

Bioinformática: aplicaciones en química medicinal

MARÍA LUISA SERRANO G.*

Resumen

La sinergia entre los métodos biológicos experimentales y la bioinformática ha proporcionado grandes beneficios a diferentes áreas de las ciencias biológicas, en particular a la química medicinal. En este trabajo se mostrara algunos ejemplos que sirven para ilustrar la utilidad de los métodos informáticos en el análisis de la estructura de las proteínas, y algunas de sus aplicaciones en modelado molecular que proporcionan información relevante acerca de sistemas de interés biomédico.

Palabras clave: bioinformática, modelado molecular.

Abstract

The synergy between experimental biology and bioinformatics has greatly benefited many different areas of the life sciences, especially medicinal chemistry. This work will describe a few examples of highlighting the effectiveness computational methods in protein structural analysis and some applications in molecular modeling that could provide relevant information about systems of biomedical interest.

Key words: bioinformatics, molecular modeling.

Introducción

Ha transcurrido más de un lustro desde que se completó la primera secuencia del genoma humano. A partir de ese momento, los Proyectos Genoma han impulsado grandes inversiones en biotecnología (IHGSC, 2004; Lander, 2001 y Venter, 2001) y aumentado la atención hacia la bioinformática, una disciplina emergente que puede ser ampliamente definida como la interfase entre dos ciencias, la biología y la computación, y una de las áreas de mayor crecimiento en la ciencia.

El desafío de esta disciplina es el facilitar la colección y organización de la información que se desprenden de los Proyectos Genoma, así como también desarrollar las herramientas mediante las cuales los investigadores puedan realizar el análisis e interpretación de varios tipos de datos, incluyendo secuencias de nucleótidos y aminoácidos, dominios y estructura de proteínas.

La primera industria que comprendió el impacto de la bioinformática fue la industria farmacéutica, que la utiliza para el diseño de nuevos fármacos a

partir de la detección de potenciales dianas o blancos terapéuticos. El trabajo multidisciplinario y el empleo de métodos informáticos permiten también estudiar la estructura y función de proteínas, evaluar sus interacciones con compuestos líderes, e inclusive estudiar interacciones entre proteínas, como las interacciones antígeno-anticuerpo, importantes para el desarrollo de vacunas.

A pesar de lo anteriormente mencionado, cabe preguntarse: ¿cuántos de los fármacos disponibles hoy en día se han descubierto gracias a la biología molecular computacional y al empleo de métodos informáticos para el diseño de fármacos? La realidad es que el uso de las herramientas computacionales ha permeado todos los aspectos del descubrimiento de nuevos fármacos. Muchos programas de investigación comienzan con la identificación de una diana molecular, de potencial valor terapéutico, a través del análisis computacional de los datos disponibles, análisis de secuencias, predicción de estructura de proteínas y búsquedas en bases de datos de la estructura, aunque su desarrollo como blanco

* Unidad de Química Medicinal, Facultad de Farmacia, Universidad Central de Venezuela, Caracas, 1041-A, Venezuela.

terapéutico se alcance luego de extensos estudios biológicos. Posteriormente, el esfuerzo se dirige hacia el descubrimiento de compuestos líderes de manera experimental, empleando el cribado masivo, que luego se mejoran con el uso de métodos informáticos como la evaluación *in silico* de librerías moleculares. Eventualmente, se puede determinar la estructura de complejos cristalinos de algunos compuestos con la diana molecular. Esto permite mejorarlos eficientemente empleando algunos de los diferentes métodos de diseño de fármacos.

En la literatura se encuentran excelentes artículos que ilustran las posibilidades de combinar el trabajo experimental y las aproximaciones computacionales (Tramontano, 2006; Chen, 2005; Fauman, 2003; Jorgensen, 2004 y Nielsen, 2005), así como también los logros alcanzados con métodos de diseño de fármacos basados en la estructura del receptor (SBDD) (Hardy, 2003 y Maryanoff, 2004) que han contribuido con unos cincuenta compuestos que se encuentran en pruebas clínicas y en la aprobación de numerosos fármacos.

Las aplicaciones de biología molecular computacional datan de los años 90, cuando se impulsa el desarrollo de métodos informáticos que utilizan la estructura del receptor o diana farmacológica, debido al incremento exponencial del número de estructuras cristalinas disponibles en las bases de datos públicas. Las bondades de emplear la estructura de las proteínas para desarrollar modelos de Relación Estructura Química-Actividad Biológica, puede ilustrarse con numerosos ejemplos de la literatura.

Los métodos indirectos que se basan en la estructura de los ligandos permiten generar una hipótesis de farmacóforo 3D (un arreglo único tridimensional de los grupos funcionales comunes a la mayoría de los ligandos del receptor). En muchos casos, cuando los ligandos no son estructuralmente muy diferentes y se cuenta con un compuesto rígido desde el punto de vista conformacional, se puede obtener un farmacóforo razonable por métodos tanto automáticos como semiautomáticos.

Los métodos directos, por el contrario, requieren del conocimiento previo de la estructura tridimensional del receptor o diana farmacológica, y cuentan con algoritmos capaces de estudiar el posicionamiento molecular de los ligandos en el sitio activo. Ha resultado muy atractivo poder combinar la precisión de la información disponible en los complejos ligando-receptor con la eficiencia computacional de los métodos informáticos indirectos. Con esta aproximación, la alineación de los ligandos, para desarrollar una hipótesis de farmacóforo y establecer un

modelo cuantitativo de relación estructura actividad química-actividad biológica (QSAR), se realiza tomando como plantilla la orientación del ligando en el sitio activo bajo estudio. Esta estrategia había sido empleada de manera exitosa por Waller y colaboradores con 59 inhibidores de la HIV proteasa (Waller, 1993).

