

Diseño de Fármacos Asistido por computadora

Dr. David Ramírez
dramirezs@udec.cl

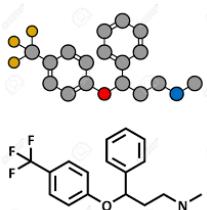
Conceptos Básicos:

Droga



Materia prima con más
De 1 principio activo

Fármaco



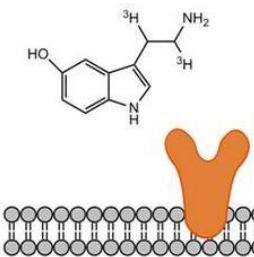
molécula bioactiva
Químicamente definida

Medicamento



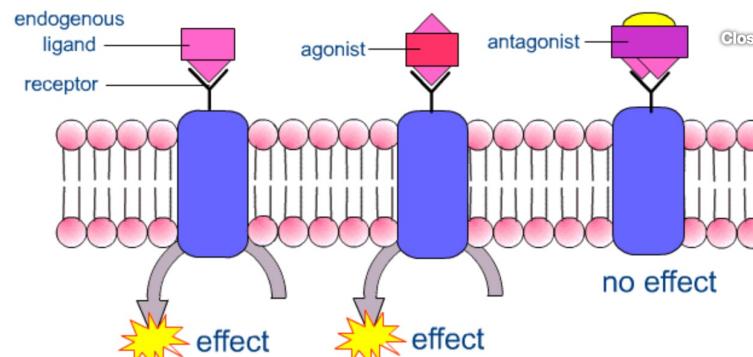
Forma farmacéutica
1 o más principios activos

Receptor



Donde se une una molécula
Respuesta fisiológica

Agonista



Antagonista

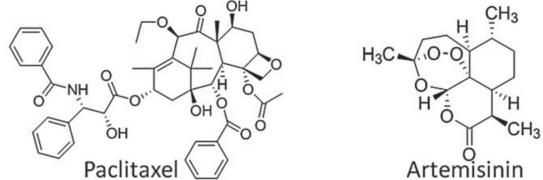
Principios generales

INTRODUCCIÓN

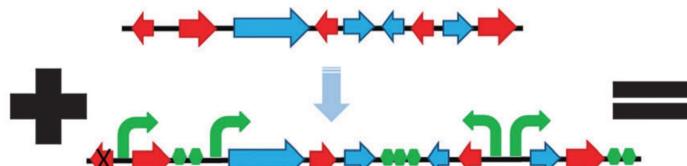
La química medicinal es la ciencia que se ocupa del diseño y desarrollo de nuevos agentes farmacéuticos.

La química Medicinal se interesa en:

