure of the aromatic system (Figure 5b).

In a recent conformational analysis performed for pyridinone derivatives with different 'Ar' substituents [50], several minimum energy conformers were obtained for each pyridinone studied. The MMFF94 force field was used. Additional calculations based on the density functional theory were performed. The study revealed that the rotational barrier of the single bonds that connect the aromatic rings is small and that there is a marked preference for a butterfly-like conformation. We reached similar conclusions in a conformational analysis performed for compound **19** with Monte Carlo simulations and systematic search using molecular mechanics and semi-empirical methods (data not shown). For example, using semi-empirical calculations we observed that the highest rotational barrier of the benzoxazol ring is less than 1 kcal/mol. However, as part of the previous docking studies we conducted with pyridinone derivatives [34] starting from different minimum energy conformers, we observed that AutoDock predicts the same docked conformation regardless of the starting conformation. This indicates that AutoDock performs an efficient conformational search inside the binding pocket and suggests that the particular conformation predicted for each 'wing 2' is due to the specific interactions with the amino acids of the binding pocket.

Interestingly, the most active compounds have a benzoxazol ring. This two-ring system is close to Tyr181 (Figure 3). The parallel positions of the benzoxazol ring and the aromatic ring of Tyr181 suggest that they may be forming π -stacking interactions [52] and thus stabilizing the complex. Similar stabilizing interactions may occur for other pyridinone derivatives with a two-ring aromatic system in wing 2.

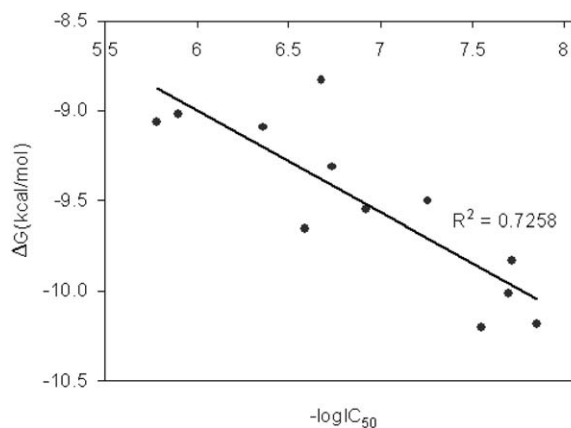


Figure 6. Plot of experimental activities vs. binding free energies (ΔG , $T=298.15$ K) for compounds with close structural similarity (**15–21**, **28**, **29**, **30**, **34**, **35**).

Compounds **11**, **12**, **23** and **39**, which have only one ring in the wing 2, adopt a docked position and conformation that is slightly different from the other pyridinone derivatives; the linker between the two wings is displaced upwards relative to the linker of the other pyridinone analogues (Figure 5b). Moreover, despite the fact that the wing 2 is close to Tyr181, the non-parallel positions of the phenyl group of Tyr181 and the wing 2 suggest weaker interactions between these systems (Figure 5b) [52]. This explains, at least in part, the low biological activity and the calculated high binding free energy of these compounds (Tables 1 and 2). Similarly, the low activity and the high binding free energy of **13** may be explained by the lack of contacts with Tyr181.

The predicted binding free energies of the pyridinone derivatives binding to RT are also listed in Tables 1 and 2. In general, the binding free energy calculated by AutoDock correlated well with the biological activity ($-\log IC_{50}$); compounds with the highest activities showed the lowest free energies. This suggests that the binding models for the pyridinone derivatives with RT are reasonable. A quantitative correlation between the binding free energy and the biological activity was better for structurally similar compounds, i.e., all compounds with a benzoxazol ring (Figure 6). All these compounds adopt a very similar binding mode.

3D-QSAR models

PLS results for the CoMFA and CoMSIA studies are summarized in Table 3. Figure 7 shows the prediction curves obtained with the CoMFA and CoMSIA mod-

Table 3. CoMFA and CoMSIA PLS statistics.^a

Charges	Step size	PLS statistics	29 Compds. model		28 Compds. model	
			CoMFA	CoMSIA	CoMFA	CoMSIA
AM1	2 Å	q^2	0.656	0.643	0.706	0.722
		PRESS	0.639	0.651	0.603	0.586
		Noc	4	4	4	4
		r^2	0.943	0.926	0.954	0.943
		S	0.259	0.295	0.239	0.266
		F	100.055	75.566	119.076	94.725
		r^2_{pred}	0.688	0.785	0.720	0.759
		Field contribution				
		Steric	0.608	0.121	0.600	0.122
	Electrostatic	0.392	0.421	0.400	0.434	
	Hydrophobic		0.457		0.445	
	1 Å	q^2	0.613	0.643	0.650	0.723
		PRESS	0.664	0.651	0.644	0.585
		Noc	3	4	3	4
		r^2	0.934	0.932	0.947	0.946
		S	0.275	0.284	0.251	0.259
		F	117.468	82.484	142.534	100.113
		r^2_{pred}	0.693	0.778	0.731	0.750
Field contribution						
Steric		0.563	0.121	0.559	0.121	
Electrostatic	0.437	0.426	0.441	0.434		
Hydrophobic		0.456		0.445		
Gasteiger-Hückel	2 Å	q^2	0.558	0.569	0.627	0.621
		PRESS	0.710	0.715	0.694	0.685
		Noc	3	4	5	4
		r^2	0.920	0.934	0.955	0.941
		S	0.303	0.280	0.241	0.269
		F	95.341	84.812	93.199	92.289
		r^2_{pred}	0.771	0.729	0.715	0.780
		Field contribution				
		Steric	0.582	0.099	0.590	0.102
		Electrostatic	0.418	0.498	0.410	0.505
		Hydrophobic		0.403		0.393

^aDefault sp³ carbon as probe atom.

els. In Table 4 are presented the residual values from the prediction of the activity values of the test set.

To further test the stability and robustness of the 3D-QSAR models, the activities of compounds in the training set were shuffled 10 times. The q^2 of the models derived from these so-called random datasets were compared to those derived with the actual activities. The average q^2 after scrambling the biological data are presented in Table 5. These results indicate the stability and robustness in the CoMFA and CoMSIA models.

CoMFA. PLS analysis using AM1 charges and default parameters resulted in a CoMFA model with a good q^2 value of 0.656 and a r^2_{pred} value of 0.688. Omission of compound **3**, which was an outlier in this analysis, resulted in an increase in the q^2 value to 0.706 and in the r^2_{pred} value to 0.720 for the remaining 28 compounds. The outlier status of compound **3** cannot be successfully explained by the present results.

The effect of the grid spacing was investigated. Reducing the lattice step size from 2 to 1 Å reduced the q^2 from 0.656 to 0.613 for the model with all 29 com-

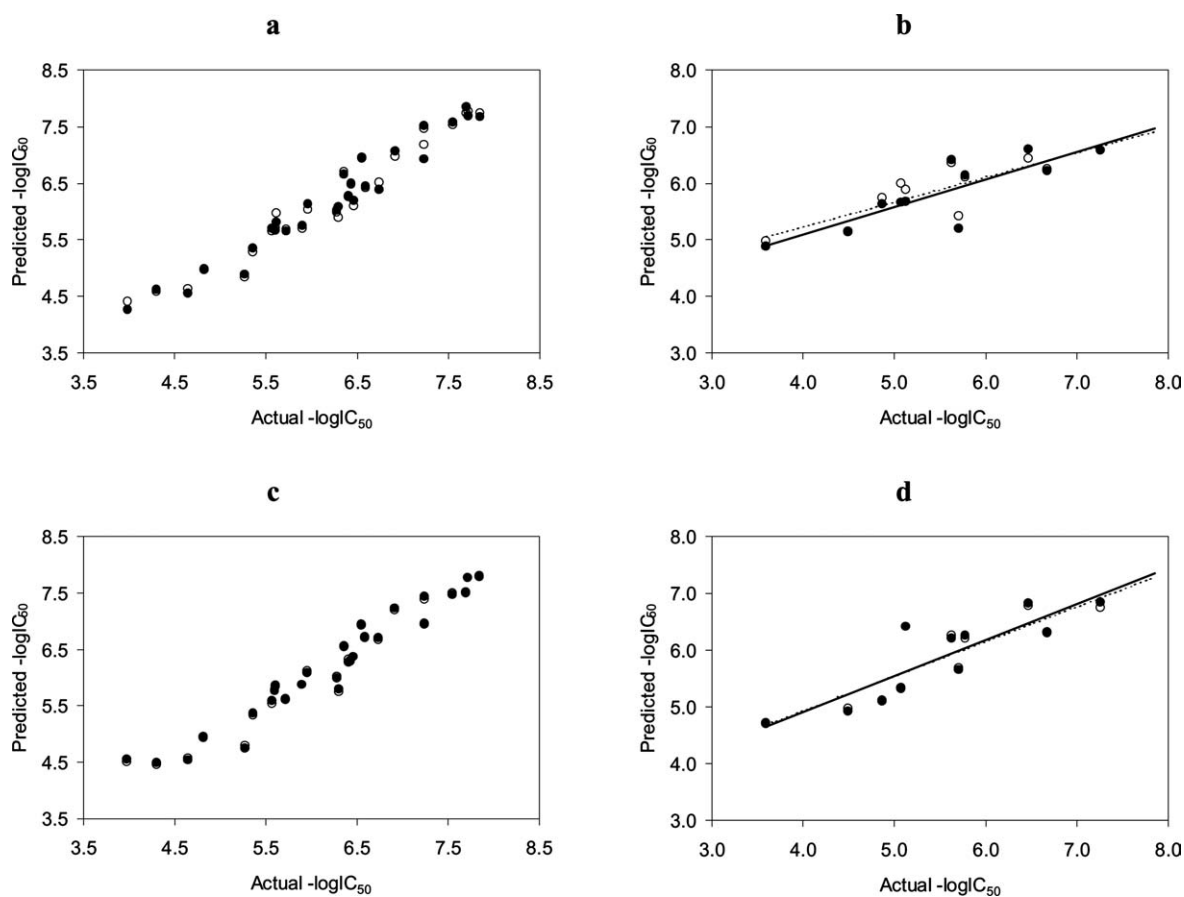


Figure 7. Predictions for the training and test sets with CoMFA and CoMSIA 28 compounds models. Predictions by the 2 Å model are in filled circles and solid lines, while predictions by the 1 Å model are in open circles and broken lines. The fitted curves for CoMFA models for the training and test sets are in panels a and b, respectively. The corresponding curves for the CoMSIA models are shown in panels c and d, respectively.

Table 4. Residuals of the predictions of the test set by the CoMFA and CoMSIA models.^a

Comp.	-LogIC ₅₀	Residuals			
		CoMFA 2 Å	CoMFA 1 Å	CoMSIA 2 Å	CoMSIA 1 Å
30	6.68	0.462	0.429	0.360	0.376
32	5.63	-0.772	-0.736	-0.583	-0.628
33	5.12	-0.563	-0.758	-1.300	-1.300
34	4.49	-0.667	-0.646	-0.429	-0.474
35	5.78	-0.362	-0.318	-0.467	-0.424
36	7.26	0.671	0.674	0.424	0.511
37	6.47	-0.136	0.036	-0.348	-0.302
38	5.71	0.514	0.287	0.061	0.035
39	5.08	-0.576	-0.910	-0.258	-0.233
40	3.59	-1.283	-1.390	-1.128	-1.101
41	4.87	-0.752	-0.867	-0.228	-0.235

^a28 compounds models.

pounds and from 0.706 to 0.650 for the model with 28 compounds.

The predictive performance of the 28 compounds models is shown in Figure 7. The 2 Å model with 28 compounds was the best CoMFA model derived, showing a high predictive performance for the training ($q^2 = 0.706$) and test sets ($r^2_{pred} = 0.720$). However, the 1 Å model with 28 compounds showed the highest predictability for the test set with a r^2_{pred} value of 0.731, despite the lower q^2 value compared to the corresponding 2 Å model. Similarly, the 1 Å model with all compounds showed a higher r^2_{pred} than the corresponding 2 Å model with all compounds, despite its lower q^2 . This means that a higher q^2 value does not necessarily make a model more predictive with reference to an external test set, as previously pointed out in other 3D-QSAR studies [53].

In the 2 Å models with 29 and 28 compounds, the contribution of the steric interactions was 60.8 and 60.0%, respectively. In the 1 Å models the steric contribution was close to 56%. Interesting to note, a larger contribution of the steric interactions over the electrostatic interactions in CoMFA studies has been observed for other NNRTIs [27–29]. Furthermore, in some of these CoMFA analyses the contribution of the steric interactions is also around 60% [28, 29].

In all cases studied, the models derived with AM1 charges showed higher q^2 than those models obtained with Gasteiger-Hückel charges. However, models obtained with Gasteiger-Hückel charges showed a good predictability for the test set (Table 3). Concerning the variation of the probe atom, all models derived with a sp^3 carbon probe atom outperformed models obtained with a sp^3 oxygen and hydrogen as probe atoms (data not shown).

The PLS stdev* coefficient contour maps for the best CoMFA model are shown in Figure 8. To aid in visualization, compound **29** is displayed in the maps. Residues within 5 Å around the inhibitor were merged into the figure. Green regions indicate areas where steric bulk is predicted to enhance biological activity, whereas yellow regions indicate areas where steric bulk is detrimental to biological activity. There is a green contour close to the wing 2 of pyridinone derivatives which is also close to the position of Tyr181. This suggests a favorable interaction between the two systems. This observation further supports the discussion above that the interaction between the aromatic rings of the pyridinone derivatives and Tyr181 may be stabilizing the complex. According to the exper-

imental results for **29** and other potent pyridinone derivatives [8, 16], the mutation of Tyr181 to Cys seems to eliminate stabilizing contacts between the aromatic ring of tyrosine and the bound inhibitor, reducing the binding affinity.

A second favorable steric interaction is also close to the wing 2 and close to the position of the amino acids Tyr188 and Leu234 (Figure 8a). This contour agrees with the docked positions of molecules **17**, **20**, **21**, **35**, and **36** where the substituent at the benzoxazol or benzofuran ring is making contacts with these amino acids. The substituent at position 4 of the benzoxazol ring of compounds **18**, **19**, **28** and **29**, which are very potent pyridinone derivatives, also makes contacts with Tyr188 and Leu234. Interestingly, the second green contour is very close to the benzene ring of compounds **27** and **31**. These compounds with a medium-to-low activity do not make contacts with Tyr181 but with Tyr188. It seems that the interactions with Tyr188 make them not as inactive as expected from the lack of interactions with Tyr181.

These observations agree with the experimental data that mutation of Tyr188 is also responsible for the loss of activity of pyridinone derivatives [15, 16], although to a lesser extent when the mutation Tyr181Cys occurs. Indeed, the amino acids Tyr181 and Tyr188 are very important in the RT inhibition by NNRTIs. The loss of aromatic ring stacking interactions with Tyr181 and Tyr188 is thought to be responsible for the lack of activity of several so-called first generation NNRTIs against RT containing mutations of these amino acids [54]. These observations support the approach to modify the wing 2 of pyridinone derivatives to overcome the problem of mutations at Tyr181 is Tyr188 [34]. Examples of these modified pyridinone derivatives are the hybrid compounds pyridinone-EBPU and pyridinone-efavirenz in Figure 1 and the recently designed molecules by our group [34].

In Figure 8a there is a prominent yellow contour close to the wing 2 and the linker, suggesting steric restriction in this region. The side chains of Lys101 and Val179 are close to this region.

In Figure 8b are represented the electrostatic contributions as contour plots. Red contours illustrate areas where electronegative groups are predicted to favor activity and blue contours illustrate areas where electronegative groups are detrimental to biological activity. A large red electrostatic contour region close to the wing 2 of pyridinone derivatives indicates that high negative charges in this area enhance the affinity. These results are confirmed by the docked position of

Table 5. Average q^2 after scrambling the biological data.

