Universidad de Alcalá

DEPARTAMENTO DE FARMACOLOGÍA

FACULTAD DE MEDICINA
Campus Universitario
28871 Alcalá de Henares
Teléfono: 91 885 45 14
Fax: 91 885 45 91

Federico Gago Badenas, Catedrático de Universidad del Departamento de Farmacología de la Universidad de Alcalá,

INFORMA:

Que el trabajo titulado "Exploración de los Sitios de Unión de Antitumorales a la Tubulina Mediante Herramientas Computacionales y Simulaciones de Dinámica Molecular" ha sido realizado, en este Departamento y bajo su dirección, por **Dña. Claire Coderch Boué** y, a su juicio, cumple todos los requisitos para proceder a su defensa pública como Tesis Doctoral.

Alcalá de Henares, a 23 de Abril de 2012

Universidad de Alcalá

DEPARTAMENTO DE FARMACOLOGÍA

FACULTAD DE MEDICINA
Campus Universitario
28871 Alcalá de Henares
Teléfono: 91 885 45 14
Fax: 91 885 45 91

Federico Gago Badenas, Catedrático de Universidad del Departamento de Farmacología de la Universidad de Alcalá,

INFORMA:

Que el trabajo titulado "Exploración de los Sitios de Unión de Antitumorales a la Tubulina Mediante Herramientas Computacionales y Simulaciones de Dinámica Molecular" ha sido realizado, en este Departamento y bajo su dirección, por **Dña. Claire Coderch Boué** y, a su juicio, cumple todos los requisitos para proceder a su defensa pública como Tesis Doctoral.

Alcalá de Henares, a 23 de Abril de 2012

Exploración de los Sitios de Unión de Antitumorales a la Tubulina Mediante Herramientas Computacionales y Simulaciones de Dinámica Molecular

Memoria para acceder al grado de doctor dirigida por el
Prof. Federico Gago Badenas
y
presentada por la
Lda. Claire Coderch Boué

enmarcada dentro de programa de doctorado del
Master en Dianas Terapéuticas y Señalización Celular,
Investigación y Desarrollo.

Alcalá de Henares, 2012

UNIVERSIDAD DE ALCALÁ
DEPARTAMENTO DE FARMACOLOGÍA

"Call it a clan, call it a network, call it a tribe, call it a family.
Whatever you call it, whoever you are, you need one."

Jane Howard

Llegados a este punto, echo la vista atrás y me da vértigo. El tiempo pasa tan deprisa que no nos damos cuenta, y parece que fue ayer cuando entré por primera vez en el laboratorio. Pero no fue ayer, fue hace cinco años, cuando yo aún no tenía la carrera acabada y esto de las "bolitas y los palitos" me llamó la atención. Si, cinco años, y si no me he dado cuenta de que han pasado es porque no se me han hecho demasiado largos. Al contrario, cortos, muy cortos. Ahora me doy cuenta de las cosas que durante todo este tiempo he dado por sentadas, que no significa que no haya valorado, pero que no todo el mundo ha tenido la suerte de tener. Por eso, con el corazón en la mano, siendo consciente de que probablemente no sea capaz de expresar todo lo que os quiero decir, y aún a riesgo de que los agradecimientos ocupen más espacio que las conclusiones de esta tesis, os quiero dar las gracias.

Federico, muchas gracias por haberme dado la oportunidad de aprender y trabajar en el laboratorio. Por haberme abierto las puertas el día en el que me planté ahí, en esa especie de cueva con los cables de red y de luz colgando del techo, y te dije que me quería quedar. Gracias por haber sido tan buen maestro, haberme enseñado a pensar y a escribir (aunque se me sigan resistiendo los which, with y witch), por haber creído en mí y no haber bajado los brazos cuando las cosas se complicaban y nos sonaban un poco a chino. Gracias Antonio Jiménez por las charlas en las comidas.

Gracias a los Dr. Fernando Díaz y al Prof. José Manuel Andreu por hacerme descubrir el maravilloso mundo de la tubulina y haber podido colaborar juntos. Gracias a los Drs. Ángeles Canales y Jesús Jiménez-Barbero por sus experimentos de RMN. Gracias Ruth y Benet por el trabajo juntos y todo lo aprendido. Muchas gracias a la empresa PharmaMar, por haber hecho posible esta colaboración, y en especial, gracias a los Drs. Carlos Galmarini y Carmen Cuevas, por su amabilidad.

Thank you very much Prof. Walter Thiel for giving me the opportunity to stay in your group and learn QM/MM. Muchísimas gracias Elsa y Mario, porque

pese a tener mucho trabajo me enseñasteis todo lo que pudisteis e hicisteis que mi estancia en Alemania fuera más agradable. No os olvidaré.

Juan, no sé por dónde empezar. Gracias por todo, porque no sólo has sido un compañero de trabajo, también eres un amigo. Gracias por los buenos días con una sonrisa, por los cafés de la mañana, por las lecciones de pádel y las clases teóricas de fútbol (aunque sigo prefiriendo el rugby). Gracias por el apoyo incondicional, por los ánimos cuando las cosas no iban bien, por aguantarme mis neuras y mi carácter (con eso creo que tienes ganado el cielo), por los abrazos y porque siempre, siempre has estado ahí, y siempre he sabido que podía contar contigo. Gracias porque el ambiente que hemos tenido en el laboratorio es impagable. ¡No sabes cuánto te echo de menos!.

Gracias Ana por esos primeros años en los que me enseñaste todo lo que sabías. Gracias porque incluso estando en NY te sigues acordando de mis "movidas", ¡eres única!. Gracias Marta por el tiempo que estuviste en el laboratorio (aunque fuese poco), por la ilusión que ponías en todo y que nos contagiabas. Gracias Irene Cuesta por la alegría que has traído todos los días y porque me conseguías alegrar en mis momentos de enfado con el mundo. Gracias Alberto porque siempre nos has llenado de optimismo, ojala te hubieses podido quedar más tiempo con nosotros.

Gracias Eva por los abrazos, los ánimos y ese dibujo de mi alter ego, el Yeti, recién levantado (a más de un visitante en el laboratorio se lo he tenido que explicar). Raúl (ilustrísimo Dr. Dr.), gracias por entender a esta "loca sin diagnosticar", por las risas y tu amistad, ¡eres genial!. Gracias David porque pese a ser una "gabacha" siempre has sido un sol conmigo. Kilian gracias por transmitir serenidad y devolverme siempre una sonrisa. Muchísimas gracias Carlos porque lo único que puedo decir de ti es que eres una bellísima persona. Gracias Miguel (Panchi) por un regalo del amigo invisible precioso y tener siempre una sonrisa y una palabra amable. Cris, gracias por ser una niña tan buena (aunque se empeñen en decir que eres malévola, eres un solete), gracias por las

risas y los pucheros. Gracias Pablo por perdonarme los desajustes en las Campanas de Gauss durante la carrera y e introducir los potatos en nuestro día a día. Muchas gracias Ana y Nadia, porque ha sido imposible no reírme con vosotras. Gracias Pedro por animar los días que estuviste con nosotros (admítelo, te vas a pasar al lado oscuro, lo veo venir). Gracias Borja por tus visitas al laboratorio y por esas prácticas del master en las que yo era capaz de pifiarla hasta en el experimento más sencillo. Mil gracias Irene Sánchez, Irene Mazo, Julia, Lauralicia, Mali, Ariel y a todos los que habéis hecho que los días fuesen más llevaderos. Gracias por las comidas y los cafés, por los picnics y los partidos de pádel, por las tardes de cañas y las cenas. Gracias a todos por vuestro apoyo, vuestra comprensión, vuestros ánimos y vuestra amistad.

Gracias Antonio Morreale por tu continua disposición a ayudarme desde el primer día. Gracias a Almudena, Javi, Rubén (¡flipáo!), David (¡en serio que me costó creer que eras biólogo!), Alfonso, Gonzalo, Helena, Álvaro, Alberto, Hugo y Raúl. Gracias porque trabajar con todos vosotros ha sido estupendo, de todos he aprendido algo diferente, gracias por vuestra continua ayuda y vuestra inconmensurable paciencia. Gracias por las cañas, por el "chino de la marmita del infierno" y el "mexicano asesino", y por todas esas cenas que hemos compartido con vosotros.

Muchísimas gracias Angélica por ayudarme en todos mis trámites y dudas existenciales con el papeleo y la burocracia, eres la eficiencia y la amabilidad personificada.

Gracias a Alberto de la Cruz, por haberme hecho descubrir lo bonita que puede llagar a ser la biología y haberme impulsado a estudiar farmacia. Gracias, porque eres uno de los mejores profesores que he tenido.

Gracias Marco, porque pese a no conocerme te has portado genial conmigo y me has salvado la vida (o los datos del disco duro, que para el caso es lo mismo). ¡Te voy a estar debiendo sushi el resto de mis días!

Mil gracias Javi, por demostrarme que era capaz de hacer cosas que jamás hubiese pensado que podría, me has enseñado a confiar en mi. Muchas gracias por convertirme en una Extramotivada más y darme la oportunidad de conocer a gente tan genial.

Gracias a mis compis del roco, por ayudarme a subirme por las paredes (en sentido literal) en este último año de tesis y descubrirme que escalar es una de las mejores terapias que existen contra el estrés (además de una buena forma de conocer gente estupenda).

Mil gracias a mi "segunda familia", a los amigos incondicionales que me habéis apoyado desde siempre. A Marta, Sara, Noe, Susana, Mercedes, Ángel, Abel, Oscar, Carlos, Ana, Antonio, María y Luis. Gracias por los años de facultad, por las comidas en la cafetería, por los cafés, por las tardes en el césped, por las tardes en la biblio, por los cumpleaños, por las innumerables cenas, por las tardes de cine, por el viaje de fin de curso, por los viajes a La Mata, por la casa rural, por haberme contagiado del gusanillo de la escalada, por las charlas trascendentales, por las clases del master, por las tardes en Madrid o en Guadalajara, por las salidas al campo, por la subida al Ocejón, por los teatros, la ópera, por acompañarme a esos "recitales" de piano, por las fiestas, las risas, los hombros sobre los que llorar, los ánimos y los abrazos. Gracias por entender a esta loca "frikicéutica". Mil gracias por vuestra amistad, porque me recordáis que hay gente buena.

Gracias a ti Georgina, porque eres un ejemplo constante de esfuerzo y porque desde hace 15 años eres la hermana que nunca he tenido.

Gracias a Papy, Mamy, Papá y Mamá, mi familia, la de verdad, la que realmente cuenta. Gracias porque sois el motor que me ha impulsado durante todo este tiempo. Gracias por enseñarme que todo se logra con esfuerzo y que al final todo llega. Gracias por animarme y apoyarme, porque me hacéis ser mejor persona y porque sin vosotros, esta tesis nunca habría visto la luz. Os quiero mucho.

A mis padres

"Now this is not the end. It is not even the beginning of the end,
but it is, perhaps, the end of the beginning."

Sir Winston Churchill

ÍNDICE

RESUMEN/SUMMARY

Cancer, in its many forms, is one of the main diseases of our time and, according to the World Health Organization, one of the major causes of premature death in industrialized countries. The pharmaceutical armamentarium to treat malignancies includes agents that interfere with DNA biosynthesis and others that alter the cellular machinery necessary for cell division. Among the therapeutic families that compose this last group one of the most extended in clinical use is the family of microtubule interfering drugs, initially represented by the Vinca alkaloids and then extended with the incorporation of paclitaxel, which was isolated from the bark of the Pacific yew tree. Nonetheless more efforts are needed aimed at obtaining new chemical entities capable of binding to microtubules with high affinity and simultaneously overcoming the described resistances to treatment. Computational methods have become increasingly useful in a number of areas such as characterization of ligand-binding sites, ligand docking and molecular dynamics simulations as they provide information that is usually beyond experimental possibilities.

The purpose of this *D. Phil.* thesis was to study, by means of molecular modelling techniques and computational methods, the binding of microtubule stabilizing and destabilizing drugs to their different target sites in tubulin and gain additional insight into the structure-activity relationships. This research was carried out in collaboration with the groups of Profs. Fernando Díaz and José Manuel Andreu at the *Centro de Investigaciones Biológicas (CSIC*), Prof. Weishuo Fang's group at the *Institute of Materia Medica of the Chinese Academy of Medical Sciences,* Dr. Antonio Morreale's group at the *Centro de Biología Molecular Severo Ochoa (CSIC)*, and *PharmaMar* researchers.

"A fool thinks himself to be wise,

but a wise man knows himself to be a fool"

William Shakespeare

1

INTRODUCCIÓN

1.1 Modelado Molecular

Cuando se habla de modelado molecular se hace referencia a la emulación del comportamiento de moléculas y sistemas biológicos. Hoy por hoy, el modelado molecular está asociado con el modelado computacional, ya que se va teniendo conocimiento de sistemas cada vez más complejos, por lo que el modelado manual con lápiz, papel y calculadora está pasando a un segundo plano. De esta forma, las técnicas computacionales nos están permitiendo enfrentarnos a sistemas cada vez más grandes sin dejar de lado los más simples, que se pueden tratar de forma más precisa.

Existe bastante confusión entre los significados de "química teórica", "química computacional" y "modelado molecular". La química teórica se suele considerar como sinónimo de mecánica cuántica, pero la química computacional engloba no solamente la mecánica quántica sino también la mecánica molecular, minimizaciones de energía, simulaciones, análisis conformacional y otros métodos computacionales utilizados para comprender y predecir el comportamiento de los sistemas moleculares. Puesto que todos estos métodos son empleados por los expertos en este ámbito, podemos definir como modelado molecular cualquier técnica computacional o teórica que permita entender el comportamiento de sistemas moleculares. La única distinción que se puede establecer es que el modelado molecular se nuclea alrededor de la representación

y manipulación de las estructuras de las moléculas, y de las características derivadas de su estructura tridimensional. El papel tan importante que han adquirido los gráficos por ordenador en el modelado molecular ha llevado a una parte de la comunidad científica a considerar que el modelado molecular es poco más o menos que "bonitas figuras". Afortunadamente, estas técnicas están fuertemente establecidas, son extensamente utilizadas y se aceptan hoy día como una disciplina por derecho propio. Aunque en la actualidad existen numerosos métodos experimentales para la determinación de la estructura molecular, la rápida evolución de los ordenadores y el notable desarrollo de la química computacional han propiciado la incorporación de modelos teóricos para el tratamiento de todos aquellos problemas relacionados con la geometría y la energía de las moléculas. En general, el modelado molecular ha supuesto una notable contribución en el campo de la investigación experimental, tanto para la interpretación de los resultados obtenidos y la planificación de futuros trabajos, como para deducir información no asequible experimentalmente (*1*).

1.2 Proteínas

Las proteínas, en su conjunto, son las biomoléculas más importantes desde un punto de vista funcional. Sus propiedades son tan útiles que se emplean como agentes terapéuticos y catalizadores, así como fuente de diversos materiales. Muchas enfermedades tienen su origen en mutaciones que se producen en las proteínas, provocando la pérdida de su función al bloquear su actividad enzimática o al alterar su estructura tridimensional. En general, las enzimas y los receptores de membrana constituyen las dianas más comunes de los medicamentos, ya sea para restaurar o modificar la función celular o para destruir agentes infecciosos y células cancerígenas.

Entre los objetivos del estudio de las proteínas se incluyen el poder predecir la estructura y la función de una molécula concreta a partir de su

secuencia de aminoácidos, así como el conocer los detalles de su interacción con pequeños ligandos u otras macromoléculas. Este conocimiento ya está permitiendo, en casos favorables, diseñar y sintetizar catalizadores y materiales de uso industrial, entender las bases moleculares de muchas enfermedades, así como el diseño de fármacos a medida para combatir ciertas enfermedades. En los últimos años se han conseguido avances significativos en este sentido ya que se ha llegado a alterar la actividad y la estabilidad de algunas proteínas por parte de grupos experimentales mientras que grupos teóricos han sido capaces de simular aspectos de su plegamiento y el mecanismo catalítico de algunas enzimas.

1.2.1 Estructura de las Proteínas

Las proteínas son polímeros formados por la concatenación de aminoácidos mediante enlaces peptídicos. Los aminoácidos que las forman son un total de 21 (si se cuenta la selenocisteína) y son los responsables de la gran versatilidad y variabilidad de las proteínas. Son *zwitteriones* formados por un grupo amino (básico) y uno carboxílico (ácido), que serán los encargados de unir los aminoácidos entre sí, y una cadena lateral, que pueden ser de carácter ácido, básico o apolar, unida al carbono Cα de tal forma que le confiere la isomería L (Figura 1).

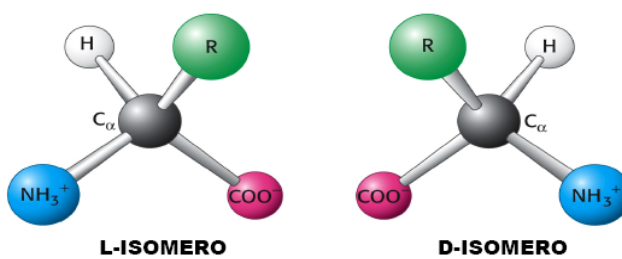


Figura 1: comparación de los dos isómeros posibles de los aminoácidos. El isómero L es el presente en las proteínas naturales mientras que el D únicamente se encuentra en algunos aminoácidos bacterianos.

Las proteínas no son polímeros de forma regular sino que presentan motivos estructurales que aparecen con frecuencia: la hélice α y la hebra β. Esta última puede dar lugar a láminas β paralelas y antiparalelas (Figura 2).

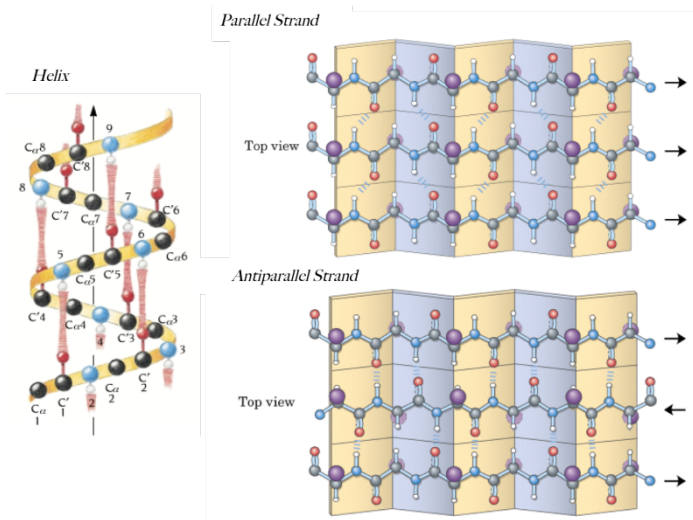


Figura 2: representación esquemática de la hélice α y los dos tipos de lámina β.

Estos motivos constituyen la estructura secundaria de la proteína, siendo la estructura primaria el orden de los aminoácidos en la secuencia. La terciaria se refiere al plegamiento de los elementos anteriores en una forma más o menos compacta, y la cuaternaria, a la asociación de subunidades proteicas en un ensamblado molecular de mayor tamaño y complejidad. Los elementos de estructura secundaria se conectan entre sí por regiones conocidas como asas o giros (*loops*), que constituyen conformaciones menos regulares (Figura 3).

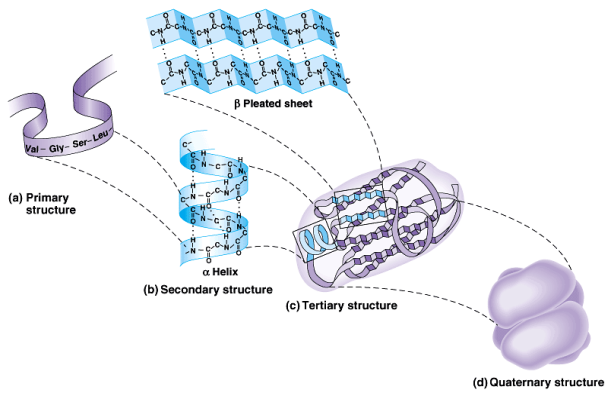


Figura 3: esquema de las diferentes estructuras proteicas.

Dadas sus diferentes características, los aminoácidos desempeñarán diferentes funciones dependiendo de su posición en la proteína plegada. Los que se encuentran en el interior son mayoritariamente apolares y son esenciales para el mantenimiento de la estructura mediante interacciones de van der Waals, mientras que los que se encuentra en la superficie son en su mayoría polares y mediante la formación de puentes de hidrógeno, participan en interacciones proteína-proteína, proteína-ligando o en reacciones enzimáticas.

Cada tipo de estructura secundaria puede ser completamente descrito por los ángulos de torsión de cada residuo. Dado que el enlace peptídico (ω) es plano y raramente se desvía de 0 ó 180^{o}, serán los ángulos φ y ψ los que determinarán la estructura secundaria, como puede verse en el mapa de Ramachandran (Figura 4).

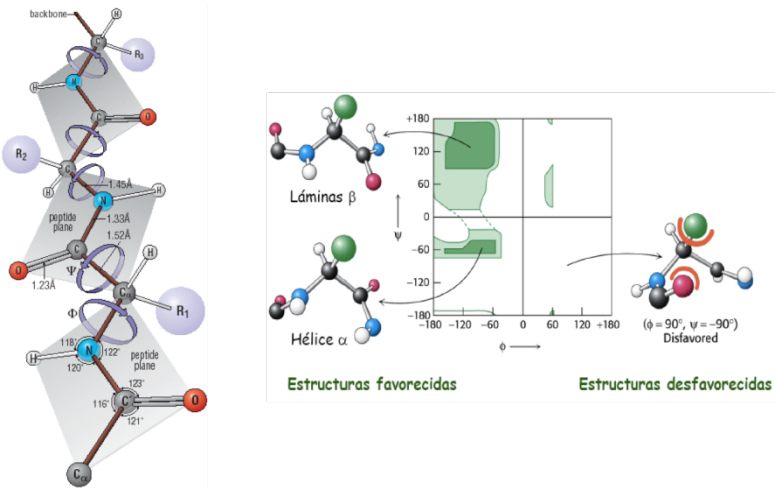


Figura 4: representación de los enlaces rotables del esqueleto peptídico, φ y ψ, y su disposición en el mapa de Ramachandran.

A medida que se fueron conociendo más estructuras de proteínas, se observó que muchas contenían más de una región diferente, a menudo con funciones separadas. A cada una de estas regiones se la conoce como dominio, que se define como una cadena polipeptídica que puede plegarse independientemente en una estructura tridimensional estable. Un principio básico de la bioquímica es que toda función fisiológica tiene una base estructural, por lo que sin estructura no hay función. Cada proteína ha evolucionado para llevar a

cabo una función específica que está asociada a su estructura. La amplia variedad de funciones biológicas requiere, por tanto, de una diversidad estructural igualmente amplia, y las mutaciones aleatorias y la selección natural han sido los grandes impulsores de la evolución que han sufrido las proteínas en el transcurso del tiempo (*2*).

1.2.2 Fuerzas de Interacción

De la misma forma que las interacciones establecidas por los diferentes grupos de aminoácidos mantienen la estructura de las proteínas, estas mismas fuerzas regirán la interacción de la proteína como macromolécula con otros componentes celulares.

En los complejos proteicos localizados en las membranas celulares predominan las interacciones hidrófobas de la superficie de la proteína en contacto con la membrana lipídica. Pero en la asociación de proteínas multiméricas el mantenimiento de la estructura depende mayoritariamente de la formación de enlaces de hidrógeno e interacciones de van der Waals. Estas interacciones, aunque débiles de modo individual, son garantes de una gran estabilidad en complejos moleculares cuando se dan de forma simultánea y en suficiente cantidad. Por lo tanto, las características superficiales de las proteínas van a ser decisivas en la formación de complejos transitorios funcionalmente activos. De hecho, la distribución superficial de carga de las proteínas desempeña un papel tan importante como la carga global en el proceso de asociación y reconocimiento entre dos proteínas. Así, las interacciones electrostáticas monopolo–monopolo de larga distancia, facilitan el acercamiento de las proteínas: cuando una proteína posee una carga global positiva y la otra una carga global negativa se produce el movimiento de atracción entre ambas. La interacción de dos proteínas se realiza como el encaje de dos piezas de un puzzle, ya que los residuos de igual naturaleza se concentran en determinadas regiones, haciendo que las proteínas presenten una distribución irregular con áreas de marcado carácter hidrófobo y otras de

naturaleza polar, que encontrarán su "negativo" en la proteína vecina. De esta forma, el reconocimiento molecular a corta distancia está controlado, fundamentalmente, por las características de la superficie de las proteínas, que a su vez vienen determinadas por la diferente naturaleza de las cadenas laterales de los aminoácidos. Las interacciones que se establecerán entre ellas son mediante la yuxtaposición de parches hidrófobos y complementariedad electrostática, que a corta distancia es de tipo iónico o dipolo–dipolo.

Por otro lado, las interacciones transitorias que se establecen entre proteínas, o entre una enzima y su sustrato, son mucho más complejas de abordar que las interacciones que conducen a complejos estables, ya que implican la formación de un complejo inestable y procesos internos de adaptación y acoplamiento mutuo de ambas unidades.

1.2.3 Movimientos de Proteínas

Hasta hace poco se decía que las proteínas con capacidad catalítica - denominadas enzimas y encargadas de acelerar las reacciones metabólicas - reconocían a los ligandos sobre los que debía actuar con enorme especificidad, al igual que una llave encaja en su cerradura. Esta manera de visualizar la relación entre las enzimas y sus sustratos tiene gran valor pedagógico a la hora de explicar cómo una proteína es capaz de reconocer e interaccionar de modo específico con una molécula concreta, pero da a entender que las proteínas son estructuras rígidas e indeformables. Hoy se sabe que las proteínas son moléculas de gran plasticidad, sobre todo en las zonas de reconocimiento con otras proteínas y en sus centros activos, mientras que las zonas internas, responsables de mantener su estructura y conformación tridimensional, presentan cierta rigidez. En efecto, las proteínas son moléculas en las que la flexibilidad estructural constituye una de las claves que explican su funcionalidad. Las proteínas pueden modificar su conformación tras unirse al sustrato o por reacciones enzimáticas específicas.

La Naturaleza aprovecha los cambios conformacionales así inducidos para controlar la actividad de las proteínas reguladoras y convertir el trabajo mecánico en energía físico–química. Así ocurre en las proteínas alostéricas con varios sitios de unión para su ligando, en los que la unión de una primera molécula del ligando provoca cambios estructurales que facilitan la entrada de las siguientes. Un ejemplo de los más representativos y mejor estudiados es el de la hemoglobina, proteína tetramérica en la que la unión de la primera molécula de oxígeno gas (O_2) a una de las subunidades favorece la entrada de las otras tres moléculas.

Los cambios estructurales asociados a la función de las proteínas pueden ser relativamente sutiles, como los que explican el alosterismo de la hemoglobina, pero en ocasiones conllevan movimientos drásticos y pronunciados de, al menos, ciertos dominios de la estructura (*3*). Hoy en día los cambios conformacionales de muchas proteínas se conocen suficientemente bien como para explicar su funcionamiento y modo de acción, pero la diversidad es tan amplia que resulta difícil analizarlos de manera homogénea y sacar conclusiones o "leyes generales" (*4*).

1.3 Interacción Ligando-Receptor

Los fenómenos de reconocimiento molecular son imprescindibles en la mayor parte de los procesos biológicos. Las proteínas llevan a cabo su función biológica interaccionando, de manera temporal o permanente, con una amplia variedad de moléculas, ya sean ligandos de bajo peso molecular, proteínas, ácidos nucleicos o membranas.

1.3.1 Equilibrio Químico

La interacción entre una proteína A y un ligando B para formar un complejo no covalente AB se representa como un equilibrio entre la asociación

para formar AB y la disociación de AB, cada uno regido por una constante de velocidad:

$$A + B \underset{Koff}{\overset{Kon}{\rightleftharpoons}} AB \tag{1.1}$$

La velocidad de la reacción de asociación será $v_{on} = K_{on} \cdot [A][B]$ y la de disociación $v_{off} = K_{off} \cdot [AB]$, que se igualarán en el estado estacionario de la reacción y se cumplirá que $K_{on} \cdot [A][B] = K_{off} \cdot [AB]$. De esta forma la relación de las concentraciones en el equilibrio se puede relacionar con la constante de asociación o con su inversa, la constante de disociación, K_d:

$$K_a = \frac{1}{K_d} \quad ; \quad \frac{K_a}{K_d} = \frac{[AB]}{[A][B]} \tag{1.2}$$

La energía libre de Gibbs del proceso de asociación, ΔG, que representa la energía libre de unión, determina la estabilidad del complejo [AB] y caracteriza la afinidad de cada uno de los componentes por el otro, está relacionada con la constante de afinidad:

$$\Delta G_a = -RT \ln K_a \tag{1.3}$$

En este caso, R es la constante de los gases (R= 1,98 cal/mol K) y T es la temperatura en grados Kelvin.

1.3.2 Componentes de la Energía Libre de Unión

La Energía libre de unión o Energía de Gibbs de un proceso se puede expresar según la segunda ley de la Termodinámica como:

$$\Delta G = \Delta H - T \cdot \Delta S \tag{1.4}$$

donde el cambio de energía del proceso (ΔG) depende del cambio en la entalpía (ΔH), que es un término energético que hace referencia a las interacciones directas ligando-proteína; y $-T \cdot \Delta S$ es el factor entrópico que hace referencia a la libertad de movimientos de ligando y receptor y al efecto hidrofóbico.

En la interacción de un ligando con una proteína, el término entálpico está descrito por dos tipos diferentes de fuerzas: atractivas y repulsivas, las cuales

intervienen en la estabilización de la unión de un ligando a la proteína. Estas fuerzas son consecuencia de interacciones de van der Waals y electrostáticas tanto favorables (atractivas) como desfavorables (repulsivas). Las interacciones por enlaces de hidrógeno son las que entálpicamente contribuyen en mayor medida a la unión frente a las interacciones de van der Waals, aunque ambas están moduladas por la desolvatación. La desolvatación es el proceso por el cual una molécula pierde parcial o totalmente su interacción con el agua del medio para interaccionar con otra molécula. Aunque este proceso suele ir acoplado a cambios entrópicos favorables, constituye un término mayoritariamente entálpico. Esto es debido a que tanto el ligando como la proteína solvatados establecen interacciones más o menos fuertes con las moléculas de agua del disolvente. Desolvatar superficies hidrófobas del ligando y del receptor es entálpicamente menos costoso que desolvatar grupos polares. Si al desolvatar donadores o aceptores de enlaces de hidrógeno del ligando, que establecen interacciones entálpicamente más fuertes con el agua del medio que los grupos apolares, estos grupos no establecen mejores interacciones con la proteína, el coste energético derivado de la desolvatación no se verá compensado (*5*).

El término entrópico es aquel que hace referencia a los movimientos tanto del ligando como de la proteína. La entropía se puede dividir en vibracional, conformacional y de desolvatación, tanto del ligando como de la proteína. La entropía vibracional, que hace referencia a la vibración intrínseca de los enlaces, no parece jugar un papel muy importante en el término entrópico. Por el contrario, los cambios en la desolvatación y los cambios conformacionales experimentados por el ligando y por la proteína son los más importantes para la entropía de unión. La desolvatación es entrópicamente favorable ya que da lugar a la salida de moléculas de agua del bolsillo de unión del ligando, que pasan de estar conformacionalmente restringidas a tener libertad plena de movimientos en el medio acuoso externo, lo que algunos autores denominan el "efecto hidrofóbico" (*6*). Ésta es la fuerza predominante que se encuentra asociada, por ejemplo, a la

energía de unión de grupos hidrófobos. También es entrópicamente favorable la liberación de moléculas de agua que están estableciendo enlaces de hidrógeno con grupos polares, tanto de la proteína como del ligando, y que están conformacionalmente restringidas. Por otra parte, los cambios de entropía conformacional del ligando y de la proteína son desfavorables, ya que la unión del ligando a su macromolécula diana implica la pérdida de grados de libertad tanto para uno como para otro (*5*). Aunque frente a la creencia inicial de que se pueden obviar los cambios conformacionales experimentados por la proteína al unirse un ligando, se están elevando voces que alertan sobre la importancia de los movimientos proteicos y las diferentes conformaciones de las proteínas que alteran la unión a ligandos (*7*).

1.3.3 Compensación Entalpía-Entropía y Diseño de Fármacos

A la hora de optimizar la unión de los ligandos a las proteínas y ganar afinidad, hay que mejorar tanto los valores entálpicos de la unión como los entrópicos. La entalpía de unión es mucho más difícil de optimizar que la entropía, debido a la contribución de las dos fuerzas contrarias previamente descritas (*8*). Y muchas veces, si se consigue una mejora entálpica, a menudo ésta no queda reflejada en una mejora en la afinidad porque la ganancia de entalpía suele ser compensada con una pérdida de entropía. El fenómeno de la compensación entalpía/entropía ha sido discutido en la literatura durante muchos años (*9-13*). Esencialmente siempre se ha dicho que un cambio en la entalpía se compensa por un cambio en la entropía y viceversa. Una mejora en la interacción ligando–receptor implica una ganancia de entalpía, pero para que la afinidad de unión mejore es necesario que esa ganancia no sea totalmente compensada por una pérdida de entropía, por lo que la ganancia de entalpía es necesaria pero no suficiente para que se produzca la ganancia de afinidad.

De esta forma, para conseguir un aumento de la afinidad del ligando por el receptor hay que conseguir reducir la compensación entálpico-entrópica. Frente a

la dificultad que presenta aumentar la entalpía de unión mediante el establecimiento de nuevas interacciones, optimizar la entropía es (en teoría) relativamente "más fácil" ya que depende fundamentalmente del "efecto hidrófobo" (*14, 15*) y de los cambios conformacionales. En esta línea, el diseño de fármacos cuenta con estrategias para restringir la conformación del ligando de forma que la forma libre sea semejante a la forma unida, para favorecer el enterramiento de grupos hidróphobos en bolsillos apolares de la proteína y así favorecer la liberación de moléculas de agua, o para introducir cambios en el ligando que interaccionen con zonas estructuradas de la proteína para así reducir el cambio en la entropía del receptor. Sin embargo, hasta la fecha, las herramientas de diseño de fármacos han fallado generalmente en predecir correctamente el componente entrópico de la unión de un ligando a su diana.

1.4 La Tubulina y los Microtúbulos como Dianas de Agentes Antitumorales

La superfamilia de la tubulina es una de las familias de GTPasas (proteínas que hidrolizan GTP a GDP) más conservadas a lo largo de la evolución ya que sus miembros juegan un papel crucial en la división celular y mantenimiento del citoesqueleto tanto de organismos procariotas como eucariotas (*16*).

En eucariotas los heterodímeros de αβ-tubulina constituyen los microtúbulos (MT) que intervienen, entre otras funciones, en la separación de los cromosomas durante la división celular (*17*). Por su parte, la γ-tubulina forma anillos que intervienen en el inicio de la nucleación de los MT (*18*). En procariotas los homólogos funcionales de la tubulina son: (i) la proteína monomérica FtsZ, acrónimo derivado de su nombre original "*Filamenting temperature-sensitive mutant Z*", que forma los anillos de extrusión que permiten la división de las bacterias (*19*); (ii) la proteína monomérica TubZ que interviene

en la división del plásmido (*20*), y (iii) la tubulina bacteriana BTubA/B, cuya presencia fue descrita en el género *Prostheobacter*, presenta una mayor homología de secuencia con la tubulina eucariótica que con FtsZ, y de la que se cree que los genes que la codifican se transfirieron horizontalmente, constituyendo un paso intermedio entre tubulina eucariótica y FtsZ (*21*).

En este apartado nos centraremos en la estructura de la tubulina eucariótica y su papel en la formación y comportamiento de los MT que resultan de su polimerización. Por último, dado su papel crucial en la división celular, trataremos de su relevancia como diana farmacológica en terapias antitumorales y describiremos las diferentes familias de compuestos que se unen a ella y los diferentes sitios de unión.

1.4.1 Estructura de la Tubulina

La tubulina existe como un heterodímero compuesto por dos proteínas globulares homólogas, la α-tubulina y la β-tubulina (*22, 23*). Pese a que presentan una homología de secuencia no muy elevada (cercana al 40%) (*24*), su alta homología estructural hace que sean perfectamente superponibles (*20*). En eucariotas hay varios genes para la α y la β tubulina que codifican proteínas con diferentes secuencias de aminoácidos. A estas proteínas se les llama isotipos de tubulina y su distribución normalmente varía en función de los tejidos (*25*).

La estructura terciaria de ambos monómeros está formada por tres dominios bien diferenciados (Figura 5):

- Dominio N-terminal de unión a nucleótido, que abarca la mayor parte de la estructura y está formada por la alternancia de seis láminas β y seis hélices α.
- Segundo dominio o dominio central, formado por tres hélices α y cuatro láminas β, separado del dominio N-terminal por la hélice central H7. En este dominio, situado entre la S7 y la H9 se encuentra el *M-loop*, un asa que se considera principalmente responsables de la capacidad de la

tubulina de eucariotas de establecer las interacciones laterales necesarias para formar los MT (*21*).

- Dominio C-terminal, formado por dos hélices largas que se cruzan sobre la superficie de la proteína, y que una vez formados los MT quedarán en la cara citoplasmática, por lo que son responsables de la unión a proteínas reguladoras (*16, 26*).

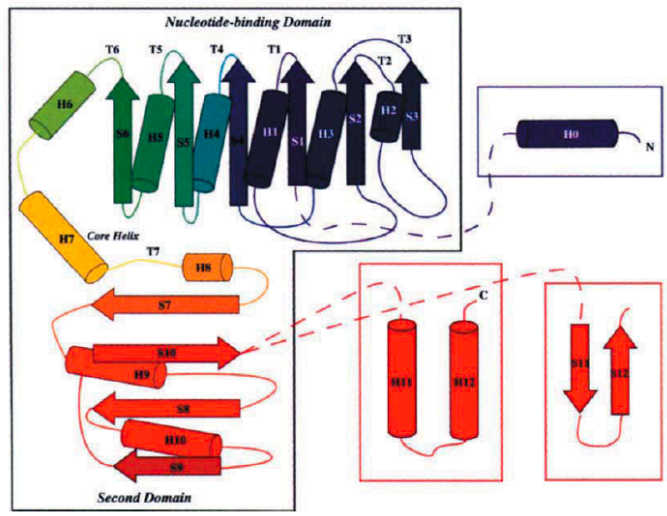


Figura 5: esquema en 2D de las estructuras secundarias y los dominios de la tubulina (imagen tomada de Nogales *et. al*) (*16*).

Ambos monómeros son capaces de unir un nucleótido de guanina (GTP), aunque difieren en la capacidad de hidrolizarlo a GDP. Esto es debido a que el sitio de unión del nucleótido en ambas subunidades se localiza en la interfaz entre monómeros, de forma que para un heterodímero se completa con uno de los monómeros del heterodímero vecino (*27*). En consecuencia, las diferencias en la secuencia de aminoácidos α y β van a provocar que haya dos tipos de dominio de unión a GTP: el dominio intercambiable (*E-site*) y el no intercambiable (*N-site*).

El *N-site* es el que se encuentra en la interfaz entre los monómeros del mismo heterodímero (α_1-β_1) y la molécula de GTP (asociada a Mg^{2+}) ahí situada no se hidrolizará ni se intercambiará por otra. Por otro lado, el *E-site* se encuentra situado en la interfaz de polimerización entre monómeros de heterodímeros

diferentes (α_2-β_1). La molécula de GTP ahí situada sí puede ser hidrolizada a GDP e intercambiada por otra molécula de GTP del medio (Figura 6) (*28*).

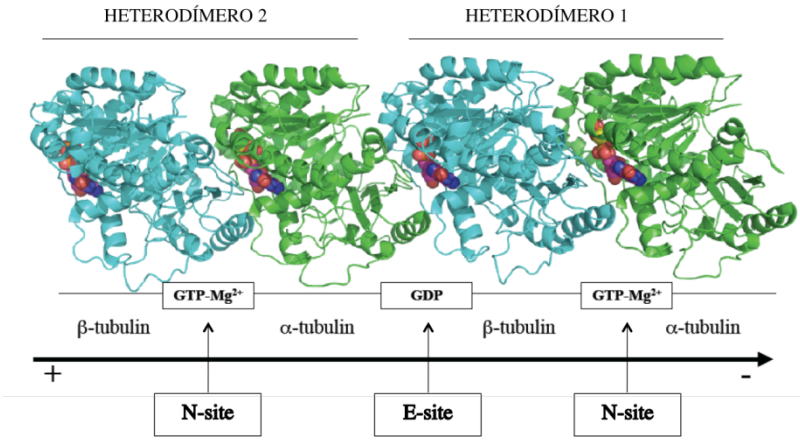


Figura 6: asociación longitudinal de dos heterodímeros de tubulina obtenida mediante representación de las estructuras de la celdilla unidad de la estructura 1JFF depositada en el Protein Data Bank (PDB).

La diferencia en la capacidad de hidrólisis del nucleótido entre los dos monómeros nace del cambio de Glu254 en la subunidad α a Lys254 en la subunidad β (*27, 29*), lo que va a provocar diferentes interacciones longitudinales (Figura 7). Entre los monómeros de un mismo heterodímero, la presencia de la Lys254 en la subunidad β va a reforzar la interacción en la interfaz entre ambos, ya que establecerá fuertes interacciones electrostáticas con los fosfatos del GTP. Por el contrario, la presencia del Glu254 en la subunidad α hará que la interacción en la interfaz longitudinal de polimerización no sea tan estable, ya que en primer lugar habrá repulsión entre las cargas de los fosfatos del GTP y la cadena lateral del Glu; y en segundo lugar, el ambiente tan ácido hará que se produzca la hidrólisis del fosfato γ del GTP por activación de una molécula de agua que solvata la región. Esto se puede ver con mucha claridad en cristales de FtsZ incubados con GTP (*30*), ya que con éstos se pueden obtener valores de resolución mucho mejores que en el caso de tubulina.

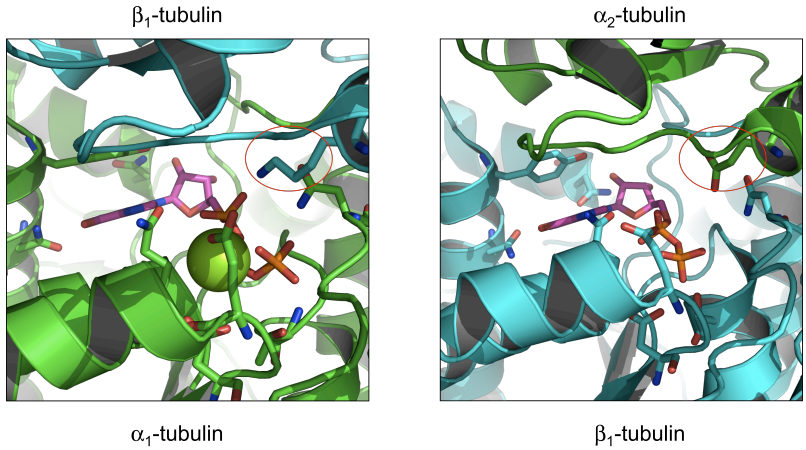


Figura 7: comparación de las dos interfaces longitudinales de los protofilamentos. Señalados con una elipse roja están la Lys254 de la subunidad α y su equivalente Glu254 de la subunidad β. Imagen obtenida mediante representación de las estructuras de la celdilla unidad de la estructura 1JFF del PDB.

En ambos dominios de unión del nucleótido se localiza el "*tubulin signature motif*", una secuencia característica (GGTG) de la familia de la tubulina, y estrictamente conservada, que se localiza en la hélice H4 y con la que interacciona el fosfato γ del GTP. Se ha demostrado que en esta secuencia la Thr es esencial para la afinidad del GTP por la tubulina y para la capacidad de hidrólisis, ya que su mutación a Gly reduce drásticamente ambas (*31*). Una vez más, la información estructural obtenida de los cristales de FtsZ, esta vez en complejo con un análogo no hidrolizable de GTP, nos permite comprender desde el punto de vista estructural la función de esta secuencia de aminoácidos (Figura 8). La cadena lateral de la Thr estabiliza el GTP mediante un puente de hidrógeno entre su OH y uno de los oxígenos del fosfato γ (*32*). Dado que esta secuencia está estrictamente conservada en la familia, es de suponer que esta función se repite en todos los miembros.

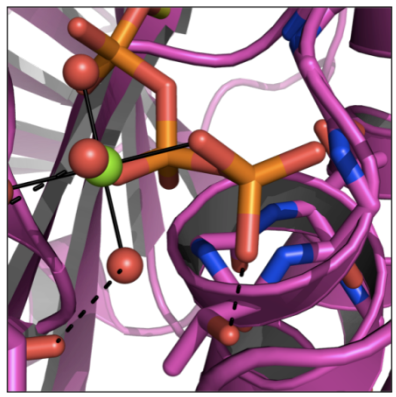


Figura 8: detalle de la unión del grupo trifosfato del 8-morpholino-GTP a FtsZ de *Aquifex aeolicus* en el cristal depositado en el PDB bajo el código 2R75. Los enlaces de coordinación entre el magnesio (esfera verde), el grupo trifosfato y las moléculas de agua (esferas rojas) están representados mediante líneas continuas, mientras que los enlaces de hidrógeno están dibujados como líneas discontinuas.

La hidrólisis del GTP a GDP en el *E-site* induce un pequeño cambio conformacional por el cual la tubulina pasa de estar recta a adquirir una curvatura de 5° (*33-36*). En este cambio conformacional, aunque pequeño, radica la dinámica de polimerización y despolimerización de los MT que se detallará a continuación (*37, 38*).

1.4.2 Estructura de los Microtúbulos

Los MT son estructuras cilíndricas que resultan de la polimerización de los heterodímeros de αβ-tubulina y de los que depende tanto el tráfico intracelular, como la separación de los cromosomas durante la mitosis y la meiosis.

Los MT tienen una polaridad bien definida, ya que sus dos extremos no son equivalentes. Hay un extremo, el llamado + formado por las subunidades β de los heterodímeros, que se alarga y acorta más rápidamente que el otro, que se denomina –, y que está formado por las subunidades α de los heterodímeros de tubulina (*39*).

El número de unidades de αβ-tubulina por vuelta varía entre 10 y 16 in vitro (*40, 41*) y, debido a las interacciones laterales entre los heterodímeros y a la curvatura intrínseca producida por la formación del túbulo, se forman dos tipos de poros a lo largo de su estructura. El primero, el poro tipo I, se encuentra en la interfaz de unión de dos heterodímeros diferentes, y está rodeado por β_1-β_1' por abajo y α_2-α_2' por arriba; mientras que el segundo, el poro tipo II, se encuentra en la interfaz de unión de los monómeros de un mismo heterodímero, y está rodeado por α_2-α_2' por abajo y β_2-β_2' por arriba (Figura 9) (*41, 42*). Los poros constituyen el conocido como sitio externo de unión de algunos fármacos a los MT (*43-47*) y la puerta de entrada de éstos para acceder al interior de los túbulos y por tanto al sitio luminal.

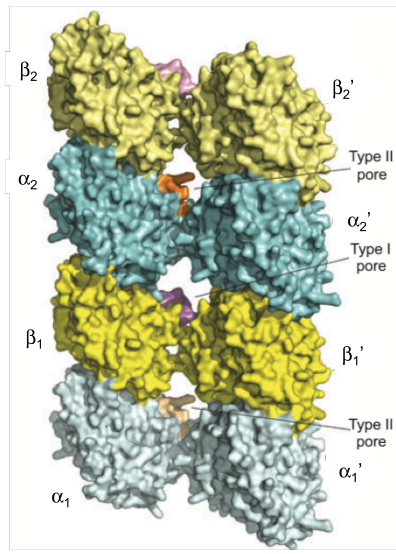


Figura 9: representación de una sección de MT y los poros generados por las interacciones laterales (imagen modificada de Magnani *et al.*) (*44*).

1.4.3 Dinámica de los Microtúbulos

La naturaleza de los MT es intrínsecamente dinámica ya que están continuamente alargándose y acortándose debido a la hidrólisis del GTP en el *E-site*. El comportamiento dinámico de los MT se puede dividir en dos tipos: el *treadmilling* y la inestabilidad dinámica.

El *treadmilling* se produce por un equilibrio entre la disociación de heterodímeros de tubulina en el extremo –, y la asociación en el extremo + de nuevos heterodímeros de tubulina, produciéndose un flujo de éstos entre ambos extremos, por lo que el tamaño de los MT permanecerá estacionario (*39*).

La inestabilidad dinámica es el máximo responsable de los cambios entre fases de crecimiento lento y fases de acortamiento rápido de los MT. Este proceso está caracterizado por cuatro variables bien diferenciadas: (i) el ratio de crecimiento del MT; (ii) la frecuencia de la transición entre la fase de crecimiento o una fase estacionaria a la de acortamiento (también conocido como catástrofe); (iii) el ratio de acortamiento; y (iv) la frecuencia de transición de la fase de catástrofe a la de crecimiento o a una estacionaria (también conocido como rescate) (Figura 10) (*39*).

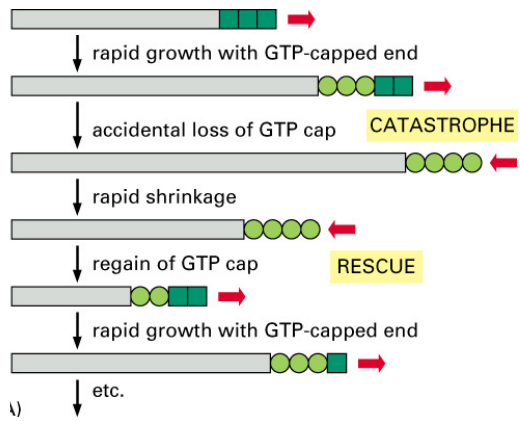


Figura 10: esquema de los cambios sufridos por los MT en el proceso de inestabilidad dinámica (imagen tomada de *Molecular Biology of the Cell, 4th Edition*).

El crecimiento de los MT comienza con la interacción "cabeza con cola" de los heterodímeros de tubulina unidos a GTP en el sitio hidrolizable que da lugar a la formación de los protofilamentos. Éstos se asocian lateralmente entre sí en forma de láminas que se irán curvando hasta cerrar el MT (*48*). El crecimiento de los MT no ocurre únicamente por adición de monómeros en el extremo positivo, sino que también se produce por la adición de oligómeros pre-

polimerizados, de forma que los MT presentan láminas abiertas en el extremo en crecimiento del MT (*48, 49*). La polimerización se inicia lentamente con la interacción de los heterodímeros unidos a GTP. Una vez la polimerización ha comenzado, la superficie de interacción lateral así formada atrae más heterodímeros, por lo que el crecimiento es exponencial (*50*).

Una vez establecida la interfaz longitudinal de polimerización α_2-β_1, el sitio hidrolizable está completo y funcional, por lo que se puede dar la hidrólisis del GTP a GDP. La capacidad GTPasa de la tubulina no es muy efectiva, por lo que quedan remanentes de tubulina unida a GTP en los MT. Esta presencia residual constituye un mecanismo de control para evitar su despolimerización completa (*51*). El proceso de despolimerización o catástrofe comienza cuando el extremo positivo deja de crecer y, tras la hidrólisis del GTP, está formado por heterodímeros de tubulina con GDP. Como la tubulina unida a GDP se curva más, provoca que los protofilamentos empiecen a curvarse hacia atrás perdiendo las interacciones laterales.

Pero la hidrólisis del GTP no es la única responsable del desensamblado de los MT. Son muchas las proteínas que se asocian a ellos y que intervienen de alguna manera en su dinámica. Son las llamadas *Microtubule Associated Proteins*, MAPs, entre las cuales las que han cobrado mayor relevancia en los últimos años son la Oncoproteína 18 (Op18), también conocida como Estathmina, y los demás miembros de su familia (*52, 53*). Se ha demostrado que esta proteína está regulada mediante fosforilación en varios residuos por varias vías de señalización de la célula. Concretamente, es sustrato de kinasas dependientes del ciclo celular (Cycle Dependent Kinases, CDK) (*54*). La Estathmina tiene dos funciones: la primera es reconocer la tubulina curvada libre en el citoplasma y secuestrarla (*55*); la segunda, acelerar los procesos de catástrofe al unirse a los protofilamentos curvados de los extremos de los MT (*52, 56*), dando lugar a la formación de anillos dobles de tubulina (*57, 58*). Una vez unidos a la Estathmina, los

heterodímeros de tubulina pueden intercambiar el GDP resultante de la hidrólisis en el *E-site* por una molécula de GTP del medio.

1.4.4 Antitumorales que se unen a la tubulina

Dado su importante papel en la división celular, los MT constituyen una de las dianas frente a las que va dirigido parte del arsenal farmacológico antitumoral (*39, 59*). Los fármacos que se unen a la tubulina soluble y a los MT lo hacen alterando el comportamiento dinámico de éstos y bloqueando las células en la metafase de la división celular. Esta familia de compuestos se puede dividir a su vez en agentes estabilizantes y agentes desestabilizantes de MT (*60*).

Los primeros agentes antimitóticos conocidos, cuya diana era el huso mitótico o la tubulina soluble, fueron obtenidos de extractos de plantas (taxanos, derivados de la colchicina y alcaloides de la Vinca). Pero la necesidad de aumentar el arsenal terapéutico, para intentar solventar los crecientes problemas de resistencia de las células tumorales a los agentes quimioterápicos, hizo que los esfuerzos se centraran en la exploración de los ecosistemas marinos de los que se tenía poca o ninguna información. Ello derivó en el descubrimiento de nuevas moléculas de origen natural, en su mayoría extraídas de invertebrados y microorganismos marinos, con propiedades citotóxicas antimitóticas (*61*).

1.4.4.1 Mecanismos de Resistencia a Agentes Reguladores de la Dinámica de los Microtúbulos

Es de gran importancia clínica la resistencia de algunos tumores al tratamiento con antimitóticos, tanto estabilizantes como desestabilizantes de MT, por lo que comprender los mecanismos moleculares de esta resistencia es crucial para el desarrollo de nuevos fármacos que permitan tratamientos quimioterápicos eficaces.

El primer factor de resistencia *in vitro* es la sobreexpresión de la Glicoproteína P (*P-Glycoprotein*, P-gp) y de otras proteínas transportadoras de la

membrana celular que actúan como bombas de expulsión que sacan del interior de la célula los agentes antitumorales, impidiéndoles así alcanzar su diana (*62*).

Otro factor es la composición isotípica de los MT. Los diferentes isotipos de tubulina α y β no producen un cambio conformacional en la tubulina, pero sí diferencias en la dinámica de los MT que los contienen debido a las pequeñas diferencias en sus secuencias de aminoácidos (*63, 64*). De todos los isotipos conocidos es la tubulina βIII la que produce mayor resistencia a agentes reguladores de la dinámica de MT en algunos tipos de tumores, especialmente a los taxanos y a los alcaloides de la vinca (*65-68*). Además, los diferentes isotipos de tubulina pueden sufrir modificaciones post-traduccionales en el extremo C-terminal que alterarán su interacción con las proteínas asociadas a microtúbulos, MAPs (*68*). Por otra parte, la sobreexpresión de algunas MAPs modifica la respuesta a los fármacos que se unen a la tubulina, ya que estas proteínas son coadyuvantes en la polimerización y en la despolimerización de los MT (*68*).

1.4.4.2 Agentes Estabilizantes de Microtúbulos

Esta familia de compuestos inducen el ensamblado de los MT e impiden su despolimerización una vez se ha hidrolizado el GTP a GDP. Dentro de este grupo, entre otros, se encuentran los taxanos, las epotilonas (EP), la Cicloestreptina, el Discodermólido, la Laulimalida y el Pelurósido A.

La unión de esta familia de compuestos se realiza en diferentes sitios, y para el único que se tiene evidencia experimental sólida es para el llamado sitio de unión de los taxanos, ya que se cuenta con las coordenadas publicadas de varios cristales. El sitio de unión de los taxanos recibe este nombre por ser el primero que se descubrió albergando una molécula de Paclitaxel (PXL) en la cara luminal de la subunidad β (*69*). Pero este sitio no es exclusivo de taxanos ya que, gracias en primer lugar a experimentos de competición y luego a la obtención de cristales, se vio que las epotilonas se unen al mismo sitio (*70*). También se piensa que este bolsillo de unión en la β-tubulina es utilizado por el

Discodermólido (*71-73*) y la Cicloestreptina (*74-77*), aunque de estos ligandos no hay publicados datos cristalográficos de unión a tubulina. Del Discodermólido se conoce su conformación unida a la tubulina gracias a datos de RMN (*78*) y se ha propuesto que compite con el PXL por el sitio de unión de taxanos (*73*). No obstante, hay varios estudios que han demostrado que es más efectivo que PXL en casos de células resistentes a tratamiento con taxanos, cuando esta resistencia no está mediada por la P-gp, y que tiene un efecto sinérgico con este taxano en células sensibles (pero no con la EPA) (*72*). Por otro lado, el caso de la Cicloestreptina es muy interesante ya que en los primeros ensayos realizados mostró polimerización débil de MT e inhibición competitiva de la unión de PXL con el que se proponía que compartía fármacoforo (*77*), aunque luego se demostró que se unía covalentemente al sitio externo de PXL, localizado en el poro de tipo I, de modo que impedía la unión del PXL y, por tanto, su paso al sitio interno (*43*). Este estudio no hizo más que reforzar la hipótesis de la existencia de un sitio externo de unión para estos ligandos, que se encontraría situado en uno de los poros de los MT y sería necesario para luego acceder al interno (*43, 79*), aunque esta propuesta ha sido contestada recientemente (*80*).

Por el contrario, el sitio de unión de Laulimalida y Pelurósido A no está tan bien definido, aunque estudios recientes apuntan a algún bolsillo de la subunidad β (*81, 82*), en contraposición con la creencia inicial de que se localizaba en la α (*83*). Se ha visto que la Laulimalida es efectiva frente a líneas celulares resistentes al tratamiento con taxanos que sobreexpresan la P-gp (*84*). Por su parte, el Pelurósido A presenta unos efectos sobre la tubulina similares a los de la Laulimalida, con la que comparte sitio de unión (*85*).

De todos estos compuestos se hará más hincapié en los dos primeros grupos (taxanos y epotilonas) ya que han constituido el principal objeto de estudio en esta tesis.

TAXANOS

A principios de los años 50, el Instituto Nacional del Cáncer de los EE.UU. de Norteamérica (*National Cancer Institute, NCI*) comenzó un programa de cribado de extractos de plantas contra células tumorales *in vivo* e *in vitro* (*86*). Pero no fue hasta 1964 que Wani *et al.* obtuvieron una pequeña cantidad de un agente muy activo a partir de la corteza del *Taxus brevifolia* al que denominaron Paclitaxel (PXL) (comercializado más tarde como Taxol®) (Figura 25) (*87*). PXL es un diterpeno compuesto por un núcleo central tetracíclico llamado baccatina III, sustituido en los carbonos C2, C7, C10 y C13, y cuyos anillos se denominan A (un ciclohexeno), B (un ciclooctano), C (un ciclohexano) y D (un oxetano).

Durante los años siguientes, diversos grupos estudiaron el mecanismo de acción de este nuevo compuesto. Fuchs *et al.* determinaron que era un agente antimitótico (*88*) y Horwitz *et al.* demostraron en diferentes estudios que el PXL promovía la polimerización irreversible de los MT, interrumpiendo así la mitosis (*89, 90*). Dada la poca cantidad de este compuesto que se podía extraer de la corteza del *Taxus brevifolia* se elaboraron procesos de semisíntesis a partir de las hojas del *Taxus baccata*, surgiendo así el primer análogo semisintético, el Docetaxel (DXL) (comercializado como Taxotere®) (Figura 11) y un nuevo campo de investigación para la síntesis de múltiples análogos de taxanos (*86*).

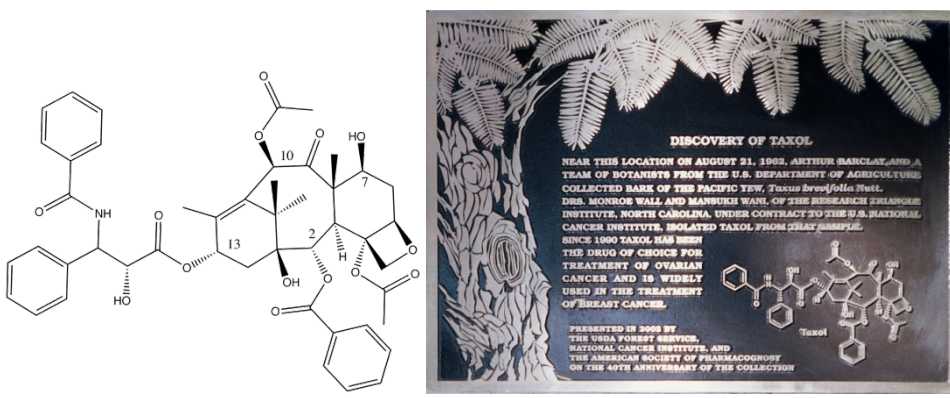


Figura 11: estructura química del PXL y placa conmemorativa de su descubrimiento.

Pero pese a haberse empleado en la clínica desde el primer momento como tratamiento de elección frente a los cánceres de ovario, mama y pulmón, no fue hasta finales de los 90 que Nogales *et al.* fueron capaces de elucidar el modo de unión del PXL a la tubulina. Tras cristalizar láminas de tubulina estabilizadas por cationes de zinc, publicaron una primera estructura tridimensional en 1998 a 3,7 Å de resolución (código PDB: 1TUB) (*69*) que fue posteriormente refinada hasta 3,5 Å y en la que fue sustituido el DXL introducido erróneamente en la anterior por PXL (código PDB: 1JFF) (*91*). Así se pudo comprobar que el PXL (y, por extensión, toda la familia de análogos de taxanos), en una conformación conocida como T-taxol (*46, 92, 93*), se une en la cara luminal de la subunidad β en un bolsillo hidrófobo que se encuentra partido en dos por la cadena lateral de la His229 de la H7 y del que forma parte de uno de sus lados el conocido como *M-loop* (*94*). No obstante, la baja resolución de estos cristales no permite identificar con precisión las interacciones importantes que establece el PXL con las cadenas laterales de los aminoácidos del bolsillo de unión (Figura 12).

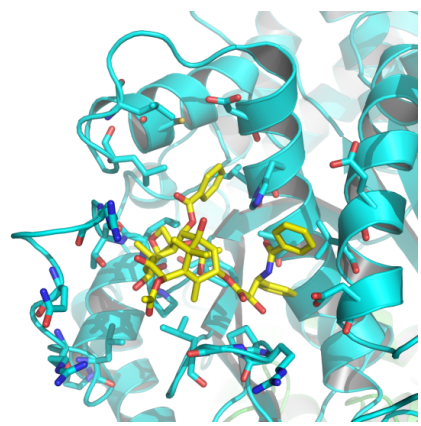


Figura 12: detalle de la unión del PXL a la β-tubulina en la estructura depositada en el PDB bajo el código 1JFF.

Sin embargo, recientes estudios realizados por Díaz *et al.* demostraron que la rápida cinética de unión de los taxanos a los MT no era compatible con la mera difusión de éstos al sitio luminal y que, probablemente, la unión al sitio interno era un evento secundario a un proceso de unión inicial a un sitio externo

(*79*). Esta hipótesis se demostró mediante estudios con Cicloestreptina, un compuesto de origen bacteriano que se une covalentemente a la Thr220 y a la Asn228 de la β-tubulina, impidiendo así la unión de los taxanos a los MT. De estos dos puntos de unión a la tubulina el primero se encuentra en el poro de tipo I de los MT, muy cerca del sitio luminal (*43-46, 95*).

EPOTILONAS

Tras el gran éxito cosechado por el PXL en el tratamiento de diversos tipos de cáncer, muchos fueron los esfuerzos realizados para desarrollar un nuevo agente con un espectro de actividad antitumoral más amplio y que fuera efectivo contra tumores resistentes al tratamiento con PXL.

Fue German Höfle, un microbiólogo del actual Helmholtz Center for Infection Research dedicado al estudio de las Myxobacterias (organismos que crecen en materia vegetal en descomposición), quien descubrió que una cepa de *Sorangium cellulosum* (So ce90), obtenida de una muestra de suelo de Sudáfrica, producía como metabolito secundario un producto antifúngico llamado Epotilona que interfería con el ensamblado de los MT (*96, 97*). Más tarde se descubrieron otros análogos con las mismas propiedades, que en un primer momento se explotaron como antifúngicos en agricultura, pero en los que pronto se descubrió un interesante potencial antitumoral debido al bloqueo de la división en G_2-M en células de mamífero (*98*).

Las epotilonas son estructuras lactámicas macrocíclicas de 15 átomos, con un epóxido en la posición 12-13 y un sustituyente insaturado con un heterociclo en la posición 15 (Figura 13). Estas moléculas también se pueden obtener por síntesis, lo que ha dado lugar a la obtención de múltiples análogos. Sin embargo, hasta la fecha sólo uno, la Ixabepilona (Ixempra®), ha llegado a la práctica clínica como antiutmoral (*99-101*).

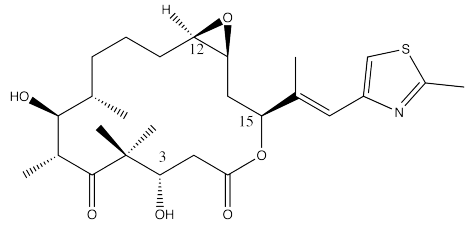


Figura 13: estructura química de la Epotilona A.

Se determinó que su actividad frente a tumores resistentes a taxanos era debida a una menor susceptibilidad a las bombas de destoxificación de la célula, incluyendo la P-gp. Por otra parte, el descubrimiento de que son capaces de inhibir la unión a β-tubulina de PXL marcado radiactivamente sugirió un mismo lugar de unión y un mismo mecanismo de acción (*98*). En 2004 Nettles *et al.* publicaron la estructura de láminas de tubulina estabilizadas por cationes de zinc unidas a Epotilona A (EPA), a 2,89 Å de resolución (código PDB: 1TVK) (*70*). En éstas se ve claramente que la EPA ocupa parte del sitio de unión del PXL, pero pronto se generaron dudas acerca de la bondad del modelo de unión ya que la conformación *endo* muy poco probable propuesta para la EPA situaba el epóxido hacia el interior del macrociclo (Figura 14). Este modo de unión no permite explicar ni la relación estructura-actividad descrita para análogos de EP (*102, 103*) ni los efectos de diferentes mutaciones puntuales presentes en β-tubulina resistente a EP (*104-106*).

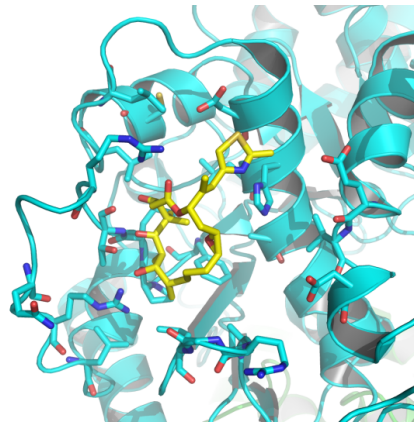


Figura 14: detalle de la unión de la EPA a la β-tubulina en la estructura depositada en el PDB bajo el código 1TVK.

Además, la conformación del *M-loop* obtenida en el cristal era diferente a la obtenida en el cristal de tubulina unida a taxanos. A raíz de esto, diversos grupos han dedicado muchos esfuerzos a determinar la conformación unida de estas moléculas en distintos complejos con tubulina mediante estudios de RMN, tanto en disolución como en estado sólido (*107-109*) y modelado molecular (*110*) pero al día de hoy sigue sin haber un consenso respecto al modo de unión.

1.4.4.3 Agentes Desestabilizantes de Microtúbulos

Estos compuestos impiden el ensamblado de los MT e inducen su despolimerización, algunas veces dando lugar a la formación de anillos de tubulina (*111-115*). Dentro de este grupo se encuentran los derivados de Colchicina, los alcaloides de la Vinca, las Dolastatinas, las Halicondrinas y su derivado semisintético Eribulina (E7389), así como las Espongiostatinas y sus análogos. Como en el grupo anterior, la unión de estos fármacos se hace a diferentes lugares, de los cuales los más conocidos son los denominados "sitio de la Vinca" y "sitio de la Colchicina".

La Colchicina en un producto natural obtenido como metabolito secundario de las plantas del género *Colchicum* (*116*). Aunque sus efectos citotóxicos y su afinidad por la tubulina se conocen desde hace mucho tiempo (*117, 118*), actualmente sólo se emplea en el tratamiento del ataque agudo de gota y de ciertas enfermedades inflamatorias pero no en el tratamiento del cáncer, ya que presenta una elevada toxicidad (*119*). En el año 2004 se publicaron las primeras coordenadas cristalográficas de la tubulina unida a Colchicina y al dominio similar a Estathmina (*Stathmin Like Domain*, SLD) de la proteína RB3 obtenidas mediante difracción de rayos X (*120*). En estos complejos la Colchicina encuentra su sitio de unión en un bolsillo hidrófobo de la β-tubulina localizado en la interfaz de dimerización αβ, justo encima del núcleotido de GTP unido al *N-site*. La presencia de este producto natural y de sus análogos impide los cambios conformacionales necesarios en la subunidad β

para que la tubulina adquiera la conformación recta requerida para la polimerización de los MT (*121*).

De los alcaloides de la Vinca de Madagascar o *Vinca rosea* (*Catharanthus roseus*), el primero en aislarse fue la Vincoleucoblastina o Vinblastina. Su sitio de unión, y por extensión el del resto de alcaloides de la Vinca, se describió en 2005 con la publicación de las coordenadas de la estructura del complejo Tubulina-Vinblastina-Colchicina-SLD-RB3. En el cristal el alcaloide se encuentra situado en la interfaz entre dos heterodímeros de tubulina asociados cabeza-cola, cerca del sitio hidrolizable de unión al nucleótido, por lo que el bolsillo de unión está consituido por aminoácidos tanto de la subunidad α como de la β (*122*). Tres años más tardese publicaron las coordenadas obtenidas por difracción de rayos X de dos complejos de Tubulina-Colchicina-SLD-RB3 unidas a Fomopsina A y a Soblidotina, dos representantes de la familia de las Dolastatinas. Esta familia está compuesta por péptidos modificados de los que la Dolastatina 10, obtenida del molusco gasterópodo *Dolabella auricularia*, fue el primero que mostró inhibición de la polimerización de tubulina (*123*). Pese a tener una estructura química muy diferente a la de los alcaloides de la Vinca, tanto la Fomopsina A como la Soblidotina se unen a la tubulina en la interfaz entre dos heterodímeros de tubulina unidos, justo en el sitio de unión de la Vinca (*124*).

Por otra parte, las Halicondrinas son una familia de macrólidos de origen marino aislados de los extractos de la esponja *Halichondria okadai*, de los que la Halicondrina B es el más potente (*125*). Pese a ser moléculas extremadamente complejas, se han obtenido por síntesis química diferentes análogos simplificados, de los que la Eribulina (E7389) es el más destacado por haber sido recientemente aprobada para el tratamiento del cáncer de mama metastático (*126*). Estos agentes se unen a la β-tubulina cerca del sitio de unión intercambiable de GTP (*127*), sin afectar a la unión de Colchicina (*128*) pero inhibiendo la unión de alcaloides de la Vinca de forma no competitiva (*129*). El

resultado de su unión a la tubulina es la formación de agregados afuncionales (*130*).

Por último, las Espongiostatinas son macrólidos policétidos, aislados de esponjas marinas por varios grupos independientes, que han demostrado un gran poder citotóxico en un cribado realizado por el NCI con células sensibles a agentes antimitóticos (*131*). Pese a que su mecanismo de acción y su sitio de unión no están del todo claros, los resultados apuntan a la inhibición de la polimerización de los MT por unión a un sitio del heterodímero diferente al de los alcaloides de la Vinca (*132*).

De esta familia de compuestos se hará más hincapié en los alcaloides de la Vinca y en un nuevo compuesto de origen marino que está siendo desarrollado por la empresa farmacéutica española PhamaMar, puesto que han sido objeto de estudio en esta tesis.

ALCALOIDES DE LA VINCA

Pese a conocerse las propiedades de los extractos de las hojas de la Vinca (*Catharanthus roseus*) desde el siglo XVII, no fue hasta finales de los años 50 que se estableció su actividad antitumoral (*133*). A partir de este momento comenzaron a emplearse tanto los alcaloides originales Vincristina (VNC) y Vinblastina (Vincaleucoblastina, VLB) (Figura 15) como los análogos semisintéticos Vindesina (VND), Vinorelbina (VNR) y Vinflunina (VFN) en el tratamiento de diferentes tipos de tumores, tanto sólidos como hematológicos (*39, 134*).