A Isolation, extraction from plants



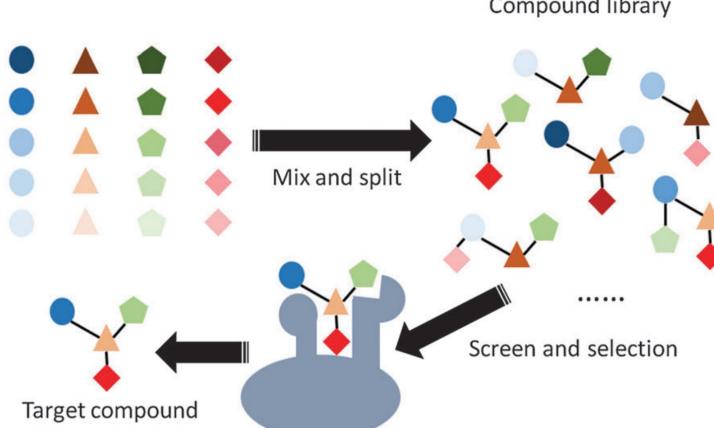
C Synthetic biology



Engineer microorganism/chassis

Engineer biosynthetic pathway parts

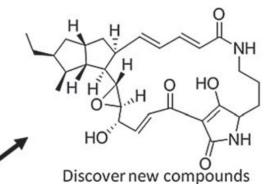
B Combinatorial chemistry



Target compound

Compound library

Screen and selection



Discover new compounds

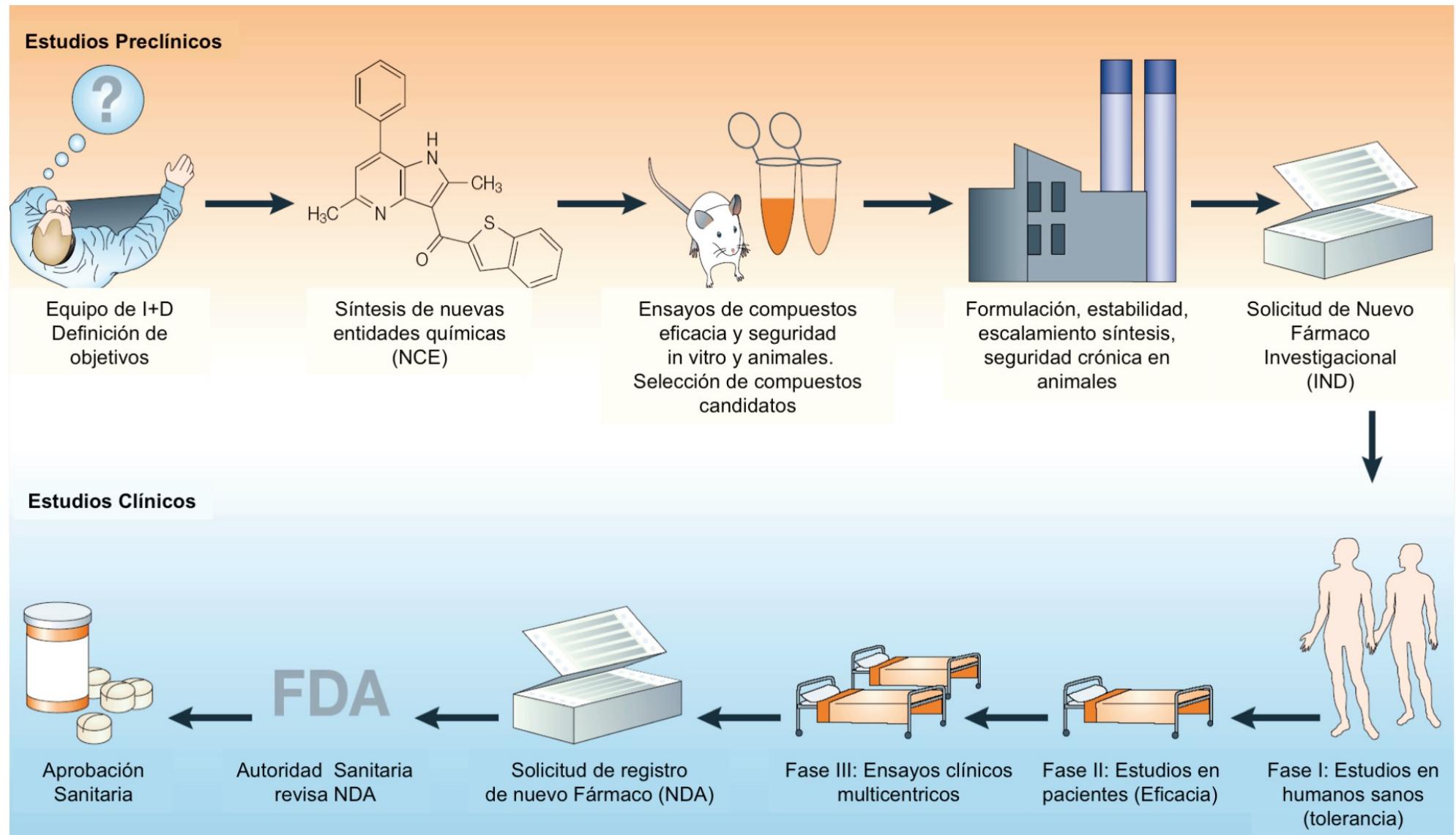


Productivity

Increase productivity of known compounds

Síntesis, Aislamiento, SAR, Interacción P-L, farmacocinética

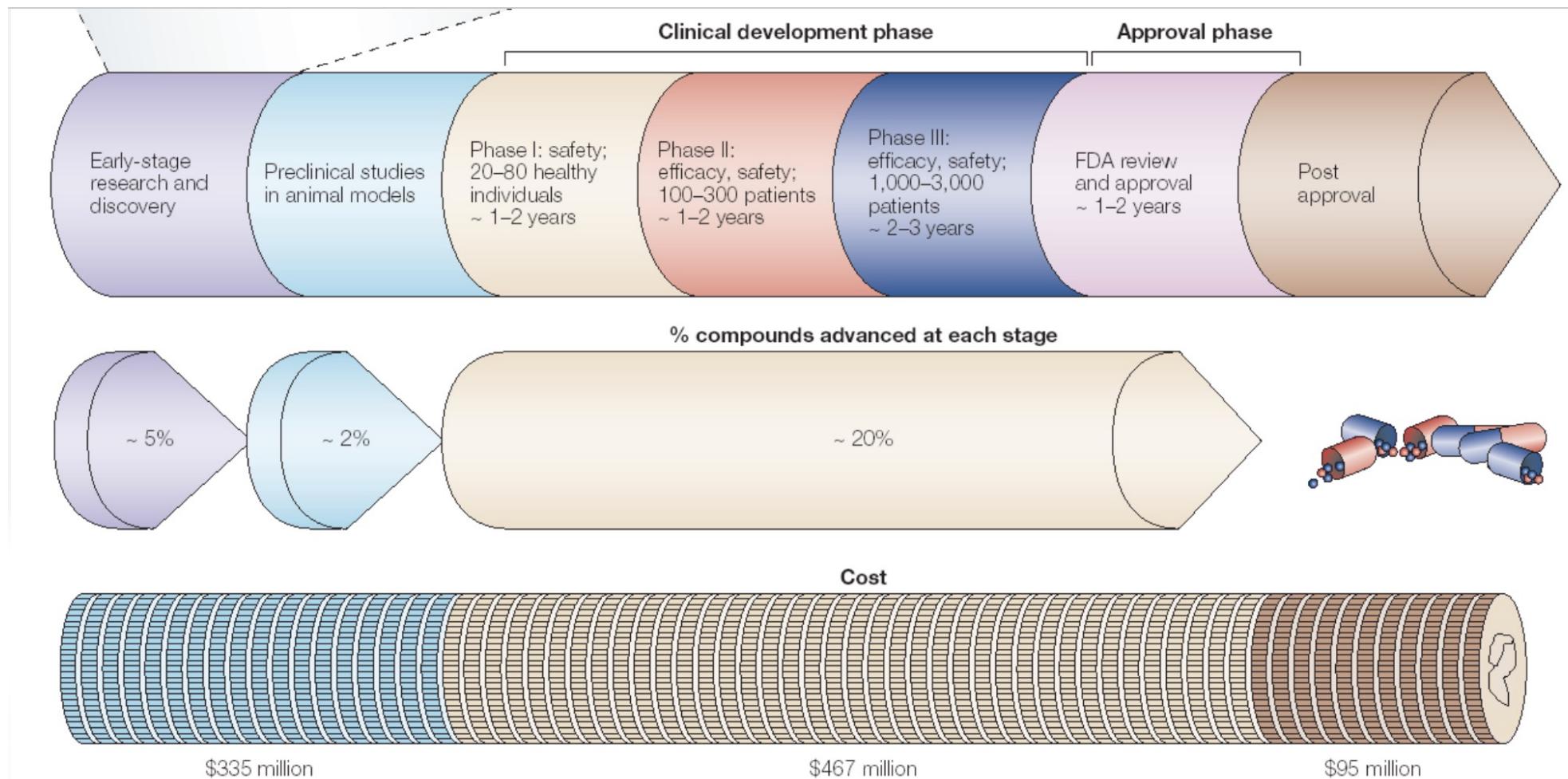
Investigación y desarrollo de nuevos medicamentos



Dos fases principales

Nat Rev Drug Discov 2004, 3, 360-4

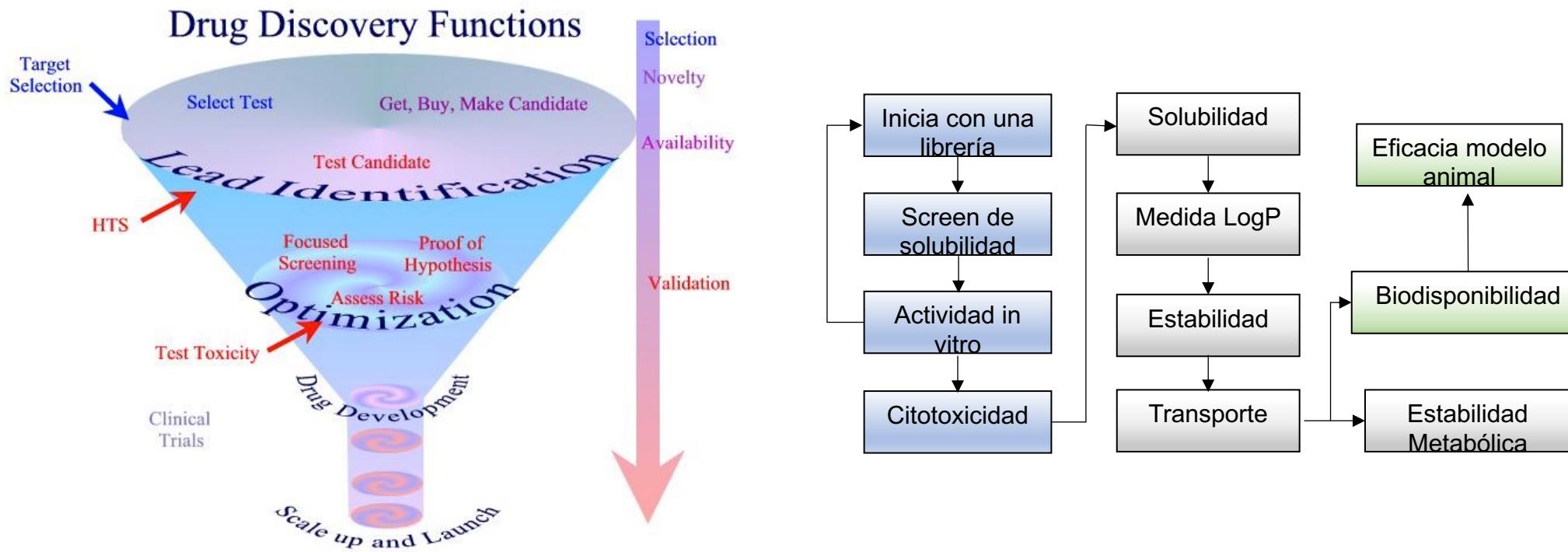
Costo para desarrollar un nuevo medicamento

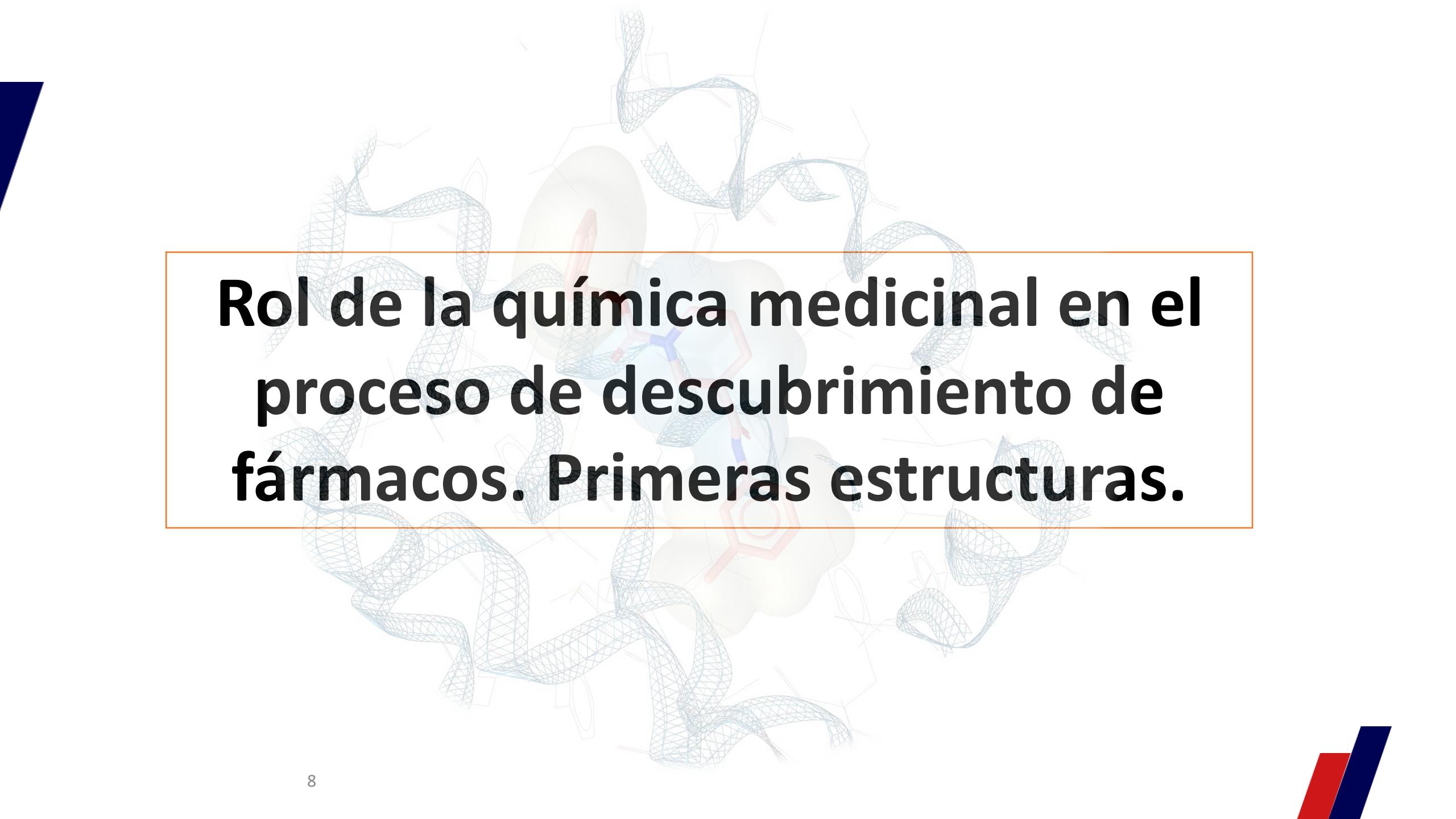


El costo aproximado de un programa de diseño de medicamentos es de US\$ 1000 millones por 8-15 años

Investigación preclínica

El descubrimiento de medicamentos es un proceso costoso y lento ... de 10000 compuestos obtenidos, solo uno llegará al mercado.





Rol de la química medicinal en el proceso de descubrimiento de fármacos. Primeras estructuras.

Descubrimiento: Inicios

Los seres humanos desde sus inicios han extraído de la naturaleza algunos principios activos que han contrarrestado los efectos lesivos de algunas patologías.

