

**SIMULACIONES POR ORDENADOR DE  
LAS INTERACCIONES FÁRMACO-ADN:  
UN VIAJE PERSONAL**

**DISCURS**

llegit a l'acte d'ingrés de l'Acadèmic Corresponent

**Il·lustre Sr. Dr. Federico Gago Badenas**

Celebrat el dia 12 de setembre de 2022

**PRESENTACIÓ**

a càrrec de l'Acadèmic Numerari

**Excel·lentíssim Sr. Dr. Oriol Valls Planells**

Barcelona

2022

*L'Acadèmia no es fa solidària de  
les opinions que s'exposen en les publicacions,  
de les quals és responsable l'autor.*



Generalitat de Catalunya  
**Departament de Justícia**

Amb la col·laboració del Departament de  
Justícia de la Generalitat de Catalunya

Dipòsit legal: B-15850-2022  
TIRO Y RETIRO

## PRESENTACIÓ

a càrrec de l'Acadèmic Numerari

**Excel·lentíssim Sr. Dr. Oriol Valls Planells**



**Excel·lentíssim Senyor President,  
Excel·lentíssims i Il·lustres Senyores i Senyors Acadèmics,  
Distingides autoritats acadèmiques i professionals,  
Estimats familiars, amics i companys,  
Senyores i Senyors,**

Voldria en primer lloc agrair a la Junta de Govern d'aquesta Reial Institució el fet d'haver-me escollit per presentar, en aquest solemne acte, a l'Il·lustre professor Dr. Federico Gago Badenas. Ha estat per a mi una gran satisfacció poder dur a terme aquesta presentació.

Federico Gago nació en Madrid el 21 de octubre de 1957, el mayor de cuatro hermanos, en un ambiente familiar ligado al mundo de la farmacia, pues su padre gestionó una farmacia comunitaria durante más de 50 años. Ello le llevó a interesarse, ya desde pequeño, por el fascinante mundo del medicamento.

Su formación primara y secundaria la realizó en el colegio de los Salesianos del Paseo de Extremadura de Madrid donde, en los últimos cursos de bachillerato, tuvo la suerte de recibir las enseñanzas de un profesor extraordinario, Don Inocencio, quien le inculcó la pasión por la química y la biología. Con él empezó, también, su afición por la botánica, que luego comentaré. En el propio colegio tuvo ocasión de recibir una sólida formación en idiomas, especialmente en inglés, que luego fue perfeccionando durante el resto de su vida.

A la hora de escoger sus estudios universitarios, no tuvo duda; el ejemplo de su padre, su pasión por la química y la biología, le dejaban una sola opción: cursar la carrera de Farmacia, opción que fue aproba-

da y apoyada por su madre (su primera maestra), quien le explicó que esa carrera le abriría también las puertas a la investigación.

Así lo hizo, cursó la licenciatura en la Facultad de Farmacia de la Universidad Complutense de Madrid. A partir del tercer curso trabajó, como alumno interno, en el Laboratorio de Técnicas Instrumentales del Departamento de Fisicoquímica, dirigido entonces por Don Manuel Ortega Mata, quien fue su mentor durante numerosos años. Él le dirigió la tesina de licenciatura, que versó sobre la purificación de lectinas de diferentes semillas vegetales con objeto de comprobar su capacidad mitogénica sobre linfocitos T. Para completar este estudio contó con colaboraciones inestimables, como las del Dr. José María Rojo del Centro de Investigaciones Biológicas del Consejo Superior de Investigaciones Científicas (CSIC), la Dra. Pajares de la Clínica Infantil de la Paz y la Dra. Julia Jordá del Hospital Clínico San Carlos de Madrid. El trabajo mereció la calificación de sobresaliente.

Terminada la carrera, y habiendo cumplido los seis primeros meses del servicio militar (tres en el Centro de Instrucción de Reclutas de Alcalá de Henares y los tres siguientes en la Academia de Infantería de Toledo), realizó el segundo ciclo de formación de la I.M.E.C. (Instrucción Militar para la Escala Complemento) en la Sección de Investigación y Control del Instituto Farmacéutico del Ejército de Madrid, como Alférez Eventual de Complemento. Este periodo formativo le familiarizó con la preparación de medicamentos e incluyó una breve estancia en el laboratorio del Dr. Juan Antonio Leal en el Consejo Superior de Investigaciones Científicas (CSIC), a donde le destinaron con la misión de producir dextranos, como sustitutivos plasmáticos, a partir de cultivos de *Leuconostoc mesenteroides* en medios libres de células.

Un hecho importante en la vida de Federico Gago tuvo lugar como consecuencia de la refundación de la Universidad de Alcalá (UAH) el año 1977 (la que con anterioridad había sido la Universidad Complutense, fundada por el Cardenal Cisneros a finales del siglo XV, antes de su traslado a Madrid a mediados del siglo XIX). Allí se incorporó, poco tiempo después de su refundación, el Dr. Manuel Ortega como Decano de la nueva Facultad de Farmacia.

Don Manuel Ortega propició que Federico Gago se incorporara, en octubre de 1980, como Profesor Encargado de Curso, al Departamento de Farmacognosia y Farmacodinamia de dicha Universidad de Alcalá, donde ha continuado la mayor parte de su actividad científica y académica. Allí estableció un contacto muy intenso con el Dr. José Ángel Fuentes. Fue este farmacólogo e investigador del CSIC, posteriormente Catedrático de Farmacología en las Facultades de Farmacia y Medicina de la Universidad Complutense de Madrid, quien le enseñó a Federico Gago los métodos de farmacología experimental y que fueron ampliados en octubre de 1983 gracias a un curso práctico al que asistió, impartido en el Departamento de Farmacognosia y Farmacodinamia de la Facultad de Farmacia de la Universidad de Barcelona. También fue con el Dr. Fuentes con quien Federico Gago realizó sus primeras tareas docentes.

Para llevar a cabo su tesis doctoral contó con la dirección, apoyo incondicional y ejemplo motivador del Dr. Julio Álvarez-Builla, en el Departamento de Química Orgánica y Farmacéutica de la Universidad de Alcalá, y con los del Profesor José Elguero, en el Instituto de Química Médica del CSIC, quienes confiaron en él e impulsaron de forma definitiva su desarrollo como docente e investigador. En abril de 1987 Federico Gago presentó su tesis doctoral sobre el tema “Medidas de la hidrofobicidad por técnicas cromatográficas y su aplicación a QSAR”, que mereció la calificación de “Apto cum laude”, la más alta posible en aquel momento.

Otro hito importante en la vida profesional de Federico Gago fue su estancia postdoctoral en el Physical Chemistry Laboratory de la Universidad de Oxford. Después de diversos intentos fallidos de realizar una estancia becada en un departamento extranjero para perfeccionar sus conocimientos farmacológicos, consiguió ser aceptado por el citado centro, aunque fuera sin ninguna dotación económica. El problema económico lo resolvió el doctor Federico Gago consiguiendo una beca del CSIC español. El citado centro de la Universidad de Oxford estaba dirigido por el Dr. W. Graham Richards, un prestigioso químico teórico, impulsor del diseño computarizado de fármacos. Allí, durante sus casi dos años de estancia, el doctor Gago se introdujo y perfeccionó sus conocimientos de Farmacología Molecular Computacional, la cual ha sido el objeto fundamental de su producción

científica desde entonces.

De vuelta a la Universidad de Alcalá, se incorporó de nuevo a la docencia e investigación, ocupando sucesivamente las plazas de profesor contratado, profesor ayudante, profesor titular interino y profesor Titular funcionario de Universidad (desde el 20 de mayo de 1989). El 15 de diciembre de 2009 obtuvo por concurso-oposición el título de catedrático de farmacología de la Universidad de Alcalá con dedicación a tiempo completo.

En el tiempo disponible resulta casi imposible exponer, ni siquiera de forma resumida, la ingente labor docente e investigadora llevada a cabo por el doctor Federico Gago durante todos estos años.

En cuanto a su labor docente destacaré que, además de las clases impartidas en los grados de Medicina y Farmacia, cuenta con más de sesenta cursos de postgrado, doctorado, máster y perfeccionamiento en varias universidades, principalmente de Madrid y Barcelona, aparte de los quince cursos de verano impartidos en diversos puntos de España.

Su labor científica es, también, muy extensa, y se puede resumir, brevemente en la publicación de 223 artículos de investigación en revistas internacionales de alto índice de impacto, además de ocho reseñas en diversas revistas internacionales y ocho artículos de divulgación científica en revistas españolas. También podemos destacar la presentación de 164 ponencias en congresos, jornadas y seminarios nacionales e internacionales, la publicación de un libro sobre “Bases de la interacción fármaco-receptor. Tratamiento y representación gráfica de los datos”, del cual se han hecho dos ediciones; la contribución a la traducción de un libro titulado “Tratado de Fisiología Médica” y la publicación de 26 capítulos de libro.

Cabe citar, finalmente, la dirección o codirección de 15 tesis doctorales, el haber obtenido 28 proyectos de investigación subvencionados por organismos oficiales y el haber realizado 40 proyectos de investigación industria-universidad. Sus trabajos le han permitido obtener dos patentes de utilidad ya registradas y dos más europeas solicitadas que están en estudio.



Su labor docente e investigadora ha sido reconocida con la obtención de 7 tramos docentes (quinquenios) y 6 tramos de investigación (sexenios), el máximo que puede obtener, actualmente, un profesor universitario en España.

A todo ello sumaremos su experiencia en la organización de Congresos y diversas actividades de I+D, como la IX Reunión de Farmacólogos de la Comunidad de Madrid el año 2000; la V Reunión de Ácidos Nucleicos y Nucleósidos en 2005 o el XIV Congreso de la Sociedad de Biofísica España en 2014. De 2001 a 2006 actuó como Director Asociado del *Screensaver-Lifesaver – NFCR Center for Computational Drug Design* de Oxford; de 2001 a 2003 coordinó la Red Nacional de Ácidos Nucleicos y Nucleósidos; de 2007 a 2011 coordinó, también, la Plataforma bioinformática para el cribado virtual de ligandos dirigidos a diferentes dianas biológicas BIPEDD-CM. De 2001 a 2015 ha sido miembro del Consejo Editorial de *Anti-Cancer Agents in Medicinal Chemistry*, (anteriormente *Current Medicinal Chemistry – Anti-Cancer Agents*); desde 2001 es Editor en Jefe de *Journal of Computer-Aided Molecular Design* (del grupo Springer Nature); en 2002 fue coordinador de la Red Temática de Farmacología y Toxicología (FARMATOXI); en el mismo 2002 actuó como coordinador de la Red Nacional de Ácidos Nucleicos y Nucleósidos; en 2003 fue nombrado miembro del Comité Asesor de *Schironia*, revista científica del Colegio Oficial de Farmacéuticos de Madrid; de 2004 a 2005 actuó como coordinador de redacción de Área Estratégica “I+D+I en moléculas de interés terapéutico o preventivo” de la Agencia Laín Entralgo de la Comunidad Autónoma de Madrid; de 2006 a 2010 fue miembro del Comité de Acceso al *Barcelona Supercomputing Centre* (Mare Nostrum); de 2006 a 2010 fue miembro del Consejo Editorial de *Journal of Medicinal Chemistry*, y un largo etcétera.

Como resultado de su trayectoria científico-profesional, el Doctor Gago ha sido premiado con diversos cargos y distinciones. Destacaremos en 1985 el 4º Accésit al premio de la Academia, concurso científico de la Real Academia de Farmacia (Madrid); en 1987 el Premio ALTER, S.A.; en 1989 el Premio Glaxo y en 1989 el Premio FAES, todos ellos de los concursos científicos de la Real Academia de Farmacia (Madrid); en 2001 el Premio Italgas, de Torino (Italia): *Italgas Prize for Research and Technological Innovation* (*Applied Molecular*

*Sciences*), concedido al Prof. Graham Richards (Oxford University) por proyecto transnacional colaborativo; en 2001 el Premio Descartes a la colaboración transnacional, concedido al Prof. Jan Balzarini (K.U. Leuven), coordinador del grupo europeo de trabajo sobre “*Development of novel drugs against human immunodeficiency virus (HIV)*”; en 2002 el Premio de Investigación 2000-2001 de la Fundación Dr. Antonio Esteve: Mención honorífica por el trabajo publicado en el *Journal of Medicinal Chemistry*, en 2001; y en 2006 el Premio 2005 del Consejo Social de la UAH a la transferencia de conocimiento Universidad-Sociedad.

Si bien lo expuesto anteriormente avala suficientemente los méritos académicos y científicos de Federico Gago, existe un motivo adicional de gran relevancia, que es el papel fundamental que ha ejercido en la implantación y promoción en España de lo que hoy se conoce como Biología Computacional. Su aceptación en nuestra comunidad científica como un área de gran dinamismo y futuro, que atrae el interés de muchos investigadores jóvenes, y con una creciente implantación en planes de estudios universitarios de tercer ciclo, se debe sin duda alguna a la labor desempeñada por el Dr. Federico Gago a lo largo de su dilatada trayectoria.

El extenso currículum científico-profesional de Federico Gago parecería indicar que no ha dispuesto de tiempo para dedicarse a otras aficiones y actividades extra-académicas, pero no es así. Ya he comentado que una de sus grandes pasiones ha sido y es la botánica y la jardinería, Desde su etapa escolar preuniversitaria y, sobre todo durante sus estudios de Farmacia, se ha ido despertando en él una fuerte afición por el mundo de las plantas. Cabe mencionar que, a pesar de presumir de poca memoria, recuerda aún los nombres botánicos de la mayoría de las plantas que estudió durante la carrera y es que, lo que se estudia con pasión, no se olvida.

Tampoco podemos dejar de comentar que Federico Gago es un buen deportista. Dedicar parte de su tiempo libre a jugar al pádel, pero, sin duda su deporte favorito son las artes marciales. Gran aficionado desde muy joven al Taekwondo, ha alcanzado el nivel de cinturón negro y es monitor nacional de esta especialidad deportiva. Siempre que sus otras actividades se lo permitían, durante su juventud se dedicó a dar

clases de artes marciales a niños y jóvenes aficionados.

Toda esta vida tan activa que ha llevado a cabo Federico Gago no habría sido posible sin la cooperación y ayuda de su esposa Oliva, a la que conoció como estudiante en la UAH y que no ha dejado de ser su más eficaz y sólida colaboradora. Oliva es farmacéutica y ejerce en farmacia comunitaria. Le ha dado tres hijos: la mayor, María, es también farmacéutica; los otros dos, Luis y Alberto, se dedican a otras actividades. Federico Gago tiene también un nieto “Fede” de 2 años, hijo de María (¿otro posible farmacéutico?).

Senyor President, vist el més que suficients mèrits del Dr. Federico Gago, li prego que, un cop llegit el seu discurs reglamentari, li imposi la medalla i li entregui el títol que l'acreditin com a nou acadèmic corresponent d'aquesta Reial Acadèmia de Farmàcia de Catalunya.

Moltes gràcies per la seva atenció.



**SIMULACIONES POR ORDENADOR DE  
LAS INTERACCIONES FÁRMACO-ADN:  
UN VIAJE PERSONAL**

Federico Gago Badenas, Área de Farmacología,  
Departamento de Ciencias Biomédicas,  
Universidad de Alcalá, E-28805 Madrid

*«Hay que marchar humildemente detrás de los sabios  
para poder marchar algún día en su compañía»*

Santiago RAMÓN Y CAJAL  
*Charlas de café*  
Madrid: Espasa-Calpe, 1978 (p. 161)



**Excel·lentíssim Senyor President,  
Excel·lentíssims i Il·lustres Senyores i Senyors Acadèmics,  
Distingides autoritats acadèmiques i professionals,  
Estimats familiars, amics i companys,  
Senyores i Senyors,**

Mis primeras palabras son para agradecer a los miembros de la *Reial Acadèmia de Farmàcia de Catalunya* mi designación como académico correspondiente de esta insigne institución. Considero -como hizo en su día Ramón y Cajal- que las Academias son instituciones que nos recuerdan constantemente el camino que nuestros predecesores siguieron en el pasado, dan voz a las personas que intervienen en el presente y, si tenemos el adecuado empeño, ponen a nuestra disposición los fundamentos para seguir afrontando el futuro con el ánimo y formación adecuados.

El tema que he escogido para este discurso de ingreso es la relación -muy personal y extendida en el tiempo- que he mantenido con el ácido desoxirribonucleico (ADN) y particularmente la forma en la que me ha resultado posible -con la ayuda de valiosos colaboradores- caracterizar mediante métodos teóricos su interacción con distintos fármacos antitumorales.

## **1. Introducción**

El ADN, como repositorio de la información genética de la célula, ha sido una de las dianas prototípicas del cáncer durante décadas (1-3). En su nivel de organización más bajo, la estructura tridimensional de esta macromolécula biológica parece engañosamente simple y monótona: dos hebras de polinucleótidos que se enrollan en orientaciones

antiparalelas para dar lugar a una doble hélice dextrógira en la que cada uno de los esqueletos de azúcar-fosfato se encuentra en la parte exterior y las bases nitrogenadas se apilan a lo largo del eje central formando parejas mediante la formación de enlaces de hidrógeno altamente complementarios (4,5). Como consecuencia de esta disposición, la capacidad de un tramo dado de ADN para actuar como diana de reconocimiento molecular parece limitarse, a primera vista, a la discriminación de grupos funcionales a lo largo de los surcos mayor y menor que se encuentran entre los enlaces fosfodiéster de ambas hebras (6). Si bien esto es cierto para muchas proteínas que se unen a secuencias específicas de ADN (7,8), así como para ligandos de bajo peso molecular (incluidas las “lexitropsinas” naturales y sintéticas (9) y muchos fármacos antitumorales de uso clínico (10)), también se acepta desde hace años que existe una microheterogeneidad de forma del ADN que depende de su secuencia y proporciona un potencial adicional para el reconocimiento específico (11,12).

El uso de modelos ayudó a Watson y Crick a sugerir una estructura para la molécula de ADN -sobre la base de la famosa fotografía 51 de Rosalind Franklin y otros datos de difracción de rayos X producidos por Maurice Wilkins- y a insinuar un “posible mecanismo de copia del material genético” (4). En las últimas décadas, la forma en que los investigadores, educadores y estudiantes perciben la estructura molecular ha cambiado drásticamente gracias a los espectaculares avances de la biología estructural (13,14) y a la disponibilidad de programas de visualización molecular de alta calidad y fácil adquisición. Merecen la pena destacarse (por orden cronológico y sin mencionar los estrictamente comerciales) Rasmol (15) -junto con su extremadamente eficaz *plug-in* Chime para los navegadores de internet (16)-, ambos actualmente en desuso y sustituidos por Jmol (17), JSMol, Jmol-Viewer (18) o el potentísimo Mol\* (19), VMD (20), Chimera (21) y PyMol (22). Estas herramientas gráficas permiten una visualización fácil y muy informativa tanto de la forma como de las propiedades de las macromoléculas biológicas y sus ligandos. Como se demostró en el caso del ADN, la mera contemplación de una sola estructura molecular puede ayudar a responder algunas de las preguntas más sencillas sobre la función, pero los problemas más complejos suelen requerir la simulación de un proceso, como el movimiento molecular o la unión/disociación de un ligando. De hecho, el comentario de Sir D. C. Phi-



llips en 1981 (“los modelos de latón del ADN y de una variedad de proteínas dominaban la escena y gran parte del pensamiento” (23)) resume la progresiva comprensión de que las moléculas biológicas no son estructuras rígidas sino sistemas dinámicos cuyos movimientos internos desempeñan importantes papeles funcionales. El objetivo a largo plazo del modelado molecular por ordenador, como enfoque “*in silico*” de la bioquímica y la farmacología (24) o “microscopio computacional” (25), es describir las interacciones biomoleculares en términos de las leyes generales de la química y la física que se han aplicado con tanto éxito a las moléculas pequeñas. Para la mayoría de fines prácticos, las herramientas elegidas para tratar grandes macromoléculas biológicas son los métodos de simulación de dinámica molecular (DM) (26,27), que hacen uso de la mecánica molecular (MM) clásica (28), y el análisis de modos normales (29).

Al generar cientos o miles de configuraciones de baja energía de las macromoléculas biológicas a lo largo de una secuencia temporal, las simulaciones de DM intentan captar con precisión “los meneos y las sacudidas”, como dijo Feynman (30), de los átomos que componen estos importantes componentes de todos los seres vivos. Los métodos de DM se han beneficiado enormemente de la reciente convergencia de avances en la parametrización de los campos de fuerzas (31-33), la mejora del tratamiento de las interacciones electrostáticas de largo alcance (34) y el gran aumento de la potencia de las computadoras, no solo por el cumplimiento de la ley de Moore (“la velocidad de la CPU suele aumentar por un factor de 2 o así cada 18 meses”), sino también por la implementación o adaptación de los códigos de software en unidades de procesamiento gráfico (GPUs) (35). Por ello, resulta ahora factible para muchos grupos de investigación con recursos modestos -como el mío en la UAH- obtener, en un tiempo razonable, trayectorias simuladas de sistemas macromoleculares solvatados durante centenares de nanosegundos. La concatenación de “instantáneas” (fotogramas) generadas sucesivamente cada cierto número de picosegundos, cada una de las cuales contiene un conjunto de coordenadas cartesianas, permite la creación de “películas moleculares” y la realización de diversos análisis de los movimientos moleculares -en detalle atómico- a lo largo del tiempo de simulación. La información derivada de estas simulaciones puede utilizarse para (i) extraer información estructural y energética que suele estar más allá de las posibi-

lidades experimentales actuales (36), (ii) proporcionar explicaciones independientes del comportamiento observado experimentalmente (37), (iii) ayudar en la interpretación de los resultados bioquímicos o farmacológicos (38), y (iv) abrir nuevas vías de investigación al plantear nuevas preguntas relevantes que pueden guiar el diseño de nuevos experimentos (por ej., mutagénesis dirigida) y, en su caso, la generación o identificación de nuevos ligandos.

En el campo del ADN, los tiempos de equilibrado y muestreo suelen oscilar actualmente entre unos pocos nanosegundos (39-41) y decenas (42-44), centenares (45) o miles de nanosegundos (46). Aunque estos tiempos todavía no son lo suficientemente largos (47,48) para monitorizar cambios estructurales a gran escala (en ausencia de restricciones de distancia o angulares) o para lograr la convergencia de las distribuciones de iones (44,49,50), se ha demostrado que los tiempos de relajación para los parámetros helicoidales son del orden de 0,5 ns (43,51). Por lo tanto, las trayectorias que duran unos pocos nanosegundos pueden proporcionar descripciones dinámicas adecuadas del ADN y de los complejos ADN-ligando porque los parámetros estructurales característicos (por ejemplo, el balanceo, la torsión y la inclinación de los pares de bases) se estabilizan mucho más rápidamente que las distribuciones iónicas (46,49,50).

La sensación inicial de que los cálculos teóricos eran una pérdida de tiempo (sobre todo por la imposibilidad material de tratar sistemas tan complejos como las macromoléculas de las células vivas con el nivel de detalle que se creía necesario para añadir algo de importancia a nuestros conocimientos (26,52)) ha dado paso, poco a poco, a una renovada ola de confianza en las metodologías de simulación. Las continuas mejoras en los campos de fuerzas de la MM (33,53,54), la incorporación sistemática de moléculas de disolvente (casi siempre agua) y contraiones en los modelos (55,56), y un tratamiento fiable de las interacciones electrostáticas de largo alcance mediante la aplicación de métodos de *particle mesh Ewald* (PME) (57) a “cajas” de disolvente -tanto paralelepípedos rectangulares o cubos como octaedros truncados- (34,35) están haciendo posible la simulación de la dinámica de sistemas relativamente grandes (incluidos los altamente cargados como los ácidos nucleicos (31,48,58)) durante tiempos considerablemente largos, a veces acercándose a (59) -o superando

(60)- la escala temporal del milisegundo, sin las inestabilidades que afectaban a las simulaciones anteriores. Además, el uso de modelos de solvatación implícita, como el Generalizado de Born (GB) (61,62) -mucho más rápido desde el punto de vista computacional-, permite excluir las moléculas de disolvente explícitas (63) y amplía aún más la duración de algunas de estas simulaciones de DM, al igual que el uso de representaciones de grano grueso de las biomoléculas (64,65).