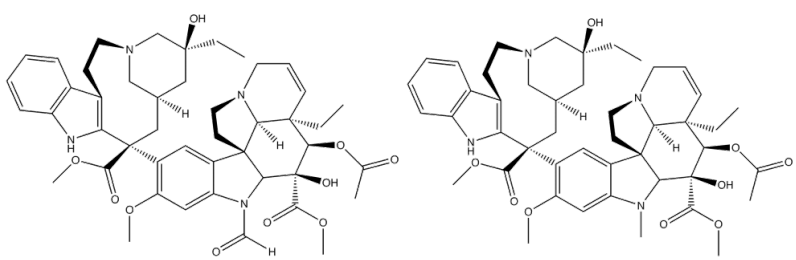


Figura 15: estructura química de la Vincristina (izda) y la Vinblastina (dcha).

Entre los años 80 y 90 se realizaron diversos estudios con el fin de establecer el mecanismo de acción de estas moléculas, que alteran la dinámica microtubular y la polimerización bloqueando el ciclo celular en la metafase. Si las concentraciones empleadas son subestequiométricas se provoca una estabilización de los MT, probablemente porque la unión del alcaloide al extremo unido a GTP impide la hidrólisis del nucleótido (*135-137*). A concentraciones más elevadas provocan la despolimerización y la formación de anillos de tubulina y estructuras paracristalinas (*138, 139*). Estudios termodinámicos de la unión de los alcaloides naturales de la Vinca y de algunos de sus análogos semisintéticos realizados por Correia *et al.* establecieron sus afinidades por la tubulina, que apenas varían en función de los isotipos de tubulina (*140*), y que en orden decreciente son VNC > VLB > VNR > VFN (*140, 141*). Con la cristalización de un complejo de Tubulina-VLB-Colchicina-SLD-RB3 en 2005 y su resolución hasta 4,10 Å, se localizó el sitio de unión de VLB en la intefaz entre dos heterodímeros de tubulina, en un lugar próximo al bolsillo de unión de la base nitrogenada del núcleotido en el sitio hidrolizable (Figura 16).

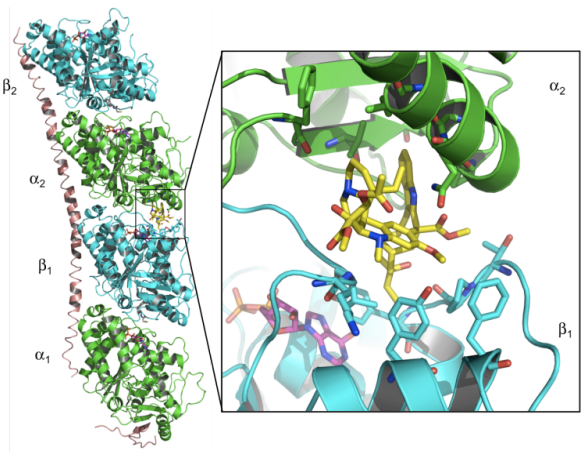


Figura 16: detalle de la unión de la VLB a la interfaz β_1-α_2 entre dos heterodímeros de tubulina en el cristal depositado en el PDB bajo el código 1Z2B. La α-tubulina está coloreada en verde, la β-tubulina en cian, las moléculas de núcleotido en magenta, las moléculas de Colchicina en blanco, la VLB en amarillo y el SLD RB3 en rosa.

La Vinblastina introduce una cuña en esta interfaz que no sólo impide la correcta interacción longitudinal de los dos heterodímeros sino que inhibe la hidrólisis del GTP al mantener alejado del fosfato gamma el Glu254 presente en la subunidad α (*122*). Recientemente se ha determinado que los alcaloides de la vinca y otros ligandos interfaciales reconocen la tubulina curvada unida a Estathmina (*55*).

"To err is human, but to really foul things up you need a computer"

Paul Ehrlich

2

MÉTODOS

2.1 Bases Teóricas del Modelado Molecular

Para el estudio de los sistemas biológicos mediante modelado molecular se tienen que emplear una serie de herramientas computacionales, como potentes computadoras y programas de gráficos moleculares, en conjunción con un buen número de reglas y procedimientos teóricos. Gracias a la rápida evolución de la potencia de los ordenadores, se han podido desarrollar técnicas computacionales cada vez más complejas que nos permiten definir y estudiar los sistemas biológicos a diferentes niveles de teoría. Sin embargo, pese a la potencia de los ordenadores actuales, el estudio de los sistemas de gran tamaño depende de limitaciones tanto de tiempo como de recursos materiales. De esta forma, es necesario emplear aproximaciones para mantener la viabilidad del experimento aplicando distintos niveles de teoría en función del tamaño del sistema que va a ser analizado ya que, a medida que aumenta el nivel teórico aumenta la calidad de los resultados pero también se incrementa notablemente el coste computacional. En este sentido se pueden emplear tres tipos de métodos: cuánticos, clásicos e híbridos, en función de la finalidad del estudio que se quiera realizar.

Los métodos cuánticos tienen en cuenta los núcleos y los electrones en el cálculo, haciendo posible el estudio de la estructura y de las propiedades que dependen de la distribución electrónica, así como el estudio de la reactividad química (formación y ruptura de enlaces). Sin embargo, su aplicabilidad está

restringida a sistemas con centenares de átomos como máximo, siendo inviable para aquellos constituidos por miles de átomos en ausencia de recursos de supercomputación. Los métodos clásicos emplean la aproximación de Born-Oppenheimer en la que, debido a su masa insignificante comparada con la del núcleo, se ignoran los movimientos propios de los electrones y se considera que se ajustan al movimiento de los núcleos atómicos, por lo que la energía de una molécula o conjunto de moléculas se calcula únicamente en función de la disposición de los núcleos atómicos. Este nivel de teoría es el que se escoge generalmente para estudiar macromoléculas biológicas. Los métodos híbridos se han desarrollado recientemente para estudiar procesos químicos que requieren de altos niveles de teoría, como reacciones químicas, dentro del seno de sistemas complejos, como proteínas en el caso de las reacciones enzimáticas. Para ello estos métodos son capaces de describir una pequeña parte del sistema mediante métodos cuánticos y el resto mediante métodos clásicos (*1*).

2.1.1 Métodos Cuánticos (QM)

Estos métodos emplean la mecánica cuántica (*Quantum Mechanics*, QM) y pueden dividirse en dos grupos: los métodos *ab initio* y los métodos semiempíricos. El término *ab initio* significa "desde el principio" o "a partir de primeros principios", y hace referencia a que estos métodos requieren únicamente de las constantes físicas, por lo que no necesitan parámetros ajustados aunque pueden llegar a ser extremadamente costosos desde el punto de vista computacional. Por el contrario, los métodos semiempíricos son aproximados e incorporan parámetros derivados de datos experimentales, por lo que pueden calcular propiedades del sistema de una forma relativamente precisa con un gasto computacional mucho menor.

Debido a estas diferentes características, estos dos métodos se emplean con fines diferentes en función del problema a resolver. Los métodos *ab initio* son capaces de proporcionar información sobre la distribución de la carga del núcleo y

de la corteza electrónica, por lo que se emplean para calcular el potencial electrostático molecular (*Molecular Electrostatic Potential*, MEP) de los ligandos (Figura 17).

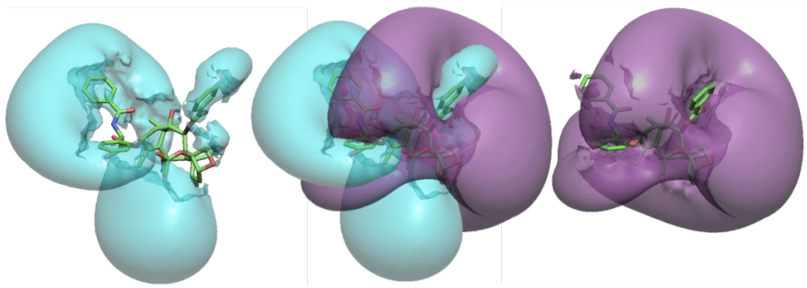


Figura 17: representación del MEP positivo (cian) y MEP negativo (magenta) del Paclitaxel (centro). Al mostrar los potenciales por separado (izquierda y derecha) se aprecia mejor la distribución del potencial en la molécula y se puede intuir el potencial complementario que será necesario en el sitio receptor.

A partir del MEP o del potencial electrostático de superficie (*Electrostatic Surface Potencial*, ESP), estos métodos son capaces de obtener las cargas parciales de cada átomo. Estas cargas atómicas se emplean posteriormente en los métodos clásicos para derivar las interacciones electrostáticas mediante un potencial generalmente coulómbico. Por esta razón, la calidad de las cargas obtenidas *ab initio* es crítica para los métodos clásicos. De hecho, el proceso de obtención de cargas atómicas se refinó en el campo de fuerzas de AMBER (*Assisted Model Building with Energy Refinment*) (*142*) dando lugar a un ajuste mediante restricción del potencial electrostático a la superficie molecular (*Restrained Electrostatic Surface Potential*, RESP) (*143-145*), obteniéndose de esta forma una distribución de la carga más uniforme.

2.1.2 Métodos Clásicos (MM)

Como se ha dicho anteriormente, estos métodos emplean la aproximación de Born-Oppenheimer, que considera la energía potencial de los átomos de manera explícita mientras que el movimiento de los electrones se trata

implícitamente. De esta forma, estos métodos van a considerar a los átomos del sistema como esferas de un radio (generalmente el de van der Waals) y una carga determinados.

2.1.2.1 El Campo de Fuerzas

Los campos de fuerzas representan el conjunto de ecuaciones que describen una serie de contribuciones energéticas enlazantes y no enlazantes a las fuerzas inter- e intramoleculares mediante expresiones empíricas relativamente sencillas que son fáciles de computar. Una de sus características importantes es la de que posibilitan la extrapolación de los parámetros calculados para un reducido número de átomos a un amplio espectro de moléculas.

En los estudios realizados en esta tesis, y siguiendo la tradición del laboratorio, se empleó el campo de fuerzas de AMBER (*142*), que se detalla a continuación.

La energía potencial total de un sistema está regida por dos términos energéticos: aquellos que dependen de las fuerzas generadas por la unión covalente de átomos entre sí (términos enlazantes) y los que dependen de la interacción de átomos que no están unidos mediante enlaces covalentes (términos no enlazantes):

$$E_{tot} = E_{enlazante} + E_{no\,enlazante} \qquad (2.1)$$

Estos parámetros se obtienen a través de datos experimentales y cálculos *ab initio* en sistemas modelo utilizados como referencia.

TÉRMINOS ENLAZANTES

Los términos enlazantes incluyen contribuciones debidas a los enlaces covalentes, ángulos de valencia y ángulos torsionales propios e impropios.

- Término de Enlace (*Bond Stretching*)

La energía potencial para un enlace covalente entre dos átomos se define mediante una función de Morse que incluye tres parámetros: energía de ruptura del enlace, valor de referencia y constante de fuerza del enlace. Sin embargo, el potencial de Morse no se emplea en campos de fuerza de mecánica molecular dado que raras veces los enlaces se desvían significativamente de sus valores de equilibrio. Por lo tanto se consideran expresiones computacionalmente más eficientes, como el potencial de Hooke (Figura 18):

$$E_{enlace} = \left(k_i/2\right)\left(I_i - I_{eq}\right)^2 \tag{2.2}$$

donde I_i corresponde a la distancia de enlace, I_{eq} al valor de equilibrio y k_i a la constante de fuerza del enlace.

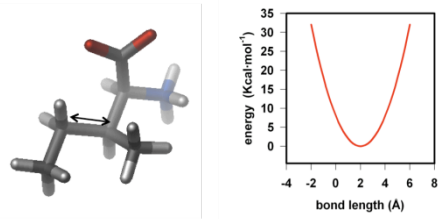


Figura 18: perfil energético de la distorsión, por acortamiento o estiramiento, de un enlace covalente.

- Término de Ángulo (*Angle Bending*)

La desviación de los ángulos de sus valores de referencia también se describe mediante un potencial armónico de Hooke (Figura 19):

$$E_{ángulo} = \left(k_i/2\right)\left(\theta_i - \theta_{eq}\right)^2 \tag{2.3}$$

donde k_i corresponde a la constante de fuerza angular y θ_{eq} constituye el valor de equilibrio para ese ángulo.

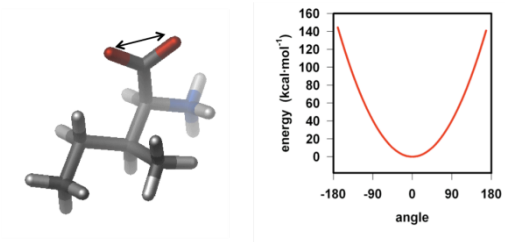


Figura 19: perfil energético de la distorsión de un ángulo de enlace.

- Término de Torsión

El término de torsión describe las barreras energéticas de rotación alrededor de un enlace. A diferencia de los anteriores, no tiene un comportamiento armónico basal, sino que presenta una periodicidad que hace necesario expresar su perfil energético como una serie de Fourier (Figura 20):

$$E_{tor} = \sum_{tor}\sum_{n}\frac{V_n}{2}\left[1+\cos\left(n\phi - n\right)\right] \tag{2.4}$$

donde *Vn* es la altura de la barrera de torsión, *n* la multiplicidad (el número de mínimos en la función cuando se rota 360°), *Φ* el ángulo diedro y *n* el ángulo de fase (indica en qué punto pasa la torsión por el mínimo energético). Este tipo de torsión se denomina torsión propia para diferenciarla de la torsión impropia que se utiliza para mantener la quiralidad sobre carbonos asimétricos y la planaridad de sistemas aromáticos y heteroaromáticos.

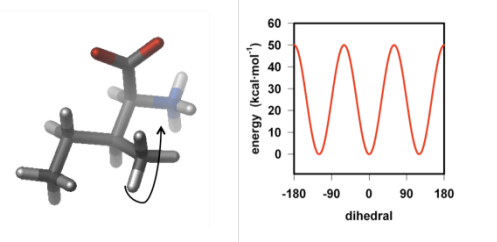


Figura 20: perfil energético de la distorsión de un ángulo diedro entre dos carbonos sp^3.

TÉRMINOS NO ENLAZANTES

Están definidos por un término de atracción-repulsión de tipo Lennard-Jones para las fuerzas de van der Waals y un término coulómbico para las interacciones electrostáticas.

- Término Electrostático

La distribución electrónica de una molécula se puede representar a partir de las cargas puntuales fraccionales centradas en los núcleos. Estas cargas reproducirán la energía electrostática del sistema que se representa habitualmente usando la ley de Coulomb:

$$E_{ele} = \sum_{i} \sum_{j>i} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \tag{2.5}$$

donde q_i y q_j son las cargas puntuales de cada átomo, r_{ij} la distancia entre ellos y ε la constante dieléctrica del medio que las separa. Si las cargas puntuales de dos átomos son contrarias, éstos se atraerán entre sí, pero si las cargas son del mismo signo se repelerán (Figura 21).

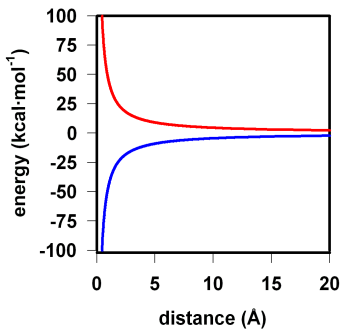


Figura 21: perfil de la interacción electrostática.

- Término de van der Waals

La interacción de van der Waals entre dos átomos se origina a partir de un balance entre fuerzas atractivas y repulsivas. Los núcleos atómicos tienden a repelerse unos a otros cuando están muy cerca y a atraerse cuando están a una distancia óptima. Este comportamiento se puede describir mediante un potencial de Lennard-Jones, cuyas exponenciales son típicamente *6* y *12*:

$$E_{Lennard\,Jones} = \sum_{i} \sum_{j>i} \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \tag{2.6}$$

donde A_{ij} y B_{ij} son constantes para cada par de átomos relacionadas con sus radios de van der Waals y r_{ij} es la distancia entre ellos (Figura 22).

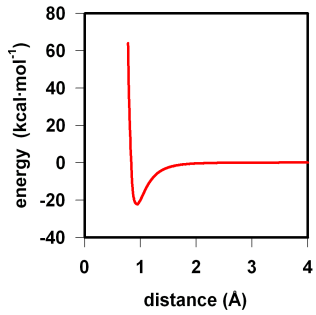


Figura 22: perfil de la interacción Lennard-Jones.

2.1.2.2 Métodos Basados en Campos de Fuerzas

2.1.2.2.1 Mecánica Molecular

La mecánica molecular (*Molecular Mechanics*, MM) se basa en la mecánica clásica o mecánica de Newton para predecir las estructuras y propiedades de las moléculas. Gracias al campo de fuerzas que se emplee, se asignarán a los átomos su carga y su radio de van der Waals correspondiente y se podrán estudiar las interacciones enlazantes y no enlazantes cuyo sumatorio representa la energía del sistema. Hay que tener en cuenta que los sistemas moleculares presentan una superficie de energía potencial muy corrugada, con multitud de máximos y mínimos locales, por lo que muchas veces no es trivial localizar el mínimo global de energía.

Al iniciar el estudio de un sistema siempre es necesario hacerlo partiendo de un mínimo energético que será normalmente relativo, ya que el global es difícil de obtener para sistemas grandes. El proceso que permite identificar las geometrías del sistema que corresponden a los puntos de mínima energía potencial se denomina minimización de la energía, y es la que se emplea en primer lugar para reducir las interacciones desfavorables presentes en el sistema de partida, antes de someterlo a otras técnicas (Figura 23).

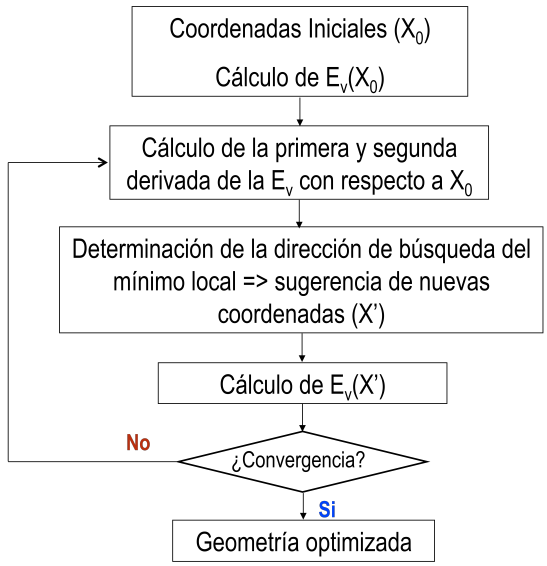


Figura 23: diagrama de flujo de trabajo de una minimización de energía.

La minimización de energía se puede describir como una función (*f*) dependiente de una o varias variables independientes. En nuestro caso, estas variables serán las coordenadas cartesianas de los átomos del sistema, para las que se quieren obtener los valores que proporcionan un mínimo para el valor de *f*. En ese mínimo, la primera derivada de la función con respecto de cada una de las variables es cero y la segunda derivada es positiva:

$$\frac{\partial f}{\partial x_i} = 0 \qquad \frac{\partial^2 f}{\partial^2 x_i} > 0 \tag{2.7}$$

Los dos métodos de minimización más empleados son el de *steepest descent* (descenso más pronunciado) y el de *conjugate gradient* (gradiente conjugado). Ambos modifican las coordenadas de los átomos hasta llegar a un mínimo energético, pero mientras el primero lo hace mediante el descenso de gradiente más rápido, el segundo lo hace conjugando el gradiente con las direcciones previas de búsqueda. De esta forma el primero se emplea cuando el sistema está alejado del mínimo y el segundo cuando está cerca del mínimo. Por

esta razón, en la práctica se emplean los dos combinados: primero el de *steepest descent* durante un número de pasos y luego el de *conjugate gradient*.

2.1.2.2.2 Dinámica Molecular

En la dinámica molecular (*Molecular Dynamics*, MD) se generan las configuraciones sucesivas del sistema integrando las leyes de Newton del movimiento. El resultado es una trayectoria en la que las posiciones y las velocidades de los átomos del sistema varían a lo largo del tiempo de simulación. La trayectoria del sistema se obtiene resolviendo la segunda derivada de la Segunda Ley de Newton, $F = ma$:

$$\frac{d^2xi}{dt^2} = \frac{F_{xi}}{m_i} \tag{2.8}$$

donde se describe el movimiento de un átomo de masa m_i a lo largo de las coordenadas (x_i) con una fuerza determinada (F_i) que viene dada por el campo de fuerzas, y un tiempo t. En el punto inicial de la dinámica, la energía total del sistema es la suma de las energías cinética y potencial para las coordenadas (x_0) y velocidades (v_0) iniciales:

$$E_{tot} = E_{pot}(x_0) + E_k(v_0) \tag{2.9}$$

Éste es un método determinista ya que el resultado final es dependiente de la posición inicial de los átomos del sistema (de ahí la recomendación de partir de varias configuraciones iniciales para comprobar que los resultados convergen, así como la necesidad de realizar una minimización previa para evitar que haya gradientes de energía muy grandes que distorsionen el comportamiento del sistema).

La evolución del sistema a lo largo del tiempo se obtendrá mediante integración numérica. Esta integración se realizará en fracciones de tiempo separadas por el valor de δt. La fuerza total de cada partícula en el tiempo t (que se asume que es constante) se calcula como la suma vectorial de las

interacciones con los demás átomos y se determina mediante la derivada de la energía potencial con respecto a las coordenadas:

$$F = -\partial E_{pot} / \partial x \tag{2.10}$$

A partir de las fuerzas se puede determinar la aceleración de los átomos de modo que, conociendo sus posiciones y velocidades a tiempo *t*, podemos calcular sus posiciones y velocidades en la siguiente fracción de tiempo, $t+\delta t$. Este proceso se repite sucesivamente para generar la trayectoria completa para todos los átomos del sistema.

Los algoritmos empleados en MD para generar las trayectorias asumen que las propiedades de éstas se pueden aproximar mediante series de Taylor:

$$\begin{aligned} x_i(t+\partial t) &= x_i(t) + \partial t v_i(t) + \tfrac{1}{2}\partial t^2 a_i(t) + \ldots \\ v_i(t+\partial t) &= v_i(t) + \partial t a_i(t) + \tfrac{1}{2}\partial t^2 b_i(t) + \ldots \\ a_i(t+\partial t) &= a_i(t) + \partial t b_i(t) + \tfrac{1}{2}\partial t^2 c_i(t) + \ldots \end{aligned} \tag{2.11}$$

donde *v* es la velocidad (primera derivada de las posiciones con respecto al tiempo), *a* es la aceleración (segunda derivada), *c* es la tercera derivada, etc. Estas ecuaciones se integran mediante el algoritmo de Verlet (*146*), que usa las posiciones y aceleraciones a tiempo *t* y las posiciones en la etapa previa, $x(t-\delta t)$, para calcular las nuevas posiciones a $t+\delta t$, $x(t+\delta)$.

El tiempo de integración δt elegido es de vital importancia ya que si es demasiado pequeño la trayectoria apenas cubrirá una parte de todo el espacio conformacional, mientras que si es demasiado grande puede provocar inestabilidades debidas a solapamientos energéticos entre átomos. Los movimientos de mayor frecuencia son aquellos que corresponden a la vibración de los enlaces con hidrógenos. Se sabe que estas vibraciones ocurren en la escala del femtosegundo (fs) y su inclusión en el cálculo no haría más que añadir una carga computacional extra innecesaria ya que no influyen en el movimiento global del sistema. Es por ello que se suele fijar la longitud de los enlaces covalentes a sus valores de equilibrio mediante el algoritmo SHAKE (*147*). De

esta forma se aumenta el tiempo de integración a 2 fs (desde 3l 0,5 o 1 fs habituales) y se acelera la velocidad del cálculo.

Por otra parte, de la misma forma que se aplican aproximaciones para el cálculo de los términos enlazantes, se aplican también para los no enlazantes, ya que son computacionalmente más costosos. En esta aproximación no se tienen en cuenta las interacciones entre dos átomos separados por una distancia mayor a un límite previamente definido (*cutoff*). Esto afecta fundamentalmente al término electrostático ya que esta interacción es de largo alcance, y la omisión de parte de los pares de átomos del sistema en la simulación puede producir artefactos y conducir a propiedades del sistema no realistas. Para evitar esto se emplean métodos como el sumatorio de Ewald (PME) que permite corregir el sistema al sumar las interacciones con infinitos sistemas réplicas del original. El método se basa en dividir la suma de las interacciones en dos sumatorios: el espacio directo y el espacio recíproco (*148*).

Puesto que los sistemas son finitos y tienen bordes hay que tener en cuenta que se pueden presentar anomalías en las simulaciones con las moléculas que se encuentran en el borde del área e simulación. Para evitar que éstas experimenten un comportamiento diferente a las demás se aplican condiciones de límite periódico (*Periodic Boundary Conditions*, PBC) que permiten realizar la simulación como si el sistema no estuviera confinado. Para ello se rodea el sistema simulado en cada una de sus caras por copias exactas de él mismo. Cada partícula en la simulación tiene un duplicado exacto, con las mismas velocidades, en cada celda que la rodea (Figura 24).

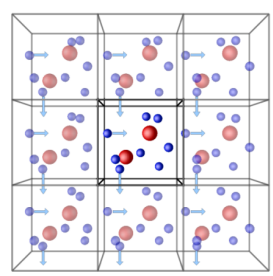


Figura 24: representación esquemática de las condiciones de límite periódico (imagen tomada de la Central Michigan University).

De esta forma, aunque un átomo se vaya de la celda de la simulación, éste se reemplaza por otro que entra por la cara contraria, exactamente con la misma velocidad. Así se conserva el número de átomos en la celda original y ninguna partícula sufre los efectos de imponer un límite físico a las caras del sistema simulado. La forma y tamaño de las cajas depende de la geometría del sistema, usándose más frecuentemente cajas cúbicas y octaédricas para la simulación de proteínas y ácidos nucleicos (Figura 25).

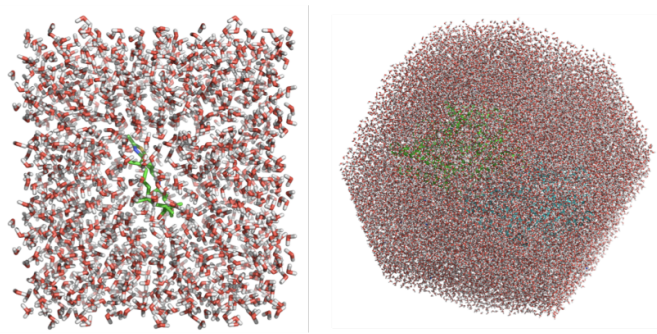


Figura 25: figuras geométricas más habituales para la simulación de sistemas macromoleculares en entornos acuosos: cubo y octaedro truncado.

El tamaño de la caja debe ser tal que el soluto no pueda interaccionar con una copia de sí mismo en ninguna de las réplicas, por lo cual debe haber suficiente espesor de disolvente alrededor del soluto en todas direcciones para que la distancia desde cualquier punto de éste hasta cualquiera de sus copias vecinas sea mayor que la distancia límite de corte (*cutoff*) establecida para el cálculo de las interacciones no enlazantes.

PREPARACIÓN Y EJECUCIÓN DE UNA SIMULACIÓN DE MD

En primer lugar, es necesario establecer una configuración inicial del sistema, que puede provenir de datos experimentales de difracción de rayos X, espectroscopía de RMN, de un modelo teórico o de una combinación de datos experimentales y teóricos.

La asignación de velocidades iniciales a los átomos del sistema, que se ha minimizado previamente mediante MM, se realiza normalmente al azar a partir de una distribución Maxwell-Boltzmann a la temperatura elegida. Finalmente, es preciso mantener fijas algunas de las condiciones de simulación: número de partículas (N), volumen (V), temperatura (T), presión (P) o energía total del sistema (E). En función de sus combinaciones se puede distinguir entre simulaciones microcanónicas (NVE), isotérmico-isobáricas (NPT) y canónicas (NVT) siendo NPT y NVT las más empleadas en MD de proteínas y ácidos nucleicos.

Si se utiliza un formalismo NPT, es necesario ajustar la temperatura y la presión, para lo cual existen distintos métodos. La manera más fácil es aplicar un escalado de velocidades (*149*) que consiste en multiplicarlas por un factor constante en cada paso de calentamiento. Un método alternativo para mantener la temperatura es acoplar el sistema a un baño térmico externo que se fija a la temperatura deseada (*150*). El baño actúa como fuente de energía térmica, añadiendo o quitando calor al sistema según sea necesario. Las velocidades se escalan en cada paso en función de la diferencia de temperatura entre el baño y el sistema. Muchos de los métodos usados para el control de la temperatura se emplean también a la hora de controlar la presión. Por lo tanto, la presión puede ser mantenida a un valor constante simplemente mediante un escalado del volumen. Y la alternativa es acoplar el sistema a un "baño de presión" análogo al baño de temperatura.

FASES DE UNA SIMULACIÓN DE MD

Las simulaciones de MD se componen de dos etapas: una fase de equilibrado y una fase de producción. El objetivo del equilibrado es llevar al sistema a un estado de equilibrio a partir de la configuración inicial. Durante esta fase se monitorizan la energía potencial, la temperatura y la densidad hasta que llegan a un valor estable. Una buena técnica para conseguir un equilibrado

óptimo es aplicar restricciones al sistema, liberándolas después lentamente para permitir su adaptación a las condiciones deseadas. Este proceso suele durar entre 200 y 500 picosegundos (ps) aunque en ciertos casos conviene equilibrar el sistema durante varios nanosegundos.

Después de un correcto equilibrado comienza la fase de producción en la que se permite la evolución del sistema, que queda recogida en forma de coordenadas y velocidades. Éstas se almacenan como una trayectoria que será utilizada para el análisis y la obtención de información (Figura 26).

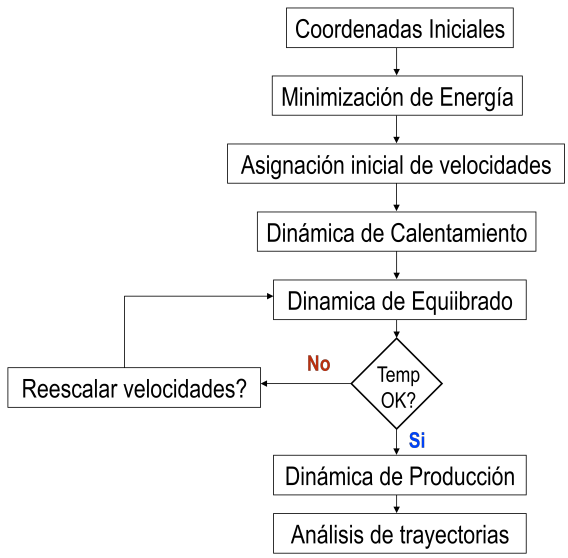


Figura 26: diagrama de flujo de trabajo de una MD.

Cuanto mayor sea el tiempo de simulación, más calidad tendrán los resultados ya que se explorará más espacio conformacional. En general se estima que el tiempo de simulación debería ser al menos 10 veces más largo que la escala temporal del proceso a estudiar. En la práctica esto significa que se tiene que hacer uso de los algoritmos de paralelización y contar con una gran capacidad de almacenamiento en discos. Por ello el uso de supercomputadores ha permitido pasar de simulaciones de 10 ns como máximo a las cada vez más habituales de 1 microsegundo (*151-153*).

Los resultados que se pueden extraer de la trayectoria de una MD son parámetros que nos darán información acerca de los cambios estructurales que afectan a las moléculas estudiadas. Los comunes a todos los sistemas son las distancias entre átomos, los ángulos, los cambios en torsionales, análisis de la superficie accesible al disolvente (*Solvent Accessible Surface Area*, SASA), la evolución de las energías y, en el caso de simulaciones de ADN, parámetros helicoidales de esta macromolécula. Pero uno de los parámetros comunes más importantes a la hora de analizar trayectorias de MD es la desviación cuadrática media (*Root Mean Square Deviation*, RMSD) respecto de una estructura de referencia, que normalmente es, o bien la estructura promedio del colectivo, o bien la estructura de partida. La función que representa la evolución del RMSD con respecto al tiempo es uno de los mejores indicadores del comportamiento global del sistema durante la simulación. Si éste se halla equilibrado, los sucesivos valores del RMSD se mantienen aproximadamente constantes, fluctuando en torno a una o varias conformaciones de equilibrio, de forma que la gráfica de la función será asintótica.

TIPOS DE SIMULACIONES DE MD

El espacio conformacional accesible a las macromoléculas biológicas es enorme y no puede ser cubierto por ninguna de las simulaciones actuales. Los datos experimentales (rayos X, RMN, etc.) indican que las biomoléculas tienden a adoptar conformaciones bastante definidas. Esta información experimental se puede utilizar en una simulación para restringir el espacio conformacional al que puede acceder el sistema, fundamentalmente dificultando la adopción de configuraciones que resultan incompatibles con la información experimental mediante una función de penalización. Cuanto más se aleje de los datos experimentales la configuración simulada del sistema, mayor será el valor de la función de penalización.

Entre los tipos de parámetros que se pueden incorporar para forzar la trayectoria de modo que se satisfaga la información experimental se encuentran:

- Los límites superiores de distancias interatómicas provenientes de medidas NOE (*Nuclear Overhauser Effect*).
- Los límites para valores que pueden adoptar los ángulos torsionales a partir de datos de acoplamiento *J* obtenidos por espectroscopía de RMN.
- Las amplitudes de los factores de estructura medidos mediante la difracción de rayos X.

Las simulaciones no se clasifican únicamente en aquéllas con restricciones (*restrained*) y sin restricciones (*unrestrained*). Existen varios tipos de protocolos empleados en el estudio de sistemas biológicos:

- TEMPLADO SIMULADO (*Simulated Annealing*) O DINÁMICA MOLECULAR "AMORTIGUADA" (*Quenched Molecular Dynamics*)

Se realiza una simulación a una temperatura elevada y después se enfría lentamente para llevar el sistema al mínimo local más próximo. La temperatura del sistema no se debe tomar como un valor real; tan solo es un parámetro de control que determina si el sistema puede escapar de ciertos mínimos locales. Este protocolo resulta útil para refinar estructuras procedentes de espectroscopía de RMN o de difracción de rayos X.

- MUESTREO AUMENTADO LOCALMENTE (*Locally Enhanced Sampling*, LES)

Esta técnica permite aumentar el muestreo de un asa (*loop*) de una proteína mediante la construcción de múltiples copias de esa región. Las copias no interaccionan entre sí pero interaccionan con el sistema de una forma promediada. Este tipo de protocolo reduce las barreras o las transiciones conformacionales y permite obtener múltiples trayectorias de esta región llevando a cabo una única simulación.

- DINÁMICA MOLECULAR DE INTERACAMBIO DE RÉPLICAS O COPIAS (*Replica Exchange Molecular Dynamics*, REMD)

Esta técnica es similar a la anterior en cuanto que permite el estudio conformacional exhaustivo de un asa o de una proteína pequeña entera. Se generan varias réplicas de un sistema que no interaccionan entre sí y se simulan a diferentes valores de una determinada variable independiente, que suele ser la temperatura. De esta forma las copias que están "atrapadas" en un mínimo local son capaces de "saltar" a otro mínimo local diferente gracias a los cambios en la temperatura. Para ello el número total de réplicas (N) se simulan en paralelo a diferentes temperaturas, que irán de forma creciente, hasta que la topología de la mitad de éstas (N/2) se intercambia por la topología de sus vecinas y se continúa la simulación bajo las condiciones de estas últimas.

- DINÁMICA MOLECULAR "ACTIVADA", "SESGADA" (*Biased*) O GUIADA (*Steered*)

Mediante la adición de fuerzas externas predeterminadas se reducen las barreras energéticas y se aumenta la probabilidad de obtener configuraciones poco probables de observar en una simulación convencional. Este procedimiento resulta adecuado para estudiar procesos que son intrínsecamente rápidos pero que constituyen acontecimientos raros porque están limitados por una o más barreras de energía.

- FORZADO SOBRE UN MOLDE O PLANTILLA (*Template Forcing*) O MD DIRIGIDA A UNA ESTRUCTURA DIANA (*Targeted Molecular Dynamics*)

En este caso se añade un término basado en la desviación cuadrática media, ponderada por la masa, a la función de energía potencial de un conjunto de átomos con respecto a los mismos en una estructura de referencia (el "blanco" o "diana"). En cada paso de MD, el algoritmo realiza un ajuste de mínimos cuadrados para los átomos seleccionados en las dos estructuras. Este método puede resultar idóneo para simular los grandes cambios conformacionales que ocurren en una escala de tiempos de milisegundos.

2.1.2.3 Cálculo de Diferencia de Energías Libres de Unión

La energía libre de unión que describe una asociación proteína–ligando o proteína–proteína es la diferencia entre la energía libre, *G*, del complejo y la de la suma de sus respectivas parejas:

$$\Delta G_{unión} = G_{complejo} - (G_{proteína} + G_{ligando})$$
$$\Delta G_{unión} = G_{complejo} - (G_{proteína} + G_{proteína}) \quad (2.12)$$

En estas ecuaciones los tres valores de *G* normalmente indican promedios sobre varios pasos de las trayectorias de MD independientes del sistema molecular en equilibrio. Pero, si queremos estudiar la evolución a lo largo del tiempo de esta energía de interacción, cada uno de los pasos de la trayectoria puede ser analizado individualmente.

$$\Delta G_{unión} = G_{complejo(i)} - \left(G_{proteína(i)} + G_{ligando(i)}\right) \quad (2.13)$$

donde *i* indica cada "*snapshot*" de la trayectoria de MD. El valor de *G* para cada especie puede descomponerse en los siguientes términos:

$$G = E_{gas} + G_{sol} - TS \quad (2.14)$$

$$E_{gas} = E_{\text{int}} + E_{ele} + E_{vdW}$$
$$E_{\text{int}} = E_{enlace} + E_{ángulo} + E_{dihedro} \quad (2.15)$$

$$G_{sol} = G_{GB} + G_{no-polar}$$
$$G_{no-polar} = \gamma \cdot SASA \quad (2.16)$$

donde E_{gas} es la energía en fase gas calculada usando el campo de fuerzas de AMBER como la suma de la energía interna (E_{int}). Esta, a su vez, es la suma de la energía de enlaces (E_{enlace}), ángulos ($E_{ángulo}$) y torsiones ($E_{dihedro}$), por una parte, y las energías no enlazadas culómbicas (E_{ele}) y de van der Waals (E_{vdW}), por otra. La energía de solvatación (G_{sol}) se descompone en contribuciones polares y no polares. La contribución polar (G_{GB}) se calcula resolviendo la ecuación del método generalizado de Born (*154*) mientras que la contribución apolar ($G_{no\text{-}polar}$) se estima a partir de la SASA (*155*) determinada con una sonda

equivalente a una molécula de agua de radio 1,4 Å, y donde γ es la constante de tensión superficial. T es la temperatura y S la entropía, que se calcula haciendo uso de la estadística clásica y análisis de modos normales de fotogramas (*snapshots*) representativos y previamente minimizados obtenidos de una trayectoria de MD.

2.1.3 Métodos Híbridos (QM/MM)

Esta técnica permite combinar potenciales de la mecánica cuántica (QM) y de la mecánica molecular (MM) en un potencial híbrido QM/MM (*1, 156, 157*). Este método combina la simplicidad y velocidad del tratamiento MM con el potencial de la QM que permite el estudio de la formación y ruptura de enlaces, así como la inclusión de la polarización electrónica debida al medio. Este tipo de aproximación nos va a permitir el estudio de reacciones químicas en el seno de grandes sistemas ya que las regiones del sistema implicados en la reacción se analizan mediante una función QM mientras que el potencial de los demás átomos del sistema se examina por cálculos clásicos de MM en los que las interacciones atómicas están regidas por el campo de fuerzas.

La energía total (E_{tot}) para este tipo de sistemas se puede escribir de la siguiente forma:

$$E_{tot} = E_{QM} + E_{MM} + E_{QM\ /\ MM} \tag{2.17}$$

donde E_{QM} y E_{MM} corresponden a la energía de aquellas partes del sistema tratadas exclusivamente con mecánica cuántica y mecánica molecular, respectivamente. Mientras que $E_{QM/MM}$ es la energía de interacción entre las partes mecánico-cuánticas y mecánico-moleculares.

Se han implementado distintos métodos con el fin de estudiar un sistema mediante QM/MM. Estos difieren entre sí (i) por el nivel de teoría utilizado para la mecánica cuántica (semiempírico, *ab initio*, enlace de valencia o funcional de densidad), (ii) por el modelo de mecánica molecular empleado o (iii) por el modo de representar el disolvente (disolvente explícito o modelo simplificado).

Un punto importante, y que también constituye una diferencia entre métodos QM/MM, es el modo de tratar la unión entre las dos regiones. En general, es preferible cortar enlaces no polares (como enlaces sencillos C–C) que cortar enlaces insaturados o polares. Existen dos métodos para abordar este problema. En una primera aproximación, se establece un orbital híbrido sp^2 que contiene un electrón a lo largo de la región QM/MM (*158*). El método alternativo y más utilizado incluye simplemente un "*link atom*" o átomo enlazante (normalmente hidrógeno) que asegura la conservación de la valencia (*1*).

2.2 Predicción del Modo de Unión de Ligandos a sus Dianas

Predecir cómo se unirá un ligando a una macromolécula de interés (ya sea una proteína o un ácido nucleico) no es trivial y la percepción subjetiva del investigador juega un papel importante. El acoplamiento (*docking*) de un ligando a una macromolécula puede realizarse de forma manual o automatizada.

Si existe una estructura cristalográfica de la macromolécula formando un complejo con uno o más ligandos, se puede utilizar la información de la posición relativa del ligando en el bolsillo de unión de la proteína o ácido nucleico para predecir el modo de unión de otros ligandos semejantes y modelarlos manualmente en el sitio de unión. En el caso de que se cuente con la estructura de la macromolécula sin ningún ligando unido, se deberá contar con información experimental para que el modelo sea razonablemente fiable, por ejemplo, el tipo de inhibición del ligando o cuáles son los residuos clave del receptor para su interacción. En este caso son de gran utilidad métodos computacionales que predicen los sitios de unión energéticamente favorables de diferentes grupos funcionales en macromoléculas de importancia biológica, como GRID (*159*). Estos grupos funcionales son llamados sondas y, seleccionando unos niveles energéticos adecuados, se pueden representar gráficamente contornos

tridimensionales que resaltan las zonas donde estas funcionalidades químicas dan lugar a una interacción favorable con la macromolécula.

Un tercer método menos sesgado consiste en explorar de forma automática y exhaustiva los posibles modos de unión de un ligando a su receptor de interés mediante programas que utilizan diferentes algoritmos que examinan el bolsillo de unión, encajan el ligando en él y evalúan las interacciones en cada posición estudiada. Para este último paso se requieren funciones de evaluación o tanteo (*scoring*) que caractericen cuantitativamente las interacciones presentes en el complejo, generalmente mediante su descomposición en contribuciones electrostáticas y de van der Waals, así como la energía de desolvatación de la superficie de contacto entre el ligando y el receptor:

$$\Delta G_{unión} = \Delta G_{vdW} + \Delta G_{ele} + \Delta G_{enlace} + \Delta G_{desolv} + \Delta G_{tor} \qquad (2.18)$$

Estos programas de *docking*, por lo tanto, realizan dos pasos consecutivos: exploración y evaluación. Como resultado de la exploración se generan un gran número de posibles modos de unión ligando–receptor ("poses"), cada uno de los cuales debe ser evaluado mediante la función de *scoring* para identificar las soluciones más favorables, desde un punto de vista energético, y las que tienen una mayor probabilidad por tanto de corresponder a la solución experimental. Estos métodos constan de tres partes: (i) la caracterización del bolsillo de unión y el precálculo de las energías de interacción con los tipos de átomo presentes en el ligando o serie de ligandos, (ii) la generación de poses de cada ligando dentro de él, y (iii) la evaluación de la energía de interacción ligando-receptor en cada complejo.

2.2.1 Caracterización del Sitio de Unión y Precálculo de las Energías de Interacción

El abordaje de esta fase es diferente dependiendo de si la estructura de la proteína carece de cualquier molécula unida o si contamos con una estructura

resuelta experimentalmente en presencia de ligandos, cofactores, moléculas de agua o iones metálicos.

En el segundo caso, habrá que caracterizar lo mejor posible las cavidades que los contienen y estudiar su modo de interacción para así intentar mimetizar estas uniones con los nuevos ligandos que se quieren estudiar, especialmente si sabemos por evidencias experimentales que se unen de la misma manera. De este modo podemos delimitar nuestra búsqueda en las regiones de interés y aumentar las probabilidades de éxito. Si la estructura de partida no tiene información de ningún ligando unido, el primer paso consiste en buscar bolsillos o cavidades donde el ligando puede supuestamente unirse al receptor. Uno de los programas más utilizados y útiles en este sentido es CASTp (*160*), que realiza una exploración de cavidades en el receptor (tanto externas, expuestas al disolvente, como internas) teniendo en cuenta que la mayoría de los centros activos y zonas de unión a ligandos están asociados a un bolsillo estructural.

Una vez definido el bolsillo del receptor donde realizar la búsqueda de poses del ligando, se define una malla tridimensional centrada en la zona de interés y, mediante sondas de diferentes tipos de átomos, se precalculan en cada nudo de la malla potenciales de afinidad atómica, tanto electrostáticos como de van der Waals, para cada tipo de átomo del ligando (Figura 27).

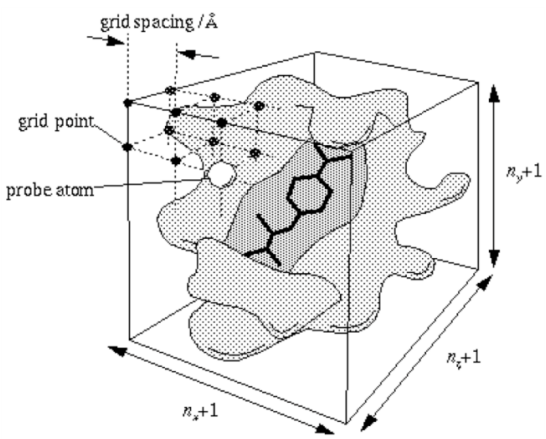


Figura 27: representación del bolsillo de unión con una malla para precalcular las energías de interacción con una sonda (imagen tomada del manual de usuario de AutoDock3.0).

De esta forma, al tener precomputados los valores de interacción entre los átomos de la proteína y los distintos tipos de átomo del ligando, resulta muy rápido calcular la energía de interacción para cada pose del ligando por interpolación trilinear.

Un factor a tener en cuenta a la hora de intentar encajar un ligando en una cavidad es que el resultado es altamente dependiente de la conformación de ésta y del resto de la proteína, ya que normalmente el bolsillo de unión se considera rígido. No obstante, hay programas como GLIDE (*161-163*) que ya incluyen en el cálculo del acoplamiento ligando–receptor una cierta flexibilidad en el centro activo al considerar diferentes conformaciones de las cadenas laterales, pudiéndose así estudiar la capacidad del receptor para adaptarse al ligando. Pero hay otras estrategias que pasan por la generación de conformaciones alternativas de la proteína mediante el análisis de modos normales (*Normal Mode Analysis*, NMA). En éste método se expresa el movimiento de una proteína en términos de movimientos colectivos de un grupo de átomos (*164*). Basándose en una aproximación armónica de la función de energía potencial del sistema en una conformación de mínima energía, el NMA proporciona una solución analítica a la ecuación del movimiento diagonalizando la segunda derivada de la matriz de energía potencial (matriz Hessiana). De esta forma, se pueden generar múltiples conformaciones del receptor que se pueden utilizar para incorporar flexibilidad en aquellos casos en que el sitio de unión se encuentra en una región "bisagra" que conecta los dominios estructurales del receptor, o cuando se sospecha que la conformación estabilizada por el ligando unido es diferente de la forma apo.

2.2.2 Generación de Poses dentro del Sitio de Unión

A la hora de buscar las posibles conformaciones y orientaciones del ligando se tienen que considerar todas las combinaciones estructurales del mismo, teniendo en cuenta su flexibilidad (ángulos torsionales activos) y sus grados de libertad (tres de rotación y tres de traslación).

2.2.3 Evaluación de la Energía

El propósito de una función de evaluación o tanteo es discriminar en un tiempo de cálculo razonable entre las soluciones correctas y el resto de las soluciones encontradas en el proceso de búsqueda. Con este objetivo se hace una estimación de la energía de unión del complejo ligando–receptor en cada posición o pose del ligando. La energía de interacción resultante de estas evaluaciones se estima utilizando campos de fuerzas, potenciales estadísticos o funciones empíricas.

En esta memoria se ha empleado el método de acoplamiento automático AUTODOCK (http://autodock.scripps.edu/) en su versión 3.0 (*165*). Éste introduce un algoritmo genético, combinado con un método de búsqueda local, y lleva a cabo la evaluación de la interacción energética mediante MM con versiones simplificadas del campo de fuerzas de AMBER. Al algoritmo genético híbrido, implementado para la búsqueda de poses en el centro activo, se le da el nombre de Algoritmo Genético Lamarckiano y es un método adaptativo basado en el proceso genético de los organismos vivos según el cual, a lo largo de las generaciones, las poblaciones evolucionan en la naturaleza de acuerdo con los principios de la selección natural y la supervivencia de los mejor dotados. Este algoritmo crea una población de posibles soluciones, cada una de las cuales es puntuada usando funciones de adaptación que evalúan su "bondad". La población cambia con el tiempo y evoluciona hacia soluciones mejores. En el caso de los confórmeros generados para el ligando, cada miembro de la población está codificado como un cromosoma y cada cromosoma está codificado por los valores de torsión de sus enlaces rotables y por su orientación dentro de la cavidad. Tanto la orientación como la conformación interna inicial variarán según evolucione la población. La población inicial se obtiene de forma aleatoria y la puntuación de cada estructura dentro del sitio de unión actúa como función de adaptación para seleccionar los individuos a partir de los cuales se generará la nueva población.

2.3 Estudio Cuantitativo de la Relación Estructura-Actividad: QSAR

Una de las principales metas a alcanzar en el diseño de fármacos es la predicción de la actividad que tendrán los nuevos compuestos antes de ser sintetizados. La cuantificación de la relación estructura-actividad (*Quantitative Structure–Activity Relationship*, QSAR) es un proceso mediante el que se relacionan ciertas propiedades cuantificables de las estructuras de una serie de moléculas con sus afinidades o actividades biológicas mediante un modelo matemático. En el desarrollo de fármacos, los métodos QSAR se han usado frecuentemente para determinar propiedades más allá de la actividad *in vivo*. Esta actividad depende en la mayoría de los casos de muchos factores. La relación entre esas propiedades numéricas y la actividad se describen mediante la siguiente ecuación:

$$v = f(p) \tag{2.19}$$

donde v es la actividad en cuestión, p son las propiedades derivadas de la estructura de la molécula y f es la función que los relaciona.

A continuación mencionaremos los tipos de métodos QSAR que existen y haremos hincapié en el método COMBINE, que ha sido el empleado en esta memoria.

2.3.1 Tipos de Métodos QSAR

2.3.1.1 Métodos de Relación Estructura-Actividad Basada en Fragmentos o Grupos (*Group Contribution QSAR, GCQSAR*)

Consiste en relacionar los fragmentos moleculares con las variaciones en las respuestas biológicas. Estos fragmentos estudiados pueden tener su origen en las diferentes sustituciones realizadas a una molécula inicial, en el caso del

estudio de un grupo de moléculas similares, o en leyes químicas predefinidas en el caso de moléculas diferentes (*166, 167*).

Un primer ejemplo de relación estructura-actividad es el que establecieron Meyer y Overton entre la potencia de diversos anestésicos y su coeficiente de reparto entre aceite y gas. Pero la primera aplicación de un método QSAR para racionalizar la actividad biológica fue realizada por Hansch (*168*) que estableció una ecuación que la relacionaba con características hidrófobas, electrónicas y estéricas de los sustituyentes sobre anillos de benceno (*1*).

2.3.1.2 Métodos de Relación Estructura-Actividad Basada en Estructuras Tridimensionales (*Three Dimensional QSAR, 3D-QSAR*)

En los últimos años, gracias a que se han obtenido estructuras tridimensionales para muchas biomoléculas y pequeños compuestos, y al avance de los gráficos moleculares y las técnicas quimiométricas, están ganando terreno estos estudios. En estos métodos se extrae información a partir de la estructura de las moléculas, ya sea por contar con datos de cristalografía, mediante superposición molecular, o mediante cálculos basados en campos de fuerzas que permiten la medición de potenciales de interacción con una malla tridimensional. Para correlacionar la estructura con la actividad medida, se emplean datos obtenidos computacionalmente como interacciones electrostáticas y de van der Waals (*1*).

2.3.1.2.1 *Comparative Molecular Field Analysis* (*CoMFA*)

El Análisis Comparativo de Campos Moleculares (CoMFA) es una técnica de 3D-QSAR que emplea datos de actividad de un grupo de ligandos conocidos y que se aplica cuando no se tiene conocimiento de la estructura tridimensional del receptor o cuando no se conoce el modo de unión de los ligandos a este receptor. Las estructuras tridimensionales de los ligandos se pueden extraer a partir de datos cristalográficos o por optimización de estructuras modeladas

mediante métodos cuánticos o de mecánica molecular.

La finalidad de este método es establecer una correlación entre la estructura tridimensional, las características electrostáticas y los patrones de enlaces de hidrógeno de un grupo de ligandos, y su actividad biológica. Esto se lleva a cabo mediante la superposición de los ligandos en lo que se supone que son sus conformaciones biológicamente activas dentro del bolsillo de unión. Cada conformación se estudia para todo el grupo de ligandos, calculándose los campos moleculares a su alrededor. Estos campos son de naturaleza estérica (interacciones de van der Waals) y electrostática (término coulómbico) y se calculan en los nudos (separados entre sí 1–2 Å) de una malla tridimensional. En cada punto se mide la interacción del ligando con un átomo de carbono de hibridación sp^3 y carga positiva, unidad que actúa como sonda. Para evitar el ruido generado por energías de interacción elevadas originadas por choques estéricos y electrostáticos se establece un corte (*cutoff*) que elimina las superiores a 30 kcal/mol. Para predecir la actividad a partir de los valores de energía se emplea el análisis de mínimos cuadrados (*Partial Least-Squares*, PLS) que se describe en el apartado dedicado al método COMBINE.

Lo más importante a la hora de emplear CoMFA con éxito es contar con una buena superposición de las estructuras de forma que las partes comunes estén bien alineadas (*169*), aunque este procedimiento no deja de ser subjetivo.

2.3.1.2.2 *COMparative BINding Energy* (*COMBINE*) *Analysis*

El análisis Comparativo de Energías de Unión (COMBINE) (*170*) es un método QSAR basado en la estructura tridimensional de los complejos ligando-receptor. Las interacciones calculadas entre los ligandos y la proteína en cada complejo se emplean para racionalizar las afinidades de unión de la serie, generalmente congenérica (*171*). Actualmente este método está integrado en una plataforma gráfica (gCOMBINE) que permite una mayor interacción y comodidad en su utilización (*172*).

El método COMBINE emplea un determinado número de energías de interacción ligando-receptor (electrostáticas y de van der Waals) calculadas en un conjunto de complejos refinados mediante técnicas de MM, para construir una matriz que posteriormente será sometida a un análisis estadístico multivariante para intentar ponderar esas interacciones y correlacionarlas con las actividades biológicas o las afinidades de unión. Opcionalmente también se pueden incorporar a la matriz los componentes electrostáticos de las energías de desolvatación de ligando y receptor calculados mediante un método continuo. La ecuación resultante tiene la siguiente forma:

$$\Delta G = \sum_{i=1}^{2n} u_i w_i + C \tag{2.20}$$

donde para cada uno de los n residuos del receptor (residuos de aminoácidos y posibles moléculas de agua implicadas en la unión) hay 2 términos, u_i, uno de van der Waals y otro electrostático, que están ponderados por los términos w_i que se determinan mediante el análisis PLS, al igual que la constante C. Hay que tener en cuenta que si una de las contribuciones medidas no varía significativamente en el conjunto de complejos, no se podrá emplear para explicar las diferencias de afinidad o actividad, por muy importante que sea para el término constante C y para la energía de unión.

Para hacer un modelo COMBINE se emplea primeramente un grupo de complejos ligando-receptor como conjunto de entrenamiento (*training set*) que se tendrá que validar de forma cruzada (*cross-validation*). Con el modelo resultante se deberían poder predecir las actividades o afinidades de otras moléculas externas al modelo, el llamado conjunto de prueba (*test set*) (Figura 28).

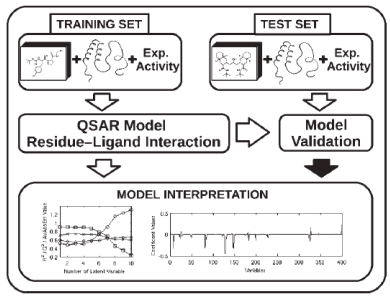


Figura 28: diagrama de flujo de trabajo de COMBINE (imagen tomada de Gil-Redondo *et. al*)(*172*).

- CONSTRUCCIÓN DE LA MATRIZ

La matriz construida contendrá todas las variables que describen las energías de interacción entre cada uno de los ligandos y los aminoácidos de la proteína. Las interacciones de van der Waals se calculan a partir del campo de fuerzas de MM (normalmente de AMBER). Las interacciones electrostáticas, por otro lado, se calculan a partir de cargas puntuales y aplicando o bien la ley de Coulomb con un dieléctrico constante o dependiente de distancia, bien mediante el método Generalizado de Born (GB) (*173*), o bien mediante métodos de Poisson-Boltzmann (*174*). En gCOMBINE basta con suministrar los archivos de entrada del programa AMBER que contienen la topología y las coordenadas de cada uno de los complejos.

Para reducir el número de variables en la matriz pero guardando la información relevante, se establece un valor mínimo de interacción (*cutoff*) que elimina aquellos residuos que no son relevantes para la unión de los ligandos, así como una desviación estándar por debajo de la cual se considera que no existen diferencias entre complejos, de modo que esas interacciones tampoco entran a formar parte del análisis. También se pueden truncar a cero los valores energéticos marcadamente positivos, que pueden deberse a inconsistencias del campo de fuerzas pero más comúnmente a errores de modelado, y realizar un escalado de las variables.

- ANÁLISIS PLS

Esta técnica combina análisis de componentes principales (*Principal Components Analysis*, PCA) y regresión linear múltiple (*Multiple Linear Regression*, MLR). Se tienen dos matrices iniciales: la matrix *X*, que contiene las variables independientes (energías de interacción y variables externas) y la matriz *Y*, que es una columna con las variables dependientes (datos biológicos):

$$X = \begin{pmatrix} E_1^1 & E_2^1 & \dots & E_M^1 & V_1^1 & V_2^1 & \dots & V_M^1 & A_1^1 & \dots & A_S^1 \\ E_1^2 & E_2^2 & \dots & E_M^2 & V_1^2 & V_2^2 & \dots & V_M^2 & A_1^2 & \dots & A_S^2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ E_1^N & E_2^N & \dots & E_M^N & V_1^N & V_2^N & \dots & V_M^N & A_1^N & \dots & A_S^N \end{pmatrix} \tag{2.21}$$

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \tag{2.22}$$

En la matriz *X*, los valores de las interacciones electrostáticas, de van der Waals y las variables externas adicionales corresponden a E_j^i, V_j^i, y A_j^i, respectivamente, mientras que *N* es el número total de compuestos, *M* el número total de residuos de la proteía y *S* el número total de variables externas. Por otra parte, en la matriz *Y* cada valor de y_i representa la actividad individual del compuesto *i*.

A partir de estas matrices, usando un algoritmo iterativo se obtiene una combinación lineal de los valores en forma de vector ortogonal. Las variables latentes (*Latent Variables*, LV) o componentes principales (*Principal Components*, PC) representan una combinación lineal de las variables originales mientras que los coeficientes extraídos dan información sobre el "peso" relativo de los diferentes términos. Esto puede reducir o ampliar la relevancia de cada interacción ligando-proteína con el fin de explicar la variación de actividad o afinidad. Es importante revisar qué compuestos de la serie tienen valores de interacción elevados o muy diferentes del resto, ya que se podrán comportar de una forma diferente al resto del grupo (*outliers*).

El programa proporciona la evolución del coeficiente de regresión (r^2) a medida que se extraen PC sucesivos y una gráfica con la recta de regresión teórica (diagonal) entre los valores determinados experimentalmente y los calculados de forma teórica que permiten la evaluación de la calidad del modelo, no solo del ajuste de la regresión del *training set*, sino de la capacidad de predicción del *test set*.

$$r^2 = \frac{\left[\sum_{i=1}^{N}(y_i - \bar{y})\left(\hat{y} - \left\langle \hat{y} \right\rangle\right)\right]^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2 \sum_{i=1}^{N}\left(\hat{y} - \left\langle \hat{y} \right\rangle\right)^2} \qquad \left\langle \hat{y} \right\rangle = \frac{\sum_{i=1}^{N} \hat{y}_i}{N} \tag{2.23}$$

- VALIDACIÓN CRUZADA

Este método se emplea para asegurarse de la robustez del modelo estadístico generado. Consiste en predecir la variable dependiente para algunos complejos que han sido excluidos del conjunto original. De esta forma, si C es el conjunto completo de N compuestos ($C = \{C_1, \ldots, C_N\}$ con actividades asociadas Y ($Y = \{Y_1, \ldots, Y_N\}$)) el programa genera subconjuntos de s elementos (si $s=1$ se conoce como *Leave One Out*) y los extrae del *training set* para convertirlos en un *test set* interno y poder predecir su actividad o afinidad. Éste es un proceso iterativo que al final obtiene una lista de actividades predichas para todos los compuestos, lo que permitirá el cálculo del coeficiente de correlación tras la validación cruzada (q^2) que describe las variaciones del conjunto de complejos que se explican con el modelo, y donde $\bar{y}$ es el valor medio de la actividad:

$$q^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{N}\left(y_i - \bar{y}_i\right)^2} \qquad \bar{y} = \frac{\sum_{i=1}^{N} y_i}{N} \tag{2.24}$$

También se extraen la desviación estándar de los errores de las predicciones (*Standard Deviation of Error of Predictions*, SDEP) y el error medio absoluto (*Average Absolute Error*, AEE):

$$SDEP = \sum_{i=1}^{N} \sqrt{\frac{\left(\hat{y}_i - y_i\right)^2}{N}} \tag{2.25}$$

$$AAE = \sum_{i=1}^{N} \frac{\left|\hat{y}_i - y_i\right|}{N} \tag{2.26}$$

Para comprobar que el ajuste de los datos al generar el modelo no se puede realizar de forma fortuita, existe la posibilidad de reasignar aleatoriamente los valores de afinidades o actividades (permutación o *Y-randomization*). De esta forma se puede comprobar que no es posible derivar un modelo si estos valores se dejan al azar (*172*).

La elección del mejor modelo de los proporcionados por el programa, pese a no existir una regla estricta, se suele hacer sobre la base de la evolución del coeficiente de validación cruzada q^2 y la desviación estándar de los errores de la predicción SDEP. La dimensionalidad óptima, que da lugar a un modelo más parsimonioso, corresponde normalmente al máximo valor de q^2 más allá del cuál el descenso de este valor va acompañado de un incremento en el número de LV extraídas.

“A goal is a dream with a deadline”

Napoleon Hill

3

OBJETIVOS

3.1 Estudio con detalle atómico de la unión de Paclitaxel a la β-tubulina. Refinado mediante técnicas de MM del cristal existente tubulina unida a Paclitaxel y propuesta del modo de unión del análogo semisintético de alta afinidad Chitax40 a la β-tubulina.

Dada la necesidad de nuevos fármacos antitumorales, muchos han sido los esfuerzos realizados para obtener nuevos análogos con mayor afinidad por la tubulina y que sean capaces de soslayar los problemas de resistencias. Con este objetivo se llevó a cabo la semisíntesis de 44 análogos de taxanos, en los que se introdujeron modificaciones en las posiciones 2, 7, 10 y 13 del núcleo central tetracíclico de la baccatina III. Dado que los efectos sobre la afinidad de las diferentes modificaciones parecían ser aditivas, se trató de encontrar la combinación óptima de éstas que confiriese al compuesto una mayor afinidad por la tubulina. La mayor potencia iría teóricamente acompañada de una menor susceptibilidad al desarrollo de resistencia debida a la sobreexpresión de la bomba de expulsión P-gp, por la que el fármaco tendría menor afinidad.

Nuestra contribución en este trabajo fue dar una explicación estructural de las diferencias de afinidad entre el cabeza de serie PXL y el análogo de mayor afinidad, Chitax40 (CTX40), se hizo preciso refinar la estructura de tubulina unida a PXL depositada en el PDB bajo el código 1JFF. Debido a su baja resolución, no

era posible extraer información precisa de las interacciones específicas con los aminoácidos que determinan el sitio de unión. Una vez refinado este complejo mediante técnicas de mecánica molecular, se empleó como molde para modelar el complejo de β-tubulina con CTX40 y así poder hacer una descomposición energética por residuo de la unión, para determinar cuáles eran las interacciones implicadas en la elevada afinidad.

3.2 Establecimiento de la relación estructura-actividad mediante un análisis COMBINE de una serie de taxanos y diseño de nuevos análogos semisintéticos modificados en las posiciones C2 y C3' del anillo de bacatina III.

La publicación de los datos termodinámicos de unión para una serie de 47 taxanos junto con el modo de unión del PXL refinado a partir de la estructura depositada en el PDB permitió abordar el estudio cuantitativo de la relación estructura-actividad de la serie. Para ello se modeló el conjunto de análogos en el bolsillo de unión de la β-tubulina bovina a partir de los complejos refinados con DXL y CTX40 obtenidos de sendas dinámicas moleculares en agua y se generó un modelo COMBINE manteniendo dos de las moléculas de agua ya que participaban en la unión de los taxanos a la β-tubulina. Al asignarse diferentes pesos a las diferentes interacciones se construyó un modelo que correlacionaba los datos teóricos de energía de unión con los valores experimentales de energías libres de unión. De esta forma, se estableció la importancia de la interacción con ciertos aminoácidos frente a otros y se encontró evidencia indirecta adicional sobre el estado de protonación de la His229, clave en la unión de los taxanos.

En base al modelo se sintetizaron, por parte de un grupo de Química Médica en Beijing (China), nuevos análogos con modificaciones en los sustituyentes de las posiciones 2 y 13 del anillo de la baccatina III. Mientras que las modificaciones en C2 exploraban interacciones en un bolsillo que ya estaban

cuantificadas, las nuevas modificaciones en C3' del sustituyente en C13 lo hacían en un bolsillo que aún no había sido explorado. Debido a esto el modelo generado a partir de los 47 taxanos predecía mejor las primeras que las segundas. Tras la incorporación de los nuevos análogos al conjunto de entrenamiento se generó un modelo COMBINE actualizado robusto que fue validado de forma cruzada.

3.3 El establecimiento de la relación estructura-actividad mediante análisis COMBINE respalda una propuesta para el modo de unión a la β-tubulina de las epotilonas.

Son muchos los trabajos experimentales y teóricos que han intentado esclarecer el modo de unión de las epotilonas a la β-tubulina. Tan pronto como se publicó la tan esperada estructura cristalográfica, depositada en el PDB con el código 1TVK, tanto la conformación propuesta para la Epotilona A como su orientación en el sitio de unión del PXL fueron puestas en duda. Mediante diversos experimentos de espectroscopía de RMN se han obtenido las conformaciones mayoritarias en disolución acuosa y en complejos con β-tubulina en diferentes estados de agregación. Cabe destacar que todos estos trabajos experimentales han ido necesariamente acompañados de estudios de modelado para proponer los modos de unión a la β-tubulina, pero ninguno de los modelos presentados proporciona una explicación de la relación estructura-actividad propuesta para varias series de análogos y de los efectos de diversas mutaciones en la proteína.

La publicación de una serie de análogos de epotilonas obtenidas mediante síntesis total, para los que se habían determinado los valores de energía libre de unión, así como las entalpías y las entropías, nos sirvió de base para realizar un análisis COMBINE con el fin de establecer si un modelo de unión alternativo a los publicados permitiría explicar la relación estructura-actividad. Para ello, mediante MD, se exploró el espacio conformacional accesible en disolución

acuosa a la Epotilona A y a un análogo representativo del grupo de análogos epimerizados en la posición 12 (EP5). A continuación se estudió el posible acoplamiento de los confórmeros resultantes en el sitio de unión de la β-tubulina y se simularon mediante MD los mejores complejos, antes y después de refinar la conformación del *M-loop* para intentar reproducir la situación en los MT estabilizados usados en los experimentos. Empleando como molde estos dos modelos ya refinados, se construyeron los restantes complejos y el conjunto se sometió a un análisis COMBINE para intentar correlacionar las entalpías de unión con las interacciones ligando-receptor.

El modelo COMBINE obtenido fue capaz, no sólo de explicar la relación estructura actividad de la serie de análogos, sino también de predecir razonablemente bien las entalpías de unión de tres análogos pertenecientes a una serie diferente. El modo de unión propuesto permite explicar además el efecto de diversas mutaciones en la tubulina sobre la afinidad de unión de las epotilonas.

3.4 Modelado y refinado de la unión de Vinblastina a la interfaz entre dos heterodímeros de tubulina asociados cabeza-cola ($\alpha_1\beta_1$-$\alpha_2\beta_2$). Propuesta del modo de unión de los análogos Vincristina, Vinorelbina y Vinflunina y establecimiento de forma semicuantitativa de la relación estructura-actividad.

Los alcaloides de la Vinca son agentes antitumorales empleados en la clínica que inhiben la polimerización de los MT. La publicación del cristal de tubulina unida a Vinblastina permitió conocer el sitio de unión de esta familia de fármacos. Pero de la misma forma que sucede con otros datos cristalográficos obtenidos con tubulina, la baja resolución obtenida no permitía el estudio en detalle del modo de unión y dificultaba un establecimiento de la relación estructura-actividad.

Para poder estudiar la unión de estos fármacos se tomaron las coordenadas de la Vinblastina y de las subunidades de tubulina que conforman su sitio de unión en el complejo depositado en el PDB como 1Z2B y se refinó mediante MD, tomándose como estructura representativa la estructura promedio de la sección más estable de la trayectoria, una vez minimizada su energía. El complejo refinado sirvió como molde para modelar los complejos de tubulina con los otros tres análogos, Vincristina, Vinorelbina y Vinflunina. Dado que se ha establecido una relación de afinidades para estos cuatro análogos, realizamos una descomposición energética por residuos que se pudo relacionar semicuantitativamente con los datos experimentales de unión.

3.5 Estudio de la unión a la tubulina de un nuevo agente desestabilizante de microtúbulos de origen marino.

El mar ha demostrado ser una fuente inagotable de nuevas moléculas. Éstas se extraen, entre otras fuentes, de esponjas marinas y de urocordados, en los que su presencia se explica a menudo como una consecuencia de la necesidad de protegerse de depredadores mediante "guerra química", debido a su reducida movilidad.

Las investigaciones en este campo realizadas por la empresa PharmaMar (Colmenar Viejo, Madrid) llevaron al descubrimiento de una nueva molécula que, en los cribados realizados, demostró poseer una actividad citostática que radica en la disrupción total del uso mitótico. Experimentos posteriores indicaron que se unía a la tubulina y competía por el sitio de unión de Vinblastina. Nuestra contribución en este trabajo fue el desarrollar un modelo molecular que explicase con detalle atómico los datos experimentales y arrojara luz sobre la unión de esta molécula y sus análogos a la tubulina.

"Cuando mayor es la dificultad, mayor es la gloria"

Marco Tulio Cicerón

4

RESULTADOS

ARTÍCULO I

"Optimization of Taxane Binding to Microtubules: Binding Affinity Dissection and Incremental Construction of a High-Affinity Analog of Paclitaxel."

Chem Biol. 2008 Jun;15(6):573-85.
[PMID: 18559268]

Se midieron las afinidades por los microtúbulos de una serie de taxanos naturales y sintéticos con el fin de determinar las contribuciones individuales de diferentes grupos. El objetivo último era obtener el conocimiento necesario para poder diseñar nuevos compuestos con mayor afinidad capaces de evitar la resistencia a fármacos. Como se observó previamente en el caso de las epotilonas, las contribuciones tanto positivas como negativas a la energía libre de unión de los diferentes sustituyentes demostraron ser acumulativas. Al combinar las sustituciones más favorables se consiguió aumentar 500 veces la afinidad del Paclitaxel. El conocimiento a nivel estructural de esta mejoría se obtuvo mediante modelado molecular y datos de RMN obtenidos de la unión de Docetaxel a microtúbulos. Los taxanos con una afinidad por los microtúbulos muy superior a la afinidad por la glicoproteína-P se muestran insensibles a la multiresistencia a fármacos. Este hallazgo indica que la optimización de la interacción ligando-diana es una buena estrategia para vencer los problemas de multiresistencia mediados por bombas de expulsión.