Aolgésicos



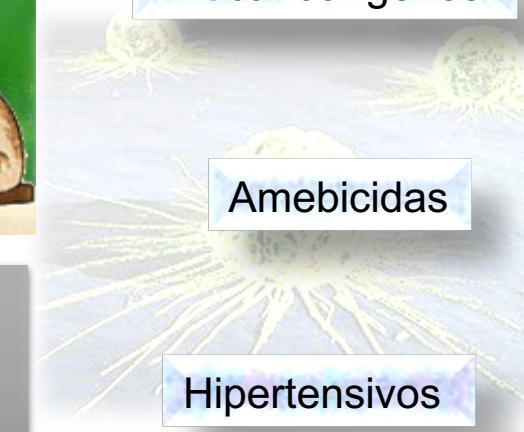
Antimicrobianos

Antivirales



Reconocimiento del poder curativo de los extractos vegetales en medicina etnobotánica

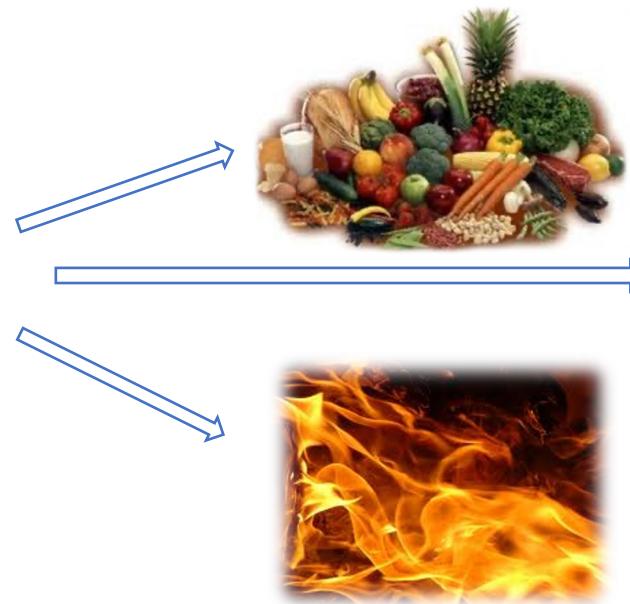
Anticancerígenos



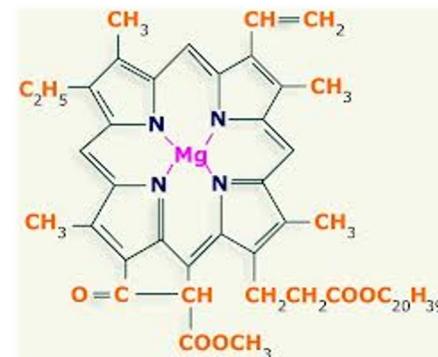
Amebicidas

Hipertensivos

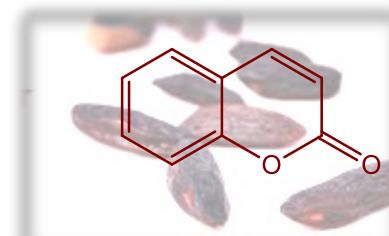
¿De dónde vinieron nuestras drogas?



Materias primas y plantas medicinales



Observación y estudio de sus transformaciones química



Descubrimiento de Drogas: Historia

Tercer milenio antes de Cristo: el ruibarbo, el opio y la ephedrina sínica.

En el caso de India estuvieron influenciados por la filosofía védica, por lo que la botánica tenía una función prominente. De acuerdo con la filosofía budista, los remedios eran dulces.



La cultura babilónica-asiria están los primeros indicios del uso de fármacos y detectaron la peligrosidad y dosificación de las sustancias administrándoselas a esclavos.



Papiro Ebers, describen la sintomatología y la prescripción para una enfermedad, así como los principios activos de plantas.



Operaciones farmaceuticas



Descubrimiento de Drogas: Historia

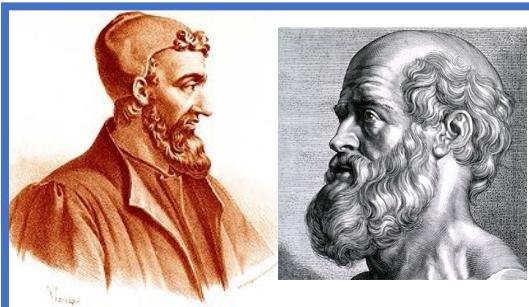


Proceso por el cual son descubiertas y/o diseñadas sustancias terapéuticamente activas para un determinado target molecular o ruta.

China: 5000 años atrás. Emperador Yan: identificación, de cientos de plantas medicinales y venenosas (probándolas consigo mismo) escribió el herbario nativo. Uso del ascaridol con propiedades antihelmínticas.



Experimentaba los efectos de los tóxicos con delincuentes convictos y consigo mismo, buscando un antídoto.



Galen (131-200 D.C): uso de productos de origen vegetal como elementos esenciales en la recuperación de la salud. Terra sigillata.

754 en Bagdad donde los árabes establecen las primeras farmacias, reguladas en el siglo IX.

Hipócrates: uso de compuestos metálicos en medicina. Té salicílico.

Descubrimiento de Drogas: Historia

America Precolombina: mezcla de prácticas empíricas y mágicas. El listado de plantas americanas se relaciona con el potencial biológico de regiones como la Amazonia que cuenta con 80.000 especies vegetales. Chamán.

Edad Media (476-1453): Bizancio conserva en fase de esterilidad la ciencia grecorromana. En este período se tienen los primeros indicios de la existencia de Farmacias como lugar físico en que se dispensan medicamentos

Renacimiento (1453-1600): Se inicia la ciencia moderna con la recuperación del mundo clásico. Vesalio, Paré y Paracelso hacen nuevas aportaciones a la anatomía humana, la cirugía y la química.

Barroco (1600-1740) Nace la ciencia moderna, por aportaciones de Bacon, Descartes y Galileo Galilei. En medicina destaca el nacimiento de la filosofía moderna y en terapéutica el auge de la yatroquímica (utilización de la química para hacer medicamentos).



Descubrimiento de Drogas: Historia



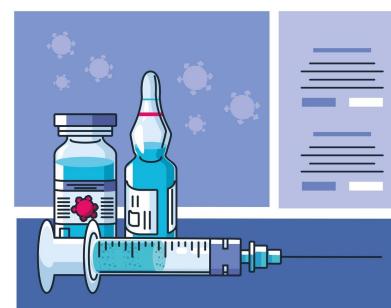
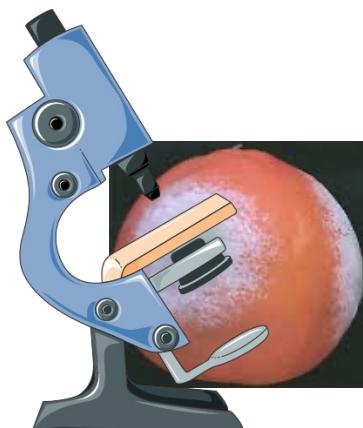
Paracelso (1493-1541): Uso de sales metálicas frente al uso de las plantas para combatir enfermedades como sífilis y bocio.
Uso del láudano analgésico (vino blanco, azafrán, canela y clavo).

"Sólo la dosis hace el veneno".

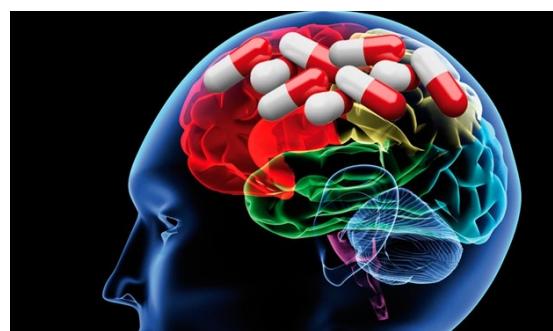


Romanticismo (1800-1848): Pasteur da un nuevo concepto de enfermedad: la microbiología médica. Nace la farmacología experimental y la química moderna que desarrolla el estudio de los gases y la Química Orgánica.
-Aislamiento de la morfina a partir del opio, en el año de 1806.

1848 y 1914



Alexander Wood



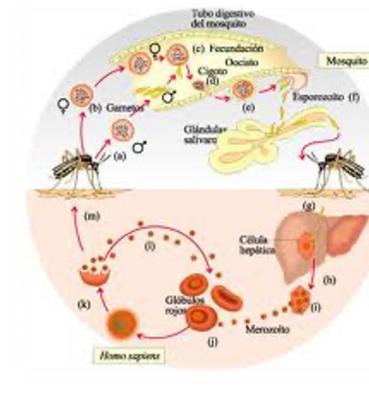
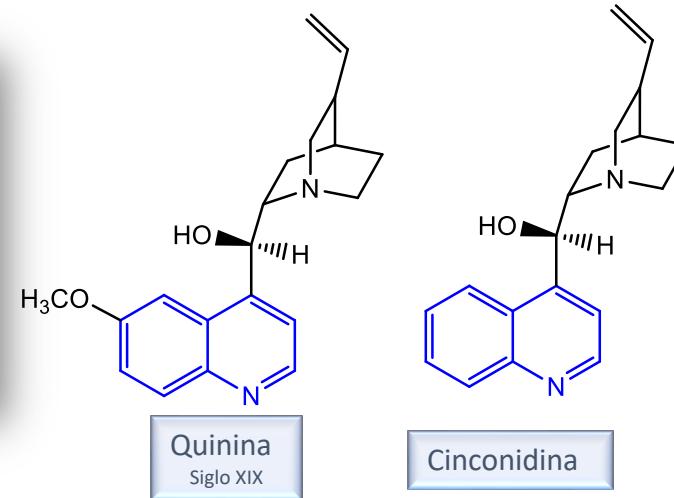
John Newport Langley