Nosotros empleamos esta aproximación conjuntamente con el método CoMFA para obtener un farmacóforo tridimensional QSAR de 60 inhibidores, estructuralmente diferentes, de la enzima acetilcolinesterasa (AChE) (Cho y col., 1996). La inhibición de esta enzima se consideraba como una de las estrategias prometedoras para el tratamiento del mal de Alzheimer (Hakansson, 1993) y con posibilidades terapéuticas para el tratamiento de la Enfermedad de Parkinson (Sirvio, 1992). Durante años habían sido sintetizados cientos de inhibidores, algunos con aplicaciones clínicas, pero de estructura química muy diversa y no había sido posible establecer un farmacóforo común a todos ellos. La estructura química de estos inhibidores es variada, desde compuestos bicuaternarios como el decametanium (DMC) hasta compuestos monocatiónicos simples como el edromorfium (EDR), y formalmente neutros como la tacrina (THA). En ese momento fue reportado el análisis por cristalografía de rayos-X de la AChE de *Torpedo californica* (EC 3.1.1.7) y seguidamente las estructuras cristalinas de tres complejos de la enzima con tres inhibidores estructuralmente diferentes, EDR, THA y DMC (Harel, 1993) que proporcionaron información, no sólo acerca de la conformación bioactiva de los ligandos, sino también de la orientación relativa de ellos en el sitio activo de la enzima. Esta aproximación permitió desarrollar un modelo CoMFA de alto valor predictivo (q^2 de 0.734) para los 60 inhibidores de la AChE.

En otros casos, además de utilizar la estructura cristalina de los complejos enzima/inhibidor para obtener la conformación bioactiva, se pueden obtener datos adicionales que de alguna manera contribuyan a establecer un modelo QSAR que correlacione la actividad biológica con algunos parámetros fisicoquímicos. Nuestro grupo ha utilizado esta estrategia recientemente en un estudio QSAR realizado para una serie de análogos de celecoxib (sulfonamidas del tipo diarilpirazoles, diarilimidazoles y diarilpirroles), inhibidores selectivos de la enzima ciclooxigenasa 2 (COX-2) (Gómez, González y Serrano, 2006). Debido al reciente interés de estos compuestos como anticancerígenos, las investigaciones en cuanto al desarrollo de este tipo de fármacos antiinflamatorios continúan, aun cuando varios de estos medicamentos han sido eliminados por la FDA del

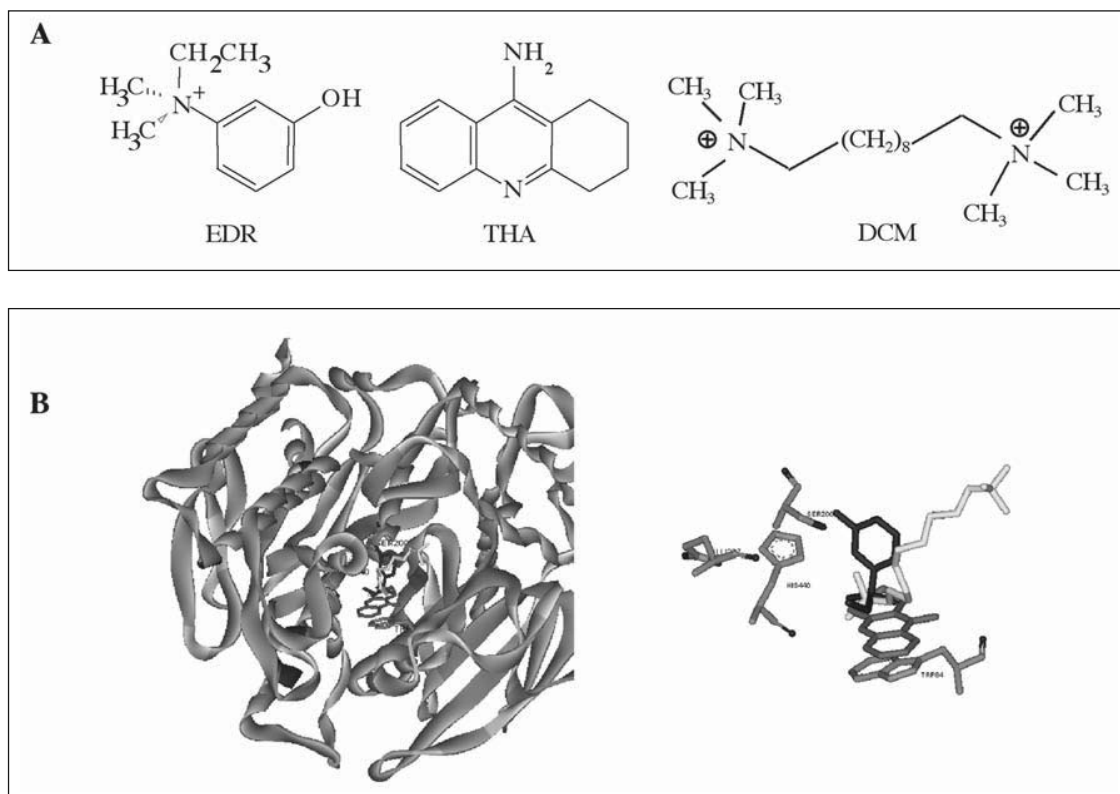


Figura 1

A. Inhibidores de la acetilcolinesterasa (AChE) presentes en las estructuras cristalinas. B. Superposición de los tres complejos cristalinos EDR/AChE, THA/AChE y DCM/AChE. En la ampliación se observa la orientación relativa en el sitio catalítico de la enzima de los tres inhibidores.

mercado farmacéutico por generar problemas a nivel cardíaco (Dogne, 2005). Algunos estudios demuestran que la COX-2 se encuentra sobre-expresada en numerosos tejidos cancerígenos humanos (Baptiste, 2006).