Step size	29 Compds.		28 Compds.	
	CoMFA	CoMSIA	CoMFA	CoMSIA
2 Å	-0.144	-0.288	-0.157	-0.295
1 Å	-0.211	-0.286	-0.212	-0.293

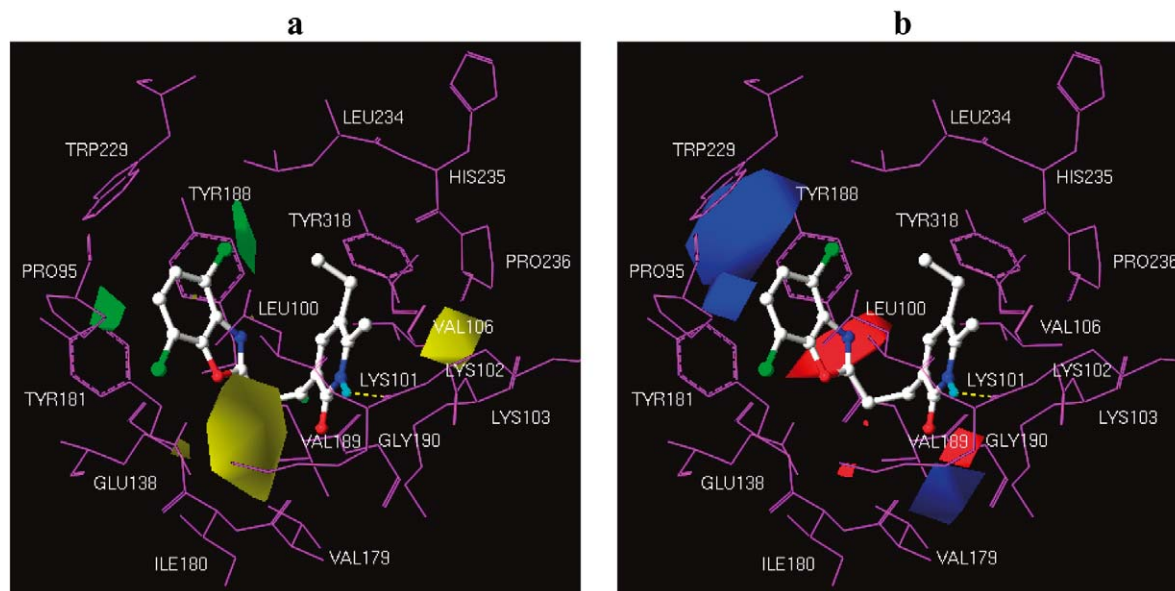


Figure 8. Stdev*coeff contour maps of best CoMFA model. Compound **29** and amino acids within 5 Å around the inhibitor are shown for clarity. (a) Steric contour maps. Sterically favored areas (contribution level of 80%) in green. Sterically unfavored areas (contribution level of 20%) in yellow. (b) Electrostatic contour maps. Negative-charge-favored areas (contribution level of 20%) in red. Negative-charge-unfavored areas (contribution level of 80%) in blue. The residues are in sticks and the inhibitor is in balls and sticks. The hydrogen bond is displayed as yellow dashes.

the wing 2 of molecules **2**, **4**, **5**, **10**, **18**, **19**, **22**, **24**, **25**, **28**, **29** and **36**. In all these cases, an atom with a high negative charge is oriented toward this region. This red contour explains the electrostatic contribution to the high activity of compounds **18**, **28** and **29**. There is a prominent blue contour close to the position of Trp229. The very low activity of compounds **14**, **33**, and **40** may be well explained by the presence of a nitrogen atom with a negative charge close to this contour. Similarly, a second but less prominent blue contour close to Pro95 explains the low activity of compound **7** that has a nitrogen atom with a negative charge close to this region.

CoMSIA. In the present study, we obtained better PLS statistics and predictive performance on the test set with the CoMSIA models than with the CoMFA models (Table 3). Using AM1 charges, default Sybyl

CoMSIA parameters and all compounds in the training set resulted in a model with a q^2 value of 0.643 and a r^2_{pred} value of 0.785 (as compared to a q^2 value of 0.656 and a r^2_{pred} value of 0.688 for CoMFA). However, elimination of the same outlier, compound **3**, brought the q^2 value up to 0.722 and the r^2_{pred} value to 0.759 (as compared to a q^2 value of 0.706 and a r^2_{pred} value of 0.720 for CoMFA). Reducing the grid spacing from 2 to 1 Å resulted in CoMFA PLS models with almost the same q^2 values (Table 3). However, the performance in predicting the activities of the test set was slightly lower (although higher than the predictive performance of the corresponding CoMFA models) (Table 3). The 1 Å model with 28 compounds was the best CoMSIA model derived with a q^2 value of 0.723 and a r^2_{pred} value of 0.750 (compare with the best CoMFA model derived with a q^2

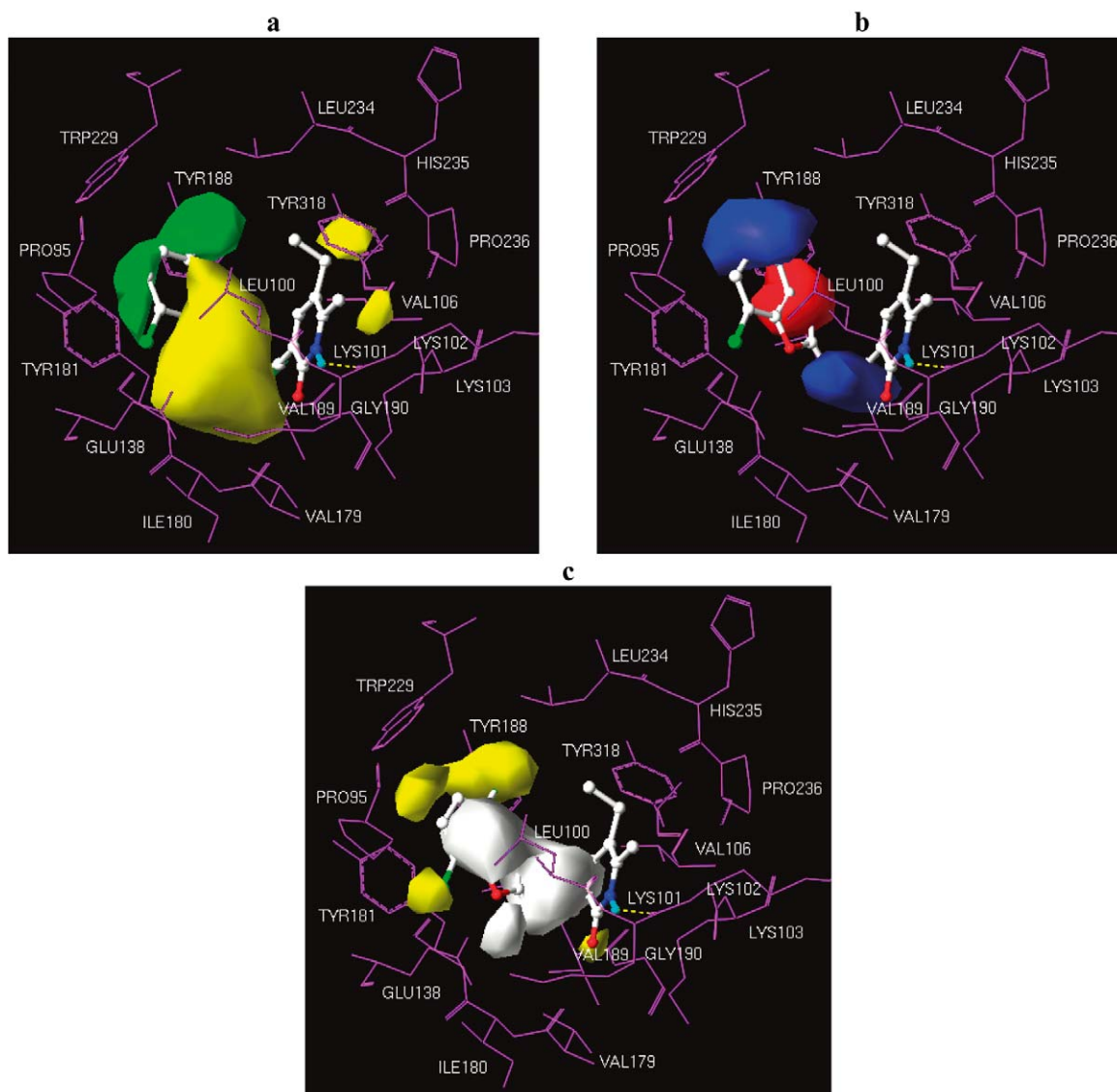


Figure 9. Stdev*coeff contour maps of best CoMSIA model. Compound **29** and amino acids within 5 Å around the inhibitor are shown for clarity. (a) Steric contour maps. (b) Electrostatic contour maps. The region colors have the same interpretation as those in Figures 8a and 8b, respectively. (c) Hydrophobic contour maps. Hydrophobically favored areas (contribution level of 80%) in yellow; hydrophobically unfavored areas (contribution level of 20%) in white. The residues, the inhibitor and the hydrogen bond are represented as in Figure 8.

value of 0.706 and a r^2_{pred} value of 0.720). The superior performance of CoMSIA relative to CoMFA, with this data set, may be attributed mainly to the contribution of the hydrophobic fields in the CoMSIA models. These types of descriptor fields are not available with standard CoMFA.

Similar strengths of steric and electrostatic fields were obtained in the best CoMFA and CoMSIA models (0.600 and 0.400, respectively, in CoMFA and 0.566 (S + H) and 0.434, respectively, in CoMSIA).

Similar to CoMFA, the models obtained with AM1 charges also showed higher q^2 than those models obtained with Gasteiger-Hückel charges. Models obtained with Gasteiger-Hückel charges showed a good predictability for the test set (Table 3).

The effect of the H-donor and H-acceptor CoMSIA fields on the PLS statistics was also investigated. Considering all five CoMSIA field and several field combinations did not improve the predictability of the obtained models (data not shown). This may be well

explained by the docking results. Only one hydrogen bond between pyridinone derivatives and the NNRTI binding site is predicted for all 29 compounds in the training set. Therefore, no differences in the binding affinities among the compounds should be expected from this hydrogen bond.

The PLS stdev* coefficient contour maps for the best CoMSIA model are shown in Figure 9. To aid in visualization, compound **29** is displayed in the maps. Residues within 5 Å around the inhibitor were merged into the figure. Color designation and interpretation for steric and electrostatic fields are the same as in the corresponding CoMFA contour maps.

Steric fields are similarly placed as those of the CoMFA model (Figure 8a). However, the two green CoMSIA contours close to the positions of Tyr181 and Tyr188 are more prominent than the corresponding contours obtained with CoMFA. This observation remarks the importance of the interactions between the pyridinone derivatives and Tyr181 and Tyr188 discussed previously.

The electrostatic fields of CoMSIA are generally in accordance with the field distribution of CoMFA maps (Figure 8b). It is interesting to note the additional blue contour surrounding the linker (see 'L' group in Tables 1 and 2) of several pyridinone derivatives, suggesting a preference for the ethylene linker over the amino methylene linker. This is confirmed by the higher activity and lower binding free energy of compounds **26**, **27** and **40** compared to their structurally related compounds **1**, **31** and **33**, respectively. The amino methylene group of **1**, **31** and **33**, where the nitrogen has a negative charge, is close to Tyr188 and Glu190, enabling negative interactions with the main-chain of these amino acids. However, compounds **23** and **39**, with an ethylene linker, are less active than their structurally related compounds **11** and **12**, respectively. This apparent exception may be explained, at least in part, because the blue contour does not surround the linker of these molecules. As discussed previously, these four molecules adopt a particular docked position such that the linker is above the linker of the other pyridinone derivatives (Figure 5b).

The hydrophobic analysis of CoMSIA could demonstrate more clearly the hydrophobic interactions between the pyridinone derivatives and RT. For the hydrophobic contour maps of CoMSIA, yellow contours indicate regions where increased hydrophobic interactions are associated with enhanced activity and white contours indicate regions where increased hydrophilic interactions are associated with enhanced activity. As

shown in Figure 9c, the yellow contours around the substituents at positions 4 and 7 of the benzoxazol ring of **29** indicate that these substituents interact with surrounding amino acids at the NNRTI binding site through hydrophobic interaction. Amino acids close to these contours are Tyr188 (close to the substituent at position 4), and the side chains of Pro95, Leu100 and Tyr181 (close to the substituent at position 7). This observation is in agreement with the SAR and 2D-QSAR studies for pyridinone derivatives with a benzoxazol ring. In these studies it was concluded that compounds di-substituted with methyl or chlorine groups at the positions 4 and 7 of the benzoxazol ring are more active than the corresponding mono- or unsubstituted compounds [6, 11, 12]. An example of these di-substituted potent inhibitors is compound **29**.

The present 3D-QSAR models did not explain the experimentally observed loss of activity due to the Lys103Asn mutation [8, 13–15]. Since all derivatives studied have the same pyridinone ring, no significant contours around this ring may be interpreted. However, from the docking studies it is observed that the pyridinone ring makes important contacts with Lys103 (Figures 3 and 4) [34].

Conclusions

We predicted the binding conformations of 40 pyridinone derivatives into the NNRTI binding pocket using AutoDock. The molecules adopt a butterfly-like conformation and most of them have a common binding mode which is similar to the binding conformation of nevirapine. The binding free energies calculated by this method correlated well with the reported inhibitory activities against RT. The correlation was better for most similar compounds. Starting from the docked positions of the pyridinone derivatives, we performed 3D-QSAR analyses. Robust CoMFA and CoMSIA models were obtained with a high predictability performance for the training and test sets. CoMSIA models performed slightly better than CoMFA models. The good 3D-QSAR models obtained further validated the reasonableness of the docking predictions by AutoDock. According to the 3D-QSAR models, steric, electrostatic and hydrophobic interactions occur between the inhibitors and the NNRTI binding pocket. Important interactions with the amino acids Tyr181 and Tyr188 were predicted which further explained the loss of activity of pyridinones against RT containing mutations. These interactions also help to rationalize

at the molecular level, at least in part, the design of hybrid pyridinone molecules, and support the approach to modify the benzoxazol ring of potent pyridinone derivatives to overcome the problem of Tyr181Cys and Tyr188Cys mutations. The 3D-QSAR models derived will be useful for the rational design of pyridinone derivatives active against RT containing mutations.

Acknowledgements

The authors are very grateful to Dr. A. Olson and his colleagues at the Scripps Research Institute for providing AutoDock and to Dr. Roman Laskowski for providing LIGPLOT. J.L. Medina-Franco is grateful to CONACyT and DGAPA, UNAM, for the Ph. D. thesis scholarships. We also would like to acknowledge CONACyT for financing project G-34851-M.