Chemistry & Biology
Article

Optimization of Taxane Binding to Microtubules: Binding Affinity Dissection and Incremental Construction of a High-Affinity Analog of Paclitaxel

Ruth Matesanz,[1,5] Isabel Barasoain,[1,5] Chun-Gang Yang,[2,3] Lei Wang,[2,6] Xuan Li,[2] Concepción de Inés,[1] Claire Coderch,[4] Federico Gago,[4] Jesús Jiménez Barbero,[1] José Manuel Andreu,[1] Wei-Shuo Fang,[2,3,7,*] and José Fernando Díaz[1,7,*]
[1]Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Ramiro de Maeztu 9, 28040 Madrid, Spain
[2]Institute of Materia Medica, Chinese Academy of Medical Sciences, 1 Xian Nong Tan Street, Beijing 100050, China
[3]Key Laboratory of Bioactive Substances and Resources Utilization of Chinese Herbal Medicine, Peking Union Medical College (Ministry of Education), Beijing 100050, China
[4]Department of Pharmacology, University of Alcalá, 28871 Alcalá de Henares, Spain
[5]These authors contributed equally to this work
[6]Present address: College of Life Sciences, Jilin University, Changchun, Jilin 130012, China
[7]This article is dedicated to the memory of our late colleague, Dr. Ángel R. Ortiz
*Correspondence: wfang@imm.ac.cn (W.-S.F.), fer@cib.csic.es (J.F.D.)
DOI 10.1016/j.chembiol.2008.05.008

SUMMARY

The microtubule binding affinities of a series of synthetic taxanes have been measured with the aims of dissecting individual group contributions and obtaining a rationale for the design of novel compounds with the ability to overcome drug resistance. As previously observed for epothilones, the positive and negative contributions of the different substituents to the binding free energies are cumulative. By combining the most favorable substitutions we increased the binding affinity of paclitaxel 500-fold. Insight into the structural basis for this improvement was gained with molecular modeling and NMR data obtained for microtubule-bound docetaxel. Taxanes with affinities for microtubules well above their affinities for P-glycoprotein are shown not to be affected by multidrug resistance. This finding strongly indicates that optimization of the ligand-target interaction is a good strategy to overcome multidrug resistance mediated by efflux pumps.

INTRODUCTION

Cancer is one of the major causes of premature death in humans, and multidrug resistance (MDR) of neoplastic tissues is a major obstacle in cancer chemotherapy. Though many tumors initially respond favorably to chemotherapeutic treatment, effectiveness at tumor regression is limited by the development of resistance. Although several primary reasons account for MDR, the predominant cause is the overexpression and drug efflux activity of several transmembrane proteins, as best exemplified by P-glycoprotein (P-gp) (Shabbits et al., 2001).

P-gp is a member of the ATP binding cassette (ABC) family, with broad substrate specificity for substances, including anticancer drugs, peptides, and HIV protease inhibitors. It has been shown that the extent of drug resistance in human tumors correlates well with P-gp expression (Tan et al., 2000).

In a previous work with a small group of C-2 substituted cephalomannines (CPHs) (Yang et al., 2007), we noticed that the resistance indexes for high-affinity taxanes in MDR cells are much lower than those for the medium-affinity taxanes, paclitaxel (TXL) and docetaxel (DXL), used in clinical practice. These results suggest that increasing the binding affinity of these compounds might be an alternative to overcome MDR, the rationale for this being that affinity for tubulin is the main force driving the entrance of the ligand into cells.

When MDR cells are exposed to taxanes, two opposite forces control ligand uptake: (1) binding to P-gp, which pumps the ligand out of the cell, and (2) binding to tubulin, which reduces the intracellular concentration of the ligand and keeps it bound inside the cell. Thus, the higher the binding affinity of the ligand for tubulin, the lower the intracellular concentration of free ligand. Because efflux relies on drug binding to P-gp, which in turn depends on free ligand concentration, at intracellular ligand concentrations far below its dissociation constant from P-gp, the efflux will be strongly decreased. In the most extreme case of a ligand that binds covalently to the taxane site, such as the natural product cyclostreptin, every molecule entering the cell will be finally trapped by tubulin, and the tumor cell's MDR phenotype will be completely circumvented (Buey et al., 2007).

Although numerous chemical and biological qualitative studies of the structure-activity relationships of taxanes have been performed (Zefirova et al., 2005; Kingston and Newman, 2007), an in-depth study of the contributions of the different substituents to the binding thermodynamics has not been conducted. We have previously shown that for epothilone (EPO) derivatives (Buey et al., 2004), the thermodynamic contributions of the substituents are accumulative, that is, the same substitution on different ligands produces a similar change on the binding affinity. The effect of a single modification can thus be quantified, and both favorable and unfavorable contributions can be combined to build tailor-made ligands with the desired affinities.

We now report on the thermodynamics of binding of a set of 44 taxanes (called Chitax [CTX]), plus the three reference compounds, TXL, DXL, and CPH (Figure 1), to crosslinked stabilized

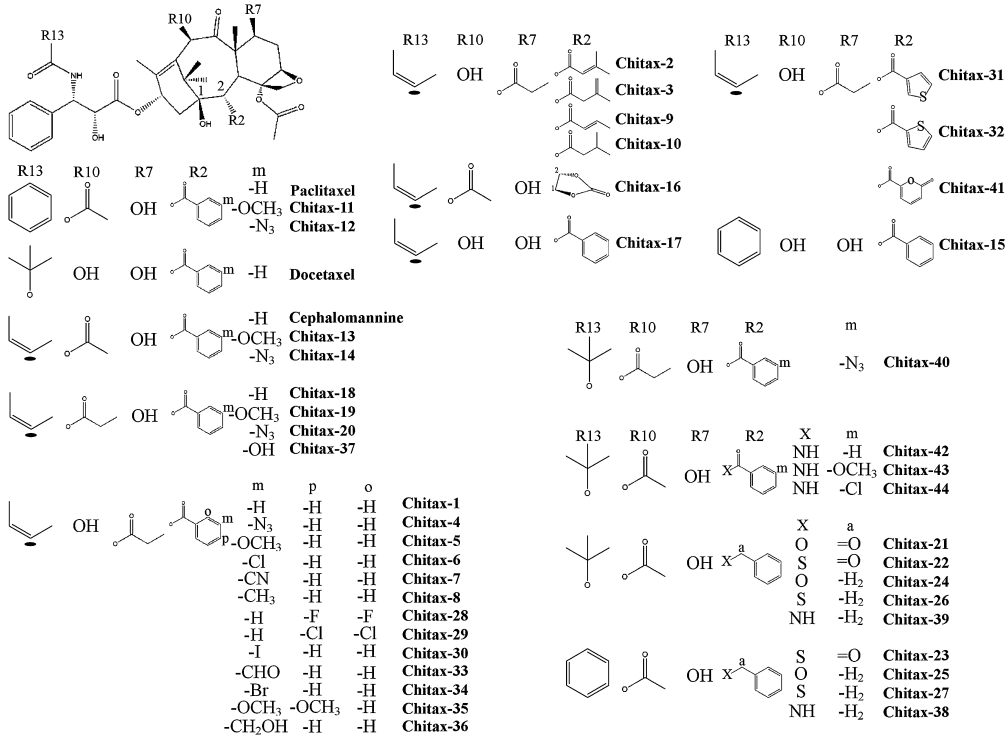


Figure 1. Chemical Structure of the Taxanes Employed in This Study

microtubules (MT) in order to quantify the contributions of single modifications at four specific locations of the taxane scaffold: two of them, C2, in the southern part of the molecule, and C13, where side chains have been described as essential for taxane activity (Chen et al., 1993b), and another two, C7 and C10, in the northern part of the molecule, where substituents have been shown as nonessential for taxane activity (Chen et al., 1993a). It is known that the side chains at these positions accept modifications that modulate the activity, favorably and unfavorably (Ojima et al., 1997; Kingston et al., 1998; Yang et al., 2007). The activity of these taxanes was investigated against the parental (A2780) and MDR P-gp-overexpressing (A2780AD) human ovarian carcinoma cell lines (Rogan et al., 1984). We were able to correlate the binding affinity of these tubulin ligands to their cytotoxicity in the resistant cells. Moreover, the resulting thermodynamic data was used to design novel high-affinity taxanes with the ability to overcome P-gp-related resistance. The higher affinity of these newly designed compounds has been rationalized by experimentally determining the tubulin bound conformation of DXL and by modeling the complexes of β-tubulin with DXL, TXL, and the best of the designed taxanes.

RESULTS

Thermodynamics of Binding of TXL Analogs to Stabilized MTs

All the compounds were first shown to be TXL-like MT-stabilizing agents (MSA) (see Table S1 available online). Then, their affinity for the taxane-binding site was measured using the competition method previously employed (Buey et al., 2004; Table 1 and Table S2). Every compound was initially measured using Flutax-2 as the competitor. Compounds 4, 11, 12, 13, 14, 19, 20, and 21 displayed very high affinities, completely displacing Flutax-2 at equimolar concentrations (Figure 2A). This indicated that they were in the limit of the range of measurement allowed by the previously employed test (Díaz and Buey, 2007).

To measure the binding affinity of these compounds more precisely, we used a direct competition experiment with a higher-affinity compound, EPO-B, whose binding affinity (7.5×10^8 M^{-1} at 35°C) has been previously determined (Buey et al., 2005; Figure 2B). This allowed the precise determination of the binding affinities of compounds 4, 11, 12, 13, 14, 19, 20, and 21, whose values range between 1.42×10^8 M^{-1} for the compound with the lowest affinity, CTX-21, and 1.51×10^9 M^{-1} for

Table 1. Apparent Binding Affinities, Thermodynamic Parameters of Binding of Taxanes for the TXL site, 10^7 M^{-1}, and Cytotoxicity of the Compounds to Nonresistant and Resistant Ovarian Carcinoma Cells

Compound	Kb 35°C	ΔG 35°C kJ/mol	ΔH kJ/mol	ΔS J/mol	IC_{50} A2780 nM	IC_{50} A2780AD nM	R/S
TXL[a]	1.43 ± 0.17	−42.1 ± 0.3	−51 ± 4	−29 ± 13	1.3 ± 0.4	980 ± 149	753
DXL[a]	3.93 ± 0.27	−44.8 ± 0.2	−53 ± 2	−26 ± 8	0.6 ± 0.2	290 ± 16	483
CPH	0.69 ± 0.08	−40.3 ± 0.3	−39 ± 6	−6 ± 18	1.5 ± 0.2	910 ± 285	606
CTX-1	0.49 ± 0.12	−39.4 ± 0.6	−56 ± 12	−56 ± 39	13.2 ± 7	1222 ± 300	92.5
CTX-2	0.043 ± 0.018	−33.2 ± 0.9	−66 ± 18	−109 ± 38	950 ± 80	10,200 ± 1900	10.7
CTX-3	0.072 ± 0.017	−34.5 ± 0.5	−32 ± 9	10 ± 29	1250 ± 200	4000 ± 700	3.2
CTX 4[b]	87 ± 19	−52.7 ± 0.5	−46 ± 13	19 ± 44	2.7 ± 0.6	14 ± 3.8	5.2
CTX-5	5.37 ± 1.39	−45.6 ± 0.6	−40 ± 4	19 ± 14	6.6 ± 1.8	160 ± 19	24.2
CTX-6	1.62 ± 0.24	−42.5 ± 0.4	−39 ± 2	13 ± 7	10 ± 2.4	274 ± 30	27.4
CTX-7	0.39 ± 0.06	−38.8 ± 0.4	−38 ± 5	4 ± 17	14.5 ± 2.9	2100 ± 660	145
CTX-8	0.492 ± 0.073	−39.4 ± 0.3	−28 ± 3	39 ± 10	22.5 ± 5	596 ± 105	26.4
CTX-9	0.028 ± 0.008	−32.1 ± 0.6	−15 ± 4	56 ± 12	3900 ± 370	15,000 ± 3500	3.8
CTX-10	0.042 ± 0.008	−33.1 ± 0.4	−12 ± 5	68 ± 18	4900 ± 600	>20000	>4
CTX 11[b]	38 ± 10	−50.5 ± 0.6	−14 ± 20	117 ± 65	1.36 ± 0.2	163 ± 37	120
CTX 12[b]	151 ± 3	−54.1 ± 0.1	−29 ± 9	84 ± 29	2.8 ± 0.38	42 ± 13	15
CTX 13[b]	16.5 ± 2.8	−48.4 ± 0.4	−28 ± 1	66 ± 1	1.3 ± 0.2	128 ± 17	98.4
CTX 14[b]	80.0 ± 2.9	−52.5 ± 0.1	−46 ± 6	21 ± 19	1.6 ± 0.3	25 ± 10	15.6
CTX-15	2.384 ± 0.53	−43.5 ± 0.5	−18 ± 11	80 ± 35	17.5 ± 2.7	5250 ± 1000	300
CTX-16	0.050 ± 0.030	−33.6 ± 1.2	−18 ± 9	49 ± 30	740 ± 100	7360 ± 750	9.9
CTX-17	0.902 ± 0.37	−41.0 ± 0.9	−24 ± 3	55 ± 10	18 ± 5.6	5412 ± 1200	301
CTX-18	1.281 ± 0.27	−41.9 ± 0.5	−22 ± 4	64 ± 14	2.1 ± 0.8	452 ± 36	215.2
CTX 19[b]	14.8 ± 0.2	−48.2 ± 0.1	−48 ± 3	0 ± 9	0.54 ± 0.07	39 ± 11	72.2
CTX 20[b]	80.6 ± 5.1	−52.5 ± 0.2	−92 ± 19	−124 ± 64	3.9 ± 1.2	27.4 ± 4	7
CTX 21[b]	14.2 ± 1.6	−48.0 ± 0.3	−26 ± 5	72 ± 16	1.9 ± 0.3	41 ± 11	21.5
CTX-22	0.013 ± 0.00	−30.1 ± 0.0	−46 ± 7	−151 ± 23	2400 ± 1000	6960 ± 670	2.9
CTX-23	0.007 ± 0.00	−28.6 ± 0.0	−37 ± 7	−123 ± 22	11,500 ± 1000	23,800 ± 2200	2.1
CTX-24	0.094 ± 0.01	−35.2 ± 0.3	−65 ± 10	−95 ± 33	353 ± 19	8600 ± 3800	24.3
CTX-25	0.008 ± 0.00	−28.9 ± 0.0	−47 ± 7	−154 ± 24	6200 ± 2600	28,500 ± 5300	4.6
CTX-26	0.018 ± 0.00	−31.0 ± 0.0	−73 ± 11	−77 ± 137	3500 ± 1400	8300 ± 2700	2.4
CTX-27	0.007 ± 0.00	−28.6 ± 0.0	−42 ± 6	−138 ± 21	10,000 ± 740	9400 ± 1700	0.94
CTX-28	0.17 ± 0.04	−36.7 ± 0.5	−17 ± 8	62 ± 25	82 ± 16	1880 ± 200	22.9
CTX-29	0.25 ± 0.08	−37.7 ± 0.7	−52 ± 9	−48 ± 30	102 ± 8.8	690 ± 60	6.8
CTX-30	1.76 ± 0.91	−42.7 ± 1.1	−11 ± 15	92 ± 49	30 ± 0.5	246 ± 27	8.2
CTX-31	0.10 ± 0.04	−35.4 ± 0.9	−44 ± 11	−30 ± 36	106 ± 4.2	2950 ± 480	27.8
CTX-32	0.24 ± 0.08	−37.6 ± 0.7	−7 ± 16	87 ± 55	62 ± 17	3200 ± 250	51.6
CTX-33	0.07 ± 0.01	−34.4 ± 0.3	−81 ± 12	−152 ± 39	69 ± 77	1500 ± 100	21.7
CTX-34	1.20 ± 0.93	−41.7 ± 1.5	−38 ± 28	14 ± 94	28.7 ± 1.9	196 ± 14	6.8
CTX 35	0.88 ± 0.77	−40.9 ± 1.6	−29 ± 33	13 ± 112	25 ± 2	153 ± 39	6.12
CTX 36	0.029 ± 0.020	−32.2 ± 1.3	−48 ± 6	−63 ± 18	1700 ± 120	>20000	>11.7
CTX 37	0.035 ± 0.010	−32.7 ± 0.6	−91 ± 7	−207 ± 24	86 ± 9.8	10,000 ± 1000	116.2
CTX 38	0.001 ± 0.001	−23.6 ± 1.8	−44 ± 9	−70 ± 34	15,400 ± 3200	20,000 ± 3000	1.3
CTX 39	0.003 ± 0.001	−26.4 ± 0.7	−157 ± 10	−461 ± 28	4200 ± 100	5700 ± 300	1.3
CTX 40[b]	628 ± 15	−57.7 ± 0.1	−26 ± 24	99 ± 80	7 ± 1	9.1 ± 0.45	1.3
CTX 41	0.021 ± 0.004	−31.4 ± 0.4	−94 ± 9	−202 ± 28	14,000 ± 2000	>20000	1.4
CTX 42	0.008 ± 0.003	−28.9 ± 0.8	−106 ± 13	−250 ± 43	192 ± 20	2750 ± 430	14.3
CTX 43	0.030 ± 0.008	−32.3 ± 0.6	−85 ± 20	−169 ± 66	69.5 ± 3.8	331 ± 70	4.8
CTX 44	0.001 ± 0.001	−23.6 ± 1.8	−177 ± 23	−493 ± 75	>20,000	>20,000	

Errors are standard errors of the mean.
[a] Binding affinity and thermodynamic parameters data are from Buey et al. (2004).
[b] Compounds measured with the EPO-B displacing method.

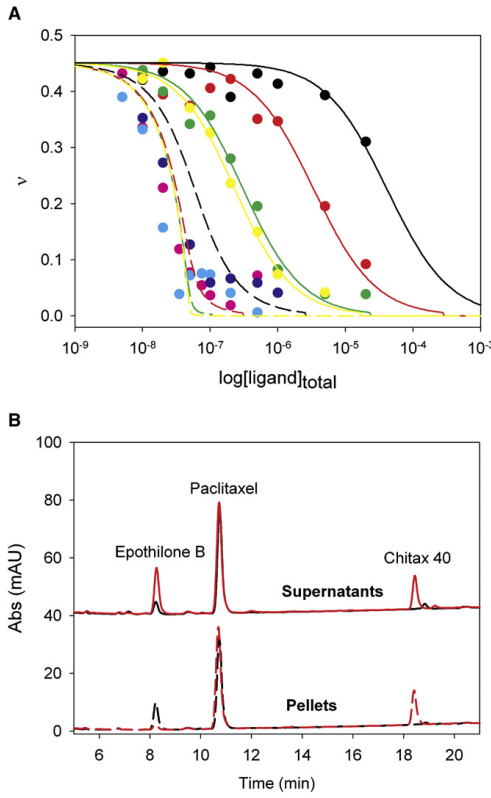


Figure 2. Binding of the Ligands to the Paclitaxel Site

(A) Displacement of the fluorescent taxane Flutax-2 bound to MT sites (50 nM) by taxanes at 35°C. The solid lines were generated with the best-fit value of the binding equilibrium constant of the competitors with binding affinities lower to 10^7 M^{-1}, assuming a one-to-one binding to the same site. Additional lines (dashed) show the expected displacement for ligands with binding constants of 10^8 M^{-1} (black), 10^9 M^{-1} (red), 10^{10} M^{-1} (green), and 10^{11} M^{-1} (yellow). Ligands binding data are as follows: green CPH, red CTX-2, yellow CTX-6, dark blue CTX-11, magenta CTX-12, light blue CTX-14, black CTX-27.

(B) Displacement of EPO-B bound to MT sites (10 μM) by CTX-40 at 35°C. One micromolar TXL-binding sites were incubated with 1.1 μM EPO-B (black lines) or with 1.1 μM EPO-B and 1.1 μM CTX-40 (red lines), MTs were harvested by sedimentation, ligands extracted from supernatants (solid lines) and pellets (dashed lines), and HPLC analyzed. One micromolar TXL was used as the internal standard. Supernatant traces were displaced 40 mAU for presentation purposes.

the highest-affinity compound, CTX-12. This method was validated when the binding affinities of compounds 11, 13, 19, and 21 (those in the range of 10^8 M^{-1}) were shown to be similar using either EPO-B or Flutax-2 (see the Supplemental Data).

To confirm that the high affinity of compounds containing 3-N_3 benzoyl at C2 does not originate from covalent binding of its reactive azido group to β-tubulin, we performed experiments in which the amounts of reversibly bound compounds 4, 12, and 14 were measured. The bound compounds could be extracted from the pellets and aqueous solutions with the aid of an organic solvent, indicating that they are not irreversibly bound.

Molecular Modeling

One conspicuous characteristic of the taxane-binding site in β-tubulin (Lowe et al., 2001) is the presence of the highly exposed side chain of His229 (located in the middle of helix 7 and positionally equivalent to Arg229 in α-tubulin) that splits the cavity into two major pockets. Because continuum electrostatic calculations predicted the imidazole ring of this residue to be doubly protonated at pH 6.5 (pK_{His229} = 7.2), this ionization state was used in subsequent work.

Molecular dynamics (MD) studies of TXL, DXL, and CTX-40 in aqueous solution provided us with a range of different conformers, the most abundant of which (Figure S1) were independently studied in a first rigid-body approach by the automated docking program. The conformation previously reported for TXL (Lowe et al., 2001; Snyder et al., 2001) using either DOCK or FlexX was also found by AutoDock among the best scoring solutions. Because similar poses were found for DXL (Figure S2) and CTX-40 as well, this common disposition of the taxane in the binding site of β-tubulin was used in the modeling of all the complexes. Noteworthy is that these conformations are not the most populated in aqueous solution (data not shown) although they are sporadically observed in the course of transitions toward other more stable and "hydrophobically collapsed" conformations. The proposed docked conformation for TXL is then coincident with that previously found in crystals of 7-mesylpaclitaxel, which was reportedly induced by specific interactions of the side chain at C13 with solvent (Gao and Chen, 1996).

The feasibility of the resulting modeled complexes was assessed by subjecting each of them to a 10 ns MD simulation followed by a simulated annealing procedure that provided us with a set of representative complex structures for further analysis and energy decomposition. In all cases, the ligand adopts a conformation in good agreement with the T-taxol geometry (Snyder et al., 2001; Figure 3) and is anchored in the binding site through a common set of well-defined interactions. Thus, the oxetane oxygen of the taxane is engaged in a good hydrogen bond with the NH of Thr276, whereas another hydrogen bond is established between the amide or carbamate carbonyl oxygen on the C13 substituent and the Nε of His229. The common phenyl ring at C13 (3′-Ph) establishes close van der Waals contacts with the hydrophobic side chains of Val23 and Ala233, whereas the benzoyl phenyl ring at C2 (2-OBz) gets lodged into another hydrophobic cavity, on the other side of His229, made up of the side chains of Leu217, Leu219, and Leu275. The offset stacking interaction of this latter phenyl with the imidazole ring of His229 (Figure S3) is improved by the substituent at the meta position whose 1,2- (methoxy, CTX-13) or 1,3-dipole (azide, CTX-40) additionally establishes a favorable electrostatic interaction with the amide dipole of the backbone peptide bond between His229 and Leu230 (Figure 3). Of the three hydroxyl groups that are common to the four ligands studied, that present on the C13 substituent is consistently engaged in a hydrogen bonding interaction with the carboxylate of Asp26 in helix 1 and the backbone nitrogen of Arg369, whereas that at C7 can

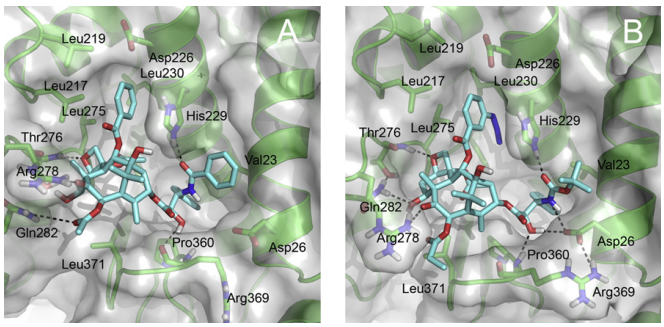


Figure 3. Model of the Ligands Bound to the Paclitaxel Site

(A) TXL and (B) CTX-40 in the binding site of β-tubulin at the end of the simulated annealing procedure. Note that the doubly protonated imidazole ring of His229 participates in (1) a hydrogen bonding interaction with the amide or carbamate carbonyl oxygen on the C13 substituent through the Nε, and (2) an offset stacking interaction with the phenyl ring on the C2 substituent.

establish transient hydrogen bonds with the carboxamide group of Gln282. The hydroxyl at C1 is permanently exposed to the solvent.

The four complexes yielded very low RMS deviations for the protein atoms with respect to the refined β-tubulin-TXL structure 1JFF (~1.3 Å on average for 400 atoms). The major differences when compared with this particular complex are the presence of a different rotameric state for His229, which we propose is protonated at physiological pH, and improved stacking and hydrogen bonding interactions between the ligand and the protein as a consequence of the mutual adaptation resulting from the simulated annealing procedure.

NMR Characterization of Bound Docetaxel

As a further step, and to provide empirical support for the modeling-derived conformations that formed the basis for the quest of the structure-activity relationship, the MT-bound conformation of DXL was elucidated, under the experimental conditions used for determining the binding constants (see the Supplemental Data file DXLNMR.pdb).

As previously shown, the transferred nuclear Overhauser enhancement (TR-NOESY) technique provides an adequate means to determine the bound conformation of ligands that exchange between free and bound states at a reasonably fast rate. TR-NOESY experiments were then performed on the DXL:MT sample at different mixing times. Negative crosspeaks were clearly observed at 310 K (Figures 4A and 4B), as expected for a ligand that binds to the assembled MTs preparation, in contrast with the lack of NOEs detected in the free state (Figure 4C).

Two control experiments were performed employing either Flutax-2 instead of DXL (the effective k_{off} of Flutax-2 release from MTs is 1.63 ± 0.18 s^{-1} (Diaz et al., 2000) or both DXL and discodermolide (DDM) at equimolecular concentrations. No TR-NOESY signals were observed in the presence of Flutax-2, indicating that the effective k_{off} of DXL is higher than that previously measured for Flutax-2, and the DXL signals were cancelled out by DDM (whose affinity for the TXL-binding site is 100× higher), indicating that DXL is effectively bound to the TXL site (Buey et al., 2005).

Cytotoxicity in Resistant and Nonresistant Tumor Cells

To check the effects of the studied modifications on the cytotoxicities of the compounds and to validate the binding affinity approach as a tool to be used in ligand optimization, we performed IC_{50} tests in A2780 human ovarian carcinoma cells and their MDR A2780AD counterparts (Table 1 and Figure 5A).

The cytotoxicities of the ligands in non-P-gp-overexpressing A2780 cells show a linear relationship (r = 0.81) with their binding affinities, but only for those compounds with a ΔG at 35°C higher than −47.5 kJ/mol^{-1} (K_b at 35°C higher than 10^8 M^{-1}), which points to a limit in the cytotoxicity that can be achieved. Thus, despite the increase in affinity of three orders of magnitude between CPH and CTX-40, the IC_{50} remained in the order of nanomolar.

In the case of the MDR A2780AD cells, a good linear relationship (r = 0.80) between cytotoxicities and binding affinities is observed for the full set of ligands, strongly suggesting that for these P-gp-overexpressing cells, tubulin binding is the main force competing with P-gp-mediated extrusion.

The best regression lines have slopes of 1.10 for A2780 cells and 0.61 for A2780AD cells, which indicates that P-gp-overexpression effectively reduces the intracellular drug concentration, thus making it necessary to increase the load to exert a cytotoxic effect.

Intake of Taxanes by Tumor Cells

The amount of compound made available for binding to the tubulin site was measured by employing radioactively labeled DXL and TXL at the concentrations needed to stop the cell cycle in G_2/M, in two leukemic cell lines (U937 and K562) and the kidney epithelial nontumor cell line PtK2 from *Potorus tridactylis*. In these conditions, the intracellular drug concentration, which ranges from 0.3 to 2.8 μM and represents a small percentage of the total drug and the cell tubulin concentrations (Table S3), increases with the total drug concentration and reaches a maximum in PtK2 cells at 300 nM DXL and 600 nM TXL. The equilibration of the ligand inside the cells is fast, with a half-life of 3 min for ^{14}C-DXL and 10 min for ^{3}H-TXL at a drug concentration of 1 μM. The drug inside the cell was found to be in the cytoplasm, with only a small fraction bound to the nucleus, as expected.

Because the total intracellular drug concentration is more than one order of magnitude above its MT dissociation constant and much lower than the total tubulin concentration (which was considered to be ~5% of the total protein measured), the mass action law dictates that most of the compound inside the cells is essentially bound to tubulin and the intracellular concentration

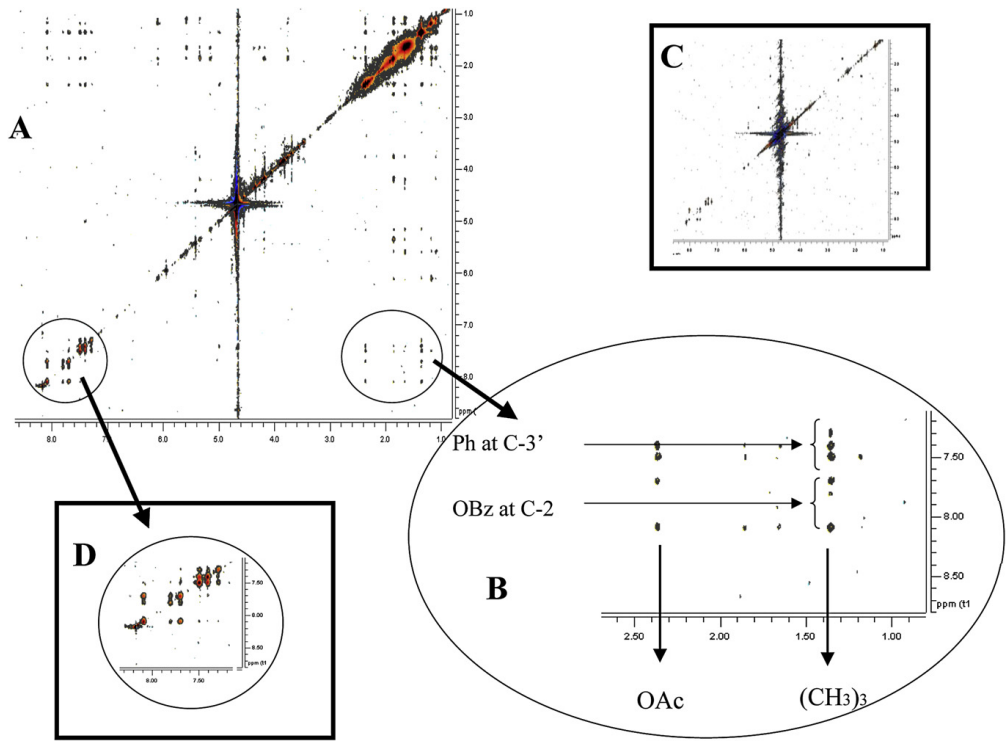


Figure 4. NMR Characterization of the Prerelease Conformation of Bound Docetaxel

(A) The 500 MHz TR-NOESY (mixing 200 ms, 310 K) spectrum of DXL in buffered water solution in the presence of MTs (20:1 molar ratio).

(B) Expansion showing the key TR-NOESY region: close contacts between the tert-butyl chain and the aromatic 2-OBz ring (besides the trivial crosspeaks with the vicinal 3′-Ph moiety).

(C) The NOESY experiment under the same experimental conditions (200 ms) in the absence of MTs did not show any crosspeak. (D) Expansion of the aromatic region in the TR-NOESY spectrum; no NOEs between the two aromatic 3′-Ph and 2-OBz moieties are observed.

of free drug is close to the dissociation constant (i.e., 70 nM for TXL and 25 nM for DXL).

DISCUSSION

The effects of modifications on the substituents attached to the baccatin scaffold on the cytotoxicity of taxanes have been qualitatively discussed in several reviews (Zefirova et al., 2005; Kingston and Newman, 2007). However, the fact that these studies were performed in different cell lines precludes a rigorous evaluation of the relationship between structural changes and cytotoxicity. To quantify the effects of substitutions at a set of specific positions in a systematic way, the binding affinity for the taxane-binding site on β-tubulin has proved to be a more precise and objective parameter (Buey et al., 2004).

With all the binding constants determined at a given temperature (35°C), it has been possible to determine the changes in apparent binding free energy caused by single-group modifications (Table 2) and to select the most favorable substituents at the positions chosen for optimization. Once this knowledge was obtained, it became feasible to design several optimized taxanes.

Effect on Binding Affinity of Changes at the C2 Position

C2 modifications have proven to be the most effective in modulating the activity of taxanes. Thus, the 2-OBz is essential as its removal (Chen et al., 1993b) or replacement with other small side chains (Ojima et al., 1994; Nicolaou et al., 1995) results in almost total loss of activity in the human colon cancer cell line tested. In contrast, changes in the structure of the ring, including its replacement with nonaromatic or heterocyclic rings, result in only moderate losses of antitumor activity (Ojima et al., 1994). Introduction of substituents on the 2-OBz ring, (Nicolaou et al., 1994; Kingston et al., 1998; Yang et al., 2007) results in increases of activity for small groups at the meta position but loss of activity

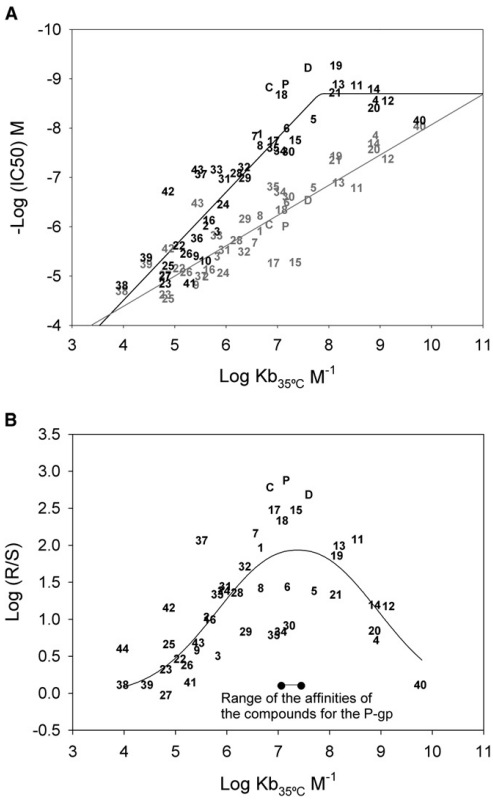


Figure 5. Comparison between Binding Affinity and Cytotoxicity in A2780 and A2780AD Cells

(A) Dependence of the IC_{50} of the ligands against A2780 (black) and A2780AD (gray) human ovarian carcinoma cells on the affinity for the TXL-binding site in MTs (binding constant, K_b). The black line represents the best fit of IC_{50} against A2780 cells versus binding affinity for ligands with binding affinity under $5 \times 10^8\ M^{-1}$. The gray line represents the best fit of IC_{50} against A2780AD cells versus binding affinity.

(B) Dependence of the resistance index of the A2780AD MDR cells on the affinity of the compounds for the taxane binding site. The range of binding affinities of taxanes for P-gp was taken from Yang et al. (2007).

for the other positions. Changes in the linker connecting the benzene ring to the taxane core also result in decreased biological activity (Wang et al., 2007).

Our results confirm and extend the qualitative data summarized above and provide a precise quantitative characterization of the effects of C2 modifications on binding affinity:

(a) Changing the nature of the linker between the benzene ring and the taxane scaffold results in a large loss of binding free energy. Thus, the replacement of the ester by an ether, thioether, or amine moiety (compounds 24, 25, 26, 27, 38, and 39) or by a thioester or an amide (compounds 22, 23, 42, 43, and 44) results in a heavy loss of binding free energy (up to 20 kJ/mol^{-1}). This stringent requirement indicates that the angle between the benzyl ring and the taxane core has to be strictly preserved and must be related to steric hindrance, as previously discussed for the thiobenzoyl compounds 22 and 23 (Wang et al., 2007), because the analogs in which the benzyl group is replaced by an alkyl ester (compounds 2, 3, 9, 10, and 16) only display a moderate loss of binding free energy (5–6 kJ/mol^{-1}), which can be assigned to the loss of interactions between the benzyl ring and the binding site.

(b) Modification of the meta substituents on the benzyl ring leads to gains of binding free energy that are the largest for $-N_3$ (-11.2 ± 1.1 kJ/mol^{-1}) and $-OCH_3$ (-7.2 ± 0.6 kJ/mol^{-1}) groups (compounds 4, 5, 11, 12, 13, 14, 20) and much smaller (-2–3 kJ/mol^{-1}) for halogen atoms (compounds 6, 30, and 34), whereas other small groups ($-CN$ and $-CH_3$) have no effect (compounds 7 and 8). Other similarly small substitutions ($-OH$ and $-CH_2OH$; compounds 36 and 37) are detrimental, resulting in a loss of 7–9 kJ/mol^{-1} of binding free energy.

(c) Introduction of double substituents at the 2,4 (compounds 28, 29) and 2,5 (35) positions results in loss of binding affinity.

(d) The thienoyl moiety (compounds 31 and 32) can effectively replace the benzoyl group.

Because previous work from our group (Buey et al., 2004) has shown that substitutions leading to gains in binding free energy also give rise to increased cytotoxicity, the $-N_3$ substituent at the meta position of 2-OBz was selected as the most suitable for molecule optimization.

Effect on Binding Affinity of Changes at the C13 Position

The side chains present at position C13 in one semisynthetic (DXL) and two natural taxanes (CPH and TXL) were evaluated to choose the optimal one for binding. Although from the direct comparison of TXL and DXL alone (Diaz and Andreu, 1993) it is not possible to assess the effect of the C13 side chain on the binding free energy due to the presence of additional differences in the substituents at C7 and C10, DXL showed a 1.9-fold larger binding affinity relative to TXL, which corresponds to a change of -1.6 kJ/mol^{-1} in free energy of binding. A similar difference (2×) was observed in the cytotoxicity on 1A9 cells (Buey et al., 2005). In contrast, the tubulin binding affinity and the cytotoxicity of CPH are about 2× lower than those of DXL (Yang et al., 2007).

In our series, by comparing compounds TXL, DXL, CPH, 11, 12, 13, 14, 15, 17, 20, 21, 22, 23, 24, 25, 26, 27, 38, 39, and 40, differing only in the side chain at C13 present in the reference molecules, we can now establish that, of the three side chains, DXL provides the largest contribution to the binding free energy, ($\Delta\Delta G$ C13-DXL$-$C13-TXL = -3.2 ± 0.9 kJ/mol^{-1}, $\Delta\Delta G$ C13-CPH$-$C13-DXL = -5.6 ± 1.1 kJ/mol^{-1}).

Effect on Binding Affinity of the Substituents Present in the Northern Side of the Taxane Ring (C7 and C10)

Modifications on the northern face of TXL at positions C7 and C10 have little effect on tubulin binding, as expected from the previously described effects on cytotoxicity (Chen et al., 1993a). From all the substituents tested at C10, the best one turns out to be the propionyl group, which provides an

Table 2. Incremental Free Energy of Binding of TXL Analogs to MTs Due to Single Group Modifications

Modification Type	Modification	Compounds	ΔΔG	Avg (Std Err)
C2	Benzoyl → benzylether	T → 25	+13.2	+13.0 ± 0.2
		21 → 24	+12.8	
	Benzoyl → benzylsulphur	T → 27	+13.6	+15.9 ± 2.3
		21 → 26	+18.1	
	Benzoyl → benzylamine	T → 38	+18.6	+20.1 ± 1.5
		21 → 39	+21.6	
	Benzoyl → thiobenzoyl	T → 23	+19.6	+15.9 ± 3.8
		21 → 22	+12.1	
	Benzoyl → benzamide	21 → 42	+19.2	
	Benzamide → 3-OCH_3-benzamide	42 → 43	−3.4	
	Benzamide → 3-Cl^-benzamide	42 → 44	+5.3	
	Benzoyl → 3-methyl- 2 butenoyl	1 → 2	+6.2	
	Benzoyl → 3-methyl- 3 butenoyl	1 → 3	+4.9	
	Benzoyl → 2(E)-butenoyl	1 → 9	+7.3	
	Benzoyl → 3-methyl- butanoyl	1 → 10	+6.3	
	Benzoyl → 2-debenzoyl-1,2-carbonate	C → 16	+5.8	
	Benzoyl → 3-N_3 Benzoyl	1 → 4	−8.0	−11.2 ± 1.3
		T → 12	−13.9	
		C → 14	−12.2	
		18 → 20	−10.6	
	Benzoyl → 3-OCH_3-benzoyl	1 → 5	−6.2	−7.2 ± 0.6
		T → 11	−8.3	
		C → 13	−8.1	
		18 → 19	−6.3	
	Benzoyl → 3-Cl^-benzoyl	1 → 6	−3.1	
	Benzoyl → 3-Br-benzoyl	1 → 34	−2.3	
	Benzoyl → 3-I-benzoyl	1 → 30	−3.3	
	Benzoyl → 3-NC-benzoyl	1 → 7	+0.6	
	Benzoyl → 3-CH3-benzoyl	1 → 8	0.0	
	Benzoyl → 3-CH_2OH-benzoyl	1 → 36	+7.2	
	Benzoyl → 3-OH-benzoyl	18 → 37	+9.2	
	3-Cl^-Benzoyl → 2,4-di-Cl^-benzoyl	6 → 29	+4.8	
	Benzoyl → 2,4-di-F-benzoyl	1 → 28	+2.7	
	3-OCH_3-Benzoyl → 2-5-di- OCH_3-benzoyl	5 → 35	+4.6	
	Benzoyl → 2-thienoyl	1 → 31	+4.1	
	Benzoyl → 3-thienoyl	1 → 32	+1.8	
	Benzoyl → 6-carboxy-pyran-2-one	1 → 41	+8.1	
C13 Side Chain	TXL → CPH	T → C	+1.9	+2.0 ± 0.2
		11 → 13	+1.9	
		12 → 14	+1.6	
		15 → 17	+2.4	
	TXL → DXL	23 → 22	−1.7	−3.2 ± 0.9
		25 → 24	−6.2	
		27 → 26	−1.3	
		38 → 39	−2.8	
		T → 21	−4.2	
	CPH → DXL	C → 21	−3.8	−5.6 ± 1.1
		17 → D	−7.7	
		20 → 40	−5.2	
C10	Acetyl → –OH	T → 15	−1.3	−1.7 ± 0.8

Table 2. *Continued*

Modification Type	Modification	Compounds	ΔΔG	Avg (Std Err)
		C → 17	−0.7	
		21 → D	−3.2	
	Propionyl → –OH	18 → 17	+0.9	+0.9
	Acetyl → propionyl	C → 18	−1.6	−0.5 ± 0.4
		13 → 19	+0.2	
		14 → 20	0.0	
C7	Propionyl → –OH	17 → 1	−1.6	−1.6

Errors are standard errors of the mean.

incremental free energy change of around −0.5 kJ/mol^{-1} over the natural C10 acetyl. In contrast, introduction of the same group at position C7 brings about a loss of 1.6 kJ/mol^{-1} of binding free energy relative to the −OH present in TXL. For this reason, a propionyl at C10 and a hydroxyl at C7 were selected as optimal substituents at these positions.

Optimal Taxane

According to the data measured, the optimal taxane should have DXL's side chain at C13, an *m*-N_3-benzoyl at C2, a propionyl at C10, and a hydroxyl at C7. Starting from compound 1, the first one in the series, with an apparent binding affinity of −39.4 kJ/mol^{-1}, the resulting molecule should gain −5.6 kJ/mol^{-1} from the replacement of the CPH side chain with that of DXL, −11.2 kJ/mol^{-1} from the substitution of *m*-N_3-benzoyl for benzoyl at C2, −1.6 kJ/mol^{-1} from the change of a propionyl at C7 to a hydroxyl, and −0.9 kJ/mol^{-1} from the change of a hydroxyl at C10 to a propionyl. Taking all of these changes together, the optimal taxane would have a predicted ΔG at 35°C of −58.7 kJ/mol^{-1}. When the compound was synthesized (CTX-40) and its binding affinity was measured using the EPO-B displacement method (K_b 35°C = 6.28 ± 0.15 × 10^9 M^{-1}; ΔG = −57.7 ± 0.1 kJ/mol^{-1}), the experimental value was found to be in good correspondence with the predicted value.

Structural Interpretation of the Binding Data

Two different reasons have been proposed for the changes in activity due to modifications at C2: (1) the need for a hydrophobic group to maintain the proper taxane conformation or (2) direct interactions of the benzoyl with hydrophobic side chains of the protein (Zefirova et al., 2005). The modeling data support the view that the higher affinity of CTX-40 relative to TXL and DXL, which is mostly conferred by the phenylazide substituent present at C2, may largely stem from the simultaneous improvement of the stacking interactions with the imidazole ring of His229(+) and a better electrostatic interaction with Asp26 and Arg369 on the opposite side of the molecule, resulting from a better anchoring of the ligand in the binding site. The same rationale applies to CTX-13, which has a methoxy substituent in place of the azide, and to a lesser extent to the derivatives containing halogen atoms. As regards the methyl, cyano, hydroxyl, or hydroxymethyl substituents, they are likely to be found facing the solvent rather than orientated toward the binding pocket, thus contributing negligibly to the binding affinity, in good accord with the experimental evidence.

Likewise, the enhanced affinity contribution of the DXL and CTX-40 side-chain at C13 relative to that of TXL arises from an improved hydrogen-bonding interaction of the carbamate NH relative to the amide NH with the carboxylate of Asp26.

Prerelease Conformation of Bound Docetaxel

Transferred NOESY signals arise from a free DXL molecule whose protons have been excited when still bound to the protein but relaxed after release from the binding site. Thus, the conformation deduced from these signals corresponds to a prerelease state of the ligand. It has been described that MT-stabilizing agents binding to the TXL site reach their luminal final location through prior transient binding to a site located in the MT pore (Diaz et al., 2003; Buey et al., 2007). Therefore the NMR structural data have to be interpreted with this caution.

Because DXT release from MT following excitation has to be fast in order to get trNOESY signals, the K_{off} of the ligand should be fast in the relaxation time scale. This is apparently in contradiction with the slow dissociation constant measured for TXL in a kinetic study, 0.091 ± 0.006 s^{-1} (Diaz et al., 2003). This observed dissociation constant does not correspond to the release step but to the rate-limiting step of the reaction. The dissociation of taxanes from MT has been studied in detail using the fluorescent taxane derivatives Flutax-1 and Flutax-2, which dissociate from MT following a two-step mechanism (Diaz et al., 2000). The first step, which is the slower one (thus the one directly observed), has a kinetic rate constant of 0.022 ± 0.001 s^{-1} (4-fold slower than that of TXL), whereas the second one (responsible for the release of the ligand to the medium and which can be measured only indirectly from the dependence of the kinetic rate constants on concentration) is nearly 100 times faster (k_{off} = 1.63 ± 0.18 s^{-1}).

In the absence of the fluorescent probe it is not possible to calculate the value of the kinetic rate constant of the release step of DXL dissociation (k_{off}). Control experiments with Flutax-2 performed in the same conditions did not show any trNOESY signal from this ligand, which indicates that its k_{off} value is not large enough to provide good trNOESY signals. Therefore, the effective kinetic rate of the release step of the dissociation process of DXL has to be higher. Because it is not unreasonable to think that the presence of the fluorescein moiety slows down the dissociation of Flutax-2 and the observed trNOESY crosspeaks for the DXL:MT ensemble are cancelled out by addition of DDM (a TXL-binding site ligand with a much higher affinity (Buey et al., 2005), it can be assumed that the detected signals arise

from DXL in the last step of dissociation from MT, which may be bound to either the external site or to a modified luminal site.

The basic features of the NMR-derived conformation might be extracted from the trNOESY crosspeaks. Clear NOEs are observed between the t-butyl protons and the 2-OBz protons (Figure 4B), whereas only extremely weak NOEs are observed between both aromatic (2-OBz and 3′-Ph) moieties (Figure 4D). The OAc-4 group also provides NOEs with both aromatic rings (Figure 4B). These experimental observations allow us to discard the presence of the so-called "polar conformation" for DXL when this molecule is bound to MT. The NMR-derived conformation is thus basically in agreement with the conformation derived from the modeling approach (see Figure S4). And though the modeled structure is in agreement with the so-called "T-taxol geometry" (Snyder et al., 2001), the NMR-derived conformation is intermediate between this one and that dubbed collapsed geometry (Vandervelde et al., 1993). That both conformations are fairly similar and resemble the T-taxol conformation possibly indicates that the prerelease step does not largely affect the conformation of DXL, and that the T-taxol conformation is stable in the protein environment. The only observed difference is likely due to the presence of His229, which in the modeled structure is found between the 2-OBz and the C13 side chain, thus further separating these two moieties. The NMR observations are in agreement with a closer proximity between the OBz and the t-Bu protons (~4–5 Å) than that suggested by the modeled DXL-β-tubulin complex (~5–6 Å). Under these constraints, the NMR-deduced prerelease bound geometry for DXL, which is close to that of T-taxol, is in agreement with that derived by the modeling procedure and resembles that described for the tubulin-bound conformation of TXL (Lowe et al., 2001), though the experimental conditions herein are markedly different.

Binding Affinity, Cytotoxicity and P-gp-Overexpression-Mediated Multidrug Resistance

The double-log plots representing cytotoxicity versus tubulin binding affinity (Figure 5A) clearly indicate that, as in the case of EPOs and other taxane-binding site ligands, both magnitudes are related, with the binding affinity behaving as a good predictor of cytotoxicity. However, a deviation from the predicted behavior can be noted from this data. There is an apparent cytotoxicity limit (IC_{50} = 1 nM) for these compounds against the non–P-gp-overexpressing cells. A review of the results from our earlier work (Buey et al., 2004, 2005) indicates that there are no MSA with an IC_{50} below nM in these cells. Despite having binding constants of the order of 10^9 M^{-1}, DDM and several EPOs have IC_{50}s in the order of 1 nM or higher. In fact, cis-CP-tmt-EPO-B (compound 19 in Buey et al. [2004]), the compound with the highest affinity for the TXL-binding site so far described (2.1×10^{10} M^{-1}), and also the most cytotoxic (IC_{50} = 0.1 nM), displays a binding affinity three orders of magnitude above that of TXL but only a 10-fold increase in cytotoxicity (Buey et al., 2004). These data suggest that a significant percentage of tubulin within the cell has to be bound to stop the cell cycle; cytotoxicity is thereby limited by the amount of compound that is needed to achieve this goal.

At the drug concentrations required to stop cell-cycle progression, the percentage of tubulin bound by the ligand is in the range of 2%–20% of the whole available protein. In the drug intake experiments, the total amount of compound available at the concentrations needed to stop the cell cycle (or at the IC_{50}) is around one-third (comparable) of the total amount of tubulin. Although this should be enough, in principle, for binding to a significant percentage of the protein, the results indicate that the amount of ligand available for binding to the sites is effectively much smaller (2%–10%). The reason for this might be that though all the binding sites are inside the small volume occupied by the cells, the drugs have to pass through the cell membranes and reach a threshold intracellular concentration that is opposed by the detoxification pumps. If a significant percentage (say 2%–5%) of cytoplasmic tubulin has to be bound for the taxane to exert its cytotoxicity, and the amount of ligand available for protein binding is a small percentage (2%–10%) of the total 1 nM concentrations—which is already one-third of the total amount of tubulin in the cells—it follows that for the MSA with a taxane way of action the 0.1–1 nM concentration is a limit for its cytotoxicity in cells. The same reasoning can be applied to a systemically distributed drug for which the minimal amount needed to kill the tumor cells is related to the amount of tubulin available for binding, which imposes a practical limit on the lowest dose that can be used.

However, if the goal is not to find a drug with the highest cytotoxicity possible (none of the newly synthesized high-affinity compounds has a remarkably better cytotoxicity on nonresistant cells than TXL and DXL) but rather to find one with the ability to overcome the main problem appearing in patients undergoing treatment with MSA—namely P-gp-mediated resistance—attempts to increase the affinity would seem to be steps in the right direction. Cells overexpressing P-gp are still sensitive to taxanes because they can still be killed by higher concentrations of either TXL or DXL. These very high concentrations affect normal nontumor cells as well, causing them to be differentially killed because of their inability to reduce the intracellular drug concentration, rather than differentially spared because of their lower division rate.

The data with A2780AD cells shows the expected correlation (although with a lower slope arising from their ability to reduce the intracellular drug concentration) between affinity and cytotoxicity that was previously observed for chemically related compounds (Buey et al., 2004) with no deviations being noted at the highest cytotoxicity values (9 nM for CTX-40). In this type of MDR cells the high-affinity drugs are nearly 100-fold more cytotoxic than the clinically employed taxanes (TXL and DXL) and display very low resistance indexes (as low as 1.3 for the highest-affinity derivative, CTX-40). This result was confirmed with LoVo human colon carcinoma cells and their MDR LoVo-Dox counterparts (Grandi et al., 1986; see the Supplemental Data).

When the resistance indexes of the compounds are represented against their binding affinities (Figure 5B), a bell-shaped curve is observed: the resistance index shows a maximum for taxanes displaying similar affinities for tubulin and P-gp, and then rapidly decreases when the affinity of the compound for tubulin either increases or decreases. An exception to this rule is found for compounds having a halogen atom (or a methoxy group) at the meta position of 2-OBz (CTX-5, 6, 30, 34, and 35) as they exhibit a much lower resistance index than that of other compounds with equivalent affinity.

The results confirm our previous data with other high-affinity ligands or with covalent binders (which can be considered to have infinite affinity) whose cytotoxicity is unaffected by P-gp overexpression (e.g., DDM: IC_{50} values of 60 nM and 53 nM [I. Barasoain, unpublished data] or cyclostreptin: IC_{50} values of 43.5 nM and 51 nM [Buey et al., 2007]) against A2780 and A2780AD cells, respectively, which is a clear indication that ligands with high affinity for the taxane-binding site can overcome the P-gp-mediated MDR phenotype. The rationale for this finding is that, in these cells, the intracellular free concentration of the high-affinity binding drugs will be low (see Figures S5–S7 for a detailed mathematical model). It is clear that for the ligand to be pumped out it first has to bind to P-gp, and assuming that the kinetics of drug efflux follows a Michaelis-Menten behavior, the ligand outflow will decrease with lower free ligand concentration. Because these ligands are tightly bound to tubulin, their intracellular free concentrations are of the order of their dissociation constants, which in the case of the high-affinity compounds (K_d of CTX-40 = 0.16 nM at 35°C) are far below their dissociation constants from P-gp (which range between 35 nM for TXL and 88 nM for CTX-7; Yang et al., 2007). This implies that, at concentrations able to exert cytotoxicity, the efficacy of P-gp to pump out the high-affinity compounds will be reduced by a factor between 200 and 1,000 (see the Supplemental Data). From a chemical standpoint, P-gp overexpression is irrelevant.

However, the low-affinity tubulin-binding ligands may escape the effect of the pump through a different mechanism. Because they need to reach concentrations that are much higher than those of either tubulin or P-gp to bind their target and thereby exert their cytotoxicity, the pump gets overloaded (saturated) and cannot effectively reduce the intracellular drug concentration. For this reason, these ligands act as MDR reversal agents by themselves (Brooks et al., 2003).

The present findings support the view (Buey et al., 2004, 2005) that binding affinity is the main variable to be maximized in attempts to increase the cytotoxicity of this type of compound (although on nonresistant cells a practical limit is observed at around 1 nM concentration). Additionally, high-affinity compounds can escape MDR due to P-gp overexpression by lowering the concentration of free ligand that can be pumped out by P-gp.

SIGNIFICANCE

The binding affinities of a series of synthetic taxanes for MTs have been measured with the aims of dissecting individual group contributions and obtaining a rationale for the design of novel compounds with the ability to overcome drug resistance. As previously observed for EPOs, the positive and negative contributions of the different substituents to the binding free energy are cumulative. By combining the most favorable substitutions in a single analog, the binding affinity was increased 500-fold over that of TXL. Insight into the structural basis for this improvement was gained when models were built that assigned an important role to the interactions of C2 and C13 substituents with the protonated side chain of His229. The relative orientation of these groups was found to be in agreement with NMR data obtained for MT-bound DXL.

The cytotoxicities of the compounds in ovarian carcinoma A2780 cells were found to correlate with their affinities, with an apparent cytotoxicity limit in the nanomolar range. A bell-shaped curve was obtained when the taxane resistance index was plotted versus the binding affinity showing that the P-gp-overexpressing multidrug-resistant A2780AD cells are sensitive to the highest and lowest affinity compounds, whereas resistance indexes in the range of 100 to 1,000 were obtained for those whose binding affinities for tubulin and P-gp are similar.

The finding that taxanes with affinities for MTs well above their affinities for P-gp are not affected by multidrug resistance strongly indicates that for a series of compounds with similar pharmacokinetic and bioavailability properties, optimization of the ligand-target interaction is a good strategy to overcome multidrug resistance mediated by efflux pumps.

EXPERIMENTAL PROCEDURES

Proteins and Ligands

Purified calf brain tubulin and chemicals were as described (Diaz and Andreu, 1993). Full details of the synthesis and characterization of the ligands employed can be found in the Supplemental Data.

Ligand-Induced Tubulin Assembly

Critical concentrations of ligand-induced tubulin assembly were measured as described (Buey et al., 2005).

Equilibrium Binding Constants of the Ligands to MTs and Tubulin

The binding constants of the ligands with apparent binding affinities below 10^8 M^{-1} for the TXL-binding site were measured as described (Buey et al., 2004). For the EPO-B method, samples of 1 ml containing 1 μM sites in glutaraldehyde-stabilized MTs, 1.1 μM EPO-B and 1.1 μM of the test compound in GAB (glycerol assembly buffer; 3.4 M glycerol, 10 mM sodium phosphate, 6 mM $MgCl_2$, and 1 mM EGTA [pH 6.7]) with 0.1 mM GTP were incubated for 30 min at 35°C in polycarbonate centrifuge tubes (Beckman Coulter, Inc., Fullerton, CA). The samples were then centrifuged at 90,000 × g for 20 min at 35°C in a TLA-100.2 rotor employing a Beckman Optima TLX ultracentrifuge. The supernatants were collected by pipetting, and the pellets were resuspended in 10 mM phosphate buffer (pH 7.0). One micromolar TXL was added as the internal standard, except for the experiments with CTX-13 in which 1 μM DXL was used instead. Both the pellets and the supernatants were extracted three times with an excess volume of dichloromethane, dried in vacuum, and dissolved in 35 μl of methanol. The samples were analyzed by HPLC.

Reversibility of binding was checked by incubating samples containing 5 μM compounds 4, 12, or 14, and 10 μM taxoid binding sites in stabilized crosslinked MT for 30 min at 25°C in polycarbonate centrifuge tubes (Beckman) in GAB with 0.1 mM GTP (DMSO concentration was always kept under 2%). The samples were then centrifuged at 90,000 g for 10 min at 25°C in a Beckman Optima TLX ultracentrifuge with a TLA100 rotor, processed, and analyzed as described.

Binding constants for compounds reversibly displacing Flutax-2 or EPO-B were calculated using Equigra v5 (Díaz and Buey, 2007). The thermodynamic parameters (apparent ΔG_o, ΔH_o, and ΔS_o) were calculated as described (Buey et al., 2005).

Binding of the compounds to unassembled dimeric tubulin was measured by centrifugation. Two hundred microliter samples containing 20 μM tubulin and 25 μM compound in 10 mM phosphate, 1 mM EDTA (pH 7.0) buffer (PEDTA) containing 1 mM GDP were incubated for 1 hr at 35°C in polycarbonate centrifuge tubes (Beckman). The samples were then centrifuged at 386,000 × g for 1 hr at 35°C in a TLA100 rotor employing a Beckman Optima TLX ultracentrifuge. The upper and lower 100 μl of the solution were carefully collected by pipetting, and the pellets were resuspended in 10 mM phosphate buffer (pH 7.0). The tubulin concentrations in the three samples were

measured by the Bradford assay, and 5 μM DXL was added as the internal standard. The samples were extracted and analyzed as described.

HPLC analysis of all samples was performed in an Agilent 1100 series instrument employing a Supercosil, LC18 DB, 250 × 4.6 mm, 5 mm bead diameter column developed in a gradient from 50% to 80% (v/v) acetonitrile in water at a flow rate of 1 ml/min^{-1}, following the absorbance at λ = 220 nm.

Cell Biology Studies

PtK2 (kidney epitelial cell from *Potorus tridactylis*), U937 (monocytic human leukemia), K562 (myelocytic human leukemia), A2780, P-gp-overexpressing A2780AD (ovarian carcinoma) cells were cultured as described previously (Buey et al., 2005). Cytotoxicity assays were performed with the MTT assay modified as described in Yang et al. (2007).

Cell intake of ^{3}H-TXL and ^{14}C-DXL were measured as reported (Manfredi et al., 1982) with modifications, using PtK2, U937, and K562 cells. These cells, and especially the PtK2 cells, were used because they are more resistant to taxanes, and more reproducible results could be obtained. Cells were grown in 24-well plates at a density of 500,000 cells/ml (PtK2) or 300,000 cells/ml (leukemic cell lines) and were incubated in 1 ml of medium with the desired drug concentration. Supernatants were collected, and cells were washed three times with 1 ml of cold PBS and incubated overnight with 0.25 ml of NaOH 0.1M, and then neutralized with 0.25 ml of HCl 0.1 M. The total protein concentration was determined by the Lowry method and the drug (both in the supernatants and incorporated into cells) was measured by liquid scintillation counting in a LKB 1219 spectrometer (GE Healthcare Bio-Sciences, Uppsala, Sweden). The correction for unspecific binding was determined by measuring the amount incorporated in cells preincubated with 10 μM colchicine for 4 hr at 37°C and washed three times prior to the incubation with ^{3}H-TXL or ^{14}C-DXL. Cell volume was calculated from the volume occupied by the pellets and found to be 2.5 ± 0.2, 3.96 ± 0.05, and 4.3 ± 0.9 μL/10^{6} cells for U937, K562, and PtK2, respectively.

The amounts of drugs bound to the nucleus and cytoplasm of the cells were determined as described in Simpson et al. (1987).

Molecular Modeling

The refined structure of the α,β-tubulin dimer at 3.5 Å resolution (Lowe et al., 2001; Protein Data Bank code: 1JFF) was used for molecular modeling and ligand docking. Addition of missing hydrogen atoms and computation of the protonation state of ionizable groups in β-tubulin at pH 6.5 were carried out using the H^{++} Web server (Gordon et al., 2005), which relies on AMBER (Cornell et al., 1995) force-field parameters and finite difference solutions to the Poisson-Boltzmann equation. The molecular graphics program PyMOL (DeLano Scientific, LLC, Palo Alto, CA) was employed for molecular visualization and representation. The charge distribution for the ligands studied was obtained by fitting the quantum mechanically calculated (RHF 6-31G*\\3-21G*) molecular electrostatic potential (MEP) to a point charge model, as implemented in Gaussian 03 (Gaussian, Inc., Wallingford, CT). Consistent bonded and nonbonded AMBER parameters were assigned to ligand atoms in the taxanes by analogy or through interpolation from those already present in the AMBER database (ff03).

The Lamarckian genetic algorithm implemented in AutoDock 3.0.5 (Morris et al., 1998) was used to generate automated docked poses of TXL, DXL, CTX-13, and CTX-40 within the taxane-binding site by randomly changing the overall orientation of conformers from the MD ensembles that were representative of the major populations, as well as the torsion angle involving the 2-OBz.

NMR Experiments

The NMR experiments were performed at 310 K in D_2O as described (Jimenez-Barbero et al., 2006), with modifications described in the Supplemental Data, on a Bruker AVANCE 500 spectrometer (Bruker, Bruker BioSpin GmbH, Rheinstetten, Germany).

SUPPLEMENTAL DATA

Supplemental Data include seven figures, three tables, Supplemental Experimental Procedures, Supplemental References, and two additional Supplemental Data files and can be found with this article online at http://www.chembiol.com/cgi/content/full/15/6/573/DC1/.

ACKNOWLEDGMENTS

We thank F. Amat-Guerri for Flutax-2, K.C. Nicolaou for EPO-B, the late M. Suffness for TXL, Rhône Poulenc Rorer Aventis for DXL, and José J. Ramírez and Tulsi Pindolia for their help in the early stages of the molecular modeling work. We also thank Carne Sierra Madrid S.A. (CIF: A78074168) for providing the calf brains for tubulin purification. This work was supported in part by grant BIO2007-61336 from M.E.C. to J.F.D., BIPPED-CM from Comunidad de Madrid to F.G., J.F.D., J.J.B., and J.M.A., and grants NSFC 20572135 and MOST 2006DFA31490 to W.S.F.

Received: November 15, 2007
Revised: May 4, 2008
Accepted: May 7, 2008
Published: June 20, 2008

REFERENCES

Brooks, T., Minderman, H., O'Loughlin, K.L., Pera, P., Ojima, I., Baer, M.R., and Bernacki, R.J. (2003). Taxane-based reversal agents modulate drug resistance mediated by P-glycoprotein, multidrug resistance protein, and breast cancer resistance protein. Mol. Cancer Ther. *2*, 1195–1205.

Buey, R.M., Diaz, J.F., Andreu, J.M., O'Brate, A., Giannakakou, P., Nicolaou, K.C., Sasmal, P.K., Ritzen, A., and Namoto, K. (2004). Interaction of epothilone analogs with the paclitaxel binding site: relationship between binding affinity, microtubule stabilization, and cytotoxicity. Chem. Biol. *11*, 225–236.

Buey, R.M., Barasoain, I., Jackson, E., Meyer, A., Giannakakou, P., Paterson, I., Mooberry, S., Andreu, J.M., and Diaz, J.F. (2005). Microtubule interactions with chemically diverse stabilizing agents: thermodynamics of binding to the paclitaxel site predicts cytotoxicity. Chem. Biol. *12*, 1269–1279.

Buey, R.M., Calvo, E., Barasoain, I., Pineda, O., Edler, M.C., Matesanz, R., Cerezo, G., Vanderwal, C.D., Day, B.W., Sorensen, E.J., et al. (2007). Cyclostreptin binds covalently to microtubule pores and lumenal taxoid binding sites. Nat. Chem. Biol. *3*, 117–125.

Chen, S.H., Huang, S., Kant, J., Fairchild, C., Wei, J.M., and Farina, V. (1993a). Synthesis of 7-deoxytaxol and 7,10-dideoxytaxol via radical intermediates. J. Org. Chem. *58*, 5028–5029.

Chen, S.H., Wei, J.M., and Farina, V. (1993b). Taxol structure-activity relationships: synthesis and biological evaluation of 2-deoxytaxol. Tetrahedron Lett. *34*, 3205–3206.

Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., and Kollman, P.A. (1995). A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J. Am. Chem. Soc. *117*, 5179–5197.

Diaz, J.F., and Andreu, J.M. (1993). Assembly of purified GDP-tubulin into microtubules induced by taxol and taxotere: reversibility, ligand stoichiometry, and competition. Biochemistry *32*, 2747–2755.

Díaz, J.F., and Buey, R.M. (2007). Characterizing ligand-microtubule binding by competition methods. Methods Mol. Med. *137*, 245–260.

Diaz, J.F., Strobe, R., Engelborghs, Y., Souto, A.A., and Andreu, J.M. (2000). Molecular recognition of taxol by microtubules. Kinetics and thermodynamics of binding of fluorescent taxol derivatives to an exposed site. J. Biol. Chem. *275*, 26265–26276.

Diaz, J.F., Barasoain, I., and Andreu, J.M. (2003). Fast kinetics of Taxol binding to microtubules. Effects of solution variables and microtubule-associated proteins. J. Biol. Chem. *278*, 8407–8419.

Gao, Q., and Chen, S.H. (1996). An unprecedented side chain conformation of paclitaxel (Taxol$^{(R)}$): crystal structure of 7-mesylpaclitaxel. Tetrahedron Lett. *37*, 3425–3428.

Gordon, J.C., Myers, J.B., Folta, T., Shoja, V., Heath, L.S., and Onufriev, A. (2005). H^{++}: a server for estimating pKas and adding missing hydrogens to macromolecules. Nucleic Acids Res. *33*, W368–371.

Grandi, M., Geroni, C., and Giuliani, F.C. (1986). Isolation and characterization of a human colon adenocarcinoma cell-line resistant to doxorubicin. Br. J. Cancer *54*, 515–518.

Jimenez-Barbero, J., Canales, A., Northcote, P.T., Buey, R.M., Andreu, J.M., and Diaz, J.F. (2006). NMR determination of the bioactive conformation of peloruside a bound to microtubules. J. Am. Chem. Soc. *128*, 8757–8765.

Kingston, D.G., and Newman, D.J. (2007). Taxoids: cancer-fighting compounds from nature. Curr. Opin. Drug Discov. Devel. *10*, 130–144.

Kingston, D.G.I., Chaudhary, A.G., Chordia, M.D., Gharpure, M., Gunatilaka, A.A.L., Higgs, P.I., Rimoldi, J.M., Samala, L., Jagtap, P.G., Giannakakou, P., et al. (1998). Synthesis and biological evaluation of 2-acyl analogues of paclitaxel (Taxol). J. Med. Chem. *41*, 3715–3726.

Lowe, J., Li, H., Downing, K.H., and Nogales, E. (2001). Refined structure of αβ-tubulin at 3.5 A resolution. J. Mol. Biol. *313*, 1045–1057.

Manfredi, J.J., Parness, J., and Horwitz, S.B. (1982). Taxol binds to cellular microtubules. J. Cell Biol. *94*, 688–696.

Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., and Olson, A.J. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J. Comput. Chem. *19*, 1639–1662.

Nicolaou, K.C., Couladouros, E.A., Nantermet, P.G., Renaud, J., Guy, R.K., and Wrasidlo, W. (1994). Synthesis of C-2 Taxol analogs. Angew. Chem. Int. Ed. Engl. *33*, 1581–1583.

Nicolaou, K.C., Renaud, J., Nantermet, P.G., Couladouros, E.A., Guy, R.K., and Wrasidlo, W. (1995). Chemical synthesis and biological evaluation of C-2 taxoids. J. Am. Chem. Soc. *117*, 2409–2420.

Ojima, I., Duclos, O., Zucco, M., Bissery, M.C., Combeau, C., Vrignaud, P., Riou, J.F., and Lavelle, F. (1994). Synthesis and structure-activity-relationships of new antitumor taxoids. Effects of cyclohexyl substitution at the C-3′ and/or C-2 of taxotere (Docetaxel). J. Med. Chem. *37*, 2602–2608.

Ojima, I., Kuduk, S.D., Pera, P., Veith, J.M., and Bernacki, R.J. (1997). Synthesis and structure-activity relationships of nonaromatic taxoids: effects of alkyl and alkenyl ester groups on cytotoxicity. J. Med. Chem. *40*, 279–285.

Rogan, A.M., Hamilton, T.C., Young, R.C., Klecker, R.W., and Ozols, R.F. (1984). Reversal of adriamycin resistance by verapamil in human ovarian cancer. Science *224*, 994–996.

Shabbits, J.A., Krishna, R., and Mayer, L.D. (2001). Molecular and pharmacological strategies to overcome multidrug resistance. Expert Rev. Anticancer Ther. *1*, 585–594.

Simpson, R.U., Hsu, T., Begley, D.A., Mitchell, B.S., and Alizadeh, B.N. (1987). Transcriptional regulation of the C-Myc protooncogene by 1,25-dihydroxyvitamin-D3 in Hl-60 promyelocytic leukemia cells. J. Biol. Chem. *262*, 4104–4108.

Snyder, J.P., Nettles, J.H., Cornett, B., Downing, K.H., and Nogales, E. (2001). The binding conformation of Taxol in β-tubulin: a model based on electron crystallographic density. Proc. Natl. Acad. Sci. USA *98*, 5312–5316.

Tan, B., Piwnica-Worms, D., and Ratner, L. (2000). Multidrug resistance transporters and modulation. Curr. Opin. Oncol. *12*, 450–458.

Vandervelde, D.G., Georg, G.I., Grunewald, G.L., Gunn, G.W., and Mitscher, L.A. (1993). "Hydrophobic collapse" of taxol and taxotere solution conformations in mixtures of water and organic solvent. J. Am. Chem. Soc. *115*, 11650–11651.

Wang, L., Alcaraz, A.A., Matesanz, R., Yang, C.G., Barasoain, I., Díaz, J.F., Li, Y.Z., Snyder, J.P., and Fang, W.S. (2007). Synthesys, biological evaluation, and tubulin binding poses of C-2a sulfur linked taxol analogues. Bioorg. Med. Chem. Lett. *17*, 3191–3194.

Yang, C.G., Barasoain, I., Li, X., Matesanz, R., Liu, R., Sharom, F.J., Diaz, J.F., and Fang, W. (2007). Overcoming tumor drug resistance mediated by P-glycoprotein overexpression with high affinity taxanes: a SAR study of C-2 modified 7-acyl-10-deacetil cephalomannines. ChemMedChem *2*, 691–701.