Si bien hoy en día estas aproximaciones siguen siendo de utilidad, los avances más recientes en el área de diseño de nuevos ligandos van dirigidos en su mayoría hacia el Cribado de Librerías de Compuestos y el diseño *de Novo*. En la etapa que corresponde a la generación de nuevos ligandos, el Cribado Altamente Eficiente (High Throughput Screening, HTS) requiere de una librería de compuestos y de los correspondientes ensayos de actividad (Herzberg, 2000); esta aproximación, que muchas veces es poco exitosa y muy costosa, se ha colocado en desventaja frente al Cribado Virtual. Este último puede realizarse sobre librerías de compuestos conocidos, o teóricos, y sólo requiere de los datos de actividad para algunos de los compuestos o de la estructura del blanco terapéutico. En este sentido, a pesar de haber notables avances en las diferentes aproximaciones de Cribado Virtual, el estudio de las interacciones ligando-receptor es actualmente una de las aproximaciones más populares con métodos como

el Docking de Alta Eficiencia (High-Throughput Docking (HTD)), que permite evaluar el posicionamiento molecular en el sitio activo de una diana farmacológica, con bases de datos de cientos o miles de compuestos, con el propósito de identificar prototipos novedosos que presenten la actividad biológica deseada (Wang, 2001). Estos compuestos se sintetizan, modifican y, en combinación con estudios teóricos de propiedades ADME (absorción, distribución, metabolismo y excreción) se puede reducir el número de compuestos que se somete a la evaluación biológica final (Shoicket, 2002).

Por otra parte, el diseño *de Novo* permite generar nuevos inhibidores en el sitio activo conocido del blanco terapéutico (Taylor, 2002). Los programas de Docking pueden utilizarse con este propósito cuando están acoplados con un generador automático de estructuras. Actualmente hay programas especializados que han sido desarrollados para construir posibles ligandos en el sitio activo de una diana, usualmente colocando y conectando fragmentos moleculares o conectando los fragmentos a un núcleo. Entre las primeras alternativas se puede mencionar a programas como LUDI (Böhm, 1992) y SPROUT (Gillet, 1994) y entre las más recientes se

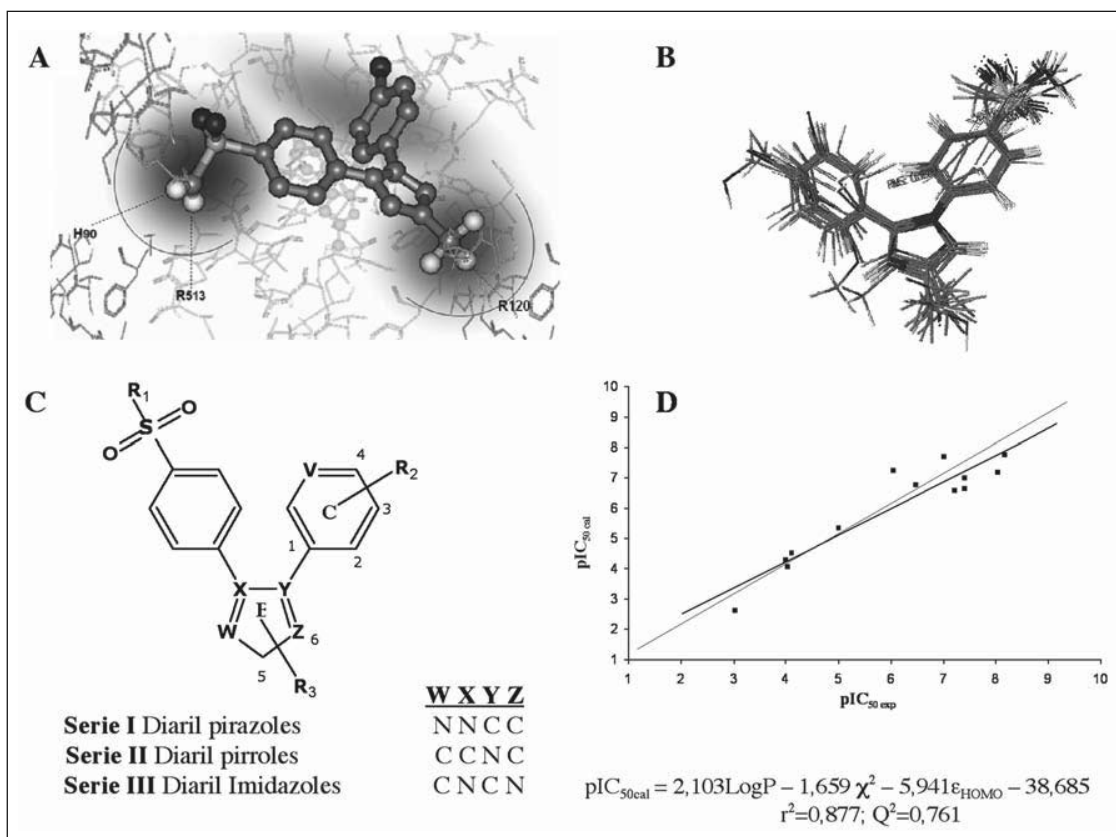


Figura 2

A. Complejo cristalino de SC-558/COX-2 donde se muestra las interacciones de puente de hidrógeno con los residuos H₉₀ y R₅₁₃ y el grupo sulfonamido, así como las del grupo trifluorometil con el residuo R₁₂₀. B. Superposición de las moléculas estudiadas. C. Inhibidores de la COX-2. D. Cálculo y evaluación de la ecuación de regresión lineal.

encuentra BOMB (Biochemical and Organic Model Builder) (Jorgensen, 2004), que puede utilizarse para construir estructuras individuales o una librería combinatoria.