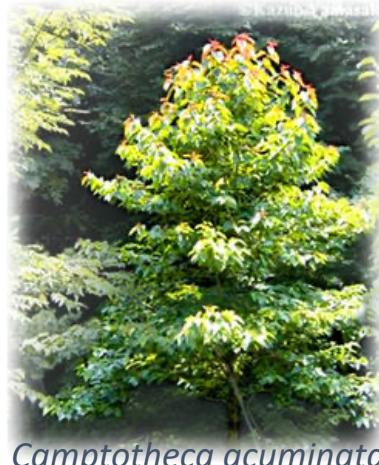
Fármacos en la naturaleza



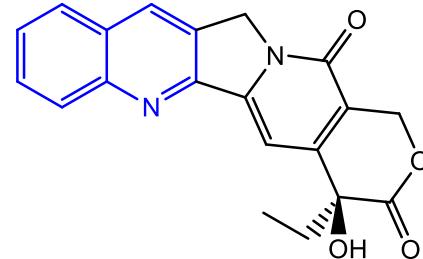
Cinchona officinalis



Importancia contra *Plasmodium falciparum*
Malaria



Camptotheca acuminata

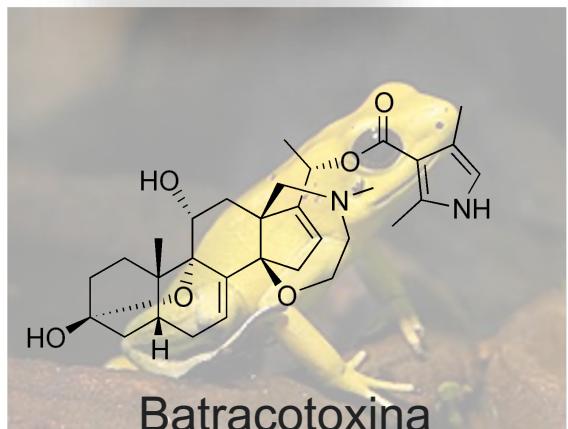
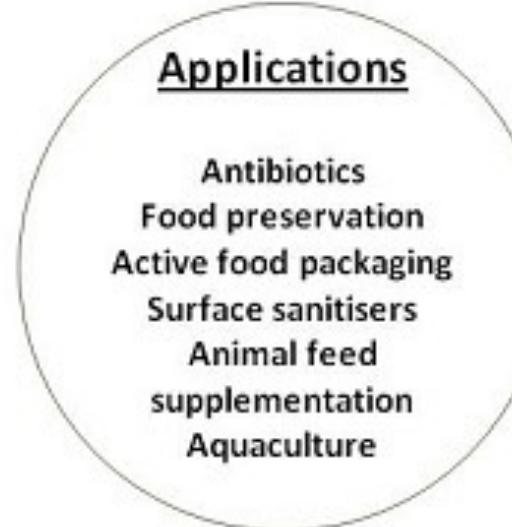
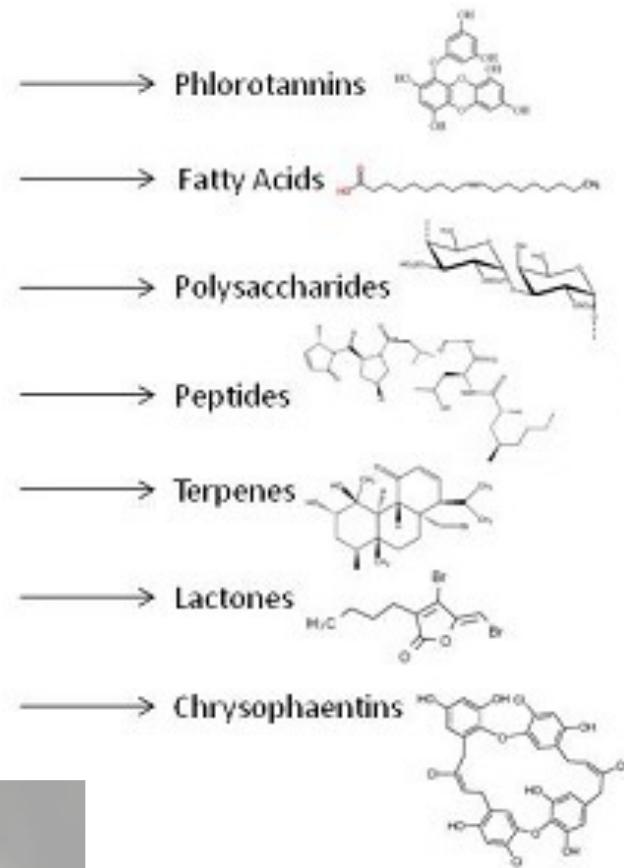
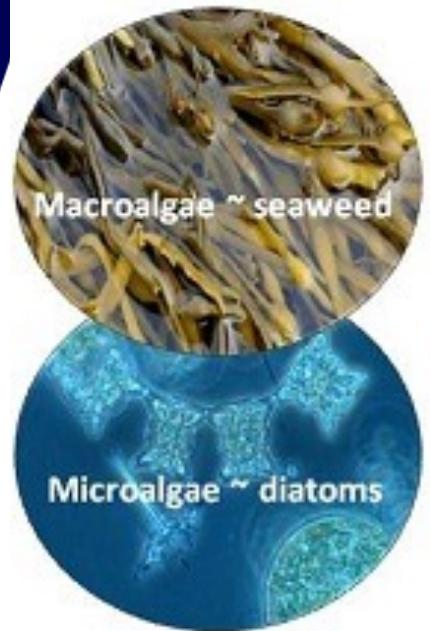


Camptotecina
Líneas de células cancerígenas

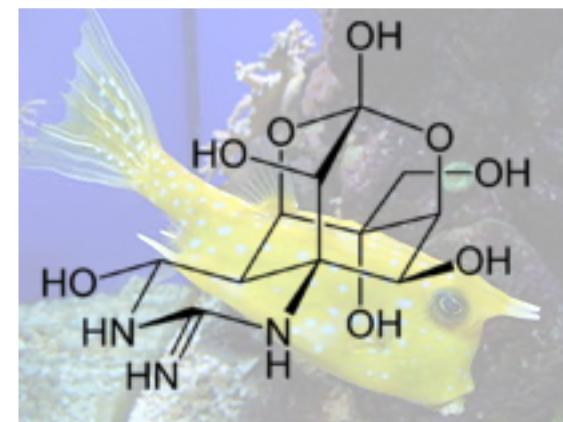


virantmicina,
antibiótico

Fármacos derivados de diversas fuentes



se superpone en parte con el sitio de acción de los anestésicos locales, antidepresivos y anticonvulsivos en el canal de Na⁺



la despolarización de membrana puede prevenirse o incluso revertirse con tetradotoxina

Fármacos derivados de diversas fuentes

- Source of Melitin, Phospholipases, Histamine, Catecholamine & Polyamines (apoptosis inducers)
- Induce apoptosis in prostate and ovarian cancer cells



- Induce apoptosis in ovarian cancer and various carcinomas and sarcomas



- Pancreastatin, an apoptosis inducer is effective against lymphomas and neuroblastomas



- Venom found selectively cytotoxic for human Hepatocellular carcinoma cells



venom



- Peptide Polybia-MP1 selectively inhibits cancer cells
- Effective against proliferation of prostate and bladder cancer cells
- Cytotoxic against leukemia T lymphocytes