Estas metodologías de simulación molecular se consideran adecuadas para estudiar los cambios conformacionales que están acoplados a la función (66) y, junto con métodos de la mecánica cuántica (QM, de *quantum mechanics*) (67), el curso de las reacciones en las que se rompen y forman enlaces covalentes, como las que tienen lugar en los sitios activos de las enzimas (68) o entre algunos fármacos reactivos y el ADN, como se verá más adelante. También existen una serie de herramientas (DM “dirigida” o “sesgada”) que pueden “bombar” una cierta cantidad de energía adicional dentro del sistema para acelerar el cruce de barreras locales, permitiendo así el estudio de procesos que normalmente se producirían con demasiada lentitud o baja frecuencia en una ejecución de DM estándar basada únicamente en fluctuaciones térmicas aleatorias y los momentos de inercia (69-71). El repertorio de macromoléculas relevantes para la investigación del cáncer que pueden estudiarse con estos métodos también se ha ampliado enormemente en los últimos 20 años, pasando de un escaso número de dianas “clásicas” (por ejemplo, la dihidrofolato reductasa o la timidilato sintasa y diversos oligonucleótidos de ADN de secuencia definida) a complejos ADN-proteína (72), sofisticadas estructuras de ARN (73) y familias enteras de proteínas (por ejemplo, receptores de factores de crecimiento (74) y enzimas como kinasas (75) y fosfatasa (76)), tanto solas como en complejos con ligandos.

Bien podría decirse (77) que los investigadores tienen hoy en día en sus ordenadores portátiles y tabletas más potencia de cálculo que la que estaba disponible en la mayoría de las grandes computadoras cuando se iniciaron las simulaciones de DM de sistemas macromoleculares a finales de la década de los 70 en el siglo pasado (52). Además, una alternativa realista a las “supercomputadoras” a gran escala consiste en utilizar la computación masivamente distribuida en red (78) para aprovechar la potencia de docenas o cientos de ordena-

dores personales (pecés) relativamente baratos en pos de un objetivo común, como ha sido en la práctica el plegamiento de proteínas (79) o el acoplamiento de ligandos putativos al sitio de unión de una proteína diana (80).

Los espectaculares avances en la biología molecular de las proteínas que reconocen y actúan sobre el ADN y los estudios cristalográficos simultáneos de complejos binarios y ternarios en los que intervienen fármacos, ADN y proteínas asociadas están reduciendo progresivamente la brecha entre nuestra percepción de los aspectos estructurales de los complejos fármaco-ADN y nuestra limitada comprensión de cómo funcionan realmente estos fármacos en las células vivas. A continuación, ofreceré una visión sucinta de algunos de estos avances en relación con mi propia investigación desde que fui estudiante postdoctoral en el laboratorio de W. Graham Richards hasta la actualidad.

Mi introducción a la modelización molecular y a las simulaciones por ordenador se originó en el Laboratorio de Química Física (hoy en día de Química Física y Teórica, *Physical and Theoretical Chemistry Laboratory*) de la Universidad de Oxford, tras completar mi doctorado en la UAH en el campo de los parámetros hidrofóbicos derivados de mediciones de cromatografía líquida. El profesor Richards era muy conocido en España, sobre todo por su libro “Quantum Pharmacology” (81), y tuve el privilegio de escuchar una de sus conferencias cuando ambos coincidimos en Barcelona con motivo de un seminario organizado por Miguel Martín en la Universidad Autónoma de esta ciudad. Por cierto, el profesor Corwin Hansch (82), del *Pomona College* de California, también asistió a esa reunión y proyectó un fascinante vídeo que mostraba detalles atómicos del acoplamiento de un inhibidor en el sitio activo de una enzima, posiblemente la alcohol deshidrogenasa. Tanto me impresionaron la potencia de estos métodos computacionales y la belleza intrínseca de los gráficos moleculares que decidí probar suerte aventurándome como novato en este campo emergente. El artículo que me dio Graham para empezar a trabajar de forma independiente cuando me aceptó en su grupo en abril de 1987 era uno en el que los autores informaban sobre la unión de algunos compuestos naturales y sintéticos a tramos de ADN ricos en adenina y timina de preferencia a regiones ricas en guaninas y citosinas (83) ¿Podrían los métodos teóricos de la MM y la QM arrojar algo de luz sobre esta cuestión?

## 2. Ligandos del surco menor del ADN.

La netropsina y la distamicina A (Figura 1) son ligandos bien caracterizados del surco menor del ADN que se han utilizado ampliamente como modelos para el estudio de las interacciones fármaco-ADN en diversos contextos (83). Ambos antibióticos se encontraban entre el conjunto de pequeñas moléculas cuya capacidad discriminadora frente a diferentes secuencias de ADN se había evaluado mediante un ensayo de desplazamiento de etidio que permitía una estimación indirecta de las constantes de asociación fármaco:ADN (83). En el caso de la netropsina, datos espectroscópicos y termodinámicos adicionales (84) mostraban que las energías libres de unión de netropsina tanto a poli[d(AT)]×poli[d(AT)] como a poli(dA)×poli(dT) eran muy similares y unas 4 kcal×mol<sup>-1</sup> más favorables que la energía libre de unión a poli[d(GC)]×poli[d(GC)].

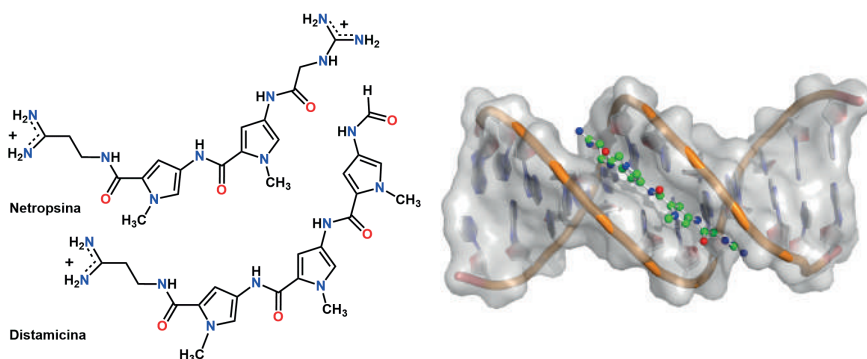


Figura 1. (A) Estructuras químicas de los ligandos del surco menor del ADN netropsina (arriba, dos cargas netas positivas) y distamicina (abajo, una sola carga positiva), ambos de origen natural. (B) Estructura cristalográfica refinada del complejo formado entre la netropsina y el dodecámero de ADN en forma B d(CGCGAATTCGCG)<sub>2</sub> (código 101D en el Protein Data Bank [PDB]).

Éramos conscientes de que un cálculo preciso de las energías libres de unión se ve obstaculizado por las incertidumbres relativas tanto a la conformación particular de las moléculas interactuantes en el estado

no unido como a la validez de los esquemas de partición de energía, así como por las limitaciones con respecto a algunas contribuciones de energía que no son fácilmente susceptibles de cálculo (por ejemplo, los efectos hidrofóbicos y entrópicos). Por esta razón, centramos nuestros primeros esfuerzos en la unión de estos y otros compuestos (incluido el colorante fluorescente Hoechst 33258) a dodecanucleótidos que contenían repeticiones de las secuencias alternantes AT y GC. Utilizando el campo de fuerzas de MM *AMBER* (28) para refinar los modelos y realizando descomposición de la energía de unión calculada, demostramos que las diferencias en las energías de interacción entre los fármacos y las dos secuencias de ADN estudiadas podían correlacionarse satisfactoriamente con la relación de selectividad proporcionada por los experimentos (85). Además, cuando sustituimos los anillos de pirrol de la netropsina y la distamicina A por imidazoles, también pudimos observar una mayor preferencia por los tramos alternos de Gs y Cs debido a los nuevos enlaces de hidrógeno formados entre los nitrógenos imidazólicos de estas “lexitropsinas” y los grupos amino de las guaninas, aunque no se detectó una pérdida significativa de afinidad por los tramos ATAT. Por lo tanto, apoyamos la opinión de que, para disponer de moléculas con el potencial de unirse de forma altamente específica a secuencias predeterminadas, había que incorporar en los ligandos no solo elementos “lectores de G,C” sino también “rechazadores de A,T”. A este respecto, un factor estructural que consideramos relevante para la especificidad de la interacción, en común acuerdo con otros autores, fue la diferente anchura del surco menor admisible o alcanzable en los dos tipos de complejos. El mismo año en que se publicó este trabajo, el descubrimiento imprevisto del primer complejo 2:1 poliamida-ADN abrió el camino a estrategias de diseño que han utilizado anillos de imidazol en lugar de pirrol para “leer” el grupo amino exocíclico de la guanina con gran eficacia (86). Resultados posteriores sugirieron que el lenguaje de los motivos diseñados para el reconocimiento de secuencias de poliamida debería ampliarse para incluir el uso de “palabras” para reconocer dos pares de bases adyacentes, en lugar de “letras” para unirse a pares de bases individuales (87,88). En definitiva, este campo del reconocimiento molecular del ADN desde la perspectiva del surco menor es un bello ejemplo de maduración desde la serendipia hasta el diseño exitoso en la interfaz de la química, la biología y (posiblemente) la medicina humana (89).

El principal problema de nuestras estimaciones de la energía de unión para la netropsina y el resto de los ligandos era que no captaban realmente la esencia del proceso de unión, es decir, la diferencia de energía entre las moléculas del fármaco en disolución, interactuando con el agua y los iones, y las moléculas del fármaco en el complejo fármaco-ADN. De hecho, las simulaciones por ordenador (ahora y entonces) no son capaces de calcular directamente esta diferencia, especialmente cuando hay grandes moléculas en juego. La dificultad radica en que el proceso de unión debe simularse con la suficiente lentitud como para alcanzar el equilibrio termodinámico y esto implica la generación de un enorme número de configuraciones representativas del sistema a lo largo de la coordenada de reacción, lo que resulta impracticable en la mayoría de las situaciones. Sin embargo, el proceso no físico de perturbación de uno de los reactantes en otro (lo que se ha denominado “alquimia computacional” (90-92)), tanto en el estado libre como en el estado ligado, es más fácilmente computable (93-95) y proporciona un cambio relativo de energía libre ( $\Delta G$ ) preciso que puede relacionarse con el físicamente relevante a través de un ciclo termodinámico (Figura 2).

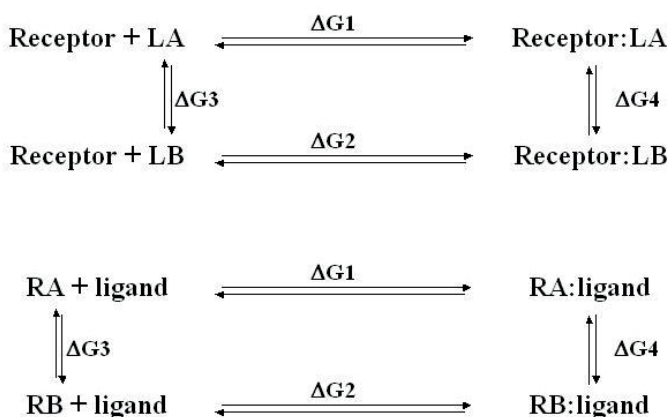


Figura 2. Ejemplos de ciclos termodinámicos típicos utilizados en las simulaciones de perturbación de la energía libre. Arriba: un ligando A se convierte en un ligando B; ambos se unen a un receptor común. Abajo: el mismo ligando se une con diferentes afinidades a los receptores A y B. Las diferencias de energía libre calculadas experimentalmente ( $\Delta\Delta_{\text{Guni3n}} = \Delta G_2 - \Delta G_1$ ) deben ser las mismas que las calculadas a partir de las perturbaciones no físicas de A en B ( $\Delta\Delta_{\text{Guni3n}} = \Delta G_4 - \Delta G_3$ ).

Estimulado por los impresionantes resultados producidos internamente por Graham y su entonces estudiante de postgrado Chris Reynolds -en la actualidad catedrático en el Departamento de Ciencias Biológicas de la Universidad de Essex- en el cálculo de potenciales redox (96), e inspirado por un artículo esclarecedor sobre la termodinámica de la unión de netropsina a varios polinucleótidos de secuencia definida (que también contenía datos para poli[d(IC)]×poly[d(IC)]) (84), me di cuenta de que podíamos intentar aplicar la misma metodología al problema de la unión ADN:ligando. ¿Por qué no “transmutar” las inosinas (Figura 3) en guanosinas (y a la inversa) tanto en el ADN libre como en su complejo con netropsina para evaluar el efecto de crecer/eliminar el grupo amino exocíclico de la guanina sobre las energías de unión calculadas? A pesar de la enormidad de la tarea que teníamos por delante y de la ausencia de precedentes para un trabajo de este tipo, me animé a seguir adelante. Una ventaja con respecto a las simulaciones de la unión de ligandos a proteínas, que generalmente necesitan abordar una reorganización molecular sustancial que puede abarcar largas escalas de tiempo, era que los cambios conformacionales implicados en los cambios alquímicos que planteábamos estudiar aquí eran relativamente menores, ya que no afectaban significativamente a la estructura helicoidal del ADN. Nos las arreglamos para pedir prestado tiempo de computación de todas las máquinas a las podíamos tener acceso, y el trabajo se completó en poco más de un año (97).

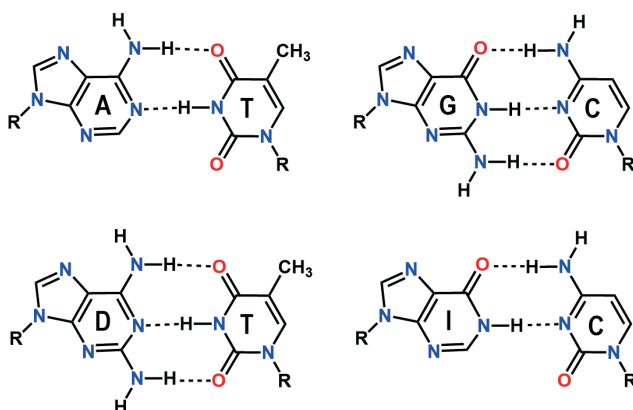


Figura 3. Estructuras químicas de los pares de bases de ADN naturales y modificadas en la disposición Watson-Crick. Los enlaces de hidrógeno se representan como líneas discontinuas. La D representa la 2,6-diaminopurina (o 2-aminoade-nina), mientras que la I representa a la hipoxantina del nucleósido inosina.



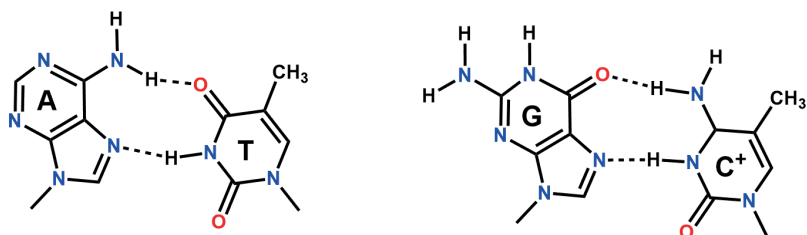
La “transmutación” del hidrógeno unido al C2 de la inosina en un grupo amino en la secuencia ICIC fue posible con la ayuda de dos “átomos ficticios” (D, de *dummy*), que se caracterizaban por tener tanto la carga puntual como el parámetro de energía no enlazada ajustados a cero. Durante las simulaciones del ADN libre y unido a la netropsina, aumentábamos linealmente en pequeños incrementos los parámetros de cada grupo HD2, de modo que al final de la perturbación alcanzaban los valores correspondientes al grupo NH2 estándar. Al hacer esto, nos dimos cuenta de que el surco menor del ADN se volvía progresivamente menos profundo, más polar y menos estrecho. En su complejo con la netropsina, estos cambios fueron acompañados por la extrusión de parte de la molécula del fármaco hacia el disolvente acuoso y por una disminución de la energía libre de unión que estaba en muy buena consonancia con el valor experimental (97). Los cálculos, por lo tanto, fueron capaces de reproducir la preferencia observada de la netropsina por unirse a ICIC sobre GCGC de una manera cuantitativa, mientras que al mismo tiempo proporcionaron una buena justificación de la especificidad de la asociación que señaló, una vez más, a la anchura del surco menor como un determinante importante de la estabilidad del complejo.

### **3. Intercalantes bifuncionales naturales y el emparejamiento de bases Hoogsteen.**

Habiendo estudiado primero los ligandos con tendencia a unirse al ADN rico en AT de preferencia a las secuencias que contienen GC, tras mi regreso a la UAH me interesé por los ligandos intercalantes, que en su mayoría se insertan en los pasos CpG (98). El proceso de intercalación en el ADN permite la inserción de un ligando plano entre pares de bases sucesivos aumentando su separación vertical (elevación, *rise*) y cambiando su grado relativo de rotación (ángulo de giro, *twist*) mediante la introducción de movimientos de biela-manivela en el esqueleto de azúcar-fosfato (como el cigüeñal de un motor) (5). El interés en los intercalantes bifuncionales (o “bisintercalantes”) no solo se debe a la posibilidad de mejorar su afinidad de unión con respecto a la de los monómeros correspondientes, sino también a las mayores oportunidades de imponer una unión selectiva a secuencias definidas que ofrece atrapar dos (¿o posiblemente más?) pares de ba-

ses entre los sitios de intercalación. En consecuencia, el tamaño del sitio ocupado por el ligando aumenta, es decir, para un sitio de unión que cubre cuatro pares de bases, el número de secuencias distinguibles es de 136 frente a solo 10 pasos únicos de dinucleótidos en los que puede tener lugar la monointercalación. El problema concreto de diseño es seleccionar un conector adecuado con la capacidad de reconocer grupos funcionales en el surco menor o en el mayor. Como suele ser el caso, la Naturaleza se anticipó proporcionándonos algunos ejemplos extremadamente elegantes y sugerentes de intercalantes bifuncionales en la familia de antibióticos de quinoxalina representada por la equinomicina (también conocida como quinomicina A) y su precursor biosintético, la triostina A. Estos agentes son producidos principalmente por *Streptomyces echinatus* y *S. triostinicus*, respectivamente, y para ellos se demostró tempranamente un tamaño del sitio de unión de cuatro pares de bases, más tarde confirmado repetidas veces (99). Se demostró que estos bisintercalantes naturales tienen una clara preferencia por unirse a un núcleo 5'-CpG-3' flanqueado por un par de bases A:T a cada lado, a diferencia de los ligandos naturales no intercalantes en forma de medialuna mencionados anteriormente (por ejemplo, la netropsina y la distamicina), que muestran una fuerte preferencia por unirse al surco menor de las regiones de ADN ricas en A,T. De forma bastante llamativa, se puso de manifiesto un reordenamiento dependiente de secuencia en los pares de bases adyacentes al núcleo del CpG en aquellos casos en que la base situada en el lado 5' del paso de unión CpG era una purina: esta base gira 180 grados sobre el enlace glicosídico para adoptar una orientación *syn* en relación con el azúcar, permaneciendo unida mediante enlaces de hidrógeno a la timina opuesta haciendo uso del llamado esquema Hoogsteen (Figura 4), como informó por primera vez un investigador alemán apellidado así cuando resolvió la estructura cristalográfica de rayos X del complejo 1:1 formado entre 1-metilimidina y 9-metiladenina (100). Curiosamente, esta disposición de emparejamiento de bases se detectó en d(ACGT)<sub>2</sub>, d(GCGC)<sub>2</sub> (a pH bajo), d(CGTACG)<sub>2</sub>, d(ACGTACGT)<sub>2</sub>, d(ACGTATACGT)<sub>2</sub> y d(GCGTACGC)<sub>2</sub> (de nuevo a pH ácido), pero no en d(AAACGTTT)<sub>2</sub>, d(TCGA)<sub>2</sub>, d(CCGG)<sub>2</sub> y d(TCGATCGA)<sub>2</sub>.

(A)



(B)

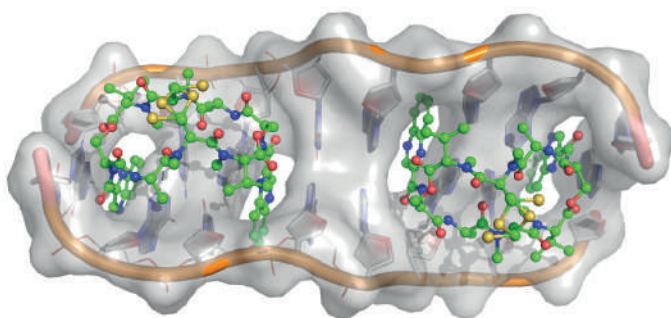


Figura 4. (A) Pares de bases adenina:timina (izquierda) y guanina:citosina (de-  
recha) en la configuración Hoogsteen. (B) Estructura cristalográfica del complejo  
formado entre el octámero de ADN d(GCGTACGC)<sub>2</sub> y dos moléculas de equino-  
micina (código PDB 1PFE). Los cuatro pares de bases G:C y A:T que flanquean  
los cromóforos intercalados están emparejados mediante el esquema Hoogsteen,  
con las purinas giradas a la orientación *syn* en relación con los azúcares.

Movido por la curiosidad de conocer el origen de este fenómeno incompletamente comprendido, propuse a mi primer y muy motivado estudiante de doctorado, José Gallego, que emprendiera una investigación teórica de los factores que podrían estar determinando estas preferencias conformacionales en los pares de bases que flanquean ambos lados del paso de unión CpG. Dado que Peter Kollman y sus colaboradores se habían centrado previamente en la estabilidad relativa del complejo d(CGTAACG)<sub>2</sub>:(trioquina A)<sub>2</sub> con los pares de bases AT centrales en Hoogsteen (como se encontró en una estructura cristalográfica publicada en la revista *Science* en 1984 (101)) o en el emparejamiento Watson-Crick (WC) utilizando MM (102) elegimos estudiar

los complejos formados entre la equinomicina y los tetrámeros de ADN  $d(\text{ACGT})_2$  y  $d(\text{TCGA})_2$ , para los que se disponía de información estructural a partir de la espectroscopía de resonancia magnética nuclear (RMN) (103). Nuestra investigación trató de abordar dos cuestiones principales: (i) por qué las bases terminales del complejo  $d(\text{ACGT})_2$  (en adelante denominado  $A_H$ ) aparecían emparejadas por Hoogsteen mientras que las del  $d(\text{TCGA})_2$  ( $T_W$ ) no, y (ii) por qué el primer complejo era más estable que el segundo. Para ello, analizamos el comportamiento de los complejos  $d(\text{ACGT})_2$ :equinomicina y  $d(\text{TCGA})_2$ :equinomicina en los que ambos pares de bases AT terminales estaban en conformación Hoogsteen ( $A_H$  y  $T_H$ ) o WC ( $A_W$  y  $T_W$ ) mediante cuatro simulaciones independientes de DM en disolución acuosa, cada una de las cuales duró 40 ps (¡un tiempo considerable en aquella época para nuestros modestos recursos computacionales!) (104).

Ninguno de los complejos sufrió grandes cambios conformacionales con respecto a la estructura inicial correspondiente durante las simulaciones en agua y las bases de ADN permanecieron emparejadas, excepto por algunos efectos de deshilachado de los extremos que fueron más notables en el complejo  $A_W$ . Las conformaciones de los complejos encontrados experimentalmente,  $A_H$  y  $T_W$ , coincidían cualitativamente con los efectos nucleares Overhauser (NOE) detectados en los espectros de RMN. Consideramos y discutimos por separado los tres principales tipos de interacción que la equinomicina establece con las moléculas de ADN: (1) interacciones van der Waals y electrostáticas entre la región depsipectídica del fármaco y el surco menor de las moléculas de ADN, (2) enlaces de hidrógeno entre los residuos de alanina de la equinomicina y las bases de guanina del ADN, y (3) interacciones de apilamiento entre los cromóforos de quinoxalina y los pares de bases adyacentes. De acuerdo con los resultados de RMN y rayos X, el modelo  $A_H$  resultó estar estabilizado por dos fuertes enlaces de hidrógeno entre los grupos amino de las alaninas y el N3 de las guaninas, y dos más entre los grupos carbonilo de las alaninas y los grupos 2-amino de las guaninas, uno de los cuales era solo ligeramente más débil debido a la unión asimétrica de la equinomicina; en los otros tres modelos, se perdió uno de los enlaces de hidrógeno entre el residuo de alanina y el fármaco. Aunque estos resultados argumentaban a favor de que la mayor estabilización tenía lugar en el

complejo  $A_H$ , no proporcionaban ninguna pista sobre el origen del reorganización WC > Hoogsteen.