La contribución de la bioinformática al desarrollo de vacunas (fármacos de naturaleza peptídica) es aun más importante, especialmente porque los prototipos son proteínas y por ello la bioinformática, la genómica y la proteómica han revolucionado la investigación en el área de diagnóstico, terapéutica y desarrollo de vacunas. Tradicionalmente, las vacunas se desarrollaban según los principios de Pasteur, aislando, inactivando e inyectando el agente causativo de una enfermedad infecciosa. Sin embargo, al disponer del genoma completo de diversos microorganismos, la aproximación se puede realizar a través de un proceso denominado *vaccinología reversa* (Rappuoli, 2000). Es un proceso que se inicia con el análisis de la información genómica que finalmente conduce, luego del análisis *in silico*, a la predicción de antígenos con posibilidades para ser desarrollados como vacunas. El nuevo

reto para este campo será combinar la vaccinología con la biología estructural.

El desarrollo de vacunas y otros compuestos de interés biológico necesita de la comprensión del papel fisiológico y funcional de las proteínas involucradas en esos procesos. Esta tarea se facilita cuando se dispone de la estructura tridimensional de las proteínas. En la actualidad, a pesar de la abundante cantidad de secuencias que se encuentra disponible en las bases de datos especializadas, como SWISS-PROT y trEMBL (Bairoch, 2004 y Boeckmann, 2005), sólo una pequeña parte de ellas tiene una estructura tridimensional conocida. La estructura 3D de una proteína, que es en definitiva el aspecto determinante para su función biológica, puede ser determinada experimentalmente, ya sea por cristalografía de rayos-X o por RMN. Sin embargo, esto no siempre es posible y, por esta razón, la determinación estructural de proteínas, empleando métodos *in silico*, ha realizado un aporte cada vez mayor gracias a los avances realizados en los programas de modelado molecular (Sali, 1997, Tramontano, 1998; Byströff, 1997 y Jones, 2003).

Las estructuras tridimensionales así determinadas y el uso de programas diseñados para la predicción de péptidos antigénicos (Tramontano, 1998) han permitido aproximarse al desarrollo de vacunas empleando péptidos sintéticos que simulen una región de la proteína nativa.

Aunque hasta la fecha no existen protocolos o programas que conduzcan a la identificación efectiva de secuencias antigénicas, hay varios métodos que se basan en algunas de las propiedades fisicoquímicas que presentan epítopos que han sido determinados experimentalmente, tales como flexibilidad, hidrofobicidad y accesibilidad. Entre los programas para la predicción de péptidos antigénicos se encuentran: Preditop (Pellequer, 1993), Antigenic Index (Jameson, 1988), Antigenic (Kolaskar, 1990) y Antheptot (Deleage, 2001).

En el área de la parasitología molecular, la bioinformática ha producido un gran impacto, especialmente en el caso de algunas enfermedades parasitarias como la malaria, debido a que el dramático incremento de la resistencia de los parásitos a los antimaláricos y a las medidas de control vectorial, ha creado la necesidad de desarrollar nuevas estrategias para atacar la enfermedad. Entre ellas el desarrollo de vacunas antimaláricas tiene, hoy en día, altísima prioridad.

El ciclo de vida del *Plasmodium* es extraordinariamente complejo, requiere de la expresión de proteínas especializadas para sobrevivir en el ambiente de sus huéspedes (el mosquito y el hombre), para la vida intracelular y extracelular, para la invasión de diferentes tipos de células y para evadir la respuesta inmune del huésped. Las estrategias para atacar esta enfermedad, ya sea con vacunas o nuevos fármacos, será mas efectiva si van dirigidas hacia un blanco o diana específica en el ciclo de vida del parásito y/o hacia proteínas que se expresan en estas etapas. Los genomas de *P. falciparum* y *P. yoelli* están completos, y actualmente se realizan estudios de secuenciación del genoma de *P. vivax* en el TIGR (The Institute for Genomic Research, Rockville MD), y permitirán a corto plazo identificar nuevos blancos o dianas que puedan emplearse para desarrollar vacunas y nuevos fármacos.

Nuestro interés en esta área se ha centrado fundamentalmente en *P. vivax*, que el parásito de mayor incidencia en Venezuela (boletines del MSDS, 2003). El objetivo de los estudios que hemos venido realizando ha sido desarrollar y utilizar los modelos tridimensionales de algunas proteínas de este parásito para definir posibles epítopos y proceder a la síntesis de péptidos con potenciales propiedades antigénicas.

Hasta el presente se ha identificado y caracterizado inmunológicamente varios antígenos de *P. vivax* del estadio asexual del parásito, muchos de ellos homólogos de *P. falciparum*, y algunos de ellos están siendo evaluados en ensayos pre-clínicos de vacunación. Entre ellos, la Proteína 1 de Superficie del Merozoito de *P. vivax* (PvMSP-1), una proteína altamente antigénica y un fuerte candidato a vacuna (Collins, 1999; Yang, 1999). Dado que el fragmento C-terminal de esta proteína es el más inmunogénico (Soares, 1997) y en vista de que hasta la fecha no ha sido reportada la estructura cristalina de este fragmento para *P. vivax*, se construyó un modelo por homología a partir de las estructuras cristalinas de otras especies de *Plasmodium* (Serrano, Pérez y Medina, 2006).