Zefirova, O.N., Nurieva, E.V., Ryzhov, A.N., Zyk, N.V., and Zefirov, N.S. (2005). Taxol: synthesis, bioactive conformations, and structure-activity relationships in its analogs. Russ. J. Org. Chem. *41*, 315–351.

Chemistry & Biology *15*

Supplemental Data

Optimization of Taxane Binding to Microtubules:

Binding Affinity Dissection and Incremental

Construction of a High-Affinity Analog of Paclitaxel

Ruth Matesanz, Isabel Barasoain, Chun-Gang Yang, Lei Wang, Xuan Li, Concepción de Inés, Claire Coderch, Federico Gago, Jesús Jiménez Barbero, José Manuel Andreu, Wei-Shuo Fang, and José Fernando Díaz

Table S1. Critical Concentrations of the Ligand-Induced Assembly Reactions

Compound	Cr (μM) in PEDTA4-GTP[a]	Cr (μM) in GAB[b]
DMSO	N.A.	3.30±0.38[d]
PTX	4.24±0.49[c]	0.63±0.06
DCX	2.35±0.25	0.62±0.08
CPH	3.50±0.94	0.64±0.15
CTX 1	3.68±0.81	0.67±0.04
CTX 2	9.73±3.94	1.41±0.06
CTX 3	11.38±2.83	2.60±0.19
CTX 4	2.87±0.36	0.50±0.10
CTX 5	2.81±0.08	0.95±0.10
CTX 6	3.07±0.26	0.76±0.06
CTX 7	3.81±0.23	1.12±0.18
CTX 8	3.92±0.19	1.37±0.06
CTX 9	N.A.	2.18±0.17
CTX 10	N.A.	1.89±0.19
CTX 11	3.52±1.52	0.72±0.04
CTX 12	3.78±2.05	0.74±0.14
CTX 13	4.03±1.15	0.48±0.13
CTX 14	3.03±1.15	0.45±0.05
CTX 15	4.27±0.30	0.48±0.04
CTX 16	N.A.	4.00±0.70
CTX 17	3.26±1.85	0.57±0.09
CTX 18	3.96±2.18	0.47±0.07
CTX 19	3.45±2.07	0.45±0.07
CTX 20	3.38±1.62	0.44±0.06
CTX 21	2.95±1.04	0.78±0.15
CTX 22	N.A.	2.36±0.09
CTX 23	N.A.	2.99±1.07

CTX 24	N.A.	1.92±0.40
CTX 25	N.A.	3.15±0.73
CTX 26	N.A.	2.51±0.04
CTX 27	N.A.	3.03±0.10
CTX 28	5.78±1.18	0.80±0.17
CTX 29	6.38±3.11	0.88±0.17
CTX 30	5.98±1.01	1.19±0.31
CTX 31	6.57±0.63	1.04±0.44
CTX 32	6.27±2.59	1.58±0.38
CTX 33	4.32±0.49	0.29±0.15
CTX 34	3.89±1.16	0.69±0.28
CTX 35	6.35±0.81	1.11±0.65
CTX 36	7.88±1.50	1.22±0.08
CTX 37	5.00±1.10	0.94±0.30
CTX 38	N.A.	2.41±0.21
CTX 39	N.A.	2.81±0.34
CTX 40	3.62± 1.1	0.93±0.23
CTX 41	9.22±2.42	1.50±0.42
CTX 42	13.26±2.83	2.26±0.21
CTX 43	12.84±2.29	2.33±0.43
CTX 44	12.59±2.57	2.30±0.99

N.A. No assembly detected.

a) PEDTA4-GTP; PEDTA with 4 mM $MgCl_2$ and 1 mM GTP

b) GAB; 3.4 M Glycerol, 10 mM phosphate, 1 mM EGTA, 6 mM $MgCl_2$, pH 6.5 buffer.

c) Data from (Buey et al. 2004)

d) Data from (Buey et al. 2005)

Errors bars are standard errors of the mean

Table S2. Apparent Binding Affinities of Taxoids for the Paclitaxel Binding Site (10^7 M^{-1})

Compounds measured with the Flutax-2 displacing method

Compound	26°C	27°C	30°C	32°C	35°C	37°C	40°C	42°C
TXL[a]	2.64±0.17	2.19±0.05	1.83±0.09	1.81±0.21	1.43±0.17	1.07±0.11	0.96±0.14	0.94±0.23
DXL[a]	6.95±0.42	6.57±0.52	5.42±0.42	4.89±0.38	3.93±0.27	3.09±0.22	2.89±0.17	2.38±0.11
CPH	N.D.	1.19±0.23	1.01±0.18	0.74±0.13	0.69±0.08	0.66±0.10	0.56±0.08	0.58±0.03
CTX-1	0.8±0.23	0.91±0.26	0.37±0.07	0.57±0.07	0.49±0.12	0.29±0.09	0.28±0.02	0.27±0.03
CTX-2	0.063±0.007	0.061±0.016	0.030±0.002	0.070±0.027	0.043±0.018	0.019±0.011	0.018±0.006	0.018±0.003
CTX-3	N.D.	0.106±0.028	0.091±0.026	0.094±0.021	0.072±0.017	0.052±0.011	0.062±0.018	0.065±0.02
CTX-5	N.D.	7.57±1.70	7.09±1.72	5.7±1.48	5.37±1.39	5.06±1.21	3.68±1.14	3.77±0.97
CTX-6	N.D.	2.53±0.19	2.23±0.12	2.07±0.04	1.62±0.24	1.55±0.27	1.40±0.10	1.21±0.07
CTX-7	N.D.	0.68±0.09	0.56±0.08	0.58±0.05	0.39±0.06	0.39±0.54	0.36±0.12	0.34±0.11
CTX-8	0.732±0.011	0.770±0.155	0.629±0.281	0.595±0.092	0.492±0.073	0.470±0.197	0.467±0.066	0.441±0.114
CTX-9	0.032±0.008	0.029±0.006	0.024±0.006	0.028±0.008	0.028±0.008	0.024±0.014	0.023±0.007	0.022±0.007
CTX-10	0.048±0.004	0.040±0.005	0.034±0.006	0.038±0.009	0.042±0.008	0.033±0.010	0.036±0.001	0.032±0.001
CTX-15	2.259±1.34	2.206±0.89	3.124±1.52	1.623±0.73	2.384±0.53	2.137±0.56	1.504±0.32	1.686±0.10
CTX-16	0.044±0.007	0.043±0.006	0.057±0.030	0.051±0.012	0.050±0.030	0.043±0.024	0.029±0.008	0.035±0.025
CTX-17	1.200±0.50	1.151±0.48	1.131±0.09	1.085±0.52	0.902±0.37	0.863±0.16	0.723±0.22	0.802±0.35
CTX-18	1.401±0.46	1.430±0.44	1.326±0.11	1.436±0.59	1.281±0.27	1.115±0.26	1.050±0.33	0.863±0.20
CTX-22	0.002±0.00	0.021±0.00	0.017±0.01	0.019±0.01	0.013±0.00	0.009±0.00	0.010±0.00	0.009±0.00
CTX-23	0.009±0.00	0.010±0.00	0.008±0.00	0.010±0.00	0.007±0.00	0.005±0.00	0.006±0.00	0.005±0.00
CTX-24	0.291±0.07	0.235±0.08	0.129±0.01	0.114±0.01	0.094±0.01	0.081±0.03	0.090±0.02	0.064±0.01
CTX-25	0.014±0.00	0.011±0.00	0.010±0.01	0.011±0.01	0.008±0.00	0.006±0.00	0.007±0.00	0.004±0.00
CTX-26	0.048±0.02	0.049±0.02	0.020±0.00	0.020±0.00	0.018±0.00	0.012±0.00	0.011±0.00	0.012±0.00
CTX-27	0.011±0.00	0.012±0.00	0.011±0.00	0.009±0.00	0.007±0.00	0.005±0.00	0.007±0.00	0.005±0.00
CTX-28	0.21±0.03	0.21±0.03	0.15±0.01	0.23±0.08	0.17±0.04	0.14±0.01	0.15±0.08	0.15±0.004
CTX-29	0.33±0.02	0.40±0.06	0.39±0	0.37±0.18	0.25±0.08	0.17±0.05	0.16±0.08	0.15±0.07

CTX-30	1.54±0.62	1.15±0.20	2.60±1.47	2.14±1.08	1.76±0.91	1.48±0.79	1.14±0.39	1.24±0.18
CTX-31	0.12±0.04	0.13±0.02	0.10±0.02	0.15±0.05	0.10±0.04	0.06±0.01	0.06±0.03	0.06±0.001
CTX-32	0.13±0.04	0.19±0.01	0.26±0.10	0.27±0.08	0.24±0.08	0.24±0.17	0.12±0.07	0.15±0.08
CTX-33	0.16±0.07	0.23±0.06	0.19±0.04	0.14±0.06	0.07±0.01	0.07±0.0001	0.06±0.02	0.04±0.0004
CTX-34	0.97±0.18	1.04±0.17	3.44±3.64	2.46±2.29	1.20±0.93	1.03±0.62	0.62±0.28	0.53±0.12
CTX 35	0.62±0.44	0.66±0.54	3.64±5.38	1.54±1.75	0.88±0.77	0.67±0.28	0.47±0.12	0.50±0.11
CTX 36	0.045±0.016	0.046±0.013	0.045±0.018	0.035±0.023	0.029±0.020	0.022±0.009	0.024±0.002	0.017±0.006
CTX 37	0.11±0.03	0.10±0.03	0.077±0.029	0.049±0.018	0.035±0.010	0.023±0.009	0.018±0.001	0.019±0.019
CTX 38	0.002±0.002	0.002±0.002	0.001±0.001	0.001±0.001	0.001±0.001	0.001±0.000	0.001±0.000	0.001±0.000
CTX 39	0.019±0.004	0.017±0.004	0.011±0.004	0.007±0.001	0.003±0.001	0.001±0.000	0.001±0.000	0.001±0.000
CTX 41	0.068±0.043	0.069±0.044	0.049±0.043	0.029±0.014	0.021±0.004	0.014±0.005	0.013±0.010	0.013±0.010
CTX 42	0.026±0.003	0.025±0.014	0.021±0.015	0.010±0.003	0.008±0.003	0.004±0.002	0.003±0.002	0.004±0.005
CTX 43	0.082±0.013	0.084±0.028	0.068±0.037	0.036±0.013	0.030±0.008	0.012±0.005	0.015±0.015	0.026±0.037
CTX 44	0.011±0.017	0.008±0.012	0.005±0.006	0.001±0.001	0.001±0.001	0.0004±0.0004	0.0003±0.0003	0.0004±0.0005

Compounds measured with the EPO-B displacing method

	26ºC	**35ºC Flutax-2**	**35 ºC**	**40ºC**
CTX 4	124±39	N.D.	87±19	52±27
CTX 11	34±19	33±24	38±10	25±2
CTX 12	239±130	N.D	151±3	147±43
CTX 13	23±6	8±3	17±3	14±2
CTX 14	126±50	N.D	80±3	55±6
CTX 19	27±2	13±5	14.8±0.2	12±2
CTX 20	310±27	N.D	81±5	63±20
CTX 21	21±4	11±4	14±2	13±2
CTX 40	612±228	N.D	628±15	355±56

a) data from (Buey et al. 2004).

Errors bars are standard errors of the mean

Table S3. Binding of ^{14}C-DXL and ^{3}H TXL to U937 and K562 Leukaemic Cells

Cell line	Drug[a]	Intracellular drug (pmol/10^6 cells)	Intracellular drug/cell tubulin (%)	Intracellular drug concentration. (nM)	Percentage of intracellular drug
U937	DXL 5 nM	1.3±0.1 (0.05)[b]	2.5±0.2	520	7.8
	TXL 10 nM	5.2±0.4(0.13)	9.5±0.7	1596	15.6
K562	DXL 10 nM	3.2±0.1 (0.15)	5.5±0.2	818	9.7
	TXL 25 nM	11.4±0.1 (0.86)	19.5±2.0	2882	13.7
Ptk2	DXL 250 nM	1.3±0.1 (1.8)	1.2±0.1	307	0.5
	TXL 1 μM	1.5±0.1 (5.8)	0.8±0.1	351	0.15

[a]Cells were incubated for 24 hours (Ptk2) or 17 h (leukemic cells) at drug concentrations that produce 90 % arrest in G_2/M.
[b]Non-specific binding.

Errors bars are standard deviations.

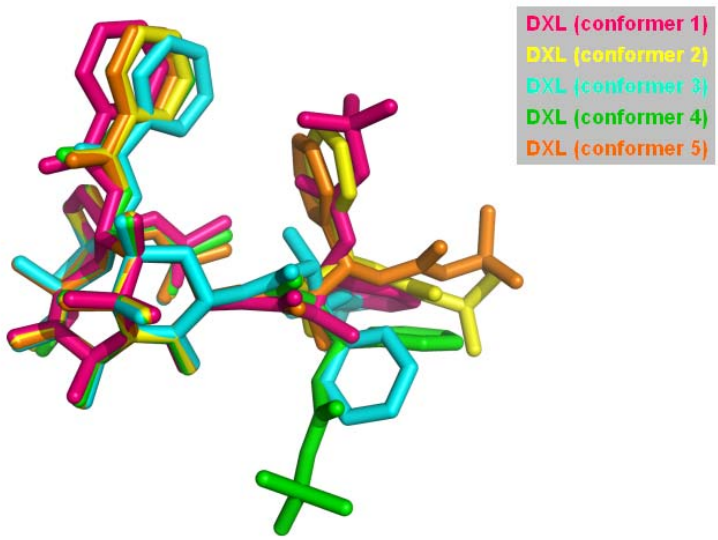


Figure S1. Best-fit Overlap (Using the Baccatin Core Atoms for the Superimposition) of the Five Major Conformers Found for DXL during the MD Simulation in Water and Then Used for Docking Purposes

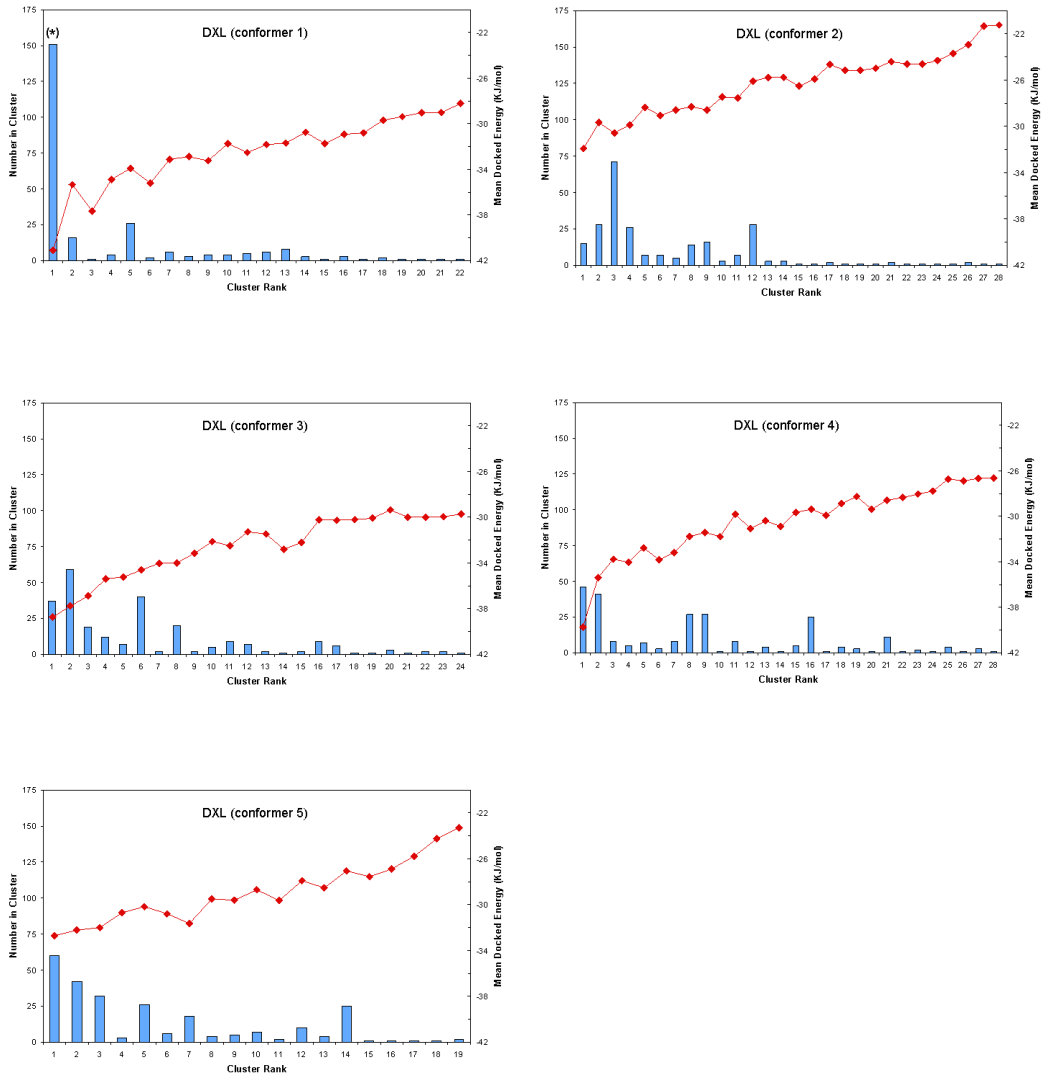


Figure S2. Docking of the Five Major DXL Conformers Obtained from the MD Simulations (Figure S4) into the Taxane-Binding Site of β-Tubulin as Present in PDB Structure 1JFF and with His229 Protonated on Nε

The 250 resulting structures were clustered according to an rmsd criterion of 1.5 Å. Bars represent the number of structures included in each cluster whereas diamonds stand for the mean docking energy of each cluster. The asterisk highlights the solution closest to the TXL conformation found in this PDB entry (Lowe et al. 2001), which happens to be both the most populated and that giving rise to the most favourable binding energy.

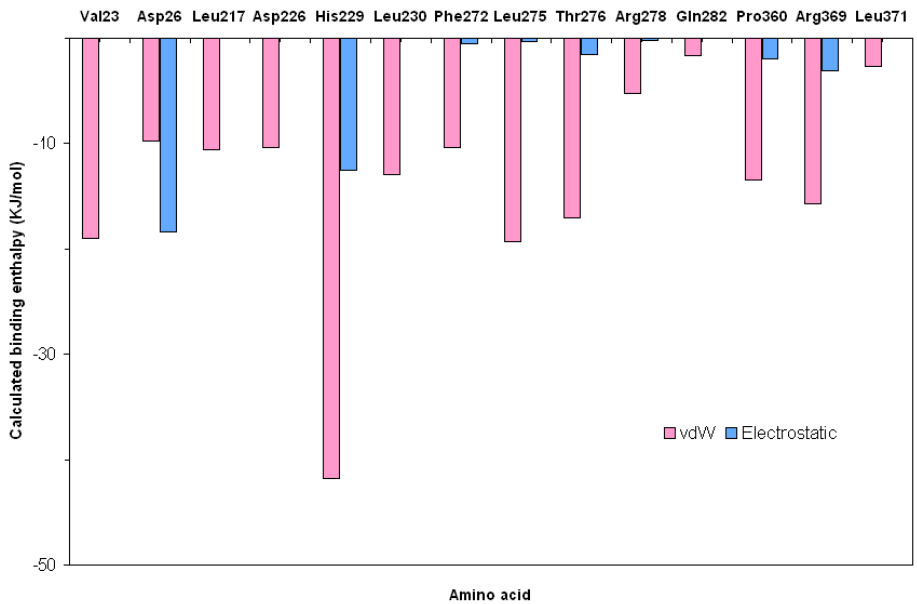


Figure S3. Decomposition of the Calculated Binding Enthalpy for the β-Tubulin–CTX40 Complex upon Completion of the Simulated Annealing Protocol
The intermolecular van der Waals term (pink) was calculated with the ANAL module in AMBER whereas the solvent-corrected electrostatic interaction (cyan) was computed using the DelPhi program, as described in (Perez et al. 1998).

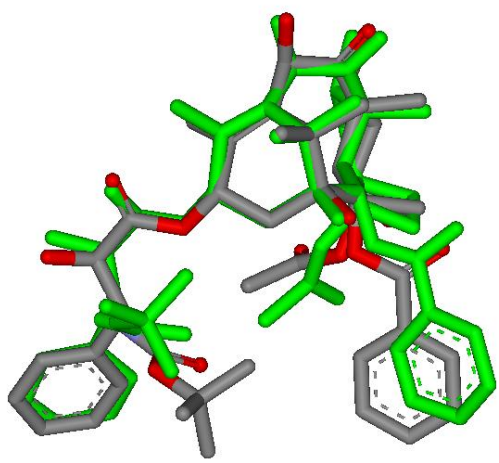


Figure S4. Best-Fit Superimposition of the T-Taxol-like Bound Structures of Docetaxel as Obtained upon Completion of the Simulated Annealing Protocol (Green) and the Independent NMR-Based Procedure (Carbon Atoms in grey).

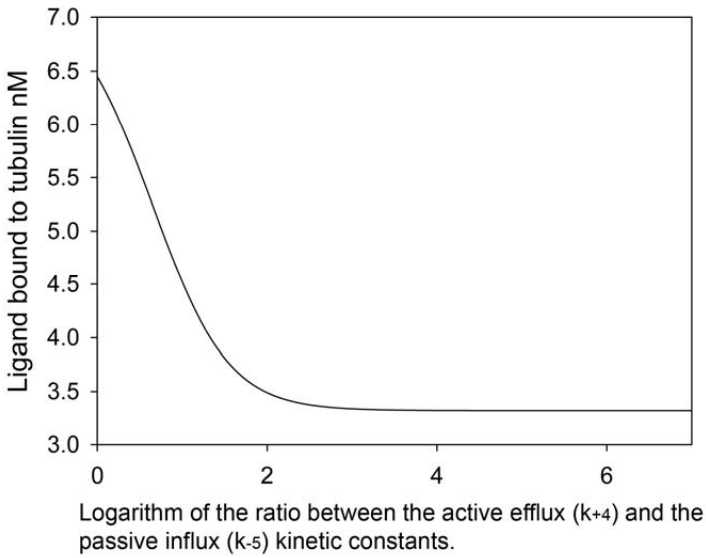


Figure S5. Dependence of the Ligand Bound to Microtubules on the Ratio between the Passive Ligand Influx/Efflux and the Active Kinetic Rate of Pumping

Concentrations and kinetic constants employed for the model are: Ligand 1 μM, Tubulin 20 μM, P-gp 4 μM, K_{+1} 10^7 M^{-1}, K_{+2} 10^7 M^{-1}.

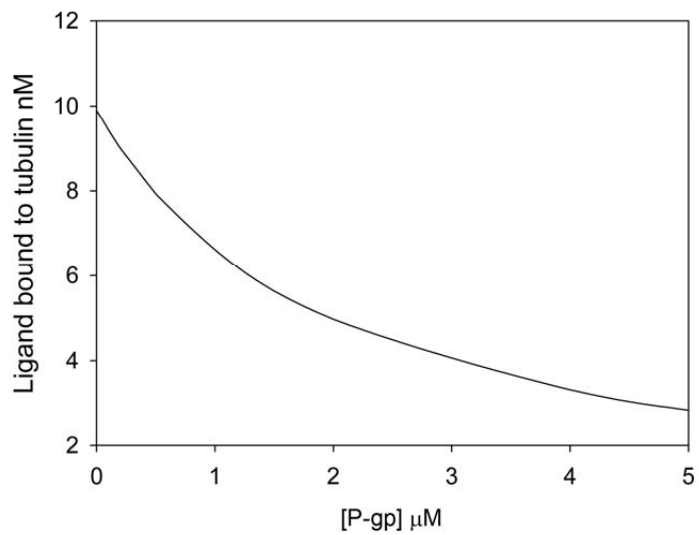


Figure S6. Dependence of the Ligand Bound to Microtubules on P-gp Concentration

Concentrations and kinetic constants employed for the model are: Ligand 1 μM, Tubulin 20 μM, K_{+1} 10^7 M^{-1}, K_{+2} 10^7 M^{-1}, k_{+3} and k_{+4} 10^6 s^{-1}.

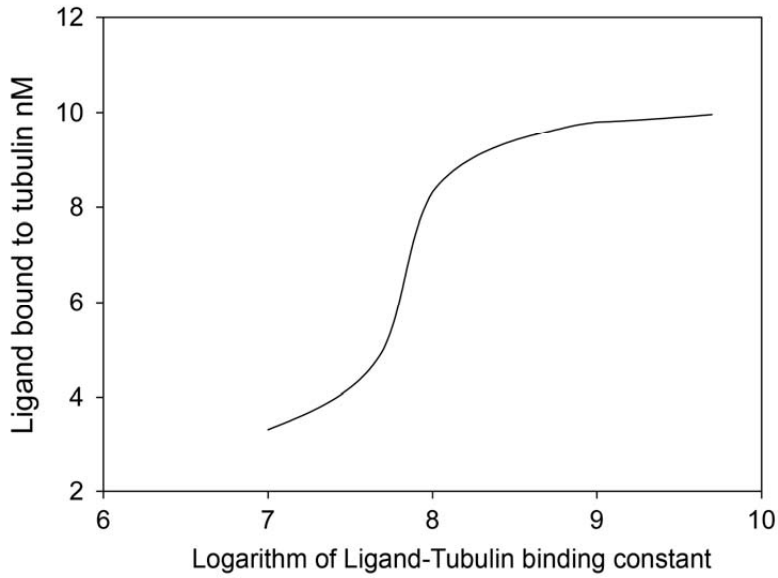


Figure S7. Dependence of the Concentration of Ligand Bound to Microtubules on Its Binding Constant to Tubulin (K_{+1})
Concentrations and kinetic constants employed for the model are Ligand 1 μM, Tubulin 20 μM, P-gp 4 μM, K_{+2} 10^7 M^{-1}, k_{+3} and k_{+4} 10^6 s^{-1}

Supplemental Results

Ligand-Induced Assembly of Tubulin

The compounds were first checked for their ability to induce tubulin assembly under conditions in which tubulin itself is not able to assemble, i.e. purified GTP-tubulin in a buffer in which tubulin is unable to self-assemble in the absence of ligand (Diaz et al. 1993). The critical concentration (Cr) of the assembly reaction was measured for all the ligands (Table S1). Compounds 9, 10, 16, 22, 23, 24, 25, 26, 27, 38 and 39 were inactive in these tests (i.e. did not induce tubulin assembly) while the remaining molecules, as well as the three reference taxanes, induced tubulin assembly. We then checked whether the inactive compounds were at least able to further stabilize microtubules in conditions under which they are able to assemble, as is the case for low-affinity taxane mimetics (Buey et al. 2005). These molecules were indeed able to

reduce the Cr in GAB buffer (Table S1), which indicates that they can be considered low-affinity microtubule-stabilizing agents (MSA).

The compounds were then evaluated for their dependence on Mg^{+2} to induce tubulin assembly as well as for their ability to bind dimeric tubulin. As was the case for TXL and its mimetics (Diaz et al. 1993) (Buey et al. 2005, Buey et al. 2004) all of the compounds required Mg^{+2} at mM concentrations in order to induce tubulin assembly and were unable to bind dimeric tubulin at 25 μM concentrations, indicating that all of them are TXL-like microtubule stabilizing agents.

Cytotoxicity of CTX-40 on LoVo Cells

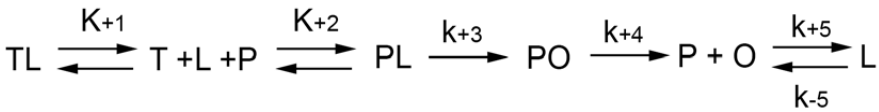
The ability of the high-affinity compound CTX-40 to overcome P-gp–mediated MDR was confirmed with LoVo human colon carcinoma cells and their MDR LoVo-Dox counterparts (Grandi et al. 1986) employing TXL, DXL and CTX-40 (TXL R/S 13.2 IC_{50LoVo}=110 nM, $IC50_{LoVo\text{-}Dox}$=1450 nM, DXL R/S 7.6 IC_{50LoVo}=60 nM, $IC50_{LoVo\text{-}Dox}$=455 nM, CTX-40 R/S 0.86 IC_{50LoVo}=4.1 nM, $IC50_{LoVo\text{-}Dox}$=3.6 nM). As observed for the ovarian carcinoma cells, the high-affinity taxane is able to overcome the P-gp–mediated MDR phenotype. However, LoVo cells require higher concentrations of the medium-affinity clinically used taxanes, TXL and DXL, than A2780 cells to die, while similar concentrations of CTX-40 ($IC50_{A2780}$=7 nm) are required for both type of cells. This result may be explained if the level of P-gp expression in LoVo cells is higher than in A2780 cells. The IC_{50}s of TXL, DXL and CTX-40 on LoVo-Dox cells are similar to those found in A2780AD cells, which display resistance to TXL and DXL but not to CTX-40.

Supplemental Discussion

Kinetics of Drug Pumping from the Cells

In order to test the influence of the different parameters controlling the influx and efflux of the ligands in the amount of bound tubulin, a simplified model of taxane binding to tubulin and P-gp was used.

Schematically, it can be described with the following equations:



where T is tubulin, L, ligand inside the cell, P, P-gp, and O, the ligand outside the cell. K_{+1} and K_{+2} the binding affinities for tubulin and P-gp, k_{+3} and k_{+4} the kinetic rate constants of the ligand being pumped out, and k_{+5} and k_{-5} the kinetic rates of passive ligand influx and efflux. The system was modelled using kinsim in order to study the effects of changes in relative binding affinities between taxanes and microtubules or P-gp.

In this simplified system, the tubulin concentration was considered to be 20 μM, and the passive influx and efflux kinetic rates (k_{+5} and k_{-5}) were set to 0.1 s^{-1} (compatible with the equilibration time determined for PtK2 cells) Given a certain arbitrary concentration of P-gp (4 μM) it is obvious that it is necessary for the facilitated rates of efflux (k_{+4}) to be much higher than the passive ones to have any effect. However, from a certain ratio between the kinetic constants (larger than 1000:1) no effect on the amount of tubulin bound to the ligand is observed (Figure S5)

The effect of P-gp overexpression is shown in Figure S6. In this simple model the amount of liganded tubulin rapidly decreases when P-gp is overexpressed.

However, as expected, the increase in the binding affinity for tubulin

compensated for the effect of P-gp overexpression (Figure S7). Increasing the binding affinity of the ligand for the target decreases the free ligand concentration accordingly, and thus, since the kinetic rate of ligand efflux is dependent on the concentration of free ligand, the ligand efflux. While a 4 μM theoretical P-gp concentration is able to reduce the amount of liganded tubulin by a factor of 4, an increase of the tubulin binding constant compensates for the effect thus restoring the levels of bound tubulin.

Supplemental Experimental Procedures.

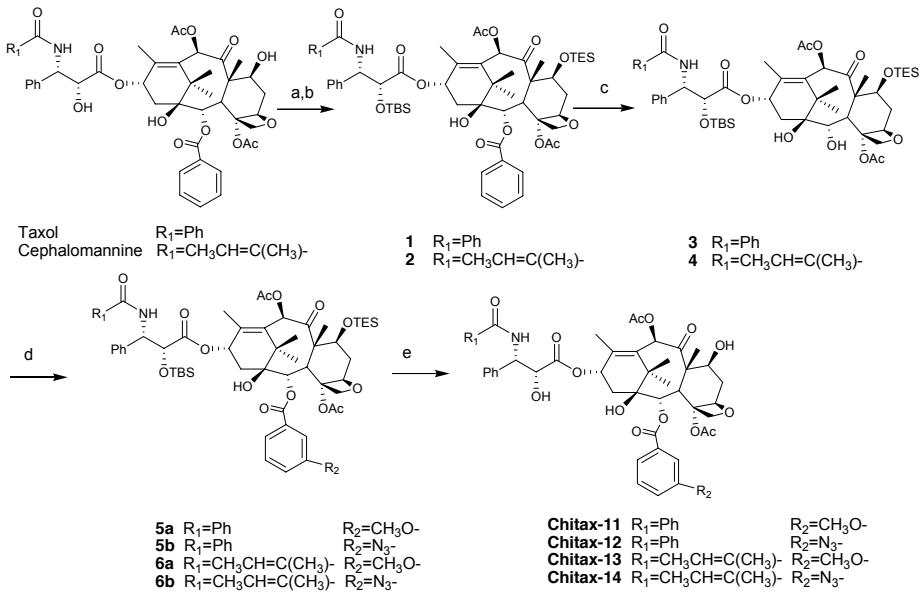
Synthesis and Characterization of the Ligands

TXL analogues carrying the CPH side chain (Figure 1) were synthesized through selective C2 debenzoylation and esterification using CPH as the starting material. The other TXL analogues studied were prepared by coupling of the C13 side chains of TXL and DXL with corresponding C2 O-, N- or S-substituted baccatins, which were obtained by multi-step reactions from 10-deacetyl baccatin (Yang et al. 2007) (See below for full details). Their purities were checked by HPLC in a C-18 column (Supercosil, LC18-DB, 250x4.6 mm, 5 μm bead size) developed with a gradient from 50% to 80% of acetonitrile in water, and were found to be greater than 98%, except for CTX-3, CTX-9, CTX-19, CTX-21, CTX-26 and CTX-27 whose purity is around 90 %. The solubility of these compounds in GAB (3.4 M Glycerol, 10 mM sodium phosphate, 6 mM $MgCl_2$, 1 mM EGTA, 0.1 mM GTP, pH 6.5 buffer) plus 0.1 mM GTP was checked by centrifugation and HPLC and was found to be greater than 50 μM.. TXL was provided by the late Dr. M. Suffness from the National Cancer Institute (Bethesda, MD). DXL was kindly provided by Rhône Poulenc Rorer, Aventis (Schiltigheim, France). Flutax-2 was provided by Dr. F. Amat-Guerri from Instituto de Química Orgánica, CSIC, (Madrid, Spain). Epothilone B (epoB)

was provided by K.C. Nicolaou from the Department of Chemistry and Biochemistry, University of California at San Diego and the Scripps Research Institute (La Jolla, CA). All the compounds were diluted in DMSO to a final concentration of 10 mM and stored at –70 ºC.

^{3}H-TXL was a gift from Drs. I. Ringel, The Hebrew University (Jerusalem) & S. B. Horwitz, Albert Einstein College of Medicine (Bronx, New York). ^{14}C-DXL (RP 71626) were kindly provided by Dr. J.-L. Fabre, Rhône-Poulenc Rorer (92165 Antony, France).

General Synthetic Procedures. All chemicals other than anhydrous solvents were obtained from Aldrich and Acros Chemical Co. and used without further purification. All anhydrous reactions were performed under N_2, and anhydrous tetrahydrofuran (THF) was dried over sodium (benzophenone as indicator). All reactions were monitored by thin layer chromatography (TLC) (silica gel, GF254) with UV light and H_2SO_4-anisaldehyde spray visualization.



Scheme 1 Reagents and conditions: a: TBSCl, imidazole, DMF; b: TESCl, imidazole, DMF; c: Triton B, CH_2Cl_2, -28°C d: DCC, pp, *m*-R_2-PhCOOH; e: 40% HF(v/v), pyridine, CH_3CN.

2'-*O*-(tert-butyldimethylsilyl)-7-*O*-triethylsilyltaxol (**1**)

Taxol (200mg, 0.234mmol) was dissolved in DMF (1.7 ml), then imidazole (84mg, 1.23mmol) and tert-butyldimethylsilyl chloride (TBSCl) (185.4mg, 1.23mmol) were added to the solution, the mixture was reacted for 5 h at room temperature. Additional imidazole (84mg, 1.23mmol) and triethylsilyl chloride (TESCl) (138μl) were added, the mixture was reacted 5 h at room temperature again. Then 100 ml saturated aq $NaHCO_3$ solution was added, the aqueous layer was extracted with EtOAc (3×150ml), the organic layer was combined, and dried over anhydrous Na_2SO_4, then the organic layer was filtered, the filtrate was concentrated to dryness in vacuo. The residue was purified on silica gel with column chromatography (gradient eluted with Petroleum ether/acetone = 9/1~4:1) to afford white solid 2'-*O*-(tert-butyldimethylsilyl)-7-*O*-triethylsilyltaxol (**1**) (229mg, 90.3%).

^{1}H NMR (300MHz, $CDCl_3$): δ 8.14 (2H, dd, *J* = 8.4, 1.5 Hz, *o*-BzO), 7.74 (2H, dd, *J* = 8.4, 1.5 Hz, *o*-BzNH), 7.31-7.63 (11H, m, BzO + BzNH + Ph), 7.07 (1H,

d, *J* = 9.0 Hz, BzNH), 6.45 (1H, s, H-10), 6.26 (1H, t, *J* = 9.3 Hz, H-13), 5.74 (1H, dd, *J* = 2.7, 8.7 Hz, H-3'), 5.70(1H, d, *J* = 7.5 Hz, H-2), 4.96 (1H, d, *J* = 7.8 Hz, H-5), 4.67 (1H, d, *J* = 2.1 Hz, H-2'), 4.48 (1H, dd, *J* = 6.6, 10.5 Hz, H-7), 4.33 (1H, d, *J* = 8.1 Hz, H-20), 4.21 (1H, d, *J* = 8.1 Hz, H-20), 3.84 (1H, d, *J* = 6.9 Hz, H-3), 2.58 (3H, s, OAc-4), 2.48-2.55 (1H, m, overlap, H-6), 2.40 (1H, dd, *J* = 9.6, 12.7 Hz, H-14), 2.18 (3H, s, OAc-10), 2.09 (1H, dd, *J* = 9.0, 15.0 Hz, H-14), 2.02 (3H, s, Me-18), 1.91 (1H, m, H-6), 1.70 (3H, s, Me-19), 1.22 (3H, s, Me-17), 1.18 (3H, s, Me-16), 0.93 (9H, t, *J* = 7.5 Hz, $SiCH_2CH_3$), 0.80 (9H, s, $SiC(CH_3)_3$), 0.54-0.65 (6H, m, $SiCH_2CH_3$), -0.02 (3H, s, $SiCH_3$), -0.30 (3H, s, $SiCH_3$).

2'-*O*-(tert-butyldimethylsilyl)-7-*O*-triethylsilylcephalomannine (**2**)

Cephalomannine (600mg, 0.721mmol) was dissolved in DMF (5 ml), then imidazole (245.5mg, 3.61mmol) and tert-butyldimethylsilyl chloride (543.5mg, 3.61 mmol) were added to the solution, the mixture was reacted for 5 h at room temperature. Additional imidazole (245.5mg, 3.61mmol) and triethylsilyl chloride (400μl) were added, the mixture was reacted 5 h at room temperature again. Then 100 ml saturated aq $NaHCO_3$ solution was added, the aqueous layer was extracted with EtOAc (3×150ml), the organic layer was combined, and dried over anhydrous Na_2SO_4, then the organic layer was filtered, the filtrate was concentrated in vacuo. The residue was purified on silica gel with column chromatography (gradient elute with Petroleum ether/acetone = 9/1~4:1) to afford white solid 2'-*O*-(tert-butyldimethylsilyl)-7-*O*-triethyl-silylcephalomannine (**2**) (760mg, 99.3%).

^{1}H NMR (400MHz, $CDCl_3$): δ 8.12 (2H, d, *J* = 6.9 Hz, *o*-Bz), 7.59 (1H, t, *J* = 7.2 Hz, *p*-Bz), 7.50 (2H, t, *J* = 7.8 Hz, *m*-Bz), 7.34 (2H, t, *J* = 7.6 Hz, Ph), 7.25-7.28 (3H, m, Ph), 6.65 (1H, d, *J* = 8.8 Hz, CONH), 6.45 (1H, s, H-10), 6.45 (1H, overlapped, CH_3CH=), 6.25 (1H, t, *J* = 8.8 Hz, H-13), 5.70 (1H, d, *J* = 6.8

Hz, H-2), 5.58 (1H, d, *J* = 8.8 Hz, H-3'), 4.94 (1H, d, *J* = 9.2 Hz, H-5), 4.59 (1H, d, *J* = 1.6 Hz, H-2'), 4.47 (1H, dd, *J* = 6.4, 10.6 Hz, H-7), 4.30 (1H, d, *J* = 8.4 Hz, H-20), 4.20 (1H, d, *J* = 8.4 Hz, H-20), 3.83 (1H, d, *J* = 7.2 Hz, H-3), 2.54 (3H, s, OAc-4), 2.51 (1H, m, H-6), 2.37 (1H, dd, *J* = 9.2, 15.4 Hz, H-14), 2.17 (3H, s, OAc-10), 2.12 (1H, dd, *J* = 8.8, 15.2 Hz, H-14), 2.04 (3H, s, =CC$\underline{H}_3$CO-), 1.91 (1H, m, H-6), 1.81 (3H, s, Me-18), 1.71 (3H, d, *J* = 8.8 Hz, $CH_3C\underline{H}$=), 1.70 (3H, s, Me-19), 1.23 (3H, s, Me-17), 1.22 (3H, s, Me-16), 0.93 (9H, t, *J* = 8.0 Hz, $SiCH_2C\underline{H}_3$), 0.78 (9H, s, $SiC(C\underline{H}_3)_3$), 0.54-0.61 (6H, m, $SiC\underline{H}_2CH_3$), -0.05 (3H, s, $SiC\underline{H}_3$), -0.33 (3H, s, $SiC\underline{H}_3$). ESI-MS: *m/z* $[M+Na]^+$ 1082.5, $[M+K]^+$ 1098.5.

2'-*O*-(tert-butyldimethylsilyl)-2-debenzoyl-7-*O*-triethylsilyltaxol (**3**)

2'-*O*-(tert-butyldimethylsilyl)-7-*O*-triethylsilyltaxol (**1**) (108mg, 0.1mmol) was dissolved in dried CH_2Cl_2 (2 ml), the obtained solution was cooled to -28°C in ice-methanol bath, then the benzyltrimethylammonium hydroxide (Triton B) (93μl, 40% w/w in methanol) was added to the solution, the mixture was stirred for 15 min at -28°C. 10 ml precooled saturated aq NH_4Cl solution was added, the aqueous layer was extracted with CH_2Cl_2 (3×50mL), the organic layer was combined, and dried over anhydrous Na_2SO_4, then the organic layer was filtered, the filtrate was concentrated to dryness in vacuo, The residue was purified on silica gel by column chromatography (hexane/EtOAc/acetone = 8/3/1) to afford white solid 2'-*O*-(tert-butyldimethylsilyl)-2-debenzoyl-7-*O*-triethylsilyltaxol (**3**) (63.6mg, 65.0%) and recover the starting material (**1**) (12mg, 12.3%). The product was used in the next step without identification by spectrums.

2'-*O*-(tert-butyldimethylsilyl)-2-debenzoyl-7-*O*-triethylsilylcephalomannine (**4**)

2'-*O*-(tert-butyldimethylsilyl)-7-*O*-triethylsilylcephalomannine (**2**) (351mg,

0.33mmol) was dissolved in dried CH_2Cl_2 (7 ml), the solution obtained was cooled to -28°C in an ice-methanol bath, then the benzyltrimethylammonium hydroxide (308μl, 40% w/w in methanol) was added to the solution, the mixture was stirred for 15 min at -28°C. 10 mL precooled saturated aq NH_4Cl solution was added, the aqueous layer was extracted with CH_2Cl_2 (3×50mL), the organic layer was combined, and dried over anhydrous Na_2SO_4, then the organic layer was filtered, the filtrate was concentrated to dryness in vacuo, The residue was purified on silica gel by column chromatography (hexane/EtOAc/acetone = 8/3/1) to afford white solid 2'-*O*-(tert-butyldimethylsilyl)-2-debenzoyl-7-*O*-triethylsilylcephalomannine (**4**) (209mg, 67.4%) and recover the starting material (**2**) (58mg, 16.5%). The product was used in the following step without identification by spectrums.

General Procedure for the Preparation of Componds 5a-b and 6a-b. An appropriate carboxylic acid (0.1 mmol) was dissolved in dried toluene (0.2 mL), and then dicyclohexylcarbodiimide (DCC) (20.6 mg, 0.1 mmol) and pyrrolidinopyridine (pp) (1.0 mg, 0.007 mmol) were added to the solution. The mixture was stirred at room temperature for 5 min, taxane **3** or **4**(0.01 mmol) was added, the mixture was stirred at 65 °C until the compound (**3** or **4)** was consumed (TLC analysis). The reaction mixture was diluted with EtOAc(10 mL), filtered through a pad of Celite, and the Celite was washed up with EtOAc (10 mL). The filtrate was concentrated in vacuo to dryness. The residue was purified using PTLC (silica gel, hexane /EtOAc/acetone = 8:3:1) to afford **5** or **6.** Although it was found that the products were contaminated with DCU by ^{1}HNMR, they were subjected to desilylation without further purification.

General Procedure for Desilylation of 5 or 6. To a solution of taxane **5** or **6** (0.022 mmol) in acetonitrile (1.06 mL) in a 5mL-plastic bottle was added pyridine (0.533 mL, 6.60 mmol) and aqueous HF (0.293 mL, 6.60 mmol) dropwise. The mixture was stirred at room temperature until the starting material was consumed, as determined by TLC monitoring. The reaction mixture was diluted with EtOAc, washed with saturated $NaHCO_3$ solution and brine, and the aqueous layers were combined and extracted three times with EtOAc. The organic layer was dried over Na_2SO_4 and concentrated to dryness. Purification of the residue was carried out using PTLC (hexane/acetone = 2:1) to give homogeneous products.

CTX-11: ^{1}H NMR (400MHz, $CDCl_3$): δ 7.74 (3H, d, *J* = 7.8 Hz, *o*-Bz + H-6 of *m*-MeO-Bz), 7.64 (1H, s, H-2 of *m*-MeO-Bz), 7.33-7.52 (9H, m, Bz + *m*-MeO-Bz + Ph), 7.15 (1H, d, *J* = 8.0 Hz, H-4 of *m*-MeO-Bz), 6.98 (1H, d, *J* = 8.8 Hz, BzNH), 6.27 (1H, s, H-10), 6.22 (1H, t, *J* = 8.4 Hz, H-13), 5.77 (1H, d, *J* =8.4 Hz, H-3'), 5.67 (1H, d, *J* = 6.8 Hz, H-2), 4.95 (1H, d, *J* = 9.2 Hz, H-5), 4.78(1H, s, H-2'), 4.40 (1H, dd, *J* = 6.8, 10.4 Hz, H-7), 4.35 (1H, d, *J* = 8.4 Hz, H-20), 4.19 (1H, d, *J* = 8.4 Hz, H-20), 3.88 (3H, s, OMe), 3.80 (1H, d, *J* = 6.8 Hz, H-3), 2.51-2.63 (1H, m, H-6), 2.36 (3H, s, OAc-4), 2.32 (2H, pseudo-t, *J* = 8.0 Hz, H-14×2), 2.24 (3H, s, OAc-10), 1.83-1.91 (1H, m, H-6), 1.79 (3H, s, Me-18), 1.68 (3H, s, Me-19), 1.24 (3H, s, Me-17), 1.15 (3H, s, Me-16). ESI-MS: *m/z* $[M+Na]^+$ 906.4. This compound is identical to compound **14bf** in [4].

CTX-12: ^{1}H NMR (300MHz, $CDCl_3$): δ 7.90 (1H, d, *J* = 7.8 Hz, H-6 of *m*-N_3-Bz), 7.82 (1H, d, *J* = 1.5 Hz, H-2 of *m*-N_3-Bz), 7.72 (2H, d, *J* = 7.2 Hz, *o*-Bz). 7.32-7.52 (9H, m, Bz + *m*- N_3-Bz + Ph), 7.22 (1H, dd, *J* = 2.4, 8.1 Hz, H-4 of *m*-N_3-Bz), 6.97 (1H, d, *J* = 8.7 Hz, BzNH), 6.50 (1H, s, H-10), 6.24 (1H, t, *J* =

9.3 Hz, H-13), 5.76 (1H, dd, *J* = 2.4, 8.7 Hz, H-3'), 5.67 (1H, d, *J* = 7.2 Hz, H-2), 4.95 (1H, d, *J* = 8.1 Hz, H-5), 4.77 (1H, d, *J* = 2.4 Hz, H-2'), 4.40 (1H, dd, *J* = 6.9, 11.0 Hz, H-7), 4.31 (1H, d, *J* = 8.4 Hz, H-20), 4.18 (1H, d, *J* = 8.7 Hz, H-20), 3.81 (1H, d, *J* = 7.5 Hz, H-3), 2.49-2.63 (1H, m, H-6), 2.37 (3H, s, OAc-4), 2.35 (2H, d, *J* = 7.8 Hz, H-14), 2.24 (3H, s, OAc-10), 1.83-1.92 (1H, m, H-6), 1.80 (3H, s, Me-18), 1.68 (3H, s, Me-19), 1.24 (3H, s, Me-17), 1.14 (3H, s, Me-16). ESI-MS: *m/z* $[M+Na]^+$ 917.4. This compound is identical to compound **14bi** in [4].

CTX-13: ^{1}H NMR (300MHz, $CDCl_3$): δ 7.70 (1H, d, *J* = 7.5 Hz, H-6 of *m*-MeO-Bz), 7.68 (1H, m, H-2 of *m*-MeO-Bz), 7.30-7.42 (6H, m, H-5 of *m*-MeO-Bz + Ph), 7.13 (1H, dd, *J* = 2.7, 7.5 Hz, H-4 of *m*-MeO-Bz), 6.55 (1H, d, *J* = 8.7 Hz, CONH), 6.42 (1H, dq, *J* = 1.2, 6.9 Hz, CH_3CH=), 6.26 (1H, s, H-10), 6.18 (1H, t, *J* = 7.8 Hz, H-13), 5.65 (1H, d, *J* = 6.9 Hz, H-2), 5.58 (1H, dd, *J* = 3.0, 8.55 Hz, H-3'), 4.93 (1H, d, *J* = 7.5 Hz, H-5), 4.69 (1H, d, *J* = 3.0 Hz, H-2'), 4.38 (1H, dd, *J* = 6.6, 11.0 Hz, H-7), 4.31 (1H, d, *J* = 8.4 Hz, H-20), 4.17 (1H, d, *J* = 8.4 Hz, H-20), 3.86 (3H, s, OMe), 3.77 (1H, d, *J* = 6.9 Hz, H-3), 2.48-2.62 (1H, m, H-6), 2.32 (3H, s, OAc-4), 2.28 (2H, d, *J* = 7.2 Hz, H-14), 2.23 (3H, s, OAc-10), 1.86-1.90 (1H, m, H-6), 1.79 (3H, s, Me-18), 1.78 (3H, s, $=CCH_3CO$-), 1.71 (3H, d, J = 6.9 Hz, CH_3CH=), 1.66 (3H, s, Me-19), 1.25 (3H, s, Me-17), 1.14 (3H, s, Me-16). ^{13}C NMR (100MHz, $CDCl_3$): δ 203.9, 173.0 171.5, 170.5, 169.4, 167.0, 159.9, 142.3, 138.4, 133.3, 132.2, 131.5, 130.6, 130.0, 129.2 (2C), 128.5, 127.2 (2C), 122.8, 120.4, 114.8, 84.6, 81.4, 79.1, 76.7, 75.8, 75.3, 73.5, 72.5, 72.4, 58.8, 55.7, 55.1, 45.8, 43.4, 35.9 (2C), 27.1, 22.8, 22.1, 21.1, 15.0, 14.2, 12.6, 9.8. ESI-MS: *m/z* $[M+Na]^+$ 884.4. HRMS (ESI, $[M+Na]^+$) Found 884.3479 Calc. for $C_{46}H_{55}NO_{15}Na$ 884.3469.

CTX-14: ^{1}H NMR (300MHz, $CDCl_3$): δ 7.89 (1H, d, *J* = 8.1 Hz, H-6 of *m*-N_3-Bz), 7.79 (1H, m, H-2 of *m*-N_3-Bz), 7.48 (1H, t, *J* = 7.8 Hz, H-5 of *m*-N_3-Bz), 7.31-7.40 (5H, m, Ph), 7.22-7.25 (1H, m, H-4 of *m*-N_3-Bz), 6.49 (1H, d, *J* = 8.7 Hz, CONH), 6.41 (1H, dq, *J* = 1.5, 6.9 Hz, CH_3CH=), 6.27 (1H, s, H-10), 6.17 (1H, t, *J* = 8.5 Hz, H-13), 5.66 (1H, d, *J* = 7.2 Hz, H-2), 5.58 (1H, dd, *J* =2.7, 8.7 Hz, H-3'), 4.94 (1H, d, *J* = 7.5 Hz, H-5), 4.68 (1H, d, *J* = 2.7 Hz, H-2'), 4.39 (1H, dd, *J* = 6.9, 9.45 Hz, H-7), 4.30 (1H, d, *J* = 8.4 Hz, H-20), 4.16 (1H, d, *J* = 8.4 Hz, H-20), 3.79 (1H, d, *J* = 7.2 Hz, H-3), 2.48-2.62 (1H, m, H-6), 2.34 (3H, s, OAc-4), 2.30 (2H, d, *J* = 9.0 Hz, H-14), 2.24 (3H, s, OAc-10), 1.82-1.91 (1H, m, H-6), 1.79 (6H, s, Me-18 + =CCH_3CO-), 1.73 (3H, d, *J* = 6.9 Hz, CH_3CH=), 1.67 (3H, s, Me-19), 1.25 (3H, s, Me-17), 1.14 (3H, s, Me-16). ^{13}C NMR (100MHz, $CDCl_3$): δ 203.6, 172.9, 171.2, 170.3, 169.1, 165.9, 142.1, 140.7, 138.1, 133.0, 131.8, 131.3, 130.9, 130.2, 128.9 (2C), 128.2, 127.0 (2C), 126.8, 124.3, 120.1, 84.4, 81.0, 78.9, 76.4, 75.5, 75.4, 73.1, 72.3, 72.1, 58.5, 54.9, 45.5, 43.1, 35.6, 35.5, 26.8, 22.6, 21.8, 20.8, 14.8, 14.0, 12.4, 9.5. ESI-MS: *m/z* $[M+Na]^+$ 895.4. HRMS (ESI, $[M+Na]^+$) Found 895.3410 Calc. for $C_{45}H_{52}N_4O_{14}Na$ 895.3378.

CTX-16 was obtained when the acid 

was used in the coupling reaction as the preparation of **5** or **6**, and subsequent desilyation.

^{1}H NMR (300MHz, $CDCl_3$): δ 7.32-7.43 (5H, m, Ph), 6.42-6.48 (2H, m, NH + CH_3CH=), 6.29 (1H, s, H-10), 6.15 (1H, t, *J* = 8.4 Hz, H-13), 5.53 (1H, dd, *J* = 2.4, 8.7 Hz, H-3'), 4.96 (1H, d, *J* = 8.4 Hz, H-5), 4.66 (1H, d, *J* = 2.7 Hz, H-2'), 4.62 (1H, d, *J* = 8.7 Hz, H-20), 4.47 (1H, d, *J* = 6.3 Hz, H-2), 4.46 (1H, d, *J* = 8.1 Hz, H-20), 4.35 (1H, dd, *J* = 7.5, 9.9 Hz, H-7), 3.42 (1H, d, *J* = 5.7 Hz, H-3), 2.73 (1H, dd, *J* = 9.6, 15.8 Hz, H-14), 2.55-2.65 (1H, m, H-6), 2.15-2.23 (1H, overlapped, H-14), 2.21 (3H, s, OAc-4), 2.17 (3H, s, OAc-10), 1.88-1.94 (1H, m,

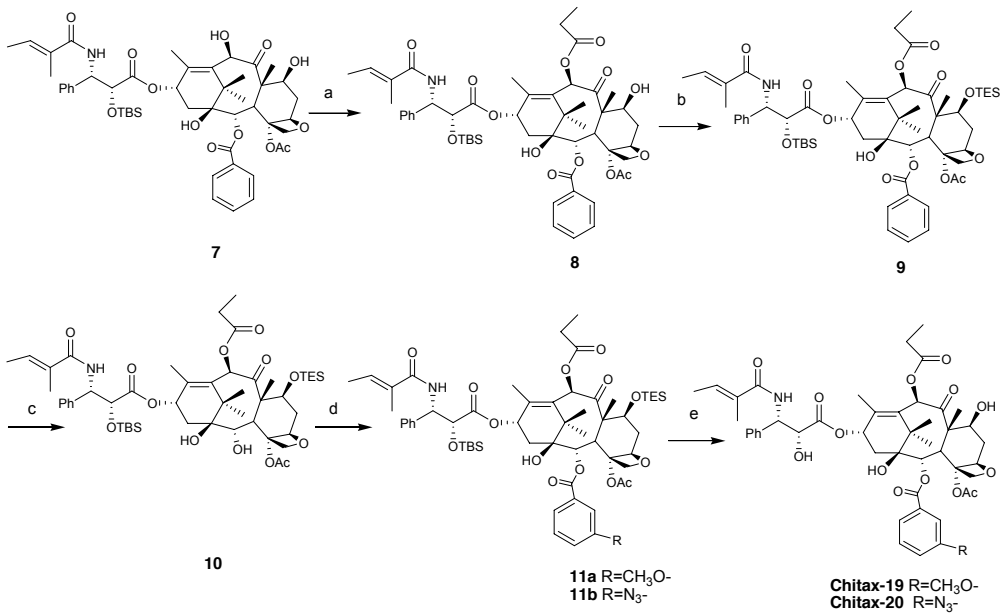
H-6), 1.83 (3H, s, Me-18), 1.81 (3H, s, =CC$\underline{H}_3$CO-), 1.76 (3H, d, *J* = 7.5 Hz, C$\underline{H}_3$CH=), 1.74 (3H, s, Me-19), 1.38 (3H, s, Me-17), 1.18 (3H, s, Me-16). ^{13}C NMR (125MHz, $CDCl_3$): δ 203.6, 172.6, 170.7, 170.6, 152.5, 145.3, 138.0, 132.0, 131.2, 131.0, 129.0 (2C), 128.4, 127.2 (2C), 89.7, 84.1, 81.2, 79.9, 76.2, 76.0, 73.0, 71.5, 71.0, 60.1, 55.0, 43.0, 41.3, 36.4, 32.6, 25.7, 22.5, 20.8, 20.5, 15.6, 14.0, 12.4, 9.5. ESI-MS: *m/z* $[M+Na]^+$ 936.3, $[M+Na+2H]^+$ 776.4. HRMS (ESI, $[M+Na]^+$) Found 776.2918 Calc. for $C_{39}H_{47}NO_{14}Na$ 776.2894.

CTX-15 was prepared from paclitaxel through 2'-OH protection with TBS, 10-deacetylation with hydrazine hydrate and then 2'-desilylation.

^{1}H NMR (300MHz, $CDCl_3$): δ 8.11 (1H, d, *J* =7.2 Hz, *o*-BzO), 7.75 (2H, t, *J* = 7.2 Hz, *o*-$\underline{Bz}$NH), 7.61 (1H, t, *J* = 7.2 Hz, *p*-BzO), 7.31-7.52 (10H, m, BzO + $\underline{Bz}$NH + Ph), 7.16 (1H, d, *J* = 8.8 Hz, BzN$\underline{H}$), 6.18 (1H, t, *J* = 8.4 Hz, H-13), 5.77 (1H, dd, *J* = 2.8, 8.8 Hz, H-3'), 5.66 (1H, d, *J* = 6.8 Hz, H-2), 5.18 (1H, s, H-10), 4.92 (1H, d, *J* = 8.0 Hz, H-5), 4.77 (1H, d, *J* = 2.8 Hz, H-2'), 4.30 (1H, d, *J* = 8.8 Hz, H-20), 4.21 (1H, overlapped, H-7), 4.20 (1H, d, *J* = 8.8 Hz, H-20), 3.88 (1H, d, *J* = 7.2 Hz, H-3), 2.53-2.62 (1H, m, H-6), 2.37 (3H, s, OAc-4), 2.27 (2H, dd, *J* = 6.0, 8.6 Hz, H-14), 1.80-1.87 (1H, m, H-6), 1.75 (3H, s, Me-18), 1.74 (3H, s, Me-19), 1.19 (3H, s, Me-17), 1.10 (3H, s, Me-16). ^{13}C NMR (100MHz, $CDCl_3$): δ 211.2, 172.5, 170.5, 167.0, 166.9, 138.1, 137.9, 136.1, 133.7, 133.7, 131.9, 130.2 (2C), 129.2, 129.0 (2C), 128.7 (2C), 128.7 (2C), 128.3, 127.1 (2C), 127.1 (2C), 84.1, 81.1, 78.7, 76.6, 74.8, 74.5, 73.3, 72.4, 72.0, 57.7, 55.1, 46.4, 43.0, 37.0, 35.9, 26.5, 22.5, 20.5, 14.3, 9.8. ESI-MS: *m/z* $[M+Na]^+$ 834.4. This compound is identical to 10-deacetyl taxol in [5].

CTX-17 was prepared from cephalomannine through 2'-OH protection with TBS, 10-deacetylation with hydrazine hydrate and then 2'-desilylation.

^{1}H NMR (300MHz, $CDCl_3$): δ 8.10 (2H, d, *J* = 7.2 Hz, *o*-Bz), 7.61 (1H, t, *J* = 7.2 Hz, *p*-Bz), 7.50 (2H, t, *J* = 7.2 Hz, *m*-Bz), 7.28-7.39 (5H, m, Ph), 6.65 (1H, d, *J* = 9.0 Hz, CONH), 6.44 (1H, dq, *J* = 1.2, 7.2 Hz, CH_3CH=), 6.16 (1H, t, *J* = 8.1 Hz, H-13), 5.66 (1H, d, *J* = 6.9 Hz, H-2), 5.59 (1H, dd, *J* = 2.7, 8.9 Hz, H-3'), 5.18 (1H, s, H-10), 4.91 (1H, d, *J* = 7.8 Hz, H-5), 4.69 (1H, brs, OH), 4.39 (1H, d, *J* = 8.7 Hz, H-20), 4.24 (1H, brs, OH), 4.19 (1H, d, *J* = 8.4 Hz, H-20), 3.87 (1H, d, *J* = 6.9 Hz, H-3), 3.72 (1H, brs, OH), 2.53-2.62 (1H, m, H-6), 2.34 (3H, s, OAc-4), 2.24 (2H, d, *J* = 9.0 Hz, H-14), 1.82-1.88 (1H, m, H-6), 1.80 (3H, s, $=CCH_3CO$-), 1.76 (3H, s, Me-18), 1.72 (3H, d, *J* = 7.5 Hz, CH_3CH=), 1.21 (3H, s, Me-17), 1.11 (3H, s, Me-16). ^{13}C NMR (100MHz, $CDCl_3$): δ 211.3, 172.6, 170.4, 169.0, 166.9, 138.2, 138.1, 136.0, 133.7, 131.9, 131.3, 130.2 (2C), 129.2, 128.9 (2C), 128.7 (2C), 128.2, 127.0 (2C), 84.1, 81.0, 78.7, 76.7, 74.8, 74.5, 73.3, 72.3, 72.0, 57.6, 54.8, 46.4, 43.0, 36.9, 35.8, 26.5, 22.5, 20.6, 14.3, 14.0, 12.4, 9.8. ESI-MS: *m/z* $[M+Na]^+$ 812.3. This compound is identical to 10-deacetyl cephalomannine in [5].



Scheme 2 Reagents and conditions: a: $(CH_3CH_2CO)_2$, $CeCl_3$; b: TESCl, imidazole, DMF; c: Triton B, CH_2Cl_2, -28°C d: DCC, pp, *m*-R-PhCOOH; e: 40% HF(v/v), pyridine, CH_3CN.

2'-*O*-(tert-butyldimethylsilyl)-10-propionyl-10-deacetylcephalomannine (**8**)

2'-(tert-butyldimethylsilyl)-10-deacetylcephalomannine (**7**) (198mg, 0.22mmol) was dissolved in THF (6 ml), then propionic anhydride (280μl, 2.2mmol) and (16mg, 0.07mmol) were added to the solution, the mixture was reacted for 2 h at room temperature. The mixture was diluted with EtOAc (30 ml), then washed by saturated aq $NaHCO_3$ solution (30 ml), the aqueous layer was extracted with EtOAc (3×50mL), the organic layer was combined, and dried over anhydrous Na_2SO_4, then the organic layer was filtered, the filtrate was concentrated in vacuo. The residue was purified on silica gel by column chromatography (Petroleum ether/EtOAc = 4/1) to afford white solid 2'-*O*-(tert-butyldimethylsilyl)-10-propionyl-10-deacetylcephalomannine (**8**) (193mg, 91.5%).

^{1}H 400M (CD_3COCD_3): δ 8.14 (2H, d, *J* = 7.2 Hz, *o*-Bz), 7.64 (1H, t, *J* = 7.2 Hz, *p*-Bz), 7.41-7.66 (4H, m, Ph), 7.40 (2H, t, *J* = 7.2 Hz, *m*-Bz), 7.29 (1H, t, *J* = 7.2 Hz, Ph), 7.23 (1H, d, *J* = 8.4 Hz, CONH), 6.43 (1H, s, H-10), 6.39 (1H, dq, *J* =

1.2, 7.2 Hz, $CH_3C\underline{H}$=), 6.22 (1H, t, *J* = 8.8 Hz, H-13), 5.72 (1H, overlapped, H-3'), 5.70 (1H, d, *J* = 7.2 Hz, H-2), 4.98 (1H, dd, *J* = 2.0, 9.2 Hz, H-5), 4.89 (1H, d, *J* = 4.0 Hz, H-2'), 4.43 (1H, m, H-7), 4.20 (1H, d, *J* = 8.4 Hz, H-20), 4.16 (1H, d, *J* = 8.0 Hz, H-20), 3.88 (1H, d, *J* = 9.2 Hz, O$\underline{H}$), 3.87 (1H, s, O$\underline{H}$), 3.53 (1H, d, *J* = 6.0 Hz, H-3), 2.61 (3H, s, OAc-4), 2.44-2.51 (4H, m, H-6 +14 + $OCOC\underline{H}_2CH_3$), 2.20 (1H, dd, *J* = 8.8, 15.2 Hz, H-14), 1.89 (3H, d, *J* = 0.8 Hz, H-18), 1.77 (3H, s, =$CC\underline{H}_3CO$-), 1.76-1.83 (1H, m, overlapped, H-6), 1.68 (3H, d, *J* = 8.0 Hz, $C\underline{H}_3CH$=), 1.67 (3H, s, Me-19), 1.23 (3H, s, Me-17), 1.20 (3H, s, Me-16), 1.16 (3H, t, *J* = 7.6 Hz, $OCOCH_2C\underline{H}_3$), 0.83 (9H, s, $SiC(C\underline{H}_3)_3$), -0.003 (3H, s, $SiC\underline{H}_3$), -0.18 (3H, s, $SiC\underline{H}_3$).

2'-*O*-(tert-butyldimethylsilyl)-7-*O*-triethylsilyl-10-propionyl-10-deacetylcephalo mannine (**9**)

2'-*O*-(tert-butyldimethylsilyl)-10-propionyl-10-deacetylcephalomannine (**8**) (193mg, 0.20mmol) was dissolved in DMF (2 ml), then imidazole (73mg, 1.07mmol) and triethylsilyl chloride(135μl, 0.80mmol) were added to the solution, the mixture was reacted for 6 h at room temperature. Additional imidazole (73mg, 1.07mmol) and triethylsilyl chloride(135μl, 0.80mmol) were added, the mixture was reacted 6 h at room temperature. 30 mL saturated aq $NaHCO_3$ solution was added, the aqueous layer was extracted with EtOAc (3×50mL), the organic layer was combined, and dried over anhydrous Na_2SO_4, then the organic layer was filtered, the filtrate was concentrated in vacuo. The residue was purified on silica gel with column chromatography (Petroleum ether/EtOAc= 5/1) to afford white solid 2'-*O*-(tert-butyldimethylsilyl)-7-*O*-triethylsilyl-10-propionyl-10-deacetylcephalomannine (**9**) (196mg, 92%).

^{1}H NMR (300MHz, $CDCl_3$): δ 8.12 (2H, d, *J* = 7.2 Hz, *o*-Bz), 7.59 (1H, t, *J* = 6.9 Hz, *p*-Bz), 7.49 (2H, t, *J* = 7.2 Hz, *m*-Bz), 7.24-7.38 (5H, m, Ph), 6.65 (1H, d,

J = 8.7 Hz, CONH), 6.46 (1H, s, H-10), 6.43 (1H, q, *J* = 6.6 Hz, CH_3CH=), 6.24 (1H, t, *J* = 9.0, H-13), 5.70 (1H, d, *J* = 6.9 Hz, H-2), 5.57 (1H, d, *J* = 8.7 Hz, H-3'), 4.94 (1H, d, *J* = 8.4 Hz, H-5), 4.58 (1H, d, *J* = 1.8 Hz, H-2'), 4.47 (1H, dd, *J* = 6.6, 10.5 Hz, H-7), 4.30 (1H, d, *J* = 8.4 Hz, H-20), 4.19 (1H, d, *J* = 8.4 Hz, H-20), 3.83 (1H, d, *J* = 6.9 Hz, H-3), 2.54 (3H, s, OAc-4), 2.32-2.52 (4H, m, H-6 + H-14 +$OCOCH_2CH_3$), 2.12 (1H, dd, *J* = 8.7, 15.3 Hz, H-14), 2.02 (3H, s, Me-18), 1.85-1.83 (1H, m, H-6), 1.80 (3H, s, $=CCH_3CO$-), 1.70 (3H, d, *J* = 7.5 Hz, CH_3CH=), 1.69 (3H, s, Me-19), 1.23 (3H, s, Me-17), 1.20 (3H, s, Me-16), 1.20 (3H, t, *J* = 7.8 Hz, $OCOCH_2CH_3$), 0.91 (9H, t, *J* = 7.8 Hz, $SiCH_2CH_3$), 0.78 (9H, s, $SiC(CH_3)_3$), 0.50-0.61 (6H, m, $SiCH_2CH_3$), -0.06(3H, s, $SiCH_3$), -0.34 (3H, s, $SiCH_3$). ESI-MS: *m/z* $[M+Na]^+$ 1096.6

2'-*O*-(tert-butyldimethylsilyl)-2-debenzoyl-7-*O*-triethylsilyl-10-propionyl-10-de acetylcephalomannine (**10**)

2'-*O*-(tert-butyldimethylsilyl)-7-*O*-triethylsilyl-10-propionyl-10-deacetylceph alomannine (**9**) (83mg, 0.077mmol) was dissolved in dried CH_2Cl_2 (2.5 mL), the obtained solution was cooled to -28°C in ice-methanol bath, then the benzyltrimethylammonium hydroxide (68μl, 40% w/w in methanol) was added to the solution, the mixture was stirred for 35 min at -28°C. 20 mL precooled saturated aq NH_4Cl solution was added, the aqueous layer was extracted with CH_2Cl_2 (3×30mL), the organic layer was combined, and dried over anhydrous Na_2SO_4, then the organic layer was filtered, the filtrate was concentrated to dryness in vacuo, The residue was purified on silica gel by column chromatography (hexane/EtOAc/acetone = 8/3/1) to afford white solid 2'-*O*-(tert-butyldimethylsilyl)-2-debenzoyl-7-*O*-triethylsilyl-10-propionyl-10-deacetylcephalomannine (**10**) (45.6mg, 61.1%) and recover the starting material (**9**) (15mg, 18%).

^{1}H 400M (CD_3COCD_3): δ 7.45 (2H, d, *J* = 7.6 Hz, *o*-Ph), 7.37 (2H, t, *J* = 7.2 Hz, *m*-Ph), 7.29 (1H, t, *J* = 7.2 Hz, *p*-Ph), 7.16 (1H, d, *J* = 9.2 Hz, CONH), 6.42 (1H, s, H-10), 6.39 (1H, dq, *J* = 1.2, 6.8 Hz, CH_3CH=), 6.14 (1H, dt, *J* = 1.2, 9.2 Hz, H-13), 5.64 (1H, dd, *J* = 3.2, 9.6 Hz, H-2'), 4.92 (1H, dd, *J* = 2.0, 8.0 Hz, H-5), 4.79 (1H, d, *J* = 2.8 Hz, H-2'), 4.63 (1H, dd, *J* = 1.2, 8.8 Hz, H-20), 4.52 (1H, d, *J* = 8.0 Hz, H-20), 4.51 (1H, dd, *J* = 6.8, 10.4 Hz, H-7), 4.11 (1H, s, OH), 3.85 (1H, s, OH), 3.95 (1H, dd, *J* = 5.2, 6.4 Hz, H-2), 3.84 (1H, d, *J* = 4.8 Hz, OH), 3.51 (1H, d, *J* = 6.4 Hz, H-3), 2.52 (1H, ddd, *J* = 6.6, 9.9, 14.1 Hz, H-6), 2.42 (2H, dq, *J* = 1.6, 7.6 Hz, $OCOCH_2CH_3$), 2.41 (3H, s, OAc-4), 2.29 (1H, dd, *J* = 9.6, 15.6 Hz, H-14), 2.08 (1H, dd, *J* = 8.8, 15.2 Hz, H-14), 1.94 (3H, d, *J* = 1.6 Hz, Me-18), 1.77-1.84 (1H, m, H-6), 1.82 (3H, t, *J* = 1.2 Hz, $=CCH_3CO$-), 1.71 (3H, dd, *J* = 1.2, 6.8 Hz, CH_3CH=), 1.61 (3H, s, Me-19), 1.19 (3H, s, Me-17), 1.14 (3H, t, *J* = 7.2 Hz, $OCOCH_2CH_3$), 1.10 (3H, s, Me-16), 0.93 (9H, t, *J* = 8.0 Hz, $SiCH_2CH_3$), 0.80 (9H, s, $SiC(CH_3)_3$), 0.56-0.63 (6H, m, $SiCH_2CH_3$), -0.03 (3H, s, $SiCH_3$), -0.23 (3H, s, $SiCH_3$).