A continuación, centramos nuestra atención en los residuos de quinoxalina que, según nuestros cálculos, aportaban casi el 50% de la energía total de interacción entre la equinomicina y los tetrámeros de ADN. Realizamos un análisis detallado de estas interacciones en los cuatro modelos estudiados porque (i) era generalmente aceptado que la especificidad de secuencia de los sitios flanqueantes de este antibiótico descansaba en gran medida sobre las interacciones de apilamiento de sus cromóforos de quinoxalina (101,105) y (ii) se había demostrado independientemente que las interacciones de apilamiento en la molécula de ADN dependían de la composición y orientación de las bases nitrogenadas y estaban ampliamente dominadas por el término electrostático (106). Consideramos todo el sistema coplanar y conjugado de la quinoxalina-2-carboxamida -en lugar de solo los anillos de quinoxalina- como la entidad molecular del fármaco responsable de las interacciones de apilamiento. A continuación, trazamos el momento dipolar de este fragmento intercalante como la forma más sencilla de representar la polaridad de la distribución de cargas e hicimos lo mismo para cada uno de los pares de bases adyacentes que componen el sitio de intercalación (Figura 4B). Dado que la suma de las cargas debe ser cero para que el momento dipolar sea independiente del origen de coordenadas, se incluyeron en los cálculos los átomos C1' del ADN y los átomos C $\alpha$  de las serinas como amortiguadores para que cada uno de estos sistemas alcanzara la neutralidad eléctrica.

Hicimos tres observaciones interesantes relacionadas con el problema en cuestión: (i) en la conformación Hoogsteen, el momento dipolar de un par de bases AT aumenta a más de 5 debye desde el valor de  $\approx 2$  debye encontrado en la conformación WC, (ii) los cromóforos de quinoxalina-2-carboxamida de la equinomicina tienen un alto momento dipolar (4-5 debye), con el polo negativo situado en la región de la pirazina-2-carboxamida y el positivo en la región del anillo benzénico, y (iii) en el complejo  $A_H$  encontrado experimentalmente, el alto momento dipolar de cada par de bases AT en una conformación Hoogsteen y el del cromóforo apilado de la equinomicina tienen direcciones opuestas, lo que da lugar a una interacción dipolar favorable; por el contrario, el momento dipolar del par de bases AT invierte su orientación en el complejo Hoogsteen  $T_H$  no experimental, de

modo que ambos vectores dipolares apuntan en la misma dirección y la interacción dipolar resultante es desfavorable. En conjunto, nuestra interpretación de estos resultados (104) fue que un esquema de pares de bases A:T en conformación Hoogsteen debería ser preferible para la unión de la equinomicina al tetrámero  $d(\text{ACGT})_2$ , mientras que la disposición clásica de enlace de hidrógeno WC sería más estable para la secuencia TCGA. Fue alentador evidenciar que esto era precisamente lo que se había detectado experimentalmente.

Cuando son los pares GC en lugar de AT los que flanquean el paso CpG al que se une (como una grapa) la equinomicina, el apareamiento de bases Hoogsteen se observa igualmente (a pH ácido) solo cuando la base de purina (guanina) está en el lado 5' de los sitios CpG, es decir, se detecta en  $d(\text{GCGC})_2$  y  $d(\text{GCGTACGC})_2$  (105,107), pero no en  $d(\text{CCGG})_2$  independientemente del pH (108). La necesidad de un pH bajo se justifica porque el esquema de enlace de hidrógeno tipo Hoogsteen en los pares de GC requiere la protonación de la base citosina (Figura 4). En un intento de racionalizar estas observaciones, ampliamos nuestro trabajo anterior modelando y estudiando mediante simulaciones de DM en disolución acuosa los complejos de equinomicina con  $d(\text{GCGC})_2$  y  $d(\text{CCGG})_2$  en los que ambos pares de bases terminales G:C adoptaban una conformación Hoogsteen (denominados  $G_H$  y  $C_H$ , respectivamente) o WC ( $G_W$  y  $C_W$ ). En este caso, descubrimos que las conformaciones de los complejos  $G_H$  y  $C_W$  detectadas experimentalmente no se alteraban de forma significativa durante el transcurso de la simulación de 70 ps, como evaluamos por la baja desviación cuadrática media (*rmsd*, de *root-mean-square deviation*) de la estructura inicial, mientras que observamos valores *rmsd* comparativamente mayores para  $C_H$  y  $G_W$  (109). Los cálculos adicionales también predijeron un gran aumento de la energía de interacción entre la equinomicina y el dúplex  $d(\text{GCGC})_2$  tras la protonación de las citosinas terminales y la posterior formación del par Hoogsteen. En concordancia con trabajos anteriores (102,104), encontramos que alrededor del 50% de la energía de interacción total en todos los complejos parecía ser aportada por las interacciones de van der Waals que involucraban únicamente el sistema bianular de la quinoxalina. Este término, sin embargo, disminuyó cuando las bases de guanina terminales en  $d(\text{CCGG})_2$  adoptaban una conformación Hoogsteen. Por el contrario, en el complejo  $G_H$  observamos una marcada ganancia de

energía de interacción electrostática de apilamiento entre los cromóforos de la equinomicina y las bases terminales emparejadas mediante el esquema Hoogsteen. Pensamos que la combinación de estos dos factores podría explicar la observación experimental de que a bajo pH la formación de pares de bases Hoogsteen solo se observa en el complejo equinomicina-d(GCGC)<sub>2</sub> (110).

En general, nuestros hallazgos confirmaron las primeras pruebas teóricas obtenidas para las bases de ADN apiladas que mostraban que la contribución electrostática juega un papel decisivo en la estabilización de la interacción de apilamiento (111) y que este término es muy sensible a la orientación relativa de los sistemas planos implicados (112). En nuestra opinión, los cambios conformacionales provocados por la unión de equinomicina y triostina A a oligonucleótidos cortos proporcionaban una prueba de concepto excelente para sondear la naturaleza y la dependencia de las interacciones de apilamiento en los complejos ADN:fármacos intercalantes utilizando métodos teóricos. Pero, ¿podría extenderse esta metodología a la cuestión de la selectividad de la unión?

#### **4. La bisintercalación de la equinomicina y agentes bifuncionales relacionados en relación con las preferencias de la secuencia de unión.**

Se aceptaba ampliamente que el origen de la selectividad de unión de la triostina A y la equinomicina por los pasos CpG era el resultado de los enlaces de hidrógeno entre la parte depsipeptídica de estos antibióticos y el grupo amino exocíclico de la guanina en el surco menor (99). Los trabajos tradicionales sobre las relaciones estructura-afinidad (“SAR”, de *Structure-Affinity Relationships*) de la familia de antibióticos de las quinoxalinas se ocupaban de los efectos que la introducción de nuevos sustituyentes o la eliminación de los existentes tenían sobre las propiedades de unión de un determinado fármaco, evaluadas principalmente mediante experimentos de “huella sobre el ADN” (*DNA footprinting*). Los ejemplos más conocidos son probablemente los proporcionados por la des-*N*-tetrametil triostina A (TANDEM) y el [*N*-MeCys<sup>3</sup>,*N*-MeCys<sup>7</sup>] TANDEM (CysMeTANDEM), análogos de la triostina A que carecen de todos o de la mitad

de los grupos *N*-metilo de las cisteínas y valinas, respectivamente, y que se unen mejor a un núcleo central de TpA. Este cambio también podría racionalizarse como debido a la falta de funcionalidades de enlace de hidrógeno convenientemente situadas en el depsi péptido modificado que pudieran interactuar con el grupo 2-amino de la guanina. Posteriormente, un elegante enfoque complementario de SAR consistió en cambiar la naturaleza de las bases que componen el sitio diana en la molécula de ADN (Figura 3) en lugar de modificar el propio ligando. Esto se consiguió en el laboratorio de Michael Waring (*Cambridge University*, Reino Unido) en una serie de elegantes experimentos consistentes en (i) sustituir la guanosina (G) por inosina (I) -lo que equivale a eliminar selectivamente el grupo amino exocíclico del surco menor- y (ii) utilizar la 2,6-diaminopurina (D, también conocida como 2-aminoadenina) en lugar de la adenina (A) para emparejarse con la timina -lo que da lugar a un ADN que contiene un grupo amino en cada uno de los pasos-. Como se esperaba, la equinomicina no se unió a los pasos CpI, lo que otorgaba aún más credibilidad al papel crucial que desempeña este grupo en la interacción de este ligando con el surco menor del ADN (113). La introducción del grupo amino extra en el surco menor de las regiones DT, por el contrario, condujo a una redistribución de los sitios de unión en relación con el ADN normal y se demostró que la equinomicina se unía a cualquier combinación pirimidina-purina que no fuera el paso CpG habitual (114). Sorprendentemente, no solo cambió drásticamente la selectividad sino que la afinidad de la equinomicina por estos nuevos sitios de unión en el ADN modificado resultó ser al menos un orden de magnitud mayor que la que presentaba por el ADN normal.

Dado que nuestros estudios anteriores habían puesto de manifiesto una interacción electrostática desfavorable entre el sistema quinoxalina-2-carboxamida de la equinomicina y los pares de bases que componen el paso central CpG (104,109), razonamos que no solo el enlace de hidrógeno sino también las interacciones de apilamiento podrían estar desempeñando un papel en la determinación de las preferencias de unión de la equinomicina. Para probar esta hipótesis, generamos 13 modelos moleculares de los complejos 1:1 de equinomicina con hexámeros de ADN tanto estándar como modificados y evaluamos las diferentes contribuciones a su estabilidad relativa (110). El conjunto de secuencias contenía todas las combinaciones de sitios de unión



sondeados experimentalmente por la equinomicina. El objetivo era obtener información tanto sobre las interacciones de apilamiento entre el sistema quinoxalina-2-carboxamida y cualquier par de bases de ADN como sobre todas las posibilidades de enlace de hidrógeno entre el depsipéptido y los átomos de ADN en el surco menor.

El análisis energético reveló que las interacciones que implican a los residuos de alanina y a los cromóforos de quinoxalina-2-carboxamida de la equinomicina podrían explicar por sí mismas las selectividades de unión observadas. Además, al representar estos componentes energéticos frente a la energía de interacción total, se detectó una buena correlación lineal y los diferentes hexámeros aparecieron agrupados en dos subconjuntos distintos. Un primer subconjunto abarcaba una familia de buenos sitios de unión que presentaban tanto un dinucleótido central dotado de plena capacidad de formación de enlaces de hidrógeno en el surco menor como una disposición de pares de bases que daba lugar a una interacción de apilamiento globalmente favorable con los cromóforos del antibiótico. Las secuencias del otro subconjunto compartían unas posibilidades de enlace de hidrógeno más pobres y unas interacciones más débiles con los agrupamientos de quinoxalina-2-carboxamida. Los sitios de unión más favorables fueron, de hecho, aquellos en los que las As del ADN habían sido sustituidas por Ds, y sobre todo la secuencia DTDT, en buena concordancia con las observaciones experimentales (110). Fue especialmente ilustrativo ver los resultados de *footprinting* a la luz de los cálculos que descomponían la entalpía de unión debida a las interacciones de apilamiento -tanto con los pares de bases atrapados entre los dos cromóforos como con los que flanqueaban el sitio de bisintercalación- en sus componentes electrostáticos y de van der Waals. Examinando en primer lugar los complejos con un paso CpG central común, pudimos comprobar que las pequeñas diferencias entre ellos se limitaban, como era de esperar, a las interacciones con las bases flanqueantes. No obstante, el componente electrostático parecía ser más discriminante que el término de van der Waals correspondiente, como se ilustra más claramente en el caso de CCGG (aparentemente, un sitio de unión de equinomicina más débil que ACGT) para el que este componente electrostático resultó ser ligeramente repulsivo. El hecho de que se calculara que la interacción electrostática entre los cromóforos del fármaco y el paso central CpG fuera desfavorable

sugería que este efecto negativo debía ser compensado por las interacciones electrostáticas y de enlace de hidrógeno muy favorables establecidas con el surco menor; además, era indicativo de que la diferente distribución de cargas de los pares de bases en un paso TpD (que también presenta dos grupos 2-amino en el surco menor) podría dar lugar a una interacción electrostática de apilamiento atractiva con la quinoxalina-2-carboxamida. Esto resultó ser así, y la mejora de la energía de unión calculada estaba en concordancia con la mayor constante de asociación experimental. Los resultados que racionalizan la mayor afinidad de la equinomicina por la secuencia DTDT respecto a ACGT se vieron reforzados por las energías de interacción obtenidas para secuencias “mixtas”, para las que se encontraron valores a medio camino entre los de cada uno de estos dos complejos. El hecho de que, en el ADN modificado empleado en los experimentos, los sitios CpG originalmente protegidos fueran adyacentes a estos sitios de alta afinidad que contienen D podría explicar por qué se excluyó la unión a los pasos CpG canónicos.

La biosíntesis dirigida ha producido una serie de análogos de la equinomicina y la triostina que contienen cromóforos alternativos a la quinoxalina. Entre ellos destaca el derivado bis-quinolínico de la equinomicina designado 2QN (115), para el que se ha encontrado una interesante correlación entre las propiedades espectroscópicas y termodinámicas (116). Este hecho hizo pensar a los investigadores que los sistemas de anillos intercalados podían controlar en gran medida la selectividad de secuencia. Sin embargo, los resultados de huella sobre el ADN no fueron significativamente diferentes de los obtenidos con la equinomicina. Así, el determinante principal de los sitios de unión preferidos por ambos fármacos fue el grupo 2-amino expuesto en el surco menor de la doble hélice de ADN y ambos fármacos se unieron a las secuencias que rodean los pasos TpD con una afinidad mucho mayor que a los pasos CpG estándar. En línea con nuestros cálculos anteriores, detectamos una mayor estabilización en un complejo 2QN:d(GDTDTC)<sub>2</sub> modelado que contenía un paso TpD central que en el complejo equivalente 2QN:d(GACGTC)<sub>2</sub> que contenía un paso CpG (115). Calculamos y mostramos el potencial electrostático molecular (PEM) para ayudar a visualizar el origen de esta diferencia. Los PEM tanto de 2QN como de equinomicina revelaron un patrón distintivo que consideramos relevante para la unión al ADN, a saber,

una región positiva alrededor de los anillos intercalantes y una región negativa (que emana de los átomos de azufre y de la mayoría de los grupos carbonilo del depsipéptido) en el lado opuesto de la molécula (Figura 5). Postulamos que la asimetría electrostática resultante ayuda a la aproximación productiva y a orientar correctamente este tipo de ligando con respecto a la molécula de ADN antes de la intercalación, compensando posiblemente la falta de una carga positiva neta (como la presente en los ligandos del surco menor discutidos más arriba). Además, la comparación de los PEM calculados para los dos hexanucleótidos de ADN en sus complejos con el fármaco utilizando un método sustractivo reveló diferencias sustanciales no solo en el surco mayor, como era de esperar por la inversión de las posiciones de los grupos O y NH<sub>2</sub> situados en él (Figura 3), sino también en los espacios entre los pares de bases que constituyen el sitio de intercalación y son ocupados por los cromóforos del fármaco durante el proceso de asociación (Figura 5). Sugerimos que estas diferencias podrían modular la especificidad de unión al ADN de la equinomicina, el 2QN y los análogos relacionados, actuando de forma concertada con las bien establecidas interacciones de enlace de hidrógeno que implican a los grupos amino exocíclicos en el surco menor.

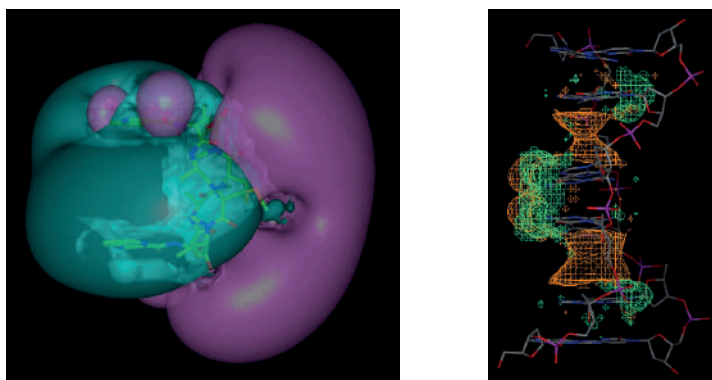


Figura 5. Izquierda: Potencial electrostático molecular (PEM) calculado para la equinomicina y mostrado como contornos sólidos coloreados en rosa y cian para las regiones negativas y positivas, respectivamente. Derecha: Mapa diferencial resultante de sustraer el PEM calculado para d(GDTDTC)<sub>2</sub> en un medio acuoso del obtenido para d(GACGTC)<sub>2</sub> con objeto de resaltar las regiones en las que el PEM difiere entre las dos moléculas de ADN. Los contornos de la malla en naranja y verde delimitan aquellas regiones en las que esta diferencia es  $\geq 0,5$  ó  $\leq -0,5$  unidades, respectivamente.

La especificidad de unión de TANDEM y CysMeTANDEM a los pasos TpA en lugar de a los pasos CpG reconocidos por la triostina A, la equinomicina y el 2QN en el ADN estándar se había racionalizado inicialmente como originada por la pérdida de posibilidades de enlace de hidrógeno con el grupo 2-amino de las guaninas debido a la formación de dos enlaces de hidrógeno intramoleculares entre el CO de las alaninas y el NH de las valinas en los análogos desmetilados. Sin embargo, la sustitución G → I en la molécula de ADN no influye en la interacción de TANDEM con los pasos TpA (es decir, TANDEM no se une a los pasos CpI) (117), y se ha demostrado que la unión de CysMeTANDEM a CpI es mucho más débil que la unión a los sitios TpA de alta afinidad (118), a pesar de que los surcos menores de los pasos CpI y TpA proporcionan las mismas disposiciones de enlace de hidrógeno (Figura 3). Además, se había demostrado que (i) la afinidad de CysMeTANDEM (118,119) y TANDEM (117) por las secuencias que contienen TpA se veía claramente afectada por las bases flanqueantes, y (ii) los tetranucleótidos ATAT y TTAA constituían los mejores y peores sitios de unión, respectivamente.

Dado que estas observaciones también apoyaban la opinión de que la especificidad de unión de este tipo de bisintercalantes depende de algunos factores adicionales más allá de las interacciones de enlace de hidrógeno entre el depsipéptido y el surco menor, intentamos comprender su origen utilizando métodos teóricos. El problema, como he explicado en la introducción, es que las energías de interacción ligando-ADN calculadas mediante MM -ya sea sobre una única estructura que se toma para representar la media del conjunto de cada complejo o sobre múltiples instantáneas extraídas periódicamente a lo largo de una trayectoria de DM con el fin de obtener un promedio- son demasiado crudas para aproximarse a las energías libres de unión que realmente se relacionan con las constantes de asociación determinadas experimentalmente. Aunque este enfoque relativamente simple ya había demostrado sus méritos en la detección de tendencias (110) y podía complementarse con métodos electrostáticos continuos que consideran los efectos de la desolvatación (120,121) de ligandos y receptores -que se oponen a las interacciones electrostáticas favorables (122)-, no hay que olvidar que descuida las contribuciones entrópicas, así como los cambios en la energía interna que sufren tanto el fármaco como la molécula de ADN al unirse. En el caso de los ligandos,

generalmente se supone que estos cambios son pequeños y de magnitud similar para todos los complejos considerados, pero la situación puede ser bastante diferente para las moléculas de ADN debido a la microheterogeneidad dependiente de secuencia mencionada también en la introducción. Por ejemplo, es concebible que la intercalación en ciertas secuencias se vea favorecida simplemente por la tendencia inherente de algunos pasos de dinucleótidos a desenrollarse parcialmente dentro de la estructura de la doble hélice de ADN (11,123).

Para corroborar la validez de nuestra hipótesis de que las interacciones de apilamiento desempeñan un papel decisivo en la modulación de la unión preferente de estos ligandos bisintercalantes, las herramientas computacionales de elección son los métodos de perturbación de la energía libre y de integración termodinámica (91,92,94). Sin embargo, estos estudios no se llevaron a cabo en mi laboratorio hasta que otra estudiante de doctorado de gran talento, Esther Marco, se unió a mi grupo y más tarde fue asistida por una nueva y entusiasta recluta, Ana Negri. Entonces conseguimos, con la inestimable ayuda de mi docto colega y amigo F. Javier Luque (y a quien sin duda debo el honor de haber sido nominado para pertenecer a la RAFC), completar varios ciclos termodinámicos en nuestro intento de comprender los orígenes tanto de la notablemente mayor afinidad de 2QN por DTDT en relación con los sitios GCGC como de la notable pérdida de afinidad de unión entre CysMeTANDEM e ICIC en comparación con ATAT. Estos estudios abarcaron la perturbación de pares de bases completos en el ADN normal y en el que contenía I o D. Además, los cálculos se realizaron de forma que las contribuciones de los pares de bases a ambos lados de los cromóforos intercalados pudieran evaluarse por separado (124). Las diferencias calculadas de energía libre de unión para los dos agentes bisintercalantes 2QN y CysMeTANDEM resultaron estar en muy buen acuerdo cualitativo con los datos experimentales. Así, la unión de 2QN a  $d(\text{GCGGCGCCGC})_2$  se vio desfavorecida en  $1,8 \text{ kcal} \times \text{mol}^{-1}$  con respecto a la unión a  $d(\text{GCGDTCGC})_2$ , de acuerdo con el aumento de  $\sim 100$  veces en la afinidad de 2QN por unirse a un sitio DTDT en comparación con un sitio GCGC en el ADN (115), mientras que la unión de CysMeTANDEM a  $d(\text{CTCICCCAG})_2$  parecía estar aún más desfavorecida (en  $\text{kcal} \times \text{mol}^{-1}$ ) sobre la unión a  $d(\text{CTCATATCAG})_2$ , en consonancia con el hecho de que este ligando se une con muy alta afinidad

a los pasos TpA pero no se une significativamente a los CpI (117). Además, nuestros cálculos apuntaban claramente a la contribución electrostática a las interacciones de apilamiento que implican a las bases distales como el determinante crítico de la selectividad de unión para estos dos ligandos. Este hallazgo reforzó la opinión de que el sitio de reconocimiento para este tipo de bisintercalantes se amplía definitivamente más allá del paso central del dinucleótido para cubrir cuatro pares de bases. Más tarde pudimos extender estas reglas al depsipéptido cíclico rico en azufre tiocoralina, un producto natural de origen marino con dos cromóforos de hidroxiquinolina que muestra una interesante actividad citotóxica a concentraciones nanomolares. Utilizando varios oligonucleótidos de secuencia definida marcados con fluoróforos (“balizas moleculares”) pudimos determinar -con la inestimable colaboración de Alberto Domingo (UAH)- la capacidad de este fármaco para estabilizar la estructura de doble hélice utilizando un instrumento de reacción en cadena de la polimerasa en tiempo real (RT-PCR) (125).

En conclusión, aunque los enlaces de hidrógeno entre el depsipéptido del fármaco bisintercalante y los grupos funcionales presentes en el surco menor del ADN deben ser correctos (porque hay una penalización cuando faltan o son erróneos), los patrones de selectividad solo pueden explicarse adecuadamente cuando también se tienen en cuenta las interacciones de apilamiento que emanan de los pares de bases que componen el sitio de intercalación a ambos lados del cromóforo.

## **5. Preferencias de unión de los bisintercalantes sintéticos de piridocarbazol.**

El ditercalinio y el Flexi-Di son los representantes más estudiados de una clase de bisintercalantes sintéticos que, a diferencia de los de origen natural que acabamos de describir, se unen al ADN desde el surco mayor. Ambos ligandos surgen de conectar dos cromóforos cuaternizados de 10-metoxi-7*H*-pirido[4,3-*c*]carbazolio (126) mediante un espaciador rígido o flexible (Figura 6). Se había demostrado mediante espectroscopía de RMN (127) y cristalografía de rayos X (128) que ambos anillos heteroaromáticos del ditercalinio se intercalaban simultáneamente entre los CpG contiguos de un oligonucleótido

d(CGCG)<sub>2</sub>, mientras que la unión del fármaco a d(GCGC)<sub>2</sub> mostraba un anillo intercalado en el paso CpG y el otro apilado sobre uno de los pares de bases externas (129). Me llamó la atención el hecho de que ninguno de estos conectores por sí mismo parecía capaz de discriminar entre los pasos GpC y CpG, ya que en ambos casos los átomos N7 y O6 de las guaninas se encuentran en la diagonal del surco mayor en posiciones equivalentes, conformando así dos regiones de PEM muy negativo. Además, la demostración en el surco menor de los pasos TpA de cadenas cargadas positivamente semejantes a las del ditercalinio y el Flexi-Di (130) ponía en entredicho el supuesto papel fundamental de los espaciadores a la hora de dictar la preferencia de estos fármacos por el surco mayor. A la luz de nuestros hallazgos anteriores, sospechamos que las interacciones de apilamiento que implican a los cromóforos cargados positivamente podrían tener una relación importante con estas observaciones experimentales y también explicar la evidencia cristalográfica adicional que sugiere una curvatura significativa del eje helicoidal del ADN hacia el surco menor. De hecho, los cristalógrafos sugirieron inicialmente que esta inusual deformación del ADN en complejo con el ditercalinio (~15° de curvatura) podría surgir de la limitada flexibilidad del espaciador que conecta los dos cromóforos del fármaco, pero se sorprendieron al encontrar que el grado de curvatura era incluso mayor en la estructura cristalográfica de la misma secuencia de ADN en complejo con Flexi-Di, el cual contiene un espaciador extremadamente flexible (131).