Bufalin ,Cinobufagin and ABG from toad venom induce anti cancer effects

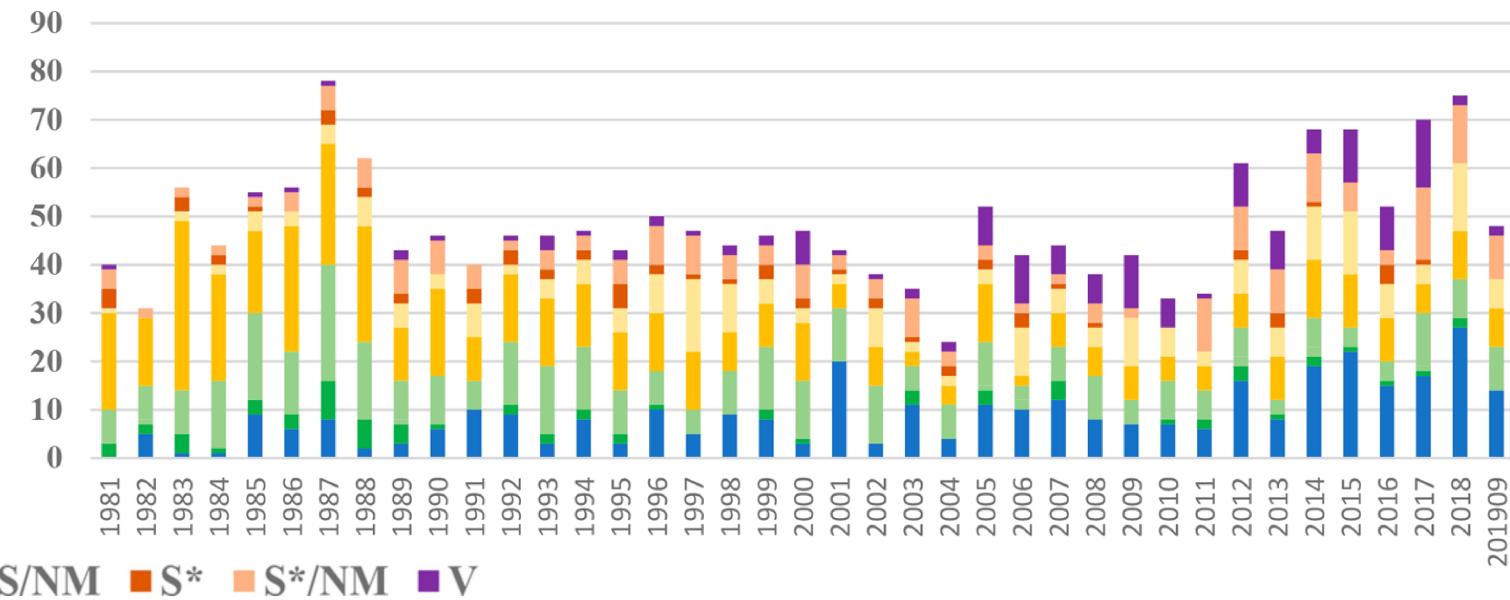
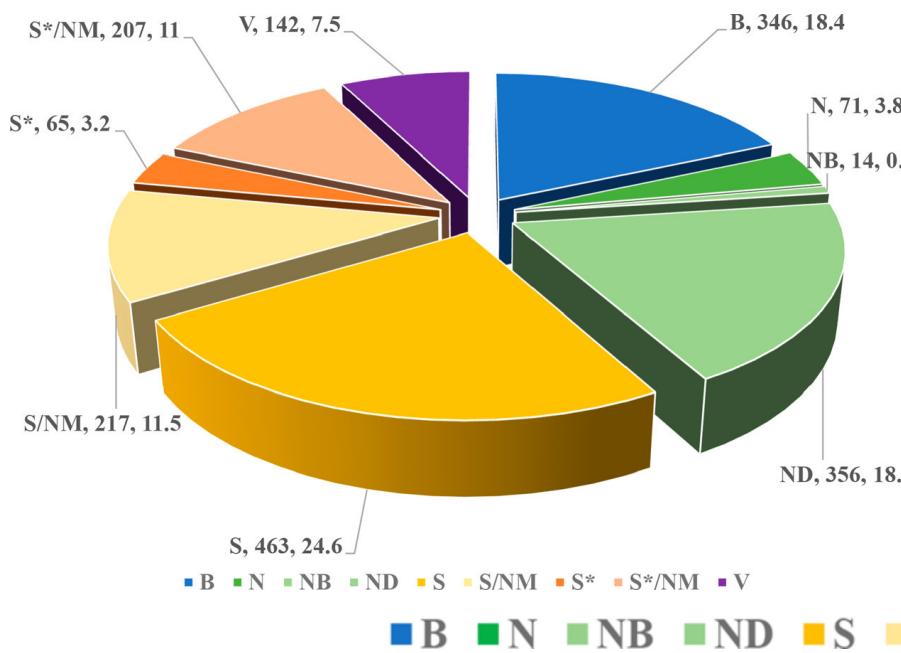
- ABG as the potent antitumor ingredient

hemilipina

- OD1 and ODK1, Chlorotoxin, Bengalin, Stoppin induce apoptosis,
- Inhibits growth of primary gliomas

Melittin
Mastoparan

Productos naturales como fuente de nuevas drogas



"B": Biológico; generalmente un péptido o proteína grande (> 50 residuos), ya sea aislado de un organismo / línea celular o producido por medios biotecnológicos en un huésped sustituto

"N": producto natural

"NB": Producto natural "Botánico" (en general estos han sido aprobados recientemente)

"ND": Derivado de un producto natural y suele ser una modificación semisintética.

"S": droga totalmente sintética, que a menudo se encuentra mediante la detección o modificación aleatoria de un agente existente

"S *": Elaborado por síntesis total, pero el farmacóforo es / era de un producto natural

"V": vacuna

"NM": imitación de producto natural

Nuevos fármacos

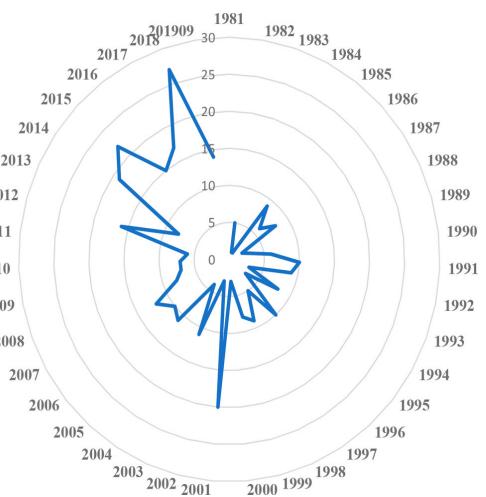
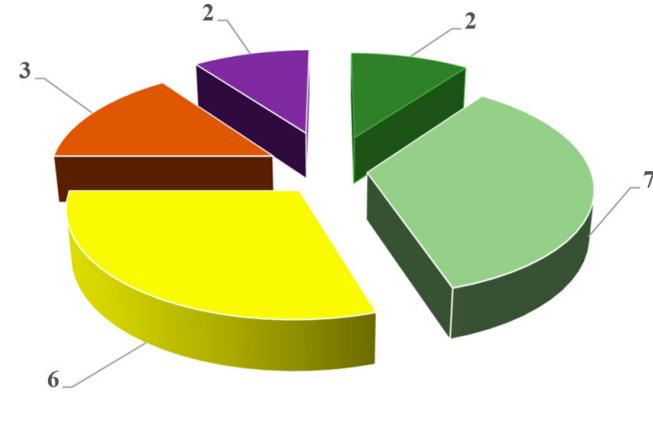


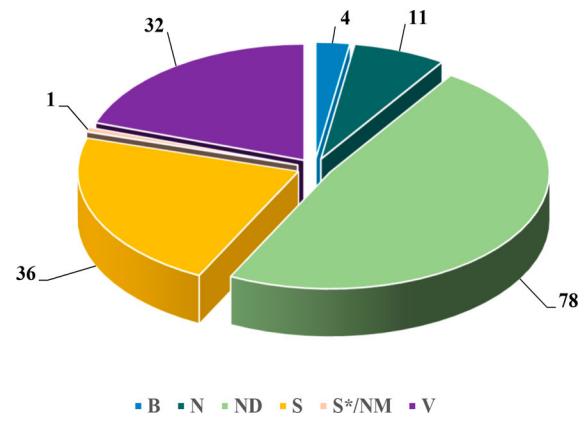
Gráfico de radar de todos los productos biológicos aprobados por año; $n = 346$.