References

- De Clercq, E., *Biochim. Biophys. Acta*, 1587 (2002) 258.
- De Clercq, E., *Antivir. Res.*, 38 (1998) 153.
- Esnouf, R., Ren, J., Ross, C., Jones, Y., Stammers, D. and Stuart, D., *Nat. Struct. Biol.*, 2 (1995) 303.
- (a) Ren, J., Esnouf, R., Garman, E., Somers, D., Ross, C., Kirby, I., Keeling, J., Darby, G., Jones, Y., Stuart, D. and Stammers, D., *Nat. Struct. Biol.*, 2 (1995) 293. (b) Hopkins, A.L., Ren, J., Esnouf, R.M., Willcox, B.E., Jones, E.Y., Ross, C., Miyasaka, T., Walker, R.T., Tanaka, H., Stammers, D.K. and Stuart, D.I., *J. Med. Chem.*, 39 (1996) 1589. (c) Ren, J., Esnouf, R., Hopkins, A., Ross, C., Jones, Y., Stammers, D. and Stuart, D., *Structure*, 3 (1995) 915. (d) Ren, J., Milton, J., Weaver, K. L., Short, S.A., Stuart, D.I. and Stammers, D.K., *Structure*, 8 (2000) 1089. (e) Ren, J., Nichols, C., Bird, L.E., Fujiwara, T., Sugimoto, H., Stuart, D.I. and Stammers, D.K., *J. Biol. Chem.*, 275 (2000) 14316. (f) Esnouf, R.M., Ren, J., Hopkins, A. L., Ross, C.K., Jones, E.Y., Stammers, D.K. and Stuart, D.I., *Proc. Natl. Acad. Sci. USA*, 94 (1997) 3984. (g) Ren, J., Diprose, J., Warren, J., Esnouf, R.M., Bird, L.E., Ikemizu, S., Slater, M., Milton, J., Balzarini, J., Stuart, D.I. and Stammers, D.K., *J. Biol. Chem.*, 275 (2000) 5633. (h) Chan, J.H., Hong, J.S., Hunter III, R.N., Orr, G.F., Cowan, J.R., Sherman, D.B., Sparks, S.M., Reitter, B.E., Andrews III, C.W., Hazen, R.J., St Clair, M., Boone, L.R., Ferris, R.G., Creech, K.L., Roberts, G.B., Short, S.A., Weaver, K., Ott, R.J., Ren, J., Hopkins, A., Stuart, D.I. and Stammers, D.K., *J. Med. Chem.*, 44 (2001) 1866. (i) Hsiou, Y., Das, K., Ding, J., Clark Jr., A.D., Kleim, J.P., Rosner, M., Winkler, I., Riess, G., Hughes, S.H. and Arnold, E., *J. Mol. Biol.*, 284 (1998) 313. (j) Ren, J., Esnouf, R.M., Hopkins, A.L., Stuart, D.I. and Stammers, D.K., *J. Med. Chem.*, 42 (1999) 3845. (k) Ren, J., Esnouf, R.M., Hopkins, A.L., Warren, J., Balzarini, J., Stuart, D.I. and Stammers, D.K., *Biochemistry*, 37 (1998) 14394. (l) Hogberg, M., Sahlberg, C., Engelhardt, P., Noreen, R., Kangasmetsa, J., Johansson, N.G., Oberg, B., Vrang, L., Zhang, H., Sahlberg, B. L., Unge, T., Lovgren, S., Fridborg, K. and Backbro, K., *J. Med. Chem.*, 43 (2000) 304.
- Goldman, M.E., Nunberg, J.H., O'Brien, J.A., Quintero, J.C., Schleif, W.A., Freund, K.F., Gaul, S.L., Saari, W.S., Wai, J.S., Hoffman, J.M., Anderson, P.S., Hupe, D.J., Emini, E.A. and Stern, A.M., *Proc. Natl. Acad. Sci. USA*, 88 (1991) 6863.
- Garg, R., Gupta, S.P., Gao, H., Babu, M.S., Debnath, A.K. and Hansch, C., *Chem. Rev.*, 99 (1999) 3525.
- Gupta, S.P., In Jucker, E. (Ed.), *Progress in Drug Research*, Vol. 58, Birkhäuser Verlag, Basel, Switzerland, 2002, pp. 252–253.
- Nunberg, J.H., Schleif, W.A., Boots, E.J., O'Brien, J.A., Quintero, J.C., Hoffman, J.M., Emini, E.A. and Goldman, M.E., *J. Virol.*, 65 (1991) 4887.
- Saari, W.S., Hoffman, J.M., Wai, J.S., Fisher, T.E., Rooney, C.S., Smith, A.M., Thomas, C.M., Goldman, M.E., O'Brien, J.A., Nunberg, J.H., Quintero, J.C., Schleif, W.A., Emini, E.A., Stern, A.M. and Anderson, P.S., *J. Med. Chem.*, 34 (1991) 2922.
- Hoffman, J.M., Wai, J.S., Thomas, C.M., Levin, R.B., O'Brien, J.A. and Goldman, M.E., *J. Med. Chem.*, 35 (1992) 3784.
- Saari, W.S., Wai, J.S., Fisher, T.E., Thomas, C.M., Hoffman, J.M., Rooney, C.S., Smith, A.M., Jones, J.H., Bamberger, D.L., Goldman, M.E., O'Brien, J.A., Nunberg, J.H., Quintero, J.C., Schleif, W.A., Emini, E.A. and Anderson, P.S., *J. Med. Chem.*, 35 (1992) 3792.
- Hoffman, J.M., Smith, A.M., Rooney, C.S., Fisher, T.E., Wai, J.S., Thomas, C.M., Bamberger, D., Barnes, J.L., Williams, T. M., Jones, J.H., Olson, B.D., O'Brien, J.A., Goldman, M.E., Nunberg, J.H., Quintero, J.C., Schleif, W.A., Emini, E.A. and Anderson, P.S., *J. Med. Chem.*, 36 (1993) 953.
- Davey, R.T., Dewar, R.L., Reed, G.F., Vasudevachari, M.B., Polis, M.A., Kovacs, J.A., Fallon, J., Walker, R.E., Masur, H., Haneiwich, S.E., O'Neill, D.G., Decker, M.R., Metcalf, J.A., Deloria, M.A., Laskin, O.L., Salzman, N. and Lane, H.C., *Proc. Natl. Acad. Sci. USA*, 90 (1993) 5608.
- Saag, M.S., Emini, E.A., Laskin, O.L., Douglas, J., Lapidus, W.I., Schleif, W.A., Whitley, R.J., Hildebrand, C., Byrnes, V.W., Kappes, J.C., Anderson, K.W., Massari, F.E., Shaw, G.M. and L-697,661 Working Group, *N. Engl. J. Med.*, 329 (1993) 1065.
- Sardana, V.V., Emini, E.A., Gotlib, L., Graham, D.J., Lineberger, D.W., Long, W.J., Schlabach, A.J., Wolfgang, J.A. and Condra, J.H., *J. Biol. Chem.*, 267 (1992) 17526.
- Byrnes, V.W., Sardana, V.V., Schleif, W.A., Condra, J.H., Waterbury, J.A., Wolfgang, J.A., Long, W.J., Schneider, C.L., Schlabach, A.J., Wolanski, B.S., Graham, D.J., Gotlib, L., Rhodes, A., Titus, D.L., Roth, E., Blahy, O.M., Quintero, J. C., Staszewski, S. and Emini, E.A., *Antimicrob. Agents Chemother.*, 37 (1993) 1576.
- Jourdan, F., Renault, J., Fossey, C., Bureau, R., Ladurée, D., Robba, M., Aubertin, A.M. and Kirn, A., *Antivir. Chem. Chemother.*, 8 (1997) 161.
- Dollé, V., Fan, E., Nguyen, C. H., Aubertin, A.M., Kirn, A., Andreola, M.L., Jamieson, G., Tarrago-Litvak, L. and Bisagni, E., *J. Med. Chem.*, 38 (1995) 4679.
- Dollé, V., Nguyen, C. H., Legraverend, M., Aubertin, A.M., Kirn, A., Andreola, M.L., Ventura, M., Tarrago-Litvak, L. and Bisagni, E., *J. Med. Chem.*, 43 (2000) 3949.
- Corbett, J.W., Kresge, K.J., Pan, S., Cordova, B.C., Klabe, R. M., Rodgers, J.D. and Erickson-Viitanen, S.K., *Bioorg. Med. Chem. Lett.*, 11 (2001) 309.
- Pungpo, P., Hannongbua, S. and Wolschann, P., *Curr. Med. Chem.*, 10 (2003) 1661, and references therein.

22. Douali, L., Villemin, D. and Cherqaoui, D., *J. Chem. Inf. Comput. Sci.*, 43 (2003) 1200.
23. Cramer III, R.D., Patterson, D.E. and Bunce, J.D., *J. Am. Chem. Soc.*, 110 (1988) 5959.
24. Klebe, G., Abraham, U. and Mietzner, T., *J. Med. Chem.*, 37 (1994) 4130.
25. Kubinyi, H. (Ed.) *3D QSAR in Drug Design: Theory, Methods and Applications*, ESCOM, Leiden, The Netherlands, 1993.
26. Barreca, M.L., Carotti, A., Carrieri, A., Chimirri, A., Monforte, A.M., Calace, M.P. and Rao, A., *Bioorg. Med. Chem.*, 7 (1999) 2283.
27. Hannongbua, S., Pungpo, P., Limtrakul, J. and Wolschann, P., *J. Comput.-Aided Mol. Des.*, 13 (1999) 563.
28. Pungpo, P. and Hannongbua, S., *J. Mol. Graph. Mod.*, 18 (2000) 581.
29. Hannongbua, S., Nivesanond, K., Lawtrakul, L., Pungpo, P. and Wolschann, P., *J. Chem. Inf. Comput. Sci.*, 41 (2001) 848.
30. Chen, H.F., Yao, X.J., Li, Q., Yuan, S.G., Panaye, A., Doucet, J.P. and Fan, B.T., *SAR QSAR Environ. Res.*, 14 (2003) 455.
31. Titmuss, S.J., Keller, P.A. and Griffith, R., *Bioorg. Med. Chem.*, 7 (1999) 1163.
32. Zhou, Z., Madrid, M. and Madura, J.D., *Proteins*, 49 (2002) 529.
33. Kontoyianni, M., McClellan, L.M. and Sokol, G.S., *J. Med. Chem.*, 47 (2004) 558.
34. Medina-Franco, J.L., Rodríguez-Morales, S., Juárez-Gordiano, C., Hernández-Campos, A., Jiménez-Barbero, J. and Castillo, R., *Bioorg. Med. Chem.*, 2004 *submitted for publication*.
35. Lozano, J.J., Pastor, M., Cruciani, G., Gaedt, K., Centeno, N. B., Gago, F. and Sanz, F., *J. Comput.-Aided Mol. Des.*, 14 (2000) 341.
36. Sippl, W., Contreras, J.-M., Parrot, I., Rival, Y.M. and Wermuth, C.G., *J. Comput.-Aided Mol. Des.*, 15 (2001) 395.
37. Sippl, W., *J. Comput.-Aided Mol. Des.*, 16 (2002) 825.
38. Sippl, W., *Bioorg. Med. Chem.*, 10 (2002) 3741.
39. Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K. and Olson, A.J., *J. Comput. Chem.*, 19 (1998) 1639.
40. Brookhaven Protein Data Bank. <http://www.rcsb.org>.
41. Sybyl 6.8, Tripos Associates, Inc., St. Louis, MO, USA.
42. Clark, M., Cramer III, R.D. and Van Opdenbosch, N., *J. Comput. Chem.*, 10 (1989) 982.
43. Streitwieser, A. (Ed.) *Molecular Orbital Theory for Organic Chemists*, Wiley, New York, 1961.
44. Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S. and Weiner, P., *J. Am. Chem. Soc.*, 106 (1984) 765.
45. Gasteiger, J. and Marsili, M., *Tetrahedron*, 36 (1980) 3219.
46. Mehler, E.L. and Solmajer, T., *Protein Eng.*, 4 (1991) 903.
47. Dewar, M.J.S., Zoebisch, E.G., Healy, E.F. and Stewart, J.J. P., *J. Am. Chem. Soc.*, 107 (1985) 3902.
48. Viswanadhan, V.N., Ghose, A.K., Revenkar, G.R. and Robins, R., *J. Chem. Inf. Comput. Sci.*, 29 (1989) 163.
49. Klebe, G., *J. Mol. Biol.*, 237 (1994) 212.
50. Parreira, R.L.T., Abrahão, O. and Galembeck, S.E., *Tetrahedron*, (2001) 3243.
51. Wallace, A.C., Laskowski, R.A. and Thornton, J.M., *Protein Eng.*, 8 (1995) 127.
52. McGaughey, G.B., Gagné, M. and Rappé, A., *J. Biol. Chem.*, 273 (1998) 15458.
53. Buolamwini, J.K. and Assefa, H., *J. Med. Chem.*, 45 (2002) 841.
54. Ren, J., Nichols, C., Bird, L., Chamberlain, P., Weaver, K., Short, S., Stuart, D.I. and Stammers, D. K., *J. Mol. Biol.*, 312 (2001) 795.

Molecular docking of the highly hypolipidemic agent α -asarone with the catalytic portion of HMG-CoA reductase

José Luis Medina-Franco,^a Fabián López-Vallejo,^a Sergio Rodríguez-Morales,^a Rafael Castillo,^{a,*} Germán Chamorro^b and Joaquín Tamariz^{c,*}

^aDepartamento de Farmacia, Facultad de Química, Universidad Nacional Autónoma de México, Circuito Exterior, Ciudad Universitaria, 04510 México, D.F., Mexico

^bLaboratorio de Toxicología Preclínica, Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional. Prol. Carpio y Plan de Ayala, 11340 México, D.F., Mexico

^cDepartamento de Química Orgánica, Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional. Prol. Carpio y Plan de Ayala, 11340 México, D.F., Mexico

Received 27 September 2004; revised 3 November 2004; accepted 15 December 2004

Abstract—Docking experiments using a number of published crystal structures of HMG-CoA reductase with the potent hypocholesterolemic agent α -asarone are described. The results indicate that α -asarone binds in the enzyme's active site. The methoxy groups play a key role in the binding and probably also in its biological activity, as shown by extensive SAR studies reported for analogues of α -asarone. The docking results will be valuable for the structure-based design of novel hypolipidemic agents.

© 2004 Elsevier Ltd. All rights reserved.

Research in new lipid-lowering drugs is very active because hypercholesterolemia¹ and high levels of serum LDL-cholesterol² have been generally recognized to contribute significantly to the progression of atherosclerosis. Cholesterol biosynthesis in the body is mainly regulated in the liver by the enzyme 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase (HMGR).³ Inhibition of this enzyme has proven to be the most efficient therapy for hyperlipidemia, since the enzymatic transformation of HMG-CoA to mevalonate represents one of the key steps in the metabolic pathway toward the biosynthesis of isoprenoids and sterols, such as cholesterol.⁴ Among the most effective hypocholesterolemic drugs for clinical use today are the statins, which possess an HMG-like moiety linked to a hydrophobic decalin core.^{5–7} Synthetic statin-like compounds including an HMG-like moiety have shown significant hypocholesterolemic activity.^{8–10}

The structures of the catalytic portion of human HMG-CoA reductase (HMGR) in complex with the substrate

and six statins have been recently determined.^{11,12} The enzyme forms tetramers and has four active sites formed by residues of two monomers. The HMG-binding pocket is characterized by the so-called *cis* loop (residues 682–694).¹¹ A proposed catalytic mechanism suggests that the residues Leu-691 and Glu-559 participate directly in the reduction of the substrate HMG-CoA (Fig. 1).¹³

α -Asarone (Fig. 1) is the main biological active component of the bark extract of *Guatteria guameri* Greenman (Annonaceae),¹⁴ a medicinal plant utilized in Mexico to treat hypercholesterolemia and cholelithiasis.¹⁵ In view of its remarkable hypolipidemic activity,^{16,17} and more recently of its use as a potential antithrombotic,¹⁸ antimicrobial, insecticidal, nematocidal, and antifeedant agent,¹⁹ α -asarone has attracted widespread interest.

The hypolipidemic action mechanism of α -asarone has recently been established in a rat model as an inhibitory effect on hepatic HMGR.²⁰ The stimulation of bile secretion was found as an additional mechanism on reducing the serum cholesterol levels and on its associated cholelitholytic activity. Although these results suggest that α -asarone inhibits cholesterol biosynthesis in parallel with the statin mechanism,¹² there is no

Keywords: α -Asarone; Docking; HMG reductase; Hypolipidemia.

* Corresponding authors. Tel.: +52 55 5729 6300/62411; fax: +52 55 5396 3503 (J.T.); e-mail: jtamariz@woodward.ench.ipn.mx

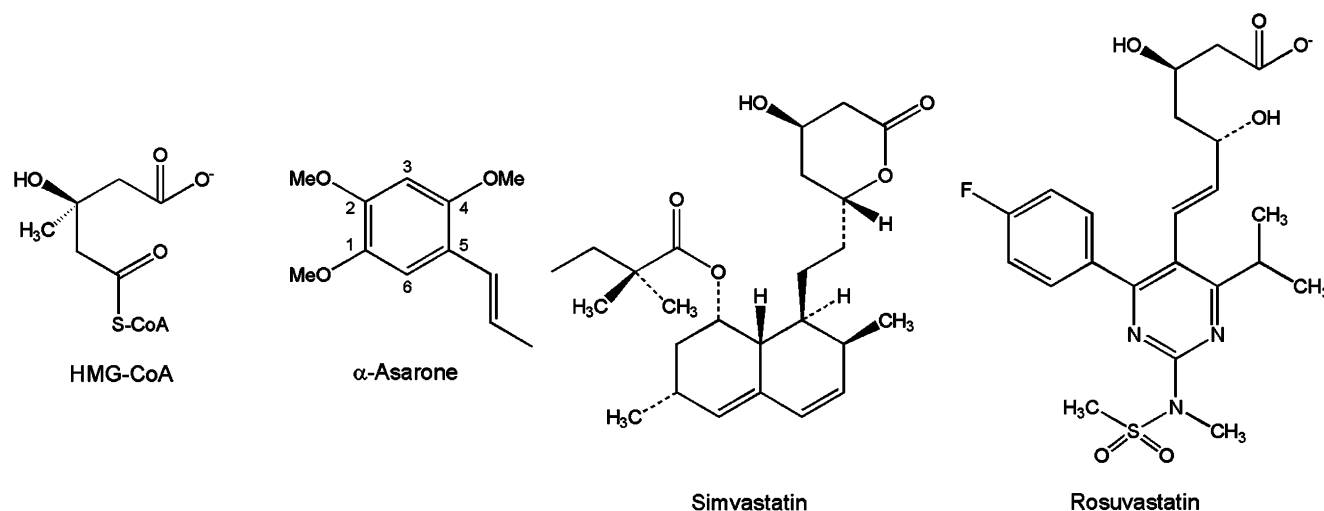


Figure 1. Chemical structures of the HMGR substrate, α -asarone and representative statins used in this study.

evidence as to how α -asarone prevents the binding of HMG-CoA to the target enzyme. In order to gain insight into this problem, to explore the interactions to the active site and consequently, to improve the development of more efficient inhibitors associated with structural features of α -asarone, a study of docking of α -asarone into HMGR and its binding pattern is hereby described.