General Procedure for the Preparation of Compounds 11a-b. An appropriate carboxylic acid (0.1 mmol) was dissolved in dried toluene (0.2 mL), and then dicyclohexylcarbodiimide (20.6 mg, 0.1 mmol) and pyrrolidinopyridine (1.0 mg, 0.007 mmol) were added to the solution. The mixture was stirred at room temperature for 5 min, taxane **10** (0.01 mmol) was added, the mixture was stirred at 65 °C until the compound **10** was consumed (TLC analysis). The reaction mixture was diluted with EtOAc(10 mL), filtered through a pad of Celite, and the Celite was washed up with EtOAc (10 mL). The filtrate was concentrated in vacuo to dryness. The residue was purified using PTLC (silica gel, hexane /EtOAc/acetone = 8:3:1) to afford **20.** Although it was found that

the products were contaminated with DCU by ^{1}H-NMR, they were subjected to desilylation without further purification.

General Procedure for Desilylation of 11. To a solution of taxane **11** (0.022 mmol) in acetonitrile (1.06 mL) in a 5mL-plastic bottle was added pyridine (0.533 mL, 6.60 mmol) and aqueous HF (0.293 mL, 6.60 mmol) dropwise. The mixture was stirred at room temperature until the starting material was consumed, as determined by TLC monitoring. The reaction mixture was diluted with EtOAc, washed with saturated $NaHCO_3$ solution and brine, and the aqueous layers were combined and extracted three times with EtOAc. The organic layer was dried over Na_2SO_4 and concentrated to dryness. Purification of the residue was carried out on silica gel by column chromatography (hexane/acetone = 2:1) to give homogeneous products.

CTX-19: ^{1}H NMR (300MHz, $CDCl_3$): δ 7.71 (1H, d, *J* = 8.1 Hz, H-6 of *m*-MeO-Bz), 7.62-7.69 (1H, m, H-2 of *m*-MeO-Bz), 7.29-7.43 (6H, m, H-5 of *m*-MeO-Bz + Ph), 7.15 (1H, ddd, *J* = 0.9, 2.7, 8.3 Hz, H-4 of *m*-MeO-Bz), 6.51 (1H, d, *J* = 8.7 Hz, CONH), 6.42 (1H, dq, *J* = 1.5, 7.2 Hz, CH_3CH=), 6.39 (1H, s, H-10), 6.19 (1H, t, *J* = 7.5 Hz, H-13), 5.66 (1H, d, *J* = 6.9 Hz, H-2), 5.59 (1H dd, *J* = 2.7, 8.6 Hz, H-3'), 4.94 (1H, d, *J* = 7.5 Hz, H-5), 4.69 (1H, brs, H-2'), 4.36-4.44 (1H, m, H-7), 4.33 (1H, d, *J* =8.4 Hz, H-20), 4.17 (1H, d, *J* = 8.4 Hz, H-20), 3.88 (3H, s, OMe), 3.79 (1H, d, *J* = 6.9 Hz, H-3), 3.63 (1H, brs, OH), 2.45-2.63 (3H, m, H-6 + $OCOCH_2CH_3$), 2.33 (3H, s, OAc-4), 2.26 (2H, dd, *J* = 3.0, 9.0 Hz, H-14), 1.86-1.92 (1H, m, H-6), 1.80 (3H, pseud-t, *J* =1.2 Hz, $=CCH_3CO$-), 1.79 (3H, d, *J* =1.2 Hz, Me-18), 1.72 (3H, dd, *J* = 1.2, 6.9Hz, CH_3CH=), 1.67 (3H, s, Me-19), 1.25 (3H, s, Me-17), 1.23 (3H, t, *J* = 7.5 Hz, $OCOCH_2CH_3$), 1.14 (3H, s, Me-16). ^{13}C NMR (100MHz, $CDCl_3$): δ 203.7, 174.6,

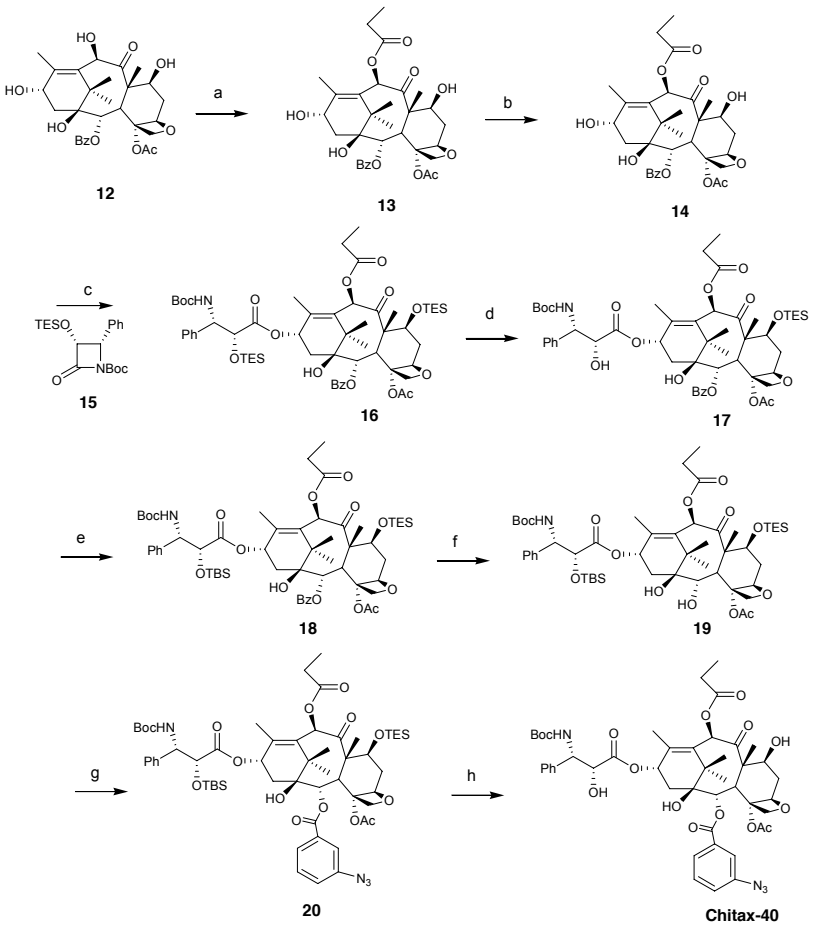
172.7, 170.3, 169.0, 166.8, 159.6, 142.0, 138.1, 133.1, 131.9, 131.3, 130.3, 129.7, 129.0 (2C), 128.3, 127.0 (2C), 122.6, 120.3, 114.5, 84.4, 81.1, 79.0, 76.5, 75.4, 75.0, 73.3, 72.2, 72.2, 58.6, 55.4, 54.9, 45.6, 43.2, 35.6 (2C), 27.6, 26.8, 22.6, 21.8, 14.8, 14.0, 12.4, 9.5, 9.0. ESI-MS: *m/z* $[M+Na]^+$ 898.4. HRMS (ESI, $[M+Na]^+$) Found 898.3647 Calc. for $C_{47}H_{57}NO_{15}Na$ 898.3626.

CTX-20: ^{1}H NMR (300MHz, $CDCl_3$): δ 7.89 (1H, d, *J* = 8.1 Hz, H-6 of *m*-N_3-Bz), 7.80 (1H, m, H-2 of *m*-N_3-Bz), 7.48 (1H, t, *J* = 7.8 Hz, H-5 of *m*-N_3-Bz), 7.31-7.42 (5H, m, Ph), 7.24 (1H, ddd, *J* = 1.2, 2.4, 8.1 Hz, H-4 of *m*-N_3-Bz), 6.49 (1H, d, *J* = 8.7 Hz, CONH), 6.40 (1H, dq, *J* = 1.5, 6.9 Hz, CH_3CH=), 6.28 (1H, s, H-10), 6.17 (1H, t, *J* = 8.1 Hz, H-13), 5.66 (1H, d, *J* = 6.9 Hz, H-2), 5.58 (1H, dd, *J* = 2.7, 8.7 Hz, H-3'), 4.94 (1H, d, *J* = 7.8 Hz, H-5), 4.68 (1H, brs, H-2'), 4.40 (1H, m, H-7), 4.30 (1H, d, *J* = 8.4 Hz, H-20), 4.17 (1H, d, *J* = 8.4 Hz, H-20), 3.80 (1H, d, *J* = 6.9 Hz, H-3), 3.63 (1H, brs, OH), 2.47-2.62 (3H, m, H-6 + $OCOCH_2CH_3$), 2.37 (3H, s, OAc-4), 2.31 (2H, overlapped, H-14), 1.82-1.95 (1H, m, H-6), 1.78-1.80 (6H, brs, $=CCH_3CO$- + Me-18), 1.72 (3H, dd, *J* = 1.5, 6.9 Hz, CH_3CH=), 1.67 (3H, s, Me-19), 1.25 (3H, s, Me-17), 1.23 (3H, t, *J* = 7.8 Hz, $OCOCH_2CH_3$), 1.14 (3H, s, Me-16). ^{13}C NMR (100MHz, $CDCl_3$): δ 203.7, 174.6, 172.9, 170.3, 169.1, 165.9, 142.1, 140.7, 138.2, 133.1, 131.8, 131.3, 130.9, 130.2, 128.9 (2C), 128.2, 127.0 (2C), 126.8, 124.3, 120.1, 84.4, 81.1, 79.0, 76.4, 75.4, 75.3, 73.1, 72.3, 72.1, 58.5, 54.9, 45.6, 43.2, 35.6, 35.5, 27.5, 26.8, 22.6, 21.8, 14.8, 14.0, 12.4, 9.5, 9.0. ESI-MS: *m/z* $[M+Na]^+$ 909.4. HRMS (ESI, $[M+Na]^+$) Found 909.3565 Calc. for $C_{46}H_{54}N_4O_{14}Na$ 909.3534.

CTX-18 was obtained by desilylation of **9** in the above process for the preparation of **CTX-19/20**.

^{1}H NMR (300MHz, $CDCl_3$): δ 8.11 (1H, d, *J* = 6.9 Hz, *o*-Bz), 7.61 (1H, t, *J* =

7.5 Hz, *p*-Bz), 7.50 (2H, t, *J* = 7.5 Hz, *m*-Bz), 7.31-7.42 (5H, m, Ph), 6.54 (1H, d, *J* = 9.0 Hz, CONH), 6.43 (1H, dq, *J* = 1.2, 6.9 Hz, CH_3CH=), 6.28 (1H, s, H-10), 6.20 (1H, t, *J* = 7.5 Hz, H-13), 5.66 (1H, d, *J* = 7.2 Hz, H-2), 5.61 (1H, dd, *J* = 3.0, 8.7 Hz, H-3'), 4.93 (1H, dd, *J* = 2.1, 9.5 Hz, H-5), 4.70 (1H, d, *J* = 2.4 Hz, H-2'), 4.36-4.42 (1H, m, H-7), 4.29 (1H, d, *J* = 8.7 Hz, H-20), 4.18 (1H, d, *J* = 8.4, H-20), 3.79 (1H, d, *J* = 7.2, H-3), 3.66 (1H, brs, OH), 2.47-2.62 (3H, m, H-6 + OCO$CH_2$$CH_3$), 2.35 (3H, s, OAc-4), 2.25-2.34 (2H, m, H-14), 1.86-1.91 (1H, m, H-6), 1.80 (3H, s, =CCH_3CO-), 1.79 (3H, s, Me-18), 1.72 (3H, dd, *J* = 1.2, 6.9 Hz, CH_3CH=), 1.67 (3H, s, Me-19), 1.25 (3H, s, Me-17), 1.23 (3H, t, *J* = 7.5 Hz, OCO$CH_2$$CH_3$), 1.14 (3H, s, Me-16). ^{13}C NMR (100MHz, $CDCl_3$): δ 203.719, 174.595, 172.759, 170.311, 169.025, 166.965, 141.953, 138.125, 133.678, 133.149, 131.841, 131.297, 130.178 (2C), 129.123, 128.933 (2C), 128.680 (2C), 128.208, 126.955 (2C), 84.397, 81.076, 79.009, 76.464, 75.358, 74.942, 73.314, 72.228, 72.174, 58.563, 54.828, 45.605, 43.159, 35.612, 35.581, 27.543, 26.818, 22.563, 21.857, 14.771, 13.961, 12.405, 9.537, 8.995. ESI-MS: *m/z* $[M+Na]^+$ 868.4. HRMS (ESI, $[M+Na]^+$) Found 868.3537 Calc. for $C_{46}H_{55}NO_{14}Na$ 868.3520.



Scheme 3 Reagents and conditions: a: $(CH_3CH_2CO)_2$, $CeCl_3$, THF; b: TESCl, imidazole, DMF; c: LHMDS, beta-lactam,THF; d: HF, pyridine,CH_3CN; e: TBSCl, imidazole, DMF; f: Triton B, CH_2Cl_2; g: DCC, m-N_3PhCOOH; h:40% HF(v/v), pyridine, CH_3CN.

10-propionyl-10-deacetylbaccatin (III) (**13**)

10-deacetylbaccatin (III) (**12**) (21.7mg, 0.04mmol) was dissolved in THF (1.2 ml), then propionic anhydride (51.4μl, 0.4mmol) and $CeCl_3$(2mg, 0.002mmol) were added to the solution, the mixture was reacted for 2 h at 35°C. 10 ml saturated aq $NaHCO_3$ solution was added, the aqueous layer was extracted with EtOAc (3×20mL), the organic layer was combined, then dried over anhydrous Na_2SO_4, and filtered, the filtrate was concentrated in vacuo. Purification of the residue was carried out using preparation thin layer chromatography (PTLC) (Petroleum ether/EtOAc = 1/3) to afford white solid

10-propionyl-10-deacetylbaccatin (III) (**13**) (22.3mg, 93.2%).

^{1}H NMR (300MHz, $CDCl_3$): δ 8.10 (2H, d, *J* = 7.2 Hz, *o*-Bz), 7.61 (1H, t, *J* = 7.2 Hz, *p*-Bz), 7.48 (2H, t, *J* = 7.2 Hz, *m*-Bz), 6.33 (1H, s, H-10), 5.62 (1H, d, *J* =7.2 Hz, H-2), 4.99 (1H, dd, *J* = 2.1, 9.8 Hz, H-5), 4.89 (1H, t, *J* = 8.4 Hz, H-13), 4.47 (1H, dd, *J* = 6.6, 10.8 Hz, H-7), 4.31 (1H, d, *J* = 8.1 Hz, H-20), 4.15 (1H, d, *J* = 8.1 Hz, H-20), 3.89 (1H, d, *J* = 7.2 Hz, H-3), 2.51-2.64 (3H, m, H-6 + $OCOC\underline{H}_2CH_3$), 2.29 (2H, overlapped, H-14), 2.28 (3H, s, OAc-4), 2.05 (3H, d, *J* = 1.2 Hz, Me-18), 1.82-1.91 (1H, m, H-6), 1.67 (3H, s, Me-19), 1.30 (3H, s, Me-17), 1.23 (3H, t, *J* = 7.5 Hz, $OCOCH_2C\underline{H}_3$), 1.10 (3H, s, Me-16). ESI-MS: *m/z* $[M+Na]^+$ 623.4, $[M+K]^+$ 639.3.

7-*O*-triethylsilyl-10-propionyl-10-deacetylbaccatin(III) (**14**)

10-propionyl-10-deacetylbaccatin(III) (**13**) (198mg, 0.33mmol) was dissolved in DMF (3 ml), then imidazole (180mg, 2.64mmol) and triethylsilyl chloride (360μl, 2.15mmol) were added to the solution, the mixture was reacted for 1.5 h at room temperature. 20 ml saturated aq NH_4Cl solution was added, the aqueous layer was extracted with EtOAc (3×50mL), the organic layer was combined, then dried over anhydrous Na_2SO_4, and filtered, the filtrate was concentrated in vacuo. Purification of the residue was carried out using PTLC (Petroleum ether/EtOAc = 4/1) to afford white solids 7-*O*-triethylsilyl-10-propionyl- 10-deacetylbaccatin(III) (**14**) (212mg, 89.9%).

^{1}H NMR (300MHz, $CDCl_3$): δ 8.11 (2H, d, *J* = 8.4 Hz, *o*-Bz), 7.60 (1H, t, *J* = 6.9 Hz, *p*-Bz), 7.48 (2H, t, *J* = 7.2 Hz, *m*-Bz), 6.48 (1H, s, H-10), 5.64 (1H, d, *J* =6.9 Hz, H-2), 4.96 (1H, d, *J* = 9.0 Hz, H-5), 4.81 (1H, t, *J* = 7.8 Hz, H-13), 4.49 (1H, dd, *J* = 6.9, 10.5 Hz, H-7), 4.31 (1H, d, *J* = 8.4 Hz, H-20), 4.15 (1H, d, *J* = 8.4 Hz, H-20), 3.89 (1H, d, *J* = 7.2 Hz, H-3), 2.38-2.56 (3H, m, H-6 + $OCOC\underline{H}_2CH_3$), 2.29 (3H, s, OAc-4), 2.28-2.29 (2H, overlapped, H-14), 2.21 (3H,

s, Me-18), 1.83-1.91 (1H, m, H-6), 1.68 (3H, s, Me-19), 1.21 (3H, s, Me-17), 1.21 (3H, t, *J* = 6.0 Hz, $OCOCH_2C\underline{H}_3$), 1.04 (3H, s, Me-16). 0.92 (9H, t, *J* = 7.8 Hz, $SiCH_2C\underline{H}_3$), 0.54-0.62 (6H, m, $SiC\underline{H}_2CH_3$), ESI-MS: *m/z* $[M+Na]^+$ 737.4.

2',7-*O*-di(triethylsilyl)-10-propionyldocetaxel (**16**)

7-*O*-triethylsilyl-10-propionyl-10-deacetylbaccatin(III) (**14**) (95mg, 0.132mmol) was dissolved in THF (3 ml), then the lithium bis(trimethylsilyl)amide (1.06M, 333μl) was added to the solution under N_2 at -50°C in the dry ice-acetone bath, the mixture was stirred for 20 min at this temperature. The β-lactam **15** (102mg, 0.27mmol)was added to the mixture, then the obtained mixture was stirred again 20 min at 50°C. 10 ml saturated aq NH_4Cl solution was added, the aqueous layer was extracted with EtOAc (3×50mL), the organic layer was combined, then dried over anhydrous Na_2SO_4, and filtered, the filtrate was concentrated in vacuo. The residue was purified on silica gel with column chromatography (Petroleum ether/EtOAc = 10/1) to afford white solids 2',7-*O*-di(triethylsilyl)-10-propionyldocetaxel (**16**) (135mg, 93.1%).

^{1}H NMR (500MHz, $CDCl_3$): δ 8.13 (2H, d, *J* = 7.5 Hz, *o*-Bz), 7.59 (1H, t, *J* = 7.0 Hz, *p*-Bz), 7.49 (2H, t, *J* = 7.5 Hz, *m*-Bz), 7.37 (2H, t, *J* = 7.5 Hz, Ph), 7.28-7.38 (3H, m, Ph), 6.49 (1H, s, H-10), 6.28 (1H, t, *J* = 9.0 Hz, H-13), 5.71 (1H, d, *J* = 6.5 Hz, H-2), 5.45 (1H, brs, BocN<u>H</u>), 5.31 (1H, brs, H-3'), 4.96 (1H, d, *J* = 9.0 Hz, H-5), 4.56 (1H, s, H-2'), 4.49 (1H, dd, *J* = 7.0, 9.5 Hz, H-7), 4.32 (1H, d, *J* = 8.0 Hz, H-20), 4.19 (1H, d, *J* = 8.0 Hz, H-20), 3.86 (1H, d, *J* = 7.0 Hz, H-3), 2.53 (3H, s, OAc-4), 2.17-2.51 (5H, m, H-6 + H-14×2 + $OCOC\underline{H}_2CH_3$), 2.03 (3H, s, Me-18), 1.88-1.93 (1H, m, H-6), 1.70 (3H, s, Me-19), 1.30 (9H, s, Me-<u>Boc</u>), 1.26 (3H, s, Me-17), 1.23 (3H, s, Me-16), 1.22 (3H, d, *J* = 6.5 Hz, $OCOCH_2C\underline{H}_3$), 0.93 (9H, t, *J* = 7.5 Hz, $SiCH_2C\underline{H}_3$), 0.78 (9H, t, *J* = 7.5 Hz, $SiCH_2C\underline{H}_3$), 0.56-0.61 (6H, m, $SiC\underline{H}_2CH_3$), 0.33-0.45 (6H, m, $SiC\underline{H}_2CH_3$). ^{13}C NMR

(125MHz, $CDCl_3$): δ 201.9, 172.7, 171.6, 170.0, 167.2, 155.2, 140.5, 139.0, 133.6, 133.5, 130.2 (2C), 129.2, 128.7 (2C), 128.5 (2C), 127.7, 126.4 (2C), 84.3, 81.1, 79.8, 79.0, 76.6, 75.2, 75.0, 74.7, 72.2, 71.3, 58.4, 56.3, 46.7, 43.4, 37.2, 35.4, 28.1 (3C), 27.6, 26.5, 23.0, 21.6, 14.1, 10.1, 9.2, 6.8 (3C), 6.5 (3C), 5.3 (3C), 4.3 (3C). ESI-MS: *m/z* $[M+Na]^+$ 1114.6.

7-*O*-triethylsilyl-10-propionyldocetaxel (**17**)

2',7-*O*-di(triethylsilyl)-10-propionyldocetaxel (**16**) (19mg, 0.017mmol) was dissolved in acetonitrile (2 ml), then pyridine (80μl, 0.99mmol) and HF(44μl, 0.99mmol) was added to the solution in turn, the mixture was stirred for 15 min at room temperature. 15 ml saturated aq $NaHCO_3$ solution was added, the aqueous layer was extracted with EtOAc (3×50mL), the organic layer was combined, then dried over anhydrous Na_2SO_4, and filtered, the filtrate was concentrated in vacuo, and the residual pyridine was removed by forming azeotropic mixture with cyclohexane. The residue was purified by PTLC (Petroleum ether/EtOAc = 3/1) to afford white solids 7-*O*-triethylsilyl-10-propionyldocetaxel (**17**) (15.6mg, 91.8%).

^{1}H NMR (300MHz, $CDCl_3$): δ 8.10 (2H, d, *J* = 7.5 Hz, *o*-Bz), 7.61 (1H, t, *J* = 7.5 Hz, *p*-Bz), 7.49 (2H, t, *J* = 7.5 Hz, *m*-Bz), 7.30-7.43 (5H, m, Ph), 6.47 (1H, s, H-10), 6.19 (1H, t, *J* = 8.7 Hz, H-13), 5.68 (1H, d, *J* = 6.9 Hz, H-2), 5.39 (1H, d, *J* = 9.6 Hz, BocN<u>H</u>), 5.26 (1H, d, *J* = 8.7 Hz, H-3'), 4.92 (1H, d, *J* = 7.8 Hz, H-5), 4.63 (1H, brs, H-2'), 4.45 (1H, dd, *J* = 6.6, 10.5 Hz, H-7), 4.30 (1H, d, *J* = 8.4 Hz, H-20), 4.16 (1H, d, *J* = 8.4 Hz, H-20), 3.82 (1H, d, *J* = 7.2 Hz, H-3), 2.39-2.58 (3H, m, H-6 + OCOC<u>H</u>$_2$CH$_3$), 2.36 (3H, s, OAc-4), 2.29 (2H, d, H-14×2), 2.04 (3H, s, Me-18), 1.84-1.99 (1H, m, H-6), 1.69 (3H, s, Me-19), 1.34 (9H, s, Me-<u>Boc</u>), 1.24 (3H, s, Me-17), 1.21 (3H, d, *J* = 6.5 Hz, OCOCH$_2$C<u>H</u>$_3$), 1.19 (3H, s, Me-16), 0.90 (9H, t, *J* = 7.5 Hz, SiCH$_2$C<u>H</u>$_3$), 0.53-0.64 (6H, m, SiC<u>H</u>$_2$CH$_3$). ^{13}C

NMR (125MHz, $CDCl_3$): δ 201.8, 172.8, 172.6, 170.2, 167.1, 155.3, 139.8, 138.4, 133.9, 133.6, 130.2 (2C), 129.2, 128.8 (2C), 128.6 (2C), 128.0, 126.8 (2C), 84.2, 81.2, 80.2, 78.8, 76.6, 74.8 (2C), 73.5, 72.5, 72.3, 58.5, 56.1, 46.8, 43.3, 37.2, 35.3, 28.2 (3C), 27.6, 26.5, 22.7, 21.1, 14.3, 10.0, 9.2, 6.7 (3C), 5.3 (3C). ESI-MS: *m/z* $[M+Na]^+$ 1000.5.

2'-*O*-(tert-butyldimethylsilyl)-7-*O*-triethylsilyl-10-propionyldocetaxel (**18**)

7-*O*-triethylsilyl-10-propionyldocetaxel (**17**) (89mg, 0.091mmol) was dissolved in DMF (0.5 ml), then imidazole (43.4mg, 0.64mmol) and tert-butyldimethylsilyl chloride (68.5mg, 0.46mmol) was added to the solution in turn, the mixture was stirred for 5 h at room temperature. 30 ml saturated aq NaCl solution was added, the aqueous layer was extracted with EtOAc (3×50mL), the organic layer was combined, then dried over anhydrous Na_2SO_4, and filtered, the filtrate was concentrated to dryness in vacuo, The residue was purified on silica gel by column chromatography (Petroleum ether/EtOAc = 4/1) to afford white solids, 2'-*O*-(tert-butyldimethylsilyl)-7-*O*-triethylsilyl-10-propionyldocetaxel (**18**) (91mg, 91.5%).

^{1}H NMR (500MHz, $CDCl_3$): δ 8.12 (2H, d, *J* = 7.5 Hz, *o*-Bz), 7.58 (1H, t, *J* = 7.0 Hz, *p*-Bz), 7.48 (2H, t, *J* = 7.0 Hz, *m*-Bz), 7.37 (2H, t, *J* = 7.5 Hz, Ph), 7.26-7.30 (3H, m, Ph), 6.49 (1H, s, H-10), 6.31 (1H, t, *J* = 9.0 Hz, H-13), 5.71 (1H, d, *J* = 7.5 Hz, H-2), 5.43 (1H, d, *J* = 8.5 Hz, BocN<u>H</u>), 5.35 (1H, brs, H-3'), 4.96 (1H, d, *J* = 9.5 Hz, H-5), 4.54 (1H, s, H-2'), 4.49 (1H, dd, *J* = 7.0, 11.0 Hz, H-7), 4.31 (1H, d, *J* = 8.5 Hz, H-20), 4.16 (1H, d, *J* = 8.0 Hz, H-20), 3.86 (1H, d, *J* = 7.2 Hz, H-3), 2.57 (3H, s, OAc-4), 2.38-2.55 (4H, m, H-6 + H-14 + OCOC<u>H</u>$_2$CH$_3$), 2.13-2.17 (1H, m, H-14), 2.02 (3H, s, Me-18), 1.89-1.93 (1H, m, H-6), 1.70 (3H, s, Me-19), 1.30 (9H, s, Me-<u>Boc</u>), 1.20-1.24 (9H, m, Me-17 + Me-16 + OCOCH$_2$C<u>H</u>$_3$), 0.93 (9H, t, *J* = 7.5 Hz, SiCH$_2$C<u>H</u>$_3$), 0.74 (9H, s,

SiC($C\underline{H}_3$)$_3$), 0.55-0.63 (6H, m, SiC$\underline{H}_2$CH$_3$), -0.10 (3H, s, SiC$\underline{H}_3$), -0.32 (3H, s, SiC$\underline{H}_3$). ^{13}C NMR (125MHz, $CDCl_3$): δ 201.8, 172.6, 171.4, 170.1, 167.1, 162.5, 155.1, 140.5, 138.9, 133.5 (2C), 130.2 (2C), 129.2, 128.6 (2C), 128.5, 127.6, 126.4 (2C), 84.2, 81.1, 79.9, 79.0, 76.5, 75.6, 75.0, 74.7, 72.2, 71.2, 58.3, 56.6, 46.6, 43.3, 37.2, 35.4, 28.1 (3C), 27.6, 26.5, 25.4 (3C), 23.0, 21.6, 18.1, 14.1, 10.1, 9.1, 6.7 (3C), 5.3 (3C), -5.4, -6.0. ESI-MS: *m/z* $[M+Na]^+$ 1114.7.

2'-*O*-(tert-butyldimethylsilyl)-2-debenzoyl-7-*O*-triethylsilyl-10-propionyldocetaxel (**19**)

2'-*O*-(tert-butyldimethylsilyl)-7-*O*-triethylsilyl-10-propionyldocetaxel (**18**) (50mg, 0.046mmol) was dissolved in dried CH_2Cl_2 (2.4 ml), the solution obtained was cooled to -28°C in ice-methanol bath, then the benzyltrimethylammonium hydroxide (45μl, 40% w/w in methanol) was added to the solution, the mixture was stirred for 15 min at -28°C. 10 ml precooled saturated aq NH_4Cl solution was added, the aqueous layer was extracted with EtOAc (3×30mL), the organic layer was combined, then dried over anhydrous Na_2SO_4, and filtered, the filtrate was concentrated to dryness in vacuo, The residue was purified on silica gel by column chromatography (Petroleum ether/EtOAc/acetone = 8/3/1) to afford white solid 2'-*O*-(tert-butyldimethylsilyl)-2-debenzoyl-7-*O*-triethylsilyl-10-propionyldocetaxel (**19**) (24mg, 53.1%) and recover the starting material (**18**) (12mg, 24%)

^{1}H NMR (400MHz, $CDCl_3$): 7.34 (2H, t, *J* = 7.2 Hz, *m*-Ph), 7.28 (1H, t, *J* = 6.8 Hz, *p*-Ph), 7.20 (2H, t, *J* = 7.2 Hz, *o*-Ph), 6.41 (1H, s, H-10), 6.23 (1H, t, *J* = 8.8 Hz, H-13), 5.48 (1H, d, *J* = 10.0 Hz, BocN$\underline{H}$), 5.22 (1H, d, *J* = 9.6 Hz, H-3'), 4.97 (1H, d, *J* = 8.8 Hz, H-5), 4.65 (1H, d, *J* = 8.2 Hz, H-20), 4.61 (1H, d, *J* = 8.2 Hz, H-20), 4.45 (1H, s, H-2'), 4.41 (1H, dd, *J* = 6.8, 10.4 Hz, H-7), 3.94 (1H, d, *J* = 7.2 Hz, H-3), 3.46 (1H, d, *J* = 6.8 Hz, H-3), 2.37-2.53 (3H, m, H-6 +

OCOC$\underline{H}_2$CH$_3$), 2.42 (3H, s, OAc-4), 2.17-2.24 (1H, m, H-14), 2.04-2.10 (1H, m, H-14), 1.96 (3H, s, Me-18), 1.89-1.92 (1H, m, H-6), 1.65 (3H, s, Me-19), 1.42 (9H, s, Me-Boc), 1.22 (3H, s, Me-17), 1.21 (3H, d, *J* = 8.0 Hz, OCOCH$_2$C$\underline{H}_3$), 1.12 (3H, s, Me-16), 0.91 (9H, t, *J* = 8.0 Hz, SiCH$_2$C$\underline{H}_3$), 0.73 (9H, s, SiC(C$\underline{H}_3$)$_3$), 0.53-0.60 (6H, m, SiC$\underline{H}_2$CH$_3$), -0.13 (3H, s, SiC$\underline{H}_3$), -0.34 (3H, s, SiC$\underline{H}_3$). ESI-MS: *m/z* $[M+Na]^+$ 1010.7.

2'-*O*-(tert-butyldimethylsilyl)-2-debenzoyl-2-*m*-N_3-benzoyl-7-*O*-triethylsilyl-10-propionyldocetaxel (**20**)

The *m*-N_3-benzoic acid (20.6mg, 0.16mmol) was dissolved in dried toluene (0.8 ml), dicyclohexylcarbodiimide (33.0mg, 0.16mmol) and 4-pyrrolidinopyridine (2.4mg, 0.016mmol) was added to the solution, the obtained mixture was stirred for 5 min at room temperature, then 2'-*O*-(tert-butyldimethylsilyl)-2-debenzoyl-7-*O*-triethylsilyl-10-propionyl-10-deacetyldocetaxel (**29**) (16mg, 0.016mmol) was added, and the mixture was stirred for 12 h at 65°C. The reaction mixture was diluted with EtOAc(10 mL), filtered through a pad of Celite, and the Celite was washed up with EtOAc (10 mL). The filtrate was concentrated in vacuo to dryness. The residue was purified with PTLC. Although it was found that the products were contaminated with *N,N*-dicyclohexylurea by ^{1}HNMR, they were subjected to desilylation without further purification.

2-debenzoyl-2-*m*-N_3-benzoyl-10-propionyl-docetaxel (**CTX-40)**

2'-*O*-(tert-butyldimethylsilyl)-2-debenzoyl-2-*m*-N_3-benzoyl-7-*O*-triethylsilyl-10-propionyldocetaxel (**20**) obtained in the above step was dissolved in acetonitrile (1 ml), then pyridine (0.533ml, 6.60mmol) and HF (0.293ml, 6.60mmol) was added to the solution in turn, the mixture was reacted 15 min at

room temperature until the protecting groups were removed completely. 50 ml EtOAc was added, the organic layer was washed with saturated aq $NaHCO_3$ solution(2×20mL), the aqueous layer was extracted with EtOAc (3×50mL), the organic layer was combined, then dried over anhydrous Na_2SO_4, and filtered, the filtrate was concentrated in vacuo, and the residual pyridine was removed by forming azeotropic mixture with cyclohexane. The residue was purified on silica gel by column chromatography (*n*-hexane/acetone = 2/1) to afford white solids **CTX-40** (7.6mg, 51%, two steps).

^{1}H NMR (300MHz, $CDCl_3$): δ 7.89 (1H, d, *J* = 7.8 Hz, H-6 of *m*-N_3-Bz), 7.81 (1H, s, H-2 of *m*-N_3-Bz), 7.48 (1H, t, *J* = 7.8 Hz, H-5 of *m*-N_3-Bz), 7.22-7.43 (6H, m, H-4 of *m*-N_3-Bz + Ph), 6.30 (1H, s, H-10), 6.21 (1H, t, *J* = 8.7 Hz, H-13), 5.66 (1H, d, *J* = 6.9 Hz, H-2), 5.34 (1H, d, *J* = 9.0 Hz, BocN<u>H</u>), 5.24 (1H, d, *J* = 8.1 Hz, H-3'), 4.97 (1H, dd, *J* = 1.8, 9.3 Hz, H-5), 4.60 (1H, s, H-2'), 4.42 (1H, dd, *J* = 6.6, 11.1 Hz, H-7), 4.32 (1H, d, *J* = 8.1 Hz, H-20), 4.16 (1H, d, *J* = 8.4 Hz, H-20), 3.82 (1H, d, *J* = 6.9 Hz, H-3), 3.32 (1H, brs, O<u>H</u>), 2.46-2.63 (3H, m, H-6 + OCOC<u>H</u>$_2$CH$_3$), 2.38 (3H, s, OAc-4), 2.30 (2H, d, *J* = 9.0 Hz, H-14), 1.85 (3H, d, *J* = 1.2 Hz, Me-18), 1.83-1.93 (1H, m, H-6), 1.67 (3H, s, Me-19), 1.33 (9H, s, Me-<u>Boc</u>), 1.26 (3H,s, Me-17), 1.24 (3H, t, *J* = 7.5 Hz, OCOCH$_2$C<u>H</u>$_3$), 1.45 (3H, s, Me-16). ^{13}C NMR (125MHz, $CDCl_3$): δ 203.7, 174.6, 173.1, 170.2, 166.1, 155.3, 142.3, 140.8, 138.3, 133.0, 131.0, 130.2, 128.9 (2C), 128.1, 126.8 (3C), 124.4, 120.2, 84.5, 81.1, 80.2, 79.1, 76.4, 75.4, 75.3, 73.5, 72.5, 72.2, 58.6, 56.2, 45.6, 43.2, 35.5, 35.4, 28.1 (3C), 27.5, 26.7, 22.6, 21.9, 14.9, 9.5, 9.0. ESI-MS: *m/z* $[M+Na]^+$ 927.4, $[M+K]^+$ 943.4. HRMS (ESI, $[M+Na]^+$) Found 927.3641 Calc. for $C_{46}H_{56}N_4O_{15}Na$ 927.3640.

For the synthesis and characterizations of CTX-1 to -10, -28 to -36, -41, please refer to (Yang et al. 2007).

For the synthesis and characterizations of CTX-22, -23, -26 and -27, please refer to the supplemental materials in (Wang et al. 2007).

Cell Biology Studies

LoVo and P-gp-overexpressing LoVo-Dox (Grandi et al. 1986) cells were cultured as previously described (Buey et al. 2005, Buey et al. 2007, de Ines et al. 1994). LoVo-Dox cells were maintained in 1 μg/ml doxorubicin and were grown in the absence of the drug for a week prior to the experiments.

Molecular Modelling

For the molecular dynamics (MD) simulations of the ligands (TXL, DXL and CTX-40) each molecule was immersed in a cubic box of ~1350 TIP3P water molecules under periodic boundary conditions (PBC) for 50 ns. Five structures representative of the most populated conformations for each ligand were selected for the automated docking study.

Default settings were used in AutoDock, except for number of runs, population size, and maximum number of energy evaluations, which were fixed at 250, 100, and 250000, respectively. Rapid intra- and intermolecular energy evaluation of each configuration was achieved by having the receptor's atomic affinity potentials for aliphatic and aromatic carbon, oxygen, nitrogen, and hydrogen atoms precalculated in a three-dimensional grid with a spacing of 0.375 Å. A distance-dependent dielectric function was employed in the computation of electrostatic interactions.

For the MD simulations of the complexes, the GDP-bound β-tubulin molecule from PDB entry 1JFF with the inhibitor (DXL, TXL or CTX40) docked in the TXL-binding site was neutralized with 8 sodium ions and immersed in a truncated octahedron of ~8500 TIP3P water molecules under PBC. The

GDP-tubulin interactions were improved on the basis of the contacts observed in the crystal structure of human γ-tubulin solved in complex with 5'-guanosine-diphosphate-monothio-phosphate (PDB id. 1Z5V) at 2.71 Å resolution (Aldaz et al. 2005). Solvent molecules and counterions were initially relaxed by energy minimization and allowed to redistribute around the positionally restrained solute (105 kJ mol^{-1} $Å^{-2}$) during 50 ps of MD at constant temperature (300 K) and pressure (1 atm). These initial harmonic restraints were gradually reduced in a series of progressive energy minimizations until they were completely removed. The resulting systems were heated again from 100 to 300 K during 20 ps in the absence of any restraints, using an integration step of 2 fs and updating the nonbonded pair list every 20 steps. The simulations then run for 10 ns during which a 21 kJ mol^{-1} $Å^{-2}$ restraint was applied on the Cα atoms of residues 1-211, 221-268, 286-353, and 364-426, which make up the α-helices of β-tubulin. At the end of this period, each system was gradually cooled down from 300 K to 273 K over a time length of 2 ns and finally subjected to energy minimization until the root-mean-square deviation of the energy gradient was less than 4.2 kJ mol^{-1} $Å^{-1}$.

The coordinates of the final complex between β-tubulin–GDP and CTX-40 are provided as supplemental data (TUBCTX40.pdb). Intermolecular van der Waals interaction energies between CTX-40 and individual residues were calculated with the ANAL module of AMBER, whereas the solvent-corrected electrostatic interaction was computed using the DelPhi program, as described in (Perez et al. 1998).

NMR Experiments

TR-NOESY experiments were performed with mixing times of 50 and 200 ms. To fit the experimental TRNOE intensities, off-rate constants between

10–500 s^{-1} were tested. Optimal agreement was achieved for k_{off} values ranging from 15 to 150 s^{-1}.

The fitting of the theoretical NOE values to the experimental NOE ones was performed using the CORCEMA programme, which considers a full relaxation matrix for the ligand protons in the presence of exchange between the free and bound forms. Thus, different average rotational correlation times (from 10 to 1000 ns) were used to fit the data and 100 ns was found to provide the best match between the experimental and expected values. It has to be considered that this value is in fact an "effective" correlation time, which contains contributions from global and internal motions.

Supplemental References

1. Aldaz, H., Rice, L.M., Stearns, T., and Agard, D.A. (2005). Insights into microtubule nucleation from the crystal structure of human gamma-tubulin. Nature 435, 523-527.
2. Buey, R.M., Barasoain, I., Jackson, E., Meyer, A., Giannakakou, P., Paterson, I., Mooberry, S., Andreu, J.M., and Diaz, J.F. (2005). Microtubule interactions with chemically diverse stabilizing agents: Thermodynamics of binding to the paclitaxel site predicts cytotoxicity. Chem Biol 12, 1269-1279.
3. Buey, R.M., Calvo, E., Barasoain, I., Pineda, O., Edler, M.C., Matesanz, R., Cerezo, G., Vanderwal, C.D., Day, B.W., Sorensen, E.J., Lopez, J.A., Andreu, J.M., Hamel, E., and Diaz, J.F. (2007). Cyclostreptin binds covalently to microtubule pores and lumenal taxoid binding sites. Nat Chem Biol 3, 117-125.
4. Buey, R.M., Diaz, J.F., Andreu, J.M., O'Brate, A., Giannakakou, P., Nicolaou, K.C., Sasmal, P.K., Ritzen, A., and Namoto, K. (2004). Interaction of Epothilone Analogs with the Paclitaxel Binding Site; Relationship between Binding Affinity, Microtubule Stabilization, and Cytotoxicity. Chem Biol 11, 225-236.
5. de Ines, C., Leynadier, D., Barasoain, I., Peyrot, V., Garcia, P., Briand, C., Rener, G.A., and Temple, C., Jr. (1994). Inhibition of microtubules and cell cycle arrest by a new 1-deaza-7,8-dihydropteridine antitumor drug, CI 980, and by its chiral isomer, NSC 613863. Cancer Res 54, 75-84.

6. Diaz, J.F., and Andreu, J.M. (1993). Assembly of purified GDP-tubulin into microtubules induced by taxol and taxotere: reversibility, ligand stoichiometry, and competition. Biochemistry 32, 2747-2755.

7. Grandi, M., Geroni, C., and Giuliani, F.C. (1986). Isolation and characterization of a human-colon adenocarcinoma cell-line resistant to doxorubicin. Brit J Cancer 54, 515-518.

8. Lowe, J., Li, H., Downing, K.H., and Nogales, E. (2001). Refined structure of alpha beta-tubulin at 3.5 Å resolution. J Mol Biol 313, 1045-1057.

9. Perez, C., Pastor, M., Ortiz, A.R., and Gago, F. (1998). Comparative binding energy analysis of HIV-1 protease inhibitors: incorporation of solvent effects and validation as a powerful tool in receptor-based drug design. J Med Chem 41, 836-852.

10. Wang, L., Alcaraz, A.A., Matesanz, R., Yang, C.G., Barasoain, I., Díaz, J.F., Li, Y.Z., Snyder, J.P., and Fang, W.S. (2007). Synthesis, biological evaluation, and tubulin binding poses of C-2a sulfur linked taxol analogues. Bioorg Med Chem Lett 17, 3191-3194.

11. Yang, C., Barasoain, I., Li, X., Matesanz, R., Liu, R., Sharom, F.J., Diaz, J.F., and Fang, W. (2007). Overcoming Tumor Drug Resistance Mediated by P-glycoprotein Overexpression with high–affinity taxanes: A SAR study of C-2 Modified 7-Acyl-10-Deacetyl Cephalomannines. Chem Med Chem 2, 691-701.

ARTÍCULO II

"COMBINE-guided design of new C2- and C3'-substituted taxanes: tubulin binding affinities and quantitative structure-activity relationships"

(en preparación)

Se generaron modelos moleculares de los complejos de β-tubulina con 47 taxanos para los que se habían determinado experimentalmente las energías libres de unión. Para cada complejo se calcularon tanto las interacciones ligando-receptor descompuestas por residuo como las contribuciones electrostáticas a las desolvataciones de ligando y receptor. Las afinidades de unión de 44 compuestos se pudieron expresar mediante un sumatorio de términos energéticos adecuadamente ponderados y seleccionados por un método basado en la proyección a variables latentes y mínimos cuadrados parciales (análisis COMBINE). Los ligandos de los tres complejos que no consiguieron ajustarse a esta correlación se caracterizan por presentar una contribución entrópica muy desfavorable a la energía libre de unión. Este comportamiento anómalo puede reflejar la importancia de la preorganización para la unión a la tubulina, ya que la conformación preferida de estos tres taxanos en disolución difiere significativamente de la conformación unida, como demuestran experimentos de RMN y cálculos de dinámica molecular. Con el fin de validar y ampliar el modelo COMBINE se sintetizaron y ensayaron nuevos taxanos modificados en C2 y en la posición C3' del sustituyente en C13. El modelo actualizado para 50 taxanos conservó su robustez estadística y las relaciones cuantitativas estructura-actividad así derivadas mejoraron nuestra comprensión de los factores determinantes de la afinidad para este tipo de agentes antitumorales.

COMBINE-guided design of new C2- and C3′-substituted taxanes: tubulin binding affinities and quantitative structure-activity relationships

Claire Coderch,[1] Yong Tang,[3] Shu-En Zhang,[3] Yun-Tao Ma,[3] Wang Shaorong, [3] Javier Klett,[2] Antonio Morreale,[2] Ruth Matesanz,[4] Benet Pera, [4] Angeles Canales,[4] Jesús Jiménez-Barbero,[4] José-Fernando Díaz,[4] Wei-Shuo Fang,[3*] and Federico Gago[1*]

[1] Departamento de Farmacología, Universidad de Alcalá, Alcalá de Henares, E-28871 Madrid, Spain.

[2] Unidad de Bioinformática, Centro de Biología Molecular Severo Ochoa (CSIC/UAM), Campus de Cantoblanco UAM, E-28049 Madrid, Spain.

[3] Institute of Materia Medica, Chinese Academy of Medical Sciences, 100050 Beijing, China.

[4] Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Campus Complutense UCM, E-28040 Madrid, Spain.

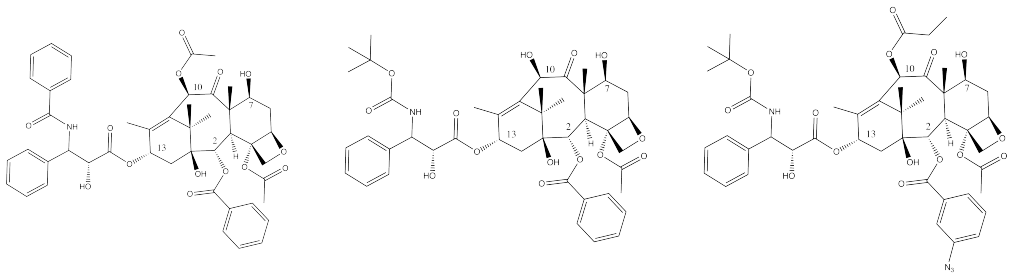
Corresponding authors: Dr. Wei-Shuo Fang (wfang@imm.ac.cn) and Prof. Federico Gago (e-mail: federico.gago@uah.es).

Running title: COMBINE-guided design of new taxanes.

ABSTRACT: Molecular models were built for the complexes of β-tubulin with 47 taxanes for which binding free energies have been experimentally measured. Residue-based ligand-receptor interaction energies were calculated in each complex as well as the electrostatic contributions to the desolvation of both ligand and receptor. By assigning weights to a selection of these energy terms, a projection to latent structures regression method (COMBINE) succeeded in providing a good prediction of the binding affinities for all but three compounds. The fact that these three outliers display large unfavorable entropic contributions to the binding free energies highlights a limitation of this approximation and may attest to the importance of taxane preorganization for tubulin binding, a possibility that is in good accord with results from molecular dynamics simulations. Two C2- and three C3'-substituted new analogues were synthesized and tested for external validation and model updating. The expanded COMBINE model conserved its robustness and yielded receptor-based quantitative structure-activity relationships that shed additional light on the determinants of tubulin-binding affinity for this class of antitumor agents and pave the way to further structural modifications.

KEYWORDS: Paclitaxel; taxanes; anticancer agents; tubulin; microtubules; molecular dynamics simulations; structure-activity relationships.

Paclitaxel (PXL, a.k.a. Taxol®) was originally discovered in the bark of the Pacific yew tree (***Taxus brevifolia***) and found to promote the assembly of microtubules (MT) and their stabilization by preventing tubulin depolymerization in dividing mammalian cells.[1] This increased stability results in the inhibition of the normal dynamic reorganization of the MT network that is essential for vital interphase and mitotic cellular functions. In clinical use since 1992, PXL became a first-in-class anti-mitotic drug that is currently indicated in the chemotherapy of a variety of human tumors including ovarian, breast and non-small cell lung carcinomas. Notwithstanding the remarkable potency of this antineoplastic agent, its limited availability from its natural source, its poor aqueous solubility and the emergence of resistance in cancer cells, partly due to overexpression of multidrug efflux transporters,[2] made it desirable from the start to obtain new analogues (taxanes) with an improved pharmacological profile.[3-7] This goal has been partially achieved with the clinical development and marketing of the semisynthetic derivatives docetaxel (DXL, ***Taxotere***®) and, more recently, cabazitaxel (***Jevtana***®), both of which partially overcome some of PXL's limitations. Nevertheless, structure-based design efforts in the search for novel taxanes[8] were hampered for many years because of difficulties associated with tubulin crystallization, the low affinity of PXL for unpolymerized tubulin subunits and the poor aqueous solubility of the drug.[9]



PXL **DXL** **CTX40**

Figure 1. Chemical structures of paclitaxel (PXL), docetaxel (DXL) and CTX40.

The first reliable three-dimensional model of tubulin in atomic detail was reported in 1998.[10] This long sought-after structure, deposited in the Protein Data Bank (PDB) as entry 1TUB, was obtained by electron crystallography of zinc-induced sheets of bovine tubulin bound to PXL. At this point it became clear that PXL binds to the lumenal side of the β subunit in the α,β-tubulin heterodimer. Subsequent refinement improved the resolution to 3.5 Å (PDB entry 1JFF)[11] but neither the acidic C-terminal tails nor the N-terminal loop including residues 35–60 in the α-subunit could still be traced accurately. Moreover, due to the limited certainty about side-chain orientation for many residues at the taxane-binding site (and elsewhere), the precise set of interactions between β-tubulin and bound PXL[12] (and related taxanes or other ligands such as the epothilones)[13] could not be entirely clarified from this structure.

The initial pharmacophore that was established for taxanes based on the so-called T-Taxol binding conformation[13-16] has been the foundation for several structure-guided modifications[17, 18] and structure-activity relationship (SAR) studies.[6, 15, 19, 20] Recent collaborative work from our laboratories resulted in the report of an extended series of 44 semisynthetic taxanes (CTX), which have in common the baccatin core present in PXL, DXL and cephalomannine (CPH) but additionally incorporate a variety of changes at positions C2, C7, C10 and C13.[21] A corollary of this endeavor was the finding of an optimal combination of substituents that yielded the high-affinity taxane CTX40 (Figure 1), the cytotoxicity of which is unaffected by P-glycoprotein (P-gp) overexpression. Furthermore, the binding of all of these compounds to glutaraldehyde-stabilized bovine MT was measured and an energy-refined binding mode for PXL, DXL and CTX40 to β-tubulin was proposed and then explored using molecular dynamics (MD) simulations in explicit solvent. Support for the conformation of bound DXL was provided by the excellent agreement obtained with data acquired from transfer nuclear Overhauser effect spectroscopy (TR-NOESY) measurements using MT-bound DXL.[21] The thermodynamic characterization of the binding of these taxanes to MT revealed not only a range of ~10 kcal·mol^{-1} in binding free energies (ΔG) but also distinct enthalpic (ΔH) and entropic (–TΔS) contributions. The existence of rather large enthalpy/entropy compensations[22-25] within this series (Figure 2)[21] attests to the difficulties that can be encountered when theoretical methods are used to try and rationalize the quantitative SAR (QSAR) for these ligands.[26]

In the following we explore the QSAR for the binding to β-tubulin of these taxanes using a chemometric tool known as

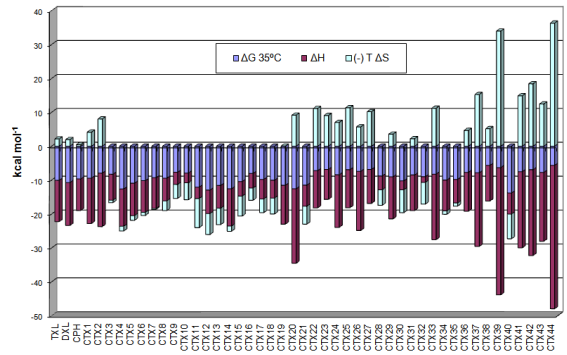


Figure 2. Thermodynamic parameters (kcal·mol^{-1}) for the binding of the taxanes studied to the PXL site in bovine β-tubulin, as reported in Table 1 of reference 21. Note that actual values correspond only to the length of the colored block, i.e. are not cumulative.

Comparative Binding Energy (COMBINE) analysis,[27] which was recently implemented in a user-friendly graphical interface (gCOMBINE) and released under a scientific/academic nonprofit and noncommercial license.[28] We show that robust cross-validated models are produced that correlate the ΔG values within the series (with only three exceptions) with a set of weighted selected interactions within the binding site. Additional support is gained for the proposed protonation state of His229,[21, 29] a residue that is shown to play a major role in modulating the affinity differences. The QSAR model was then externally challenged with several new C2- and C3'-modified taxanes that were synthesized and experimentally tested.

Methodology

Theoretical Methods. The refined structure of the α,β-tubulin heterodimer at 3.5 Å resolution (Protein Data Bank code 1JFF) was used for molecular modeling and automated docking of DXL and CTX40, as already described.[21] In brief, addition of missing hydrogen atoms and computation of the protonation state of ionizable groups in β-tubulin at pH 6.5 were carried out using the H++ Web server,[30] which relies on AMBER force-field parameters[31, 32] and finite difference solutions to the Poisson-Boltzmann equation. The GDP molecule was conserved and the protein side-chains in the nucleotide binding site were slightly reoriented so as to reproduce the same interactions that are observed in the 2.51 Å resolution X-ray crystal structure of γ-tubulin bound to 5'-guanosine-diphosphate-monothio-phosphate[33]

(PDB entry 1ZSV). Following energy minimization, MD simulations at 300 K were run for these two ligand-protein complexes during 10 ns under periodic boundary conditions in truncated octahedra each containing ~10,000 TIP3P water molecules and 11 randomly placed sodium ions to neutralize the total charge of the systems. After this time, the temperature was gradually decreased from 300 K to 273 K over 1200 ps. Energy minimization of the resulting "cooled" structures employing 5000 cycles of steepest descent followed by 4000 cycles of conjugate gradient provided the "frozen" templates that were used as representative structures to build the remaining complexes for all the CTX derivatives. The need for two templates arose from the fact that the C2 substituent can be placed slightly differently inside the hydrophobic binding pocket, depending on the absence or presence of substituents on the phenyl ring, in such a way that the stacking interaction of the latter with the side chain of His229 is also different. Therefore the bound conformation of DXL was used as the template for model building those taxanes that have no substituent on this ring whereas those bearing bulky substituents on this phenyl were modeled from CTX40. Of note, during the MD simulations we observed the long residence times of two water molecules that bridged a hydrogen bonding interaction between the carbonyl oxygen present in the C2 linker and the side-chains of Asp226 and His229. Consequently, we decided to keep these two water molecules (WAT1 and WAT2) as part of the ligand-receptor complexes (Figure 3). These two extra residues were later shown to be instrumental for an accurate prediction of taxanes CTX22–27, which have in common the presence of non-ester C2 linkers.

The remaining taxanes bearing modifications on the C2, C7, C10 and C13 substituents were built within the molecular graphics program PyMOL[28] using as templates the common baccatin cores of the refined CTX40 and DXL structures in their respective complexes with β-tubulin. The geometries of the isolated new analogues were fully optimized at the Hartree-Fock level using the Gaussian 03 program[34] and the 3-21G basis set. Charge distributions for all the ligands were then obtained by fitting the quantum mechanically calculated (HF/6-31G*//HF/3-21G*) molecular electrostatic potential (MEP) to a RESP point charge model.[35] Covalent and non-bonded parameters for the inhibitors were derived, by analogy or through interpolation, from those already present in the AMBER[31, 32] database (ff03). The tubulin-taxane complexes were refined by energy minimization using only the steepest descent algorithm until the root-mean-square (rms) value of the potential energy gradient was below 0.1 kcal·mol^{-1}·Å^{-1}.

The set of 47 refined ligand-receptor complexes (including WAT1 and WAT2) was then used as input to the gCOMBINE program,[28] which calculates all the van der Waals (AMBER force field) and electrostatic interaction energies between each ligand and every protein residue. The electrostatic contributions were calculated using Goodford's implementation of the images method[36] and a uniform dielectric constant of 4.0 Debye. The desolvation energies of both the protein and the ligand, which were used as additional descriptors (a.k.a. external variables), were calculated. To this end the binding process is described as consisting of first desolvating the apposing surfaces of both ligand and receptor and then letting the charges of the two molecules interact.[37] All these calculations were performed by numerically solving the linear Poisson-Boltzmann equation using the finite difference method as implemented in program DelPhi.[38] Atomic charges and radii were assigned according to the AMBER force field (ff03). Each complex was immersed in a cubic box occupying

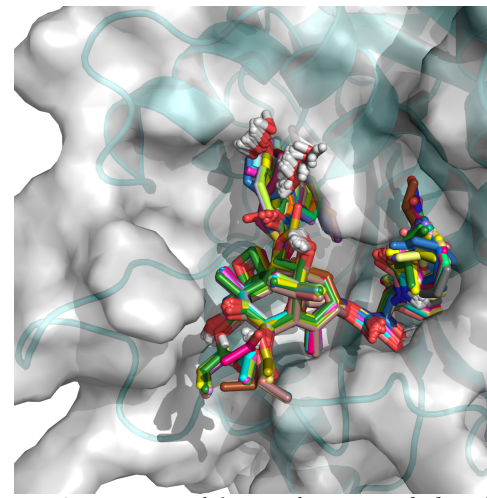


Figure 3. Superposition of the set of taxanes studied inside the PXL-binding site of β-tubulin. For simplicity, only the cartoon representation of the protein (from the DXL-tubulin complex) enveloped by a semitransparent grey van der Waals surface is shown.

65% of the total volume, with a grid spacing of 0.5 Å. The solute dielectric constant was set to 4, while the solvent dielectric medium was set to 80. The dielectric boundary was calculated using a solvent probe radius of 1.4 Å. A minimum separation of 11 Å was allowed between any solute atom and the box walls. The potentials at the grid points delimiting the box were calculated analytically by treating each charge atom as a Debye–Hückel sphere.

The data matrix containing the computed energy components (X variables) was then subjected to a multivariate statistical analysis using projection to latent structures (a.k.a. partial least squares, PLS)[39] in order to find a correlation with the experimental binding free energies (Y variable). For cross-validation purposes, a group consisting of 5 randomly chosen complexes was excluded from the training set and used as a test set, as described,[28] and this procedure was repeated a total of 10 times. This provides a more rigorous estimation of the robustness of the model than the very conservative leave-one-out method. The dimensionality of the model (i.e. the number of extracted latent variables, LV) was chosen as that corresponding to the peak cross-validated regression coefficient (q^2) value.

To study the conformational preferences of these taxanes in aqueous solution, CTX40, CTX42, CTX43 and CTX44 were selected as representative compounds. Each of them was extracted from its respective complex and immersed in a cubic box of approximately 1,400 TIP3P water molecules. Following energy minimization, a 100-ns MD simulation at 300 K was carried out for each compound. The resulting trajectories were processed with the ***ptraj*** module of AMBER to monitor the rms deviations from the initial T-taxol geometry. Desolvation energies of representative structures of each conformation were obtained using the Generalized Born method that is implemented in the AMBER 10 distribution. The relative energies of each conformer were used to calculate the probability of the microstates ($\boldsymbol{p}$) as defined by a Boltzmann distribution:

$$\boldsymbol{p}_i = \boldsymbol{n}\, e^{-Ei/RT} / \Sigma_i \boldsymbol{n}\, e^{-Ei/RT}$$

where $\boldsymbol{n}$ is the number of structures belonging to each cluster $\boldsymbol{i}$, $\boldsymbol{E}_i$ is the energy of the average structure extracted from each cluster, $\boldsymbol{T}$ is

the temperature in Kelvin, ***R*** is the gas constant expressed in kcal·mol^{-1} and ***j*** is the total number of clusters.

The molecular graphics program PyMOL (v. 0.99rc6, DeLano Scientific, LLC, Palo Alto, CA)[28] was employed for molecular editing, visualization and representation.

NMR experiments. The samples of the MT-bound CTX42 were prepared using a 300 μM concentration of CTX42 and 20 μM of tubulin in D_2O, 10 mM KPi, 0.1 mM GMPCPP, 6 mM $MgCl_2$ pH 6.7. The tubulin samples were prepared by removing sucrose, Mg^{2+}, and H_2O from the storage buffer of a 20 mg sample of frozen tubulin using a two-step procedure by chromatography in a drained centrifuge column of Sephadex G-25 medium (6 × 1 cm) equilibrated in D_2O, 10 mM KPi, 10 μM GTP pH 7.0 in the cold, followed by a second chromatography using another Sephadex G-25 medium column (15 × 0.9 cm) equilibrated in D_2O, 10 mM KPi, pH 7.0. Tubulin was centrifuged, and its concentration was determined spectrophotometrically by employing an extinction coefficient of 107,000 M^{-1} cm^{-1} in 10 mM phosphate buffer containing 1% SDS. Tubulin was diluted to 20 μM and GMPCPP 0.1 mM and 6 mM $MgCl_2$ (final pH 6.7) were added prior to drug addition. The samples were then incubated at 37 °C for 30 min prior to measurement.

The increased solubility of CTX42 relative to the other amide-containing taxanes allowed preparation of a sample of this free compound in the same buffer described above. TR-NOESY information was acquired for CTX42 in the presence of MT with a mixing time of 200 ms at 310 K. 2D NOE in the rotating frame (ROESY) data were acquired for CTX42 in the free state because the NOESY cross peaks for the free molecule were essentially zero at 310 K. All the spectra were obtained in a Bruker 500 MHz spectrometer.

Synthetic Procedures. For C2 modification, the 2-ethyl baccatins were prepared from 10-deacetyl baccatin III (10-DAB) by a procedure similar to that described in the literature,[42] in which the 2-debenzoyl baccatin intermediates were reacted with mesyl chloride and then transformed into the C2 ethers after treatment with sodium phenolate. Subsequent reduction of the 13-oxo group, incorporation of the C13 side chain bearing the 3'N-Boc group, and further desilylation led to the final products CTX55 and CTX56. As it is known that chloroformates derived from tertiary alcohols bearing electron-donating groups (EDGs) are unstable, whereas those bearing electron-withdrawn groups (EWGs) are relatively stable,[40] ***p***-nitrobenzoyl carbonates were used instead of chloroformates during the carbamoylation of the 3'-NH_2 for some taxanes. For the syntheses of EWG-bearing analogs CTX57 and CTX60, the 3'-debenzoyl PXL was transformed into a carbamate by reacting with the corresponding chloroformates, whereas for EDG-bearing 3'N-modified taxanes CTX58 and CTX59, the 3'-debenzoyl PXL was transformed into a carbamate by reacting first with the corresponding ***p***-nitrobenzoyl carbonates [41], and then with the corresponding alcohols. By subsequent desilylation the 3'N-modified DXL analogues CTX57 to CTX60 were obtained.

Measurement of the Equilibrium Binding Constants and Cytotoxicity. Stabilized, moderately crosslinked MT were prepared from purified calf brain tubulin as reported before.[42] Binding constants and thermodynamic parameters were measured by displacement of the fluorescent taxane derivative Flutax-2 from the PXL-binding site in these MT.[43] Wild-type A2780 ovarian carcinoma cells and their P-gp-overexpressing counterparts A2780AD were cultured as described previously.[43] Cell proliferation was measured using a modified (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium (MTT) bromide assay, as previously reported.[6]

Results and Discussion

The 47 taxanes studied in our training set span a range of binding energies of ~10 kcal·mol^{-1} and differ from one another in the nature of the groups attached to the C2, C7, C10 and C13 positions of the common baccatin scaffold (Figure 3). Earlier energy decomposition analyses on a limited set of refined taxane-tubulin complexes highlighted the amino acids that are thought to be key for binding and helped to understand, albeit in a qualitative way, the observed SAR.[21] When the whole set of complexes was subjected to a COMBINE analysis, the PLS result showed the clearly outlying behavior of CTX42, CTX43 and CTX44, whose common characteristic is the presence of an amide linker in the C2 substituent instead of an ester, as in CTX21, CTX11, and CTX6, which are their most similar counterparts, respectively. That this sole modification is translated into a largely decreased binding free energy is in stark contrast with the facts that (i) the binding modes appear very similar, and (ii) no protein contacts are apparent in this region that could account for these differences, in contrast to what has already been described for C2 S-linked taxanes.[44] However, when these three grossly overpredicted compounds were removed from the training set, a robust cross-validated COMBINE model was obtained (r^2=0.94 and q^2=0.89 for the best model; average r^2=0.94±0.03 and q^2=0.87±0.02 for the 10 runs, using 5±1 LV). To assess whether the affinity of the most tightly binding compound, CTX40, could have been predicted using the remaining set of analogues and/or whether this molecule was having a high leverage on the model, we left it out and run gCOMBINE with the remaining 40 complexes. The quality of the resulting PLS equation (r^2=0.94 and q^2=0.86 for the best model; average r^2=0.94±0.03 and q^2=0.84±0.02 for the 10 runs, using 5±1 LV) proved to be similar to that obtained previously and the binding free energy of CTX40 was predicted with high accuracy (Figure 4). Of note, the

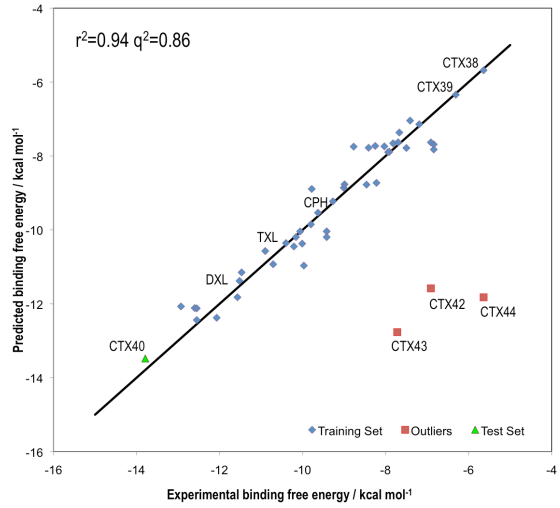


Figure 4. Correlation between the binding free energies calculated in the cross-validated COMBINE model and the experimental values for compounds in the training set (◆) upon exclusion of the three outliers (■). This model, with only 5 LV, accurately predicts the affinity of CTX40 (▲), which was not included in model derivation.

outstanding affinity of CTX40 was also predicted earlier on the basis of additivity of contributions from optimal substituents.[21]

The relative importance of the β-tubulin residues that best explain the differences in binding free energies, as judged from the assigned PLS pseudocoefficients (Supporting Information, Table S1), depends on the actual magnitude of the calculated interaction energies because these signed coefficients are used to properly weigh these terms.[27, 45] Since the energy values are usually negative (i.e. attractive), a positive coefficient will make the calculated binding free energy more negative, and vice versa. Thus, improved van der Waals interactions between the C13 moiety and both Val23 and Asp26, as well as between the baccatin core and the C2 substituent with Asp226, His229, Ala233, Phe272 and Leu275, are predicted to result in gains in binding affinity. On the contrary, the more the ligands interact with other residues, particularly Cys213, Leu217 and Leu219, the lower the affinity, which most likely reflects the existence of steric clash.[45]

From the electrostatics point of view, the importance of the interactions with His229, Arg278 and WAT2 is reflected in the positive coefficients assigned to the terms involving these two protein residues and one of the water molecules. Likewise, the negative PLS coefficient that modulates the electrostatic contribution to ligand desolvation strongly suggests that this factor plays an important role in the binding affinity differences, in contrast to the small positive coefficient that is assigned to the desolvation of the binding pocket. Moreover, mutation of this Arg to Ser confers resistance to PXL.[46, 47]

These computational studies were undertaken considering a pH of 6.5 for the calculation of the protonation state of ionizable groups in β-tubulin to mimic the experimental conditions.[21] At this pH the side-chain imidazole of His229 (found in the middle of helix 7 and positionally equivalent to Arg229 in α-tubulin) is likely to be protonated on both Nδ and Nε giving rise to (i) a MEP in the taxane-binding site that is quite distinct from that obtained with a neutral imidazole side chain,[21] and (ii) different hydrogen-bonding possibilities. For this reason, we also refined the whole set of complexes with a β-tubulin in which His229 was protonated exclusively on Nε. The corresponding COMBINE model (Supporting Information Figure S1) turned out to be of inferior quality (average r^2 and q^2 over 10 runs of 0.75±0.02 and 0.63±0.02, respectively, using 2±1 LV) and was unable to predict accurately the binding free energy of CTX40 when this compound was taken out of the training set. The largest change in the PLS pseudocoefficients affected precisely the electrostatic interaction energy with His229 (Supporting Information Table S2).

With a view to expanding the exploration of the taxane-binding pocket and validating our best COMBINE model with some external molecules, several modifications were introduced at the C2 (CTX55 and CTX56) and C3' positions (CTX57-CTX60) (Figure 5). The rationale for synthesizing CTX55 and CTX56 was to explore the effect of replacing the 2-OBz bearing the azide at the meta position[21] with an extended N-containing aromatic ring system (i.e. a quinoline or an isoquinoline, respectively) with a view to further improving the stacking interactions with the imidazole of His229. On the other hand, the replacement of one of the methyl groups in the 3'N-Boc ***tert***-butyl moiety with a small substituent, as in CTX57 to CTX60, attempted to assess the effect of putatively filling a protein pocket in the vicinity of the 3'N-Boc group.

The binding affinities of the new C2- and C3'-substituted analogs (Supporting Information Tables S3 and S4) were overpredicted, and CTX59 clearly behaved as an outlier (Figure 6).

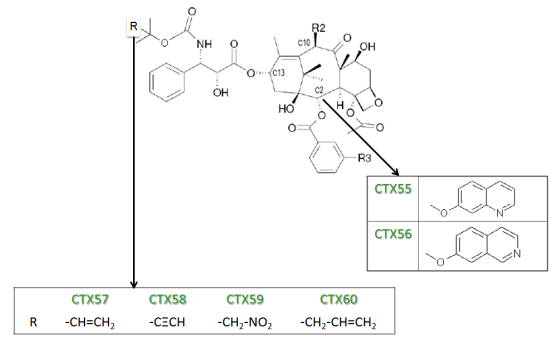


Figure 5. New synthesized taxanes showing the modification introduced at positions C2 and C13.

This result is likely a consequence of the more thorough exploration of the C2-binding pocket relative to the site harboring the C3' substituents because of the smaller chemical variation at the latter position within the original series. As discussed before, lead optimization is a cyclic process involving repeated rounds of compound design, synthesis and testing, and QSAR models derived at one stage need to be continually updated incorporating as much new information as possible.[27] Therefore, we added the new C3'-substituted analogues to the original training set and tried to predict the affinity of the C2-modified compounds. This updated model showed good figures of merit (r^2=0.93, q^2=0.75 for the best model and average r^2=0.94±0.1, q^2=0.73±0.1 over 10 runs using 6±1 LV) and was able to predict the binding free energies of CTX55 and CTX56 remarkably well (Figure 7). The PLS coefficients in this new model barely differed from those in the original one except for those assigned to Lys19 and Glu22 that became larger than before and of opposite signs for the van der Waals and electrostatic terms (Table1). Finally, when the set was expanded to contain all 53 taxanes and subjected to COMBINE analysis using random groups for cross-validation, a very similar high-quality model was obtained (r^2=0.93, q^2=0.80 for the best model and average r^2=0.92±0.1, q^2=0.77±0.1 over 10 runs using 5±1 LV).