Los datos estructurales así obtenidos permitieron realizar la comparación entre el modelo para la PvMSP-1₁₉ y las estructuras cristalinas previamente reportadas para otras especies del parásito. De igual forma permitió la identificación de una cavidad en el primer dominio de la proteína adecuada para interacciones proteína-proteína (Figura 3).

Empleando el modelo se planificó una estrategia para la selección de los péptidos a sintetizar, usando de manera conjunta algoritmos matemáticos de predicción de epítopos como ANTHEPROT 4.0 y el método de Hoops y Woods para el análisis

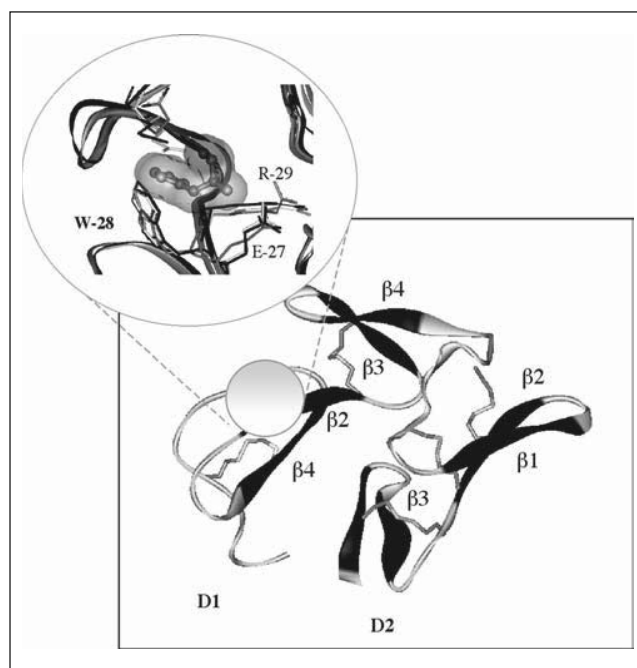


Figura 3

Modelo final de la PvMSP-1₁₉. La representación en cinta incluye los enlaces disulfuro, muestra la cadena C? y las láminas antiparalelas ? en cada uno de los dominios EGF. Detalle donde se muestra la cavidad ubicada en el primer dominio de la proteína y se compara el modelo PvMSP-1₁₉ con las estructuras cristalinas de PmMSP-1₁₉ y PkMSP-1₁₉.

de las regiones hidrofílicas de la proteína (Hoop, 1981). Los péptidos fueron evaluados frente al plasma de individuos con anticuerpos a *P. vivax* y varios de ellos resultaron altamente reactivos (Serrano, 2004). Actualmente nuestro grupo continúa trabajando en este sentido y pensamos extendernos hacia el estudio de otras proteínas.

El conocimiento, aunque sea aproximado, de la estructura tridimensional de las proteínas es esencial para conocer los detalles de su función molecular y proporciona información valiosa para el desarrollo de estrategias racionales que permitan identificar dianas farmacológicas, estudiar el posicionamiento molecular de nuevos ligandos, plantear mejores modelos teóricos de relación estructura química-actividad biológica (QSAR) y finalmente diseñar nuevos fármacos y vacunas.

Los ejemplos que hemos mencionado sólo describen una pequeña muestra de las enormes posibilidades de las aplicaciones de la bioinformática y de la gran diversidad de métodos informáticos que de ella se nutren.