The docking experiments²¹ were performed using the structures of HMGR complexed with the substrate¹¹ and complexed with statins.¹² The HMG structures were retrieved from the RCSB protein data base (PDB entries 1DQ8, 1HW8, 1HW9, 1HWI, 1HWJ, 1HWK, and 1HWL).²⁹ Before docking α -asarone, the docking protocol was validated. HMG and the statins were removed from the active site and docked back into the binding site. The root mean square deviation (RMSD) between the predicted conformation and the observed X-ray crystallographic conformation led to RMSD values below or near 1 Å. The statins with the lowest RMSD values were simvastatin and rosuvastatin (Fig. 1). Validation results for the substrate and the statins with the lowest RMSD values are summarized in Table 1. The docked energy, the binding free energy (ΔG), and the inhibition constant (K_i) calculated by AutoDock are also reported, including the experimental biological activity.

The interaction energies calculated for simvastatin and rosuvastatin (Table 1) are in agreement with the recent published experimental results, which shows that rosuvastatin is a more potent inhibitor than simvastatin.^{30,31}

The docked energy and the binding free energy calculated by AutoDock for rosuvastatin are more favorable than that calculated for simvastatin (Table 1). A comparison with the calculated interaction energy for HMG is not representative since the crystallographic structure of HMG corresponds to the hydrolyzed molecule and not to HMG-CoA.¹¹

The above results indicate that the parameters used for AutoDock successfully reproduced the X-ray structures. The ability of AutoDock to reproduce the binding conformation of statins, which have several flexible bonds (Fig. 2) is remarkable.

α -Asarone was first docked with the structure of HMGR in complex with the substrate and in complex with simvastatin and rosuvastatin, the statins for which the lowest RMSD values were obtained in the validation of the docking protocol. The docking results are summarized in Table 2.

The docked energy and the binding free energy estimated for α -asarone when docked into the binding site of simvastatin and rosuvastatin are almost identical, which is not surprising since the binding pocket geometries are very similar (Table 2).¹² The slightly lower energy calculated with the binding pocket of the substrate may be explained by favorable interactions with the residue Leu-862. The latter is absent in the pocket geometry of simvastatin and rosuvastatin (cf. Fig. 4).¹² The differences in the conformation of the side chains of some residues of the binding pocket such as Asp-690, Arg-590, Leu-857, and Leu-853 (cf. Fig. 6)

Table 1. Docking of HMG and statins

Compd	Docked energy (kcal/mol)	ΔG (kcal/mol)	K_i (at 298.15 K)	RMSD	Experimental IC_{50} (M) ^a
HMG	-11.27	-10.74	23.5×10^{-9}	1.19	4000×10^{-9b}
Simvastatin	-15.29	-11.70	2.6×10^{-9}	0.89	11.2×10^{-9}
Rosuvastatin	-15.81	-12.34	0.9×10^{-9}	0.81	5.4×10^{-9}

^a Average IC_{50} of inhibition of HMGR.³⁰

^b This is the Michaelis constant for HMG-CoA.¹²

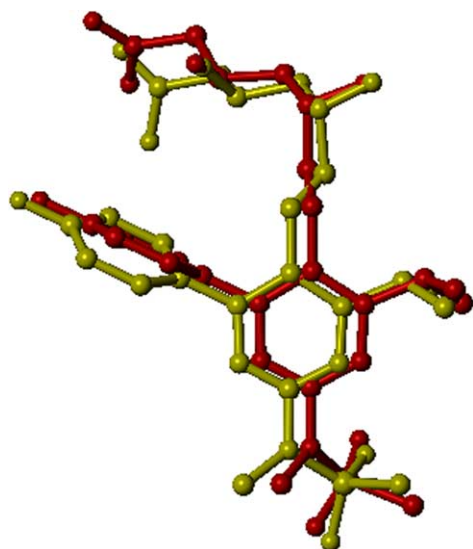


Figure 2. Comparison between the binding position of rosuvastatin found within the crystal structure (yellow) and the conformation predicted by AutoDock (red).

Table 2. Results of docking experiments of α -asarone with the catalytic portion of human HMGR

Binding pocket	Docked energy (kcal/mol)	ΔG (kcal/mol)	K_i (at 298.15 K)
HMG-CoA	-6.60	-6.02	3.9×10^{-5}
Simvastatin	-6.29	-5.68	6.8×10^{-5}
Rosuvastatin	-6.24	-5.66	7.1×10^{-5}

may also contribute to this difference. The much less favorable binding free energy calculated for α -asarone compared with the free energy of binding calculated for simvastatin and rosuvastatin (cf. Tables 1 and 2) is quite noteworthy. This is in good agreement with the experimental HMGR inhibition trend determined for α -asarone and the statins that show inhibition constants (K_i) and IC_{50} in the nanomolar range for rosuvastatin and simvastatin³⁰ and IC_{50} in the millimolar range for α -asarone.²⁰ These observations suggest that the binding mode predicted for α -asarone is reasonable.

AutoDock found one main binding mode for α -asarone in all binding pockets used. α -Asarone binds in the active site of HMGR occupying the region where the 3-hydroxymethylglutaryl moiety of the substrate binds (Fig. 3).

According to the derived docking model, α -asarone interacts with two monomers of the HMGR tetramer. Polar interactions with residues Ser-684, Asp-690, Lys-691, and Lys-692 of the *cis* loop and with Arg-590 of one monomer are observed (Fig. 3). The C-4 methoxy group of α -asarone makes polar contacts with Glu-559 of the second monomer and hydrogen bonds were formed with the side chains of Lys-691 and Asn-755. Additional hydrogen bonds were formed between C-1 and C-2 methoxy groups and Ser-684 and Arg-590 (Fig. 3). The propenyl side chain of α -asarone is close

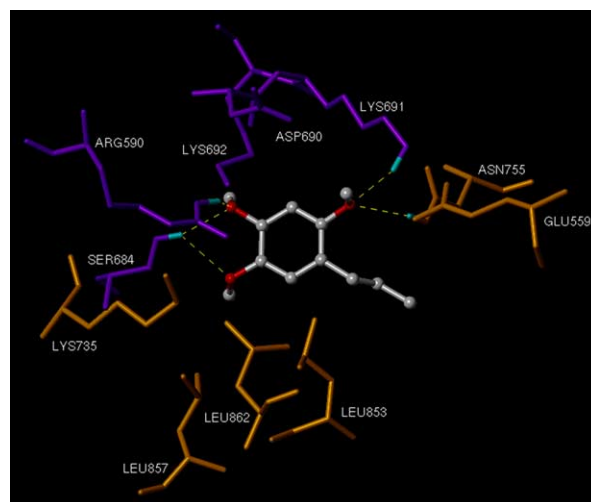


Figure 3. Docking model derived for α -asarone with the catalytic portion of HMGR. The optimized complex of α -asarone with the binding pocket of the substrate is shown. Hydrogen bonds are displayed as yellow dashes. Representative residues within 3.6 Å of α -asarone are shown. Residues from one monomer are orange and those from the other monomer are purple. Nonpolar hydrogens are omitted for clarity except the hydrogens involved in hydrogen bonding.

to Glu-559, Leu-562, His-752, Asn-755, and Leu-853, all residues being of one monomer. The aromatic ring is close to Leu-853 and Leu-862, with which van der Waals contacts are made. These interactions seem to be responsible for the affinity of α -asarone with HMGR, thus sterically preventing the substrate from binding. Similar conclusions were obtained when the docking experiments were conducted using the structures of HMGR in complex with compactin, fluvastatin, cerivastatin, and atorvastatin.

The interactions of α -asarone with the catalytic site of HMGR resemble the interactions observed in the substrate and statins complexes. The C-4 methoxy group of α -asarone occupies a region similar to the C-5 carbonyl oxygen of HMG-CoA and to the C-5 hydroxyl group of statins, the latter interacting with the side chains of Lys-691 and Glu-559, which seem to participate directly in the catalysis (Fig. 4).¹³ According to the mechanism of catalysis proposed for the reduction of HMG-CoA,¹³ which suggests a Glu-559 protonated under physiological conditions, the C-4 oxygen of the methoxy group in α -asarone may also form a hydrogen bond with the carboxylic group of Glu-559.

The C-1 and C-2 methoxy groups of α -asarone have also similar interactions to those of the C-1 carboxylate group of both the substrate and the statins, including hydrogen bonds with Ser-684 and interactions with the polar side chains of Lys-692 and Lys-735.^{11,12} The C-1 methoxy group is located in the binding region of the butyryl group of simvastatin and the fluorophenyl group of rosuvastatin (Fig. 4b and c, respectively). The methyl group of the C-1 methoxy group makes van der Waals contacts with the side chain of Leu-857. The propenyl side chain of α -asarone is located near the hydrophobic group of the statins.

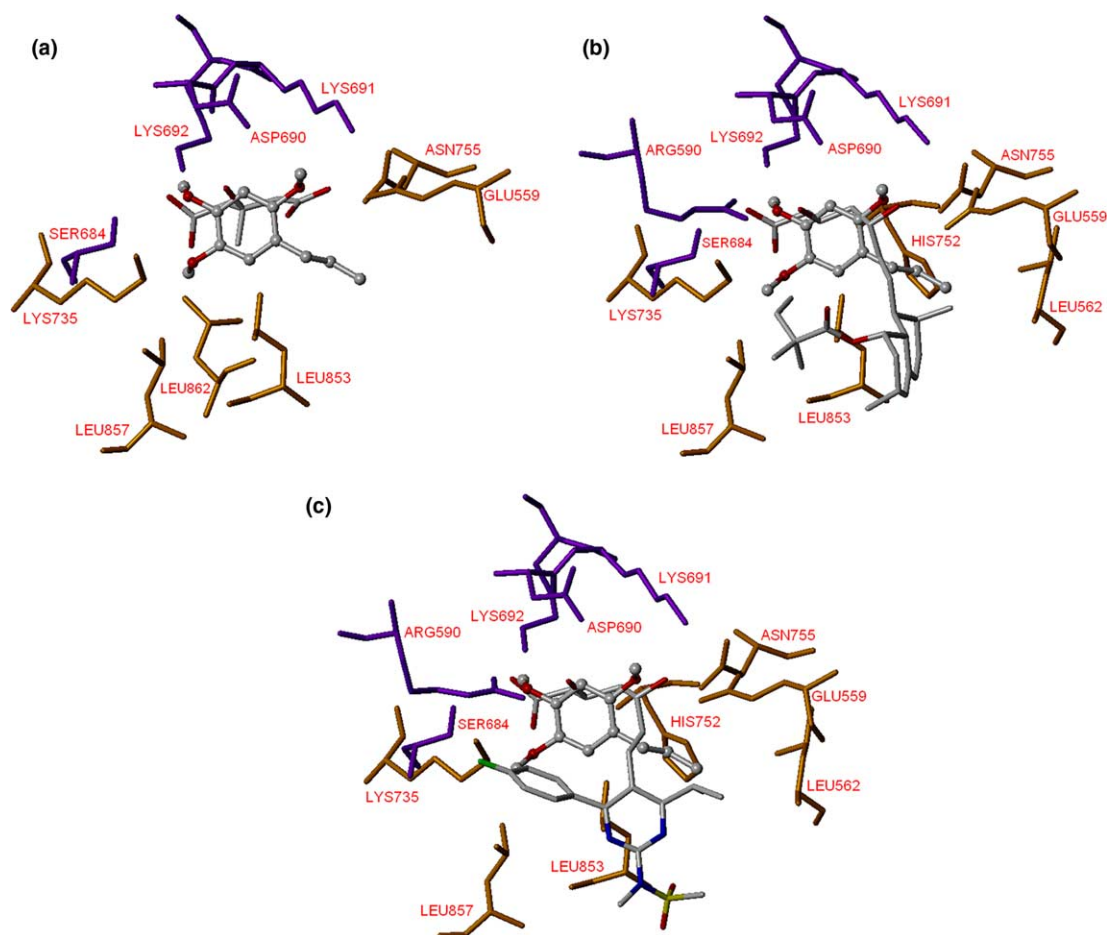


Figure 4. Comparison of the predicted binding mode of α -asarone with the actual binding mode of (a) the substrate; (b) simvastatin, and (c) rosuvastatin. Representative residues within 3.6 Å of α -asarone are shown. Residues are colored as in Figure 3. Nonpolar hydrogens are omitted for clarity.

The above observations suggest that the interactions of the C-1, C-2, and C-4 oxygen atoms on α -asarone are very important in the α -asarone binding to HMGR and play a similar role to that of the HMG-like moiety present in the substrate and the statins (Fig. 5). This is in agreement with the experimental observation that the trimethoxybenzene core seems to be fundamental for the hypocholesterolemic activity of α -asarone analogues.³² The docking model also helps to visualize, at the molecular level, the experimental results that indicate that the side chain of α -asarone should be at the *meta* and *para* positions of the C-1 and C-2 methoxy groups, respectively.³³

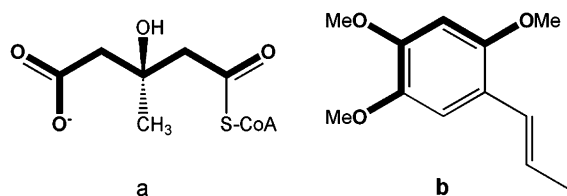


Figure 5. (a) HMG-like moiety of the substrate (in bold). (b) Structure fragment of α -asarone (in bold) that play a similar role to that of the HMG-like moiety in the binding to HMGR.

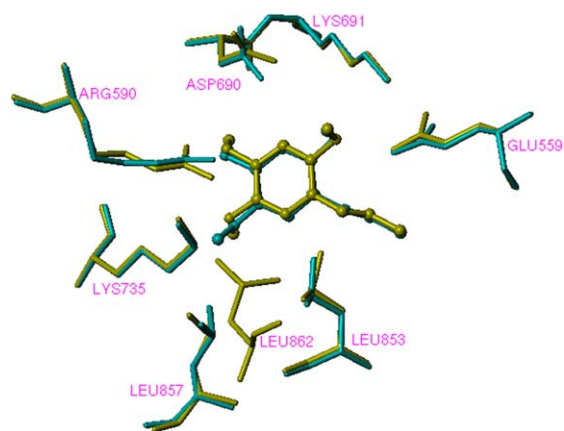


Figure 6. Predicted binding conformations for the C-1 methoxy group of α -asarone when docked into the binding pocket of the substrate (yellow) and into the binding pocket of rosuvastatin (cyan). Representative residues are shown. Nonpolar hydrogens are omitted for clarity.

AutoDock found two possible binding conformations for the C-1 methoxy group depending on the binding site. In the binding pocket of the substrate, the methoxy group was in a *cis* position relative to the vicinal C-2

methoxy group (dihedral angle of 281.6°) while, in the binding pocket of the statins the relative position was *trans* (dihedral angle of 72.6°). According to the ab initio (HF/6-31G**) energy calculations, the conformation of α -asarone with the C-1 methoxy group in a *cis* position was more stable by ca. 2 kcal/mol than the conformation with the C-2 methoxy group in the *trans* position. The different predicted conformation for the C-1 methoxy group might be explained by the presence of the residue Leu-862 in the binding pocket of the substrate that is lacking in the binding geometry of the statins (Fig. 6).¹² Thus, in the binding site of the substrate, the C-1 methoxy group, possibly having steric hindrance with the side chain of Leu-862, prefers the *cis* position. The slightly different conformation of the side chains of residues Leu-857 and Arg-590 in the pocket geometries may also contribute to the different orientation of the methoxy group. Despite the different conformations predicted for the C-1 methoxy group, the orientation of the benzene ring of α -asarone and the conformation of both C-2 and C-4 methoxy groups and the propenyl side chain were predicted to be the same in all binding geometries (Fig. 6).