Perhaps the most striking feature of our overall results is the outlying behavior (Figure 4) of taxanes CTX42, CTX43, and CTX44, which have in common not only an amide linker between the C2-substituent and the baccatin core, instead of an ester, but also a thermodynamic binding profile that is characterized by a large unfavorable entropic contribution to the Gibbs free energy of binding (Figure 2). It is usually agreed that solvation entropy changes originating from the release of water molecules brought about by the burial of a significant hydrophobic surface are favorable for binding[48] whereas the loss of conformational degrees of freedom by the inhibitor and/or some residues in the protein is unfavorable.[49, 50] Thus, a common goal in molecular design is to minimize ligand entropy loss by restraining conformational flexibility and promoting a configuration in the unbound state that optimally matches the binding site,[51] even though doubts have been cast on this assertion.[52] This "preorganization" may lead to ligands with affinities in the nanomolar range even in the presence of unfavorable binding enthalpies.[53] Conversely, favorable binding enthalpies can be largely offset by unfavorable entropic contributions, as clearly seen in the series presented here. In this case, although we have tried to take into account, at least in part, the cost of ligand and binding-site desolvation as well as some water-mediated interactions, we have neglected the configurational

entropies of the ligands, the water molecules and the protein.[25,52] Therefore, we speculated that CTX42, CTX43, and CTX44 might be stabilized in aqueous solution in a conformation that is not compatible with binding to the site. To test this hypothesis, NMR experiments were carried out in order to characterize the conformation of these derivatives both in the free and in the bound states (Supplementary Information Figure S2). In addition, we also performed MD simulations in explicit water to sample the conformational space of these amide-containing taxanes and also CTX40 as a representative example of an ester-containing taxane.

The conformation of MT-bound CTX42 was deduced by analysis of TR-NOESY cross peaks. The strong negative cross peaks observed for CTX42 in the presence of MT indicate binding of the taxane to the tubulin preparation (NOESY cross peaks were essentially zero for the free ligand). The fact that the TR-NOESY spectrum of CTX42 in the presence of MT shows an NOE pattern similar to that described for DXL[21] points to a similar conformation for both molecules in the bound state. Clear NOEs are observed between the C13 ***tert***-butoxy and C2 aromatic ring protons. The 4-OAc group also gives NOEs with both aromatic rings, and they are stronger with the protons of the phenyl ring on the C13 substituent which also gives an NOE contact with the methyl group at C12.

In contrast, no ROE contacts were detected between the C3' ***tert***-butoxy protons and those of the C2 aromatic ring in the ROESY of free CTX42. On the contrary, weak ROEs were observed between the 4-OAc group and the C13 phenyl ring protons whereas no ROEs were observed between this latter moiety and the C12 methyl group. All these data support the hypothesis of different conformations for CTX42 in the free and MT-bound states.

To compare with the NMR data obtained for the more water-soluble CTX42, MD simulations were carried out with CTX42, CTX43 and CTX44. For comparison, the most potent analog from the series (CTX40), whose binding is both enthalpically and entropically favored, was also simulated under the same conditions. The MD trajectories showed that the probabilities of the microstate corresponding to the bound conformation were 0.78 for CTX40 but only 0.0005, 0.0001, and 0.03, respectively, for CTX42, CTX43 and CTX44. These amide derivatives appeared to be locked in a hydrophobically collapsed conformation that was energetically more favorable by ~4.0 kcal·mol^{-1} for the first two and ~6.0 kcal·mol^{-1} for the third one (Supplementary Information Figure S3). This conformation is in agreement with the ROESY spectrum of free CTX42. Collectively, these results support the view that the entropic penalty associated to tubulin binding (Figure 2) may arise from an energetically unfavorable conformational change. On the contrary, for taxanes that are preorganized for binding to the site, e.g. CTX40, the ligand-dependent entropic component is likely to be favorable.

Conclusions

Robust QSAR models have been derived using the COMBINE methodology that support a well-defined orientation for a large series of taxanes in the PXL-binding site of β-tubulin and highlight the relative importance for binding affinity of a number of protein residues as well as the electrostatic contribution to ligand desolvation. Our results favor the view that His229 is doubly protonated on the imidazole ring and also that some water molecules get trapped in the binding process. When new molecules were synthesized and tested to enrich the chemical diversity of the series and to challenge the model we found it necessary to expand

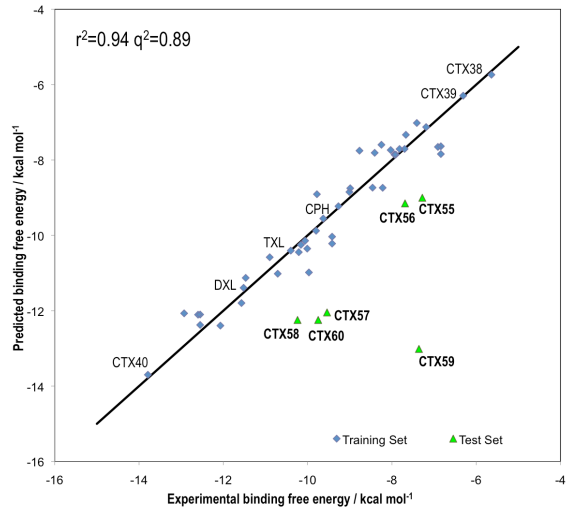


Figure 6. Correlation between the binding free energies calculated in the cross-validated COMBINE model and the experimental values for compounds in the training set (◆) and for the external set consisting of the newly synthesized taxanes (▲).

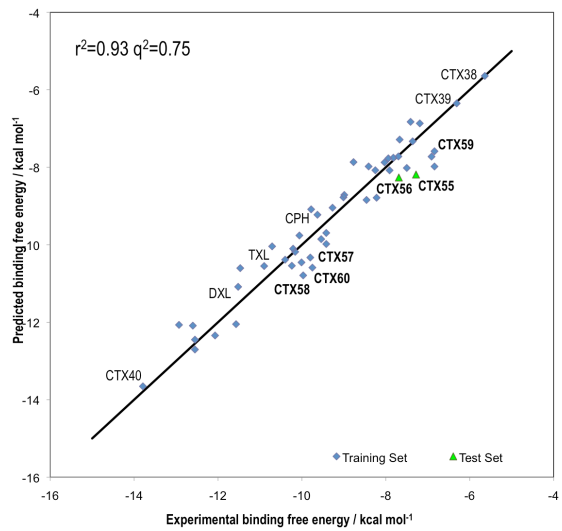


Figure 7. Correlation between calculated and experimental binding free energies in the updated COMBINE model that now includes in the training set (◆) all the new C3'-substituted derivatives and correctly predicts the external set (▲) consisting only of the new C2-substituted taxanes.

the training set to strengthen the predictive ability of the PLS model. A serious limitation was found when it was realized that the affinities of a few analogues whose binding is characterized by a large unfavorable entropic component could not be correlated with the calculated interaction and desolvation energies. To account for this anomalous behavior we hypothesized, on the basis of results from MD simulations in aqueous solution, that the most populated conformation in the free state is unsuitable for binding and is

Table 1. Selected PLS pseudocoefficients (absolute value ≥ |0.1|) for the amino acid residues that contribute the most to explaining the predicted binding free energy differences.

Residue	vdW	Ele	Residue	vdW	Ele
Lys19	-0.70	-0.68	Phe272	0.94	
Glu22	-0.97	0.28	Pro274	0.18	
Val23	0.45		Leu275	0.45	
Glu27	0.22	0.12	Thr276		0.20
Cys213	-0.18		Ser277	0.19	0.16
Leu217	-0.41		Arg278	-0.23	0.36
Leu219	-0.31		Gln282	-0.20	0.20
Asp226	0.20	-0.46	Arg320	-0.23	
Leu227	0.20		Pro360	0.1	
Hip229	0.43	0.61	Arg369	-0.10	0.36
Leu230	0.12		Gly370	-0.49	-0.25
Ala233	0.55	-0.14	WAT1		0.19
Gly237	-0.14		WAT2	-0.27	0.38
			DesolvL		-0.17

stabilized by hydrophobic collapse. NMR data obtained for the only representative taxane that could be studied tubulin-free in a water-DMSO mixture provided qualitative support to this hypothesis. Attempts are being made to use this recently acquired information to further expand the series and synthesize more analogues that, for example, incorporate the bridging water's hydrogen-bonding functionality into their own structure.

Acknowledgments

The authors thankfully acknowledge Prof. José Manuel Andreu (CIB, CSIC) for advice and useful discussions. C.C. enjoys an FPU grant awarded by the Spanish Ministry of Science and Innovation (AP2007-01225) and A. M. acknowledges Comunidad de Madrid for financial support through the AMAROUTO program to the Fundación Severo Ochoa. This work was financed by the National Natural Science Foundation of China (NSFC grant 30930108 to W.S.F.), the Spanish Ministerio de Ciencia e Innovación (BIO2008-04384 and BIO-2010-16351 to F.D.) and Comisión Interministerial de Ciencia y Tecnología (SAF2009-13914-C02-02 to F.G.), and Comunidad de Madrid (S-BIO/0214/2006 and S2010-BMD-2457 to F.D. and F.G.). The computer resources, technical expertise and assistance provided by the Barcelona Supercomputing Center (Centro Nacional de Supercomputación) are gratefully acknowledged.

Supporting Information. Supplementary figures, tables and all synthetic procedures are available free of charge via the Internet at http://pubs.acs.org.

References

1. Schiff, P. B.; Fant, J.; Horwitz, S. B. Promotion of microtubule assembly in vitro by taxol. ***Nature*** **1979,** 277, 665-7.

2. Breedveld, P.; Beijnen, J. H.; Schellens, J. H. Use of P-glycoprotein and BCRP inhibitors to improve oral bioavailability and CNS penetration of anticancer drugs. ***Trends Pharmacol. Sci.*** **2006,** 27, 17-24.

3. Ojima, I.; Slater, J. C.; Michaud, E.; Kuduk, S. D.; Bounaud, P. Y.; Vrignaud, P.; Bissery, M. C.; Veith, J. M.; Pera, P.; Bernacki, R. J. Syntheses and structure-activity relationships of the second-generation antitumor taxoids: exceptional activity against drug-resistant cancer cells. ***J. Med. Chem.*** **1996**, 39, 3889-96.

4. Wang, Y. F.; Shi, Q. W.; Dong, M.; Kiyota, H.; Gu, Y. C.; Cong, B. Natural taxanes: developments since 1828. ***Chem. Rev.*** **2011**.

5. Ojima, I.; Slater, J. C.; Kuduk, S. D.; Takeuchi, C. S.; Gimi, R. H.; Sun, C. M.; Park, Y. H.; Pera, P.; Veith, J. M.; Bernacki, R. J. Syntheses and structure-activity relationships of taxoids derived from 14 beta-hydroxy-10-deacetylbaccatin III. ***J. Med. Chem.*** **1997**, 40, 267-78.

6. Yang, C. G.; Barasoain, I.; Li, X.; Matesanz, R.; Liu, R.; Sharom, F. J.; Yin, D. L.; Diaz, J. F.; Fang, W. S. Overcoming tumor drug resistance with high-affinity taxanes: a SAR study of C2-modified 7-acyl-10-deacetyl cephalomannines. ***ChemMedChem*** **2007**, 2, 691-701.

7. Ojima, I.; Chen, J.; Sun, L.; Borella, C. P.; Wang, T.; Miller, M. L.; Lin, S.; Geng, X.; Kuznetsova, L.; Qu, C.; Gallager, D.; Zhao, X.; Zanardi, I.; Xia, S.; Horwitz, S. B.; Mallen-St Clair, J.; Guerriero, J. L.; Bar-Sagi, D.; Veith, J. M.; Pera, P.; Bernacki, R. J. Design, synthesis, and biological evaluation of new-generation taxoids. ***J. Med. Chem.*** **2008,** 51, 3203-21.

8. Geney, R.; Sun, L.; Pera, P.; Bernacki, R. J.; Xia, S.; Horwitz, S. B.; Simmerling, C. L.; Ojima, I. Use of the tubulin bound paclitaxel conformation for structure-based rational drug design. ***Chem. Biol.*** **2005**, 12, 339-48.

9. Suffness, M. ***Taxol: science and applications***. CRC Press: Boca Raton, Florida, 1995; p 426 p.

10. Nogales, E.; Wolf, S. G.; Downing, K. H. Structure of the alpha beta tubulin dimer by electron crystallography. ***Nature*** **1998**, 391, 199-203.

11. Lowe, J.; Li, H.; Downing, K. H.; Nogales, E. Refined structure of alpha beta-tubulin at 3.5 A resolution. ***J. Mol. Biol.*** **2001**, 313, 1045-57.

12. Sun, L.; Simmerling, C.; Ojima, I. Recent advances in the study of the bioactive conformation of taxol. ***ChemMedChem*** **2009**, 4, 719-31.

13. Forli, S.; Manetti, F.; Altmann, K. H.; Botta, M. Evaluation of novel epothilone analogues by means of a common pharmacophore and a QSAR pseudoreceptor model for taxanes and epothilones. ***ChemMedChem*** **2010**, 5, 35-40.

14. Snyder, J. P.; Nettles, J. H.; Cornett, B.; Downing, K. H.; Nogales, E. The binding conformation of Taxol in beta-tubulin: a model based on electron crystallographic density. ***Proc. Natl. Acad. Sci. U. S. A.*** **2001**, 98, 5312-6.

15. Kingston, D. G.; Bane, S.; Snyder, J. P. The taxol pharmacophore and the T-taxol bridging principle. ***Cell Cycle*** **2005**, 4, 279-89.

16. Alcaraz, A. A.; Mehta, A. K.; Johnson, S. A.; Snyder, J. P. The T-Taxol conformation. ***J. Med. Chem.*** **2006**, 49, 2478-88.

17. Walters, W. P.; Green, J.; Weiss, J. R.; Murcko, M. A. What do medicinal chemists actually make? A 50-year retrospective. ***J. Med. Chem.*** **2011**, 54, 6405-16.

18. Ojima, I.; Chakravarty, S.; Inoue, T.; Lin, S.; He, L.; Horwitz, S. B.; Kuduk, S. D.; Danishefsky, S. J. A common pharmacophore for cytotoxic natural products that stabilize microtubules. ***Proc. Natl. Acad. Sci. U. S. A.*** **1999**, 96, 4256-61.

19. Ojima, I.; Das, M. Recent advances in the chemistry and biology of new generation taxoids. ***J. Nat. Prod.*** **2009**, 72, 554-65.

20. Ganesh, T.; Yang, C.; Norris, A.; Glass, T.; Bane, S.; Ravindra, R.; Banerjee, A.; Metaferia, B.; Thomas, S. L.; Giannakakou, P.; Alcaraz, A. A.; Lakdawala, A. S.; Snyder, J. P.; Kingston, D. G. Evaluation of the tubulin-bound paclitaxel conformation: synthesis, biology, and SAR studies of C-4 to C-3' bridged paclitaxel analogues. ***J. Med. Chem.*** **2007**, 50, 713-25.

21. Matesanz, R.; Barasoain, I.; Yang, C. G.; Wang, L.; Li, X.; de Ines, C.; Coderch, C.; Gago, F.; Barbero, J. J.; Andreu, J. M.; Fang, W. S.; Diaz, J. F. Optimization of taxane binding to microtubules: binding affinity dissection and incremental construction of a high-affinity analog of paclitaxel. ***Chem. Biol.*** **2008**, 15, 573-85.

22. Breslauer, K. J.; Remeta, D. P.; Chou, W. Y.; Ferrante, R.; Curry, J.; Zaunczkowski, D.; Snyder, J. G.; Marky, L. A. Enthalpy-entropy compensations in drug-DNA binding studies. ***Proc. Natl. Acad. Sci. U. S. A.*** **1987**, 84, 8922-6.

23. Lobert, S.; Vulevic, B.; Correia, J. J. Interaction of vinca alkaloids with tubulin: a comparison of vinblastine, vincristine, and vinorelbine. ***Biochemistry*** **1996**, 35, 6806-14.

24. Lobert, S.; Frankfurter, A.; Correia, J. J. Energetics of vinca alkaloid interactions with tubulin isotypes: implications for drug efficacy and toxicity. ***Cell Motil Cytoskeleton*** **1998**, 39, 107-21.

25. Singh, N.; Warshel, A. A comprehensive examination of the contributions to the binding entropy of protein-ligand complexes. ***Proteins*** **2010**, 78, 1724-35.

26. Coderch, C. K., J.; Morreale, A.; Díaz, J.-F.; Gago, F. Comparative Binding Energy (COMBINE) analysis supports a proposal for the binding mode of epothilones to β-tubulin. ***ChemMedChem*** **2012**, in press.

27. Perez, C.; Pastor, M.; Ortiz, A. R.; Gago, F. Comparative binding energy analysis of HIV-1 protease inhibitors: incorporation of solvent effects and validation as a powerful tool in receptor-based drug design. ***J. Med. Chem.*** **1998**, 41, 836-52.

28. Gil-Redondo, R.; Klett, J.; Gago, F.; Morreale, A. gCOMBINE: A graphical user interface to perform structure-based comparative binding energy (COMBINE) analysis on a set of ligand-receptor complexes. ***Proteins*** **2010**, 78, 162-72.

29. Reese, M.; Sanchez-Pedregal, V. M.; Kubicek, K.; Meiler, J.; Blommers, M. J.; Griesinger, C.; Carlomagno, T. Structural basis of the activity of the microtubule-stabilizing agent epothilone A studied by NMR spectroscopy in solution. ***Angew. Chem. Int. Ed. Engl.*** **2007**, 46, 1864-8.

30. Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. H++: a server for estimating pKas and adding missing hydrogens to macromolecules. ***Nucleic Acids Res.*** **2005**, 33, W368-71.

31. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. ***J. Am. Chem. Soc.*** **1995**, 117, 5179-5197.

32. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. ***J. Am. Chem. Soc.*** **1996**, 118, 2309-2309.

33. Aldaz, H.; Rice, L. M.; Stearns, T.; Agard, D. A. Insights into microtubule nucleation from the crystal structure of human gamma-tubulin. ***Nature*** **2005**, 435, 523-7.

34. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Laham, A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian 03, Revision B.04. In Gaussian, Inc.: Pittsburgh, PA., 2003.

35. Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. ***J. Phys. Chem.*** **1993**, 97, 10269-10280.

36. Goodford, P. J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. ***J. Med. Chem.*** **1985**, 28, 849-57.

37. Cuevas, C.; Pastor, M.; Perez, C.; Gago, F. Comparative binding energy (COMBINE) analysis of human neutrophil elastase inhibition by pyridone-containing trifluoromethylketones. ***Comb Chem High Throughput Screen*** **2001**, 4, 627-42.

38. Nicholls, A.; Honig, B. A rapid finite difference algorithm, utilizing successive over-relaxation to solve the Poisson–Boltzmann equation. ***Journal of Computational Chemistry*** **1991**, 12, 435-445.

39. Abdi, H. Partial least squares regression and projection on latent structure regression (PLS Regression). ***Wiley Interdisciplinary Reviews: Computational Statistics*** **2010**, 2, 97-106.

40. Simila, S. T. M.; Martin, S. F. Toward the Total Synthesis of FR901483: Concise Synthesis of the Azatricyclic Skeleton. ***The Journal of Organic Chemistry*** **2007**, 72, 5342-5349.

41. Bertrand, P.; Gesson, J. P. Click Chemistry with O-Dimethylpropargylcarbamate for Preparation of pH-Sensitive Functional Groups. A Case Study. ***The Journal of Organic Chemistry*** **2007**, 72, 3596-3599.

42. Diaz, J. F.; Barasoain, I.; Andreu, J. M. Fast kinetics of Taxol binding to microtubules. Effects of solution variables and microtubule-associated proteins. ***J. Biol. Chem.*** **2003**, 278, 8407-19.
43. Buey, R. M.; Barasoain, I.; Jackson, E.; Meyer, A.; Giannakakou, P.; Paterson, I.; Mooberry, S.; Andreu, J. M.; Diaz, J. F. Microtubule interactions with chemically diverse stabilizing agents: thermodynamics of binding to the paclitaxel site predicts cytotoxicity. ***Chem. Biol.*** **2005**, 12, 1269-79.
44. Wang, L.; Alcaraz, A. A.; Matesanz, R.; Yang, C. G.; Barasoain, I.; Diaz, J. F.; Li, Y. Z.; Snyder, J. P.; Fang, W. S. Synthesis, biological evaluations, and tubulin binding poses of C-2alpha sulfur linked taxol analogues. ***Bioorg Med Chem Lett*** **2007**, 17, 3191-4.
45. Rodriguez-Barrios, F.; Gago, F. Chemometrical identification of mutations in HIV-1 reverse transcriptase conferring resistance or enhanced sensitivity to arylsulfonylbenzonitriles. ***J. Am. Chem. Soc.*** **2004**, 126, 2718-9.
46. Freedman, H.; Huzil, J. T.; Luchko, T.; Luduena, R. F.; Tuszynski, J. A. Identification and characterization of an intermediate taxol binding site within microtubule nanopores and a mechanism for tubulin isotype binding selectivity. ***J Chem Inf Model*** **2009**, 49, 424-36.
47. Ferlini, C.; Raspaglio, G.; Mozzetti, S.; Cicchillitti, L.; Filippetti, F.; Gallo, D.; Fattorusso, C.; Campiani, G.; Scambia, G. The seco-taxane IDN5390 is able to target class III beta-tubulin and to overcome paclitaxel resistance. ***Cancer Res*** **2005**, 65, 2397-405.
48. Snyder, P. W.; Mecinovic, J.; Moustakas, D. T.; Thomas, S. W., 3rd; Harder, M.; Mack, E. T.; Lockett, M. R.; Heroux, A.; Sherman, W.; Whitesides, G. M. Mechanism of the hydrophobic effect in the biomolecular recognition of arylsulfonamides by carbonic anhydrase. ***Proc. Natl. Acad. Sci. U. S. A.*** **2011**, 108, 17889-94.
49. Williams, D. H.; Cox, J. P. L.; Doig, A. J.; Gardner, M.; Gerhard, U.; Kaye, P. T.; Lal, A. R.; Nicholls, I. A.; Salter, C. J.; Mitchell, R. C. Toward the semiquantitative estimation of binding constants. Guides for peptide-peptide binding in aqueous solution. ***J. Am. Chem. Soc.*** **1991**, 113, 7020-7030.
50. Snyder, P. W.; Mecinovic, J.; Moustakas, D. T.; Thomas, S. W., 3rd; Harder, M.; Mack, E. T.; Lockett, M. R.; Heroux, A.; Sherman, W.; Whitesides, G. M. Mechanism of the hydrophobic effect in the biomolecular recognition of arylsulfonamides by carbonic anhydrase. ***Proceedings of the National Academy of Sciences USA*** **2011**.
51. Todd, M. J.; Luque, I.; Velazquez-Campoy, A.; Freire, E. Thermodynamic basis of resistance to HIV-1 protease inhibition: calorimetric analysis of the V82F/I84V active site resistant mutant. ***Biochemistry*** **2000**, 39, 11876-83.
52. Marshall, G. R. Limiting assumptions in structure-based design: binding entropy. ***J. Comput. Aided Mol. Des.*** **2012**, 26, 3-8.
53. Luque, I.; Todd, M. J.; Gomez, J.; Semo, N.; Freire, E. Molecular basis of resistance to HIV-1 protease inhibition: a plausible hypothesis. ***Biochemistry*** **1998**, 37, 5791-7.

Table S1. Selected PLS pseudocoefficients (absolute value ≥ |0.1|) for the energy terms that contribute the most to explaining the predicted binding free energy differences.

Residue	vdW	Ele	Residue	vdW	Ele
Lys19		0.11	Pro274	0.1	
Glu22		0.16	Leu275	0.24	
Val23	0.15		Thr276	-0.1	0.1
Asp26	0.30		Ser277		
Glu27		0.14	Arg278		0.31
Cys213	-0.21		Gln282	-0.1	0.1
Leu217	-0.33		Arg320		
Leu219	-0.26		Pro360	0.1	-0.1
Asp226	0.25	-0.49	Arg369	0.1	
Hip229	0.55	0.50	Gly370	-0.23	-0.15
Leu230	0.14		WAT1		
Ala233	0.48		WAT2	-0.52	0.62
Gly237			DesolvR		
Phe272	0.60		DesolvL		-0.19

Table S2. Selected PLS pseudocoefficients (absolute value ≥ |0.1|) for the energy terms that contribute the most to explaining the predicted binding free energy differences.

Residue	vdW	Ele	Residue	vdW	Ele
Lys19			Pro274		
Glu22		-0.1	Leu275	0.19	
Val23	-0.1		Thr276		
Asp26	0.1	-0.16	Ser277		
Glu27		0.1	Arg278		0.1
Cys213			Gln282		
Leu217	-0.12		Arg320		
Leu219	-0.1		Pro360		
Asp226	0.1	-0.39	Arg369		0.12
Hip229	0.64		Gly370		
Leu230	0.27		WAT1		
Ala233	0.1		WAT2	-0.15	0.13
Gly237			DesolvR		0.13
Phe272	0.23		DesolvL		

Table S3. Binding constants measured at different temperatures of the newly synthesized taxanes.

($x10^7$ M^{-1})	26°C	27°C	30°C	32°C	35°C	37°C	40°C	42°C
CTX55	0.018±0.007	0.018±0.006	0.019±0.003	0.018±0.005	0.015±0.003	0.015±0.002	0.015±0.004	0.014±0.003
CTX56	0.036±0.012	0.035±0.009	0.036±0.005	0.035±0.012	0.029±0.005	0.026±0.004	0.024±0.006	0.029±0.007
CTX57	0.751±0.161	0.661±0.128	0.630±0.135	0.674±0.131	0.595±0.042	0.571±0.081	0.561±0.068	0.569±0.081
CTX58	2.11±0.39	1.80±0.35	1.84±0.61	1.67±0.47	1.87±0.09	1.61±0.30	1.69±0.28	1.60±0.33
CTX59	0.022±0.003	0.019±0.004	0.021±0.001	0.019±0.002	0.017±0.002	0.017±0.004	0.015±0.003	0.013±0.003
CTX60	1.03±0.09	0.90±0.20	0.90±0.12	0.98±0.10	0.85±0.15	0.71±0.15	0.77±0.07	0.82±0.16

Table S4. Cytotoxicity of known and new Taxanes against nonresistant and resistant ovarian carcinoma cells.

Taxane	A2780 (nM)	A2780AD (nM)	R/S
PXL	0.6 ± 0.2	955 ± 356	1516
CTX1	20 ± 1.8	767 ± 508	38.3
CTX55	2920 ± 323	6918 ± 958	2.4
CTX56	2507 ± 107	10250 ± 2185	4.1
CTX57	1.3 ± 0.6	5.8 ± 0.8	4.5
CTX58	0.7 ± 0.1	0.3 ± 0.1	0.4
CTX59	3297 ± 200	5632 ± 914	1.7
CTX60	1.6 ± 0.5	5 ± 3	3.1

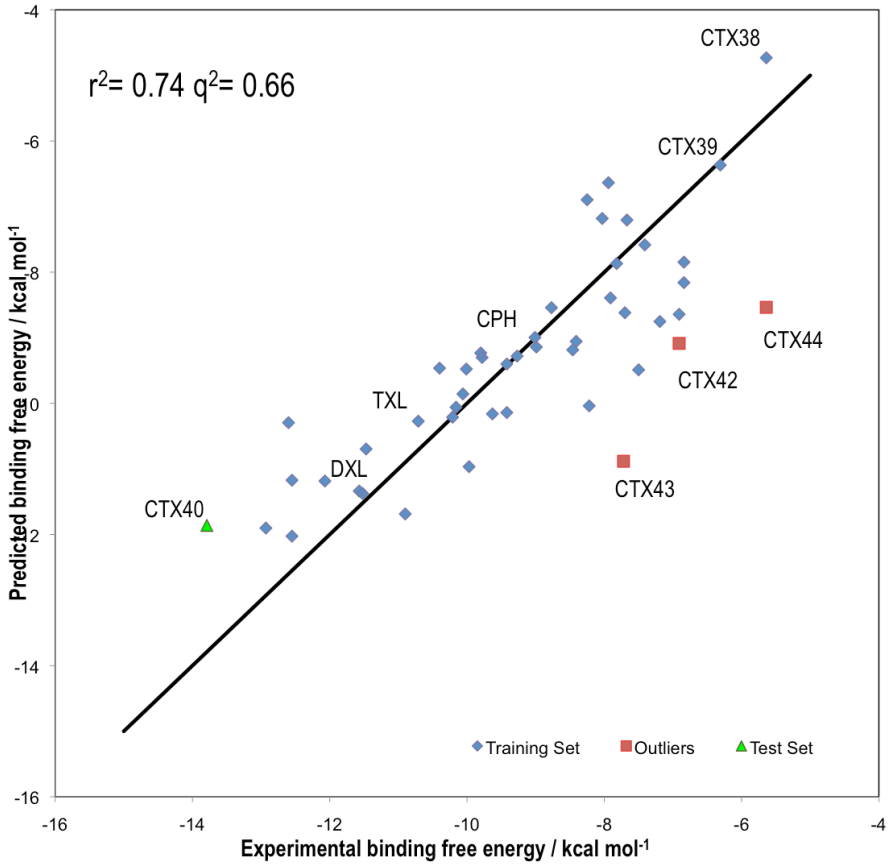


Figure S1. Correlation between the binding free energies calculated in the cross-validated COMBINE model and the experimental values for compounds in the training set (◆) upon exclusion of the three outliers (■). This model, with only 5 LV, accurately predicts the affinity of CTX40 (▲), which was not included in model derivation.

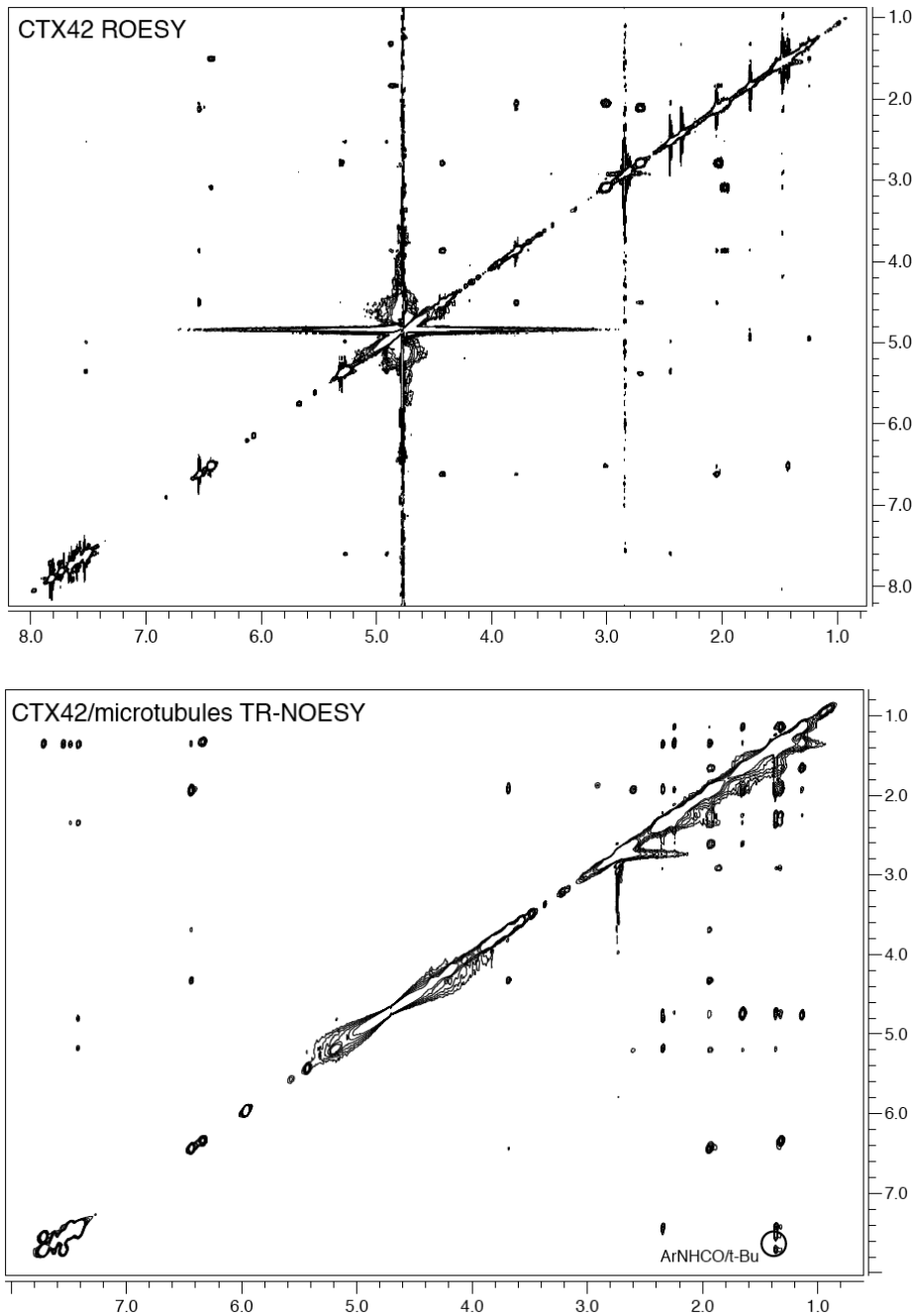


Figure S2. ROESY and TR-NOESY spectra acquired for free and microtubule bound CTX42, respectively.

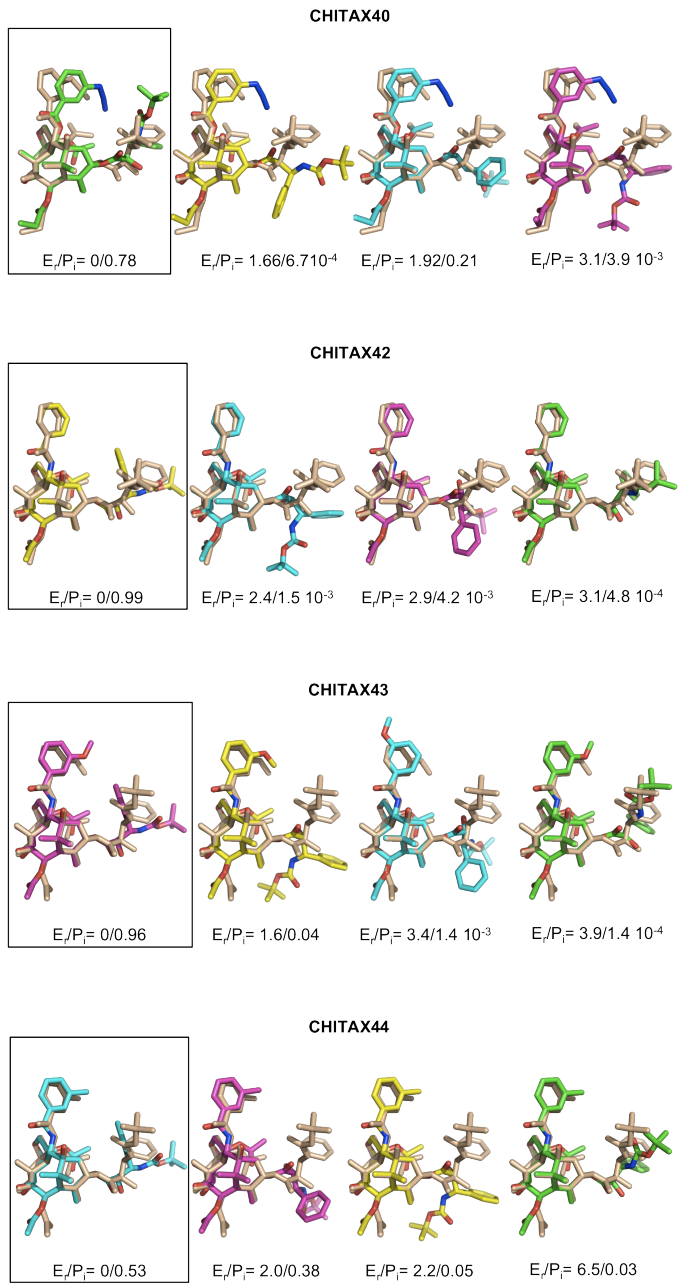


Figure S3. PyMOL stick representation of the superimposition of the four different conformations of CTX40 and CTX42-44 found by free MD simulation in water solvent with the bound conformation as found in the COMBINE model (pale sticks). The free conformers that correspond to the bound one are colored in green, and the most populated ones that have the highest microstate probability and contribute most to the free conformation are highlighted with a black frame.

ARTÍCULO III

"Comparative Binding Energy (COMBINE) analysis supports a proposal for the binding mode of epothilones to β-tubulin."

ChemMedChem 2012 May;7(5):836-43.
doi:10.1002/cmdc.201200065.[PMID:22431398]

Se exploraron las preferencias conformacionales de la Epotilona A (EPA) y de un análogo epimerizado en C12 con un ciclopropilo en 12,13, empleando simulaciones de dinámica molecular en agua, y se seleccionaron los confórmeros que presentaban un acoplamiento óptimo en el sitio de unión de Paclitaxel en la β-tubulina. Los complejos resultantes se refinaron, tras una reconstrucción del M-loop, para comprobar su estabilidad y optimizar la complementariedad. La conformación de la EPA unida a tubulina resultó ser diferente a la previamente propuesta en un modelo construido sobre la base de datos cristalográficos, de espectroscopía de RMN y modelado molecular pero estaba en concordancia con una propuesta basada en datos de RMN en disolución y en el estado sólido, aunque la pose adoptada en el sitio de unión era diferente. Usando estos dos complejos como moldes, se construyeron modelos moleculares para los complejos entre β-tubulina y otros 14 derivados de epotilonas para los que se contaba con datos termodinámicos de afinidad.. A continuación el método de proyección a estructuras latentes (PLS) implementado en COMBINE consiguió obtener una buena correlación entre las entalpías de unión experimentales y una selección de términos de energía de interacción adecuadamente ponderados. Esta relación estructura-actividad cuantitativa apoya el modelo de unión propuesto, contribuye a arrojar luz sobre el efecto de diversas mutaciones en la β-tubulina y puede potencialmente dirigir nuevos estudios experimentales.

CHEMMEDCHEM

ChemPubSoc Europe

DOI: 10.1002/cmdc.201200065

Comparative Binding Energy (COMBINE) Analysis Supports a Proposal for the Binding Mode of Epothilones to β-Tubulin

Claire Coderch,[a] Javier Klett,[b] Antonio Morreale,[b] J. Fernando Díaz,[c] and Federico Gago*[a]

The conformational preferences of epothilone A (EPA) and a 12,13-cyclopropyl C12-epimerized analogue were explored in aqueous solution using molecular dynamics simulations. The simulated conformers that provided an optimal fit in the paclitaxel binding site of mammalian β-tubulin were then selected. The resulting modeled complexes were simulated before and after refinement of the M-loop to improve the fitting and assess ligand stability within the binding pocket. The tubulin-bound conformation of EPA was found to be unlike a previously reported solution obtained through mixed crystallographic/NMR/modeling studies. However, our conformation was in agreement with an NMR-based proposal although the exact binding pose within the site was different. Molecular models were built for the complexes of 14 epothilone derivatives with β-tubulin. A projection to latent structures regression method succeeded in providing a good prediction of the experimentally measured binding enthalpies for the whole set of ligands by assigning weights to a selection of interaction energy terms. These receptor-based, quantitative structure–activity relationships support the proposed binding mode, help confirm and interpret previously acquired experimental data, shed additional light on the effect of several β-tubulin mutations on ligand binding, and can potentially direct further experimental studies.

Introduction

Epothilones A (EPA) and B (EPB, patupilone) were discovered as secondary metabolites in a strain of the soil myxobacterium *Sorangium cellulosum*.[1] They were shown first to have antifungal properties and then to be able to induce cell cycle arrest in the G_2/M transition in mammalian cells.[2] Indeed, EPs bind to polymerized microtubules (MT), the intrinsic dynamics of which are disrupted in a manner similar to that observed with paclitaxel (Taxol), a clinically used antitumor agent that also targets MT.[3] Although many different semisynthetic and synthetic EP analogues have been prepared and tested,[4–6] so far only the EPB lactam derivative ixabepilone (Ixempra)[7,8] has been approved for use in human cancer chemotherapy.[9] The early finding that EPs can displace tubulin-bound paclitaxel[2] suggested a common or overlapping binding site in the β subunit of this heterodimeric protein target. Furthermore, the existence of some similar chemical moieties in both types of natural product was taken as an indication that taxanes and EPs could also share a common pharmacophore,[10,11] a hypothesis that is still supported by some recent studies.[12,13] Nonetheless, EPs promote the assembly and stabilization of purified MT from *Saccharomyces cerevisiae*,[14] whereas paclitaxel does not, although a number of mutations have been identified that can effectively create a paclitaxel binding site on yeast β-tubulin that also accepts EPs as ligands.[15,16]

Establishing the precise binding mode in mammalian β-tubulin of these MT-stabilizing drugs has been hampered by many failed crystallization attempts[17] and by the limited resolution of available two-dimensional crystal structures containing ligand-bound tubulin heterodimers that form Zn^{2+}-stabilized sheets.[18] One of these is the structure deposited in the Protein Data Bank (PDB) with code 1TVK,[19] which does unequivocally display an EPA molecule in the same binding site of β-tubulin that paclitaxel occupies in PDB entry 1JFF.[20] However, whereas independent evidence has supported the binding pose proposed for paclitaxel,[21] doubts have arisen over the bound conformation and orientation of EPA,[22] firstly because of the unlikely *endo* configuration of its epoxide ring, and secondly because it could not account for some known structure–activity relationships (SAR) elucidated for EPs,[4,23] including the effect of several point mutations on β-tubulin from EP-resistant cancer cells.[24–26] In fact, the unique conformation proposed in the joint crystallographic, NMR and modeling work that led to the EPA–β-tubulin cocrystal structure (PDB: 1TVK) is unlike any other described for 1) free EPs (be it in solu-

[a] *C. Coderch, Prof. F. Gago*
Departamento de Farmacología, Universidad de Alcalá
Edificio Medicina, Campus Universitario, 28871 Alcalá de Henares, Madrid (Spain)
E-mail: federico.gago@uah.es

[b] *J. Klett, Dr. A. Morreale*
Unidad de Bioinformática, Centro de Biología Molecular Severo Ochoa (CBMSO)
Consejo Superior de Investigaciones Científicas/Universidad Autónoma de Madrid (CSIC/UAM)
C/ Nicolás Cabrera 1, Campus de la UAM, Cantoblanco, 28049 Madrid (Spain)

[c] *Dr. J. F. Díaz*
Centro de Investigaciones Biológicas (CIB)
Consejo Superior de Investigaciones Científicas (CSIC)
C/ Ramiro de Maeztu 9, 28040 Madrid (Spain)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cmdc.201200065.

tion,[27,28] solvent-derived crystals,[29] an amorphous powder,[30] or a polycrystalline state);[31] 2) EPA bound to nonpolymerized tubulin;[12,32] 3) MT-bound EPB;[30] or 4) EPB and EPD within the substrate (EPD) binding site of a cytochrome P450 enzyme.[33] Furthermore, C6–C8-bridged EP derivatives designed to mimic this binding pose through conformational restriction were recently shown to be devoid of any significant ability to stabilize MT.[34]

NMR spectroscopy stands out as a very useful alternative to electron crystallography but its applicability to EP–tubulin complexes is also limited mostly due to the large size of these systems.[10,18] Despite this caveat, elegant experiments both in solution[12,32] and in the solid state[30] using different preparations of nonpolymerized and polymerized tubulin have been instrumental to unraveling the bound conformation of EPA and EPB. Unfortunately, these state of the art techniques can identify the ligand atoms that undergo chemical shifts upon complex formation but not the interacting protein residues so that the proposed binding poses arise from theoretical docking solutions that best fulfill a number of experimental restraints. Therefore, despite intense experimental efforts, the precise orientation of EPs within the binding site of β-tubulin and the conformation of the residues making up this site in both the α/β-tubulin heterodimer and MT are still a matter of debate.

On the other hand, attempts to derive quantitative SAR (QSAR) for the EPs have been hampered by the lack of reliable structures for the complexes, hence the use of receptorless, ligand-based comparative molecular field analysis (CoMFA),[35] as well as by the nature and quality of available experimental binding data, which most often relied on tubulin depolymerization and/or cytotoxicity assays rather than on direct determination of the true ligand binding affinities. A few years ago, however, the expedient setup of a competition method that measures the displacement of fluorescent paclitaxel analogues, such as Flutax-2,[36,37] from stabilized assembled MT in solution allowed the measurement of equilibrium binding constants at different temperatures and the calculation of incremental changes in Gibbs free energy not only for taxanes[21] but also for a series of chemically modified EPs.[23] A word of caution was then issued in view of the observed enthalpy–entropy compensations throughout the series, in good accord with previous experience in other medicinal chemistry projects.[21,38–40] This observation was not entirely unexpected given the conformational degrees of freedom of these molecules and the solvent accessibility of their binding pocket in β-tubulin. Since complex formation was shown to be dominated by entropy for some compounds and by enthalpy for others, deriving SAR using binding free energies (ΔG) and traditional forcefield methods is likely to be unsuccessful because the configurational entropies ($-T\Delta S$) of the ligand and the water molecules[41] are not taken into account. In this regard, it has recently been proposed that changes in enthalpy (ΔH) could provide a valuable, complementary addition to established tools for selecting compounds in lead discovery and as an aid in lead optimization.[42]

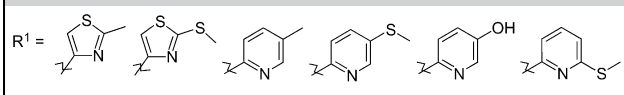
Following this reasoning, our "first-principles" approach to the biologically active conformation and orientation of EPs in the paclitaxel binding site in β-tubulin was 1) to sample the conformations available to EPA and a 12,13-cyclopropyl C12-epimerized analogue (EP5) in aqueous solution using molecular dynamics (MD) simulations; 2) to choose the EPA and EP5 conformers that provide optimal fits in the binding site; 3) to refine the M-loop[43] and equilibrate the resulting solvated EP–β-tubulin complexes in aqueous solution by means of MD; 4) to compare our results with those obtained by NMR for tubulin-bound EPA and MT-bound EPB; and 5) to build the complexes of a series of 16 EPs with β-tubulin with the aim of using receptor-based comparative binding energy (COMBINE) analysis[44,45] to correlate, by means of the projection to latent structures (PLS) regression method,[46] the experimentally measured binding enthalpies[23] to a combination of weighted interaction energies between the ligands and selected protein residues.[47] We believed that if a robust high-quality QSAR model can be derived, indirect support for the proposed binding modes will be obtained, and this knowledge could be of help to improve our understanding of some of the determinants of binding affinity and possibly guide further chemical modifications in the future.

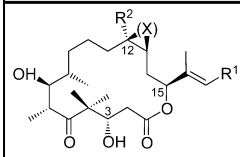
Results and Discussion

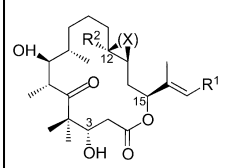
The EP analogues studied in our training set (Table 1) span a range of binding enthalpies of approximately 14 kcal mol^{-1} and differ from one another in the nature of the aromatic group attached to the substituent present at the C15 position, in the conformation of the macrolide ring through *S* to *R* epimerization at C12, and in the nature of the atoms bonded to C12 and C13.[23] A 100 ns MD simulation in water carried out for EPA, as a representative EP containing the native macrolide ring, predicted four major conformations that had a similar outward orientation of the C3 hydroxy group but only occasionally gave a predicted conformation that adopts the inward conformation determined by NMR in aqueous solution for tubulin-bound EPA or that found in the X-ray crystal structure of cytochrome P450-bound EPB.[31] On the contrary, this latter conformation was predicted for the C12-epimerized analogue EP5 during most of the simulation under identical conditions (Figure S1 in the Supporting Information). This proposed preorganization of the C3 region of the macrolide could presumably account for the less unfavorable entropic contribution measured for EP5 relative to EPA.[23]

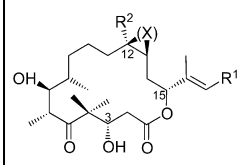
From the poses provided by the automated docking program for EPA and EP5, we prioritized a common solution in which the macrolide ring interacts with the M-loop while the thiazole substituent at C15 is placed inside the hydrophobic pocket lined by the side chains of helices H6 and H7. It is worth noting that in our β-tubulin model, the imidazole ring of His 229 in H7 is protonated on both Nδ and Nε, as proposed previously for the binding of taxanes to the same site[21] at the experimental pH value of 6.5 used in both studies and also in accordance with previous independent work.[12] This modeled binding mode was stable for both EPA and EP5 throughout

Table 1. Structures and tubulin binding enthalpies of the epothilones that provided the training set used for the COMBINE study.[23]









Compd	R^1	R^2	X	ΔH [kcal mol^{-1}]
EPA	A	H	O	−15.5
EPB	A	Me	O	−16.7
EP3	B	Me	O	−8.8
EP4	C	H	CH_2	−12.9
EP14	A	H	CH_2	−13.6
EP19	B	Me	CH_2	n.d.
EP20	C	Me	CH_2	−11.0
EP5	A	H	CH_2	−13.7
EP6	A	Me	O	−8.4
EP7	C	H	O	−12.2
EP8	C	Me	CH_2	−4.8
EP9	F	Me	CH_2	−10.3
EP10	B	H	CH_2	−5.0
EP11	B	Me	CH_2	−2.9
EP12	D	Me	CH_2	−8.6
EP13	E	Me	CH_2	−5.5
EP15	A	H	$(CH_2)_2$	−11.7
EP16	A	H	CH_2	−7.4
EP17	C	H	CH_2	n.d.
EP18	C	H	CH_2	n.d.

n.d.: not determined.

the initial 10 ns MD simulation. To enhance the fitting and to take into account the presence of a neighboring subunit, each EP–tubulin complex was simulated again upon remodeling[48] of the M-loop so as to improve the lateral interactions taking place in MT.[49] The new conformation chosen had a defined helical structure from Arg 278 to Tyr 283 that orientates the latter residue, which is highly conserved in eukaryotes, towards the adjacent protofilament (Figure 1). This putative nucleation of a short α-helix is reminiscent of that recently published for the Ser 174(β)–Ser 178(β) stretch in the vicinity of the vinca domain at the interdimer (β_1–α_2) interface in the apo form of a guanosine triphosphate (GTP)–tubulin complex at 2.1 Å resolution.[50]

Following cooling and energy minimization, the resulting predicted structure for the EPA–tubulin complex revealed a water molecule located inside the binding pocket (WAT1) that bridged a hydrogen-bonding interaction between the epoxide oxygen and the side chain hydroxy of Ser 374. Of note, in the EP5–tubulin complex, this water molecule cannot form a hydrogen bond with the ligand because this oxygen in the drug is replaced with a methylene group (Table 1). Our proposed binding mode for EPA is qualitatively compatible with recent solid-state NMR results showing that the largest chemical shifts upon binding of EPB to β-tubulin are experienced by C1, C15, C17 and C18 (in our model all of them close to His 229), C3 (close to Thr 276 in the M-loop), C9 (close to Leu 371), and C22 (close to Thr 276, Ser 277 and Gln 281 in the M-loop).[30] Therefore, we built and refined the remaining set of EP–tubulin complexes using these two templates (incorporating WAT1) and subjected them to a COMBINE analysis (Figure 2). A cross-validated PLS model with just four latent variables (LV) revealed that the experimental binding enthalpies could indeed be expressed as a function of a weighted subset of ligand–residue interactions (r^2 = 0.94, q^2 = 0.77 for the best model and an average of $r^2 = 0.93 \pm 0.3$, $q^2 = 0.65 \pm 0.1$ and 4 ± 1 LV for the remaining nine runs) (Figure 3).

The β-tubulin residues that are assigned the largest PLS pseudo-coefficients for the van der Waals and electrostatic blocks are shown in Table 2. Their relative importance for explaining the differences in binding enthalpies depends on the actual magnitude of the calculated interaction energies, because these signed coefficients are used to properly weigh these terms.[44,51] Since the energy values are usually negative (i.e., attractive), a positive coefficient will make the calculated binding enthalpy more negative, and vice versa. Thus, the importance of the direct hydrogen bonds established between the hydroxy groups of EPs and the carbonyl groups of Thr 276 and Gly 370 is reflected in the positive coefficients assigned to the electrostatic terms involving these two residues. Likewise, an improved steric interaction between the macrocycle and Pro 274 and Pro 360 results in a gain in binding enthalpy. The substituents on the thiazole ring, on the other hand, modulate the affinity through favorable van der Waals interactions with Cys 213 and His 229, but the electrostatic term only favors the former residue and not the latter, which is positively charged. Conversely, increased ligand interactions with other residues, particularly Leu 219, Leu 227 and Leu 230, leads to a loss of binding efficiency, which most likely reflects the existence of steric clash.[51]

Table 2. Selected projection to latent structure (PLS) pseudo-coefficients for the amino acids that contribute the most to explaining the predicted binding enthalpy differences.[a]

Residue	vdW	Ele	Residue	vdW	Ele
Leu 209	−0.49	–	Leu 275	0.21	–
Ile 212	−0.26	–	Thr 276	−0.40	0.12
Cys 213	0.36	0.66	Ser 277	−0.47	−0.12
Leu 217	−0.38	–	Gln 281	–	0.12
Leu 219	−1.77	–	Arg 284	0.60	–
Leu 227	−2.82	–	Pro 360	1.39	–
His 229	2.16	−0.86	Arg 369	−0.71	–
Leu 230	−3.07	–	Gly 370	0.88	0.94
Ala 233	0.76	–	Leu 371	−0.41	–
Phe 272	−1.18	–	Ser 374	−0.53	–
Pro 274	2.74	–	WAT1	–	0.63

[a] Only PLS pseudo-coefficients with absolute values $\geq |0.1|$ for the van der Waals (vdW) and electrostatic (Ele) interactions are shown.

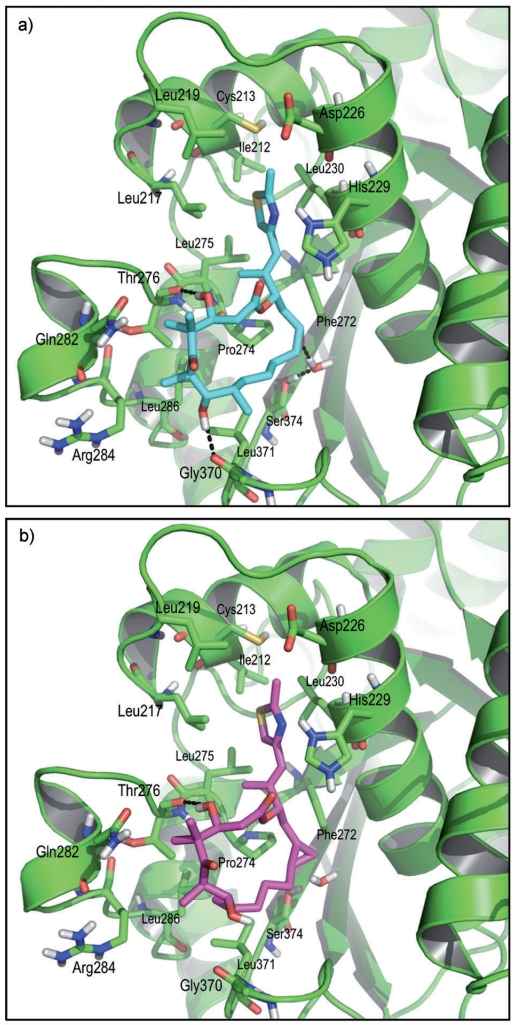


Figure 1. a) EPA and b) EP5 docked in the taxane binding site of a β-tubulin model in which the M-loop (Pro 274–Pro 289) has been remodeled to improve drug fitting and better account for the lateral interactions present in the MT. Dotted lines represent hydrogen bonds. Note the water molecule (WAT1) adjacent to the hydroxy of Ser 374. Some relevant residues have been numbered according to PDB entry 1JFF and label size is inversely proportional to distance from the viewer.

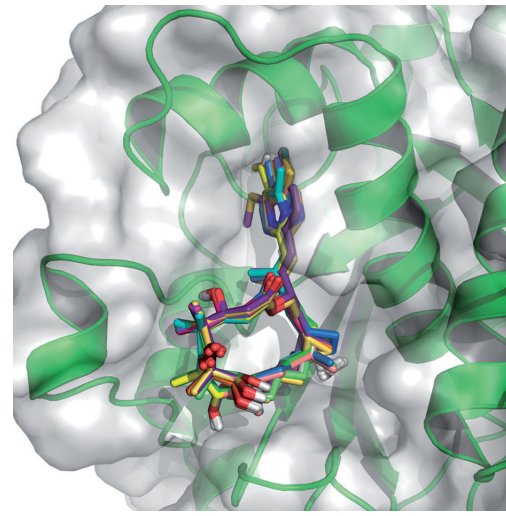


Figure 2. Superposition of the set of EPs studied in the paclitaxel binding site of β-tubulin. For simplicity, only a cartoon representation of the protein enveloped by a semitransparent grey van der Waals surface is shown.

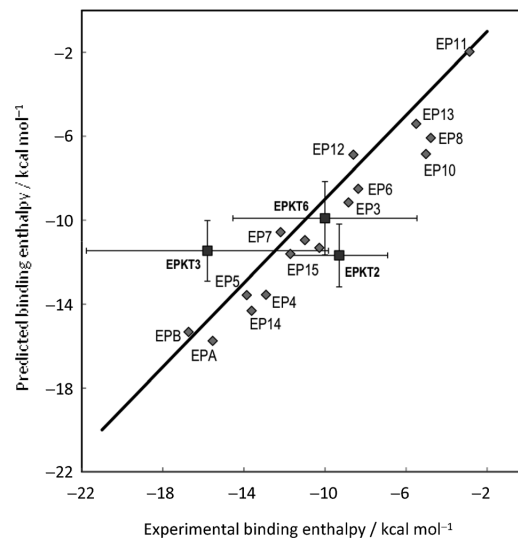
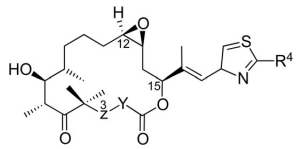


Figure 3. Correlation between experimental binding enthalpies and those calculated using the COMBINE model: training set (◆); prediction set (■). For simplicity, error bars are shown only for the prediction set.

To further validate this COMBINE model, we chose three EPA analogues that were not part of the training set and for which the binding enthalpies have been recently determined:[52] two of them are the 3-deoxy (EPKT2) and 3-deoxy-2,3-didehydro (EPKT3) derivatives of EPA for which only minor conformational differences in aqueous solution have been previously reported,[28] whereas the other one (EPKT6) bears an *n*-propyl group on the thiazole ring in place of the methyl (Table 3). The decreased binding enthalpies of EPKT2 and EPKT6 relative to EPA were reasonably well estimated but the binding enthalpy for EPKT3 was underpredicted. This analogue reportedly displays an affinity similar to that of EPA (although the experimental standard error is much larger), despite the absence of the hydroxy group at C3, and adopts a planar arrangement of its C1, C2 and C3 carbon atoms due to the conjugation of the C2–C3 double bond with the C1 carbonyl. As a consequence, in the complex, the carbonyl oxygen appears slightly reoriented towards the imidazole ring of His 229. This finding made us think

Table 3. Structures and tubulin binding enthalpies of the epothilone prediction set.[52]



Compd	R^4	Y–Z	ΔH [kcal mol^{-1}]
EPKT2	Me	CH_2–CH_2	-9.3 ± 2.4
EPKT3	Me	CH=CH	-15.8 ± 6.0
EPKT6	Pr	CH_2–CHOH	-10.0 ± 4.5

of the possibility of a hydrogen bond being established between this oxygen and Nε of His 229, provided the imidazole ring of this amino acid side chain is rotated 180° about the Cβ–Cγ bond. Accordingly, a new set of complexes was built, refined, and used to generate an alternative COMBINE model, the quality of which was only slightly worse than that of the preceding model ($r^2 = 0.86$, $q^2 = 0.68$, 3 LV and an average of $r^2 = 0.85 \pm 0.03$, $q^2 = 0.60 \pm 0.07$ and 3 ± 1 LV for the remaining nine runs). This second model predicted EPKT3 more accurately, provided a similar value for the binding enthalpy of EPKT6, but failed to predict the binding enthalpy of EPKT2. The major difference in the weights assigned to the ligand–residue interactions affected His 229 (Figure S2 in the Supporting Information). Nonetheless, since the rotameric state of the imidazole ring reversed to the original one when MD simulations were run to test this alternative (data not shown), we do not favor this second binding mode and suggest that the loss of binding enthalpy caused by the lack of the hydroxy group at C3 is partly compensated by improved van der Waals interactions within the binding pocket.

The side chain of His 229 divides the ligand binding cavity into two major pockets, both of which are occupied in the case of taxanes (Figure 4).[21] In contrast, in our proposal, EPs fill only one half-site and establish interactions with the amino acids present in helices H7 and H6 (Ile 212, Cys 213, Leu 217, Leu 219, His 229, Leu 230), part of the M-loop (Phe 272, Pro 274, Leu 275, Thr 276, Ser 277, Gln 281), and the loop linking strands S9 and S10 (Gly 370, Leu 371, Ser 374) (Figure 1). Consequently, the reported existence of a "common pharmacophore"[10–13,25] appears to be limited to a good general overlap of both types of compounds in the commonly filled half-site and the presence of two similar anchoring moieties, namely the C3-OH/oxetane and the C15-thiazole/C2-benzoyl pairs in EPs and paclitaxel, respectively. Thus, we also favor the view that the thiazole side chain of EPs corresponds to the C-2 side chain of taxanes.[11]

When the three-dimensional coordinates of the NMR-based model of the EPA–tubulin complex were used as the template for model building the whole set of complexes, the COMBINE method could not find a correlation between the calculated binding energies and the experimental binding enthalpies (Figure S3 in the Supporting Information). This means that the calculated ligand–tubulin interactions are either not sufficiently

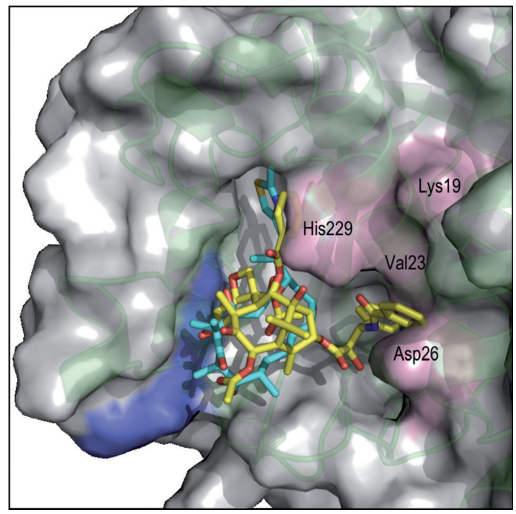


Figure 4. Taxane binding site in β-tubulin enveloped by a semitransparent grey van der Waals surface displaying bound EPA (carbon atoms in cyan) and a superimposed paclitaxel (carbon atoms in yellow) for reference (as found in PDB entry 1JFF). Shaded regions correspond to residues that confer resistance to epothilones (blue) or taxanes (pink) upon mutation.

different among the series of analogues or do not correlate with the observed changes in ΔH values. We take this result as an indication that this binding pose might be a snapshot representing an intermediate state between the free and bound forms for this ligand and/or that the non-microtubular form of tubulin that was employed in these experiments differs from that found in MT.

The orientation that we suggest for EP binding to β-tubulin on mammalian MT also rationalizes the observation that this association process is hampered in the presence of Thr 276 Ile, Arg 284 Gln, and Gln 294 Glu point mutations to the extent that they cause drug resistance.[24,25] The increase in bulk and hydrophobicity brought about by the former substitution (Thr 274 in the original publication[19]) should be detrimental for affinity, in good accord with the negative PLS coefficient calculated for this residue, while the reverse is true for Arg 284 (published as Arg 282[25]) (Table 2). The latter residue (Gln 292 in the original publication[19]) does not appear to interact with EP directly but its mutation is likely to affect the conformation of the M-loop. The proposed binding pose also accounts for the fact that EPs bind to yeast tubulin without requiring any of the mutations that have to be introduced into this orthologous protein to generate a taxane binding site,[15] particularly Val 23 and Asp 26[53] (Figure S4 in the Supporting Information).

Conclusions

A robust target-based QSAR model has been derived using the COMBINE method on a set of modeled complexes that supports a well-defined orientation for EP in the taxane binding site of β-tubulin that is in good agreement with the qualitative

data obtained from solid-state NMR experiments on MT-bound EPB.[30] This was achieved by leaving aside the yet computationally elusive entropic component ($-T\Delta S$) from the Gibbs free energies (ΔG) and focusing on the experimental enthalpic contributions (ΔH). The fact that incorporation of ligand and receptor desolvation terms[44] neither improved the quality of the models nor was sufficient in this case to provide a correlation with ΔG values could indicate that most of the entropic component arises from the configurational entropy of the ligands and/or from the water molecules that are released/rearranged upon complex formation. In this respect, it is also likely that, in the chemically heterogeneous environment of this binding pocket, different water molecules might have distinct thermodynamic signatures, as recently described for the binding of a series of arylsulfonamides to human carbonic anhydrase.[54]

Further support to the binding mode proposed herein—or any other—will undoubtedly require more experimental evidence including further refinement of the M-loop and additional data obtained with new analogues. Work in this latter direction is already underway in collaboration with a synthetic chemistry group, but we believe that sharing our findings at the present time might also help other researchers in their own endeavors.

Experimental Section

The starting geometry for EPA was taken from the X-ray crystal structure deposited in the Cambridge Structural Database (CSD) with reference code XAPDIC.[22] The models of the remaining EPs published in Ref. [23] for which ΔH values are available were built using EPA as a template, with the exception of EP16, EP17 and EP18 that were not considered because the change in geometry at C15 from *S* to *R* results in a complete loss of binding affinity.[23] EP19 could not be used either because its high affinity prevented the accurate determination of its binding enthalpy.[23] The charge distribution for all the ligands was obtained by fitting the quantum mechanically calculated (HF/6-31G*//HF/3-21G*) molecular electrostatic potential (MEP), as implemented in Gaussian 03,[55] to a restrained electrostatic potential (RESP) point-charge model.[56] The general AMBER (http://ambermd.org/) force field (GAFF) was used to assign bonded and nonbonded parameters (parm03) to EP atoms. EPA and EP5 were immersed in a cubic box of ~1000 TIP3P water molecules[57] and subjected to molecular dynamics (MD) simulations at 300 K for 100 ns to assess their conformational variability.

The three-dimensional structure of β-tubulin as found in the α,β-tubulin dimer in complex with EPA at 2.9/4.2 Å resolution (PDB: 1TVK)[19] was used as the receptor upon removal of EPA. The guanosine diphosphate (GDP) molecule was conserved, and the protein side chains in the nucleotide binding site were slightly reoriented so that they established the same interactions that are observed in the 2.51 Å resolution X-ray crystal structure of γ-tubulin bound to 5′-guanosine-diphosphate-monothiophosphate (PDB: 1Z5V). Addition of missing hydrogen atoms and computation of the protonation state of ionizable groups at pH 6.5 were carried out using the H^{++} Web server,[58] which relies on AMBER parameters and finite difference solutions to the Poisson–Boltzmann equation. Automated docking of EPA and EP5 was performed using the Lamarckian genetic algorithm, implemented in AutoDock 3.0.5,[59] by allowing random changes of overall orientation within the taxane binding site of all the different conformers identified in the MD ensembles as well as torsional freedom about the C15–C16 and C17–C18 bonds. The complexes with the most plausible poses were then simulated under periodic boundary conditions for 10 ns at 300 K in explicit solvent (a truncated octahedron containing ~12000 TIP3P water molecules plus 11 randomly placed Na^+ ions to achieve electroneutrality) using a harmonic restraint of 5 kcal mol^{-1} $Å^{-2}$ on the protein Cα atoms. Electrostatic interactions were treated using the smooth particle mesh Ewald method,[60] with a grid spacing of 1 Å; the cutoff distance for the nonbonded interactions was 9 Å; the SHAKE algorithm[61] was applied to all bonds; and an integration step of 2.0 fs was used throughout. Subsequent gradual cooling followed by energy minimization provided representative structures for the EPA– and EP5–tubulin complexes.

At this point, we found it necessary to optimize the conformation of the M-loop as this stretch is known to be involved not only in lateral interactions between MT protofilaments,[49] but also in EP binding. To achieve this goal, the Robetta web server[48] was used to provide different conformations for the M-loop (from Pro 274 to Pro 289), and one was selected that placed Tyr 283 in an outward orientation suitable for interaction with the neighboring monomer and maintained Thr 276, the NH of which is known to form a hydrogen bond with the oxetane oxygen of taxanes,[21] roughly in the same position as in PDB codes 1TVK and 1JFF. After an initial energy minimization of the water molecules and counter ions, the system was heated to 300 K in 25 ps after which the solvent was allowed to redistribute around the position-restrained solute for 220 ps. After this time, only the protein Cα atoms were restrained with a harmonic force constant of 5 kcal mol^{-1} $Å^{-2}$ during the subsequent 10 ns MD simulation that explored the mutual adaptation between ligand and amino acid side chains without altering the overall conformation of the M-loop. In polymerized tubulin, the conformation of this loop is stabilized by the presence of a neighbor subunit that is absent in our calculations for computational efficiency. The systems were then gradually cooled from 300 K to 273 K over 1200 ps, as reported previously for several taxane–tubulin complexes.[21] Energy minimization of the resulting "cooled" EPA–tubulin and EP5–tubulin complexes employing 5000 cycles of steepest descent followed by 4000 cycles of conjugate gradient provided the representative templates for the rest of the complexes.

The remaining EP in the training set[23] were modeled inside the binding pocket of β-tubulin using either EPA (EPB, EP3, EP4 and EP14) or EP5 (EP6-EP13 and EP15) for molecular editing. All the complexes were refined by energy minimization using the steepest descent algorithm until the root-mean-square of the potential energy gradient was below 0.1 kcal mol^{-1} $Å^{-1}$. The set of refined ligand-receptor complexes (including one water molecule, as explained below) was then used as input to the gCOMBINE program,[45] which automatically calculated all the van der Waals (AMBER force field) and electrostatic interaction energies (Goodford's implementation of the images method[62] and a uniform dielectric constant of 4.0 Debye) between each ligand and every protein residue. The data matrix containing the computed energy components (*x* variables) was then subjected to multivariate statistical analysis using partial least squares (PLS) in order to find a correlation with the experimental binding enthalpies (*y* variable). For cross-validation purposes and to assess the robustness of the model, instead of the very conservative leave-one-out method, random groups of five elements per group were excluded from

the training set each time, as previously described,[45] and the procedure was repeated ten times.

The molecular graphics program PyMOL (v. 0.99rc6, DeLano Scientific, LLC, Palo Alto, CA) was employed for molecular editing, visualization and representation.

Acknowledgements

We thank Dr. Teresa Carlomagno (European Molecular Biology Laboratory, Heidelberg, Germany) for providing the coordinates of the model reported in Ref. [12] and also Dr. Ángeles Canales and Prof. José Manuel Andreu (CIB/CSIC, Madrid, Spain) for useful discussions. C.C. acknowledges the financial support of the Ministerio de Ciencia e Innovación de España (FPU grant AP2007-01225) and A.M. acknowledges the Comunidad de Madrid (Spain) for financial support to the Fundación Severo Ochoa through the AMAROUTO program. This work was funded by the Comisión Interministerial de Ciencia y Tecnología de España (SAF2009-13914-C02-02), the Comunidad de Madrid (Spain) (BIPEDD: S-BIO/0214/2006 and BIPEDD-2: S2010-BMD-2457), and the Ministerio de Ciencia e Innovación de España (BIO2008-04384 and BIO-2010-16351). The computer resources, technical expertise and assistance provided by the Barcelona Supercomputing Center (Centro Nacional de Supercomputación) are gratefully acknowledged.