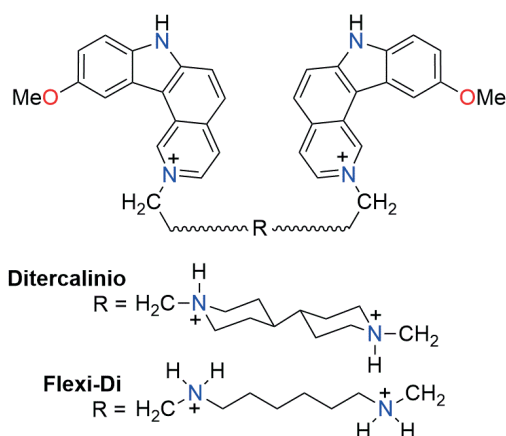


Figura 6. Estructuras químicas de los bisintercalantes sintéticos ditercalinio y Flexi-Di.

Para investigar los orígenes tanto de las preferencias de secuencias de unión como de los cambios conformacionales producidos en la molécula de ADN al formar complejos con estos fármacos, encargué este proyecto a Beatriz de Pascual-Teresa, que acababa de incorporarse a mi grupo tras una exitosa estancia postdoctoral con Ken Houk en la Universidad de California (UCLA). Los cálculos preliminares sobre las estructuras cristalográficas de rayos X mostraron que el componente de energía electrostática de las interacciones de apilamiento entre los anillos de piridocarbazolio y las bases de ADN era favorable para los pares de bases interiores comprendidos entre ellos, pero desfavorable para los pares de bases que componían los límites exteriores del sitio de bisintercalación (132). Además, la interacción electrostática resultó ser atractiva para las guaninas internas pero repulsiva para las citosinas internas, lo que explicaba muy bien los pobres contactos de van der Waals observados experimentalmente entre estas últimas y los cromóforos del fármaco (128,131). También establecimos que, como resultado de la unión del fármaco, el paso constituido por los dos pares de bases centrales se encontraba parcialmente desenrollado en relación con un paso GpC en una hélice de ADN estándar.

A continuación, construimos modelos computacionales de los complejos formados entre ditercalinio y Flexi-Di con el hexanucleótido de secuencia alternante  $d(GCGCGC)_2$ , en los que cada fármaco abrazaba el sitio de unión GpC central. Los complejos se sumergieron en una caja de agua, se neutralizaron mediante la adición de iones sodio, y se sometieron a simulaciones de DM durante 0,5 ns. Para evaluar el efecto de los espaciadores cargados positivamente, se realizaron dos simulaciones adicionales, en las que se “apagaron” las cargas de los enlazadores, es decir, se pusieron a cero. Nuestros resultados reforzaron la idea de que el potencial de formación de enlaces de hidrógeno entre los conectores y el ADN se reducía efectivamente en el medio acuoso debido a las interacciones competitivas con las moléculas de agua. Por esta razón confirmamos que su ubicación en el surco mayor es esencialmente el resultado de las marcadas preferencias orientacionales -dirigidas electrostáticamente- que muestran los cromóforos cuaternizados de piridocarbazolio con respecto a los pares de bases G:C altamente polarizados (133).

En su conjunto, los resultados teóricos de nuestro trabajo sobre los



bisintercalantes proporcionaron un apoyo cuantitativo a las descripciones cualitativas formuladas anteriormente basadas en la visualización de las características electrostáticas de los sistemas apilados (133). Así, la representación gráfica de los PEM calculados mediante métodos de la mecánica cuántica en un plano cercano a la superficie de reconocimiento resulta de gran ayuda para entender la complementariedad electrostática y las interacciones repulsivas. Para un par G:C, la región de PEM más negativo se encuentra en la vecindad de los átomos N7 y O6 de la guanina, mientras que la región de PEM más positivo se encuentra alrededor del anillo de citosina (Figura 7).

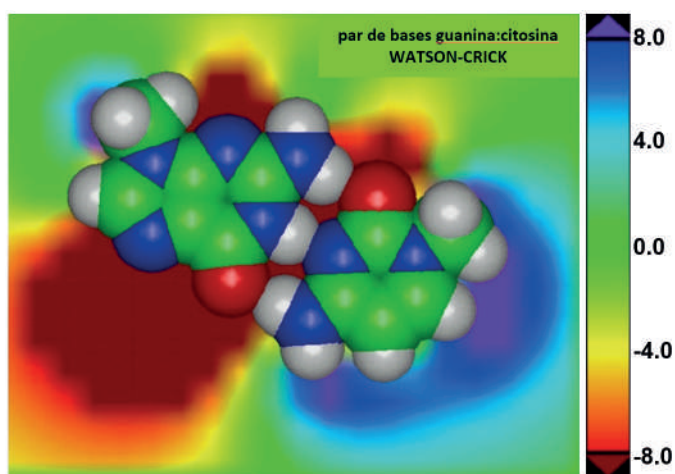


Figura 7. Potencial electrostático molecular (PEM) de un par de bases G:C calculado en un plano situado 1,7 Å por debajo de los centros de los átomos. Este plano se aproxima a la superficie de reconocimiento y el PEM está codificado por colores, de modo que las regiones de 8 kcal×mol<sup>-1</sup> aparecen en azul y las regiones de -8 kcal×mol<sup>-1</sup> están coloreadas en rojo, con los valores intermedios en una rampa suave de colores tipo arcoíris. Las regiones más positivas se colorean en violeta y las más negativas, en marrón. Los átomos se representan como esferas CPK: C, verde; O, rojo; N, azul; H, blanco.

## 6. Selectividad de secuencia de la actinomicina D.

La importancia clínica de la actinomicina D (Figura 8A) como agente antitumoral y su marcada e inusual preferencia por la unión a pasos 5'-GpC-3' (GpC) y otros sitios secundarios como GpG (CpC) y

GpT (ApC) hicieron de este fármaco otro modelo atractivo para nuestros estudios computacionales de las interacciones ligando-ADN. La unión de la actinomicina D al ADN implica la intercalación del cromóforo plano de 2-aminofenoxazina-3-ona entre dos pares de bases y el alojamiento de los dos pentadepsipéptidos cíclicos en el surco menor, uno a cada lado del sitio de intercalación, cubriendo dos pares de bases (Figura 8B).

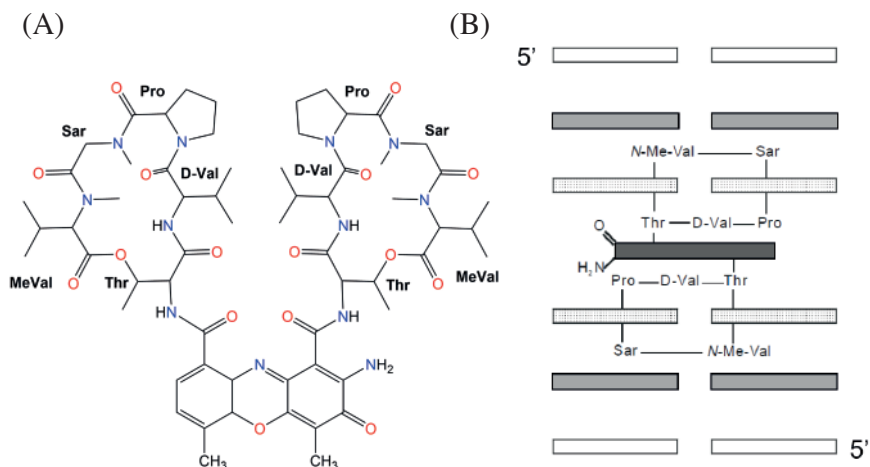


Figura 8. Modo de unión al ADN de la actinomicina D. (A) Estructura química y composición de aminoácidos de la actinomicina D (Sar = sarcosina). (B) Esquema simplificado del modo de unión de la actinomicina D (cromóforo en gris oscuro) al surco menor de un hexanucleótido de ADN visto desde el surco mayor. Los pares de bases inmediatamente adyacentes al sitio de intercalación se denominan centrales y están punteados; los pares de bases contiguos a estos se denominan flanqueantes y están sombreados.

La resolución de varias estructuras cristalográficas (134-137) reforzó la creencia existente de que los enlaces de hidrógeno guanina-treonina eran los factores responsables de la especificidad de unión de la actinomicina a los pasos GpC, una hipótesis que había sido apoyada por estudios teóricos (138). Sin embargo, esta especificidad se mantenía tanto en la actinomicina -un análogo de la actinomicina en el que los pentapéptidos cíclicos se sustituyen por cadenas laterales de *N,N*-dietililenodiamina- como en la 2-aminofenoxazina-3-ona, dos moléculas que no pueden formar enlaces de hidrógeno con los

grupos 2-amino de la guanina. Estas observaciones daban crédito a los primeros informes que apuntaban a las interacciones electrónicas entre el cromóforo de fenoxazona y el par de bases G:C relativamente polarizado como fuente de la especificidad por guanina sobre las otras bases comunes (139). Además, cuando se utilizó 2,6-diaminopurina (D) en lugar de adenina en un fragmento de ADN, los experimentos de *footprinting* mostraron que la especificidad de secuencia de la actinomicina se alteraba drásticamente y el fármaco se unía a cualquier paso 5'-purina-pirimidina-3' diferente del sitio GpC estándar (114,140). Dado que los surcos menores de los pasos GpC y DpT presentan capacidades de enlace de hidrógeno similares, como se ha explicado anteriormente (Figura 3), esta interacción por sí sola no podía explicar adecuadamente el cambio observado en las preferencias de unión.

Para obtener una comprensión más detallada de las interacciones que gobiernan la unión específica de la actinomicina al ADN, construimos modelos moleculares de 14 hexámeros de ADN de doble cadena que contenían una variedad de sitios de unión en la región central, a saber, el paso canónico GpC, sitios secundarios como GpG y GpT, secuencias no preferidas como CpG, ApT, ApA y TpA, y el sitio óptimo DpT. Además, los pasos GpC y DpT se incorporaron en diferentes contextos para evaluar la influencia de los pares de bases flanqueantes (141). El conjunto de complejos contemplaba así todas las posibilidades de formación de enlaces de hidrógeno entre los depsipéptidos y el surco menor del ADN, así como todas las posibles combinaciones de interacciones de apilamiento. Tras el refinamiento energético, calculamos las energías de las interacciones intermoleculares y las descompusimos en diferentes componentes electrostáticos y de van der Waals, sometiendo la matriz de datos resultante a un análisis de componentes principales ("PCA", *Principal Component Analysis*) como medio para destacar las SAR de este importante antibiótico antitumoral.

Comprobamos que las energías de interacción globales reflejaban bien las afinidades relativas de la actinomicina D por los diferentes sitios: el paso canónico GpC se veía energéticamente favorecido sobre cualquier otro paso central en la familia de secuencias naturales, seguido de los sitios secundarios GpG y GpT, aunque las energías de

unión más favorables en todo el conjunto fueron las correspondientes a secuencias en las que las adeninas habían sido sustituidas por D, de acuerdo con la evidencia experimental (140). En cuanto a las secuencias flanqueantes, se prefirieron las pirimidinas a las purinas en el lado 5' del sitio de intercalación GpC, y GGCC apareció como la secuencia menos favorecida. Quizás lo más interesante fue el hallazgo de que los dos primeros componentes principales explicaban fácilmente más del 90% de la varianza contenida en la matriz de datos original y reproducían de forma notable las preferencias de secuencia de la actinomicina D observadas experimentalmente. Los diferentes grupos de sitios se discriminaban fácilmente a lo largo del eje del primer PC, que estaba formado por las interacciones electrostáticas entre los péptidos y los pares de bases centrales, pero también por las interacciones electrostáticas y de van der Waals entre el cromóforo de fenoxazona y los mismos pares de bases. Por lo tanto, demostramos que la actinomicina D discrimina entre las diferentes secuencias mediante una combinación de interacciones de apilamiento y enlaces de hidrógeno. La importancia del término de apilamiento aumenta a la hora de distinguir entre los sitios que contienen G:C -como GpC, GpG y CpG- y es crucial para favorecer los pasos centrales DpT frente a los GpC. Estos resultados indican que los grupos 2-amino no solo alteran el potencial de formación de enlaces de hidrógeno o la geometría del surco menor, como se ha discutido anteriormente, sino que también imparten distintas propiedades de apilamiento a los pares de bases, por motivos tanto estéricos como electrostáticos (106,112). El modelo simple e intuitivo de que las cargas similares se repelen y las cargas opuestas se atraen (ley de Coulomb) también es válido para moléculas y fragmentos moleculares, y se puede considerar que el mínimo de energía encontrado en cada uno de los complejos ADN-ligando intercalado discutidos hasta ahora resulta de un equilibrio entre minimizar las interacciones de apilamiento repulsivas y maximizar las atractivas entre los momentos cuadrupolares eléctricos (142). Invitado por el Profesor Pier de Benedetti (*Università degli Studi di Modena, Italia*), en 1998 contribuí con una revisión sobre las interacciones de apilamiento y la unión de intercalantes al ADN a un número especial que se publicó en *Methods*, un complemento de *Methods in Enzymology* (143).

## **7. Unión del potente agente antitumoral trabectedina al ADN y las inesperadas semejanzas con la unión de la cromomicina A<sub>3</sub>.**

El grupo amino de la guanina volvió a ser el centro de mi atención cuando, a principios de este siglo, me involucré en el estudio de los mecanismos de acción de varios antitumorales de origen marino bajo los auspicios de la empresa farmacéutica y biotecnológica PharmaMar, con sede en Madrid (España) y Cambridge (EE.UU.). El compuesto que más había avanzado hacia la clínica en ese momento era la trabectedina (Figura 9A), una potente ecteinascidina antitumoral -aislada originalmente del tunicado marino *Ecteinascidia turbinata* (144,145)- que obtuvo por primera vez el estatus de medicamento huérfano para el tratamiento de sarcomas de tejidos blandos tanto en Europa como en EE.UU. y que actualmente está aprobada globalmente para uso humano y comercializada con el nombre de Yondelis®. Las ecteinascidinas, que están estructuralmente relacionadas con las safracinas y saframycinas de origen microbiano, sufren un ataque nucleófilo por el grupo 2-amino exocíclico de la guanina en el surco menor del ADN de doble cadena (146) cuando esta base se encuentra en el contexto de un triplete adecuado (147). Así, Lawrence Hurley (*University of Texas, Austin*) había demostrado que la selectividad de secuencia operaba predominantemente a través de un conjunto de reglas de enlace de hidrógeno bien definidas, de manera que los tripletes preferidos como diana son 5'-RGC y 5'-YGG, donde R e Y representan la purina y la pirimidina, respectivamente, y la base subrayada es la guanina a la que el fármaco está unido covalentemente.

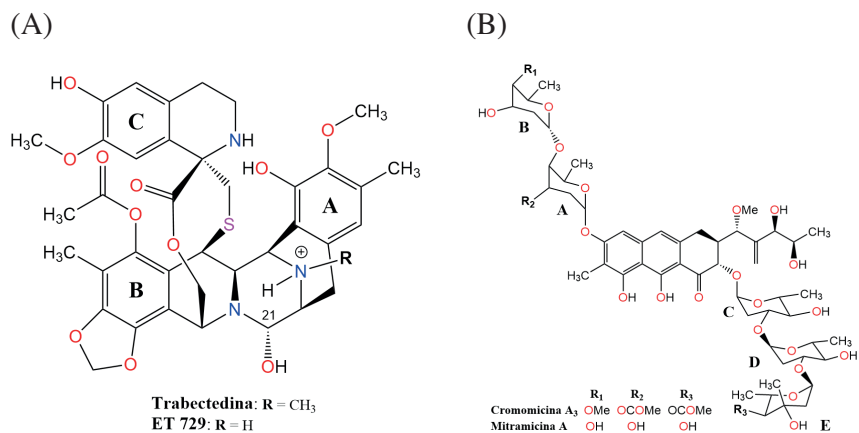


Figura 9. (A) Estructuras químicas de la trabectedina (Yondelis®) y de su principal metabolito en los seres humanos, el ET-729. Los tres anillos de tetrahydroisoquinolina fusionados en el esqueleto policíclico están etiquetados como A, B y C. Nótese las posiciones del grupo reactivo 21-carbinolamina (hemiaminal) y el N12 protonado que se utiliza para establecer un enlace de hidrógeno con un átomo aceptor (O2 o N3) situado en posición 5' con respecto a la guanina modificada en la cadena opuesta. (B) Estructura química de los derivados del ácido aureólico cromomicina A<sub>3</sub> y mitramicina A con los anillos de azúcar etiquetados de la A a la E.

La exploración de la estabilidad -mediante simulaciones de DM no restringida en disolución acuosa- y el estudio conformacional de los complejos precovalentes entre la trabectedina y dos nonámeros de ADN que contenían 5'-AGC o 5'-CGA en la región central constituyeron el primer proyecto de mi estudiante de doctorado Raquel García-Nieto (38). El complejo que contenía AGC resultó muy estable durante toda la simulación (1,4 ns en total) y dio lugar a una geometría adecuada para la formación de aductos. Por el contrario, el complejo que contenía el triplete CGA desfavorecido (que contiene un único desajuste respecto al CGG) no dio lugar a una asociación estable, ya que el fármaco se deslizó un paso “aguas arriba” a lo largo del surco menor. Estos resultados teóricos apoyaban el papel propuesto de una red de enlaces de hidrógeno bien definida en la estabilización de estos complejos (147).

A continuación, simulamos los complejos covalentes entre la trabectedina y los nonámeros de ADN que contenían los sitios diana

5'-AGC y 5'-CGG. En ambos casos, el fármaco enlazado mostró el modo de unión predicho, dio lugar a un ensanchamiento del surco menor, redujo el ángulo de torsión en los pasos ApG y CpG, e introdujo un balanceo (*roll*) positivo significativo en el par de bases implicado en la formación del enlace covalente que iba acompañado de una suave curvatura de la hélice hacia el surco mayor, en muy buena concordancia cualitativa y cuantitativa con resultados independientes de experimentos de electroforesis en gel (148). Encontramos que los llamados “diales de curvatura” (*bending dials*) (149), que son gráficos polares que muestran tanto la magnitud del ángulo de desviación del eje ( $\theta$ ) como su orientación en relación con el surco mayor ( $\phi$ ), son la representación pictórica más conveniente (Figura 10) de los parámetros helicoidales locales balanceo e inclinación, tal y como los calcula el programa CURVES desarrollado en el grupo de Richard Lavery (*Université de Lyon, France*) (150) (y hoy ampliado como CURVES+ e implementado en un servidor web (151)). Los puntos en el hemisferio norte de la esfera reflejan claramente la curvatura que comprime el surco mayor mientras que los que se encuentran en el lado izquierdo de la línea vertical indican una inclinación negativa que comprime el esqueleto de azúcar-fosfato de la cadena complementaria.

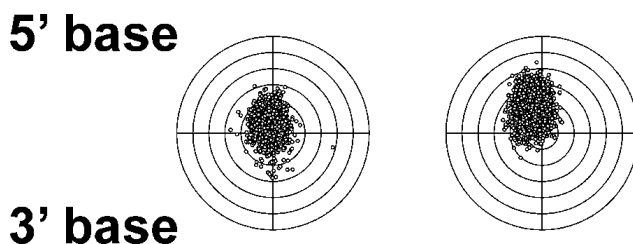


Figura 10. Ejemplo de “diales de curvatura” típicos que muestran la introducción de balanceo positivo e inclinación negativa en un paso de ADN unido a trabectedina (derecha) en relación con el mismo paso en un dúplex de ADN libre (izquierda). Las coordenadas radiales y angulares son, respectivamente,  $\theta$  y  $\phi$  (en el sentido de las agujas del reloj desde la parte superior) y cada anillo concéntrico indica una desviación de  $10^\circ$  del eje helicoidal. Los puntos individuales corresponden a estructuras separadas por 1 ps a lo largo de una trayectoria de MD de 1,5 ns, de modo que cada conjunto de 1500 puntos constituye una densidad de probabilidad.

Comparando el comportamiento dinámico de estos dúplex en ausencia y en presencia del fármaco covalentemente unido, pudimos eva-

luar las distintas curvaturas de los tramos AGC y CGG, las cuales resultaron estar en consonancia con sus propensiones intrínsecas deducidas a partir de estructuras cristalográficas (11). Así, el significativo sesgo reportado para los pasos CpG de curvarse hacia el surco mayor ya era claramente aparente en la simulación del ADN libre, mientras que el desenrollamiento parcial descrito para los pasos ApG era mucho mayor en presencia del fármaco, el cual también indujo curvatura mediante la introducción de un balanceo positivo que no existía previamente (Figura 10).

De forma bastante inesperada, descubrimos más tarde que la peculiar distorsión que la trabectedina inflige a la molécula de ADN es sorprendentemente similar a la observada en los complejos formados entre el ADN y los dedos de zinc de factores de transcripción típicos como EGR-1 (también conocido como Zif268) (152). En nuestra opinión, esto planteaba la interesante posibilidad de que la trabectedina pudiera dirigirse preferentemente al surco menor de aquellas regiones de ADN en las que una proteína que contiene uno o más dedos de zinc en su dominio de unión al ADN ya está unida en el surco mayor. El razonamiento era que la reestructuración del ADN inducida por la proteína podría estar facilitando la unión de trabectedina de preferencia a un sitio de ADN desnudo. Incidentalmente, esta hipótesis nos llevó a modelar el dominio de unión al ADN del ubicuo factor de transcripción Sp1, que consiste en un conjunto de tres módulos de dedos de zinc Cys<sub>2</sub>-His<sub>2</sub> en tándem (72), así como al estudio de un complejo covalente entre tres moléculas de trabectedina y un oligonucleótido que contiene tres sitios diana adyacentes (153). La secuencia representativa elegida fue TGGCGGCGG, que se encuentra en la región promotora “aguas arriba” del gen humano *P73* que codifica una proteína con homología estructural y funcional con la proteína supresora de tumores p53.

Nuestros resultados de DM sugirieron que la unión en tándem de varias moléculas de trabectedina a sitios adecuados que se encuentran adyacentes en una secuencia de ADN es estérica y energéticamente factible, lo que hace pensar en la posibilidad (aún no probada) de una unión cooperativa. Este complejo mostró un comportamiento muy estable, de acuerdo con el hecho de que los tripletes TGG y CGG representan sitios de unión igualmente óptimos para el fármaco. El



esquema intermolecular de enlaces de hidrógeno esperado entre cada molécula de trabectedina y los sitios YGG se mantuvo durante toda la simulación. El surco menor del ADN era notablemente más ancho que en la simulación paralela del oligonucleótido libre y la consiguiente compresión del surco mayor era comparable a la encontrada en los complejos estudiados anteriormente: se observaron aumentos del balanceo positivo en los pasos sucesivos -solo interrumpidos en los pasos GpC- y valores de inclinación negativos para los pasos CpG y TpG -indicativos de curvatura hacia el esqueleto azúcar-fosfato de su cadena complementaria-. Tanto los componentes de balanceo como los de inclinación contribuían a la curvatura de la doble hélice pero, dada la adyacencia de los sitios de unión alrededor de una vuelta de hélice, la alternancia de balanceo positivo en los pasos Y/R y R/R y balanceo cero en los pasos GpC en el lado 3' condujo a un suave y continuo desenrollamiento que no cambió la dirección general de la hélice. Esto significa que la curvatura local inducida por una sola molécula -como habíamos publicado con anterioridad (38,152)- se cancela por la unión de las otras, dando lugar a una conformación del ADN que se describe mejor como intermedia entre las formas A y B del ADN, precisamente la misma que se encuentra en los híbridos de ADN:ARN y en los complejos de ADN con dedos de zinc. Para evaluar la posible dependencia de la conformación observada para el oligonucleótido libre con respecto a la geometría de partida, este se construyó en las formas A y B (154), y ambas se simularon en condiciones idénticas. Curiosamente, la forma A evolucionó hacia una conformación notablemente similar a la obtenida cuando la simulación partió del ADN B, la cual era intermedia entre A y B en ambos casos (155).