The docking model proposed for α -asarone prompts for modifications that could improve its affinity for HMGR and increase the hypercholesterolemic activity. The C-4 methoxy group may be replaced by a hydroxyl group to increase the polar interactions with the side chains of the catalytic important residues Lys-691 and Glu-559. Also, a cavity formed by the residues of the *cis* loop may be occupied by a larger group attached to the C-2 oxygen atom of α -asarone that could make polar interactions with these residues.

A further possible modification is to increase the size of the side chain (hydrophobic group) of α -asarone to resemble the hydrophobic group of statins. It has been suggested that the contacts associated with this hydrophobic group of statins, such as rosuvastatin, are predominantly responsible for their nanomolar K_i values.¹²

In conclusion, we have developed a docking model for α -asarone with HMGR, using an automated docking approach. To the best of our knowledge, this is the first automated docking study reported with the recently resolved structure of HMGR. The docking program used, AutoDock, successfully reproduced the binding modes of crystal structures. The docking model suggests that, similar to statins of type 1 (simvastatin) and type 2 (rosuvastatin), α -asarone inhibits HMGR by binding to the enzyme's active site, thus sterically preventing the substrate from binding. The binding energies calculated by the docking program are in good agreement with the observed inhibitory abilities of α -asarone, simvastatin, and rosuvastatin. According to the docking model, the C-1, C-2, and C-4 oxygen atoms make polar contacts with residues of the binding pocket and may form several hydrogen bonds. Their interaction with the catalytic important residues Lys-691 and Glu-559 is remarkable. The aromatic ring and the propenyl side chain of α -asarone makes van der Waals contacts with residues of the binding pocket. The present results help

to explain, at the molecular level, several previous SAR studies for α -asarone analogues which indicate the importance of the C-1 and C-2 oxygens of the methoxy groups and a hydrocarbon side chain in the *meta* and *para* positions to these groups. The docking results suggest modifications to the α -asarone structure to improve its HMGR inhibitory activity, which will be valuable for the structure-based design of further α -asarone analogues.

Acknowledgements

The authors are very grateful to Dr. A. Olson and his colleagues at the Scripps Research Institute for providing AutoDock, and to Dr. Guillermo Ceballos for providing the X-ray file of the enzyme. We thank CONACyT for financing project G-34851-M and DGAPA for project NI 202101. J.L.M.-F. is grateful to CONACyT and DGAPA, UNAM, for the PhD scholarships. G.C. and J.T. are fellows of the EDI/IPN and COFAA/IPN programs.

References and notes

- Kreeman, V.; Shottova, N.; Walterová, D.; Ulrichová, J.; Simánek, V. *Planta Med.* **1998**, *64*, 138.
- Lusis, A. J. *Nature* **2000**, *407*, 233.
- Goldstein, J. L.; Brown, M. S. *Nature* **1990**, *343*, 425.
- Jakobisiak, M.; Golab, J. *Int. J. Oncol.* **2003**, *23*, 2055.
- Wierzbicki, A. J. *Expert Opin. Pharmacother.* **2001**, *2*, 819.
- Miyazaki, A.; Koieyuma, T.; Shimada, Y.; Kikuchi, T.; Ito, K.; Kasanuki, N.; Koga, T. *J. Atheroscler. Thromb.* **2004**, *11*, 22.
- Vivancos-Mora, J.; Leon-Colombo, T.; Monforte-Dupret, C. *Rev. Neurol.* **1999**, *29*, 857.
- Sit, S. Y.; Parker, R. A.; Motoc, Y.; Han, W.; Balasubramanian, N.; Catt, J. D.; Brown, P. J.; Harte, W. E.; Thompson, M. D.; Wright, J. *J. Med. Chem.* **1990**, *33*, 2982.
- Roth, B. D.; Blankley, C. J.; Chucholowski, A. W.; Ferguson, E.; Hoefle, M. L.; Ortwine, D. F.; Newton, R. S.; Sekerke, C. S.; Sliskovic, D. R.; Stratton, C. D.; Wilson, M. W. *J. Med. Chem.* **1991**, *34*, 357.
- Beck, G.; Kessler, K.; Baader, E.; Bartmann, W.; Bergmann, A.; Granzer, E.; Jendralla, H.; Kerkjart, B. V.; Krause, R.; Paulul, E.; Shubert, W.; Wess, G. *J. Med. Chem.* **1990**, *33*, 52.
- Istvan, E. S.; Palnitkar, M.; Buchanan, S. K.; Deisenhofer, J. *EMBO J.* **2000**, *19*, 819.
- Istvan, E. S.; Deisenhofer, J. *Science* **2001**, *292*, 1160.
- Istvan, E. S.; Deisenhofer, J. *Biochim. Biophys. Acta* **2000**, *1529*, 9.
- Enríquez, R. G.; Chávez, M. A.; Jaúregui, F. *Phytochemistry* **1980**, *19*, 2024.
- Martínez, M. *Las Plantas Medicinales de México*; Botas: México City, 1992; pp 125–127.
- Garduño, L.; Salazar, M.; Salazar, S.; Morelos, M. E.; Labarríos, F.; Tamariz, J.; Chamorro, G. A. *J. Ethnopharmacol.* **1997**, *55*, 161.
- Chamorro, G.; Salazar, M.; Salazar, S.; Mendoza, T. *Rev. Inv. Clin.* **1993**, *45*, 597.
- Poplawski, J.; Lozowicka, B.; Dubis, A. T.; Lachowska, B.; Witkowski, S.; Siluk, D.; Petruszewicz, J.; Kaliszán, R.;

- Cybulski, J.; Strzalkowska, M.; Chilmonczyk, Z. *J. Med. Chem.* **2000**, *43*, 3671.
19. Momin, A. A.; Nair, M. G. *J. Agric. Food Chem.* **2002**, *50*, 4475.
 20. Rodríguez-Páez, L.; Juárez-Sánchez, M.; Antúnez-Solís, J.; Baeza, I.; Wong, C. *Phytomedicine* **2003**, *10*, 397.
 21. Automated docking was used to locate the appropriate binding orientation and conformation of α -asarone with HMGR. The powerful genetic algorithm method implemented in the program AutoDock 3.0²² was employed. The structures of α -asarone and the proteins were prepared using Sybyl 6.8.²³ All water, adenosine-5'-diphosphate, and 2,3-dihydroxy-1,4-dithiobutane molecules when present were removed from the original RCSB protein data base files. Polar hydrogen atoms were added and Kollman charges,²⁴ atomic solvation parameters and fragmental volumes were assigned to the protein using AutoDock tools (ADT). For docking calculations, Gasteiger–Marsili partial charges²⁵ were assigned to the ligands and nonpolar hydrogen atoms were merged. All torsions were allowed to rotate during docking. The auxiliary program AutoGrid generated the grid maps. Each grid was centered at the crystal structure of the substrate or the corresponding statin. The grid dimensions were $23 \times 23 \times 23 \text{ \AA}^3$ with points separated by 0.375 \AA . The grid dimensions were large enough to cover the substrate, the statins and the NADP(H) binding sites. Lennard–Jones parameters 12–10 and 12–6, supplied with the program, were used for modeling H-bonds and van der Waals interactions, respectively. The distance-dependent dielectric permittivity of Mehler and Solmajer²⁶ was used for the calculation of the electrostatic grid maps. For all ligands, random starting positions, random orientations, and torsions were used. The translation, quaternion, and torsion steps were taken from default values in AutoDock. The Lamarckian genetic algorithm and the pseudo-Solis and Wets methods were applied for minimization, using default parameters. The number of docking runs was 100. The population in the genetic algorithm was 50, the energy evaluations were 250,000, and the maximum number of iterations 27,000. The complexes of α -asarone with HMGR resulting from molecular docking were further structurally optimized with the Tripos force field.²⁷ During minimization, atoms within 6 \AA from the ligand were free to move (other atoms were fixed). For complex optimization, α -asarone with Mülliken charges obtained from ab initio (HF/6-31G**) single-point energy calculation was used. For the protein, Kollman all charges encoded in Sybyl 6.8 were employed. The ab initio calculations were conducted with Spartan'02.²⁸
 22. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comp. Chem.* **1998**, *19*, 1639.
 23. Sybyl, version 6.8; Tripos, Inc.: St. Louis, MO, 2001.
 24. Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765.
 25. Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219.
 26. Mehler, E. L.; Solmajer, T. *Protein Eng.* **1991**, *4*, 903.
 27. Purcel, W. P.; Singer, J. A. *J. Chem. Eng. Data* **1967**, *12*, 235–246. Details of the implementation are given in *Sybyl 6.6 Force Field Manual*; Tripos: St. Louis, MO, 1999; p 80.
 28. Kong, J.; White, C. A.; Krylov, A. I.; Sherrill, C. D.; Adamson, R. D.; Furlani, T. R.; Lee, M. S.; Lee, A. M.; Gwaltney, S. R.; Adams, T. R.; Ochsenfeld, C.; Gilbert, A. T. B.; Kedziora, G. S.; Rassolov, V. A.; Maurice, D. R.; Nair, N.; Shao, Y.; Besley, N. A.; Maslen, P. E.; Dombroski, J. P.; Daschel, H.; Zhang, W.; Korambath, P. P.; Baker, J.; Byrd, E. F. C.; Van Vooris, T.; Oumi, M.; Hsu, C.-P.; Ishikawa, N.; Florian, J.; Warshel, A.; Johnson, B. G.; Gill, P. M. W.; Head-Gordon, M.; Pople, J. A. *J. Comput. Chem.* **2000**, *21*, 1532.
 29. Brookhaven Protein Data Bank. <http://www.rcsb.org>.
 30. McTaggart, F.; Buckett, L.; Davidson, R.; Holdgate, G.; McCormick, A.; Scheneck, D.; Smith, G.; Warwick, M. *Am. J. Cardiol.* **2001**, *87*(suppl.), 28B.
 31. Holdgate, G. A.; Ward, W. H. J.; McTaggart, F. *Biochem. Soc. Trans.* **2003**, *31*, 528.
 32. Cruz, A.; Garduño, L.; Salazar, M.; Martínez, E.; Jiménez-Vázquez, H. A.; Díaz, F.; Chamorro, G.; Tamariz, J. *Arzneimittel-Forsch.* **2001**, *51*, 535.
 33. Cruz, M. C.; Salazar, M.; Garciafigueroa, Y.; Hernández, D.; Díaz, F.; Chamorro, G.; Tamariz, J. *Drug Dev. Res.* **2003**, *60*, 186.

Quantitative structure–activity relationship analysis of pyridinone HIV-1 reverse transcriptase inhibitors using the k nearest neighbor method and QSAR-based database mining

Jose Luis Medina-Franco^a, Alexander Golbraikh^b, Rafael Castillo^a & Alexander Tropsha^{b,*},

^aDepartamento de Farmacia, Facultad de Química, Universidad Nacional Autónoma de México, 04510, Mexico City, Mexico; ^bDivision of Medicinal Chemistry and Natural Products, School of Pharmacy, University of North Carolina, CB# 7360, Chapel Hill, NC, 27599-7360, USA

Received 23 November 2004; accepted 29 March 2005

© Springer 2005

Key words: database mining, docking, HIV-1, k NN-QSAR, non-nucleoside inhibitors, reverse transcriptase

Summary

We have developed quantitative structure–activity relationship (QSAR) models for 44 non-nucleoside HIV-1 reverse transcriptase inhibitors (NNRTIs) of the pyridinone derivative type. The k nearest neighbor (k NN) variable selection approach was used. This method utilizes multiple descriptors such as molecular connectivity indices, which are derived from two-dimensional molecular topology. The modeling process entailed extensive validation including the randomization of the target property (Y-randomization) test and the division of the dataset into multiple training and test sets to establish the external predictive power of the training set models. QSAR models with high internal and external accuracy were generated, with leave-one-out cross-validated R^2 (q^2) values ranging between 0.5 and 0.8 for the training sets and R^2 values exceeding 0.6 for the test sets. The best models with the highest internal and external predictive power were used to search the National Cancer Institute database. Derivatives of the pyrazolo[3,4-*d*]pyrimidine and phenothiazine type were identified as promising novel NNRTIs leads. Several candidates were docked into the binding pocket of nevirapine with the AutoDock (version 3.0) software. Docking results suggested that these types of compounds could be binding in the NNRTI binding site in a similar mode to a known non-nucleoside inhibitor nevirapine.

Abbreviations: AIDS – Acquired immunodeficiency syndrome; HIV – Human immunodeficiency virus; k NN – k Nearest neighbor; MZ – MolConnZ; NNRTIs – Non-nucleoside reverse transcriptase inhibitors; QSAR – Quantitative structure–activity relationships; RT – Reverse transcriptase

Introduction

The reverse transcriptase (RT) of human immunodeficiency virus type 1 (HIV-1) is an attractive target to treat the acquired immune deficiency syndrome (AIDS) for which no universal successful

chemotherapy is available to date [1, 2]. RT inhibitors are classified as nucleoside or non-nucleoside depending on their mechanism of action. One advantage of the latter group is that they lack toxic effects associated with the nucleosides. To date, three non-nucleoside RT inhibitors have been approved for the clinical use including nevirapine (VIRAMUNE), delavirdine (RESCRIPTOR) and efavirenz (SUSTIVA) [1].

*To whom correspondence should be addressed. Tel.: +1-919-966-2955; Fax: +1-919-966-0204; E-mail: alex_tropsha@unc.edu

47 However, all these and other approved drugs
48 induce drug resistant variants of HIV-1.

49 Pyridinone derivatives [3] (cf., Table 1) are a
50 class of non-nucleoside reverse transcriptase inhib-
51 itors (NNRTIs). Kinetic studies [3] and the
52 analysis of resistance mutations [4] indicated that
53 these derivatives should be binding at the same
54 active site of HIV RT as other NNRTIs. Extensive
55 structure–activity relationships (SAR) studies have
56 been conducted for these compounds leading to
57 potent inhibitors [5–8]. Compound 27 (cf.,
58 Table 1) showed good activity in clinical studies
59 but resistant strains of the virus emerged [9, 10].
60 Several molecular modeling studies of the pyridi-
61 none derivatives have been reported including
62 QSAR modeling using Hansch method [11, 12],
63 3D-QSAR approach [13], and docking [14].

64 Several crystal structures of HIV-1 RT in the
65 apo form [15] and in complex with different
66 NNRTIs were published [16–28] and successful
67 structure-based design studies leading to the
68 development of active NNRTIs have been re-
69 ported [29]. However, no crystal structure for the
70 RT-pyridinone complex is available. The great
71 flexibility of the NNRTIs binding pocket is a
72 major problem to apply this technique [17], and it
73 is feasible that chemical structures that belong to
74 different congeneric series may bind to the active
75 site in different modes. Thus, ligand based
76 approaches such as quantitative structure–activity
77 relationships (QSAR) modeling may present a
78 plausible alternative to the structure-based design
79 of NNRTIs. In fact, recent studies in our group
80 have established that thoroughly validated QSAR
81 models could be used to mine chemical databases
82 to discover novel compounds with the desired
83 biological activity. For instance, the k -nearest-
84 neighbor (k NN) variable selection QSAR
85 approach [30, 31] has been applied successfully to
86 the discovery of novel anticonvulsants [32, 33].