Referencias bibliográficas

- BAIROCH A, BOECKMANN B, FERRO S, GASTEIGER E. 2004. Swiss-Prot: Juggling between evolution and stability. *Brief Bioinform.* 5:39-55.
<http://www.expasy.org/sprot/>
- BOECKMANN B, BLATTER MC, FAMIGLIETTI L, HINZ U, LANE L, ROECHERT B, BAIROCH A. 2005. Protein variety and functional diversity: Swiss-Prot annotation in its biological context *Comptes Rendus Biologies* 328:882-99.
- BÖHM HJ. 1992. The computer program LUDI: a new method for the de novo design of enzyme inhibitors. *J Comput Aided Mol Des* 6: 61.
- BYSTROFF C, BAKER D. 1997. Blind predictions of local protein structure in CASP2 targets using the I-sites library. *Proteins, Suppl.* 1: 167-171.
- CHEN I, NEAMATI N, MACKERELL A. 2002. Structure-based inhibitor design targeting HIV-1 integrase. *Curr Drug Targets Infect Disord* 2: 217-234.
- CHO SJ, SERRANO ML, BIER J, TROPSHA A. 1996. Structure-Based Alignment and Comparative Molecular Field Analysis of Acetylcholinesterase inhibitors. *J Med Chem* 39: 5064-5071.
- COLLINS WE, KASLOW DC, SULLIVAN JS, MORRIS CL, GALLAND GG, YANG C, SAEKHO AM, XIAO L, LAL AA. 1999. Testing the efficacy of a recombinant merozoite surface protein (MSP-1(19) of *Plasmodium vivax* in *Saimiri boliviensis* monkeys. *Am J Trop Med Hyg* 60: 350-356.
- DELEAGE G, COMBET C, BLANCHET C, GEOURJON C. 2001. ANTHE-PROT: An integrated protein sequence analysis software with client/server capabilities. *Computers Biol Med* 31: 259-267.
- FAUMAN E, HOPKINS A, GROOM C. 2003. Structural bioinformatics in drug discovery. *Methods Biochem Anal* 44: 477-497.
- FLORENS L, WASHBURN MP, RAINE JD, ANTHONY RM, GRAINGER M, HAYNES JD, MOCH JK, MUSTER N, SACCI JB, TABB DL, WITNEY AA, WOLTERS D, WU Y, GARDNER MJ, HOLDER AA, SINDEN RE, YATES JR, CARUCCI DJ. 2002. A proteomic view of the *Plasmodium falciparum* life cycle. *Nature* 419: 520-526.
- GILLET VJ, NEWELL W, MATA P, MYATT G, SIKE S, ZSOLDOS Z, JOHNSON AP. 1994. SPROUT: recent developments in the de novo design of molecules. *J Chem Inf Comput Sci* 34: 207-17.
- HAKANSSON L. 1993. Mechanism of Action of Cholinesterase Inhibitors in Alzheimer's Disease. *Acta Neurol. Scand. Suppl.* 149, 7-9.
- HARDY LW, MALIKAYIL A. 2003. The impact of structure-guided drug design on clinical agents. *Curr Drug Discov* 15: 15-20.
- HAREL M, SCHALK I, EHRET-SABATIER L, BOUET F, GOELDNER M, HIRTH C, AXELSEN PH, SILMAN I, SUSSMAN JL. 1993. Quaternary ligand binding to aromatic residues in the active-site gorge of acetylcholinesterase. *Proc. Natl. Acad. Sci. U.S.A.* 90: 9031-9035.
- HERZBERG RP, POPE AJ. 2000. High-throughput screening: new technology for the 21st century. *Curr Opin Chem Biol* 4: 445-451.
- HOOP TP, WOODS, KR. 1981. Prediction of protein antigenic determinants from amino acid sequences. *Proc Natl Acad* 78: 3824-3828.
- INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM. 2004. Finishing the euchromatic sequence of the human genome. *Nature* 431:931-945.
- JAMESON BA, WOLF H. 1988. The antigenic index: a novel algorithm for predicting antigenic determinants. *Comput Appl Biosci.* 41:181-186.
- JONES DT, MCGUFFIN LJ. 2003. Assembling novel protein folds from supersecondary structural fragments. *Proteins* 5, Suppl. 6: 480-485.
- JORGENSEN W. 2004. The many roles of computation in drug discovery. *Science* 303: 1813-1818.
- KOLASKAR AS, TONGAONKAR PC. 1990. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett.* 276: 172-174.
- LANDER ES, LINTO LM, BIRREN B, NUSBAUM C, ZODY MC, BALDWIN J, DEVON K, DEWAR K, DOYLE M, FITZHUGH W, FUNKE R, GAGE D, HARRIS K, HEAFORD A, HOWLAND J, KANN L, LEHOCZKY J, LEVINE R, MCEWAN P, MCKERNAN K, MELDRIM J, MESIROV JP, MIRANDA C, MORRIS W, NAYLOR J, RAYMOND C, ROSETTI M, SANTOS R, SHERIDAN A, SOUGNEZ C, STANGE-THOMANN N, STOJANOVIC N, SUBRAMANIAN A, WYMAN D, ROGERS J, SULSTON J, AINSCOUGH R, BECK S, BENTLEY D, BURTON J, CLEE C, CARTER, N,