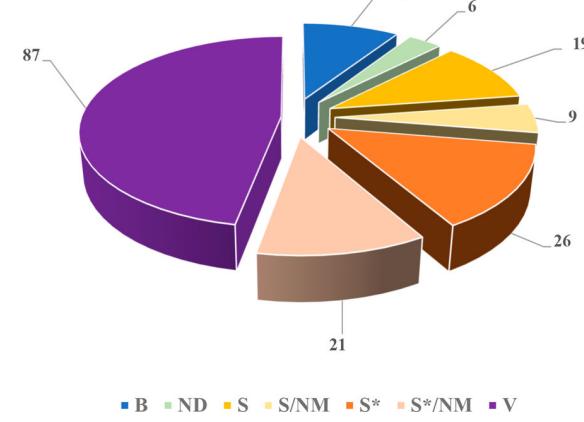
Fármacos antiparasitarios



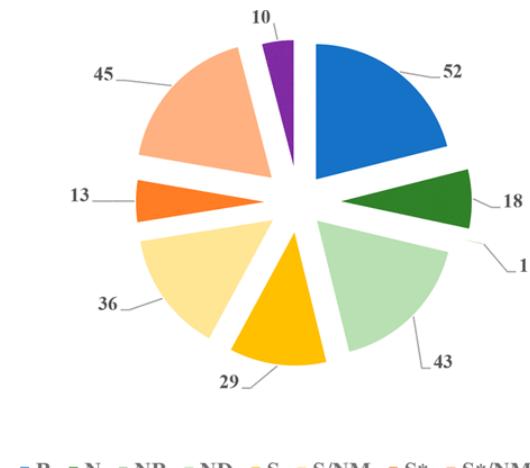
Fármacos antimicóticos



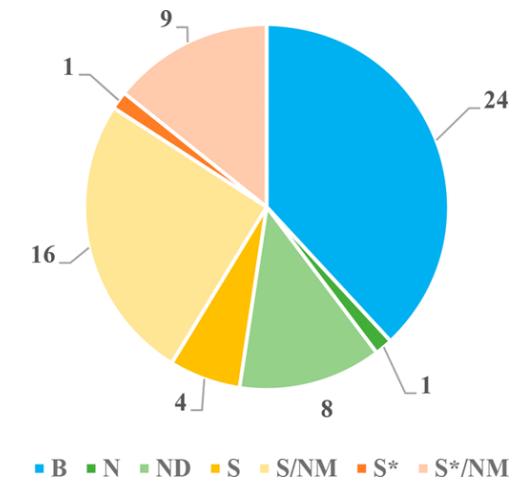
Medicamentos antibacterianos



Medicamentos antivirales

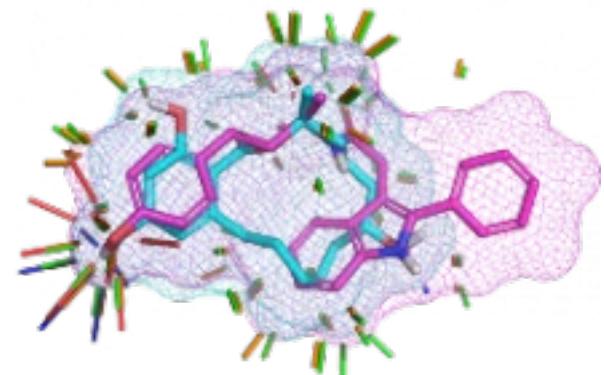


Fármacos anti-cáncer



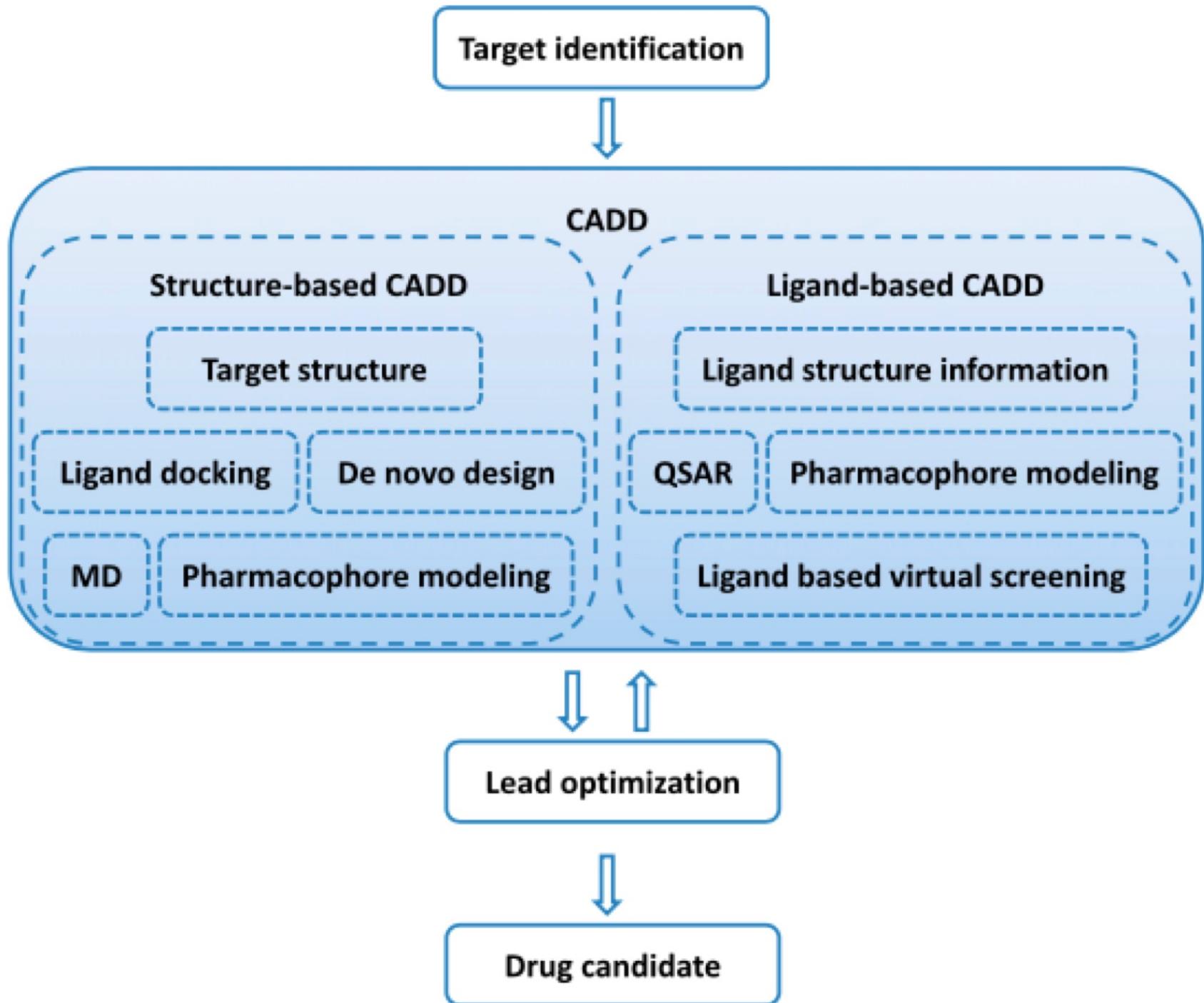
Fármacos anti-diabetes

Diseño de Fármacos Basado en el ligando (LBDD)



Computer-aided drug design (CADD)

	Known ligand(s)	No known ligand
Known protein structure	Structure-based drug design (SBDD) Protein-ligand docking	<i>De novo</i> design
Unknown protein structure	Ligand-based drug design (LBDD) <i>1 or more ligands</i> <ul style="list-style-type: none">• Similarity searching <i>Several ligands</i> <ul style="list-style-type: none">• Pharmacophore searching <i>Many ligands (20+)</i> <ul style="list-style-type: none">• Quantitative Structure-Activity Relationships (QSAR)	CADD of no use Need experimental data of some sort

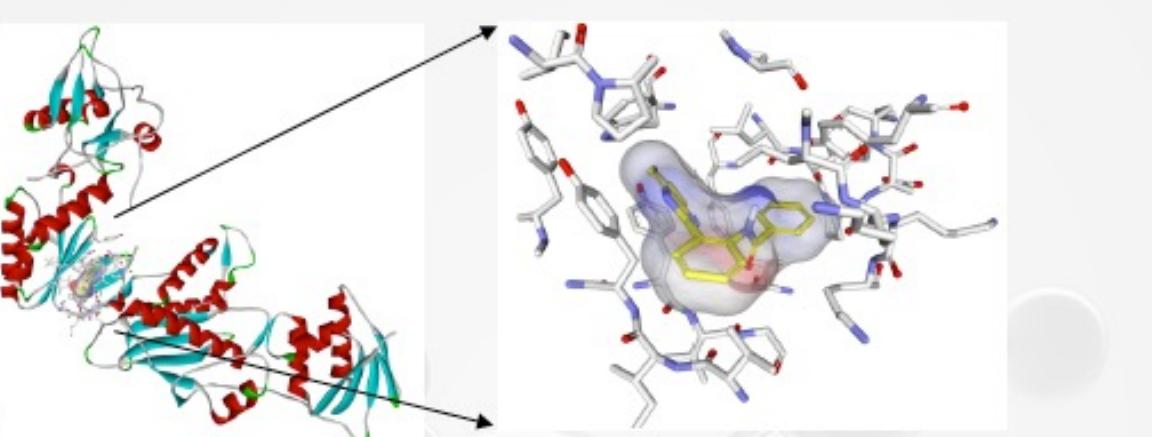


De Identificación Asistidos por computadora

Métodos Directos:

Estructura 3D del blanco terapéutico existe.

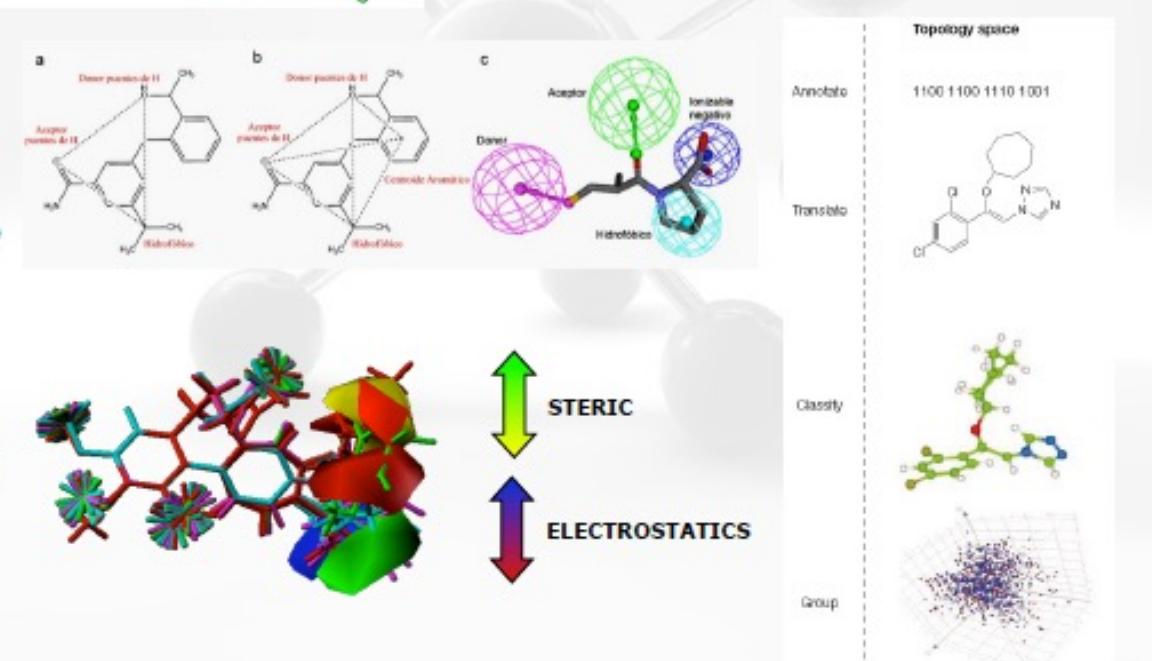
- **Structure-Based Drug Design SBDD**
- Rayos X or RMN
- Modelamiento comparativo de proteínas



Métodos Indirectos:

Estructura 3D del blanco terapéutico desconocida

- **Ligand-Based Drug Design LBDD**
- Diseño basado en análogos
- Relaciones Estructura-Actividad (QSAR)
- Farmacóforos
- Similaridad