87 In this paper, we have applied the k NN QSAR
88 approach to a dataset of 44 pyridinone derivatives
89 (Table 1) [7, 8]. Our objective was to develop
90 robust, validated QSAR models and search, with
91 these models, the National Cancer Institute data-
92 base NCISMA99 [34] for novel NNRTIs. Several
93 compounds from this database have already been
94 implicated as anti-AIDS agents [35], suggesting
95 that NCISMA99 may continue to provide novel
96 compounds with the anti-HIV activity. Screening
97 the NCISMA99 database with our QSAR models

identified several potential leads with high-pre- 98
dicted activity. These compounds were docked 99
into the binding pocket of nevirapine in the HIV-1 100
RT, the structure with the highest resolution 101
available to date, to assess the putative binding 102
mode of these computational datamining hits 103
within the HIV-1 RT binding site. 104

Computational methods 105

Dataset and biological data 106

The structures and *in vitro* activity of 44 pyridi- 107
none derivatives used in this study are listed in 108
Table 1. Activities were converted into the corre- 109
sponding $-\log_{10}IC_{50}$ values (pIC_{50}), where IC_{50} is 110
the effective concentration of compound required 111
to achieve 50% of inhibition of RT. The biological 112
data was taken from the literature [7, 8]. 113

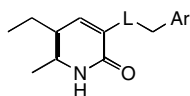
Generation of molecular descriptors 114

All chemical structures were generated using SYB- 115
YL software [36]. Molecular topological indices [37, 116
38] were generated with the MolConnZ program 117
(MZ descriptors) [39]. Overall, MolconnZ generates 118
over 400 different descriptors. Most of these 119
descriptors characterize chemical structure, but 120
several depend on the arbitrary numbering of atoms 121
in a molecule and are introduced solely for book- 122
keeping purposes. In our study, only 193 chemical 123
descriptors were eventually used (after deleting 124
these bookkeeping descriptors as well as those with 125
zero value or zero variance). The descriptors were 126
ranged-scaled prior to subsequent calculations (see 127
[30] for details), since the absolute scales for MZ 128
descriptors can differ by orders of magnitude. 129
Accordingly, our use of range scaling avoided 130
giving descriptors with significantly higher ranges 131
a disproportional weight upon distance calculations 132
as part of k NN procedure in multidimensional MZ 133
descriptor space. All calculations were performed 134
on an SGI Octane at the University of North 135
Carolina's Laboratory for Molecular Modeling. 136

kNN QSAR method 137

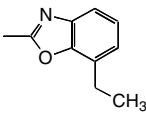

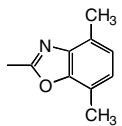
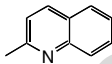
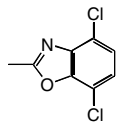
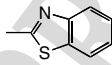
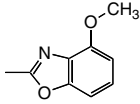
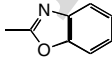
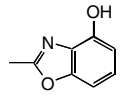
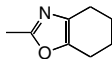
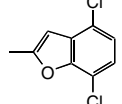
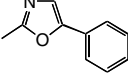
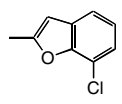
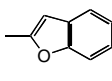
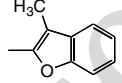
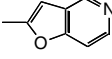

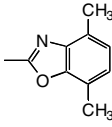

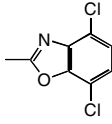
The k NN QSAR method [30] employs the k NN 138
classification principle [40] and the variable selec- 139
tion procedure. Briefly, a subset of $nvar$ (number of 140

Table 1. Structures, HIV-1 RT inhibitory activity, experimental and calculated by *k*NN QSAR models, of 44 pyridinone derivatives used in this study.



Compd.	L	Ar	Experimental pIC ₅₀	Calculated pIC ₅₀	Compd.	L	Ar	Experimental pIC ₅₀	Calculated pIC ₅₀
1	NH		6.68	6.49	13	NH		5.57	5.57
2	NH		6.48	6.39	14	NH		5.36	5.33
3	NH		6.55	6.70	15	NH		5.27	4.91
4	NH		6.47	6.11	16	NH		5.12	5.54
5	NH		6.46	6.54	17	NH		4.82	4.44
6	NH		6.34	6.04	18	NH		4.65	5.07
7	NH		6.30	6.31	19	NH		4.49	4.54
8	NH		6.28	6.10	20	NH		3.98	4.54
9	NH		5.96	5.78	21	NH		6.92	6.80
10	NH		5.72	5.51	22	NH		5.90	5.83
11	NH		5.63	5.79	23	NH		5.78	5.81
12	NH		5.60	5.68	24	NH		7.26	7.21

Table 1. Continued.

Compd.	L	Ar	Experimental pIC ₅₀	Calculated pIC ₅₀	Compd.	L	Ar	Experimental pIC ₅₀	Calculated pIC ₅₀
25	NH		6.59	6.43	35	CH ₂		3.59	4.10
26	NH		7.70	7.65	36	CH ₂		5.61	5.82
27	NH		7.72	7.61	37	CH ₂		6.43	6.43
28	NH		6.74	6.56	38	CH ₂		7.64	7.18
29	NH		6.36	6.43	39	CH ₂		7.24	7.02
30	NH		7.24	7.22	40	CH ₂		6.41	6.12
31	NH		6.47	6.51	41	CH ₂		6.89	6.72
32	NH		5.71	6.51	42	CH ₂		4.87	5.19
33	CH ₂		4.30	4.64	43	CH ₂		7.55	7.65
34	CH ₂		5.08	5.43	44	CH ₂		7.85	7.67

141 selected variables) descriptors is selected randomly
 142 as a hypothetical descriptor pharmacophore
 143 (HDP). The *nvar* is set to different values to obtain
 144 the best q^2 possible. The HDP is validated by leave-
 145 one-out cross-validation, where each compound is
 146 eliminated from the training set and its biological
 147 activity is predicted as the average activity of k
 148 most similar molecules ($k = 1-5$). The similarity is
 149 characterized by the Euclidean distance between
 150 compounds in multidimensional descriptor space.
 151 A method of simulated annealing with the Metrop-
 152 olis-like acceptance criteria is used to optimize the
 153 variable selection. Further details of the k NN
 154 method implementation, including the description
 155 of the simulated annealing procedure used for
 156 stochastic sampling of the descriptor space, are
 157 given elsewhere [30].

158 Robustness of QSAR models

159 The q^2 values for the models for experimental
 160 training sets were compared to those derived for
 161 so-called random datasets, which are generated by
 162 random shuffling of compounds' activities. The
 163 statistical significance of QSAR models for train-
 164 ing sets was evaluated with the standard hypoth-
 165 esis testing method [41].

166 In this approach, two alternative hypothesis are
 167 formulated: (1) for H_0 , $h = \mu$; (2) for H_1 , $h < \mu$,
 168 where μ is the average value of q^2 for random
 169 datasets and h is the q^2 value for the actual dataset.
 170 Thus, the null hypothesis, H_0 , states that the
 171 QSAR model for the actual dataset is not signifi-
 172 cantly better than random models, whereas the
 173 alternative hypothesis, H_1 , assumes the opposite
 174 (i.e., that the actual model is significantly better
 175 than random models). The decision-making is
 176 based on a standard one-tail test, which involves
 177 the following procedure.

- 178 (1) Determine the average value of q^2 (μ) and its
 179 standard deviation (σ) for random datasets.
- 180 (2) Calculate the Z score that corresponds to the
 181 q^2 value for the actual dataset: $Z = (h - \mu) / \sigma$
- 182 (3) Compare this Z score with the tabular critical
 183 values of Z_c at different levels of significance
 184 (α) [41] to determine the level at which H_0
 185 should be rejected. If the Z score is higher
 186 than tabular values of Z_c (cf., Table 2), one
 187 concludes that at the level of significance that
 188 corresponds to that Z_c , H_0 should be rejected

Table 2. Frequently used α values and the corresponding critical values of Z_c for the one-tail test [41].

α	Z_c
0.10	1.28
0.05	1.64
0.01	2.33

and, therefore, H_1 should be accepted. In this case, it is concluded that the result obtained for the actual dataset is statistically better than those obtained for random datasets at the given level of significance.

Model validation: training and test set compound selection 194 195

To obtain training and test sets, we used a sphere exclusion algorithm described in detail elsewhere [42]. This algorithm allows construction of training and test sets of various sizes that cover the entire descriptor space of all compounds. To estimate the predictive power of a QSAR model, the following parameters were used [43]. (i) Correlation coefficient R^2 between the predicted and observed activities for the test set; (ii) Coefficients of determination [43] (predicted versus observed activities R^2 and observed versus predicted activities $R_0'^2$); (iii) slopes k and k' of regression lines (predicted versus observed activities, and observed versus predicted activities) through the origin. We suggest that a QSAR model has an acceptable predictive power if the following conditions are satisfied [42]: 213

$$q^2 > 0.5; \quad (1)$$

$$R^2 > 0.6; \quad (2) \quad 215$$

$$|R_0^2 - R_0'^2| < 0.3 \quad (3) \quad 217$$

$$0.85 \leq k \leq 1.15 \text{ or } 0.85 \leq k' \leq 1.15. \quad (4) \quad 219$$

221

NCI database mining 222

Our database-mining strategy makes use of validated QSAR models developed for the available 223 224

225 series of biologically active compounds. The
 226 general flowchart of the database-mining proce-
 227 dure is shown in Figure 1 and includes the
 228 following major steps; a similar approach was
 229 successfully used recently for the discovery of
 230 novel anticonvulsant agents [33].

- 231 1. Develop validated variable selection QSAR
 232 models for a dataset of compounds with
 233 known structures and activities; define
 234 descriptor pharmacophores [44] and
 235 applicability domains [45] for all models.
- 236 2. Compute chemical descriptors used in
 237 QSAR model development for all com-
 238 pounds in the available chemical data-
 239 base(s); in our studies we have used
 240 molecular topological indices calculated with
 241 the MolConnZ program [39].
- 242 3. Calculate chemical similarity values (we use
 243 the Euclidean distance in the descriptor phar-
 244 macophore space) between all active probes
 245 (i.e., molecules used for QSAR model devel-
 246 opment) and every structure in the database.
- 247 4. Rank all database structures by their similar-
 248 ity to a probe(s), and select M structures
 249 within certain similarity threshold.
- 250 5. Predict biological activity values for these M
 251 structures based on preconstructed QSAR
 252 models using applicability domain.
- 253 6. Select structures to have high values of bio-
 254 logical activity as *computational hits*.

Flexible docking

255

256 The geometry of chemical structures selected
 257 from the database mining were optimized in
 258 Sybyl with the Powell method [46] using the
 259 Tripos force field [47]. A gradient convergence
 260 criterion with a value of 0.05 kcal/mol was used.
 261 The docking experiments were performed with
 262 AutoDock 3.0 [48]. In short, AutoDock performs
 263 an automated docking of the ligand with user-
 264 specified dihedral flexibility within a protein rigid
 265 binding site. The program performs several runs
 266 in each docking experiment. Each run provides
 267 one predicted binding mode with the highest
 268 interaction energy.

269 The crystal structure of RT in complex with
 270 nevirapine [16] (cf. Figure 5) was obtained from
 271 the Protein Data Bank (PDB code 1vrt) [49]. For
 272 the validation of the docking protocol, the nevi-
 273 rapine coordinates were removed from the crystal
 274 structure and the bond orders were checked.
 275 Gasteiger charges [50] were assigned to the struc-
 276 ture and the non-polar hydrogens were treated
 277 implicitly. The cyclopropyl group was allowed to
 278 rotate during docking. For the subsequent docking
 279 studies using compounds identified from database
 280 mining, nevirapine, magnesium ion and water
 281 molecules were removed from the original struc-
 282 ture. With the AutoDock Tools [51], polar
 283 hydrogen atoms were added and Kollman charges
 284 [52], atomic solvation parameters and fragmental

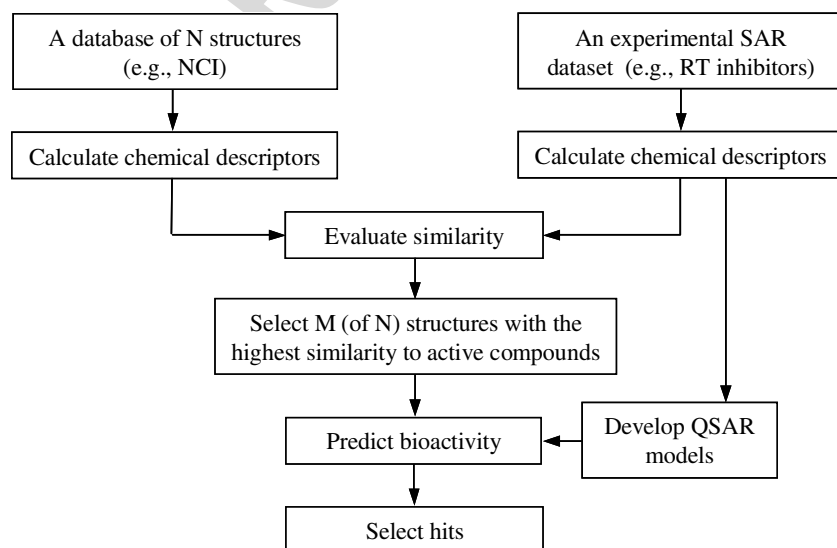


Figure 1. Flowchart of database mining that employs predictive QSAR models.

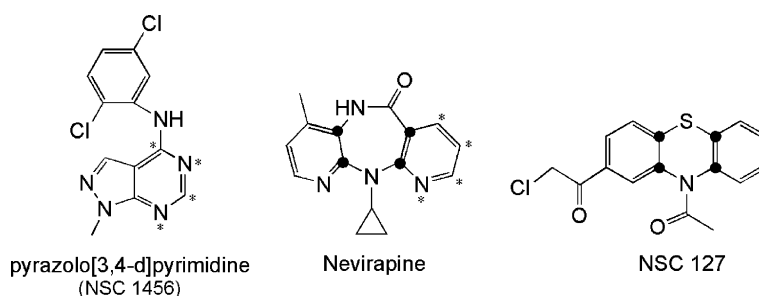


Figure 2. Atoms used to superimpose the structure of nevirapine with the pyrazolo[3,4-d]pyrimidines (marked with asterisks) and with NSC 127 (marked with filled circles) before docking.

285 volumes were assigned to the protein. The starting
 286 orientation for docking studies was obtained by
 287 superimposing the structure of the molecule to be
 288 docked with that of nevirapine using the FIT
 289 ATOMS command in Sybyl. The atoms consid-
 290 ered for the superimposition are showed in
 291 Figure 2.

292 The grid maps were generated with the Auto-
 293 Grid module. Each grid was centered at the crystal
 294 structure of the corresponding NNRTI. Their
 295 dimensions were $23 \times 23 \times 23 \text{ \AA}^3$ with points
 296 separated by 0.375 \AA . The default Lennard-Jones
 297 parameters were used for modeling H-bonds and
 298 van der Waals interactions. The distance-depend-
 299 ent dielectric permittivity of Mehler and Solmajer
 300 [53] was used for the calculation of the electrostatic
 301 grid maps. For all docked molecules, random
 302 starting positions, random orientations and tor-
 303 sions were used. The Lamarckian genetic algo-
 304 rithm and the pseudo-Solis and Wets methods
 305 were applied for minimization using default
 306 parameters. The number of docking runs was
 307 100. The number of populations for the genetic
 308 algorithm optimization was 50, the energy evalu-
 309 ations parameter was set to 250,000 and the
 310 maximum number of iterations was 27,000.
 311 Following docking calculations, the 100 best poses
 312 were clustered into groups with RMS deviations
 313 within each group lower than 1.0 \AA . The clusters
 314 were ranked by the lowest energy representative of
 315 each cluster.

316 Results and discussion

317 The results obtained with *k*NN QSAR analysis are
 318 discussed in terms of the optimized q^2 values,
 319 variable selection, actual versus predicted

activities, and statistical significance of the result-
 ing QSAR models.

QSAR models and their robustness

In the *k*NN QSAR method, *nvar* can be set to any
 value that is less than the total number of
 descriptors. Since the optimal number for *nvar* is
 not known *a priori*, multiple models have to be
 generated for different values of *nvar*. As previ-
 ously discussed, the robustness of a QSAR model
 should be established by comparing results for the
 actual dataset with those for datasets with the
 randomized activity values. Figure 3 shows a plot
 of q^2 versus *nvar* for the actual and random data
 sets. Every q^2 value is the average of 10 indepen-
 dent computations. Overall, we obtained signifi-
 cantly higher q^2 values for the actual dataset
 compared to those for randomized activity data-
 sets (see Figure 3).

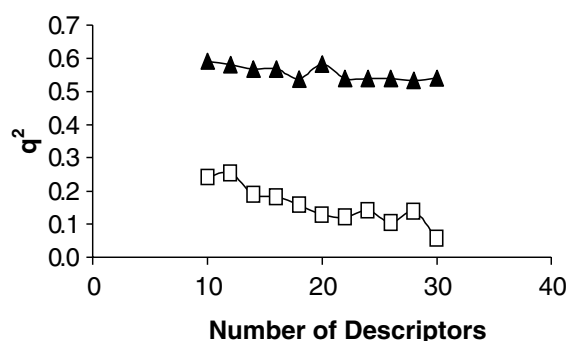


Figure 3. Plots of q^2 versus the number of descriptors selected for the best *k*NN QSAR models for 44 pyridinone HIV-1 RT inhibitors. The results for both actual and random (with shuffled activity values) datasets are shown. Every q^2 value is the average of 10 independent calculations. Triangles represent the actual dataset, and squares represent the random dataset.