Los tripletes reconocidos por los dedos de zinc de EGR1 y otros factores de transcripción emparentados con Sp1 incluyen GCG, GAG, GGG y TGG, mientras que los sitios favorecidos por la trabectedina son CGG, TGG, AGC y GGC (147). Aunque diferentes combinaciones de estos tripletes ofrecen sitios de unión tanto a las proteínas como a los fármacos, la mayoría de los sitios óptimos para la trabectedina estarían localizados en las uniones entre los tripletes que son reconocidos por los sucesivos dedos de zinc y no en los propios tripletes (TGG sería la excepción más obvia). Sorprendentemente, la superposición directa de los tripletes TGG, CGG y CGG con la trabectedina

unida sobre los correspondientes tripletes de ADN consenso reconocidos por cada dedo de zinc de EGR1 (es decir, GCG, GGG y GCG) mostró una coincidencia muy notable ( $rmsd = 1,1 \text{ \AA}$  utilizando los átomos C1') para los tres tripletes. Además, se podía obtener una semejanza estructural igualmente buena ( $rmsd = 0,9 \text{ \AA}$ ) cuando un complejo se desplazaba un par de bases hacia arriba o hacia abajo al hacer la superposición. La mejor coincidencia depende en última instancia de las secuencias particulares que pueden ser reconocidas simultáneamente por el fármaco y los diferentes dedos de zinc de un factor de transcripción o de la proteína concreta que se une al ADN. En cualquier caso, nuestros descubrimientos apoyaron aún más la existencia putativa de un complejo ternario proteína-ADN-trabectedina (152) y abrieron nuevas vías de investigación.

En el último caso presentado, la unión de la trabectedina tuvo lugar de forma directa en una sola de las cadenas de ADN. Sin embargo, cuando fuimos conocedores de que los parámetros de actividad de la trabectedina en el panel de 60 líneas celulares de tumores humanos del *Anticancer Drug Screen* del Instituto Nacional del Cáncer (*National Cancer Institute*, NCI) de los EE.UU. de Norteamérica estaban altamente correlacionados -utilizando el algoritmo COMPARE (156)- con los de la cromomicina  $A_3$  (Figura 9B), decidimos modelar y simular la estructura del complejo formado entre el dodecanucleótido autocomplementario  $d(\text{GTATGGCCATAC})_2$  y dos moléculas de trabectedina, cada una unida covalentemente a una hebra diferente, en disposición "cola a cola" (157). Seleccionamos esta secuencia en particular porque se había resuelto por cristalografía de rayos X la estructura de un complejo entre el octanucleótido  $d(\text{TTGGCAA})_2$  y dos moléculas de cromomicina  $A_3$  (cuya estructura tiene poco que ver con la de las ecteinascidinas) queladas por un ión  $\text{Mg}^{2+}$  de tal forma que el dímero reconocía secuencias TGG adyacentes en ambas cadenas (158). Para comparar y completar, también simulamos la misma secuencia en condiciones idénticas en ausencia de cualquier fármaco. Esto nos permitió evaluar las preferencias conformacionales de (i) el triplete TGG, que está omnipresente en secuencias de eucariotas y virus humanos, (ii) su secuencia complementaria, es decir, CCA, que es un componente importante de la llamada "caja CCAAT" presente en un gran número de promotores de genes (79), y (iii) la yuxtaposición de los dos, es decir, la secuencia TGGCCA que ha sido denominada por algunos cristalógrafos de rayos

X como un “elemento natural de curvatura “ (159).

En consonancia con observaciones anteriores (147), la unión de trabectedina dio lugar a una mayor estabilización del ADN, puesta de manifiesto por las menores fluctuaciones cuadráticas medias (*rmsd*) durante la simulación del complejo en relación con la del oligonucleótido libre. En el complejo 2:1, (i) cada subunidad A de la trabectedina sobresale perpendicularmente de la hélice frente a la guanina a la que está unida, (ii) las subunidades B se apilan sobre los anillos de azúcar de las citosinas complementarias de una manera que recuerda a la de los típicos ligandos del surco menor no intercalantes, y (iii) las subunidades C exponen un lado plano al disolvente, mientras que el otro lado hace amplios contactos con el esqueleto azúcar-fosfato de los dos nucleósidos que se encuentran “aguas abajo” de cada guanina covalentemente unida. Además, en este tipo de disposición “cola a cola”, ambas subunidades B establecen contactos de van der Waals favorables en la zona de conexión de ambos tripletes de ADN.

El mayor desenrollamiento de la hélice tras la unión de trabectedina tuvo lugar en los pasos TpG (=CpA), que también mostraron los mayores valores de ángulo de balanceo, en buena concordancia con la conocida relación inversa entre este parámetro y el giro helicoidal [10]. Por el contrario, el paso GpC mostró un balanceo negativo y también la mayor torsión helicoidal dentro de la secuencia TGGCCA. En general, las características conformacionales del complejo recordaban a las encontradas en el complejo 3:1 cabeza-cola descrito anteriormente y corresponden, una vez más, a una forma intermedia entre el ADN-A y el ADN-B (153). Sorprendentemente, los complejos 2:1 de trabectedina:ADN y cromomicina:ADN revelaron varias notables características en común, tales como (i) un ADN desenrollado con un surco menor considerablemente ensanchado en la región diana TGGCCA prácticamente superponible (*rmsd* de tan solo 2.1 Å sobre los átomos del esqueleto azúcar-fosfato); (ii) un patrón similar de incremento del balanceo (*roll*) positivo en los pasos TpG y un pequeño balanceo positivo en los pasos GpG, junto con un balanceo negativo en el paso GpC central, que era más marcado en el complejo de trabectedina; y (iii) un deslizamiento (*slide*) negativo en ambos pasos GpG, en buena concordancia con el hallazgo de que los ángulos de torsión del enlace glicosídico,  $\chi$ , eran más cercanos a los del ADN de forma A.

Nuestra simulación del dodecámero libre mostró dos curvaturas distintas producidas por el balanceo en cada paso TpG, pero ninguna en la unión entre TGG y CCA, en concordancia con los resultados de un estudio de espectroscopía de RMN bidimensional que informaba de que GGC en el decámero autocomplementario relacionado  $d(\text{CATGGCCATG})_2$  forma una pila apretada con bases paralelas en la que ambos pasos TpG presentan un alto balanceo positivo (160). Este resultado, sin embargo, discrepaba de observaciones anteriores de cristalografía de rayos X que mostraban que las dobles hélices aplastadas de moléculas de ADN tipo B de secuencia general CCAxxxTGG (159) exhiben la misma curvatura de  $23^\circ$  a través de la unión no enlazada TGG CCA en la red cristalina que la encontrada en el centro del decámero CATGCATG (161). Dado que el alto balanceo en GGC no se observa en los pasos GpC de otras secuencias, la explicación más probable es que este efecto estructural, junto con el inusualmente alto deslizamiento positivo adoptado por los pasos CpA(=TpG) y la conformación  $B_{II}$  no estándar del esqueleto azúcar-fosfato en estas formas cristalográficas, se deben a fuertes interacciones intermoleculares en la red cristalina (162).

Como se ha demostrado que tanto la trabectedina como la cromomicina  $A_3$  ejercen su citotoxicidad -al menos en parte- interfiriendo en la replicación y la transcripción celular, propusimos -a la luz de la información estructural obtenida- que estos efectos pueden deberse en gran medida a la paralización de las horquillas de replicación y transcripción mediante la estabilización de la estructura de doble hélice del ADN y la prevención de la separación de las hebras, de modo semejante a como lo consiguen los clásicos entrecruzadores de cadenas de ADN (*interstrand crosslinkers*, ICL). A este respecto, varias propuestas de la probable existencia de uno o varios complejos ternarios citotóxicos de trabectedina:ADN:proteína que implican a miembros de la maquinaria de reparación del ADN (163) y a la propia ARN polimerasa II (164) nos impulsaron a modelar algunos de estos ensamblados macromoleculares (165,166). Creo que estos (y otros) resultados del trabajo interdisciplinar son una buena muestra de que la colaboración entre experimentalistas y modeladores moleculares puede ser muy enriquecedora y rendir beneficios mutuos.

## 8. Otros ejemplos de antitumorales tetrahydroisoquinólicos que se unen al surco menor del ADN.

Cuando estudiamos las secuencias preferidas de unión de ET729 -el análogo N12-desmetilado de la trabectedina (Figura 9A)-, encontramos que este compuesto tiene un repertorio ampliado de sitios de unión en comparación con la trabectedina (ET-743), incluyendo el triplete CGA (167). Esta observación resultaba de interés, no solo porque ET729 es uno de los principales metabolitos de Yondelis® en los seres humanos (a través de la acción del CYP3A4) -y por lo tanto puede contribuir a algunos aspectos de la actividad clínica de la molécula madre-, sino también porque su estudio podría ayudar a aclarar aspectos mecanísticos de la reacción que conduce a la unión covalente. En efecto, encontramos que los tripletes CGG proporcionan buenos sitios de unión para la trabectedina porque el anclaje del grupo metilendioxi por el grupo amino presente en el surco menor del lado 3' del triplete determina un sutil cambio en la orientación general que es suficiente para contrarrestar el choque estérico entre el grupo metilo sobre N12 y el grupo amino del par de bases situado en el lado 5'. Cuando falta esta interacción de enlace de hidrógeno porque se encuentra un par de bases A:T en el lado 3' (por ejemplo, en el triplete CGA), la trabectedina discrimina en contra de G:C en el lado 5' de la guanina modificada covalentemente. Esta interpretación también proporciona una justificación al hallazgo de que CGT, GGT y GGA son sitios desfavorecidos por la trabectedina. Por el contrario, la falta del grupo metilo en la posición N12 de ET729 permite a este compuesto alcanzar la yuxtaposición óptima necesaria para promover su activación por guanina en un sitio CGA. El hecho de que estas diferencias -hasta entonces subestimadas- en la orientación y geometría de la unión fueran suficientes para racionalizar los comportamientos encontrados experimentalmente reveló, en nuestra opinión, el poder de las simulaciones de DM de última generación para sondear los detalles de las interacciones ligando-ADN en ausencia de restricciones experimentales.

El PM00104 (Zalypsis®) es un análogo de la trabectedina diseñado por científicos de PharmaMar que contiene dos anillos de tetrahydroisoquinolina fusionados y un sustituyente más lineal en lugar del anillo C. Al igual que Yondelis®, también forma enlaces covalentes con el

grupo amino de guaninas seleccionadas en el ADN de doble cadena, tanto en ausencia como en presencia de citosinas metiladas, gracias a un grupo hemiaminal reactivo. Los experimentos de desnaturalización térmica de oligonucleótidos de secuencia definida demostraron que los tripletes de ADN más favorables para la formación de aductos covalentes son TGG, CGG, GGC, AGC y AGG, y también que el aumento medio de la temperatura de fusión del ADN que provoca su aducto en la doble hélice es aproximadamente la mitad del logrado con la trabectedina y comparable al alcanzado con la equinomicina o la tiocoralina. Pudimos racionalizar las preferencias del sitio de unión y la distinta estabilización de la doble hélice de ADN mediante la modelización molecular de los complejos precovalentes entre Zalypsis<sup>®</sup> y una serie de oligonucleótidos de secuencia general 5'-d(ATAXYZATAATA)/5'-d(TATTATZ'Y'X'TATTAT), donde XYZ/Z'Y'X' representan AGA/TCT, AGG/CCT y AGC/GCT (168).

En la lurbinectedina (antes conocida como PM01183 y hoy día comercializada como Zepzelca<sup>®</sup> -previamente Zepsyre<sup>®</sup>-) la sustitución de la subunidad C de la tetrahydroisoquinolina de la trabectedina por una tetrahydrocarbolina confiere a este análogo algunas ventajas farmacocinéticas y de actividad intrínseca (169). La modelización y simulación mediante DM de los complejos precovalentes y covalentes entre PM01183 y un oligonucleótido de secuencia 5-d(CAATACGATAAG)/5-d(CTTATCCGTATTG) reveló un surco menor ensanchado y un surco mayor comprimido en la región central debido a los aumentos significativos del balanceo en los pasos C6/G7 y G7/G8. Se pensó que la ocupación del surco menor y la exposición al disolvente de algunos grupos funcionales podrían tener un impacto en la capacidad de este complejo binario ligeramente doblado para asociarse o interferir con una o más proteínas de unión al ADN implicadas en la transcripción, la replicación y/o la reparación dentro de la célula, pero no necesariamente las mismas que participan en el reconocimiento de los aductos trabectedina-ADN.

La unión de este tipo de antitumorales tetrahydroisoquinolínicos al ADN celular acaba dando lugar a roturas de la doble hebra que, si no son reparadas, conducen a la muerte de la célula. Utilizando ensayos *in vitro* e *in vivo* bien definidos, se ha demostrado que los aductos resultantes en el ADN (i) estimulan, de forma dependiente de la con-

centración, la escisión por parte de la nucleasa XPF/ERCC1 de la hebra de ADN opuesta a la unida por el fármaco, (ii) impiden la unión al ADN de factores de transcripción como Sp1, y (iii) detienen a la ARN polimerasa II en la fase de elongación en la misma posición nucleotídica, independientemente de la hebra en la que se encuentren. Nuestros modelos estructurales proporcionan una justificación para estos hallazgos y destacaron la semejanza entre este tipo de modificación del ADN y un enlace cruzado entre hebras (169).

## 9. Fusión simulada del ADN en el ordenador.

Ya he comentado brevemente que la diferencia en la temperatura de fusión de una molécula de ADN de doble cadena en ausencia y en presencia de ligandos unidos de forma no covalente (o covalente) puede proporcionar información experimental sobre la estabilización térmica provocada por el ligando. Una pregunta pendiente en mi mente era si esto también podría lograrse *in silico* (es decir, en el ordenador). Una búsqueda en la literatura reveló que el grupo de B. Montgomery (“Monti”) Pettitt, en la Universidad de Houston (Texas), había realizado simulaciones de DM a 400 K de un dúplex de ADN homopolimérico de secuencia d(A<sub>12</sub>)-d(T<sub>12</sub>) en agua salada explícita durante 100 ns (170). La elección de la secuencia obedeció a consideraciones tanto de simplicidad como de índole práctica, ya que se estimó en 318 K la temperatura de fusión de este dodecámero en las mismas condiciones de la simulación, lo que hacía pensar en la posibilidad de éxito en un marco de tiempo computacionalmente factible. Las dos simulaciones solo diferían en la forma de asignar las velocidades aleatorias, pero en ambos casos fueron evidentes: (i) la pérdida de los enlaces de hidrógeno WC en los extremos del dúplex (“deshilachado”), (ii) movimientos mediante los cuales aumenta el grado de desenrollamiento debido a la formación de enlaces de hidrógeno no canónicos con el esqueleto azúcar-fosfato de una hebra (“despegado”), (iii) intercambios incrementales de los enlaces de hidrógeno WC entre las hebras (que dan como resultado cambios de registro de una secuencia con relación a la otra), y (iv) eventos de rehibridación de las dos hebras previamente separadas.

Animado por este precedente, pero ligeramente preocupado por si

nuestros recursos computacionales -que incluían un acceso limitado al superordenador de Barcelona “Mare Nostrum”- estarían a la altura de la tarea, entregué este reto a mi nuevo estudiante de doctorado Juan A. Bueren-Calabuig, que se encargó con gusto de la simulación del comportamiento dinámico -y, con suerte, de la desnaturalización- de un dúplex de secuencia 5'-d(TAATAACGGATTATT):5'-d(AATAATCCGTTATTA) en disolución acuosa de NaCl 0,1 M. Afortunadamente, en algo menos de 200 ns de simulación fuimos capaces de caracterizar con detalle atómico el perfil completo de desnaturalización térmica de este dúplex de secuencia mixta A:T que contiene un triplete central CGG adecuado para la unión de diversos fármacos. Esto significó que estábamos en una posición adecuada para intentar simular los monoadductos formados entre la guanina subrayada en este oligonucleótido y diferentes ligantes covalentes, a saber, Yondelis<sup>®</sup>, Zalypsis<sup>®</sup> y PM01183 (Zepzelca<sup>®</sup>), así como el bisadducto formado en el paso CpG central con la mitomicina C (MMC), agente prototípico del grupo de los ICL.

Observamos una sorprendente asimetría del proceso de fusión a ambos lados del triplete CGG central y comprobamos que el proceso de separación de las hebras se veía inequívocamente afectado por la presencia de los adductos de los fármacos. Así, mientras que la descompresión completa de la hélice y la pérdida del emparejamiento de bases WC fue evidente para el 15-mero libre de fármacos después de 190 ns de simulación, los eventos de rehibridación fueron mucho más conspicuos y duraderos en los dúplex que contenían monoadductos, cuyas regiones centrales todavía permanecían unidas por enlaces de hidrógeno de modo significativo al final de los 200 ns de simulación (171). En nuestra opinión, estas diferencias significativas proporcionan evidencia adicional de la capacidad demostrada de estos fármacos para estabilizar el ADN de doble cadena (ADNdc; *double-stranded DNA*, dsDNA), detener las horquillas de replicación y transcripción, y reclutar proteínas de reparación del ADN. Esta estabilización, cuantificada aquí en términos de pares de bases correctamente apareados, también apoyó nuestra hipótesis de que un monoadducto de este tipo puede imitar funcionalmente un enlace cruzado entre las dos hebras de la hélice de ADN (como hacen los ICL), a pesar de que estos fármacos están unidos covalentemente a una sola hebra.



Sin embargo, en contra de nuestras expectativas, el oligonucleótido que contenía el bisaducto de MMC siguió un curso muy similar al del ADN libre de fármacos, siendo la principal diferencia que las dos hebras de ADN no podían separarse físicamente por completo debido a la existencia del entrecruzamiento de hebras por medio de las dos guaninas que componen el paso central CpG. Así, al final de la simulación solo permanecían unidos por hidrógeno WC estos dos pares de bases G:C. También resultó interesante la caracterización en este complejo de una serie de intermedios únicos colapsados hidrofóbicamente en los que el sistema de anillos de la MMC aparecía pseudointercalado en estructuras secundarias transitorias del ADN. Estas observaciones sugerían de forma clara que estos aductos podrían plantear retos distintos a la maquinaria de reparación del ADN, lo cual podría explicar algunas de las diferencias observadas experimentalmente en células de mamíferos entre las acciones de la MMC y las tetrahidroisoquinolinas antitumorales que se unen covalentemente al surco menor del ADN (172).

## **10. La formación de bisaductos de mitomicina como caso de prueba para los métodos QM/MM.**

A medida que los métodos híbridos QM/MM seguían madurando y se implementaban en el conjunto de programas AMBER, tradicionalmente empleados en mi laboratorio, me sentí irremediamente tentado a intentar simular la formación de un enlace covalente entre un ligando y una molécula de ADN. ¿Qué mejor opción que la MMC, un profármaco antitumoral de producción microbiana estudiado experimentalmente a lo largo de los años -sobre todo por Maria Tomasz, Don Crothers y sus respectivos colaboradores- y utilizado regularmente como el ICL tradicional? Además, la sensibilidad a la MMC ha servido para clasificar a los pacientes con anemia de Fanconi (AF) en diferentes grupos de complementación y ha sido fundamental para delinear la llamada vía AF/BRCA de la respuesta al daño sobre el ADN (173).

La capacidad única de la MMC para entrecruzar las hebras complementarias de la doble hélice específicamente en los pasos 5-CG-3 (CpG) en las células hipóxicas es especialmente notable -dado el pe-

queño tamaño de la molécula y su arquitectura restringida- y se basa en la siguiente secuencia de eventos: (i) reducción intracelular de su anillo de quinona y apertura de su anillo de aziridina alquilante con la participación de los átomos C1 y C2, (ii) formación de un enlace covalente con el grupo 2-amino de guaninas seleccionadas en el surco menor del ADNc, y (iii) evolución del aducto monofuncional, pero exclusivamente en los pasos CpG, hacia una lesión bifuncional (o intercatenaria) más deletérea conocida como enlace cruzado entre hebras (Figura 11). El examen de la literatura relevante reveló otros dos aspectos intrigantes que también llamaron nuestra atención, a saber, (i) la reactividad mejorada o disminuida de la MMC cuando las bases de citosina están metiladas o sustituidas por 5-fluorocitosina (5FC), respectivamente, y (ii) las estereoquímicas opuestas de los principales aductos de ADN producidos en células de mamífero por la MMC y su análogo cercano, la 10-descarbamoil-mitomocina C (DMC).

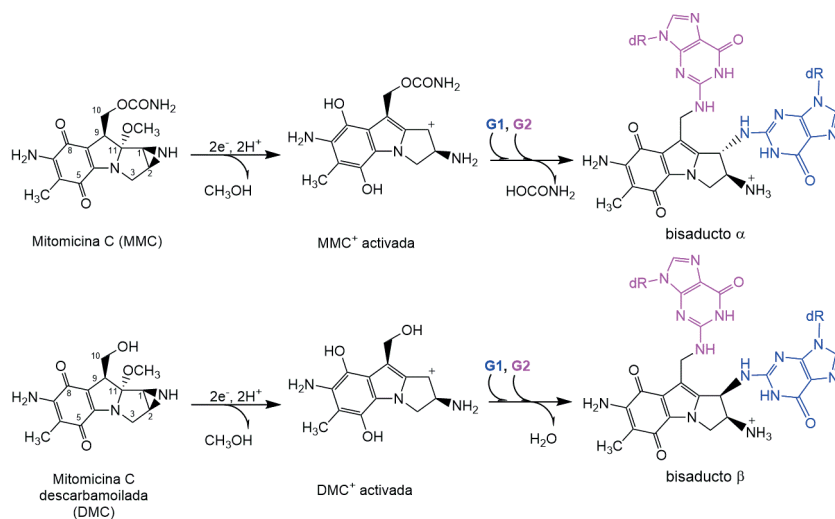


Figura 11. Esquemas que muestran las rutas de reacción que conducen a la formación de los monoadductos ADN:MMC $\alpha$  (arriba) y ADN:DMC $\beta$  (abajo) y sus correspondientes bisadductos (o entrecruzamientos intercatenarios, *interstrand crosslinks*) de estereoquímica opuesta. Obsérvese que las funcionalidades alquilantes del ADN de la MMC y la DMC se desenmascaran en las células vivas solo tras la activación biorreductora.

Nuestra aproximación a este problema, de nuevo en manos de Juan

A. Bueren-Calabuig (entre paréntesis, hijo de mi amigo, colega y antiguo compañero de la Facultad de Farmacia, Juan. A. Bueren, uno de los mejores expertos españoles en AF) fue estudiar los complejos precovalentes formados entre las formas activadas de MMC y DMC y tres dúplex de ADN de fórmula general d(CGATAAXGCTAACG) en una hebra -en la que X significa C, 5-metilcitosina (5MC) o 5FC- y la secuencia complementaria en la hebra opuesta. En aras de la exhaustividad, estos tres oligodesoxinucleótidos totalmente solvatados también se simularon en sus estados libres. Todas las trayectorias generadas se analizaron en términos de parámetros estructurales, geometrías adecuadas para el ataque nucleófilo, PEM y energías de unión (que incluían un término electrostático corregido por el efecto del medio acuoso). A continuación, simulamos, de forma escalonada, la formación de los respectivos monoadductos para MMC y DMC con la guanina subrayada en el sitio CGC de alta afinidad presente en el centro del 14-mero estándar y su posterior evolución a los (distintos) bisadductos finales acoplando una función QM a un potencial MM clásico. La principal ventaja de este enfoque híbrido para los complejos fármaco-ADN solvatados respecto a estudios anteriores que se centraban en sistemas modelo simplificados es que las redistribuciones de carga y los cambios en las conectividades de los átomos (es decir, la ruptura y la creación de enlaces) se realizan en un contexto dinámico que permite de forma realista que los solutos, las moléculas de agua y los contraiones se reorganicen a lo largo de todo el proceso (174).