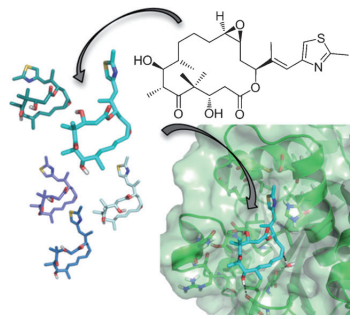
Keywords: anticancer agents · epothilones · molecular dynamic simulations · quantitative structure–activity relationships · tubulin

[1] K. Gerth, N. Bedorf, G. Hofle, H. Irschik, H. Reichenbach, *J. Antibiot.* **1996**, *49*, 560 – 563.
[2] D. M. Bollag, P. A. McQueney, J. Zhu, O. Hensens, L. Koupal, J. Liesch, M. Goetz, E. Lazarides, C. M. Woods, *Cancer Res.* **1995**, *55*, 2325 – 2333.
[3] C. Dumontet, M. A. Jordan, *Nat. Rev. Drug Discovery* **2010**, *9*, 790 – 803.
[4] K. C. Nicolaou, D. Vourloumis, T. Li, J. Pastor, N. Winssinger, Y. He, S. Ninkovic, F. Sarabia, H. Vallberg, F. Roschangar, N. P. King, M. R. V. Finlay, P. Giannakakou, P. Verdier-Pinard, E. Hamel, *Angew. Chem.* **1997**, *109*, 2181 – 2187; *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 2097 – 2103.
[5] K. H. Altmann, B. Pfeiffer, S. Arseniyadis, B. A. Pratt, K. C. Nicolaou, *ChemMedChem* **2007**, *2*, 396 – 423.
[6] K. H. Altmann, F. Z. Gaugaz, R. Schiess, *Mol. Diversity* **2011**, *15*, 383 – 399.
[7] J. T. Hunt, *Mol. Cancer Ther.* **2009**, *8*, 275 – 281.
[8] R. M. Borzilleri, G. D. Vite in *Annual Reports in Medicinal Chemistry*, Vol. 44 (Ed.: J. E. Macor), Academic Press, **2009**, pp. 301 – 322.
[9] J. J. Lee, S. M. Swain, *Clin. Cancer Res.* **2008**, *14*, 1618 – 1624.
[10] I. Ojima, S. Chakravarty, T. Inoue, S. Lin, L. He, S. B. Horwitz, S. D. Kuduk, S. J. Danishefsky, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 4256 – 4261.
[11] L. He, P. G. Jagtap, D. G. Kingston, H. J. Shen, G. A. Orr, S. B. Horwitz, *Biochemistry* **2000**, *39*, 3972 – 3978.
[12] M. Reese, V. M. Sanchez-Pedregal, K. Kubicek, J. Meiler, M. J. Blommers, C. Griesinger, T. Carlomagno, *Angew. Chem.* **2007**, *119*, 1896 – 1900; *Angew. Chem. Int. Ed.* **2007**, *46*, 1864 – 1868.
[13] S. Forli, F. Manetti, K. H. Altmann, M. Botta, *ChemMedChem* **2010**, *5*, 35 – 40.
[14] C. J. Bode, M. L. Gupta, Jr., E. A. Reiff, K. A. Suprenant, G. I. Georg, R. H. Himes, *Biochemistry* **2002**, *41*, 3870 – 3874.
[15] M. L. Gupta, Jr., C. J. Bode, G. I. Georg, R. H. Himes, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 6394 – 6397.
[16] R. A. Entwistle, R. D. Winefield, T. B. Foland, G. H. Lushington, R. H. Himes, *FEBS Lett.* **2008**, *582*, 2467 – 2470.
[17] A. Cormier, M. Knossow, C. Wang, B. Gigant, *Methods Cell Biol.* **2010**, *95*, 373 – 390.
[18] J. Jiménez-Barbero, F. Amat-Guerri, J. P. Snyder, *Curr. Med. Chem.: Anti-Cancer Agents* **2002**, *2*, 91 – 122.
[19] J. H. Nettles, H. Li, B. Cornett, J. M. Krahn, J. P. Snyder, K. H. Downing, *Science* **2004**, *305*, 866 – 869.
[20] J. Lowe, H. Li, K. H. Downing, E. Nogales, *J. Mol. Biol.* **2001**, *313*, 1045 – 1057.
[21] R. Matesanz, I. Barasoain, C. G. Yang, L. Wang, X. Li, C. de Ines, C. Coderch, F. Gago, J. J. Barbero, J. M. Andreu, W. S. Fang, J. F. Diaz, *Chem. Biol.* **2008**, *15*, 573 – 585.
[22] D. W. Heinz, W. D. Schubert, G. Hofle, *Angew. Chem.* **2005**, *117*, 1324 – 1327; *Angew. Chem. Int. Ed.* **2005**, *44*, 1298 – 1301.
[23] R. M. Buey, J. F. Díaz, J. M. Andreu, A. O'Brate, P. Giannakakou, K. C. Nicolaou, P. K. Sasmal, A. Ritzén, K. Namoto, *Chem. Biol.* **2004**, *11*, 225 – 236.
[24] L. He, C. P. Yang, S. B. Horwitz, *Mol. Cancer Ther.* **2001**, *1*, 3 – 10.
[25] P. Giannakakou, R. Gussio, E. Nogales, K. H. Downing, D. Zaharevitz, B. Bollbuck, G. Poy, D. Sackett, K. C. Nicolaou, T. Fojo, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 2904 – 2909.
[26] M. Wartmann, K.-H. Altmann, *Curr. Med. Chem.: Anti-Cancer Agents* **2002**, *2*, 123 – 148.
[27] R. E. Taylor, Y. Chen, G. M. Galvin, P. K. Pabba, *Org. Biomol. Chem.* **2004**, *2*, 127 – 132.
[28] M. Erdélyi, B. Pfeiffer, K. Hauenstein, J. Fohrer, J. Gertsch, K. H. Altmann, T. Carlomagno, *J. Med. Chem.* **2008**, *51*, 1469 – 1473.
[29] G. Höfle, N. Bedorf, H. Steinmetz, D. Schomburg, K. Gerth, H. Reichenbach, *Angew. Chem.* **1996**, *108*, 1671 – 1673; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1567 – 1569.
[30] A. Kumar, H. Heise, M. J. Blommers, P. Krastel, E. Schmitt, F. Petersen, S. Jeganathan, E. M. Mandelkow, T. Carlomagno, C. Griesinger, M. Baldus, *Angew. Chem.* **2010**, *122*, 7666 – 7669; *Angew. Chem. Int. Ed.* **2010**, *49*, 7504 – 7507.
[31] A. Lange, T. Schupp, F. Petersen, T. Carlomagno, M. Baldus, *ChemMedChem* **2007**, *2*, 522 – 527.
[32] T. Carlomagno, M. J. Blommers, J. Meiler, W. Jahnke, T. Schupp, F. Petersen, D. Schinzer, K. H. Altmann, C. Griesinger, *Angew. Chem.* **2003**, *115*, 2615 – 2619; *Angew. Chem. Int. Ed.* **2003**, *42*, 2511 – 2515.
[33] S. Nagano, H. Li, H. Shimizu, C. Nishida, H. Ogura, P. R. Ortiz de Montellano, T. L. Poulos, *J. Biol. Chem.* **2003**, *278*, 44886 – 44893.
[34] W. Zhan, Y. Jiang, S. Sharma, P. J. Brodie, S. Bane, D. G. Kingston, D. C. Liotta, J. P. Snyder, *Chem. Eur. J.* **2011**, *17*, 14792 – 14804.
[35] K. W. Lee, J. M. Briggs, *J. Comput. Aided Mol. Des.* **2001**, *15*, 41 – 55.
[36] J. F. Diaz, R. Strobe, Y. Engelborghs, A. A. Souto, J. M. Andreu, *J. Biol. Chem.* **2000**, *275*, 26265 – 26276.
[37] I. Barasoain, J. F. Diaz, J. M. Andreu, *Methods Cell Biol.* **2010**, *95*, 353 – 372.
[38] T. S. Olsson, J. E. Ladbury, W. R. Pitt, M. A. Williams, *Protein Sci.* **2011**, *20*, 1607 – 1618.
[39] P. A. Borea, A. Dalpiaz, K. Varani, P. Gilli, G. Gilli, *Biochem. Pharmacol.* **2000**, *60*, 1549 – 1556.
[40] A. Memic, M. R. Spaller, *ChemBioChem* **2008**, *9*, 2793 – 2795.
[41] N. Singh, A. Warshel, *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 1724 – 1735.
[42] J. E. Ladbury, G. Klebe, E. Freire, *Nat. Rev. Drug Discovery* **2010**, *9*, 23 – 27.
[43] E. Nogales, M. Whittaker, R. A. Milligan, K. H. Downing, *Cell* **1999**, *96*, 79 – 88.
[44] C. Pérez, M. Pastor, A. R. Ortiz, F. Gago, *J. Med. Chem.* **1998**, *41*, 836 – 852.
[45] R. Gil-Redondo, J. Klett, F. Gago, A. Morreale, *Proteins* **2010**, *78*, 162 – 172.
[46] H. Abdi, *Wiley Interdiscip. Rev. Comput. Statistics* **2010**, *2*, 97 – 106.
[47] G. H. Lushington, J.-X. Guo, J. L. Wang, *Curr. Med. Chem.* **2007**, *14*, 1863 – 1877.
[48] D. E. Kim, D. Chivian, D. Baker, *Nucleic Acids Res.* **2004**, *32*, W526 – 531.
[49] P. Meurer-Grob, J. Kasparian, R. H. Wade, *Biochemistry* **2001**, *40*, 8000 – 8008.
[50] A. Nawrotek, M. Knossow, B. Gigant, *J. Mol. Biol.* **2011**, *412*, 35 – 42.
[51] F. Rodríguez-Barrios, F. Gago, *J. Am. Chem. Soc.* **2004**, *126*, 2718 – 2719.

[52] M. Erdélyi, A. Navarro-Vázquez, B. Pfeiffer, C. N. Kuzniewski, A. Felser, T. Widmer, J. Gertsch, B. Pera, J. F. Díaz, K. H. Altmann, T. Carlomagno, *ChemMedChem* **2010**, *5*, 911 – 920.
[53] M. Hari, F. Loganzo, T. Annable, X. Tan, S. Musto, D. B. Morilla, J. H. Nettles, J. P. Snyder, L. M. Greenberger, *Mol. Cancer Ther.* **2006**, *5*, 270 – 278.
[54] P. W. Snyder, J. Mecinovic, D. T. Moustakas, S. W. Thomas III, M. Harder, E. T. Mack, M. R. Lockett, A. Heroux, W. Sherman, G. M. Whitesides, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 17889 – 17894.
[55] Gaussian 03, Revision B.04; M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, A. Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, J. A. Pople, Gaussian, Inc., Pittsburgh, PA., **2003**.
[56] C. I. Bayly, P. Cieplak, W. Cornell, P. A. Kollman, *J. Phys. Chem.* **1993**, *97*, 10269 – 10280.
[57] W. Jorgensen, J. Chandrasekhar, J. Madura, R. Impey, M. Klein, *J. Chem. Phys.* **1983**, *79*, 926 – 935.
[58] J. C. Gordon, J. B. Myers, T. Folta, V. Shoja, L. S. Heath, A. Onufriev, *Nucleic Acids Res.* **2005**, *33*, W368 – 371.
[59] G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, *J. Comput. Chem.* **1998**, *19*, 1639 – 1662.
[60] T. A. Darden, D. York, L. G. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10089 – 10092.
[61] J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327 – 341.
[62] P. J. Goodford, *J. Med. Chem.* **1985**, *28*, 849 – 857.

Received: January 31, 2012
Revised: March 2, 2012
Published online on March 16, 2012

An elusive binding pose: The results of receptor-based, quantitative structure–activity relationship studies lend further support to one of the proposed binding modes for epothilones to β-tubulin and help interpret previously acquired experimental binding enthalpies.



*C. Coderch, J. Klett, A. Morreale, J. F. Díaz, F. Gago**

■■ – ■■

Comparative Binding Energy (COMBINE) Analysis Supports a Proposal for the Binding Mode of Epothilones to β-Tubulin

CHEMMEDCHEM
Chemistry & Drug Discovery

Supporting Information

Comparative Binding Energy (COMBINE) Analysis Supports a Proposal for the Binding Mode of Epothilones to β-Tubulin

Claire Coderch,[a] Javier Klett,[b] Antonio Morreale,[b] J. Fernando Díaz,[c] and Federico Gago*[a]

cmdc_201200065_sm_miscellaneous_information.pdf

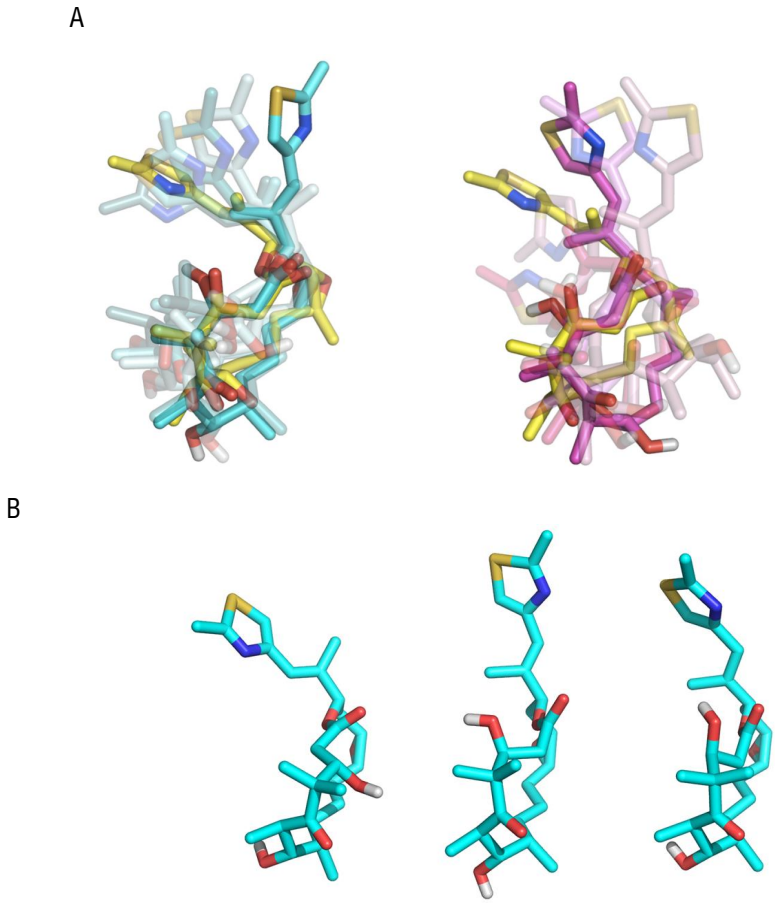


Figure S1. A. Best-fit superimposition of the C1-C6 region of both EPA (*left*) and EP5 (*right*), as found in their complexes with tubulin (solid cyan and magenta sticks, respectively), onto the structure of CytP450-bound EPB (yellow, PDB code 1Q5D) and the different conformers (semitransparent sticks) identified during the MD simulations. The most populated conformers of EPA oriented their C3 hydroxyl moiety outwards whereas only one minor conformation (the one used for modelling the complex) oriented it inwards in the same manner as the CytP450-bound EPB. On the other hand, the epimerization of the C12 in EP5 favored that the most populated conformers had an inward orientation of the C3 hydroxyl while the largest fluctuations were observed from C6 to C12 and the thiazole side chain at C15. B. *Left:* EPA conformation in Cambridge Structural Database entry XAPDIC and also found during most of the molecular dynamics simulation in water. *Middle:* conformation of EPA that was only rarely seen during the molecular dynamics simulation in water but was selected for model building the complex with tubulin. *Right:* NMR-derived conformation of tubulin-bound EPA as reported in ref. 12.

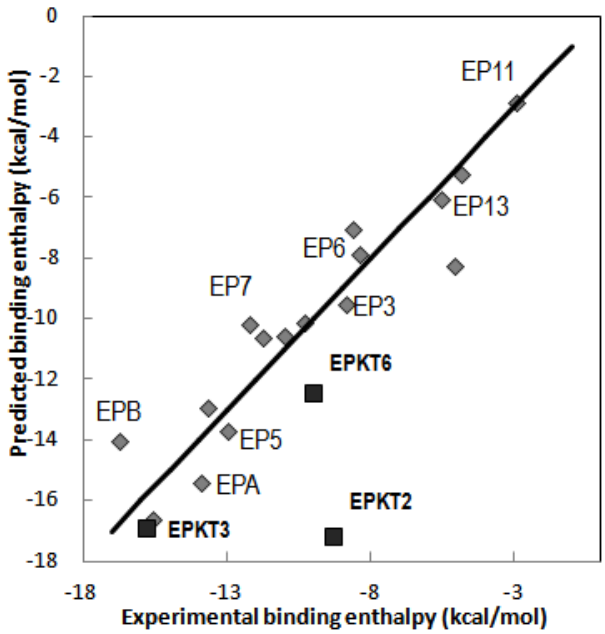


Figure S2. A. Alternative COMBINE model generated from the complexes in which the imidazole ring of His229 was rotated 180° about the Cβ–Cγ bond. Grey diamonds and filled squares stand for compounds in the training and prediction set, respectively.

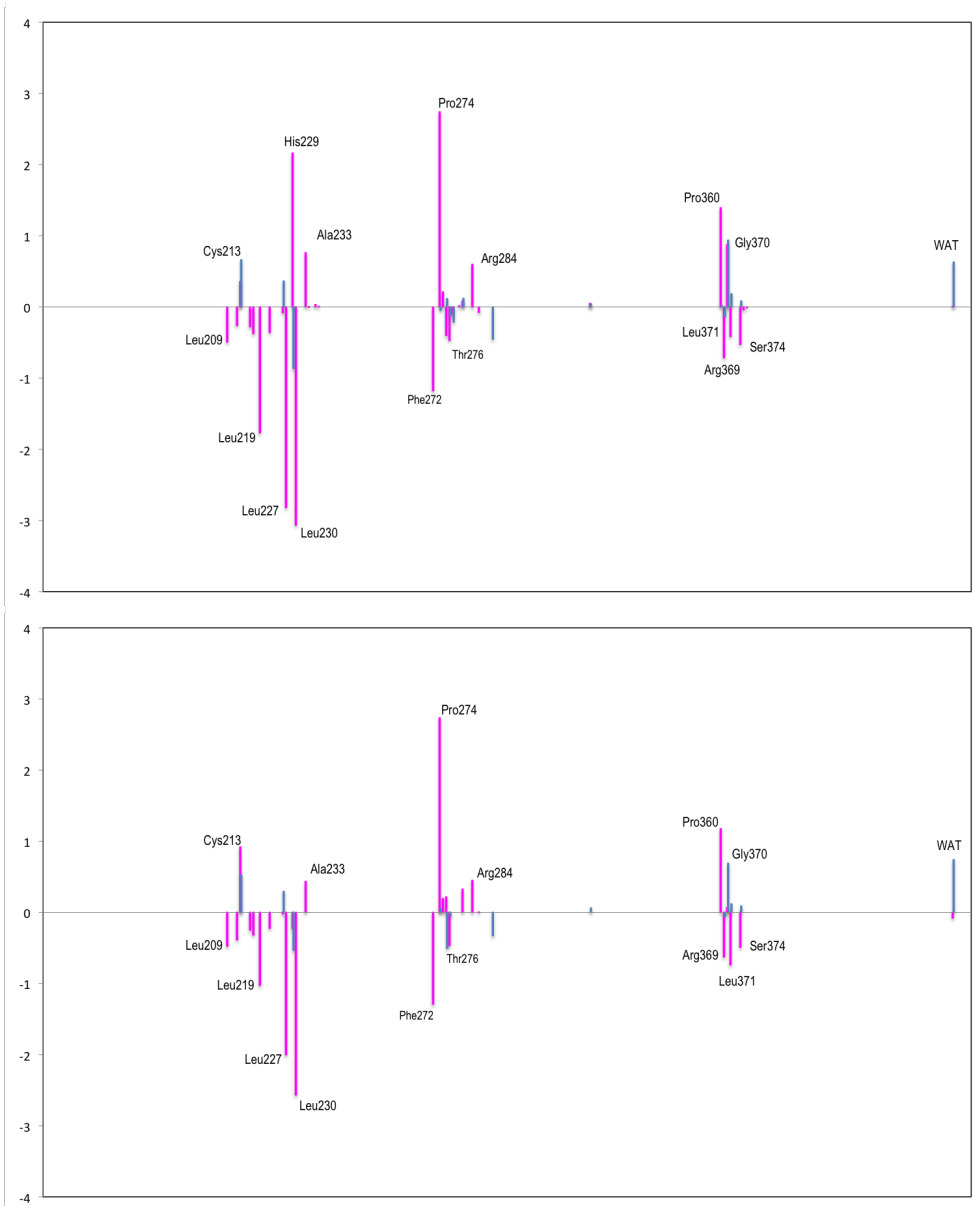


Figure S2. B. PLS coefficients for the two COMBINE models discussed in the main text, i.e. the original one (*top*) and that corresponding to complexes in which the imidazole ring of His229 has been rotated 180° about the Cβ–Cγ bond (*bottom*). Van der Waals and electrostatic interactions are colored pink and blue, respectively. The amino acids lining the binding pocket that contribute the most to explaining the differences in binding enthalpies have been labeled.

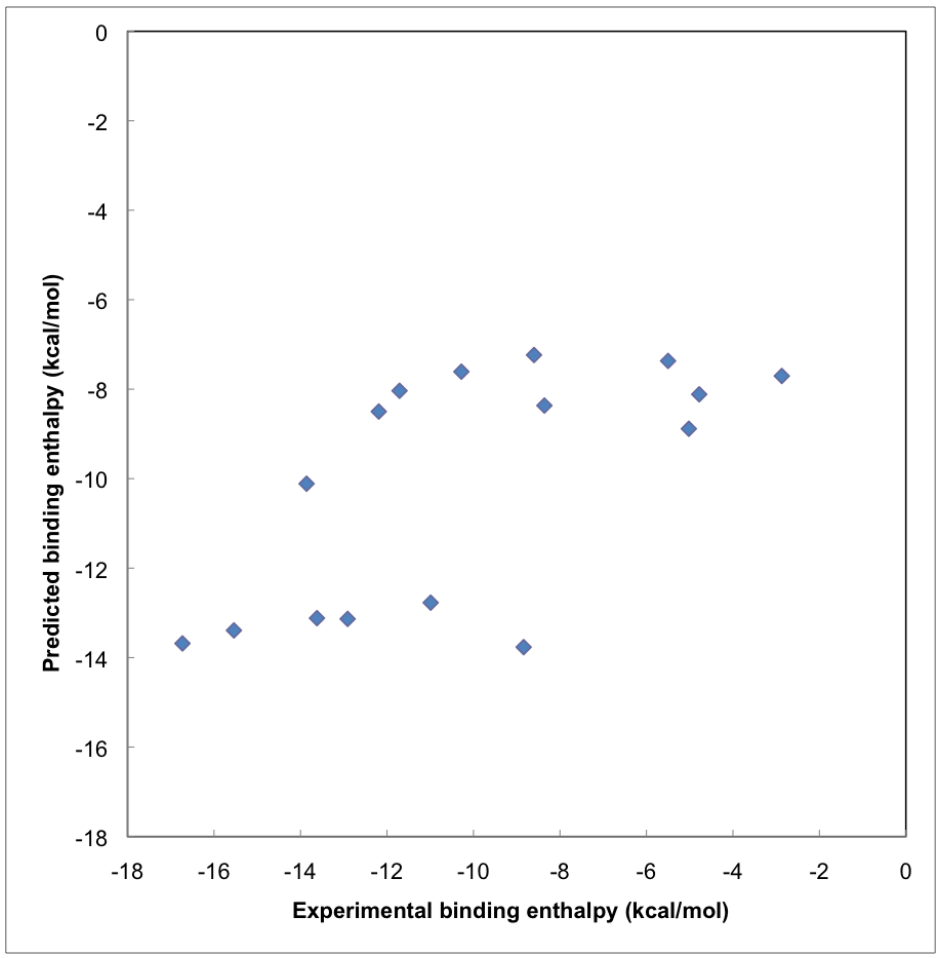


Figure S3. Plot of calculated binding enthalpies in the COMBINE model (1 LV, $r^2 = 0.422$, $q^2 = 0.371$; average r^2=0.422±0, q^2=0.198±0.08 and 1±0 LV for the remaining 9 runs) obtained using the complex reported in ref. 12 as the template vs. experimental values.

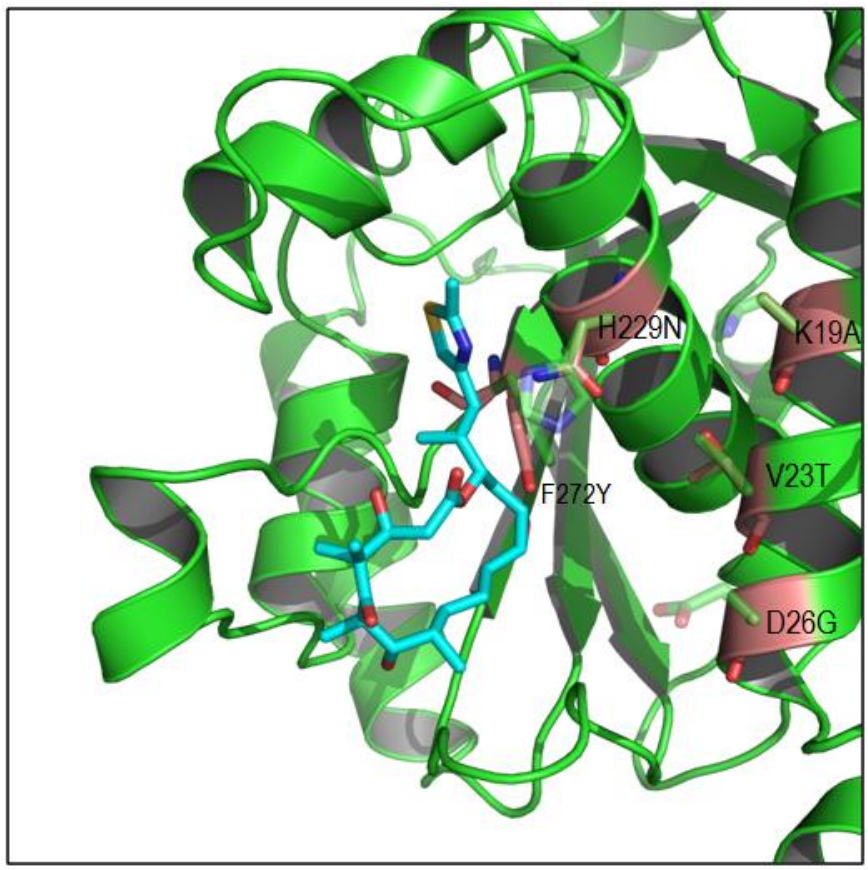


Figure S4. Cartoon representation of mammalian β-tubulin (green) as found in our modeled complex with EPA (C atoms in cyan). Labeled side chains correspond to amino acids that have been mutated in yeast tubulin (brown) to their counterparts in mammalian tubulin to make binding of taxanes possible. On the one hand, mutations located on H1 (i.e. Lys19←Ala, Val23←Thr and Asp26←Gly) do not affect EP binding because none of these residues directly participates in ligand stabilization. On the other hand, the His229←Asn and Phe272←Tyr mutations are either neutral or they could help stabilize the binding mode by means of hydrogen bonding interactions with the thiazole and the epoxide oxygen, respectively. In fact, the Nδ atoms of His229 and its Asn counterpart nicely superimpose, as do the aromatic rings of Phe272 and Tyr.

ARTÍCULO IV

"Tubulin-based Structure-affinity Relationships for Antimitotic Vinca Alkaloids."

Anticancer Agents Med Chem. 2012 Mar 1;12(3):219-25.
[PMID: 22044006]

Los alcaloides de la Vinca, originalmente extraídos de la Vinca de Madagascar, son un grupo de agentes antitumorales ampliamente empleados que alteran la dinámica de los microtúbulos en células de mamífero al interferir en el ensamblado de los heterodímeros de α,β-tubulina. Estos alcaloides promueven la formación de tubulina curvada que desestabiliza los microtúbulos e induce la aparición de agregados en forma de espiral. Sus perfiles de energía de unión han sido caracterizados mediante ensayos de velocidad de sedimentación y el sitio de unión de Vinblastina en la interfaz entre dos dimeros de tubulina ($\alpha_1\beta_1$-$\alpha_2\beta_2$) se ha determinado mediante cristalografía de rayos X en complejos de tubulina unido a un dominio similar a estathmina de la proteína RB3, aunque a baja resolución. Hemos utilizado técnicas de modelado molecular y simulación para construir, refinar y analizar comparativamente las estructuras tridimensionales de los complejos de Vinblastina, Vincristina, Vinorelbina y Vinflunina con la interfaz $\beta_1\alpha_2$-tubulina en disolución acuosa para intentar racionalizar las diferencias en afinidad de unión en términos tanto estructurales como energéticos. Nuestros resultados arrojan algo de luz adicional sobre los factores determinantes de la unión y la relación estructura-actividad de estos agentes de gran utilidad clínica.

Tubulin-based Structure-affinity Relationships for Antimitotic *Vinca* Alkaloids

Claire Coderch[1], Antonio Morreale[2] and Federico Gago[1,*]

[1]*Departmento de Farmacología, Universidad de Alcalá, 28871 Alcalá de Henares, Madrid, Spain;* [2]*Unidad de Bioinformática, Centro de Biología Molecular Severo Ochoa, 28049 Tres Cantos, Madrid, Spain*

Abstract: The *Vinca* alkaloids are a group of widely used anticancer drugs, originally extracted from the Madagascar periwinkle, that disrupt microtubule dynamics in mammalian cells by interfering with proper assembly of α,β-tubulin heterodimers. They favor curved tubulin assemblies that destabilize microtubules and induce formation of spiral aggregates. Their binding energy profiles have been characterized by means of sedimentation velocity assays and the binding site of vinblastine at the interface between two tubulin dimers ($\alpha_1\beta_1$–$\alpha_2\beta_2$) has been ascertained by X-ray crystallographic studies on a complex of tubulin with the stathmin-like domain of protein RB3, albeit at relatively low resolution. Here we use molecular modeling and simulation techniques to build, refine and perform a comparative analysis of the three-dimensional complexes of vinblastine, vincristine, vinorelbine and vinflunine with a $\beta_1\alpha_2$-tubulin interface in explicit water to rationalize the binding affinity differences in structural and energetic terms. Our results shed some more light into the binding determinants and the structure-activity relationships of these clinically useful agents.

Keywords: Tubulin, antimitotic drugs, *Vinca* alkaloids, computer simulations, binding energy analysis, molecular dynamics.

[#]**Author Profile:** Federico Gago studied Pharmacy at Complutense University, Madrid, and followed post-doctoral studies at the Physical Chemistry Laboratory, Oxford University. He is currently a Full Professor in the Department of Pharmacology at the University of Alcalá, near Madrid, where he pursues research work in the areas of computer simulations of biomolecular systems and structure-based drug design.

INTRODUCTION

Microtubules (MT) are key cell structures found in eukaryotes that play a vital role in the organization of the spatial distribution of organelles throughout interphase and of chromosomes during cell division. MT assemble from protofilaments built through longitudinal head-to-tail juxtaposition of α,β-tubulin heterodimers. α- and β-tubulin are the most conserved proteins in eukaryotes and they belong, along with their prokaryotic ancestors, to the distinct tubulin superfamily of GTPases, whose common structure is an N-terminal, or nucleotide binding domain, connected by a core helix (H7) to a C-terminal domain [1]. Each monomer binds a GTP molecule [2] but, whereas the nucleotide bound to the α subunit is nonhydrolysable and nonexchangeable, that bound to the β subunit is hydrolysable and exchangeable in unassembled tubulin heterodimers. This is so because, upon head-to-tail assembly of two heterodimers, residues from the α subunit complete the functional architecture of the GTP hydrolysis site in the β subunit. Subsequently, once the γ-phosphate group has been cleaved, GDP is sequestered at the nucleotide exchangeable site and MT disassemble by peeling outwards. Catastrophe ensues when the disassembly rate dominates over assembly formation and this occurs because the once straight GTP-bound tubulin heterodimers curve following hydrolysis of the GTP nucleotide [3]. Since GTP hydrolysis accompanies MT assembly, and GDP-tubulin is released upon depolymerization, tubulin has to exchange its bound GDP with GTP in order to assemble again. These two properties, GTP hydrolysis and nucleotide exchange, are therefore intrinsic to tubulin. MT dynamics is also affected by the presence of the cell-cycle regulating protein stathmin, which interacts with two α,β-tubulin heterodimers to form a tight ternary complex, the so-called T2S complex, that is unsuitable for MT formation. Stathmin phosphorylation on Ser16, Ser25, Ser38 and Ser63 by cell cycle kinases at the onset of mitosis weakens this association and the increased concentration of tubulin available in the cytoplasm allows the assembly of the mitotic spindle [4]. Another regulatory mechanism in cell proliferation is the alteration of the rate of catastrophe brought about by the existence of different β-tubulin isotypes [5, 6]. Moreover, recent evidence has demonstrated the presence of unhydrolysed GTP-bound tubulin not only at the growing end of MT but also in older parts where it could play a role in the rescue events that recover these structures from catastrophe [7].

MT dynamics, and therefore cell division, can also be perturbed by small molecules, and this interference can lead to cell death [8]. This antimitotic property is, in fact, exploited by a number of potent anticancer chemotherapeutic drugs, which are usually divided into two groups: (i) MT-stabilizing agents that prevent depolymerization, and (ii) MT-depolymerizing agents that inhibit MT formation. The former bind to polymerized MT, stabilize the M-loop that is responsible for lateral interactions between neighboring protofilaments, and prevent depolymerization even after GTP hydrolysis [9]. A common characteristic of these drugs is that they have two binding sites: an external site of lower affinity located at the MT pore to which they bind before being internalized to the lumen, and a higher-affinity luminal site located inside the MT [10-12]. The best known MT-stabilizing drugs are the clinically used taxanes, paclitaxel (Taxol®) and docetaxel (Taxotere®) [13], and the epothilone B analogue ixabepilone (IXEMPRA®) [14].

MT-destabilizing agents bind to unpolymerized tubulin and block MT formation. Two distinct binding sites were early recognized, with affinities for colchicine and the so-called *Vinca* alkaloids, respectively. Colchicine is a secondary metabolite produced by plants of the genus *Colchicum* that has been known for centuries for its pain-relieving effects in acute gout flares and has also been used to induce polyploidy in plant cells because it inhibits chromosome segregation during meiosis. Vinblastine (a.k.a. vincaleukoblastine, VLB) and vincristine (VNC), on the other hand, were originally isolated from the Madagascar periwinkle (*Catharanthus roseus*, basionym *Vinca rosea*) and have been in clinical use for the chemotherapy of a number of hematological and solid tumors for many years [15]. The capacity of these alkaloids to arrest cells in metaphase is due to the fact that they inhibit the assembly and dynamics of MT.

*Address correspondence to this author at the Departmento de Farmacología, Universidad de Alcalá, 28871 Alcalá de Henares, Madrid, Spain;
Tel: +34-918854514; Fax: +34-918854591;
E-mail: federico.gago@uah.es

1871-5206/12 $58.00+.00

The colchicine-binding site is a pocket located within β-tubulin at the interface with the α-subunit of the same heterodimer [16]. Addition of colchicine to steady-state MT inhibits their growth and induces their disassembly into heterodimers [17]. Several crystal structures (Protein Data Bank [PDB] entries 1SA0, 1SA1, 3HKC, 3HKD, 3HKE, 3N2K, and 3N2G) have been solved of protein complexes comprising the stathmin-like domain of protein RB3 (RB-SLD) and two tubulin heterodimers with ligands bound at the colchicine-binding site. These complexes show a curved arrangement of tubulin that prevents it from adopting the straight conformation needed for MT formation [18]. Soaking of some of these crystals with VLB [19] and use of synchroton radiation allowed Knossow *et al*. to determine the binding site for this alkaloid at 4.1 Å resolution: VLB is located at the longitudinal interface between two tubulin heterodimers next to the exchangeable GTP-binding site (E-site) [20]. This location explains why the *Vinca* alkaloids inhibit GTP hydrolysis [21], a bimolecular process in which residues from the α-tubulin subunit are used to catalyze the cleavage of the β-γ phosphodiester linkage of the GTP molecule bound to the neighboring β-tubulin subunit. The observed bent structure also accounts for the tendency of these drugs to induce self-association of tubulin into spiral aggregates at the expense of MT growth [22]. Interestingly, a bent conformation of free tubulin dimers has recently been demonstrated, which suggests that stathmin evolved to recognize curved structures in unassembled and disassembling tubulin [23].

Since the binding of natural *Vinca* alkaloids, as well as that of their semi-synthetic analogs vinorelbine (VNR) and vinflunine (VFN)[24], is linked to tubulin self-association, their affinities have been determined by sedimentation velocity experiments at different temperatures. These studies have concluded that their binding is entropically driven and that the overall affinities decrease in the order VNC > VLB > VNR > VFN, with the latter inducing the shortest spirals [25-27]. However, the affinity of the four compounds for tubulin heterodimers appears to be almost identical, the major differences among them being due to the distinct affinities of the resulting liganded heterodimers for polymerized spirals. Likewise, no preference for binding to certain tubulin isotypes over others has been observed although the binding was reported to be enhanced by the presence of GDP rather that GTP in the nucleotide binding site [28]. On the other hand, the origin of the positive binding enthalpy (ΔH) and/or the contributions to the observed enthalpy-entropy (ΔS) compensations [29, 30] remain to be established.

The binding of VLB to tubulin can be inhibited non-competitively by yet another class of compounds represented by halichondrin B, which was first extracted from a marine Japanese sponge, *Halichondria okadai* [31], and by its synthetic analogue, eribulin, which is approved for the treatment of metastatic breast cancer [32]. These agents bind to β-tubulin near the exchangeable GTP-binding site [33] but do not affect the binding of colchicine [34]. They suppress MT growth and sequester tubulin into nonfunctional aggregates [35] but, unlike the *Vinca* alkaloids, they have no effect on MT shortening.

In this study we have tried to shed some more light into the binding of VLB, VNC, VNR and VFN (Fig. **1**) to tubulin heterodimers using molecular modeling techniques, molecular dynamics (MD) simulations, continuum electrostatics calculations and energy decomposition analysis.

METHODOLOGY

The crystal structure of the VLB-colchicine-tubulin-RB-SLD assembly at 4.1 Å resolution (PDB entry 1Z2B) was used as a template to build molecular models of the whole set of complexes. For computational limitations, our systems comprised the GDP-bound β-tubulin subunit of the "bottom" heterodimer (β_1), the GTP-Mg^{2+}-bound α-tubulin subunit of the "top" heterodimer (α_2) and the respective *Vinca* alkaloid bound at the interface between them (Fig. **2**). VNC, VNR and VFN were modeled inside the binding site by editing and refining the three-dimensional structure of VLB. The molecular graphics program PyMOL version 0.99 (DeLano Scientific, LLC, Palo Alto, CA) was employed for visualization and model building.

Addition of missing hydrogen atoms and computation of the protonation state of titratable groups in α and β-tubulin at pH 6.5 were carried out using the H++ Web server [36], which relies on AMBER [37] force-field parameters and finite difference solutions

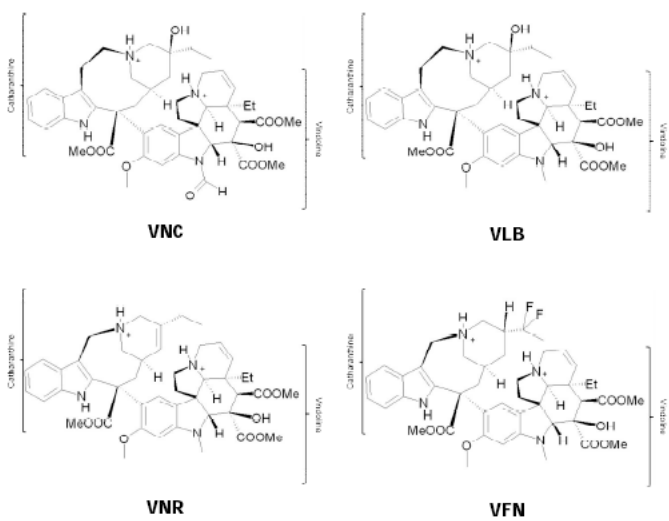


Fig. (1). Chemical formulae of the four *Vinca* alkaloids studied: vinblastine (VLB), vincristine (VNC), vinorelbine (VNR) and vinflunine (VFN).

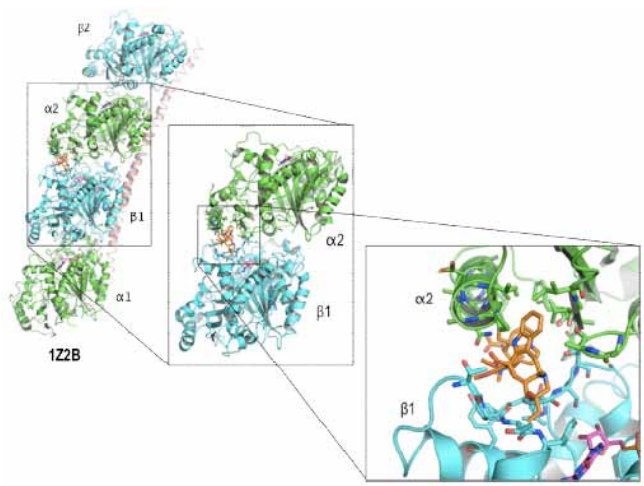


Fig. (2). PyMOL representation of the two αβ-tubulin heterodimers (α-tubulin in green, β-tubulin in cyan, and GTP and GDP nucleotides in magenta) stabilized by colchicine (represented in grey sticks) and bound to VLB (represented as orange sticks) and the stathmin-like domain of protein RB3 (pink helix), as found in PDB entry 1Z2B. The central figure represents the system that was extracted and solvated for our simulations.

to the Poisson-Boltzmann equation [38]. The charge distribution for the ligands studied was obtained by fitting the quantum mechanically calculated (RHF/6-31G*//RHF/3-21G*) molecular electrostatic potential (MEP) of the geometry-optimized molecule to a point charge model, as implemented in Gaussian 03 (Gaussian, Inc., Wallingford, CT). Consistent bonded and non-bonded AMBER parameters for the *Vinca* alkaloids were assigned by analogy or through interpolation from those already present in the AMBER database (ff03). Each molecular system was immersed in a truncated octahedron containing ~32,300 TIP3P water molecules [39] and 23 Na^+ ions [40] to achieve system electroneutrality. The *sander* and *pmemd* modules from the AMBER10 suite (http://ambermd.org/) were used for the restrained and unrestrained MD simulations, respectively. Periodic boundary conditions were applied and electrostatic interactions were treated using the smooth particle mesh Ewald method [41] with a grid spacing of 1 Å. The cutoff distance for the non-bonded interactions was 9 Å, the SHAKE [42] algorithm was applied to all bonds, and an integration step of 2.0 fs was used throughout. After an initial energy minimization of the water molecules and counterions, the system was heated to 300 K in 25 ps after which the solvent was allowed to redistribute around the positionally restrained solute for 220 ps. After this time, only the protein Cα atoms were restrained with a harmonic force constant of 10 kcal/mol·$Å^2$ so as to explore the mutual adaptation between ligand and amino acid side chains without altering the overall conformation of the dimer. Snapshots from each 10-ns MD trajectory were collected every 20 ps for further analysis. For visualization purposes, a representative average structure for each complex was calculated and refined using energy minimization.

To get an estimate of the free energy change that describes tubulin-ligand binding we calculated the difference between the free energy of the complex and that of the respective binding partners using a hybrid molecular mechanics (MM)/generalized Born-surface area (GBSA) approach, as implemented in AMBER 10 [43, 44]. Energy values were calculated as the averages over 200 snapshots from the middle part of the production phase of the MD trajectory for each complex. The van der Waals contribution to the binding energy was represented as the sum of the total ΔG_{vdW} calculated in the gas phase and the total ΔG_{surf} calculated from the solvent-accessible surface area (SASA) [45], which was determined using a water probe of radius 1.4 Å. The electrostatic contribution to the binding energy was estimated as the sum of the total ΔG_{ele} calculated in the gas phase and the total ΔG_{GB} calculated by solving the generalized Born equation [46] using dielectric constants of 1 and 78.5 for solute and solvent, respectively.

To study the conformational space of free VLB in aqueous solution, the drug was extracted from the complex and, in its bound conformation, immersed in a cubic box containing ~1,400 TIP3P water molecules and a chloride ion to achieve electroneutrality. The energy of this system was minimized in the AMBER force field prior to running unrestrained MD simulations for 100 ns using the same conditions as for the complexes. The trajectories were then processed with the *ptraj* module in the AMBER 11 suite to estimate the root-mean-square deviation (RMSD) from the bound initial geometry and to cluster the resulting conformers of the drug in bulk solvent. The relative energies of each conformer were extracted and used to calculate the probability of the microstates as defined by the Boltzmann distribution formula:

$$p_i = n\, e^{-Ei/RT} / \Sigma_j\, n\, e^{-Ei/RT}$$

where n is the number of structures belonging to each cluster i, E_i is the energy of the average structure extracted from i, T is the temperature in Kelvin, R is the gas constant expressed in kcal/mol, and j is the total number of clusters.

RESULTS AND DISCUSSION

The 100-ns MD simulation of VLB in aqueous solution provided conformations that did not differ from the initial structure by more than 1.5 Å of RMSD. The clustering analysis identified three main conformers differing in just the value of the torsional angle relating the catharanthine and vindoline domains. The most populated one, by far, was precisely that found in the complex with tubulin in the crystal structure. Thus, we can state that the VLB structure, and by extension that of the other *Vinca* alkaloids, is quite rigid and that the conformation of the ligand in aqueous solution is likely to be the same as that present in the VLB-tubulin complex (Fig. **3**). Therefore we can safely assume that these compounds are

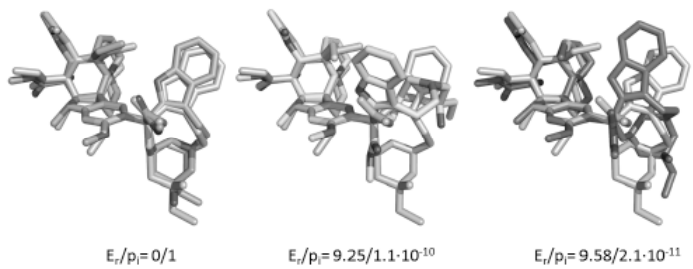


Fig. (3). Best-fit superposition of the three conformers of VLB extracted from the MD simulation in aqueous solution over the initial tubulin-bound structure (light gray). The ratio between relative energy of the conformer (E_r) and the probability of the microstate according to the Boltzmann distribution (p_i) is shown.

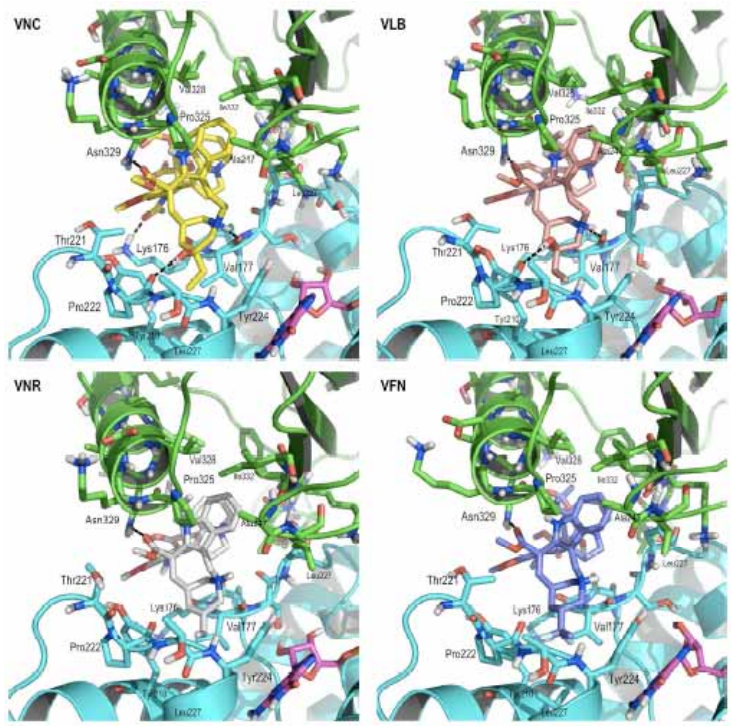


Fig. (4). Detail of the drug-binding site at the $\beta_1\alpha_2$-tubulin (cyan/green) interface in each of the complexes studied after the MD simulation and subsequent energy refinement of the average structure. VNC (yellow), VLB (pink), VNR (white), VFN (blue), and GDP (magenta) are shown as sticks as well as the amino acids charted in the energy decomposition analyses. Only polar hydrogens are displayed and hydrogen-bonding interactions are depicted as broken lines. For clarity only the main drug-interacting amino acids have been labeled.

highly preorganized for binding to curved tubulin at the $\beta_1\alpha_2$ interface between heterodimers.

In their respective complexes with $\alpha_2\beta_1$-tubulin, mutual adaptation during the MD simulations between the drugs and the side chains of the amino acids making up the binding site improved the intermolecular interactions and provided distinct details for each complex. The four *Vinca* alkaloids studied (Fig. **1**) exhibit overall similar van der Waals and electrostatic interaction energy terms (Fig. **4**) thereby defining a common pharmacophore. Their binding site is lined by the side chains of a number of mostly non-polar amino acids from both tubulin monomers. Thus, in the β subunit, $Val_{\beta1}177$, $Tyr_{\beta1}210$, $Thr_{\beta1}221$, $Pro_{\beta1}222$, $Thr_{\beta1}223$, $Tyr_{\beta1}224$ and $Leu_{\beta1}227$ are in close contact with the catharanthine domain whereas $Pro_{\beta1}175$ and $Lys_{\beta1}176$ provide an interacting surface for the vindoline domain. In the α subunit, the side chains of $Leu_{\alpha2}248$ and $Pro_{\alpha2}325$ establish contacts mainly with the indole ring in

the catharanthine domain while $Val_{\alpha 2}328$, $Asn_{\alpha 2}329$, $Ile_{\alpha 2}332$, $Ala_{\alpha 2}333$, and $Val_{\alpha 2}353$ interact with both drug domains. There are two common intermolecular hydrogen bonding interactions: one between the carboxamide of $Asn_{\alpha 2}329$ and the methyl-ester in the catharanthine domain, and another between the charged amino group in the catharanthine domain and the backbone carbonyl of $Val_{\beta 1}177$. But while this latter interaction is a direct hydrogen bond in the complexes with VNC and VLB, the slight change of geometry induced by the shortening by one carbon of the seven-member ring in the catharanthine domain of VNR and VFN causes this interaction to be water-mediated, which could be translated into a lower affinity. Importantly, the formyl group on the indole ring of VNC that replaces the methyl in VLB is able to engage in a hydrogen-bonding interaction with the positively charged amino group of $Lys_{\beta 1}176$. This must be relevant for the *in vivo* action of these two drugs because VNC and VLB, despite this minor structural variation, are endowed with distinct spectra of pharmacological activity and dose-related toxicities, as well as differential cellular uptake and retention characteristics [47, 48].

The binding features of the four structurally similar analogs can be compared pairwise because the differences among them are mostly confined to the number of carbon atoms in the central saturated ring of the catharanthine domain and the presence or absence of a hydroxyl group in the common six-membered ring of this same domain. Thus, VNR and VFN lack the hydroxyl in their six-membered ring (Fig. **1**) whereas VNC and VLB have a seven-membered ring with an attached hydroxyl that acts as a hydrogen bond donor to the carbonyl oxygen of $Pro_{\beta 1}222$. On the other hand, the saturated double bond and the two fluorine atoms that differentiate VFN from VNR provide only a marginally improved electrostatic interaction with $Val_{\beta 1}177$ and $Tyr_{\beta 1}224$ (Fig. **5**). Taken together, our results are therefore in overall semiquantitative agreement with the binding affinities measured experimentally (VNC> VLB > VNR > VFN).

It must be realized that in curved tubulin the binding site we have studied is exposed to the solvent and displacement of any bound water molecules by drug binding must be accompanied by the release of these water molecules into the bulk solvent, which is thought to result in a significant entropic contribution to the binding energy. The same can be said about the water molecules in contact with the hydrophobic surfaces of the drugs that are released upon complex formation. These two processes are likely to underlie the experimental observation that binding of these drugs to tubulin is an entropically driven process [25, 26, 49]. For the better studied HIV protease, it was concluded for some ligands that the origin of an unfavorable enthalpy, despite the fact that intrinsic interactions were favorable, was the energy cost of rearranging the flap region in the enzyme and that the dominant binding force was the increase in solvent entropy that accompanies the burial of a significant hydrophobic surface [30].

One limitation of our first approximation to the binding of these agents to tubulin is, of course, the restraints that we imposed on the Cα trace of the protein. This was done to prevent any major artifactual distortions on tubulin in the absence of protein RB3, which was not included in our simulations for computational efficiency. This limitation may have prevented a deeper burial of both VNR and VFN in the binding site so as to establish a direct hydrogen bond between the charged amino group in the catharanthine domain and the backbone carbonyl of $Val_{\beta 1}177$. If this were the case, the wedge effect at the $\beta_1\alpha_2$-tubulin interface upon drug binding would be greater for VNR and VFN than that observed for VLB and VNC, and this may well account for the finding that the spiraling potential of VFN, a property that happens to be inversely related to the clinically used dosage, is the lowest of all [25, 50].

Since it is well known that MT's behavior depends on β-tubulin isotype composition [5], one could have expected differential interactions between these drugs and tubulins of diverse amino acid composition at both the inter-dimer and intra-dimer interfaces [6]. However, when we performed a multiple sequence alignment (Fig. **6**) and mapped the drug pharmacophoric region onto it no differences were found. Therefore, and taking into account that these alkaloids bind to unpolymerized curved tubulin, their different

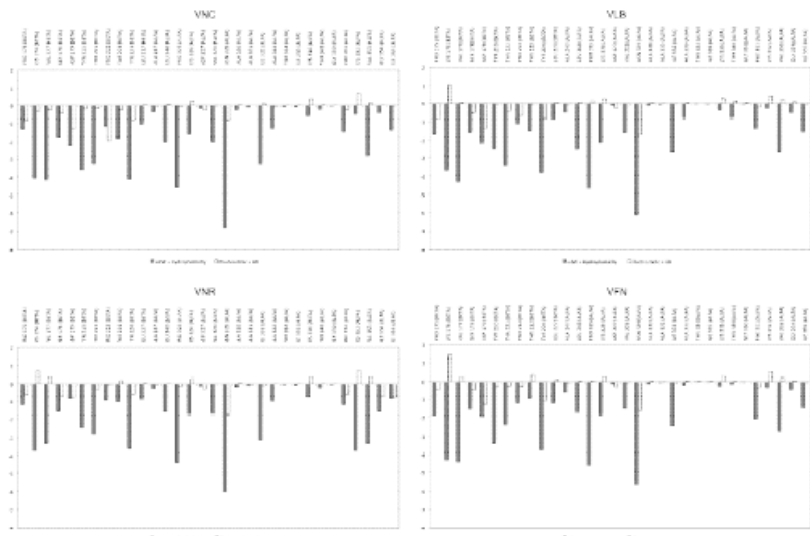


Fig. (5). Residue-based free energy decomposition for the binding of the four *Vinca* alkaloids to the $\beta_1\alpha_2$-tubulin interface. The filled bars represent van der Waals energy + hydrophobicity (SASA) while the empty bars stand for the sum of electrostatic (ΔG_{ele} in the gas phase) and desolvation (total ΔG_{GB}) energies.

Fig. (6). Multiple sequence alignment of the β-tubulin isotypes β_I (Q9H4B7), β_{IIA} (Q13885), β_{IIB} (Q9BVA1), β_{IIC} (P68371), β_{III} (Q13509) and βIV (P04350). The shaded boxes highlight the β-tubulin regions that interact with the *Vinca* alkaloids, which happen to be of identical sequence in all isotypes. Asterisks, colons, dots and blanks indicate, respectively, identical residues and conservative, semi-conservative, or non-conservative mutations.

binding affinities must surely arise from distinct kinetics of MT formation and catastrophe being exhibited by the tubulin isotypes and their mixtures rather than from changes in drug-protein interaction patterns.

CONCLUSIONS

An MD simulation of VLB in aqueous solution showed a conformational rigidity in this alkaloid that is translated into a high degree of preorganization for tubulin binding and can be safely assumed for the remaining members of the *Vinca* family of drugs. By extracting the relevant interdimer interface from the only available crystal structure of VLB-bound tubulin and using model building and simulation techniques, we explored the conformational behavior of the drug-binding site in the presence of VLB, VNC, VNR and VFN. The resulting complexes showed a distinct and consistent pattern of intermolecular interactions that were analyzed using the MM/GBSA approach. An extensive number of van der Waals contacts and a number of hydrogen-bonding interactions were apparent that nevertheless could not account for the entropy-driven free energy changes measured experimentally [25, 26]. This is probably due to enthalpy-entropy compensations arising from polar group desolvation, structural rearrangement of the binding site, and the burial of hydrophobic surfaces upon ligand binding. Nonetheless, we detected the possibility of an extra hydrogen bonding interaction between VNC and $Lys_{\beta1}176$ that is not possible for VLB and could account for the improved binding energy of the former over the latter and for the distinct biological profiles of these otherwise almost identical drugs [25]. Likewise, the absence in VNR of the interaction involving the hydroxyl group in the catharanthine domain that is present in the other alkaloids might well be responsible for the loss of binding free energy in this derivative compared to VLB [25]. Lastly, it seems that the introduction of the two fluorine atoms in VFN is not translated into an improved thermodynamic profile, in agreement with experimental data showing a worse binding enthalpy for this derivative relative to VNR and consequently to VNC and VLB [26]. The resulting pharmacophore for drug binding cannot explain any differences regarding β-tubulin isotypes, given the fact that the residues in contact with the drugs are highly conserved in the tubulin family. Hence, it is more likely that any differences observed result from variations in the kinetics of MT formation and catastrophe rather than from a distinct set of drug-protein interactions.

ACKNOWLEDGEMENTS

The authors thankfully acknowledge financial support from Comunidad Autónoma de Madrid (S-BIO/0214/2006) and Comisión Interministerial de Ciencia y Tecnología (SAF2009-13914-C02-02) as well as the computer resources, technical expertise and assistance provided by the Barcelona Supercomputing Center (Centro Nacional de Supercomputación). C.C. enjoys an FPU grant awarded by the Spanish Ministry of Science and Innovation (AP2007-01225). A.M. is grateful to Comunidad Autónoma de Madrid (Consejería de Educación) and Fundación Severo Ochoa for financial support through the AMAROUTO program.

REFERENCES

[1] Nogales, E.; Downing, K.H.; Amos, L.A.; Lowe, J. Tubulin and FtsZ form a distinct family of GTPases. *Nat. Struct. Biol.,* **1998**, *5*, 451-458.

[2] Wade, R.H. Microtubules: an overview. *Methods Mol. Med.,* **2007**, *137*, 1-16.

[3] Nogales, E.; Wang, H.W. Structural intermediates in microtubule assembly and disassembly: how and why? *Curr. Opin. Cell Biol.,* **2006**, *18*, 179-184.

[4] Rubin, C.I.; Atweh, G.F. The role of stathmin in the regulation of the cell cycle. *J. Cell Biochem.,* **2004**, *93*, 242-250.

[5] Panda, D.; Miller, H.P.; Banerjee, A.; Luduena, R.F.; Wilson, L. Microtubule dynamics *in vitro* are regulated by the tubulin isotype composition. *Proc. Natl. Acad. Sci. U S A,* **1994**, *91*, 11358-11362.

[6] Rezania, V.; Azarenko, O.; Jordan, M.A.; Bolterauer, H.; Luduena, R.F.; Huzil, J.T.; Tuszynski, J.A. Microtubule assembly of isotypically purified tubulin and its mixtures. *Biophys. J.,* **2008**, *95*, 1993-2008.

[7] Dimitrov, A.; Quesnoit, M.; Moutel, S.; Cantaloube, I.; Pous, C.; Perez, F. Detection of GTP-tubulin conformation *in vivo* reveals a role for GTP remnants in microtubule rescues. *Science,* **2008**, *322*, 1353-1356.

[8] Jordan, A.; Hadfield, J.A.; Lawrence, N.J.; McGown, A.T. Tubulin as a target for anticancer drugs: agents which interact with the mitotic spindle. *Med. Res. Rev.,* **1998**, *18*, 259-296.

[9] Wade, R.H. On and around microtubules: an overview. *Mol. Biotechnol.,* **2009**, *43*, 177-191.

[10] Buey, R.M.; Calvo, E.; Barasoain, I.; Pineda, O.; Edler, M.C.; Matesanz, R.; Cerezo, G.; Vanderwal, C.D.; Day, B.W.; Sorensen, E.J.; Lopez, J.A.; Andreu, J.M.; Hamel, E.; Diaz, J.F. Cyclostreptin binds covalently to microtubule pores and lumenal taxoid binding sites. *Nat. Chem. Biol.,* **2007**, *3*, 117-125.

[11] Magnani, M.; Maccari, G.; Andreu, J.M.; Diaz, J.F.; Botta, M. Possible binding site for paclitaxel at microtubule pores. *FEBS J.,* **2009**, *276*, 2701-2712.

[12] Barasoain, I.; Garcia-Carril, A.M.; Matesanz, R.; Maccari, G.; Trigili, C.; Mori, M.; Shi, J.Z.; Fang, W.S.; Andreu, J.M.; Botta, M.; Diaz, J.F. Probing the pore drug binding site of microtubules with fluorescent taxanes: evidence of two binding poses. *Chem. Biol.,* **2010**, *17*, 243-253.

[13] Lowe, J.; Li, H.; Downing, K.H.; Nogales, E. Refined structure of alpha beta-tubulin at 3.5 A resolution. *J. Mol. Biol.,* **2001**, *313*, 1045-1057.

[14] Hunt, J.T. Discovery of ixabepilone. *Mol. Cancer Ther.,* **2009**, *8*, 275-281.

[15] Seam, P.; Janik, J.E.; Longo, D.L.; Devita, V.T., Jr. Role of chemotherapy in Hodgkin's lymphoma. *Cancer J.,* **2009**, *15*, 150-154.

[16] Ravelli, R.B.; Gigant, B.; Curmi, P.A.; Jourdain, I.; Lachkar, S.; Sobel, A.; Knossow, M. Insight into tubulin regulation from a complex with colchicine and a stathmin-like domain. *Nature,* **2004**, *428*, 198-202.

[17] Vandecandelaere, A.; Martin, S.R.; Engelborghs, Y. Response of microtubules to the addition of colchicine and tubulin-colchicine: evaluation of models for the interaction of drugs with microtubules. *Biochem. J.,* **1997**, *323 (Pt 1)*, 189-196.

[18] Dorleans, A.; Gigant, B.; Ravelli, R.B.; Mailliet, P.; Mikol, V.; Knossow, M. Variations in the colchicine-binding domain provide insight into the structural switch of tubulin. *Proc. Natl. Acad. Sci. U S A,* **2009**, *106*, 13775-13779.

[19] Cormier, A.; Knossow, M.; Wang, C.; Gigant, B. The binding of vinca domain agents to tubulin: structural and biochemical studies. *Methods Cell Biol.,* **2010**, *95*, 373-390.

[20] Gigant, B.; Wang, C.; Ravelli, R.B.; Roussi, F.; Steinmetz, M.O.; Curmi, P.A.; Sobel, A.; Knossow, M. Structural basis for the regulation of tubulin by vinblastine. *Nature,* **2005**, *435*, 519-522.

[21] Wang, C.; Cormier, A.; Gigant, B.; Knossow, M. Insight into the GTPase activity of tubulin from complexes with stathmin-like domains. *Biochemistry,* **2007**, *46*, 10595-10602.

[22] Himes, R.H. Interactions of the *Catharanthus* (*Vinca*) alkaloids with tubulin and microtubules. *Pharmacol. Ther.,* **1991**, *51*, 257-267.

[23] Barbier, P.; Dorleans, A.; Devred, F.; Sanz, L.; Allegro, D.; Alfonso, C.; Knossow, M.; Peyrot, V.; Andreu, J.M. Stathmin and interfacial microtubule inhibitors recognize a naturally curved conformation of tubulin dimers. *J. Biol. Chem.,* **2010**, *285*, 31672-31681.

[24] Fahy, J.; Duflos, A.; Ribet, J.-P.; Jacquesy, J.-C.; Berrier, C.; Jouannetaud, M.-P.; Zunino, F. Vinca alkaloids in superacidic media: a method for creating a new family of antitumor derivatives. *J. Am. Chem. Soc.,* **1997**, *119*, 8576-8577.

[25] Lobert, S.; Vulevic, B.; Correia, J.J. Interaction of vinca alkaloids with tubulin: a comparison of vinblastine, vincristine, and vinorelbine. *Biochemistry,* **1996**, *35*, 6806-6814.

[26] Lobert, S.; Ingram, J.W.; Hill, B.T.; Correia, J.J. A comparison of thermodynamic parameters for vinorelbine- and vinflunine-induced tubulin self-association by sedimentation velocity. *Mol. Pharmacol.,* **1998**, *53*, 908-915.

[27] Taylor, R.E.; Daly, E.M. Entropy and enthalpy in the activity of tubulin-based antimitotic agents. *Curr. Chem. Biol.,* **2009**, *3*, 47-59.

[28] Lobert, S.; Frankfurter, A.; Correia, J.J. Energetics of vinca alkaloid interactions with tubulin isotypes: implications for drug efficacy and toxicity. *Cell Motil. Cytoskeleton,* **1998**, *39*, 107-121.

[29] Ferenczy, G.G.; Keseru, G.M. Thermodynamics guided lead discovery and optimization. *Drug Discov. Today,* **2010**, *15*, 919-932.

[30] Luque, I.; Todd, M.J.; Gomez, J.; Semo, N.; Freire, E. Molecular basis of resistance to HIV-1 protease inhibition: a plausible hypothesis. *Biochemistry,* **1998**, *37*, 5791-5797.

[31] Uemura, D.; Takahashi, K.; Yamamoto, T.; Katayama, C.; Tanaka, J.; Okumura, Y.; Hirata, Y. Norhalichondrin-A - an antitumor polyether macrolide from a marine sponge. *J. Am. Chem. Soc.,* **1985**, *107*, 4796-4798.

[32] Fornier, M.N. Approved agents for metastatic breast cancer. *Semin. Oncol.,* **2011**, *38 Suppl 2*, S3-10.

[33] Bai, R.; Nguyen, T.L.; Burnett, J.C.; Atasoylu, O.; Munro, M.H.; Pettit, G.R.; Smith, A.B.; Gussio, R.; Hamel, E. Interactions of halichondrin B and eribulin with tubulin. *J. Chem. Inf. Model.,* **2011**.

[34] Luduena, R.F.; Roach, M.C.; Prasad, V.; Pettit, G.R. Interaction of halichondrin B and homohalichondrin B with bovine brain tubulin. *Biochem. Pharmacol.,* **1993**, *45*, 421-427.

[35] Jordan, M.A.; Kamath, K.; Manna, T.; Okouneva, T.; Miller, H.P.; Davis, C.; Littlefield, B.A.; Wilson, L. The primary antimitotic mechanism of action of the synthetic halichondrin E7389 is suppression of microtubule growth. *Mol. Cancer Ther.,* **2005**, *4*, 1086-1095.

[36] Gordon, J.C.; Myers, J.B.; Folta, T.; Shoja, V.; Heath, L.S.; Onufriev, A. H++: a server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Res.,* **2005**, *33*, W368-371.

[37] Cornell, W.D.; Cieplak, P.; Bayly, C.I.; Gould, I.R.; Merz, K.M.; Ferguson, D.M.; Spellmeyer, D.C.; Fox, T.; Caldwell, J.W.; Kollman, P.A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.,* **1995**, *117*, 5179-5197.

[38] Bashford, D.; Karplus, M. pKa's of ionizable groups in proteins: atomic detail from a continuum electrostatic model. *Biochemistry,* **1990**, *29*, 10219-10225.

[39] Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys,* **1983**, *79*, 926–935.

[40] Aqvist, J. Ion-water interaction potentials derived from free energy perturbation simulations. . *J. Phys. Chem.,* **1990**, *94*, 8021-8024.

[41] Hansch, C. The advent and evolution of QSAR at Pomona College. *J. Comput. Aided Mol. Des.,* **2011**, *25*, 495–507.

[42] Lill, M.A. Efficient incorporation of protein flexibility and dynamics into molecular docking simulations. *Biochemistry,* **2011**, *50*, 6157-6169.

[43] Srinivasan, J.; Cheatham, T.E.; Cieplak, P.; Kollman, P.A.; Case, D.A. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate - DNA helices. *J. Am. Chem. Soc.,* **1998**, *120*, 9401-9409.

[44] Kollman, P.A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D.A.; Cheatham, T.E., 3rd Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.,* **2000**, *33*, 889-897.

[45] Sitkoff, D.; Sharp, K.A.; Honig, B. Accurate calculation of hydration free-energies using macroscopic solvent models. *J. Phys. Chem.,* **1994**, *98*, 1978-1988.

[46] Onufriev, A.; Bashford, D.; Case, D.A. Modification of the generalized Born model suitable for macromolecules. *J. Phys. Chem. B,* **2000**, *104*, 3712-3720.

[47] Ferguson, P.J.; Phillips, J.R.; Selner, M.; Cass, C.E. Differential activity of vincristine and vinblastine against cultured cells. *Cancer Res.,* **1984**, *44*, 3307-3312.

[48] Ferguson, P.J.; Cass, C.E. Differential cellular retention of vincristine and vinblastine by cultured human promyelocytic leukemia HL-60/Cl cells: the basis of differential toxicity. *Cancer Res.,* **1985**, *45*, 5480-5488.

[49] Lobert, S.; Ingram, J.W.; Correia, J.J. The thermodynamics of *Vinca* alkaloid-induced tubulin spirals formation. *Biophys. Chem.,* **2007**, *126*, 50-58.

[50] Lobert, S.; Fahy, J.; Hill, B.T.; Duflos, A.; Etievant, C.; Correia, J.J. *Vinca* alkaloid-induced tubulin spiral formation correlates with cytotoxicity in the leukemic L1210 cell line. *Biochemistry,* **2000**, *39*, 12053-12062.

TRABAJO EN PREPARACIÓN

"Molecular modeling of the interaction of PM060327 with tubulin"

(trabajo no concluido)

PM050489 y PM060184 son miembros de una nueva familia de compuestos antimitóticos aislados de la esponja marina Lithoplocamia lithistoides, y PM060327 es un análogo más soluble sintetizado por PharmaMar S.A.U. (Colmenar Viejo, Madrid, Spain). PM050489 y PM060184 inducen disrupción del ciclo celular en la fase G_2/M con valores de IC_{50} en el rango de subnanomolar, y su dominio de unión solapa con el de los alcaloides de la Vinca. Se han establecido modelos moleculares plausibles para los complejos de PM060327 con tubulina mediante estudios de acoplamiento y simulación por dinámica molecular.

Molecular modeling of the interaction of PM060327 with tubulin.

Introduction

PM050489 and PM060184 (Figure 1) are members of a new family of antimitotic compounds originally isolated from the marine sponge *Lithoplocamia lithistoides*. PM060327 (Figure 1) is a more water-soluble analog that has been synthesized at PharmaMar S.A.U. (Colmenar Viejo, Madrid, Spain).

PM050489 and PM060184 induce disruption of the microtubule (MT) network and arrest cells at the G_2/M phase of the cell cycle with IC_{50} values in the subnanomolar range.[1] The binding site for these drugs was previously shown to overlap with that used by vinblastine (VLB).[1] Molecular models for the complex of tubulin with PM060327 were derived using docking tools and molecular dynamics simulations.

Methods.

Modeling of the binding mode to the β_1:α_2-tubulin interdimer interface. Our computer modeling and simulation studies started with the high-affinity analogue PM050489. Atomic point charge calculations and geometry optimizations for all ligands were carried out as previously described.[2,3] The conformational space available to PM050489 in aqueous solution was studied by means of a 50-ns molecular dynamics (MD) simulation at 300 K. The resulting conformations were automatically docked at the β_1:α_2-tubulin interdimer interface using the Lamarckian genetic algorithm implemented in AutoDock 3.0.5 [4] by allowing random changes of overall orientation. This interface was generated from the phomopsin A-colchicine-tubulin-RB-SLD assembly solved at 4.1 Å resolution (PDB entry 3DU7) upon removal of colchicine and phomopsin A. The latter drug shares the binding site with the vinca alkaloids.[5] The GDP molecule was conserved and the protein side-chains in the nucleotide-binding site were slightly reoriented so that they established the same interactions that are observed in the 2.1 Å resolution X-ray crystal structure of tubulin-RB3-SLD complex (PDB entry 3RYC). To explore the mutual

adaptation between ligand and amino acid side-chains without altering the overall conformation of the dimer, to improve the intermolecular interactions and to provide distinct details for the binding of PM050489, 20-ns MD simulations were carried out for several docking solutions. At the beginning protein C_α atoms were restrained with a harmonic force constant of 5 kcal·mol^{-1}·Å^{-2}. Thereafter all of the restraints were removed and 80-ns MD simulations were carried out under the same conditions reported for several vinca alkaloid-tubulin complexes.[3] The binding mode for one of the complexes remained stable after a total of 100 ns of simulation, and this system was gradually cooled down from 300K to 273K over 1.2 ns, as reported previously for several taxane-tubulin complexes.[2] Energy minimization of the resulting cooled complex employing 10000 cycles of steepest descent followed by 40,000 cycles of conjugate gradient provided the representative structure that was then used as template to model the complex of PM060327 with tubulin. This new system was simulated for 50 ns and cooled under the same conditions of free MD simulation as the previous to provide representative structures. The molecular graphics program PyMOL (v. 0.99rc6, DeLano Scientific, LLC, Palo Alto, CA) was used for visualization and molecular editing.