338 The statistical examination of the results was
 339 performed with one-tail hypothesis testing as
 340 described in the Computational methods. The
 341 average q^2 value for 30-descriptor models obtained
 342 for 10 different randomized datasets was 0.06 with
 343 a standard deviation of 0.18. The Z score for the
 344 most significant 30-descriptor model for the actual
 345 dataset was 2.71. This indicates that the probabili-
 346 ty that the k NN QSAR model constructed for the
 347 real dataset is random is less than 0.01.

348 k NN QSAR model validation

349 To obtain reliable and truly predictive QSAR
 350 models, it is necessary to demonstrate that the
 351 training set models can accurately predict activi-
 352 ties of compounds in external datasets. We
 353 accept models with q^2 values for the training set
 354 greater than 0.5 and R^2 values for predicted
 355 versus actual activities of the test set compounds
 356 greater than 0.6 [42, 54]. Table 3 presents the 10
 357 best models obtained from multiple k NN analy-
 358 ses. Figure 4 shows actual versus calculated
 359 activity values for the training and test sets based
 360 on the best model. The number of descriptors for
 361 the 10 best models varied between 10 and 26.
 362 Even for a test set as large as 19 compounds a
 363 good model with $q^2 = 0.52$ and $R^2 = 0.75$ was
 364 still obtained.

365 Database mining

366 Several compounds from the NCI database whose
 367 anti-AIDS activity has not been yet documented

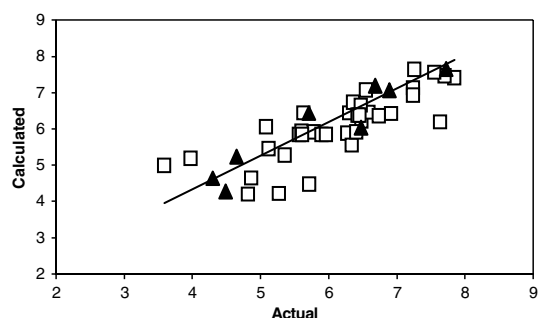


Figure 4. Actual versus calculated activity values for the training (squares) and test (triangles) set using k NN QSAR model 1 ($q^2 = 0.62$, $R^2 = 0.89$).

were predicted to be active. The structures of some
 of them are shown in Figure 5. It was especially
 interesting to observe the structural similarity of
 the compound NSC 127 and nevirapine. The core
 of both structures is a three-ring system involving
 two six-member aromatic rings. Both compounds
 have a small substituent (methylketone or
 cyclopropyl group) attached to a nitrogen located
 at the central ring (Figure 5).

Flexible docking. Validation of the docking protocol

The docking protocol was validated by removing
 nevirapine from the crystal structure and then
 predicting the binding mode with the AutoDock.
 All 100 docked energy conformations formed a
 single cluster within 1.0 Å as described in
 Computational methods. The RMSD between the
 lowest docked energy conformation found (docked
 energy = -9.67 kcal/mol; average docked energy

Table 3. Ten best k NN QSAR models.

Model number	Size of the training set (number of cpds.)	Size of the test set (number of cpds.)	Number of descriptors	q^2 (training set)	R^2 (test set)	R_0^2	R_0^2
1	36	8	14	0.62	0.89	0.88	0.89
2	36	8	14	0.60	0.87	0.85	0.87
3	36	8	12	0.61	0.75	0.70	0.75
4	36	8	26	0.55	0.74	0.48	0.72
5	36	8	10	0.58	0.72	0.68	0.71
6	31	13	10	0.59	0.72	0.72	0.68
7	31	13	20	0.63	0.71	0.60	0.71
8	28	16	16	0.63	0.75	0.46	0.72
9	27	17	24	0.53	0.71	0.48	0.70
10	25	19	10	0.52	0.75	0.72	0.42

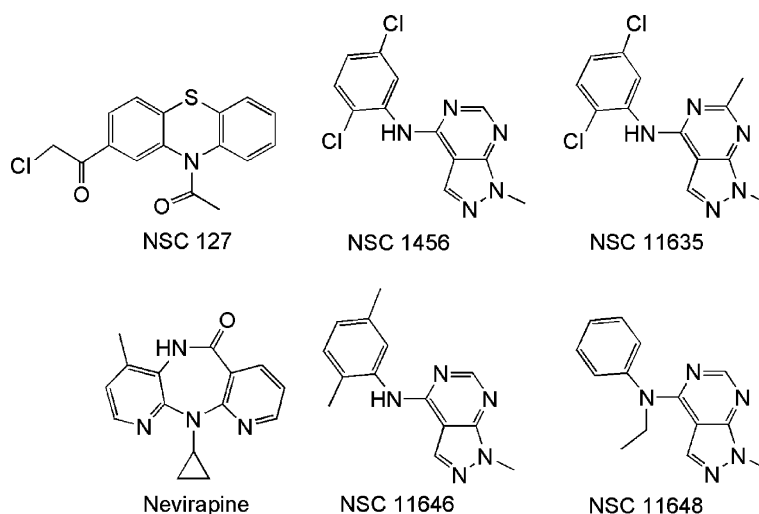


Figure 5. Chemical structures of the compounds found in the database mining and nevirapine.

386 for the cluster = -9.66 kcal/mol) and the actual
 387 coordinates was 0.70. This demonstrated that the
 388 protocol successfully predicted the binding position
 389 of nevirapine.

390 *Docking of selected molecules resulting from the* 391 *database mining*

392 The compounds predicted most active from the
 393 database mining were visually inspected. Based on
 394 some common features with the known NNRTIs,
 395 e.g., the possibility to adopt a butterfly-like con-
 396 formation within the NNRTIs binding pocket
 397 [16–20], the selected compounds were docked into
 398 the binding site of nevirapine. The results are
 399 summarized in Table 4.

400 The pyrazo[3, 4-*d*]pyrimidines NSC 1456,
 401 11635, 11646 and 11648 [55] were found to adopt

a butterfly-like geometry inside the nevirapine
 binding pocket. Moreover, the predicted binding
 mode is similar to the binding mode of nevirapine.
 The binding pocket of most pyrazo[3,4-*d*]pyrimi-
 dines is formed by amino acids Pro95, Leu100,
 Lys101, Lys102, Lys103, Val106, Tyr181, Tyr188,
 Phe227, Trp229, Leu234, His235, Pro236, Asp237,
 Tyr318 of the p66 subunit and Glu138 of the p51
 subunit [16]. Most of these amino acids are
 common to the binding pocket of other NNRTIs
 [16–28]. The binding mode of the predicted most
 active compound, NSC 11635, into the NNRTIs
 binding pocket is depicted in Figure 6. The bind-
 ing mode of nevirapine is shown for comparison.
 The methyl group at position 6 of the pyrazolo[3,4-*d*]pyrimidine ring of NSC 11635 showed
 additional contacts with Pro225 of the p66 sub-
 unit. These additional contacts can explain the

Table 4. The results of docking of several selected compounds found by the means of database mining into the binding pocket of nevirapine.

Compd.	Cluster ^a	Members	Lowest energy (kcal/mol) ^b	Average energy (kcal/mol) ^c
NSC 11635	1	93	-10.03	-10.02
NSC 11646	1	100	-9.37	-9.36
NSC 1456	1	62	-9.34	-9.29
NSC 11648	1	9	-8.61	-8.57
NSC 127	2	43	-10.28	-10.16
	4	40	-9.74	-9.71

^aThe best clusters of a total 100 runs are shown.

^bValue for the optimal structure in the cluster. Docking energies are as determined by AutoDock.

^cAverage value of docking energies of all members in the cluster.

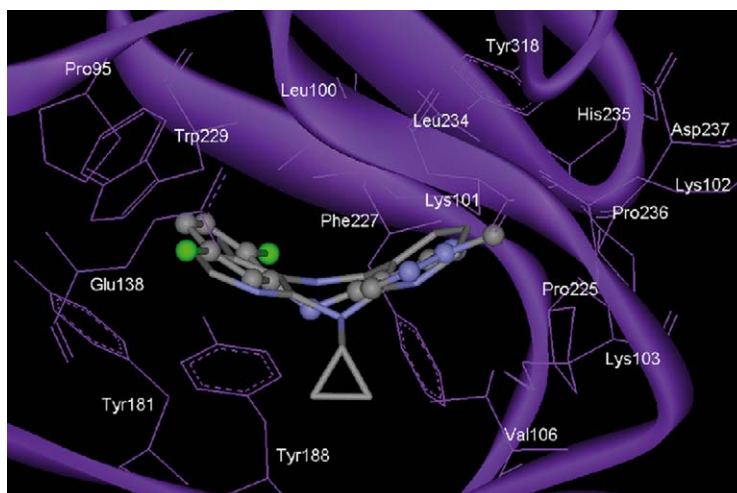


Figure 6. Predicted binding mode of NSC 11635 (balls and sticks) in the binding pocket of HIV-1 RT. Nevirapine (sticks) is displayed for comparison. Hydrogens are omitted for clarity.

420 lowest energy interaction calculated by AutoDock
 421 for NSC 11635 compared with the energy
 422 interactions calculated for NSC 11646 and NSC
 423 1456 (Table 4). This methyl group, that is missing
 424 in the structures of NSC 11646 and NSC 1456
 425 (Figure 5), could be an important structural fea-
 426 ture in the putative activity of NSC 11635. NSC
 427 11648 presented additional interactions between
 428 the ethyl group and the amino acids Val179 and
 429 Gly190 of the p66 subunit (cf., Figure 7d). How-
 430 ever, for this compound contacts with Asp237 and
 431 Glu138 were not observed.

432 In order to compare the binding mode pre-
 433 dicted for the pyrazo[3,4-*d*]pyrimidines with the
 434 predicted for the pyridinone derivatives, the bind-
 435 ing mode of the former compounds was superim-
 436 posed with the binding mode of the potent
 437 inhibitor **27** previously reported [14]. We have
 438 shown that the predicted binding mode of **27** is
 439 representative of most potent pyridinone deriva-
 440 tives [13, 14]. The comparisons are shown in
 441 Figure 7. The predicted binding mode of all
 442 compounds is similar to the binding mode of the
 443 pyridinone derivative. The phenyl ring of NSC
 444 11635, NSC 11646 and NSC 1456, all with a 2,5-
 445 dichloro or 2,5-dimethyl substitution, may be
 446 roughly overlapped with the benzoxazol ring of
 447 **27**. Noteworthy, both chlorine or methyl groups of
 448 the phenyl ring of these pyrazo[3,4-*d*]pyrimidines
 449 occupies almost the same binding position that the
 450 chlorine groups of **27** (Figure 7a–c) interacting

mainly with amino acids Pro95, Tyr181, Tyr188
 and Trp229. Interesting to note, it has been shown
 that pyridinone derivatives di-substituted with
 methyl or chlorine groups at the positions 4 and
 7 of the benzoxazol ring are more active than the
 corresponding mono or unsubstituted compounds
 [7, 8, 11]. These observations suggest the putative
 biological importance of the chloro or methyl
 groups at the phenyl ring of compounds NSC
 11635, NSC 11646 and NSC 1456.

451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461 For NSC 11648, which lacks of substituents at
 the phenyl ring, a rotated conformation was
 predicted if compared to the binding orientation
 of the former three pyrazo[3,4-*d*]pyrimidines. The
 pyrazolo[3,4-*d*]pyrimidine ring occupies a similar
 binding region that the benzoxazol ring of **27**
 (Figure 7d). The ethyl group was predicted to
 occupy the same pocket region that the aminom-
 ethylene group of the pyridinone derivative.

462
 463
 464
 465
 466
 467
 468
 469
 470 Table 5 presents a comparison between the
 activity predicted by the *k*NN QSAR models, the
 docked energy and the free energy of binding as
 estimated by AutoDock for four compounds.
 Obviously, there is a good agreement between
 these predictions, with similar relative ranking of
 all compounds by all computational methods.
 Noteworthy, there is a good agreement between
 the activity predicted by *k*NN models for all the
 pyridinone derivatives in the dataset, and the
 experimental activities (Table 1). In docking stud-
 ies for several pyridinone derivatives, we have

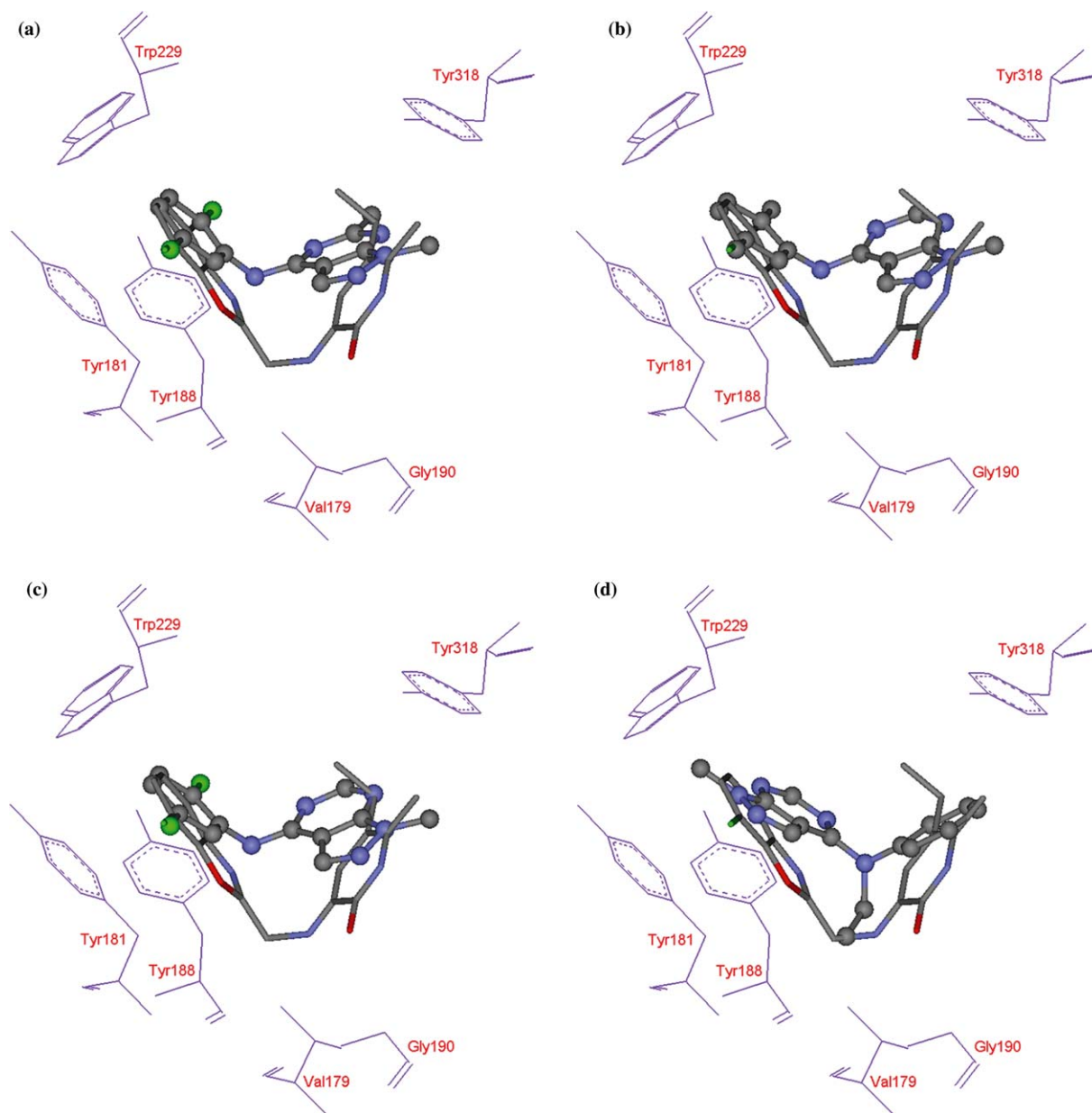


Figure 7. Comparison of the binding conformations of the pyazo[3,4-*d*]pyrimidines (balls and sticks) with **27** (sticks) [13]. (a) NSC 11635. (b) NSC 11646. (c) NSC 1456. (d) NSC 11648. The position of Val179, Tyr181, Tyr188, Gly190, Trp229 and Tyr318 is displayed for reference. Hydrogens are omitted for clarity.