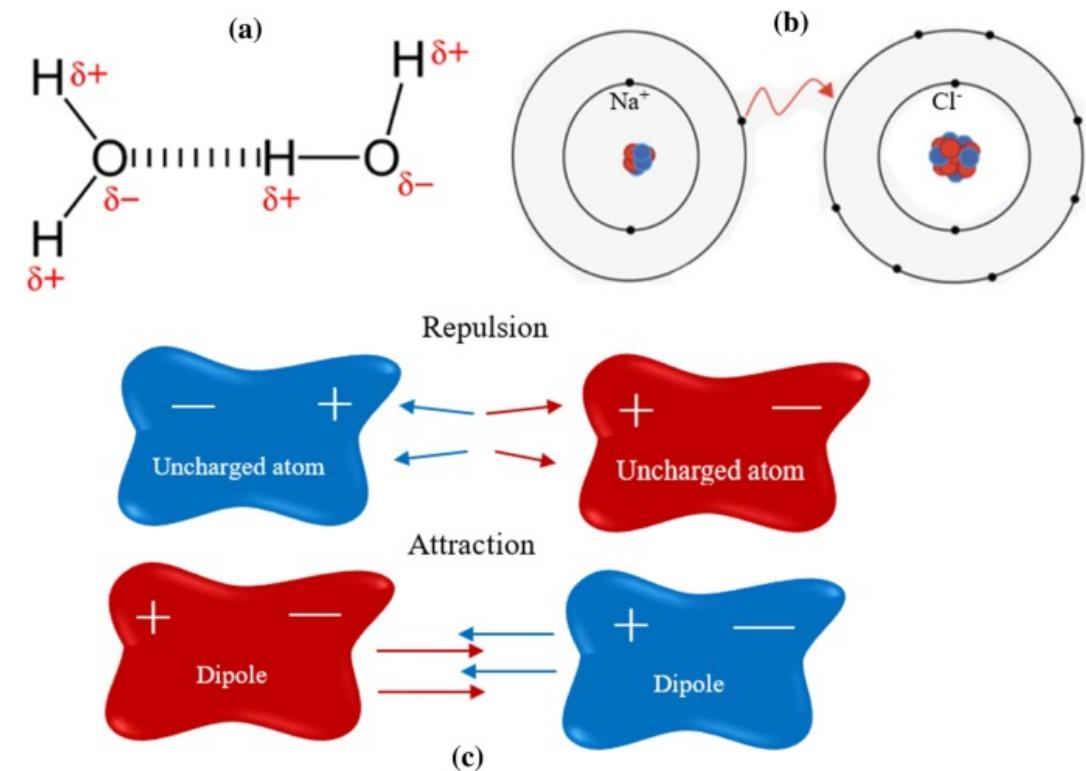
Bleicher, K. H.; Bohm, H. J.; Muller, K.; Alanine, A. I. Hit and lead generation: beyond high-throughput screening. *Nat Rev Drug Discov* 2003, 2, 369-78.

Sali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 1993, 234, 779-815.

Jorgensen, W. L. The many roles of computation in drug discovery. *Science* 2004, 303, 1813-8.

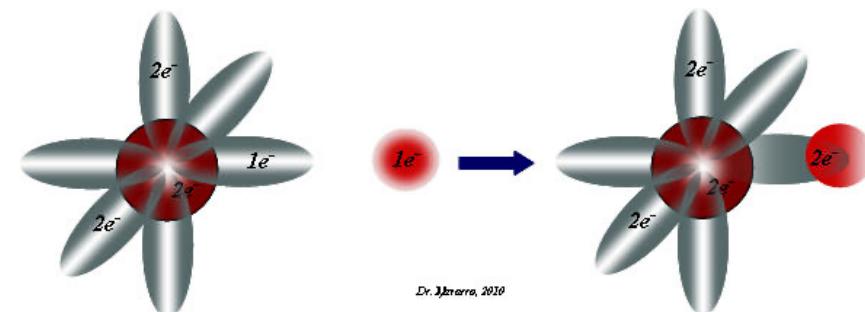
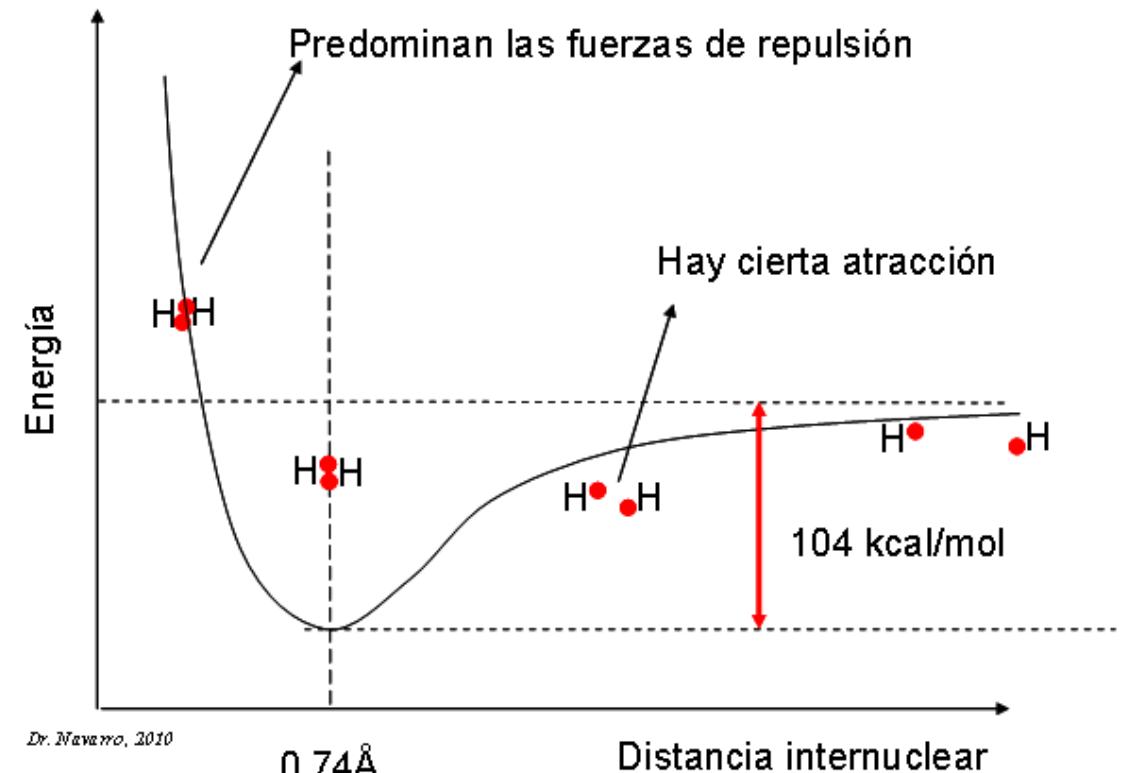
Base física de la unión de ligandos

- Unión no covalente enzimas/sustratos: La afinidad y especificidad de unión a menudo se obtienen mediante interacciones no covalentes.
- Las moléculas pequeñas pueden tener superficies de energía complejas, con grupos polares, no polares y polarizables, grados de libertad, estados de protonación múltiples, iones coenlazados, los cuales pueden reorganizarse al unirse.

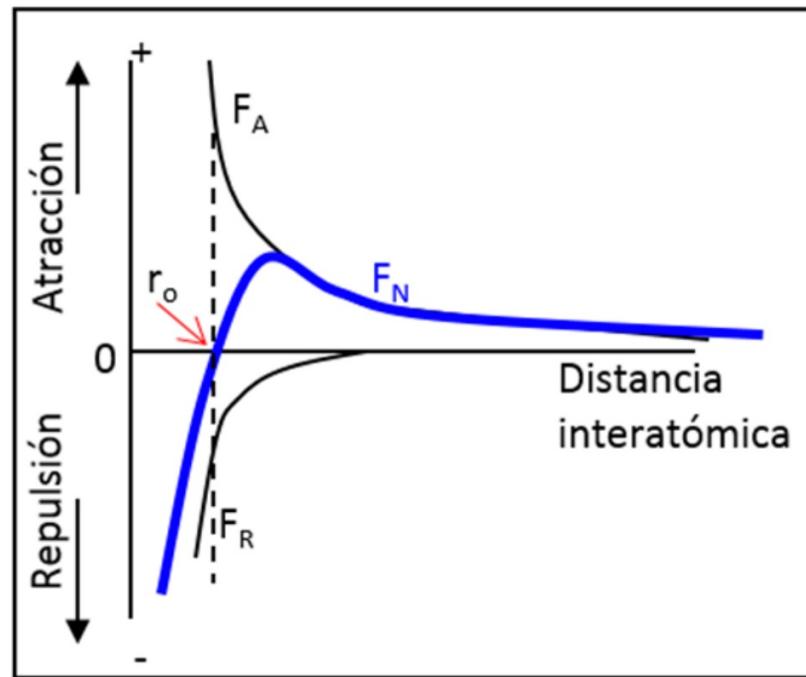


Estado de enlace

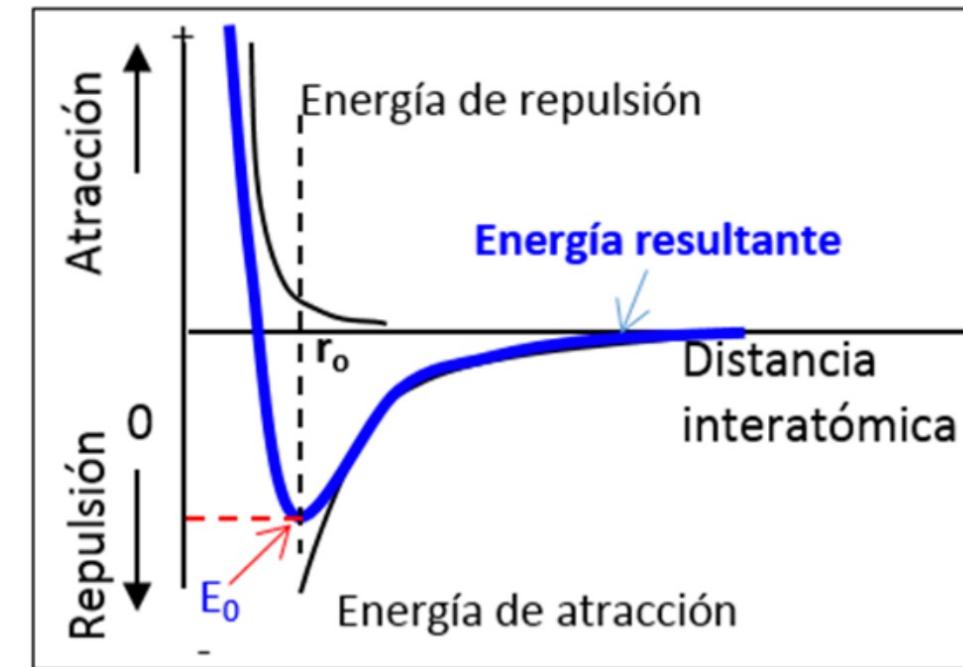
- Cuando dos átomos de elevada electronegatividad se unen, lo hacen compartiendo un par de electrones. Ese par situado entre los dos átomos ejerce una atracción sobre los núcleos de ambos que así se mantienen unidos.
- Las distancias de enlace quedan determinadas por las interacciones entre las fuerzas de atracción y de repulsión, lo cual es un problema energético.
- La diferencia de energía entre el estado de no interacción (cuando los dos átomos se encuentran a una distancia infinita) y el estado de máxima estabilidad, cuando se equilibraran las fuerzas de atracción y repulsión, es la energía de enlace.