- COULSON A, DEADMAN R, DELOUKAS P, DUNHAM A, DUNHAM I, DURBIN R, FRENCH L, GRAFHAM D, GREGORY S, HUBBARD T, HUMPHRAY S, HUNT A, JONES M, LLOYD C, McMURRAY A, MATTHEWS L, MERCER S, MILNE S, MULLIKIN JC, MUNGALL A, PLUMB R, ROSS M, SHOWNKEEN R, SIMS S, WATERSTON RH, WILSON RK, HILLIER LW, McPHERSON JD, MARRA MA, MARDIS ER, FULTON LA, CHINWALLA AT, PEPIN KH, GISH WR, CHISSOE SL, WENDL MC, DELEHAUNTY KD, MINER TL, DELEHAUNTY A, KRAMER JB, COOK LL, FULTON RS, JOHNSON DL, MINX PJ, CLIFTON SW, HAWKINS T, BRANSCOMB E, PREDKI P, RICHARDSON P, WENNING S, SLEZAK T, DOGGETT N, CHENG JF, OLSEN A, LUCAS S, ELKIN C, UBERBACHER E, FRAZIER M, GIBBS RA, MUZYNY DM, SCHERER SE, BOUCK JB, SODERGREN EJ, WORLEY KC, RIVES CM, GORRELL JH, METZKER ML, NAYLOR SL, KUCHERLAPATI RS, NELSON DL, WEINSTOCK GM, SAKAKI Y, FUJIYAMA A, HATTORI M, YADA T, TOYODA A, ITOH T, KAWAGOE C, WATANABE H, TOTOKI Y, TAYLOR T, WEISSENBACH J, HEILIG R, SAURIN W, ARTIGUENAVE F, BROTTIER P, BRULS T, PELLETIER E, ROBERT C, WINCKER P, SMITH DR, DOUCETTE-STAMM L, RUBENFIELD M, WEINSTOCK K, LEE HM, DUBOIS J, ROSENTHAL A, PLATZER M, NYAKATURA G, TAUDIEN S, RUMP A, YANG H, YU J, WANG J, HUANG G, GU J, HOOD L, ROWEN L, MADAN A, QIN S, DAVIS RW, FEDERSPIEL NA, ABOLA AP, PROCTOR MJ, MYERS RM, SCHMUTZ J, DICKSON M, GRIMWOOD J, COX DR, OLSON MV, KAUL R, RAYMOND C, SHIMIZU N, KAWASAKI K, MINOSHIMA S, EVANS GA, ATHANASIOU M, SCHULTZ R, ROE BA, CHEN F, PAN H, RAMSER J, LEHRACH H, REINHARDT R, McCOMBIE WR, DE LA BASTIDE M, DEDHIA N, BLÖCKER H, HORNISCHER K, NORDSIEK G, AGARWALA R, ARAVIND L, BAILEY JA, BATEMAN A, BATZOGLIOU S, BIRNEY E, BORK P, BROWN DG, BURGE CB, CERUTTI L, CHEN HC, CHURCH D, CLAMP M, COPLEY RR, DOERKS T, EDDY SR, EICHLER EE, FUREY TS, GALAGAN J, GILBERT JG, HARMON C, HAYASHIZAKI Y, HAUSSLER D, HERMJAKOB H, HOKAMP K, JANG W, JOHNSON LS, JONES TA, KASIF S, KASPRYZK A, KENNEDY S, KENT WJ, KITTS P, KOONIN EV, KORF I, KULP D, LANCET D, LOWE TM, McLYSAGHT A, MIKKELSEN T, MORAN JV, MULDER N, POLLARA VJ, PONTING CP, SCHULER G, SCHULTZ J, SLATER G, SMIT AF, STUPKA E, SZUSTAKOWSKI J, THIERRY-MIEG D, THIERRY-MIEG J, WAGNER L, WALLIS J, WHEELER R, WILLIAMS A, WOLF YI, WOLFE KH, YANG SP, YEH RF, COLLINS F, GUYER MS, PETERSON J, FELSENFELD A, WETTERSTRAND KA, PATRINOS A, MORGAN MJ, DE JONG P, CATANESE JJ, OSOEGAWA K, SHIZUYA H, CHOI S, CHEN YJ; INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM (2001). Initial sequencing and analysis of the human genome. *Nature* 409: 860-921.
- MARYANOFF BE. 2004. Inhibitors of Serine Proteases as Potential Therapeutic Agents: The Road from Thrombin to Trypsin to Cathepsin G. *J Med Chem* 47: 769-787.
- NIELSEN U, SCHOEBERL B. 2005. Using computational modeling to drive the development of targeted therapeutics. *Drugs* 8: 822-826.
- PELLEQUER JL, WESTHOF E. 1993. PREDITOP: a program for antigenicity prediction. *J Mol Graph.* 11:204-10, 191-192.
- Plasmodium vivax Genome Project. The Institute for Genomic Research (TIGR) <http://www.tigr.org>
- PROGRAMA DE ERRADICACIÓN DE LA MALARIA. Boletines del Ministerio de Salud y Desarrollo Social (2003). República Bolivariana de Venezuela.
- RAPPUOLI R. 2000. Reverse vaccinology. *Curr Opin Microbiol* 3: 445-450.
- SÁNCHEZ R, SALI A. 1997. Advances in comparative protein-structure modeling. *Curr Opin Struct Biol* 7: 206-214.
- SERRANO ML. 2004. Determinación de la Estructura Tridimensional de Moléculas Peptídicas a Través de Modelado Molecular. Aplicación a la Porción C-terminal de la MSP-1 de *Plasmodium vivax*. Selección y Síntesis de Péptidos con Posible Actividad Antigénica. Tesis Doctoral. Universidad Central de Venezuela.
- SERRANO ML, PÉREZ HA, MEDINA JD. 2006. Structure of C-terminal fragment of merozoite surface protein-1 from *Plasmodium vivax* determined by homology modeling and molecular dynamics refinement, *Bioorganic & Medicinal Chemistry* 14: 8359-8365.
- SHOICKET BK, McGOVERN SL, WEI B, IRWIN JJ. 2002. Lead discovery using molecular docking. *Curr Opin Chem Biol* 6: 439.
- SIRVIO J, RIEKKINEN PJ. 1992. Brain and Cerebrospinal Fluid Cholinesterases in Alzheimer's Disease, Parkinson's Disease and Aging. A Critical Review of Clinical and Experimental Studies. *J. Neural Transm.: Parkinson's. Dis Dementia Sect.* 4: 337-358.
- SOARES IS, LEVITUS G, SOUSA JM, DEL PORTILLO HA, RODRIGUES MM. 1997. Acquired immune responses to the N- and C-terminal regions of *Plasmodium vivax* merozoite surface protein 1 in individuals exposed to malaria. *Infect Immun* 65: 1606-1614.
- TAYLOR RD, JEWSEBURY PJ, ESSEX JW. 2002. A review of protein-small molecule docking methods. *J Comput Aided Mol Des* 16: 151-166.
- TRAMONTANO A. 1998. Homology modelling with low sequence identity. *Methods (San Diego, CA)* 14: 293-300.
- TRAMONTANO A. 2006. The role of molecular modeling in biomedical research. *FEBS Lett* 2928-2934.
- VENTER JC, ADAMS MD, MYERS EW, LI PW, MURAL RJ, SUTTON GG, SMITH HO, YANDELL M, EVANS CA, HOLT RA, GOCAYNE JD, AMANATIDES P, BALLEW RM, HUSON DH, WORTMAN JR, ZHANG Q, KODIRA CD, ZHENG XH, CHEN L, SKUPSKI M, SUBRAMANIAN G, THOMAS PD, ZHANG J, GABOR MIKLOS GL, NELSON C, BRODER S, CLARK AG, NADEAU J, McKUSICK VA, ZINDER N, LEVINE AJ, ROBERTS RJ, SIMON M, SLAYMAN C, HUNKAPILLER M, BOLANOS R, DELCHER A, DEW I, FASULO D, FLANIGAN M, FLOREA L, HALPERN A, HANNENHALLI S, KRAVITZ S, LEVY S, MOBARRY C, REINERT K, REMINGTON K, ABU-THREIDEH J, BEASLEY E, BIDDICK K, BONAZZI V, BRANDON R, CARGILL M, CHANDRAMOULISWARAN I, CHARLAB R, CHATURVEDI K, DENG Z, DI FRANCESCO V, DUNN P, EILBECK K, EVANGELISTA C, GABRIELIAN AE, GAN W, GE W, GONG F, GU Z, GUAN P, HEIMAN TJ, HIGGINS ME, JI RR, KE Z, KETCHUM KA, LAI Z, LEI Y, LI Z, LI J, LIANG Y, LIN X, LU F, MERKULOV GV, MILSHINA N, MOORE HM, NAIK AK, NARAYAN VA, NEELAM B, NUSSKERN D, RUSCH DB, SALZBERG S, SHAO W, SHUE B, SUN J, WANG Z, WANG A, WANG X, WANG J, WEI M, WIDES R, XIAO C, YAN C, YAO A, YE J, ZHAN M, ZHANG W, ZHANG H, ZHAO Q, ZHENG L, ZHONG F,