482 reported an acceptable quantitative correlation
 483 between the predicted binding free energy and the
 484 biological activity especially for structural similar
 485 compounds, i.e., all compounds with a benzoxazol
 486 ring [13]. Although a quantitative correlation
 487 between experimental activities and predictions
 488 were better with QSAR, all the above observations

489 suggest that the pyazo[3,4-*d*]pyrimidines represent
 490 promising novel NNRTIs leads.

491 Compound NSC 127 [56] was also found to
 492 adopt a butterfly-like conformation in the nevira-
 493 pine binding site. The two predicted binding
 494 modes, represented by the most populated clusters
 495 2 and 4 (Table 4), are similar to the binding mode

Table 5. Predicted activity by *k*NN best models, docking energy and free energy of binding as calculated by AutoDock for selected molecules.

Compd.	Predicted activity by <i>k</i> NN models (pIC ₅₀) ^a	Docking energy (kcal/mol) ^b	Free energy of binding (kcal/mol) ^b
NSC 11635	6.44	-10.03	-9.40
NSC 11646	6.34	-9.37	-8.72
NSC 1456	6.12	-9.34	-8.65
NSC 11648	5.89	-8.61	-7.98

^aAverage of predictions.

^bValue for the optimal structure.

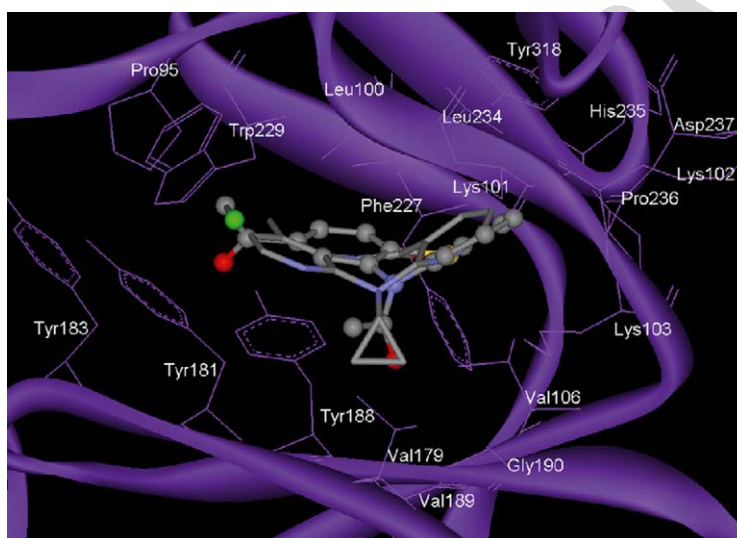


Figure 8. Top ranked binding mode (cluster 2) of NSC 127 (balls and sticks) in the binding pocket of HIV-1 RT. Nevirapine (sticks) is displayed for comparison. Hydrogens are omitted for clarity.

496 of nevirapine. The binding pocket of NSC 127 is
 497 formed by amino acids Pro95, Leu100, Lys101,
 498 Lys102, Lys103, Val106, Val179, Tyr181, Tyr183,
 499 Tyr188, Val189, Gly190, Phe227, Trp229, Leu234,
 500 His235, Pro236, Asp237 and Tyr318 (Figure 8).
 501 As expected, the methylketone moiety of NSC 127
 502 could be occupying the same pocket region that
 503 the cyclopropyl group of nevirapine making
 504 contacts with Val106, Val179, Tyr181, Tyr188,
 505 Val189 and Gly190.

506 Conclusions and future studies

507 We have developed and validated QSAR models
 508 for a series of NNRTIs. The models were further
 509 used to search the NCI database for novel
 510 inhibitors. Several classes of compounds were
 511 predicted to be active. The most promising novel
 512 leads are pyrazolo[3,4-*d*]pyrimidines and pheno-

thiazine derivatives. Flexible docking studies of
 these types of compounds to HIV-1 RT have
 suggested that they could be binding to the
 NNRTIs binding site in a very similar mode to
 other NNRTIs, especially nevirapine.

Biological testing of these types of derivatives,
 e.g., NSC 11635 and NSC 127, is the next step in
 our effort to discover novel anti-HIV-1 com-
 pounds. The approaches employed in this paper
 for the rational discovery of HIV RT inhibitors
 could be adapted for similar studies of many
 biologically active series of compounds.

Acknowledgements

The authors thank Dr. Arthur Olson and his col-
 leagues at the Scripps Research Institute for pro-

528 viding AutoDock and auxiliary programs. J. L.
529 M-F. is grateful to CONACyT and DGAPA,
530 UNAM, for the PhD scholarships. AT acknowl-
531 edges the support from NIH Grant GM066940.

532 References

- 533 1. De Clercq, E., *Int. J. Biochem. Cell Biol.*, 36 (2004) 1800.
534 2. Jacobo-Molina, A. and Arnold, E., *Biochemistry*, 30 (1991)
535 6351.
536 3. Goldman, M.E., Nunberg, J.H., O'Brien, J.A., Quintero,
537 J.C., Schleif, W.A., Freund, K.F., Gaul, S.L., Saari, W.S.,
538 Wai, J.S., Hoffman, J.M., Anderson, P.S., Hupe, D.J.,
539 Emini, E.A. and Stern, A.M., *Proc. Natl. Acad. Sci. USA*,
540 88 (1991) 6863.
541 4. Nunberg, J.H., Schleif, W.A., Boots, E.J., O'Brien, J.A.,
542 Quintero, J.C., Hoffman, J.M., Emini, E.A. and Goldman,
543 M.E., *J. Virol.*, 65 (1991) 4887.
544 5. Saari, W.S., Hoffman, J.M., Wai, J.S., Fisher, T.E.,
545 Rooney, C.S., Smith, A.M., Thomas, C.M., Goldman,
546 M.E., O'Brien, J.A., Nunberg, J.H., Quintero, J.C., Schleif,
547 W.A., Emini, E.A., Stern, A.M. and Anderson, P.S.,
548 *J. Med. Chem.*, 34 (1991) 2922.
549 6. Hoffman, J.M., Wai, J.S., Thomas, C.M., Levin, R.B.,
550 O'Brien, J.A. and Goldman, M.E., *J. Med. Chem.*, 35
551 (1992) 3784.
552 7. Saari, W.S., Wai, J.S., Fisher, T.E., Thomas, C.M.,
553 Hoffman, J.M., Rooney, C.S., Smith, A.M., Jones, J.H.,
554 Bamberger, D.L., Goldman, M.E., O'Brien, J.A., Nunberg,
555 J.H., Quintero, J.C., Schleif, W.A., Emini, E.A. and
556 Anderson, P.S., *J. Med. Chem.*, 35 (1992) 3792.
557 8. Hoffman, J.M., Smith, A.M., Rooney, C.S., Fisher, T.E.,
558 Wai, J.S., Thomas, C.M., Bamberger, D., Barnes, J.L.,
559 Williams, T.M., Jones, J.H., Olson, B.D., O'Brien, J.A.,
560 Goldman, M.E., Nunberg, J.H., Quintero, J.C., Schleif,
561 W.A., Emini, E.A. and Anderson, P.S., *J. Med. Chem.*, 36
562 (1993) 953.
563 9. Davey, R.T., Dewar, R.L., Reed, G.F., Vasudevachari,
564 M.B., Polis, M.A., Kovacs, J.A., Fallon, J., Walker, R.E.,
565 Masur, H., Hanevich, S.E., O'Neill, D.G., Decker, M.R.,
566 Metcalf, J.A., Deloria, M.A., Laskin, O.L., Salzman, N. and
567 Lane, H.C., *Proc. Natl. Acad. Sci. USA*, 90 (1993) 5608.
568 10. Saag, M.S., Emini, E.A., Laskin, O.L., Douglas, J., Lapi-
569 dus, W.I., Schleif, W.A., Whitley, R.J., Hildebrand, C.,
570 Byrnes, V.W., Kappes, J.C., Anderson, K.W., Massari,
571 F.E., Shaw, G.M. and L-697,661 Working Group., *N.*
572 *Engl. J. Med.*, 329 (1993) 1065.
573 11. Garg, R., Gupta, S.P., Gao, H., Babu, M.S., Debnath,
574 A.K. and Hansch, C., *Chem. Rev.*, 99 (1999) 3525.
575 12. Gupta, S.P., In Jucker, E. (Ed.), *Progress in Drug*
576 *Research*. Birkhäuser Verlag, Basel, Switzerland, 2002, vol.
577 58, pp. 252–253.
578 13. Medina-Franco, J.L., Rodríguez-Morales, S., Juárez-Gor-
579 diano, C., Hernández-Campos, A. and Castillo, R., *J.*
580 *Comput. Aid. Mol. Des.*, 18 (2004) 345.
581 14. Medina-Franco, J.L., Rodríguez-Morales, S., Juárez-Gor-
582 diano, C., Hernández-Campos, A., Jiménez-Barbero, J. and
583 Castillo, R., *Bioorg. Med. Chem.*, 12 (2004) 6085.
584 15. Esnouf, R., Ren, J., Ross, C., Jones, Y., Stammers, D. and
585 Stuart, D., *Nat. Struct. Biol.*, 2 (1995) 303.
16. Ren, J., Esnouf, R., Garman, E., Somers, D., Ross, C.,
Kirby, I., Keeling, J., Darby, G., Jones, Y., Stuart, D. and
Stammers, D., *Nat. Struct. Biol.*, 2 (1995) 293.
17. Hopkins, A.L., Ren, J., Esnouf, R.M., Willcox, B.E.,
Jones, E.Y., Ross, C., Miyasaka, T., Walker, R.T., Tanaka,
H., Stammers, D.K. and Stuart, D.I., *J. Med. Chem.*, 39
(1996) 1589.
18. Ren, J., Esnouf, R., Hopkins, A., Ross, C., Jones, Y.,
Stammers, D. and Stuart, D., *Structure*, 3 (1995) 915.
19. Ren, J., Milton, J., Weaver, K.L., Short, S.A., Stuart, D.I.
and Stammers, D.K., *Structure*, 8 (2000) 1089.
20. Ren, J., Nichols, C., Bird, L.E., Fujiwara, T., Sugimoto, H.,
Stuart, D.I. and Stammers, D.K., *J. Biol. Chem.*, 275
(2000) 14316.
21. Esnouf, R.M., Ren, J., Hopkins, A.L., Ross, C.K., Jones,
E.Y., Stammers, D.K. and Stuart, D.I., *Proc. Natl. Acad.*
Sci. USA, 94 (1997) 3984.
22. Ren, J., Diprose, J., Warren, J., Esnouf, R.M., Bird,
L.E., Ikemizu, S., Slater, M., Milton, J., Balzarini, J.,
Stuart, D.I. and Stammers, D.K., *J. Biol. Chem.*, 275
(2000) 5633.
23. Chan, J.H., Hong, J.S., Hunter, III R.N., Orr, G.F.,
Cowan, J.R., Sherman, D.B., Sparks, S.M., Reitter, B.E.,
Andrews, III C.W., Hazen, R.J., St Clair, M., Boone, L.R.,
Ferris, R.G., Creech, K.L., Roberts, G.B., Short, S.A.,
Weaver, K., Ott, R.J., Ren, J., Hopkins, A., Stuart, D.I.
and Stammers, D.K., *J. Med. Chem.*, 44 (2001) 1866.
24. Hsiou, Y., Das, K., Ding, J., Clark, A.D. Jr., Kleim, J.P.,
Rosner, M., Winkler, I., Riess, G., Hughes, S.H. and
Arnold, E., *J. Mol. Biol.*, 284 (1998) 313.
25. Ren, J., Esnouf, R.M., Hopkins, A.L., Stuart, D.I. and
Stammers, D.K., *J. Med. Chem.*, 42 (1999) 3845.
26. Ren, J., Esnouf, R.M., Hopkins, A.L., Warren, J.,
Balzarini, J., Stuart, D.I. and Stammers, D.K., *Biochem-*
istry, 37 (1998) 14394.
27. Hogberg, M., Sahlberg, C., Engelhardt, P., Noreen, R.,
Kangasmetsa, J., Johansson, N. G., Oberg, B., Vrang, L.,
Zhang, H., Sahlberg, B.L., Unge, T., Lovgren, S., Fridborg,
K. and Backbro, K., *J. Med. Chem.*, 43 (2000) 304.
28. Das, K., Clark, A.D. Jr., Lewi, P.J., Heeres, J., Jonge,
M.R.De, Koymans, L.M.H., Vinkers, H.M., Daeyaert, F.,
Ludovici, D.W., Kukla, M.J., Corte, B.De, Kavash, R.W.,
Ho, C.Y., Ye, H., Lichtenstein, M.A., Andries, K., Pauw-
els, R., Béthune, M.-P.De, Boyer, P.L., Clark, P., Hughes,
S.H., Janssen, P.A.J. and Arnold, E., *J. Med. Chem.*, 47
(2004) 2550.
29. Mao, C., Subeck, E.A., Venkatachalam, T.K. and Uckun,
F.M., *Biochem. Pharmacol.*, 60 (2000) 1251.
30. Zheng, W. and Tropsha, A., *J. Chem. Inf. Comput. Sci.*, 40
(2000) 185.
31. Tropsha A., Cho S.J. and Zheng, W., In Parrill, A.L. and
Reddy, M.R., (Eds.), *ACS Symposium Series 719*,
American Chemical Society, Washington, DC, 1999, pp.
198–211.
32. Shen, M., LeTiran, A., Xiao, Y., Golbraikh, A., Kohn, H.
and Tropsha, A., *J. Med. Chem.*, 45 (2002) 2811.
33. Shen, M., Béguin, C., Golbraikh, A., Stables, J.P., Kohn,
H. and Tropsha, A., *J. Med. Chem.*, 47 (2004) 2356.
34. <http://search.nci.nih.gov/search97/cgi/s97.cgi>.
35. http://dtp.nci.nih.gov/docs/aids/aids_data.html.
36. The program Sybyl is available from Tripos Associates, St.
Louis, MO.
37. Randic, M., *J. Am. Chem. Soc.*, 97 (1975) 6609.

- 649 38. Kier, L.B. and Hall, L.H. Molecular Connectivity in
650 Chemistry and Drug Research. Academic Press, New York,
651 1976.
- 652 39. Molconn-Z, version 4.05, Hall Associates Consulting,
653 Quincy, MA.
- 654 40. Sharaf, M.A., Illman, D.L. and Kowalski, B.R. Chemo-
655 metrics. John Wiley & Sons, New York, 1986.
- 656 41. Gilbert, N., Statistics, W.B. Saunders, Co., Philadelphia,
657 PA, 1976.
- 658 42. Golbraikh, A. and Tropsha, A., J. Comput. Aid. Mol.
659 Des., 16 (2002) 357.
- 660 43. Sachs, L., Applied Statistics, A Handbook of Techniques.
661 Springer, Verlag, 1984, pp. 349.
- 662 44. Tropsha, A. and Zheng, W., Curr. Pharm. Design, 7 (2001)
663 599.
- 664 45. Tropsha, A., Gramatica, P. and Gombar, V.K., Quant.
665 Struct. Act. Relat. Comb. Sci., 22 (2003) 69.
- 666 46. Powell, M.J.D., Math. Program., 12 (1977) 241.
- 667 47. Clark, M. and Cramer, R.D., N., J. Comp. Chem., 10
668 (1989) 982.
48. Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R.,
Hart, W.E., Belew, R.K. and Olson, A.J., J. Comp. Chem.,
19 (1998) 1639.
49. <http://www.rcsb.org>.
50. Gasteiger, J. and Marsili, M., Tetrahedron, 36 (1980) 3219.
51. [http://www.scripps.edu/pub/olson-web/doc/autodock/
tools.html](http://www.scripps.edu/pub/olson-web/doc/autodock/tools.html).
52. Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C.,
Ghio, C., Alagona, G., Profeta, S. and Weiner, P., J. Am.
Chem. Soc., 106 (1984) 765.
53. Mehler, E.L. and Solmajer, T., Protein Eng., 4 (1991)
903.
54. Golbraikh, A. and Tropsha, A., J. Mol. Graph. Model., 20
(2002) 269.
55. Cheng, C.C. and Robins, R.K., J. Org. Chem., 21 (1956)
1240.
56. Chan, C., Yin, H., Garforth, J., McKie, J.H., Jaouhari,
R., Speers, P., Douglas, K.T., Rock, P.J., Yardley, V.,
Croft, S.L. and Fairlamb, A.H., J. Med. Chem., 41 (1998)
4910.
- 669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689