Nuestro primer descubrimiento fue que el enlace de hidrógeno entre el oxígeno del carbamato unido a C10 y el grupo amino exocíclico de G22 (la guanina emparejada con C7), que supuestamente dominaba el reconocimiento de la secuencia en el complejo precovalente, era débil y de corta duración en presencia de moléculas de agua competidoras. Además, observamos que la orientación propuesta para el fármaco en el surco menor era incompatible con un ataque nucleófilo por parte del grupo amino de G8 sobre la cara  $\alpha$  del anillo de tetrahidropirrol de la MMC. Por lo tanto, exploramos modos alternativos de enlace de hidrógeno y descubrimos que es más probable que el oxígeno unido a C10 establezca un enlace de hidrógeno interno con el grupo fenol en C8 mientras que el oxígeno carbonílico participa en una interacción de enlace de hidrógeno con el grupo amino de G22. Esta nueva pose de unión (y “conformación cercana al ataque”, *near-attack confor-*

*mation*) no solo se mantuvo estable debido a las interacciones intermoleculares adicionales, sino que también proporcionó una distancia adecuada (es decir,  $\leq 4$  Å) para el ataque nucleófilo de N2(G8) sobre el C1 de la MMC activada en la cara del anillo opuesta a la del grupo amino primario en C2 (Figura 11). En el caso de la DMC activada, que carece del oxígeno carbonílico, la orientación del fármaco era tal que el ataque de N2(G8) se produciría preferentemente -aunque no exclusivamente sobre la cara  $\beta$  del anillo de tetrahidropirrol del fármaco. De este modo, nuestros resultados apoyaban la idea de que la aparición relativa muy distinta de los monoadductos de ADN tipo  $\alpha$  y  $\beta$  causados por la MMC y la DMC está dictada en gran medida por la naturaleza del enlace de hidrógeno en el que está implicado el oxígeno del carbamato o del hidroxilo de ambos fármacos en su forma activa.

Una vez identificado el modo de unión putativo para ambos fármacos en el surco menor de la región central C7:G22/G8:C21, realizamos simulaciones de DM similares con los oligodesoxinucleótidos en los que C7 y C21 habían sido sustituidas por 5MC o 5FC. El cálculo de las energías de unión a lo largo de las trayectorias utilizando el método MM-GBSA, tal y como se implementa en AMBER (175), y los análisis de descomposición de la energía mostraron que un C21 metilado, es decir, la citosina complementaria a la guanina sometida a alquilación, mejoraba significativamente su interacción electrostática con el fármaco cargado positivamente en comparación con una C estándar o una 5FC, en buena consonancia con los mapas de PEM manifiestamente distintos. La inspección de las geometrías de apilamiento entre los pares de bases vecinas G8:C21 y C9:G20 también reveló un desplazamiento significativo hacia valores de inclinación negativos en relación con el ADN libre en los complejos precovalescentes, una característica que está presente de forma natural -en ausencia de fármaco unido- cuando C9 y C21 están metilados y argumenta a favor de una preorganización estructural intrínseca para la unión de MMC a los sitios 5MCpG. En su conjunto, estos resultados de modelado molecular estaban en muy buena consonancia con la evidencia experimental de que el aumento de la velocidad de monoalquilación de la G inducido por la 5MC se debe a la base 5MC emparejada más que a la que flanquea el lado 5' (176).

La coordenada de reacción QM que condujo a la formación del enla-

ce covalente consistió en acortar progresivamente la distancia entre el N2 de G8 y el C1 del fármaco utilizando como punto de partida distintos “fotogramas” extraídos de las trayectorias de DM de los complejos precovalentes. La imagen consistente obtenida fue que la estereoquímica  $\alpha$  o  $\beta$  opuesta en el C1 del aducto resultante (Figura 11) es totalmente dependiente del posicionamiento de una determinada conformación del fármaco activado en el surco menor del ADN, como se ha comentado anteriormente. Relajamos las estructuras de los complejos MMC-ADN y DMC-ADN resultantes -que contienen, respectivamente, los monoaductos  $\alpha$  y  $\beta$ -, tras reubicar el H22 extra del N2(G8) sobre el grupo amino del fármaco enlazado, y las equilibramos durante 10 ns de simulaciones de DM sin restricciones. Las principales características del complejo MMC-ADN fueron (i) la ausencia de desviaciones significativas en los parámetros helicoidales con respecto a los del ADN libre, y (ii) dos posibles ubicaciones descentradas del fármaco unido covalentemente dentro de un surco menor ligeramente ensanchado, dependiendo de la orientación del carbamato de la cadena lateral. Por el contrario, el surco menor del ADN de la región central que alberga el aducto  $\beta$  de DMC estaba ligeramente comprimido debido a la interacción simultánea del fármaco con el esqueleto de azúcar-fosfato de ambas hebras y a la interacción sostenida de enlace de hidrógeno entre el O2(T23) y el grupo hidroxilo en C10 de la DMC.

Una vez equilibradas las estructuras de los dúplex que contienen los monoaductos  $\alpha$  y  $\beta$ , procedimos a (i) generar la segunda especie alquilante, eliminando del fármaco bien el carbamato (en el caso de la MMC), bien el hidroxilo (en el caso de la DMC), (ii) equilibrar durante 10 ns utilizando la DM estándar, y (iii) cambiar al método QM/MM para acortar la distancia N2-C10 y crear el segundo enlace covalente. Equilibramos los bisaductos resultantes igualmente durante 20 ns y las respectivas trayectorias mostraron que el ADN puede acomodar esta modificación covalente dual (sobre G8 en una hebra y G22 en la hebra opuesta) sufriendo una distorsión mínima, como se informó originalmente (178), solo en el caso del bisaducto C1 $\alpha$ . En el C1 $\beta$ , por el contrario, encontramos que los parámetros de elevación (*rise*), desplazamiento (*shift*) y balanceo (*roll*) en el paso C7pG8 se desviaban significativamente de los valores en el ADN libre. Curiosamente, el balanceo negativo indica curvatura de la hélice hacia el

surco menor, en contraposición con lo observado en los dúplex de ADN que contienen aductos monofuncionales con la trabectedina o la lurbinectedina (véase la sección 8), fármacos que parecen comportarse funcionalmente como imitadores o miméticos de los ICL [140].

Concluimos este trabajo afirmando que las diferencias conformacionales reportadas para estos aductos, que aparecen en distintas proporciones en el ADN de las células que han sido expuestas a la MMC o a la DMC, es probable que planteen desafíos diferentes a las maquinarias de reparación del ADN. Esto puede explicar los hallazgos experimentales de que las células AF son hipersensibles a la MMC pero tienen una sensibilidad normal a la DMC, mientras que en otros tipos celulares la DMC es más citotóxica que la MMC (178).

## **11. Curvatura del ADN en relación con la acción antitumoral de ciertos fármacos citotóxicos.**

La curvatura del ADN es el resultado de una combinación de características intrínsecas de la doble hélice y de distorsiones introducidas por la unión de diversas proteínas o ciertos fármacos. La propensión de ciertas secuencias de ADNdc a curvarse es esencial en procesos biológicos esenciales, como son la replicación y la transcripción, durante los cuales se sabe que algunas proteínas reconocen las conformaciones no canónicas del ADN o promueven su formación tras unirse a él. Entre todos los elementos que se han estudiado en relación con la curvatura del ADNdc, los “tractos A” destacan como uno de los más controvertidos. La teoría de la “cuña ApA” fue refutada cuando se demostró que (i) una serie de polinucleótidos lineales que contenían tramos  $5'-A_4T_4-3'$  presentaban curvatura mientras que los que contenían tramos  $5'-T_4A_4-3'$  eran rectos (179), y (ii) la estructura cristalográfica de un dodecámero de ADN tipo B que contenía un tracto homopolimérico de seis pares de bases A:T era esencialmente recto. Además, algunos de los minicírculos de ADNdc más pequeños descritos hasta la fecha (de un tamaño de unos 100 pares de bases, pb) carecen de tractos A y están sometidos a distintos niveles de tensión torsional (180).

Alberto Mills, en mi laboratorio, fue el encargado de modelar con

detalle atómico secuencias de ADN representativas de estos experimentos publicados y de simular su comportamiento dinámico durante cientos de nanosegundos utilizando el campo de fuerzas ParmBSC1 (33) implementado en AMBER. El análisis posterior de las trayectorias resultantes nos permitió (i) establecer de forma inequívoca la localización de la curvatura en todos los casos; (ii) identificar los elementos estructurales que son directamente responsables de la curvatura detectada macroscópicamente; y (iii) revelar la importancia no solo de sumar coherentemente los efectos de los elementos de curvatura cuando están en sincronía con la repetición natural de la hélice (es decir, separados por un número entero de vueltas helicoidales), sino también cuando se alternan con una separación de efectos opuestos cada media vuelta de hélice (181). Conseguimos demostrar de forma independiente que el principal determinante de la curvatura observada macroscópicamente es el agrupamiento “en fase” del balanceo (*roll*) positivo impuesto por los pasos YR y el balanceo negativo o nulo característico de los pasos RY y los tractos A, respectivamente. Esta conclusión concuerda muy bien con los parámetros estructurales derivados experimentalmente para moléculas de ADN mucho más pequeñas, ya sea solas o en forma de complejos ADN:proteína (182).

A continuación, como sabíamos que los aductos de G:trabectedina (Yondelis<sup>®</sup>) inducen curvatura en el ADN y aumentan -tras la acción de una ADN-ligasa- la ratio de circularización de varias construcciones de ADNdc -pero no de otras- (148), extendimos nuestras simulaciones de DM sin restricciones a modelos atómicos explícitamente solvatados de (i) una serie de oligonucleótidos de 21, 42 y 63 pb (es decir, múltiplos de 2, 4 ó 6 repeticiones de una vuelta de hélice), tanto en presencia como en ausencia de trabectedina, y (ii) minicírculos de 6×21 pb formados por el ligado de dos tipos de 21-meros tanto “desnudos” como con un aducto de G:trabectedina en la región central. Estas longitudes representaban un compromiso entre la posibilidad de tratarlos computacionalmente durante un tiempo suficiente y el objetivo de disponer de información estructural sobre cada paso de dinucleótido presente en las secuencias estudiadas, incluyendo los que componen las uniones (que se descuidan cuando solo se estudian los oligos de 21 pb de forma aislada). Los resultados de este trabajo nos permitieron explicar en términos estructurales detallados (i) la tendencia intrínseca a doblarse de aquellas secuencias en las que

los elementos de curvatura están dispuestos de forma sincrónica con respecto a la repetición natural de la hélice y (ii) por qué la disposición asíncrona de estos componentes impide la ciclación de dímeros y trímeros y solo permite la ligadura lineal (183). Adicionalmente, la inclusión de la propia ADN-ligasa T4 en algunas de las simulaciones nos permitió arrojar luz sobre el papel crucial, aunque posiblemente subestimado, que desempeña la propia enzima en la estabilización de una conformación de ADN curvada antes de la ciclación. En su conjunto, estos resultados recientes amplían nuestra comprensión actual sobre cómo la modificación de la forma del ADN por la trabectedina puede afectar tanto al reconocimiento específico de secuencia por parte de algunos factores de transcripción que se unen a los promotores como a la unión de la ARN polimerasa II durante el proceso de transcripción.

Posteriormente hemos aplicado los mismos métodos teóricos al estudio de las características estructurales y dinámicas asociadas a la selección de secuencias diana en el ADN por parte de las calicheamicinas, otro grupo de potentes antitumorales de origen natural (184,185). La unión de la calicheamicina  $\gamma_1^I$  y análogos reactivos al surco menor de secuencias bien definidas del ADN acaba provocando roturas de la doble hélice que provocan la muerte de la célula. De forma complementaria, el análogo no reactivo calicheamicina  $\epsilon$  (Figura 12) ha servido para demostrar que estas moléculas estructuralmente sofisticadas curvan la estructura helicoidal de ciertas secuencias de ADN hasta el punto de aumentar su ratio de circularización al concatenarse en multímeros ( $n = 5-13$ ) por acción de una ADN ligasa (186). En primer lugar modelamos y simulamos mediante DM un 21-mero de ADNdc (CA21) que contenía dos sitios diana, uno de alta afinidad (5'-TCCT-3') y otro de baja afinidad (5'-TTGT-3'). Pudimos demostrar que las diferentes afinidades observadas para ambos tetranucleótidos se deben a la ausencia de un grupo amino exocíclico en la cadena complementaria del segundo sitio que es crucial para la formación de un enlace de hidrógeno con el átomo de azufre del grupo tiobenzato.



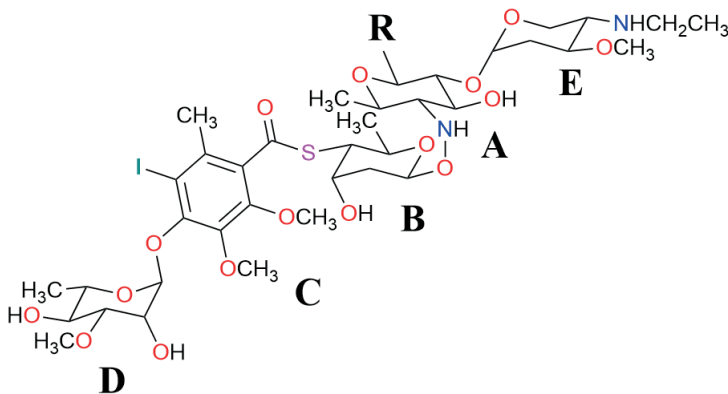
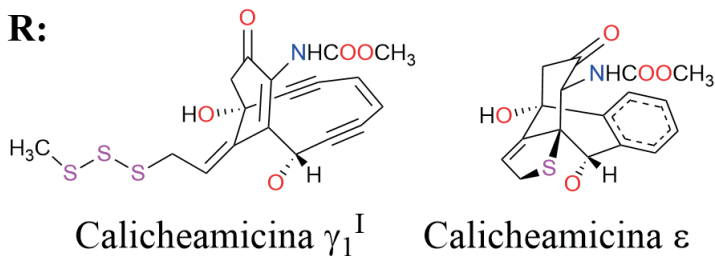


Figura 12. Estructuras químicas de la calicheamicina  $\gamma_1^I$  (izquierda) y del análogo no reactivo calicheamicina  $\epsilon$  (derecha), con los anillos etiquetados con letras mayúsculas.

A continuación estudiamos el dímero CA21x2 lineal -en presencia y ausencia de cuatro moléculas de calicheamicina  $\epsilon$  unidas en el surco menor- y un multímero circularizado de 126 pb (CA21x6) -tanto en ausencia como en presencia de doce moléculas de calicheamicina  $\epsilon$  en el surco menor-. Este multímero contenía un corte en cada hebra para simular la construcción de ADN sobre la que la ADN ligasa realizaría la última ligadura conducente al círculo covalentemente cerrado. Aunque los dímeros lineales (notablemente más cortos que la longitud de persistencia de la molécula de ADN, es decir,  $\sim 150$  pb (187)) no mostraron una tendencia intrínseca a curvarse -como se esperaba-, observamos que la hélice de ADNc sufre distorsiones locales tras la unión de CA $\epsilon$  que dan lugar a un patrón de valores de balanceo (*roll*) positivo adecuadamente "en fase" que propician un aumento de curvatura. Resultó particularmente llamativo el hecho de que los patrones de balanceo (*roll*), giro (*twist*) e inclinación (*tilt*)

observados para el ADN circular unido al fármaco no coincidían con los mostrados por los círculos de ADN desnudo. Esto significa que CAL $\epsilon$  no refuerza una conformación curvada predeterminada de esta secuencia, sino que da lugar a una combinación diferente de ángulos de balanceo debido a las distorsiones introducidas en el ADN tras el reconocimiento y firme fijación a sus secuencias diana. Es, por tanto, la reorganización de estos elementos lo que fuerza la mayor curvatura observada en presencia de las moléculas de fármaco (188).

## 12. Las lamelarinas como venenos de la topoisomerasa I.

Uno de los objetivos de los estudios de modelización molecular durante muchos años ha sido comprender la forma en que la camptotecina (CPT) y otros compuestos relacionados, como el agente antineoplásico de uso clínico topotecán (Figura 13), bloquean específicamente el paso de religado que sigue al corte de una sola hebra de ADN efectuado por la topoisomerasa I (Top1) eucariótica, una enzima que cataliza la relajación del ADN superenrollado en las células. Sin embargo, a pesar de la existencia previa de una variedad de modelos propuestos (189) y de la disponibilidad de estructuras cristalográficas de rayos X de Top1 interactuando tanto covalentemente como no covalentemente con el ADN (190), solo se comprendieron plenamente los detalles atómicos precisos de la interacción entre CPT y Top1 cuando se cocristalizó el topotecán con el “complejo binario covalente” (191). Estos notables avances en biología estructural avivaron mi interés por esta diana anticancerosa y abrieron el camino a los estudios teóricos sobre un gran ensamblado proteína:ADN:ligando que podría simularse en presencia de moléculas de agua explícitas gracias a las mejoras concomitantes en la capacidad y potencia de las computadoras. Así, cuando Christian Bailly (a la sazón en el *Laboratoire de Pharmacologie Antitumorale du Centre Oscar Lambret*, en Lille, Francia) me propuso colaborar en la búsqueda de algunos agentes novedosos que también fueran capaces de convertir a Top1 en un veneno celular, tuve mucho interés en aceptar. Las moléculas pertenecían a la clase de alcaloides 6*H*-[1]benzopirano-[4,3;4,5]pirrolo[2,1-*a*]isoquinolina conocidos como lamelarinas, que se habían aislado originalmente de un molusco prosobranquio (*Lamellaria sp.*), pero que posteriormente se identificaron también en otros organismos marinos. Uno de los

compuestos más citotóxicos de esta serie resultó ser la lamelarina D (LMD) (Figura 13), algunos de cuyos derivados fueron seleccionados por PharmaMar para su desarrollo preclínico (192).

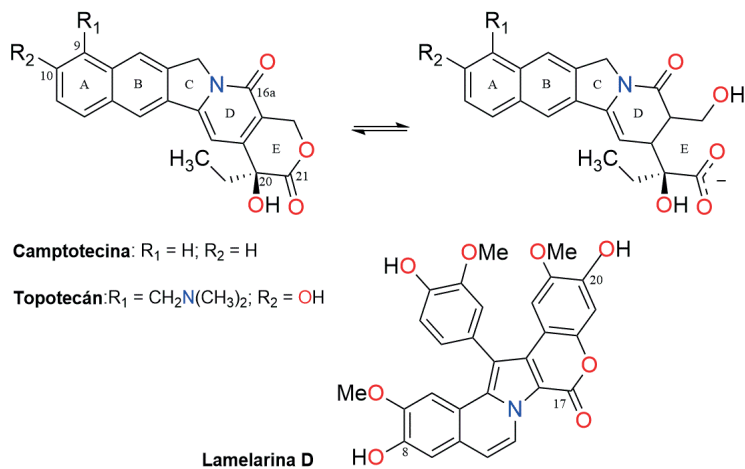


Figura 13. Estructuras químicas de camptotecina (con sus anillos etiquetados de la A a la E), topotecán y lamelarina D.

Usamos la estructura cristalográfica de la Top1 humana unida covalentemente a una molécula de ADN de doble cadena en complejo con topotecán, a una resolución de 2,10 Å (código PDB: 1K4T) (191), para modelar los complejos covalentes sin fármacos entre Top1 y dos 15-meros de ADN diferentes que contenían un sitio de intercalación CpG o TpG. Estos fueron los sitios de corte inducidos por LMD y/o CPT en los ensayos de escisión de ADN mediados por Top1 (193). Nuestro trabajo tenía como objetivo validar un modo de unión de LMD al complejo binario previamente propuesto (194) utilizando técnicas de acoplamiento automatizado y simulaciones de DM, para después poder delinear las relaciones estructura-actividad de las lamelarinas. A efectos de comparación, aplicamos el mismo protocolo de DM a un complejo similar que contenía CPT y del que no se disponía de información en el momento de iniciar nuestros estudios.

Para evitar cualquier tipo de sesgo, la LMD y la CPT se acoplaron en

el sitio de intercalación dentro del complejo ADN-Top1 utilizando el algoritmo genético Lamarckiano implementado en el programa Auto-Dock 3.0 (195). En la orientación mejor puntuada por el programa, el esqueleto del anillo de CPT era prácticamente superponible (*rmsd* de 0,3 Å) al del topotecán, tal y como se encuentra en las estructuras cristalográficas, y el 20-OH se encontraba interaccionando mediante enlace de hidrógeno con el carboxilo de la cadena lateral de Asp533. La orientación elegida para la LMD, como representativa del modo de unión de este tipo de fármacos al complejo ADN-Top1, presentaba (i) el anillo fenilo exocíclico en el surco mayor, (ii) los grupos hidroxilo C8 y C20 a distancias de enlace de hidrógeno de las cadenas laterales de Asn722 y Glu356, respectivamente, y (iii) el grupo ceto (O17) orientado hacia el grupo guanidinio de Arg364. Así, el C10 y el C16a de CPT parecen ser posicionalmente equivalentes al C20 y al C8 de LMD, a pesar de que los lados cóncavos y convexos de ambos fármacos se encuentran en surcos diferentes de la molécula de ADN. Este resultado sugiere marcadamente que los enfoques de acoplamiento y modelización basados simplemente en la similitud molecular pueden ser muy engañosos, especialmente en ausencia de información estructural 3D sobre el sitio de unión del receptor.

Curiosamente, cuando se publicaron las estructuras cristalográficas del complejo Top1:ADN humano unido a la forma lactónica de CPT y a miembros representativos de venenos de la Top1 de las clases indenoisoquinolina e indolocarbazol (196), la superposición de estos complejos ternarios reveló (i) la presencia común de un aceptor de enlaces de hidrógeno en el lado del surco menor de las moléculas del fármaco (en el caso de CPT, entre el átomo N1 de piridina y el Nε de Arg364), (ii) la presencia de sustituyentes en el lado del surco mayor, (iii) importantes contactos con los aminoácidos Asn352 y Glu356, y también (iv) que las interacciones de apilamiento de bases entre las moléculas del fármaco y el lado de la hebra cortada del dúplex de ADN no se conservan espacialmente. Comprobamos la viabilidad de las orientaciones de unión propuestas para CPT y LMD realizando simulaciones de DM en disolución acuosa de los complejos, los cuales mostraron un comportamiento notablemente estable durante todas las trayectorias de 1,5 ns, con los fármacos totalmente intercalados dando lugar a interacciones de apilamiento con las bases del ADN y a otras interacciones estabilizadoras tanto con la proteína como con

las moléculas de agua circundantes. Observamos que los enlaces de hidrógeno que implicaban a los hidroxilos en C8 y C20 de LMD se rompían y volvían a formar con frecuencia al intercambiarse con interacciones similares con las moléculas de agua vecinas. El análisis de la energía de unión del complejo Top1:ADN:CPT y de los dos complejos Top1:ADN:LMD (con el fármaco intercalado en TpG o en CpG) reveló un predominio de las interacciones de van der Waals que implican a las bases nitrogenadas que constituyen el sitio de intercalación y de las interacciones electrostáticas con algunos residuos proteicos cruciales (Arg488, Lys532, Asp533, Arg364, Glu356 y Asn722). En los tres casos estudiados, la interacción más favorable y cuantitativamente más importante fue con  $G_{+1}$ , aunque en el complejo CG-LMD este término de van der Waals era de magnitud similar al existente entre LMD y la guanina complementaria a  $C^{-1}$ . Sorprendentemente, la interacción con  $T_{-1}$  fue más favorable para CPT (en unas  $3 \text{ kcal} \times \text{mol}^{-1}$ ) que para LMD. Además, la ausencia de sitios  $C \downarrow G$  detectables inducidos por Top1 en presencia de CPT podría explicarse por el choque estérico que surgiría entre el grupo 4-amino exocíclico de  $C_{-1}$  y el carbono  $sp^3$  del anillo C. Por el contrario, tanto los sitios TpG como CpG proporcionaban sitios de unión igualmente buenos para LMD, en buena concordancia con los hallazgos experimentales.

Dado que nuestras simulaciones mostraban de forma consistente que el 20-OH de LMD participaba en un enlace de hidrógeno directo o mediado por agua con el carboxilato de Glu356, decidimos comprobar si la pérdida de esta interacción podría explicar el efecto deletéreo que la eliminación del grupo 20-OH en LMD tiene tanto en la inhibición de Top1 como en la citotoxicidad (193). Para ello, hicimos uso de un ciclo termodinámico (Figura 2) que nos permitió calcular la diferencia entre las energías libres de unión de estas dos moléculas a las mismas secuencias de ADN utilizando la información de mecánica estadística generada a partir de un conjunto de simulaciones de DM durante las cuales convertíamos este hidroxilo en un hidrógeno cuando el fármaco estaba libre en solución y también cuando formaba parte de cada uno de los dos complejos ternarios estudiados. Las diferencias de energía libre calculadas fueron de 2,6 y 2,9  $\text{kcal} \times \text{mol}^{-1}$  para los complejos CG y TG Top1-ADN, respectivamente, favoreciendo a la LMD frente al análogo sin el 20-OH, lo que explicaba satisfactoriamente la menor potencia detectada en los ensayos experimentales.