- ZHONG W, ZHU S, ZHAO S, GILBERT D, BAUMHUETER S, SPIER G, CARTER C, CRAVCHIK A, WOODAGE T, ALI F, AN H, AWE A, BALDWIN D, BADEN H, BARNSTEAD M, BARROW I, BEESON K, BUSAM D, CARVER A, CENTER A, CHENG ML, CURRY L, DANAHER S, DAVENPORT L, DESILETS R, DIETZ S, DODSON K, DOUP L, FERRIERA S, GARG N, GLUECKSMANN A, HART B, HAYNES J, HAYNES C, HEINER C, HLADUN S, HOSTIN D, HOUCK J, HOWLAND T, IBEGWAM C, JOHNSON J, KALUSH F, KLINE L, KODURU S, LOVE A, MANN F, MAY D, MCCAWLEY S, MCINTOSH T, McMULLEN I, MOY M, MOY L, MURPHY B, NELSON K, PFANNKOCHE C, PRATTS E, PURI V, QURESHI H, REARDON M, RODRIGUEZ R, ROGERS YH, ROMBLAD D, RUHFEL B, SCOTT R, SITTER C, SMALLWOOD M, STEWART E, STRONG R, SUH E, THOMAS R, TINT NN, TSE S, VECH C, WANG G, WETTER J, WILLIAMS S, WILLIAMS M, WINDSOR S, WINN-DEEN E, WOLFE K, ZAVERI J, ZAVERI K, ABRIL JF, GUIGO R, CAMPBELL MJ, SJOLANDER KV, KARLAK B, KEJARIWAL A, MI H, LAZAREVA B, HATTON T, NARECHANIA A, DIEMER K, MURUGANUJAN A, GUO N, SATO S, BAFNA V, ISTRAIL S, LIPPERT R, SCHWARTZ R, WALENZ B, YOUSEPH S, ALLEN D, BASU A, BAXENDALE J, BLICK L, CAMINHA M, CARNES-STINE J, CAULK P, CHIANG YH, COYNE M, DAHLKE C, MAYS A, DOMBROSKI M, DONNELLY M, ELY D, ESPARHAM S, FOSLER C, GIRE H, GLANOWSKI S, GLASSER K, GLODEK A, GOROKHOV M, GRAHAM K, GROPMAN B, HARRIS M, HEIL J, HENDERSON S, HOOVER J, JENNINGS D, JORDAN C, JORDAN J, KASHA J, KAGAN L, KRAFT C, LEVITSKY A, LEWIS M, LIU X, LOPEZ J, MA D, MAJOROS W, MCDANIEL J, MURPHY S, NEWMAN M, NGUYEN T, NGUYEN N, NODELL M, PAN S, PECK J, PETERSON M, ROWE W, SANDERS R, SCOTT J, SIMPSON M, SMITH T, SPRAGUE A, STOCKWELL T, TURNER R, VENTER E, WANG M, WEN M, WU D, WU M, XIA A, ZANDIEH A, ZHU X. (2001). The sequence of the human genome. *Science* 291: 1304-1351.
- WALLER CL, OPREA TI, GIOLITTI A, MARSHALL GR. 1993. Three-Dimensional QSAR of Human Immunodeficiency Virus (I) Protease Inhibitors. 1. A CoMFA Study Employing Experimentally-Determined Alignment Rules. *J. Med. Chem.* 36: 4152-4160.
- WANG R, WANG S. 2001. How does consensus scoring work for virtual library screening? An idealized computer experiment. *J Chem Inf Comput Sci* 41: 1422-1426.
- WELLING GW, WIEJER WJ, VAN DER ZEE R, WELLING-WEBSTER S. 1985. Prediction of sequential antigenic regions in proteins. *FEBS Letters* 188: 215-218.
- WHO, 2001. Informe de la Situación de los Programas Regionales en las Américas.
- YANG C, COLLINS WE, SULLIVAN JS, KASLOW DC, XIAO L, LAL AA. (1999). Partial protection against *Plasmodium vivax* blood-stage infection in *Saimiri* monkeys by immunization with a recombinant C-terminal fragment of merozoite surface protein 1 in block copolymer adjuvant. *Infect Immun* 67: 342-349.