Modeling of the binding mode to α_2-tubulin. Removal of the solvent, counterions and the β_1-tubulin subunit from the previous refined complex provided the starting structure for the α_2-tubulin-PM050489 model, which was immersed in a truncated octahedron of water molecules, neutralized and equilibrated. The system was first simulated for 10 ns during which the protein C_α atoms were restrained with a harmonic force constant of 5 kcal·mol^{-1}·Å^{-2} so as to reproduce the conformational constraints brought about by the presence of the upper β_2-tubulin monomer that was not included in the simulations for computational efficiency. After this time the restraints were removed only in the protein segments that were in direct contact with the ligand so as to study the mutual adaptation without altering the overall conformation of the α_2-tubulin monomer. Once the binding mode remained stable, the same procedure as the one used for the dimer was carried out to produce the template to model the complex of PM060327 with α_2-tubulin. This new system was simulated for 50 ns and cooled as before to provide representative structures.

Modeling of the binding mode to β_1-tubulin. Removal of the solvent, counterions and the α_2-tubulin subunit from the previously refined PM060327-β_1:α_2-tubulin complex provided the starting structure, which was immersed in a truncated octahedron of water molecules and neutralized. The complex was then refined and equilibrated using the procedures described above.

Results

To account for the fact that binding of PM060327 and its analogues to tubulin displaces bound VLB,[1] we built three different complexes: (1) an $\alpha_1\beta_1$:$\alpha_2\beta_2$ dimer of dimers (tetramer) with the ligand bound at the β_1:α_2 interface (Figure 2A), (2) an αβ dimer with the ligand bound to the α-subunit (Figure 2B), and (3) an αβ dimer with the ligand bound to the β-subunit (Figure 3).

(1) PM060327 is proposed to adopt an extended conformation at the β_1:α_2-interdimer interface and to occupy most of the VLB-binding site[3,6] in such a way that the *tert*-butyl moiety interacts mainly with the α_2 subunit, the free hydroxyl with the β_1 subunit, and only the lactone end interacts with both α_2 and β_1 subunits. This binding site is mainly made up of the side chains of non-polar amino acids from both tubulin monomers. Thus, $Leu_{\alpha2}248$, $Pro_{\alpha2}325$, $Val_{\alpha2}328$, $Ile_{\alpha2}332$, $Asn_{\alpha2}329$, $Lys_{\alpha2}336$, $Phe_{\alpha2}351$, $Val_{\alpha2}353$ and $Ile_{\alpha2}355$ from α_2-tubulin line the hydrophobic binding pocket in which the *tert*-butyl group is lodged. In β_1-tubulin, $Pro_{\beta1}175$, $Lys_{\beta1}176$, $Val_{\beta1}177$ and $Ser_{\beta1}178$ provide an interacting surface for the terminal lactone whereas $Tyr_{\beta1}210$, $Phe_{\beta1}214$, $Pro_{\beta1}222$, $Thr_{\beta1}223$, $Tyr_{\beta1}224$ and $Leu_{\beta1}227$ make up a hydrophobic pocket where the terminal Cl-butenyl moiety is buried. The two central amide bonds play a key role in the stabilization of this binding pose because the NH at position 15 and the carbonyl oxygen at position 14 establish hydrogen bonds, respectively, with the $Asn_{\alpha2}329$ carboxamide oxygen and the backbone NH of $Ser_{\beta1}178$. In addition, both ends of the ligand are stabilized by electrostatic interactions: the lactone carbonyl oxygen acts as a hydrogen bond acceptor for the side-chain amino group of $Lys_{\alpha2}336$, and the hydroxyl at position 22 establishes transient hydrogen bonding interactions with the backbone carbonyl

oxygen of $Pro_{\beta1}222$ and the NH of $Thr_{\beta1}223$. This binding mode we propose for PM060327 shares some features with that of VLB insofar as it involves similar electrostatic interactions with $Asn_{\alpha2}329$ and $Pro_{\beta1}222$ and also van der Waals interactions with the side chains lining the hydrophobic pockets that lodge both the *tert*-butyl group (in α_2-tubulin) and the terminal Cl-butenyl (in β_1-tubulin).[3] This docking pose accounts for the structure-activity relationship for PM060327, PM060184 and PM050489. The change from a hydroxyl at position 22 in PM060327 to a carbamate in PM050489 and PM060184 increases the affinity because the NH of this group establishes an additional hydrogen bond with the backbone carbonyl oxygen of $Pro_{\beta1}222$, while its carbonyl oxygen interacts with $Tyr_{\alpha2}357$ through a bridging water molecule that is always present during the simulations of the β_1:α_2-tubulin–PM050489 complex. Likewise, the loss of binding affinity of PM060184 compared to PM050489 is probably due to the presence of a hydrogen atom instead of a chlorine at position 25, which reduces the van der Waals interactions with the hydrophobic pocket in the β_1 subunit. When this model was compared with the high-resolution tubulin-RB3-SDL complex at 2.10 Å resolution (PDB entry 3RYC),[7] no gross structural changes were found except for small mutual adaptations. The main differences are due to the lodging of the Cl-butenyl moiety between helix H6 and the N-terminus of H7 in β_1-tubulin, and the burial of the *tert*-butyl group into the cavity present between S9 and H10 in α_2-tubulin. This pocket is occupied by similar hydrophobic groups present in VLB, phomopsin A and soblidotin in their respective complexes with colchicin-tubulin-RB3-SLD (Figure 4).[5]

(2) In the absence of the interdimer β_1-subunit, the PM060327-α_2-tubulin complex evolved in such a way that the Cl-butenyl moiety of the ligand induced a slight rearrangement of helix H8 that led to the creation and occupancy of a hydrophobic pocket made up of the side chains of $Leu_{\alpha2}248$, $Val_{\alpha2}250$, $Glu_{\alpha2}254$ (the residue responsible for GTP hydrolysis in polymerized MT),[8,9] $Phe_{\alpha2}255$, $Cys_{\alpha2}316$ and $Lys_{\alpha2}352$ (Figure 5). On the other hand, the lactone end and the *tert*-butyl group remained lodged in the hydrophobic cleft between S9 and H10, and the two hydrogen-bonding interactions involving $Asn_{\alpha2}329$ and $Lys_{\alpha2}336$ were maintained. In addition, reorientation of the Cl-butenyl group favored the formation of three additional hydrogen bonds between: (i) the carbonyl oxygen at position 14 and the backbone NH of $Val_{\alpha2}353$ in strand S9 through a

bridging water molecule, (ii) the NH at position 18 and the backbone carbonyl oxygen of this same valine, and (iii) the hydroxyl at position 22 and the side chains of $Lys_{\alpha 2}352$ and $Glu_{\alpha 2}254$ although solvation makes these latter interactions quite transient. In this proposed binding mode, PM060327 then mimics the N-terminal β-hairpin of RB3-SLD protein by occupying the shallow groove lined by strand S9 and helix H10 in α_2-tubulin in such a way that the *tert*-butyl moiety nicely superimposes with the side chains of either Phe20 or Trp20. [5,6,10-13]

(3) In the absence of the interdimer α_2-subunit, the PM060327-β_1-tubulin complex did not display any stable intermolecular hydrogen bonds that act as ligand-anchoring points and the hydrophobic interactions were limited to the burial of the Cl-butenyl into the vinca alkaloid-binding subsite (Figure 3).

Conclusions

From the molecular modeling point of view, we favor models (1) and (2) in preference to model (3). Significantly, the indole ring of Trp20 from RB3-SDL occupies the same hydrophobic pocket as the indole ring present on the catharantine domain of VLB and the 3,4-didehydrovaline and isoleucine moieties of phomopsin A and soblidotin, respectively[5] (Figure 6). This structural evidence persuades us to favor a unique binding site for PM060327 and related analogues in α-tubulin rather than an alternative, low-affinity binding subsite in β-tubulin, as defined for phomopsin A by Krebs *et al.*[5] and also proposed for halichondrin A and ERI by Hamel *et al.* [14]

***Disclaimer:** The project has been partially funded by PharmaMar, which has commercial interests in these compounds.

References

1. Pera-Gresely, B. Universidad Complutense de Madrid (2011).
2. Matesanz, R. et al. Optimization of taxane binding to microtubules: binding affinity dissection and incremental construction of a high-affinity analog of paclitaxel. *Chem Biol* **15**, 573-85 (2008).
3. Coderch, C., Morreale, A. & Gago, F. Tubulin-based Structure-affinity Relationships for Antimitotic Vinca Alkaloids. *Anticancer Agents Med Chem* (2011).
4. Morris, G. et al. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **19**, 1639-1662 (1998).
5. Cormier, A., Marchand, M., Ravelli, R.B., Knossow, M. & Gigant, B. Structural insight into the inhibition of tubulin by vinca domain peptide ligands. *EMBO Rep* **9**, 1101-6 (2008).
6. Gigant, B. et al. Structural basis for the regulation of tubulin by vinblastine. *Nature* **435**, 519-22 (2005).
7. Nawrotek, A., Knossow, M. & Gigant, B. The determinants that govern microtubule assembly from the atomic structure of GTP-tubulin. *J Mol Biol* **412**, 35-42 (2011).
8. Nogales, E., Downing, K.H., Amos, L.A. & Lowe, J. Tubulin and FtsZ form a distinct family of GTPases. *Nat Struct Biol* **5**, 451-8 (1998).
9. Anders, K.R. & Botstein, D. Dominant-lethal alpha-tubulin mutants defective in microtubule depolymerization in yeast. *Mol Biol Cell* **12**, 3973-86 (2001).
10. Barbier, P. et al. Stathmin and interfacial microtubule inhibitors recognize a naturally curved conformation of tubulin dimers. *J Biol Chem* **285**, 31672-81 (2010).
11. Dorleans, A. et al. Variations in the colchicine-binding domain provide insight into the structural switch of tubulin. *Proc Natl Acad Sci U S A* **106**, 13775-9 (2009).
12. Ravelli, R.B. et al. Insight into tubulin regulation from a complex with colchicine and a stathmin-like domain. *Nature* **428**, 198-202 (2004).
13. Gigant, B. et al. The 4 A X-ray structure of a tubulin:stathmin-like domain complex. *Cell* **102**, 809-16 (2000).
14. Bai, R. et al. Interactions of halichondrin B and eribulin with tubulin. *J Chem Inf Model* **51**, 1393-404 (2011).

Figures.

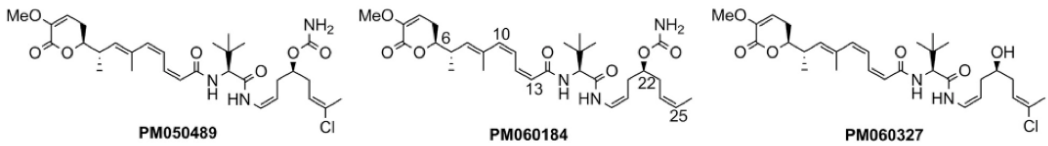


Figure 1- PM050489 is a potent microtubule depolymerizing agent. Chemical structures of PM05489 and the employed analogs.

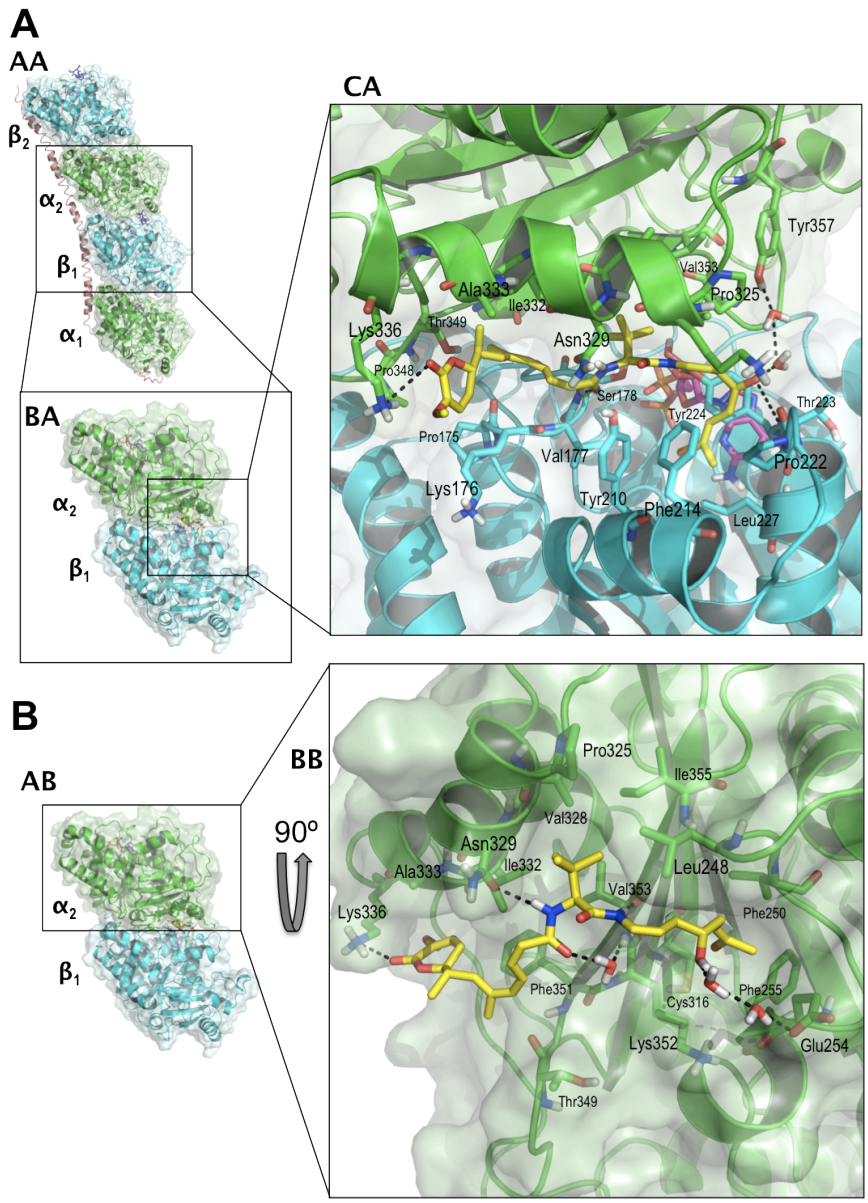


Figure 2.- High resolution molecular model of the two possible interactions of PM060327 with tubulin. A. β_1:α_2-interdimer interface model. **AA.** PyMOL representation of the two αβ-tubulin heterodimers (α-tubulin in green, β-tubulin in cyan, and GTP and GDP nucleotides in magenta) stabilized by colchicine (represented in grey sticks) and bound to phomopsin A (represented as blue sticks) and the stathmin-like domain of protein RB3 (pink helix), as found in PDB entry 3DU7. **AB.** The central segment of the previous structure was extracted and, upon removal of phomopsin A, used to build and simulate the complex with PM060327. A**C.** Detail of the β_1:α_2 interface showing stably bound PM060327 (yellow sticks). Tubulin residues relevant for binding are displayed as sticks and dashed lines represent hydrogen bonds. **B. Single site model. BA.** PyMOL representation of the β_1:α_2-interdimer interface model out of which the α_2-tubulin (in green with the GTP nucleotide represented in magenta sticks) subunit in complex with PM060327 (represented as yellow sticks) was extracted. **BB.** A 90º rotated image of α-tubulin showing stably bound PM600327. Important amino acids for drug binding are displayed in sticks and dashed lines represent hydrogen bonds.

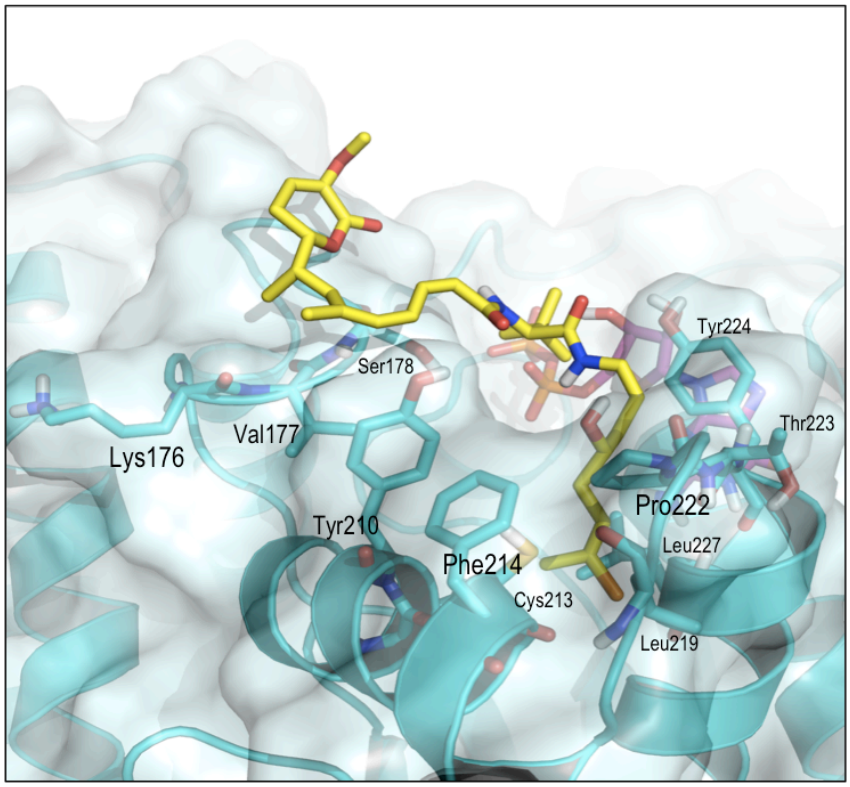


Figure 3. PyMOL representation of PM060327 (yellow sticks) bound to β_1-tubulin (cyan cartoons), with the important amino acids for drug stabilization displayed as sticks. For reference, the GDP molecule is shown as sticks colored in magenta.

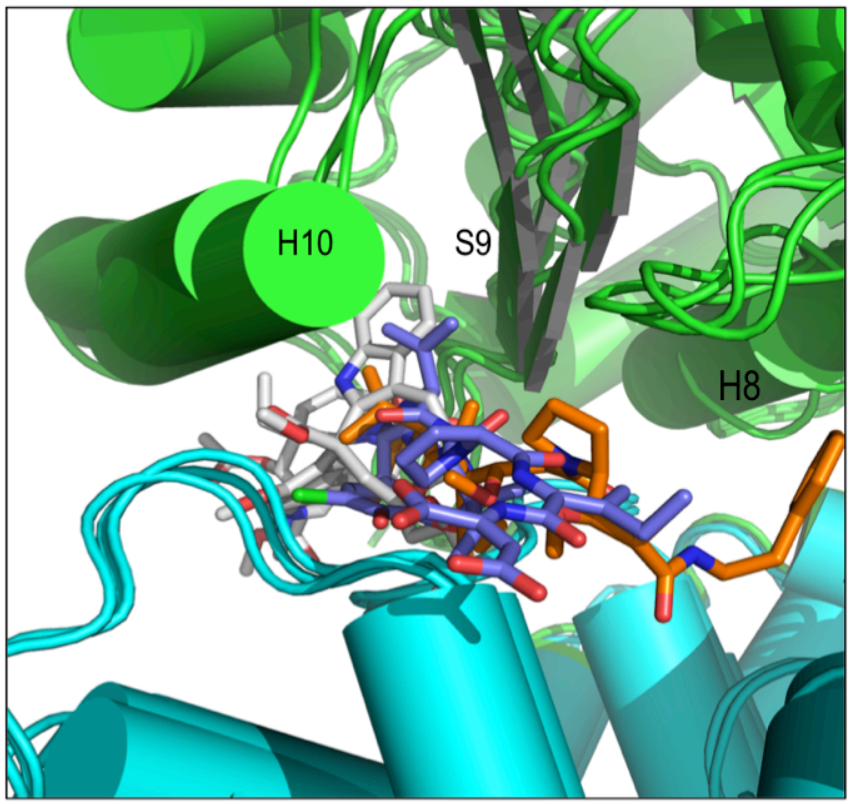


Figure 4. Superimposition of the β_1:α_2-interdimer interfaces of the X-ray crystal structures (1Z2B, 3DU7 and 3E22) containing VLB (grey sticks), phomopsin A (blue sticks) and soblidotin (orange sticks), respectively. The β_1- and α_2-subunits are represented as cyan and green cartoons, and the S9 β-sheet and H8 and H10 α-helices have been labeled.

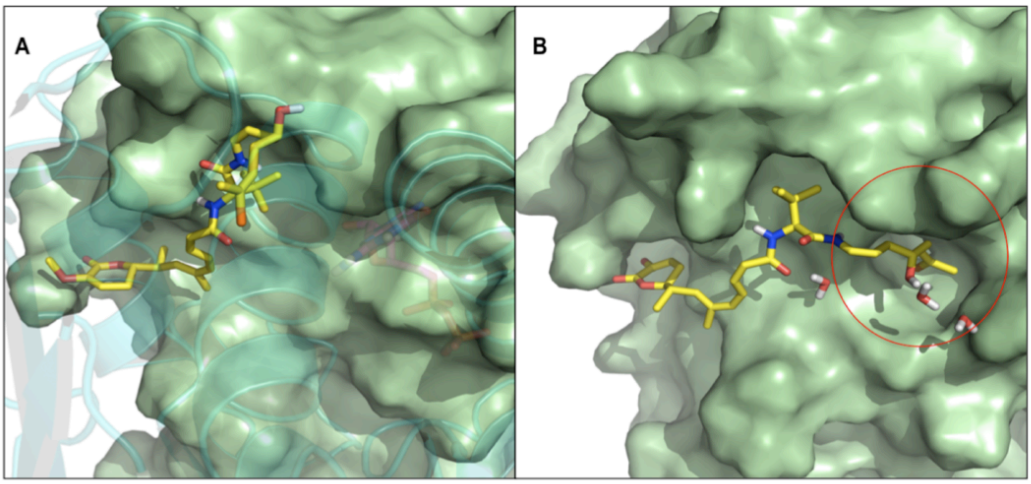


Figure 5. A. PM060327 (yellow sticks) bound at the β_1:α_2-interdimer interface model as seen from the β subunit (semitransparent cartoon in cyan); for reference, the bound GDP nucleotide is also shown as magenta sticks. **B.** PM060327 (yellow sticks) docked with the same orientation in the single site model in α-tubulin. The red circle indicates the location of the cryptic pocket that opens up during the simulation of the α-tubulin-PM060327 complex.

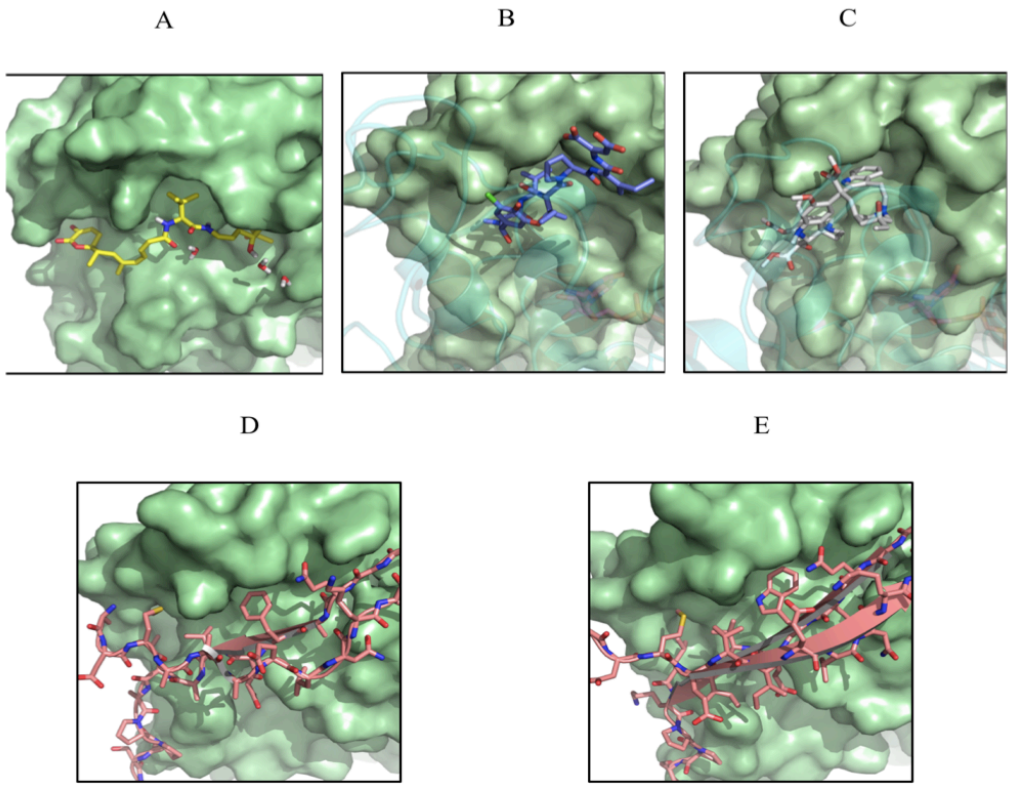


Figure 6. Molecular surface representation (green) of the hydrophobic cleft in α-tubulin in the complexes with A) PM060327 (yellow sticks) in our single site model; B) phomopsin A (blue sticks) in the 4.10 Å resolution structure in PDB code 3DU7; C) VLB (grey sticks) in the 4.10 Å resolution structure in PDB code 1Z2B; D) RB3-SLD (pink cartoons and sticks) protein in the 3.65 Å resolution structure in PDB code 3HKB; and E) RB3-SLD protein (pink cartoons and sticks) in the 2.10 Å resolution structure in PDB code 3RYC.

“L’expérience est une lanterne que l’on porte sur le dos
et qui n’éclaire jamais que le chemin parcouru”
Confucius

5

CONCLUSIONS

5.1 *Optimization of Taxane Binding to Microtubules: Binding Affinity Dissection and Incremental Construction of a High-Affinity Analog of Paclitaxel.*

1. Mutual adaptation of both ligand and protein during refinement of the crystallographic structure supports a different rotameric state for the imidazole side chain of His229 in the PXL-β-tubulin complex relative to the reported X-ray crystal structure.
2. The PXL-binding site in β-tubulin is divided into two sub-pockets by the side chain of His229 that is exposed to bulk solvent and for which a double protonation state with a formal positive charge is proposed.
3. The proposed bound conformation for PXL corresponds to that described as the T-taxol geometry.
4. Common hydrogen bonds that stabilize the binding mode of PXL, DXL and CTX40 are established between (i) the oxetane oxygen of the baccatine ring and the backbone NH of Thr276, and (ii) the carbonyl oxygen of the C13 substituent and Nε in the imidazole side chain of His229.
5. The main van der Waals interactions that stabilize taxane binding are due to the burial of the C2 benzoyl moiety within the hydrophobic pocket next

to His229 and of the common phenyl ring in C13 inside the hydrophobic pocket located at the base of the binding site.

5.2 *COMBINE-guided design of new C2- and C3'-substituted taxanes: tubulin binding affinities and quantitative structure-activity relationships.*

1. A robust quantitative structure-activity relationship has been established for a group of semisynthetic taxanes for which the thermodynamic contribution to stabilized microtubules had already been measured.
2. A COMBINE chemometric model able to correlate experimental ΔG values with a series of weighted interaction energy terms has been built.
3. The improved predictive ability of the COMBINE model built from complexes that incorporate a doubly protonated imidazole side chain at His229 support the proposed protonation state for this crucial tubulin residue.
4. In order to validate the model externally new taxanes with modified moieties at the C2 and C3' substituents were designed, synthesized and tested by our collaborators in this project.
5. The outlying behaviour of the three taxanes that largely deviated from the regression line could be accounted for by an exceedingly unfavourable entropic component that is not currently incorporated into the COMBINE methodology.
6. Incorporation of the new modified taxanes into the training series updated the COMBINE model and provided more accurate guidelines for further modifications.

5.3 *Comparative Binding Energy (COMBINE) analysis supports a proposal for the binding mode of epothilones to β-tubulin.*

1. Molecular dynamics simulations in aqueous solvent were used to sample the conformational space of free epothilone A (EPA) and captured, albeit in a very low percentage, the β-tubulin-bound conformation as determined by NMR-determined.
2. The proposed binding mode for EPA to β-tubulin obtained by an automated ligand-receptor docking program remained stable during the molecular dynamics simulations and proved to be totally different from that reported in the crystallographic structure deposited in the PDB under the code 1TVK.
3. The bound conformation obtained in our model is in consonance with a previously reported NMR-derived conformation although the proposed orientation within the binding site is different.
4. The common pharmacophore for taxanes and epothilones appears to be limited to a good general overlap of both types of compounds in the commonly filled half-site and the presence of two similar anchoring moieties, namely the C3-OH/oxetane and the C15-thiazole/C2-benzoyl pairs in epothilones and taxanes, respectively.
5. The proposed binding mode accounts for the importance of the hydroxyl group at position 3 of the epothilone macrocycle and explains the effect that some tubulin mutations have on the binding affinity of this family of antitumor drugs.
6. Extension of the proposed binding mode for EPA to a series of synthetic analogues and use of the COMBINE method using the whole set of complexes allowed us to derive a quantitative QSAR model that correlated the experimentally measured binding enthalpies with a selection of weighted interaction energy terms.

7. The derived COMBINE model was able to predict reasonably well the binding enthalpies of three synthetic epothilones that were not part of the training set.
8. The fact that binding enthalpies but not binding free energies could be accurately predicted highlights the need to improve the description of entropy and its incorporation into the chemometric model.

5.4 *Tubulin-based Structure-affinity Relationships for Antimitotic Vinca Alkaloids.*

1. Molecular dynamics (MD) simulations of vinblastine (VLB) in water revealed a conformation identical to that reported for the tubulin-bound VLB during almost the whole trajectory.
2. The refined binding mode for tubulin-bound VLB after MD simulations of the complex shows improved mutual adaptation and optimization of the specific interactions of the ligand with the amino acids that line the binding pocket.
3. The other three Vinca alkaloids studied bind to the $\alpha_2\beta_1$ interdimer interface in a very similar way and therefore share a common pharmacophore.
4. The small structural differences among the four alkaloids studied translate into very subtle variations in the binding mode that, once studied by per-residue energy decomposition, could be qualitatively related to differences in experimentally measured binding energies.
5. Given that the amino acid composition of the Vinca-binding domain is the same in all tubulin isotypes, the reported different affinities of the Vinca alkaloids for these isotypes is possibly due to distinct MT kinetics depending on tubulin isotype content.

5.5 *Molecular modeling of the interaction of PM060327 with tubulin.*

1. PM060327, the more water-soluble analogue of PM050489, is proposed to bind at the VLB-binding site at the β_1:α_2-interdimer interface in an extended conformation.
2. The proposed binding mode for PM060327 shares some features with that of VLB as it involves similar electrostatic interactions with $Asn_{\alpha 2}329$ and $Pro_{\beta 1}222$ and also van der Waals contacts with the side chains lining the hydrophobic pockets that lodge both the *tert*-butyl group (in α_2-tubulin) and the terminal Cl-butenyl (in β_1-tubulin).
3. The hydrophobic pocket in α_2-tubulin is also occupied by a hydrophobic ligand group in the complexes of phomopsin A and soblidotin with colchicin-tubulin-RB3-SLD.
4. This docking pose at the β_1:α_2-interdimer interface was extended to analogues PM060184 and PM050489 and satisfactorily accounts for the structure-activity relationship.
5. Of the two possible single-site models only that between PM060327 and α_2-tubulin proved to show distinct anchoring points and to be stable over the molecular dynamics trajectory.
6. Upon rearrangement of helix H8 in the PM060327-α_2-tubulin complex, a cryptic hydrophobic pocket was revealed into which the Cl-butenyl end of the molecule was finally lodged.

"I not only use all the brains that I have,

but all that I can borrow"

Woodrow Wilson, 28th president of US (1856-1924)

6

REFERENCIAS

1. Leach, A. R. (2001) *Molecular modelling : principles and applications*, 2nd ed., Prentice Hall, Harlow, England ; New York.
2. Mathews, C. K., Van Holde, K. E., and Ahern, K. G. (2000) *Biochemistry*, 3rd ed., Benjamin Cummings, San Francisco, Calif.
3. Negri, A., Rodriguez-Larrea, D., Marco, E., Jimenez-Ruiz, A., Sanchez-Ruiz, J. M., and Gago, F. (2010) Protein-protein interactions at an enzyme-substrate interface: characterization of transient reaction intermediates throughout a full catalytic cycle of Escherichia coli thioredoxin reductase, *Proteins 78*, 36-51.
4. Marco, E., and Gago, F. (2007) Overcoming the inadequacies or limitations of experimental structures as drug targets by using computational modeling tools and molecular dynamics simulations, *ChemMedChem 2*, 1388-1401.
5. Kawasaki, Y., and Freire, E. (2011) Finding a better path to drug selectivity, *Drug Discov Today 16*, 985-990.
6. Snyder, P. W., Mecinovic, J., Moustakas, D. T., Thomas, S. W., 3rd, Harder, M., Mack, E. T., Lockett, M. R., Heroux, A., Sherman, W., and Whitesides, G. M. (2011) Mechanism of the hydrophobic effect in the biomolecular recognition of arylsulfonamides by carbonic anhydrase, *Proceedings of the National Academy of Sciences USA*.
7. Marshall, G. R. (2012) Limiting assumptions in structure-based design: binding entropy, *J Comput Aided Mol Des 26*, 3-8.
8. Ohtaka, H., and Freire, E. (2005) Adaptive inhibitors of the HIV-1 protease, *Prog Biophys Mol Biol 88*, 193-208.
9. Lumry, R., and Rajender, S. (1970) Enthalpy–entropy compensation phenomena in water solutions of proteins and small molecules: A ubiquitous properly of water, *Biopolymers 9*, 1125-1227.
10. Eftink, M. R., Anusiem, A. C., and Biltonen, R. L. (1983) Enthalpy-entropy compensation and heat capacity changes for protein-ligand interactions: general thermodynamic models and data for the binding of nucleotides to ribonuclease A, *Biochemistry 22*, 3884-3896.

11. Cooper, A., Johnson, C. M., Lakey, J. H., and Nollmann, M. (2001) Heat does not come in different colours: entropy-enthalpy compensation, free energy windows, quantum confinement, pressure perturbation calorimetry, solvation and the multiple causes of heat capacity effects in biomolecular interactions, *Biophys Chem 93*, 215-230.
12. Sharp, K. (2001) Entropy-enthalpy compensation: fact or artifact?, *Protein Sci 10*, 661-667.
13. Starikov, E. B., and Norden, B. (2007) Enthalpy-entropy compensation: a phantom or something useful?, *J Phys Chem B 111*, 14431-14435.
14. Freire, E. (2008) Do enthalpy and entropy distinguish first in class from best in class?, *Drug Discov Today 13*, 869-874.
15. Snyder, P. W., Mecinovic, J., Moustakas, D. T., Thomas, S. W., 3rd, Harder, M., Mack, E. T., Lockett, M. R., Heroux, A., Sherman, W., and Whitesides, G. M. (2011) Mechanism of the hydrophobic effect in the biomolecular recognition of arylsulfonamides by carbonic anhydrase, *Proc Natl Acad Sci U S A 108*, 17889-17894.
16. Nogales, E., Downing, K. H., Amos, L. A., and Lowe, J. (1998) Tubulin and FtsZ form a distinct family of GTPases, *Nature Structural Biology 5*, 451-458.
17. Wade, R. H. (2007) Microtubules: an overview, *Methods Mol Med 137*, 1-16.
18. Aldaz, H., Rice, L. M., Stearns, T., and Agard, D. A. (2005) Insights into microtubule nucleation from the crystal structure of human gamma-tubulin, *Nature 435*, 523-527.
19. Vollmer, W. (2008) Targeting the bacterial Z-ring, *Chem Biol 15*, 93-94.
20. Aylett, C. H., Wang, Q., Michie, K. A., Amos, L. A., and Lowe, J. Filament structure of bacterial tubulin homologue TubZ, *Proc Natl Acad Sci U S A 107*, 19766-19771.
21. Schlieper, D., Oliva, M. A., Andreu, J. M., and Lowe, J. (2005) Structure of bacterial tubulin BtubA/B: evidence for horizontal gene transfer, *Proc Natl Acad Sci U S A 102*, 9170-9175.
22. Krauhs, E., Little, M., Kempf, T., Hoferwarbinek, R., Ade, W., and Ponstingl, H. (1981) Complete Amino-Acid-Sequence of Beta-Tubulin from Porcine Brain, *P Natl Acad Sci-Biol 78*, 4156-4160.
23. Ponstingl, H., Krauhs, E., Little, M., and Kempf, T. (1981) Complete Amino-Acid-Sequence of Alpha-Tubulin from Porcine Brain, *P Natl Acad Sci-Biol 78*, 2757-2761.
24. Luduena, R. F. (1998) Multiple forms of tubulin: Different gene products and covalent modifications, *Int Rev Cytol 178*, 207-275.
25. Ludueña, R. F., and Banerjee, A. (2008) The Isotypes of Tubulin

The Role of Microtubules in Cell Biology, Neurobiology, and Oncology, (Fojo, T., Ed.), pp 123-175, Humana Press.

26. Downing, K. H., and Nogales, E. (1999) Crystallographic structure of tubulin: implications for dynamics and drug binding, *Cell Struct Funct 24*, 269-275.
27. Nogales, E., Downing, K. H., Amos, L. A., and Lowe, J. (1998) Tubulin and FtsZ form a distinct family of GTPases, *Nat Struct Biol 5*, 451-458.
28. Wade, R. H. (2009) On and around microtubules: an overview, *Mol Biotechnol 43*, 177-191.
29. Anders, K. R., and Botstein, D. (2001) Dominant-lethal alpha-tubulin mutants defective in microtubule depolymerization in yeast, *Mol Biol Cell 12*, 3973-3986.
30. Oliva, M. A., Cordell, S. C., and Lowe, J. (2004) Structural insights into FtsZ protofilament formation, *Nat Struct Mol Biol 11*, 1243-1250.
31. Dougherty, C. A., Sage, C. R., Davis, A., and Farrell, K. W. (2001) Mutation in the beta-tubulin signature motif suppresses microtubule GTPase activity and dynamics, and slows mitosis, *Biochemistry 40*, 15725-15732.
32. Lappchen, T., Pinas, V. A., Hartog, A. F., Koomen, G. J., Schaffner-Barbero, C., Andreu, J. M., Trambaiolo, D., Lowe, J., Juhem, A., Popov, A. V., and den Blaauwen, T. (2008) Probing FtsZ and tubulin with C8-substituted GTP analogs reveals differences in their nucleotide binding sites, *Chem Biol 15*, 189-199.
33. Nogales, E., and Wang, H. W. (2006) Structural mechanisms underlying nucleotide-dependent self-assembly of tubulin and its relatives, *Curr Opin Struct Biol 16*, 221-229.
34. Nogales, E., and Wang, H. W. (2006) Structural intermediates in microtubule assembly and disassembly: how and why?, *Current Opinion in Cell Biology 18*, 179-184.
35. Gebremichael, Y., Chu, J. W., and Voth, G. A. (2008) Intrinsic bending and structural rearrangement of tubulin dimer: molecular dynamics simulations and coarse-grained analysis, *Biophys J 95*, 2487-2499.
36. Buey, R. M., Diaz, J. F., and Andreu, J. M. (2006) The nucleotide switch of tubulin and microtubule assembly: a polymerization-driven structural change, *Biochemistry 45*, 5933-5938.
37. Cassimeris, L. (2009) Microtubule assembly: lattice GTP to the rescue, *Curr Biol 19*, R174-176.
38. Krebs, A., Goldie, K. N., and Hoenger, A. (2005) Structural rearrangements in tubulin following microtubule formation, *EMBO Rep 6*, 227-232.
39. Jordan, M. A., and Wilson, L. (2004) Microtubules as a target for anticancer drugs, *Nat Rev Cancer 4*, 253-265.
40. Chretien, D., and Wade, R. H. (1991) New data on the microtubule surface lattice, *Biol Cell 71*, 161-174.
41. Meurer-Grob, P., Kasparian, J., and Wade, R. H. (2001) Microtubule structure at improved resolution, *Biochemistry 40*, 8000-8008.

42. Li, H., DeRosier, D. J., Nicholson, W. V., Nogales, E., and Downing, K. H. (2002) Microtubule structure at 8 A resolution, *Structure 10*, 1317-1328.
43. Buey, R. M., Calvo, E., Barasoain, I., Pineda, O., Edler, M. C., Matesanz, R., Cerezo, G., Vanderwal, C. D., Day, B. W., Sorensen, E. J., Lopez, J. A., Andreu, J. M., Hamel, E., and Diaz, J. F. (2007) Cyclostreptin binds covalently to microtubule pores and lumenal taxoid binding sites, *Nat Chem Biol 3*, 117-125.
44. Magnani, M., Maccari, G., Andreu, J. M., Diaz, J. F., and Botta, M. (2009) Possible binding site for paclitaxel at microtubule pores, *FEBS J 276*, 2701-2712.
45. Barasoain, I., Garcia-Carril, A. M., Matesanz, R., Maccari, G., Trigili, C., Mori, M., Shi, J. Z., Fang, W. S., Andreu, J. M., Botta, M., and Diaz, J. F. (2010) Probing the pore drug binding site of microtubules with fluorescent taxanes: evidence of two binding poses, *Chem Biol 17*, 243-253.
46. Canales, A., Rodriguez-Salarichs, J., Trigili, C., Nieto, L., Coderch, C., Andreu, J. M., Paterson, I., Jimenez-Barbero, J., and Diaz, J. F. (2011) Insights into the interaction of discodermolide and docetaxel with tubulin. Mapping the binding sites of microtubule-stabilizing agents by using an integrated NMR and computational approach, *ACS Chem Biol 6*, 789-799.
47. Freedman, H., Huzil, J. T., Luchko, T., Luduena, R. F., and Tuszynski, J. A. (2009) Identification and Characterization of an Intermediate Taxol Binding Site Within Microtubule Nanopores and a Mechanism for Tubulin Isotype Binding Selectivity, *Journal of Chemical Information and Modeling 49*, 424-436.
48. Nogales, E., and Wang, H. W. (2006) Structural intermediates in microtubule assembly and disassembly: how and why?, *Curr Opin Cell Biol 18*, 179-184.
49. Kerssemakers, J. W., Munteanu, E. L., Laan, L., Noetzel, T. L., Janson, M. E., and Dogterom, M. (2006) Assembly dynamics of microtubules at molecular resolution, *Nature 442*, 709-712.
50. Rice, L. M., Montabana, E. A., and Agard, D. A. (2008) The lattice as allosteric effector: structural studies of alphabeta- and gamma-tubulin clarify the role of GTP in microtubule assembly, *Proc Natl Acad Sci U S A 105*, 5378-5383.
51. Dimitrov, A., Quesnoit, M., Moutel, S., Cantaloube, I., Pous, C., and Perez, F. (2008) Detection of GTP-tubulin conformation in vivo reveals a role for GTP remnants in microtubule rescues, *Science 322*, 1353-1356.
52. Lawler, S. (1998) Microtubule dynamics: if you need a shrink try stathmin/Op18, *Curr Biol 8*, R212-214.
53. Charbaut, E., Curmi, P. A., Ozon, S., Lachkar, S., Redeker, V., and Sobel, A. (2001) Stathmin family proteins display specific molecular and tubulin binding properties, *J Biol Chem 276*, 16146-16154.

54. Larsson, N., Marklund, U., Gradin, H. M., Brattsand, G., and Gullberg, M. (1997) Control of microtubule dynamics by oncoprotein 18: dissection of the regulatory role of multisite phosphorylation during mitosis, *Mol Cell Biol 17*, 5530-5539.
55. Barbier, P., Dorleans, A., Devred, F., Sanz, L., Allegro, D., Alfonso, C., Knossow, M., Peyrot, V., and Andreu, J. M. (2010) Stathmin and interfacial microtubule inhibitors recognize a naturally curved conformation of tubulin dimers, *J Biol Chem 285*, 31672-31681.
56. Andersen, S. S. (2000) Spindle assembly and the art of regulating microtubule dynamics by MAPs and Stathmin/Op18, *Trends Cell Biol 10*, 261-267.
57. Jourdain, L., Curmi, P., Sobel, A., Pantaloni, D., and Carlier, M. F. (1997) Stathmin: a tubulin-sequestering protein which forms a ternary T2S complex with two tubulin molecules, *Biochemistry 36*, 10817-10821.
58. Nogales, E., Wang, H. W., and Niederstrasser, H. (2003) Tubulin rings: which way do they curve?, *Curr Opin Struct Biol 13*, 256-261.
59. Dumontet, C., and Jordan, M. A. (2010) Microtubule-binding agents: a dynamic field of cancer therapeutics, *Nat Rev Drug Discov 9*, 790-803.
60. Morris, P. G., and Fornier, M. N. (2008) Microtubule active agents: beyond the taxane frontier, *Clin Cancer Res 14*, 7167-7172.
61. Molinski, T. F., Dalisay, D. S., Lievens, S. L., and Saludes, J. P. (2009) Drug development from marine natural products, *Nat Rev Drug Discov 8*, 69-85.
62. Fojo, A. T., and Menefee, M. (2005) Microtubule targeting agents: basic mechanisms of multidrug resistance (MDR), *Semin Oncol 32*, S3-8.
63. Panda, D., Miller, H. P., Banerjee, A., Luduena, R. F., and Wilson, L. (1994) Microtubule dynamics in vitro are regulated by the tubulin isotype composition, *Proc Natl Acad Sci U S A 91*, 11358-11362.
64. Rezania, V., Azarenko, O., Jordan, M. A., Bolterauer, H., Luduena, R. F., Huzil, J. T., and Tuszynski, J. A. (2008) Microtubule assembly of isotypically purified tubulin and its mixtures, *Biophys J 95*, 1993-2008.
65. De Donato, M., Mariani, M., Petrella, L., Martinelli, E., Zannoni, G. F., Vellone, V., Ferrandina, G., Shahabi, S., Scambia, G., and Ferlini, C. (2011) Class III beta-tubulin and the cytoskeletal gateway for drug resistance in ovarian cancer, *J Cell Physiol*.
66. Mozzetti, S., Ferlini, C., Concolino, P., Filippetti, F., Raspaglio, G., Prislei, S., Gallo, D., Martinelli, E., Ranelletti, F. O., Ferrandina, G., and Scambia, G. (2005) Class III beta-tubulin overexpression is a prominent mechanism of paclitaxel resistance in ovarian cancer patients, *Clin Cancer Res 11*, 298-305.
67. Ploussard, G., Terry, S., Maille, P., Allory, Y., Sirab, N., Kheuang, L., Soyeux, P., Nicolaiew, N., Coppolani, E., Paule, B., Salomon, L., Culine, S., Buttyan, R., Vacherot, F., and de la Taille, A. (2010) Class III beta-

tubulin expression predicts prostate tumor aggressiveness and patient response to docetaxel-based chemotherapy, *Cancer Res 70*, 9253-9264.
68. Kavallaris, M. (2010) Microtubules and resistance to tubulin-binding agents, *Nat Rev Cancer 10*, 194-204.
69. Nogales, E., Wolf, S. G., and Downing, K. H. (1998) Structure of the alpha beta tubulin dimer by electron crystallography, *Nature 391*, 199-203.
70. Nettles, J. H., Li, H., Cornett, B., Krahn, J. M., Snyder, J. P., and Downing, K. H. (2004) The binding mode of epothilone A on alpha,beta-tubulin by electron crystallography, *Science 305*, 866-869.
71. ter Haar, E., Kowalski, R. J., Hamel, E., Lin, C. M., Longley, R. E., Gunasekera, S. P., Rosenkranz, H. S., and Day, B. W. (1996) Discodermolide, a cytotoxic marine agent that stabilizes microtubules more potently than taxol, *Biochemistry 35*, 243-250.
72. Altmann, K. H. (2001) Microtubule-stabilizing agents: a growing class of important anticancer drugs, *Curr Opin Chem Biol 5*, 424-431.
73. Khrapunovich-Baine, M., Menon, V., Verdier-Pinard, P., Smith, A. B., 3rd, Angeletti, R. H., Fiser, A., Horwitz, S. B., and Xiao, H. (2009) Distinct pose of discodermolide in taxol binding pocket drives a complementary mode of microtubule stabilization, *Biochemistry 48*, 11664-11677.
74. Sato, B., Muramatsu, H., Miyauchi, M., Hori, Y., Takase, S., Hino, M., Hashimoto, S., and Terano, H. (2000) A new antimitotic substance, FR182877. I. Taxonomy, fermentation, isolation, physico-chemical properties and biological activities, *J Antibiot (Tokyo) 53*, 123-130.
75. Sato, B., Nakajima, H., Hori, Y., Hino, M., Hashimoto, S., and Terano, H. (2000) A new antimitotic substance, FR182877. II. The mechanism of action, *J Antibiot (Tokyo) 53*, 204-206.
76. Yoshimura, S., Sato, B., Kinoshita, T., Takase, S., and Terano, H. (2000) A new antimitotic substance, FR182877. III. Structure determination, *J Antibiot (Tokyo) 53*, 615-622.
77. Edler, M. C., Buey, R. M., Gussio, R., Marcus, A. I., Vanderwal, C. D., Sorensen, E. J., Diaz, J. F., Giannakakou, P., and Hamel, E. (2005) Cyclostreptin (FR182877), an antitumor tubulin-polymerizing agent deficient in enhancing tubulin assembly despite its high affinity for the taxoid site, *Biochemistry 44*, 11525-11538.
78. Canales, A., Matesanz, R., Gardner, N. M., Andreu, J. M., Paterson, I., Diaz, J. F., and Jimenez-Barbero, J. (2008) The bound conformation of microtubule-stabilizing agents: NMR insights into the bioactive 3D structure of discodermolide and dictyostatin, *Chemistry 14*, 7557-7569.
79. Diaz, J. F., Barasoain, I., and Andreu, J. M. (2003) Fast kinetics of Taxol binding to microtubules. Effects of solution variables and microtubule-associated proteins, *J Biol Chem 278*, 8407-8419.

80. Prussia, A. J., Yang, Y., Geballe, M. T., and Snyder, J. P. (2010) Cyclostreptin and microtubules: is a low-affinity binding site required?, *Chembiochem 11*, 101-109.
81. Bennett, M. J., Barakat, K., Huzil, J. T., Tuszynski, J., and Schriemer, D. C. (2010) Discovery and characterization of the laulimalide-microtubule binding mode by mass shift perturbation mapping, *Chem Biol 17*, 725-734.
82. Kanakkanthara, A., Wilmes, A., O'Brate, A., Escuin, D., Chan, A., Gjyrezi, A., Crawford, J., Rawson, P., Kivell, B., Northcote, P. T., Hamel, E., Giannakakou, P., and Miller, J. H. (2011) Peloruside- and laulimalide-resistant human ovarian carcinoma cells have betaI-tubulin mutations and altered expression of betaII- and betaIII-tubulin isotypes, *Mol Cancer Ther 10*, 1419-1429.
83. Pineda, O., Farras, J., Maccari, L., Manetti, F., Botta, M., and Vilarrasa, J. (2004) Computational comparison of microtubule-stabilising agents laulimalide and peloruside with taxol and colchicine, *Bioorg Med Chem Lett 14*, 4825-4829.
84. Mooberry, S. L., Tien, G., Hernandez, A. H., Plubrukarn, A., and Davidson, B. S. (1999) Laulimalide and isolaulimalide, new paclitaxel-like microtubule-stabilizing agents, *Cancer Res 59*, 653-660.
85. Gaitanos, T. N., Buey, R. M., Diaz, J. F., Northcote, P. T., Teesdale-Spittle, P., Andreu, J. M., and Miller, J. H. (2004) Peloruside A does not bind to the taxoid site on beta-tubulin and retains its activity in multidrug-resistant cell lines, *Cancer Res 64*, 5063-5067.
86. Wang, Y. F., Shi, Q. W., Dong, M., Kiyota, H., Gu, Y. C., and Cong, B. (2011) Natural taxanes: developments since 1828, *Chem Rev 111*, 7652-7709.
87. Wani, M. C., Taylor, H. L., Wall, M. E., Coggon, P., and McPhail, A. T. (1971) Plant antitumor agents. VI. The isolation and structure of taxol, a novel antileukemic and antitumor agent from Taxus brevifolia, *J Am Chem Soc 93*, 2325-2327.
88. Fuchs, D. A., and Johnson, R. K. (1978) Cytologic evidence that taxol, an antineoplastic agent from Taxus brevifolia, acts as a mitotic spindle poison, *Cancer Treat Rep 62*, 1219-1222.
89. Schiff, P. B., Fant, J., and Horwitz, S. B. (1979) Promotion of microtubule assembly in vitro by taxol, *Nature 277*, 665-667.
90. Horwitz, S. B. (2004) Personal recollections on the early development of taxol, *J Nat Prod 67*, 136-138.
91. Lowe, J., Li, H., Downing, K. H., and Nogales, E. (2001) Refined structure of alpha beta-tubulin at 3.5 A resolution, *J Mol Biol 313*, 1045-1057.
92. Alcaraz, A. A., Mehta, A. K., Johnson, S. A., and Snyder, J. P. (2006) The T-Taxol conformation, *J Med Chem 49*, 2478-2488.
93. Kingston, D. G., Bane, S., and Snyder, J. P. (2005) The taxol pharmacophore and the T-taxol bridging principle, *Cell Cycle 4*, 279-289.

94. Matesanz, R., Barasoain, I., Yang, C. G., Wang, L., Li, X., de Ines, C., Coderch, C., Gago, F., Barbero, J. J., Andreu, J. M., Fang, W. S., and Diaz, J. F. (2008) Optimization of taxane binding to microtubules: binding affinity dissection and incremental construction of a high-affinity analog of paclitaxel, *Chem Biol 15*, 573-585.
95. Calvo, E., Barasoain, I., Matesanz, R., Pera, B., Camafeita, E., Pineda, O., Hamel, E., Vanderwal, C. D., Andreu, J. M., Lopez, J. A., and Diaz, J. F. (2012) Cyclostreptin derivatives specifically target cellular tubulin and further map the Paclitaxel site, *Biochemistry 51*, 329-341.
96. Gerth, K., Bedorf, N., Hofle, G., Irschik, H., and Reichenbach, H. (1996) Epothilons A and B: antifungal and cytotoxic compounds from Sorangium cellulosum (Myxobacteria). Production, physico-chemical and biological properties, *J Antibiot (Tokyo) 49*, 560-563.
97. Höfle, G., Bedorf, N., Steinmetz, H., Schomburg, D., Gerth, K., and Reichenbach, H. (1996) Epothilone A and B—Novel 16-Membered Macrolides with Cytotoxic Activity: Isolation, Crystal Structure, and Conformation in Solution, *Angew. Chem. Int. Ed. Engl. 35*, 1567-1569.
98. Bollag, D. M., McQueney, P. A., Zhu, J., Hensens, O., Koupal, L., Liesch, J., Goetz, M., Lazarides, E., and Woods, C. M. (1995) Epothilones, a new class of microtubule-stabilizing agents with a taxol-like mechanism of action, *Cancer Res 55*, 2325-2333.
99. Moulder, S. L. (2008) Ixabepilone for the treatment of taxane-refractory breast cancer, *Future Oncol 4*, 333-340.
100. Egerton, N. (2008) Ixabepilone (ixempra), a therapeutic option for locally advanced or metastatic breast cancer, *P T 33*, 523-531.
101. Huang, H., Menefee, M., Edgerly, M., Zhuang, S., Kotz, H., Poruchynsky, M., Huff, L. M., Bates, S., and Fojo, T. (2010) A phase II clinical trial of ixabepilone (Ixempra; BMS-247550; NSC 710428), an epothilone B analog, in patients with metastatic renal cell carcinoma, *Clin Cancer Res 16*, 1634-1641.
102. Nicolaou, K. C., Vourloumis, D., Li, T., Pastor, J., Winssinger, N., He, Y., Ninkovic, S., Sarabia, F., Vallberg, H., Roschangar, F., King, N. P., Finlay, M. R. V., Giannakakou, P., Verdier-Pinard, P., and Hamel, E. (1997) Designed Epothilones: Combinatorial Synthesis, Tubulin Assembly Properties, abd Cytotoxic Action against Taxol-Resistant Tumor Cells, *Angew. Chem. Int. Ed. Engl. 36*, 2097-2103.
103. Buey, R. M., Diaz, J. F., Andreu, J. M., O'Brate, A., Giannakakou, P., Nicolaou, K. C., Sasmal, P. K., Ritzen, A., and Namoto, K. (2004) Interaction of epothilone analogs with the paclitaxel binding site: relationship between binding affinity, microtubule stabilization, and cytotoxicity, *Chem Biol 11*, 225-236.
104. Cheung, C. H., Wu, S. Y., Lee, T. R., Chang, C. Y., Wu, J. S., Hsieh, H. P., and Chang, J. Y. (2010) Cancer cells acquire mitotic drug resistance

properties through beta I-tubulin mutations and alterations in the expression of beta-tubulin isotypes, *PLoS One 5*, e12564.
105. Giannakakou, P., Gussio, R., Nogales, E., Downing, K. H., Zaharevitz, D., Bollbuck, B., Poy, G., Sackett, D., Nicolaou, K. C., and Fojo, T. (2000) A common pharmacophore for epothilone and taxanes: molecular basis for drug resistance conferred by tubulin mutations in human cancer cells, *Proc Natl Acad Sci U S A 97*, 2904-2909.
106. Wartmann, M., and Altmann, K. H. (2002) The Biology and Medicinal Chemistry of Epothilones, *Current Medicinal Chemistry -Anti-Cancer Agents 2*, 123-148.
107. Reese, M., Sanchez-Pedregal, V. M., Kubicek, K., Meiler, J., Blommers, M. J., Griesinger, C., and Carlomagno, T. (2007) Structural basis of the activity of the microtubule-stabilizing agent epothilone a studied by NMR spectroscopy in solution, *Angew. Chem. Int. Ed. Engl. 46*, 1864-1868.
108. Erdélyi, M., Pfeiffer, B., Hauenstein, K., Fohrer, J., Gertsch, J., Altmann, K. H., and Carlomagno, T. (2008) Conformational preferences of natural and C3-modified epothilones in aqueous solution, *J Med Chem 51*, 1469-1473.
109. Kumar, A., Heise, H., Blommers, M. J., Krastel, P., Schmitt, E., Petersen, F., Jeganathan, S., Mandelkow, E. M., Carlomagno, T., Griesinger, C., and Baldus, M. (2010) Interaction of epothilone B (patupilone) with microtubules as detected by two-dimensional solid-state NMR spectroscopy, *Angew. Chem. Int. Ed. Engl. 49*, 7504-7507.
110. Forli, S., Manetti, F., Altmann, K. H., and Botta, M. (2010) Evaluation of novel epothilone analogues by means of a common pharmacophore and a QSAR pseudoreceptor model for taxanes and epothilones, *ChemMedChem 5*, 35-40.
111. Hamel, E. (1992) Natural products which interact with tubulin in the vinca domain: maytansine, rhizoxin, phomopsin A, dolastatins 10 and 15 and halichondrin B, *Pharmacol Ther 55*, 31-51.
112. Smith, C. D., and Zhang, X. (1996) Mechanism of action cryptophycin. Interaction with the Vinca alkaloid domain of tubulin, *J Biol Chem 271*, 6192-6198.
113. Bai, R., Durso, N. A., Sackett, D. L., and Hamel, E. (1999) Interactions of the sponge-derived antimitotic tripeptide hemiasterlin with tubulin: comparison with dolastatin 10 and cryptophycin 1, *Biochemistry 38*, 14302-14310.
114. Watts, N. R., Cheng, N., West, W., Steven, A. C., and Sackett, D. L. (2002) The cryptophycin-tubulin ring structure indicates two points of curvature in the tubulin dimer, *Biochemistry 41*, 12662-12669.
115. Lobert, S., Fahy, J., Hill, B. T., Duflos, A., Etievant, C., and Correia, J. J. (2000) Vinca alkaloid-induced tubulin spiral formation correlates with cytotoxicity in the leukemic L1210 cell line, *Biochemistry 39*, 12053-12062.

116. Dixon, W. E., and Malden, W. (1908) Colchicine with special reference to its mode of action and effect on bone-marrow, *J Physiol 37*, 50-76.
117. Borisy, G. G., and Taylor, E. W. (1967) The mechanism of action of colchicine. Binding of colchincine-3H to cellular protein, *J Cell Biol 34*, 525-533.
118. Borisy, G. G., and Taylor, E. W. (1967) The mechanism of action of colchicine. Colchicine binding to sea urchin eggs and the mitotic apparatus, *J Cell Biol 34*, 535-548.
119. Sengupta, S., and Thomas, S. A. (2006) Drug target interaction of tubulin-binding drugs in cancer therapy, *Expert Rev Anticancer Ther 6*, 1433-1447.
120. Ravelli, R. B., Gigant, B., Curmi, P. A., Jourdain, I., Lachkar, S., Sobel, A., and Knossow, M. (2004) Insight into tubulin regulation from a complex with colchicine and a stathmin-like domain, *Nature 428*, 198-202.
121. Dorleans, A., Gigant, B., Ravelli, R. B., Mailliet, P., Mikol, V., and Knossow, M. (2009) Variations in the colchicine-binding domain provide insight into the structural switch of tubulin, *Proc Natl Acad Sci U S A 106*, 13775-13779.
122. Gigant, B., Wang, C., Ravelli, R. B., Roussi, F., Steinmetz, M. O., Curmi, P. A., Sobel, A., and Knossow, M. (2005) Structural basis for the regulation of tubulin by vinblastine, *Nature 435*, 519-522.
123. Bai, R. L., Pettit, G. R., and Hamel, E. (1990) Binding of dolastatin 10 to tubulin at a distinct site for peptide antimitotic agents near the exchangeable nucleotide and vinca alkaloid sites, *J Biol Chem 265*, 17141-17149.
124. Cormier, A., Marchand, M., Ravelli, R. B., Knossow, M., and Gigant, B. (2008) Structural insight into the inhibition of tubulin by vinca domain peptide ligands, *EMBO Rep 9*, 1101-1106.
125. Uemura, D., Takahashi, K., Yamamoto, T., Katayama, C., Tanaka, J., Okumura, Y., and Hirata, Y. (1985) Norhalichondrin-a - an Antitumor Polyether Macrolide from a Marine Sponge, *Journal of the American Chemical Society 107*, 4796-4798.
126. Fornier, M. N. (2011) Approved agents for metastatic breast cancer, *Semin Oncol 38 Suppl 2*, S3-10.
127. Bai, R., Nguyen, T. L., Burnett, J. C., Atasoylu, O., Munro, M. H., Pettit, G. R., Smith, A. B., 3rd, Gussio, R., and Hamel, E. (2011) Interactions of halichondrin B and eribulin with tubulin, *J Chem Inf Model 51*, 1393-1404.
128. Luduena, R. F., Roach, M. C., Prasad, V., and Pettit, G. R. (1993) Interaction of halichondrin B and homohalichondrin B with bovine brain tubulin, *Biochem Pharmacol 45*, 421-427.
129. Dabydeen, D. A., Burnett, J. C., Bai, R., Verdier-Pinard, P., Hickford, S. J., Pettit, G. R., Blunt, J. W., Munro, M. H., Gussio, R., and Hamel, E.

(2006) Comparison of the activities of the truncated halichondrin B analog NSC 707389 (E7389) with those of the parent compound and a proposed binding site on tubulin, *Mol Pharmacol 70*, 1866-1875.
130. Jordan, M. A., Kamath, K., Manna, T., Okouneva, T., Miller, H. P., Davis, C., Littlefield, B. A., and Wilson, L. (2005) The primary antimitotic mechanism of action of the synthetic halichondrin E7389 is suppression of microtubule growth, *Mol Cancer Ther 4*, 1086-1095.
131. Qi, Y., and Ma, S. (2011) The medicinal potential of promising marine macrolides with anticancer activity, *ChemMedChem 6*, 399-409.
132. Bai, R., Taylor, G. F., Cichacz, Z. A., Herald, C. L., Kepler, J. A., Pettit, G. R., and Hamel, E. (1995) The spongistatins, potently cytotoxic inhibitors of tubulin polymerization, bind in a distinct region of the vinca domain, *Biochemistry 34*, 9714-9721.
133. Noble, R. L., Beer, C. T., and Cutts, J. H. (1959) Further biological activities of vincaleukoblastine,Äîan alkaloid isolated from Vinca rosea (L.), *Biochemical Pharmacology 1*, 347-348.
134. Seam, P., Janik, J. E., Longo, D. L., and Devita, V. T., Jr. (2009) Role of chemotherapy in Hodgkin's lymphoma, *Cancer J 15*, 150-154.
135. Jordan, M. A., and Wilson, L. (1990) Kinetic analysis of tubulin exchange at microtubule ends at low vinblastine concentrations, *Biochemistry 29*, 2730-2739.
136. Jordan, M. A., Thrower, D., and Wilson, L. (1991) Mechanism of inhibition of cell proliferation by Vinca alkaloids, *Cancer Res 51*, 2212-2222.
137. Toso, R. J., Jordan, M. A., Farrell, K. W., Matsumoto, B., and Wilson, L. (1993) Kinetic stabilization of microtubule dynamic instability in vitro by vinblastine, *Biochemistry 32*, 1285-1293.
138. Amos, L. A., Jubb, J. S., Henderson, R., and Vigers, G. (1984) Arrangement of protofilaments in two forms of tubulin crystal induced by vinblastine, *J Mol Biol 178*, 711-729.
139. Nogales, E., Medrano, F. J., Diakun, G. P., Mant, G. R., Towns-Andrews, E., and Bordas, J. (1995) The effect of temperature on the structure of vinblastine-induced polymers of purified tubulin: detection of a reversible conformational change, *J Mol Biol 254*, 416-430.
140. Lobert, S., Frankfurter, A., and Correia, J. J. (1998) Energetics of vinca alkaloid interactions with tubulin isotypes: implications for drug efficacy and toxicity, *Cell Motil Cytoskeleton 39*, 107-121.
141. Lobert, S., Vulevic, B., and Correia, J. J. (1996) Interaction of vinca alkaloids with tubulin: a comparison of vinblastine, vincristine, and vinorelbine, *Biochemistry 35*, 6806-6814.
142. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) A Second Generation Force Field for the

Simulation of Proteins, Nucleic Acids, and Organic Molecules, *Journal of the American Chemical Society 117*, 5179-5197.
143. Singh, U. C., and Kollman, P. A. (1984) An approach to computing electrostatic charges for molecules, *Journal of Computational Chemistry 5*, 129-145.
144. Bayly, C. I., Cieplak, P., Cornell, W., and Kollman, P. A. (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model, *The Journal of Physical Chemistry 97*, 10269-10280.
145. Cornell, W. D., Cieplak, P., Bayly, C. I., and Kollmann, P. A. (1993) Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation, *Journal of the American Chemical Society 115*, 9620-9631.
146. Schiff, D., and Verlet, L. (1967) Ground State of Liquid Helium-4 and Helium-3, *Physical Review 160*, 208-218.
147. Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *Journal of Computational Physics 23*, 327-341.
148. Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An N [center-dot] log(N) method for Ewald sums in large systems, *The Journal of Chemical Physics 98*, 10089-10092.
149. L.V, W. (1971) Isothermal molecular dynamics calculations for liquid salts, *Chemical Physics Letters 10*, 257-261.
150. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath, *The Journal of Chemical Physics 81*, 3684-3690.
151. Monticelli, L., Sorin, E. J., Tieleman, D. P., Pande, V. S., and Colombo, G. (2008) Molecular simulation of multistate peptide dynamics: a comparison between microsecond timescale sampling and multiple shorter trajectories, *J Comput Chem 29*, 1740-1752.
152. Cino, E. A., Wong-Ekkabut, J., Karttunen, M., and Choy, W. Y. (2011) Microsecond molecular dynamics simulations of intrinsically disordered proteins involved in the oxidative stress response, *PLoS One 6*, e27371.
153. Bueren-Calabuig, J. A., Giraudon, C., Galmarini, C. M., Egly, J. M., and Gago, F. (2011) Temperature-induced melting of double-stranded DNA in the absence and presence of covalently bonded antitumour drugs: insight from molecular dynamics simulations, *Nucleic Acids Res 39*, 8248-8257.
154. Onufriev, A., Bashford, D., and Case, D. A. (2000) Modification of the Generalized Born Model Suitable for Macromolecules, *The Journal of Physical Chemistry B 104*, 3712-3720.
155. Sitkoff, D., Sharp, K. A., and Honig, B. (1994) Correlating solvation free energies and surface tensions of hydrocarbon solutes, *Biophys Chem 51*, 397-409.

156. Seabra, G. d. M., Walker, R. C., Elstner, M., Case, D. A., and Roitberg, A. E. (2007) Implementation of the SCC-DFTB Method for Hybrid QM/MM Simulations within the Amber Molecular Dynamics Package,Ä†, *The Journal of Physical Chemistry A 111*, 5655-5664.
157. Senn, H. M., and Thiel, W. (2009) QM/MM methods for biomolecular systems, *Angew Chem Int Ed Engl 48*, 1198-1229.
158. Warshel, A., and Levitt, M. (1976) Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme, *J Mol Biol 103*, 227-249.
159. Goodford, P. J. (1985) A computational procedure for determining energetically favorable binding sites on biologically important macromolecules, *Journal of Medicinal Chemistry 28*, 849-857.
160. Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., and Liang, J. (2006) CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues, *Nucleic Acids Res 34*, W116-118.
161. Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., Shelley, M., Perry, J. K., Shaw, D. E., Francis, P., and Shenkin, P. S. (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, *J Med Chem 47*, 1739-1749.
162. Halgren, T. A., Murphy, R. B., Friesner, R. A., Beard, H. S., Frye, L. L., Pollard, W. T., and Banks, J. L. (2004) Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening, *J Med Chem 47*, 1750-1759.
163. Friesner, R. A., Murphy, R. B., Repasky, M. P., Frye, L. L., Greenwood, J. R., Halgren, T. A., Sanschagrin, P. C., and Mainz, D. T. (2006) Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes, *J Med Chem 49*, 6177-6196.
164. Tirion, M. M. (1996) Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis, *Phys Rev Lett 77*, 1905-1908.
165. Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., and Olson, A. J. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, *Journal of Computational Chemistry 19*, 1639-1662.
166. Ajmani, S., Jadhav, K., and Kulkarni, S. A. (2009) Group-Based QSAR (G-QSAR): Mitigating Interpretation Challenges in QSAR, *QSAR & Combinatorial Science 28*, 36-51.
167. Manoharan, P., Vijayan, R. S., and Ghoshal, N. (2010) Rationalizing fragment based drug discovery for BACE1: insights from FB-QSAR, FB-QSSR, multi objective (MO-QSPR) and MIF studies, *J Comput Aided Mol Des 24*, 843-864.
168. Hansch, C. (1969) Quantitative approach to biochemical structure-activity relationships, *Accounts of Chemical Research 2*, 232-239.

169. Kubinyi, H. (2008) Comparative Molecular Field Analysis (CoMFA), In *Handbook of Chemoinformatics*, pp 1555-1574, Wiley-VCH Verlag GmbH.
170. Ortiz, A. R., Pisabarro, M. T., Gago, F., and Wade, R. C. (1995) Prediction of drug binding affinities by comparative binding energy analysis, *J Med Chem 38*, 2681-2691.
171. Kubinyi, H., Folkers, G., and Martin, Y. C. (1998) *3D QSAR in drug design*, Kluwer Academic, Dordrecht ; Boston, Mass.
172. Gil-Redondo, R., Klett, J., Gago, F., and Morreale, A. (2010) gCOMBINE: A graphical user interface to perform structure-based comparative binding energy (COMBINE) analysis on a set of ligand-receptor complexes, *Proteins 78*, 162-172.
173. Still, W. C., Tempczyk, A., Hawley, R. C., and Hendrickson, T. (1990) Semianalytical treatment of solvation for molecular mechanics and dynamics, *Journal of the American Chemical Society 112*, 6127-6129.
174. Honig, B., and Nicholls, A. (1995) Classical electrostatics in biology and chemistry, *Science 268*, 1144-1149.