Este hallazgo dio credibilidad al modo de unión propuesto, que además fue apoyado por la observación de que la incorporación de un hidroxilo en la posición C10 de CPT (posicionalmente equivalente al 20-OH de LMD), como en topotecán (Figura 13) y otros análogos más simples, resulta en una notable mejora de su actividad inhibidora de Top1.

### **13. Observaciones finales**

Creo firmemente que las técnicas de modelización molecular y simulación por ordenador han alcanzado la mayoría de edad y han demostrado su valor en una enorme cantidad de áreas de investigación, incluido el estudio de las interacciones ADN-ligando y ADN-proteína. En este discurso he intentado ofrecer un resumen de mi actividad investigadora en este área concreta, que ha evolucionado de forma natural en consonancia con (i) la mayor disponibilidad de datos estructurales sobre macromoléculas biológicas, (ii) los avances en los campos de fuerzas (incluido el tratamiento mejorado de las interacciones electrostáticas de largo alcance) y (iii) los sorprendentes avances en las tecnologías informáticas, sobre todo en lo que respecta a la velocidad de procesamiento de datos, la programación para procesadores múltiples y GPUs, la capacidad de almacenamiento de datos y las prestaciones de los programas de gráficos moleculares interactivos. Me considero muy afortunado por el hecho de que Graham me ofreciera -como aprendiz de modelador molecular- estudiar la doble hélice del ADN, la cual -a pesar de su evidente simplicidad- sigue dando sorpresas y aguarda la plena comprensión de sus mensajes codificados en todos los organismos vivos. También fue una feliz coincidencia que el área de investigación en la que me involucré después de mi doctorado en cromatografía líquida -tras haberlo intentado infructuosamente durante varios años en farmacología experimental- se haya beneficiado continuamente desde entonces de una amplia gama de programas independientes, servidores web, conocimientos interdisciplinarios y una suma de experiencias aportadas por muchos científicos de talento de todo el mundo.

Durante varias décadas también he tenido el privilegio de ayudar a mis compañeros del Departamento de Química Orgánica y Farma-

céutica de la UAH en sus intentos de sintetizar nuevos compuestos con potencial actividad anticancerosa que interactúan con el ADN (197-200). Esta colaboración -al igual que otras con valiosos colegas y amigos del Instituto de Química Médica del CSIC en diferentes áreas- ha tenido una gran repercusión en mi carrera porque ha puesto de manifiesto los tremendos retos que conlleva el desarrollo de nuevos fármacos. Paralelamente, me ha maravillado la gran diversidad química y el tremendo potencial de los productos naturales, especialmente los procedentes del mar, que pueden dar lugar a fármacos muy prometedores con mecanismos de acción únicos. Si bien es verdad que hemos aprendido muchas lecciones de los estudios pasados, no es menos cierto que los descubrimientos de hoy plantean preguntas aún más desafiantes para mañana. Solo puedo confiar en que seguiremos formando nuevas generaciones de estudiantes bien preparados y entusiastas, capaces de responder a estas cuestiones pendientes y de pasar el testigo a los científicos del siguiente relevo generacional.

## **AGRADECIMIENTOS**

Para el trabajo que describo en esta memoria he contado con la ayuda de un grupo talentoso y dedicado de estudiantes, compañeros y colaboradores que han enriquecido todo este esfuerzo de investigación. A todos ellos les agradezco su entusiasmo y sus útiles debates a lo largo de los años. También agradezco la financiación recibida para conseguir estos resultados a la Comisión Interministerial de Ciencia y Tecnología del Gobierno de España, a la Comunidad Autónoma de Madrid, a la *National Foundation for Cancer Research* (EE.UU.) y a los Laboratorios PharmaMar S.A.U.

## Bibliografía

1. Hurley, L.H. and Boyd, F.L. (1988) DNA as a target for drug action. *Trends Pharmacol Sci*, **9**, 402-407.
2. Hurley, L.H. (1989) DNA and associated targets for drug design. *J Med Chem*, **32**, 2027-2033.
3. Hurley, L.H. (2002) DNA and its associated processes as targets for cancer therapy. *Nat Rev Cancer*, **2**, 188-200.
4. Watson, J.D. and Crick, F.H. (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, **171**, 737-738.
5. Saenger, W. (1984) *Principles of Nucleic Acid Structure*. Springer-Verlag, New York.
6. Hurley, L.H. (2001) Secondary DNA structures as molecular targets for cancer therapeutics. *Biochem Soc Trans*, **29**, 692-696.
7. Luscombe, N.M., Laskowski, R.A. and Thornton, J.M. (2001) Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res*, **29**, 2860-2874.
8. Suzuki, M., Brenner, S.E., Gerstein, M. and Yagi, N. (1995) DNA recognition code of transcription factors. *Protein Eng*, **8**, 319-328.
9. Goodsell, D.S. (2001) Sequence recognition of DNA by lexitropsins. *Curr Med Chem*, **8**, 509-516.
10. Yang, X.L. and Wang, A.H. (1999) Structural studies of atom-specific anticancer drugs acting on DNA. *Pharmacol Ther*, **83**, 181-215.
11. Gorin, A.A., Zhurkin, V.B. and Olson, W.K. (1995) B-DNA twisting correlates with base-pair morphology. *J Mol Biol*, **247**, 34-48.
12. Young, M.A., Ravishanker, G., Beveridge, D.L. and Berman, H.M. (1995) Analysis of local helix bending in crystal structures of DNA oligonucleotides and DNA-protein complexes. *Biophys J*, **68**, 2454-2468.



13. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. and Bourne, P.E. (2000) The Protein Data Bank. *Nucleic Acids Res*, **28**, 235-242.
14. Berman, H.M., Olson, W.K., Beveridge, D.L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S.H., Srinivasan, A.R. and Schneider, B. (1992) The Nucleic Acid Database. A comprehensive relational database of three-dimensional structures of nucleic acids. *Biophys J*, **63**, 751-759.
15. Sayle, R.A. and Milner-White, E.J. (1995) RASMOL: biomolecular graphics for all. *Trends Biochem Sci*, **20**, 374.
16. Liang, M.P., Banatao, D.R., Klein, T.E., Brutlag, D.L. and Altman, R.B. (2003) WebFEATURE: An interactive web tool for identifying and visualizing functional sites on macromolecular structures. *Nucleic Acids Res*, **31**, 3324-3327.
17. McMahon, B. and Hanson, R.M. (2008) A toolkit for publishing enhanced figures. *J Appl Crystallogr*, **41**, 811-814.
18. Hanson, R.M. (2010) Jmol– a paradigm shift in crystallographic visualization. *J Appl Crystallogr*, **43**, 1250-1260.
19. Sehnal, D., Bittrich, S., Deshpande, M., Svobodova, R., Berka, K., Bazgier, V., Velankar, S., Burley, S.K., Koca, J. and Rose, A.S. (2021) Mol\* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures. *Nucleic Acids Res*, **49**, W431-W437.
20. Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD: visual molecular dynamics. *J Mol Graph*, **14**, 33-38, 27-38.
21. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera a visualization system for exploratory research and analysis. *J Comput Chem*, **25**, 1605-1612.
22. DeLano, W.D. (2006) In LLC, D. S. (ed.). 0.99 rc6 ed, South San Francisco.
23. Phillips, D.C. (1981) In Sarma, R. H. (ed.). Adenine Press, Guilderland, New York, Vol. II, pp. 497.
24. Gago, F. (2004) Modelling and simulation: a computational perspective in anticancer drug discovery. *Curr Med Chem Anti-*

- cancer Agents*, **4**, 401-403.
25. Dror, R.O., Dirks, R.M., Grossman, J.P., Xu, H. and Shaw, D.E. (2012) Biomolecular simulation: a computational microscope for molecular biology. *Annu Rev Biophys*, **41**, 429-452.
  26. Karplus, M. (2002) Molecular dynamics simulations of biomolecules. *Acc Chem Res*, **35**, 321-323.
  27. Karplus, M. and McCammon, J.A. (2002) Molecular dynamics simulations of biomolecules. *Nat Struct Biol*, **9**, 646-652.
  28. Weiner, P.K. and Kollman, P.A. (1981) AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions. *J Comput Chem*, **2**, 287-303.
  29. Ma, J. (2005) Usefulness and limitations of normal mode analysis in modeling dynamics of biomolecular complexes. *Structure*, **13**, 373-380.
  30. Feynman, R.P., Leighton, R.B. and Sands, M. (1963) *The Feynman Lectures in Physics*. Addison-Wesley, Reading, MA.
  31. Cheatham, T.E., 3rd, Cieplak, P. and Kollman, P.A. (1999) A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J Biomol Struct Dyn*, **16**, 845-862.
  32. Pérez, A., Marchan, I., Svozil, D., Spöner, J., Cheatham, T.E., 3rd, Laughton, C.A. and Orozco, M. (2007) Refinement of the AMBER force field for nucleic acids: improving the description of a/g conformers. *Biophys J*, **92**, 3817-3829.
  33. Ivani, I., Dans, P.D., Noy, A., Perez, A., Faustino, I., Hospital, A., Walther, J., Andrio, P., Goni, R., Balaceanu, A. *et al.* (2016) Parmbsc1: a refined force field for DNA simulations. *Nat Methods*, **13**, 55-58.
  34. Cheatham, T.E., III, Miller, J.L., Fox, T., Darden, T.A. and Kollman, P.A. (2002) Molecular dynamics simulations on solvated biomolecular systems: The particle mesh Ewald method leads to stable trajectories of DNA, RNA, and proteins. *J Am Chem Soc*, **117**, 4193-4194.
  35. Salomon-Ferrer, R., Götz, A.W., Poole, D., Le Grand, S. and

- Walker, R.C. (2013) Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent Particle Mesh Ewald. *J Chem Theor Comput*, **9**, 3878-3888.
36. Kazlauskas, R. (2001) Modeling – A tool for experimentalists. *Science*, **293**, 2277-2279.
  37. Strahs, D. and Schlick, T. (2000) A-Ttract bending: insights into experimental structures by computational models. *J Mol Biol*, **301**, 643-663.
  38. García-Nieto, R., Manzanares, I., Cuevas, C. and Gago, F. (2000) Bending of DNA upon binding of ecteinascidin 743 and phthalascidin 650 studied by unrestrained molecular dynamics simulations. *J Am Chem Soc*, **122**, 7172-7182.
  39. Barsky, D., Foloppe, N., Ahmadi, S., Wilson, D.M., 3rd and MacKerell, A.D., Jr. (2000) New insights into the structure of abasic DNA from molecular dynamics simulations. *Nucleic Acids Res*, **28**, 2613-2626.
  40. Giudice, E., Varnai, P. and Lavery, R. (2003) Base pair opening within B-DNA: free energy pathways for GC and AT pairs from umbrella sampling simulations. *Nucleic Acids Res*, **31**, 1434-1443.
  41. Spackova, N., Cubero, E., Sponer, J. and Orozco, M. (2004) Theoretical study of the guanine to 6-thioguanine substitution in duplexes, triplexes, and tetraplexes. *J Am Chem Soc*, **126**, 14642-14650.
  42. Varnai, P. and Zakrzewska, K. (2004) DNA and its counterions: a molecular dynamics study. *Nucleic Acids Res*, **32**, 4269-4280.
  43. Beveridge, D.L., Barreiro, G., Byun, K.S., Case, D.A., Cheatham, T.E., 3rd, Dixit, S.B., Giudice, E., Lankas, F., Lavery, R., Maddocks, J.H. *et al.* (2004) Molecular dynamics simulations of the 136 unique tetranucleotide sequences of DNA oligonucleotides. I. Research design and results on d(CpG) steps. *Biophys J*, **87**, 3799-3813.
  44. Ponomarev, S.Y., Thayer, K.M. and Beveridge, D.L. (2004) Ion motions in molecular dynamics simulations on DNA. *Proc Natl Acad Sci U S A*, **101**, 14771-14775.

45. Banavali, N.K. and Roux, B. (2005) Free energy landscape of A-DNA to B-DNA conversion in aqueous solution. *J Am Chem Soc*, **127**, 6866-6876.
46. Pasi, M., Maddocks, J.H., Beveridge, D., Bishop, T.C., Case, D.A., Cheatham, T., 3rd, Dans, P.D., Jayaram, B., Lankas, F., Laughton, C. *et al.* (2014) mABC: a systematic microsecond molecular dynamics study of tetranucleotide sequence effects in B-DNA. *Nucleic Acids Res*, **42**, 12272-12283.
47. Schlick, T., Beard, D.A., Jing, H., Strahs, D.A. and Xiaoliang, Q. (2000) Computational challenges in simulating large DNA over long times. *Computing in Science & Engineering*, **2**, 38-51.
48. Cheatham, T.E., 3rd. (2004) Simulation and modeling of nucleic acid structure, dynamics and interactions. *Curr Opin Struct Biol*, **14**, 360-367.
49. Lavery, R., Maddocks, J.H., Pasi, M. and Zakrzewska, K. (2014) Analyzing ion distributions around DNA. *Nucleic Acids Res*, **42**, 8138-8149.
50. Pasi, M., Maddocks, J.H. and Lavery, R. (2015) Analyzing ion distributions around DNA: sequence-dependence of potassium ion distributions from microsecond molecular dynamics. *Nucleic Acids Res*, **43**, 2412-2423.
51. Dixit, S.B., Beveridge, D.L., Case, D.A., Cheatham, T.E., 3rd, Giudice, E., Lankas, F., Lavery, R., Maddocks, J.H., Osman, R., Sklenar, H. *et al.* (2005) Molecular dynamics simulations of the 136 unique tetranucleotide sequences of DNA oligonucleotides. II: sequence context effects on the dynamical structures of the 10 unique dinucleotide steps. *Biophys J*, **89**, 3721-3740.
52. McCammon, J.A., Gelin, B.R. and Karplus, M. (1977) Dynamics of folded proteins. *Nature*, **267**, 585-590.
53. Minhas, V., Sun, T., Mirzoev, A., Korolev, N., Lyubartsev, A.P. and Nordenskiöld, L. (2020) Modeling DNA flexibility: comparison of force fields from atomistic to multiscale levels. *J Phys Chem B*, **124**, 38-49.
54. MacKerell, A.D. and Banavali, N.K. (2000) All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution. *J Comput.*

- Chem*, **21**, 105-120.
55. Feig, M. and Pettitt, B.M. (1999) Sodium and chlorine ions as part of the DNA solvation shell. *Biophys J*, **77**, 1769-1781.
  56. Cheatham, T.E., 3rd and Kollman, P.A. (2000) Molecular dynamics simulation of nucleic acids. *Annu Rev Phys Chem*, **51**, 435-471.
  57. Darden, T., York, D. and Pedersen, L. (1993) Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*, **98**, 10089-10092.
  58. Giudice, E. and Lavery, R. (2002) Simulations of nucleic acids and their complexes. *Acc Chem Res*, **35**, 350-357.
  59. Duan, Y. and Kollman, P.A. (1998) Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science*, **282**, 740-744.
  60. Dror, R.O., Arlow, D.H., Borhani, D.W., Jensen, M.O., Piana, S. and Shaw, D.E. (2009) Identification of two distinct inactive conformations of the b2-adrenergic receptor reconciles structural and biochemical observations. *Proc Natl Acad Sci U S A*, **106**, 4689-4694.
  61. Bashford, D. and Case, D.A. (2000) Generalized born models of macromolecular solvation effects. *Annu Rev Phys Chem*, **51**, 129-152.
  62. Onufriev, A., Bashford, D. and Case, D.A. (2000) Modification of the Generalized Born Model suitable for macromolecules. *J Phys Chem B*, **104**, 3712-3720.
  63. Jorgensen, W., Chandrasekhar, J., Madura, J., Impey, R. and Klein, M. (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys*, **79**, 926-935.
  64. Snodin, B.E.K., Schreck, J.S., Romano, F., Louis, A.A. and Doye, J.P.K. (2019) Coarse-grained modelling of the structural properties of DNA origami. *Nucleic Acids Res*.
  65. Tozzini, V. (2005) Coarse-grained models for proteins. *Curr Opin Struct Biol*, **15**, 144-150.
  66. Young, M.A., Gonfloni, S., Superti-Furga, G., Roux, B. and Kuriyan, J. (2001) Dynamic coupling between the SH2 and

- SH3 domains of c-Src and Hck underlies their inactivation by C-terminal tyrosine phosphorylation. *Cell*, **105**, 115-126.
67. Truhlar, D.G., Gao, J., Alhambra, C., Garcia-Viloca, M., Corchado, J., Sanchez, M.L. and Villa, J. (2002) The incorporation of quantum effects in enzyme kinetics modeling. *Acc Chem Res*, **35**, 341-349.
  68. Mendieta, J., Martin-Santamaria, S., Priego, E.M., Balzarini, J., Camarasa, M.J., Perez-Perez, M.J. and Gago, F. (2004) Role of histidine-85 in the catalytic mechanism of thymidine phosphorylase as assessed by targeted molecular dynamics simulations and quantum mechanical calculations. *Biochemistry*, **43**, 405-414.
  69. Isralewitz, B., Gao, M. and Schulten, K. (2001) Steered molecular dynamics and mechanical functions of proteins. *Curr Opin Struct Biol*, **11**, 224-230.
  70. Ma, J. and Karplus, M. (1997) Molecular switch in signal transduction: reaction paths of the conformational changes in ras p21. *Proc Natl Acad Sci U S A*, **94**, 11905-11910.
  71. Rodríguez-Barríos, F., Balzarini, J. and Gago, F. (2005) The molecular basis of resilience to the effect of the Lys103Asn mutation in non-nucleoside HIV-1 reverse transcriptase inhibitors studied by targeted molecular dynamics simulations. *J Am Chem Soc*, **127**, 7570-7578.
  72. Marco, E., García-Nieto, R. and Gago, F. (2003) Assessment by molecular dynamics simulations of the structural determinants of DNA-binding specificity for transcription factor Sp1. *J Mol Biol*, **328**, 9-32.
  73. Zaman, G.J., Michiels, P.J. and van Boeckel, C.A. (2003) Targeting RNA: new opportunities to address drugless targets. *Drug Discov Today*, **8**, 297-306.
  74. Hunter, T. (2000) Signaling—2000 and beyond. *Cell*, **100**, 113-127.
  75. Hubbard, S.R. (2002) Protein tyrosine kinases: autoregulation and small-molecule inhibition. *Curr Opin Struct Biol*, **12**, 735-741.

76. Klumpp, S. and Krieglstein, J. (2002) Serine/threonine protein phosphatases in apoptosis. *Curr Opin Pharmacol*, **2**, 458-462.
77. Gago, F. (2012) Reflections on the past 25 years. *J Comput Aided Mol Des*, **26**, 45.
78. Chien, A., Foster, I. and Goddette, D. (2002) Grid technologies empowering drug discovery. *Drug Discov Today*, **7**, S176-180.
79. Folding@home project. <https://foldingathome.org/>.
80. Davies, E.K. and Richards, W.G. (2002) The potential of Internet computing for drug discovery. *Drug Discov Today*, **7**, S99-103.
81. Richards, W.G. (1983) *Quantum Pharmacology*. 2nd ed. Butterworth-Heinemann.
82. Martin, Y.C. (2011) Remembrances of Corwin Hansch. *J Comput Aided Mol Des*, **25**, 519-523.
83. Baguley, B.C. (1982) Nonintercalative DNA-binding antitumour compounds. *Mol Cell Biochem*, **43**, 167-181.
84. Marky, L.A. and Breslauer, K.J. (1987) Origins of netropsin binding affinity and specificity: correlations of thermodynamic and structural data. *Proc Natl Acad Sci U S A*, **84**, 4359-4363.
85. Gago, F., Reynolds, C.A. and Richards, W.G. (1989) The binding of nonintercalative drugs to alternating DNA sequences. *Mol Pharmacol*, **35**, 232-241.
86. Dervan, P.B. and Edelson, B.S. (2003) Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr Opin Struct Biol*, **13**, 284-299.
87. Buchmueller, K.L., Staples, A.M., Howard, C.M., Horick, S.M., Uthe, P.B., Le, N.M., Cox, K.K., Nguyen, B., Pacheco, K.A., Wilson, W.D. *et al.* (2005) Extending the language of DNA molecular recognition by polyamides: unexpected influence of imidazole and pyrrole arrangement on binding affinity and specificity. *J Am Chem Soc*, **127**, 742-750.
88. Dervan, P.B. (2001) Molecular recognition of DNA by small molecules. *Bioorg Med Chem*, **9**, 2215-2235.
89. Edayathumangalam, R.S., Weyermann, P., Gottesfeld, J.M.,

- Dervan, P.B. and Luger, K. (2004) Molecular recognition of the nucleosomal “super groove”. *Proc Natl Acad Sci U S A*, **101**, 6864-6869.
90. Straatsma, T.P. and McCammon, J.A. (1992) Computational alchemy. *Annu Rev Phys Chem*, **43**, 407-435.
91. Kaus, J.W., Pierce, L.T., Walker, R.C. and McCammon, J.A. (2013) Improving the efficiency of free energy calculations in the AMBER molecular dynamics package. *J Chem Theor Comput*, **9**.
92. Lee, T.S., Allen, B.K., Giese, T.J., Guo, Z., Li, P., Lin, C., McGee, T.D., Jr., Pearlman, D.A., Radak, B.K., Tao, Y. *et al.* (2020) Alchemical binding free energy calculations in AMBER20: Advances and best practices for drug discovery. *J Chem Inf Model*, **60**, 5595-5623.
93. Kollman, P.A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W. *et al.* (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc Chem Res*, **33**, 889-897.
94. Kollman, P. (2002) Free energy calculations: Applications to chemical and biochemical phenomena. *Chem Rev*, **93**, 2395-2417.
95. Rodinger, T. and Pomes, R. (2005) Enhancing the accuracy, the efficiency and the scope of free energy simulations. *Curr Opin Struct Biol*, **15**, 164-170.
96. Reynolds, C.A., King, P.M. and Richards, W.G. (1988) Computed redox potentials and the design of bioreductive agents. *Nature*, **334**, 80-82.
97. Gago, F. and Richards, W.G. (1990) Netropsin binding to poly[d(IC)]·poly[d(IC)] and poly[d(GC)]·poly[d(GC)]: a computer simulation. *Mol Pharmacol*, **37**, 341-346.
98. Pullman, B.; Jortner, J., eds. (1990) *Molecular Basis of Specificity in Nucleic Acid-Drug Interactions*. Kluwer Academic Publishers, Dordrecht.
99. Waring, M.J. (1990) In Pullman, B. J., J. (ed.), *Molecular Basis*



*of Specificity in Nucleic Acid-Drug Interactions*. Kluwer Academic Publishers, Dordrecht, pp. 225–245.

100. Hoogsteen, K. (1959) The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine. *Acta Crystallographica*, **12**, 822-823.
101. Wang, A.H., Ughetto, G., Quigley, G.J., Hakoshima, T., van der Marel, G.A., van Boom, J.H. and Rich, A. (1984) The molecular structure of a DNA-triostin A complex. *Science*, **225**, 1115-1121.
102. Singh, U.C., Pattabiraman, N., Langridge, R. and Kollman, P.A. (1986) Molecular mechanical studies of d(CGTACG)<sub>2</sub>: complex of triostin A with the middle A-T base pairs in either Hoogsteen or Watson-Crick pairing. *Proc Natl Acad Sci U S A*, **83**, 6402-6406.
103. Gao, X.L. and Patel, D.J. (1988) NMR studies of echinomycin bisintercalation complexes with d(A1-C2-G3-T4) and d(T1-C2-G3-A4) duplexes in aqueous solution: sequence-dependent formation of Hoogsteen A1.T4 and Watson--Crick T1.A4 base pairs flanking the bisintercalation site. *Biochemistry*, **27**, 1744-1751.
104. Gallego, J., Ortiz, A.R. and Gago, F. (1993) A molecular dynamics study of the bis-intercalation complexes of echinomycin with d(ACGT)<sub>2</sub> and d(TCGA)<sub>2</sub>: rationale for sequence-specific Hoogsteen base pairing. *J Med Chem*, **36**, 1548-1561.
105. Wang, A.H., Ughetto, G., Quigley, G.J. and Rich, A. (1986) Interactions of quinoxaline antibiotic and DNA: the molecular structure of a triostin A-d(GCGTACGC) complex. *J Biomol Struct Dyn*, **4**, 319-342.
106. Rein, R. (1978) In Pullman, B. (ed.), *Perspectives in Quantum Chemistry and Biochemistry*. John Wiley & Sons, New York, Vol. II, pp. 307–362.
107. Cuesta-Seijo, J.A. and Sheldrick, G.M. (2005) Structures of complexes between echinomycin and duplex DNA. *Acta Crystallogr D Biol Crystallogr*, **61**, 442-448.
108. Gao, X.L. and Patel, D.J. (1989) Antitumour drug-DNA interactions: NMR studies of echinomycin and chromomycin com-

- plexes. *Q Rev Biophys*, **22**, 93-138.
109. Gallego, J., Luque, F.J., Orozco, M. and Gago, F. (1994) Binding of echinomycin to d(GCGC)<sub>2</sub> and d(CCGG)<sub>2</sub>: distinct stacking interactions dictate the sequence-dependent formation of Hoogsteen base pairs. *J Biomol Struct Dyn*, **12**, 111-129.
  110. Gallego, J., Luque, F.J., Orozco, M., Burgos, C., Alvarez-Builla, J., Rodrigo, M.M. and Gago, F. (1994) DNA sequence-specific reading by echinomycin: role of hydrogen bonding and stacking interactions. *J Med Chem*, **37**, 1602-1609.
  111. Aida, M. and Nagata, C. (1986) An *ab initio* molecular orbital study on the stacking interaction between nucleic acid bases: Dependence on the sequence and relation to the conformation. *Int. J. Quantum Chem*, **29**, 1253-1261.
  112. Ornstein, R.L., Rein, R., Breen, D.L. and Macelroy, R.D. (1978) An optimized potential function for the calculation of nucleic acid interaction energies I. Base stacking. *Biopolymers*, **17**, 2341-2360.
  113. Marchand, C., Bailly, C., McLean, M.J., Moroney, S.E. and Waring, M.J. (1992) The 2-amino group of guanine is absolutely required for specific binding of the anti-cancer antibiotic echinomycin to DNA. *Nucleic Acids Res*, **20**, 5601-5606.
  114. Bailly, C., Marchand, C. and Waring, M.J. (2002) New binding sites for antitumor antibiotics created by relocating the purine 2-amino group in DNA. *J Am. Chem Soc*, **115**, 3784-3785.
  115. Bailly, C., Echepare, S., Gago, F. and Waring, M.J. (1999) Recognition elements that determine affinity and sequence-specific binding to DNA of 2QN, a biosynthetic bis-quinoline analogue of echinomycin. *Anticancer Drug Des*, **14**, 291-303.
  116. Alfredson, T.V., Maki, A.H. and Waring, M.J. (1991) Optically detected triplet-state magnetic resonance studies of the DNA complexes of the bisquinoline analogue of echinomycin. *Biochemistry*, **30**, 9665-9675.
  117. Bailly, C. and Waring, M.J. (1998) DNA recognition by quinoxaline antibiotics: use of base-modified DNA molecules to investigate determinants of sequence-specific binding of triostin A and TANDEM. *Biochem J*, **330** ( Pt 1), 81-87.

118. Address, K.J. and Feigon, J. (1994) Sequence specificity of quinoxaline antibiotics. 2. NMR studies of the binding of [*N*-MeCys3,*N*-MeCys7]TANDEM and triostin A to DNA containing a CpI step. *Biochemistry*, **33**, 12397-12404.
119. Lavesa, M. and Fox, K.R. (2001) Preferred binding sites for [*N*-MeCys3, *N*-MeCys7]TANDEM determined using a universal footprinting substrate. *Anal Biochem*, **293**, 246-250.
120. Checa, A., Ortiz, A.R., de Pascual-Teresa, B. and Gago, F. (1997) Assessment of solvation effects on calculated binding affinity differences: trypsin inhibition by flavonoids as a model system for congeneric series. *J Med Chem*, **40**, 4136-4145.
121. Pérez, 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.
122. Honig, B. and Nicholls, A. (1995) Classical electrostatics in biology and chemistry. *Science*, **268**, 1144-1149.
123. Olson, W.K., Gorin, A.A., Lu, X.J., Hock, L.M. and Zhurkin, V.B. (1998) DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc Natl Acad Sci U S A*, **95**, 11163-11168.
124. Marco, E., Negri, A., Luque, F.J. and Gago, F. (2005) Role of stacking interactions in the binding sequence preferences of DNA bis-intercalators: insight from thermodynamic integration free energy simulations. *Nucleic Acids Res*, **33**, 6214-6224.
125. Negri, A., Marco, E., García-Hernández, V., Domingo, A., Llamas-Saiz, A.L., Porto-Sanda, S., Riguera, R., Laine, W., David-Cordonnier, M.H., Bailly, C. *et al.* (2007) Antitumor activity, X-ray crystal structure, and DNA binding properties of thio-coraline A, a natural bisintercalating thiodepsipeptide. *J Med Chem*, **50**, 3322-3333.
126. Pelaprat, D., Delbarre, A., Le Guen, I., Roques, B.P. and Le Pecq, J.B. (1980) DNA intercalating compounds as potential antitumor agents. 2. Preparation and properties of 7*H*-pyrido-carbazole dimers. *J Med Chem*, **23**, 1336-1343.
127. Delbarre, A., Delepierre, M., Garbay, C., Igolen, J., Le Pecq,

- J.B. and Roques, B.P. (1987) Geometry of the antitumor drug ditercalinium bisintercalated into d(CpGpCpG)<sub>2</sub> by <sup>1</sup>H NMR. *Proc Natl Acad Sci U S A*, **84**, 2155-2159.
128. Gao, Q., Williams, L.D., Egli, M., Rabinovich, D., Chen, S.L., Quigley, G.J. and Rich, A. (1991) Drug-induced DNA repair: X-ray structure of a DNA-ditercalinium complex. *Proc Natl Acad Sci U S A*, **88**, 2422-2426.
129. Delepierre, M., Milhe, C., Namane, A., Dinh, T.H. and Roques, B.P. (1991) <sup>1</sup>H- and <sup>31</sup>P-NMR studies of ditercalinium binding to a d(GCGC)<sub>2</sub> and d(CCTATAGG)<sub>2</sub> minihelices: a sequence specificity study. *Biopolymers*, **31**, 331-353.
130. Spielmann, H.P., Wemmer, D.E. and Jacobsen, J.P. (1995) Solution structure of a DNA complex with the fluorescent bis-intercalator TOTO determined by NMR spectroscopy. *Biochemistry*, **34**, 8542-8553.
131. Peek, M.E., Lipscomb, L.A., Bertrand, J.A., Gao, Q., Roques, B.P., Garbay-Jaureguiberry, C. and Williams, L.D. (1994) DNA distortion in bis-intercalated complexes. *Biochemistry*, **33**, 3794-3800.
132. de Pascual-Teresa, B., Gallego, J., Ortiz, A.R. and Gago, F. (1996) Molecular dynamics simulations of the bis-intercalated complexes of ditercalinium and Flexi-Di with the hexanucleotide d(GCGCGC)<sub>2</sub>: theoretical analysis of the interaction and rationale for the sequence binding specificity. *J Med Chem*, **39**, 4810-4824.
133. Gallego, J., de Pascual-Teresa, B., Ortiz, A.R., Pisabarro, M.T. and Gago, F. (1995) In Sanz, F., Giraldo, J. and Manaut, F. (eds.), *QSAR and Molecular Modelling: Concepts, Computational Tools and Biological Applications*. Barcelona, J. R. Prous, pp. 274-281.
134. Sobell, H.M., Jain, S.C., Sakore, T.D. and Nordman, C.E. (1971) Stereochemistry of actinomycin-DNA binding. *Nat New Biol*, **231**, 200-205.
135. Kamitori, S. and Takusagawa, F. (1992) Crystal structure of the 2:1 complex between d(GAAGCTTC) and the anticancer drug actinomycin D. *J Mol Biol*, **225**, 445-456.

136. Takusagawa, F., Dabrow, M., Neidle, S. and Berman, H.M. (1982) The structure of a pseudo intercalated complex between actinomycin and the DNA binding sequence d(GpC). *Nature*, **296**, 466-469.
137. Kamitori, S. and Takusagawa, F. (2002) Multiple binding modes of anticancer drug actinomycin D: X-ray, molecular modeling, and spectroscopic studies of d(GAAGCTTC)<sub>2</sub>-actinomycin D complexes and its host DNA. *J Am Chem Soc*, **116**, 4154-4165.
138. Lybrand, T.P., Brown, S.C., Creighton, S., Shafer, R.H. and Kollman, P.A. (1986) Computer modeling of actinomycin D interactions with double-helical DNA. *J Mol Biol*, **191**, 495-507.
139. Müller, W. and Crothers, D.M. (1968) Studies of the binding of actinomycin and related compounds to DNA. *J Mol Biol*, **35**, 251-290.
140. Jennewein, S. and Waring, M.J. (1997) Footprinting of echinomycin and actinomycin D on DNA molecules asymmetrically substituted with inosine and/or 2,6-diaminopurine. *Nucleic Acids Res*, **25**, 1502-1510.
141. Gallego, J., Ortiz, A.R., de Pascual-Teresa, B. and Gago, F. (1997) Structure-affinity relationships for the binding of actinomycin D to DNA. *J Comput Aided Mol Des*, **11**, 114-128.
142. Williams, J.H. (2002) The molecular electric quadrupole moment and solid-state architecture. *Acc Chem Res*, **26**, 593-598.
143. Gago, F. (1998) Stacking interactions and intercalative DNA binding. *Methods*, **14**, 277-292.
144. Rinehart, K.L. (2000) Antitumor compounds from tunicates. *Med Res Rev*, **20**, 1-27.
145. Manzanares, I., Cuevas, C., Garcia-Nieto, R., Marco, E. and Gago, F. (2001) Advances in the chemistry and pharmacology of ecteinascidins, a promising new class of anti-cancer agents. *Curr Med Chem Anticancer Agents*, **1**, 257-276.
146. Pommier, Y., Kohlhagen, G., Bailly, C., Waring, M., Mazumder, A. and Kohn, K.W. (1996) DNA sequence- and structure-selective alkylation of guanine N2 in the DNA minor groove

- by ecteinascidin 743, a potent antitumor compound from the Caribbean tunicate *Ecteinascidia turbinata*. *Biochemistry*, **35**, 13303-13309.
147. Gago, F. and Hurley, L.H. (2002) In Demeunynck, M., Bailly, C., Wilson, W. D. (ed.), *Small Molecule DNA and RNA Binders: From Synthesis to Nucleic Acid Complexes*. Wiley-VCH: Weinheim, Germany, pp. 643-675.
  148. Zewail-Foote, M. and Hurley, L.H. (1999) Ecteinascidin 743: a minor groove alkylator that bends DNA toward the major groove. *J Med Chem*, **42**, 2493-2497.
  149. Ravishanker, G., Swaminathan, S., Beveridge, D.L., Lavery, R. and Sklenar, H. (1989) Conformational and helicoidal analysis of 30 ps of molecular dynamics on the d(CGCGAATTCGCG) double helix: "curves", dials and windows. *J Biomol Struct Dyn*, **6**, 669-699.
  150. Lavery, R. and Sklenar, H. (1988) The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J Biomol Struct Dyn*, **6**, 63-91.
  151. Blanchet, C., Pasi, M., Zakrzewska, K. and Lavery, R. (2011) CURVES+ web server for analyzing and visualizing the helical, backbone and groove parameters of nucleic acid structures. *Nucleic Acids Res*, **39**, W68-73.
  152. García-Nieto, R., Manzanares, I., Cuevas, C. and Gago, F. (2000) Increased DNA binding specificity for antitumor ecteinascidin 743 through protein-DNA interactions? *J Med Chem*, **43**, 4367-4369.
  153. Marco, E., Garcia-Nieto, R., Mendieta, J., Manzanares, I., Cuevas, C. and Gago, F. (2002) A 3·(ET743)-DNA complex that both resembles an RNA-DNA hybrid and mimicks zinc finger-induced DNA structural distortions. *J Med Chem*, **45**, 871-880.
  154. Arnott, S. and Hukins, D.W. (1972) Optimised parameters for A-DNA and B-DNA. *Biochem Biophys Res Commun*, **47**, 1504-1509.
  155. Ng, H.L., Kopka, M.L. and Dickerson, R.E. (2000) The structure of a stable intermediate in the A <-> B DNA helix transition. *Proc Natl Acad Sci U S A*, **97**, 2035-2039.

156. Paull, K.D., Shoemaker, R.H., Hodes, L., Monks, A., Scudiero, D.A., Rubinstein, L., Plowman, J. and Boyd, M.R. (1989) Display and analysis of patterns of differential activity of drugs against human tumor cell lines: development of mean graph and COMPARE algorithm. *J Natl Cancer Inst*, **81**, 1088-1092.
157. Marco, E. and Gago, F. (2005) DNA structural similarity in the 2:1 complexes of the antitumor drugs trabectedin (Yondelis) and chromomycin A<sub>3</sub> with an oligonucleotide sequence containing two adjacent TGG binding sites on opposing strands. *Mol Pharmacol*, **68**, 1559-1567.
158. Hou, M.H., Robinson, H., Gao, Y.G. and Wang, A.H. (2004) Crystal structure of the [Mg<sup>2+</sup>-(chromomycin A<sub>3</sub>)<sub>2</sub>]-d(TTGGCCAA)<sub>2</sub> complex reveals GGCC binding specificity of the drug dimer chelated by a metal ion. *Nucleic Acids Res*, **32**, 2214-2222.
159. Grzeskowiak, K., Goodsell, D.S., Kaczor-Grzeskowiak, M., Cascio, D. and Dickerson, R.E. (1993) Crystallographic analysis of C-C-A-A-G-C-T-T-G-G and its implications for bending in B-DNA. *Biochemistry*, **32**, 8923-8931.
160. Dornberger, U., Flemming, J. and Fritzsche, H. (1998) Structure determination and analysis of helix parameters in the DNA decamer d(CATGGCCATG)<sub>2</sub> comparison of results from NMR and crystallography. *J Mol Biol*, **284**, 1453-1463.
161. Goodsell, D.S., Kopka, M.L., Cascio, D. and Dickerson, R.E. (1993) Crystal structure of CATGGCCATG and its implications for A-tract bending models. *Proc Natl Acad Sci U S A*, **90**, 2930-2934.
162. Dickerson, R.E., Goodsell, D.S. and Neidle, S. (1994) "...the tyranny of the lattice...". *Proc Natl Acad Sci U S A*, **91**, 3579-3583.
163. Takebayashi, Y., Pourquier, P., Zimonjic, D.B., Nakayama, K., Emmert, S., Ueda, T., Urasaki, Y., Kanzaki, A., Akiyama, S.I., Popescu, N. *et al.* (2001) Antiproliferative activity of ecteinascidin 743 is dependent upon transcription-coupled nucleotide-excision repair. *Nat Med*, **7**, 961-966.
164. Larsen, A.K., Galmarini, C.M. and D'Incalci, M. (2016) Uni-



- que features of trabectedin mechanism of action. *Cancer Chemother Pharmacol*, **77**, 663-671.
165. Feuerhahn, S., Giraudon, C., Martinez-Diez, M., Bueren-Calabuig, J.A., Galmarini, C.M., Gago, F. and Egly, J.M. (2011) XPF-dependent DNA breaks and RNA polymerase II arrest induced by antitumor DNA interstrand crosslinking-mimetic alkaloids. *Chem Biol*, **18**, 988-999.
  166. Herrero, A.B., Martin-Castellanos, C., Marco, E., Gago, F. and Moreno, S. (2006) Cross-talk between nucleotide excision and homologous recombination DNA repair pathways in the mechanism of action of antitumor trabectedin. *Cancer Res*, **66**, 8155-8162.
  167. Marco, E., David-Cordonnier, M.H., Bailly, C., Cuevas, C. and Gago, F. (2006) Further insight into the DNA recognition mechanism of trabectedin from the differential affinity of its demethylated analogue ecteinascidin ET729 for the triplet DNA binding site CGA. *J Med Chem*, **49**, 6925-6929.
  168. Leal, J.F., Garcia-Hernandez, V., Moneo, V., Domingo, A., Bueren-Calabuig, J.A., Negri, A., Gago, F., Guillen-Navarro, M.J., Aviles, P., Cuevas, C. *et al.* (2009) Molecular pharmacology and antitumor activity of Zalypsis in several human cancer cell lines. *Biochem Pharmacol*, **78**, 162-170.
  169. Leal, J.F., Martinez-Diez, M., Garcia-Hernandez, V., Moneo, V., Domingo, A., Bueren-Calabuig, J.A., Negri, A., Gago, F., Guillen-Navarro, M.J., Aviles, P. *et al.* (2010) PM01183, a new DNA minor groove covalent binder with potent in vitro and in vivo anti-tumour activity. *Br J Pharmacol*, **161**, 1099-1110.
  170. Wong, K.Y. and Pettitt, B.M. (2008) The pathway of oligomeric DNA melting investigated by molecular dynamics simulations. *Biophys J*, **95**, 5618-5626.
  171. 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.
  172. Martinez, S., Perez, L., Galmarini, C.M., Aracil, M., Tercero,



- J.C., Gago, F., Albella, B. and Bueren, J.A. (2013) Inhibitory effects of marine-derived DNA-binding anti-tumour tetrahydroisoquinolines on the Fanconi anaemia pathway. *Br J Pharmacol*, **170**, 871-882.
173. Mirchandani, K.D. and D'Andrea, A.D. (2006) The Fanconi anemia/BRCA pathway: a coordinator of cross-link repair. *Exp Cell Res*, **312**, 2647-2653.
174. Bueren-Calabuig, J.A., Negri, A., Morreale, A. and Gago, F. (2012) Rationale for the opposite stereochemistry of the major monoadducts and interstrand crosslinks formed by mitomycin C and its decarbamoylated analogue at CpG steps in DNA and the effect of cytosine modification on reactivity. *Org Biomol Chem*, **10**, 1543-1552.
175. Kollman, P.A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W. *et al.* (2000) Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Acc Chem Res*, **33**, 889-897.
176. Johnson, W.S., He, Q.Y. and Tomasz, M. (1995) Selective recognition of the m<sup>5</sup>CpG dinucleotide sequence in DNA by mitomycin C for alkylation and cross-linking. *Bioorg Med Chem*, **3**, 851-860.
177. Tomasz, M., Lipman, R., Chowdary, D., Pawlak, J., Verdine, G.L. and Nakanishi, K. (1987) Isolation and structure of a covalent cross-link adduct between mitomycin C and DNA. *Science*, **235**, 1204-1208.
178. Paz, M.M., Ladwa, S., Champeil, E., Liu, Y., Rockwell, S., Boamah, E.K., Bargonetti, J., Callahan, J., Roach, J. and Tomasz, M. (2008) Mapping DNA adducts of mitomycin C and decarbamoyl mitomycin C in cell lines using liquid chromatography/ electrospray tandem mass spectrometry. *Chem Res Toxicol*, **21**, 2370-2378.
179. Hagerman, P.J. (1986) Sequence-directed curvature of DNA. *Nature*, **321**, 449-450.
180. Vafabakhsh, R. and Ha, T. (2012) Extreme bendability of DNA less than 100 base pairs long revealed by single-molecule cycli-

- zation. *Science*, **337**, 1097-1101.
181. Mills, A. and Gago, F. (2020) Atomistic insight into sequence-directed DNA bending and minicircle formation propensity in the absence and presence of phased A-tracts. *J Comput Aided Mol Des*, **34**, 253-265.
  182. Crothers, D.M. (1998) DNA curvature and deformation in protein-DNA complexes: a step in the right direction. *Proc Natl Acad Sci U S A*, **95**, 15163-15165.
  183. Mills, A. and Gago, F. (2021) Insight into the sequence-specific elements leading to increased DNA bending and ligase-mediated circularization propensity by antitumor trabectedin *J Comput Aided Mol Des*, **35**, 707-719.
  184. Lee, M.D., Ellestad, G.A. and Borders, D.B. (1991) Calicheamicins: discovery, structure, chemistry, and interaction with DNA. *Acc Chem Res*, **24**, 235-243.
  185. Nicolaou, K.C. and Dai, W.M. (1991) Chemistry and biology of the enediyne anticancer antibiotics. *Angew Chem Int Ed Engl*, **30**, 1387-1416.
  186. Salzberg, A.A. and Dedon, P.C. (1997) An improved method for the rapid assessment of DNA bending by small molecules. *J Biomol Struct Dyn*, **15**, 277-284.
  187. Manning, G.S. (2006) The persistence length of DNA is reached from the persistence length of its null isomer through an internal electrostatic stretching force. *Biophys J*, **91**, 3607-3616.
  188. Mills, A. and Gago, F. (2021) Structural and mechanistic insight into DNA bending by antitumour calicheamicins. *Org Biomol Chem*, **19**, 6707-6717.
  189. Thomas, C.J., Rahier, N.J. and Hecht, S.M. (2004) Camptothecin: current perspectives. *Bioorg Med Chem*, **12**, 1585-1604.
  190. Redinbo, M.R., Stewart, L., Kuhn, P., Champoux, J.J. and Hol, W.G. (1998) Crystal structures of human topoisomerase I in covalent and noncovalent complexes with DNA. *Science*, **279**, 1504-1513.
  191. Staker, B.L., Hjerrild, K., Feese, M.D., Behnke, C.A., Burgin,

- A.B., Jr. and Stewart, L. (2002) The mechanism of topoisomerase I poisoning by a camptothecin analog. *Proc Natl Acad Sci U S A*, **99**, 15387-15392.
192. Tardy, C., Facompré, M., Laine, W., Baldeyrou, B., García-Grávalos, D., Francesch, A., Mateo, C., Pastor, A., Jiménez, J.A., Manzanares, I. *et al.* (2004) Topoisomerase I-mediated DNA cleavage as a guide to the development of antitumor agents derived from the marine alkaloid lamellarin D: triester derivatives incorporating amino acid residues. *Bioorg Med Chem*, **12**, 1697-1712.
193. Marco, E., Laine, W., Tardy, C., Lansiaux, A., Iwao, M., Ishibashi, F., Bailly, C. and Gago, F. (2005) Molecular determinants of topoisomerase I poisoning by lamellarins: comparison with camptothecin and structure-activity relationships. *J Med Chem*, **48**, 3796-3807.
194. Facompré, M., Tardy, C., Bal-Mahieu, C., Colson, P., Pérez, C., Manzanares, I., Cuevas, C. and Bailly, C. (2003) Lamellarin D: a novel potent inhibitor of topoisomerase I. *Cancer Res*, **63**, 7392-7399.
195. Morris, G.M., Huey, R. and Olson, A.J. (2008) Using AutoDock for ligand-receptor docking. *Curr Protoc Bioinformatics*, **Chapter 8**, Unit 8 14.
196. Staker, B.L., Feese, M.D., Cushman, M., Pommier, Y., Zembower, D., Stewart, L. and Burgin, A.B. (2005) Structures of three classes of anticancer agents bound to the human topoisomerase I-DNA covalent complex. *J Med Chem*, **48**, 2336-2345.
197. Molina, A., Vaquero, J.J., García-Navío, J.L., Alvarez-Builla, J., de Pascual-Teresa, B., Gago, F., Rodrigo, M.M. and Ballesteros, M. (1996) Synthesis and DNA binding properties of  $\gamma$ -carbolinium derivatives and benzologues. *J Org Chem* **61**, 5587-5599.
198. Pastor, J., Siro, J.G., García-Navío, J.L., Vaquero, J.J., Alvarez-Builla, J., Gago, F., de Pascual-Teresa, B., Pastor, M. and Rodrigo, M.M. (1997) Azino-Fused Benzimidazolium Salts as DNA Intercalating Agents. 2. *J Org Chem*, **62**, 5476-5483.
199. Molina, A., Vaquero, J.J., García-Navío, J.L., Alvarez-Builla,

- J., de Pascual-Teresa, B., Gago, F. and Rodrigo, M.M. (1999) Novel DNA Intercalators Based on the Pyridazino[1',6':1,2]pyrido[4,3-b]indol-5-inium System. *J Org Chem*, **64**, 3907-3915.
200. Martínez, V., Burgos, C., Alvarez-Builla, J., Fernández, G., Domingo, A., García-Nieto, R., Gago, F., Manzanares, I., Cuevas, C. and Vaquero, J.J. (2004) Benzo[f]azino[2,1-a]phthalazinium cations: novel DNA intercalating chromophores with antiproliferative activity. *J Med Chem*, **47**, 1136-1148.