Distancia de enlace



Cuando la fuerza resultante es cero tenemos un estado de equilibrio. A la distancia en que esto ocurre, r_0 , se alcanza el equilibrio. Esta es la distancia de enlace.

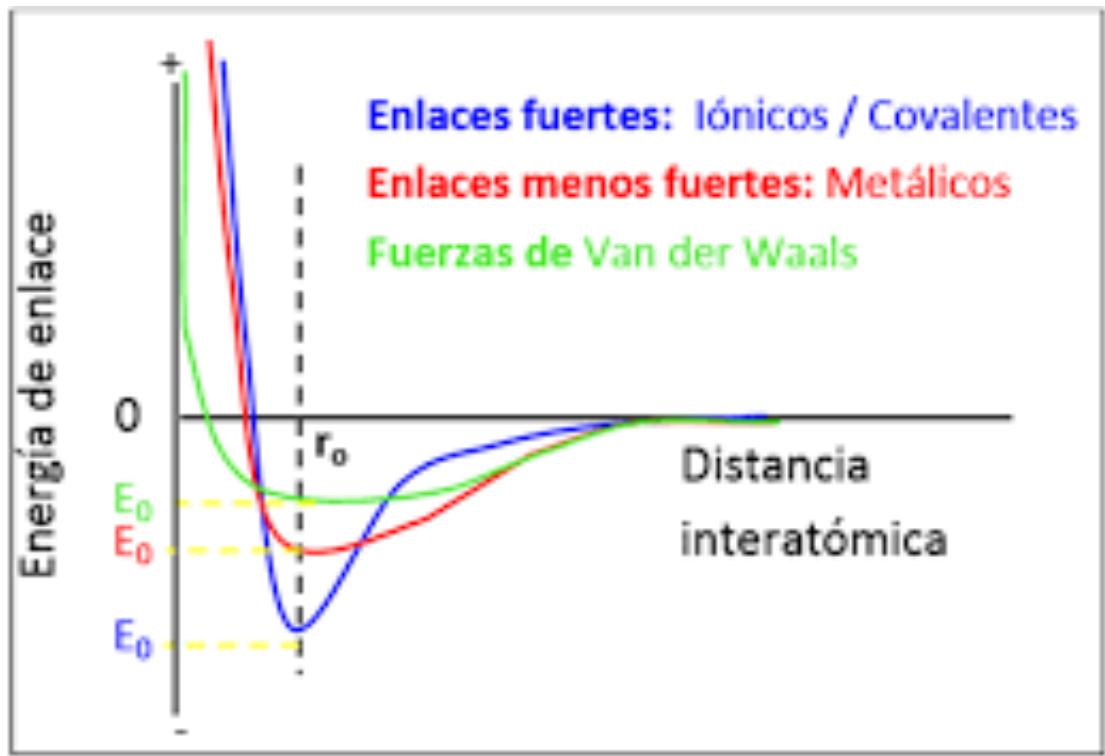


Fuerza y energía están relacionadas de modo que la energía es la derivada de la fuerza respecto de la distancia. Por tanto, cuando la FUERZA es CERO, corresponde a la mínima energía del sistema: es un pozo de energía potencial.

Tipos de enlace

- Los electrones son partículas con una onda asociada, en continuo movimiento alrededor del núcleo.
- Las distancias interatómicas de enlace y la energía es diferente.

Tipo de enlace	Energía de enlace (kcal/mol)
Iónico	150-370
Covalente	125-300
Metálico	25-200
Van der Waals	<10



VdW: El pozo de energía es pequeño y ancho, y por ello la distancia interatómica puede presentar un rango amplio.

Enlace en proteína-ligando

- Las conformaciones en las que el ligando está dentro de un bolsillo de unión bien definido se etiquetarían como "unido". * En algunos casos, el bolsillo de unión corresponderá a un pozo de energía profunda, de modo que las conformaciones de ligando cerca del límite del bolsillo tendrán un alto energías y pesos estadísticos bajos.
- Experimentalmente se mide una señal física, como la liberación de calor o la absorción de energía óptica, y debemos considerar qué conformaciones contribuyen a la señal experimental y usarlas como base para la comparación.

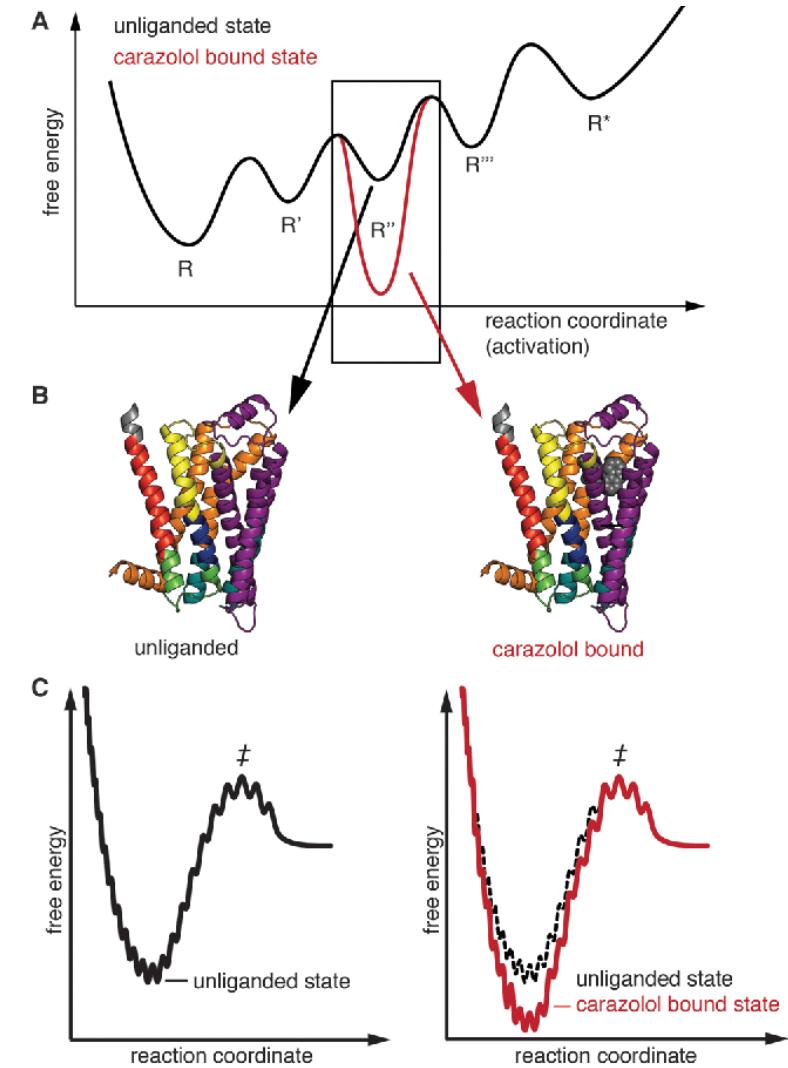
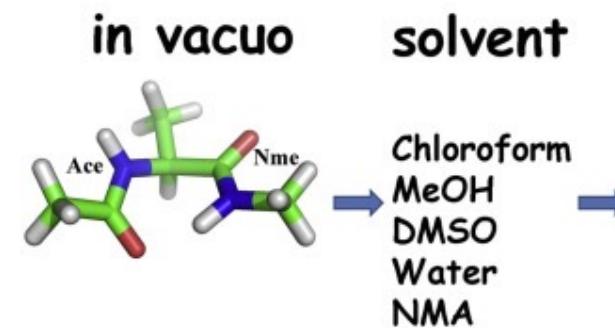
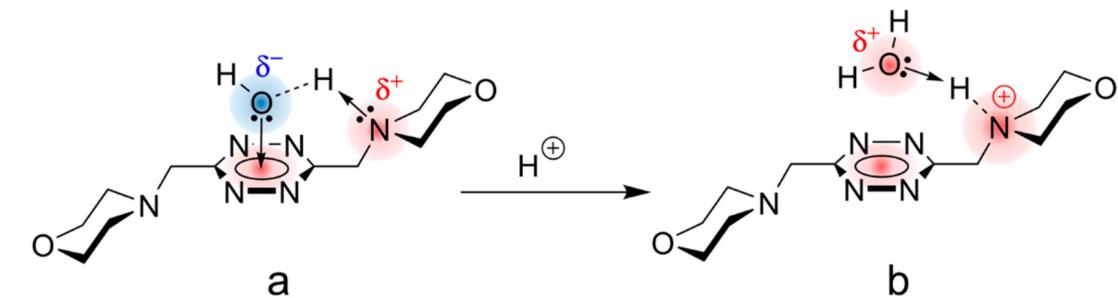
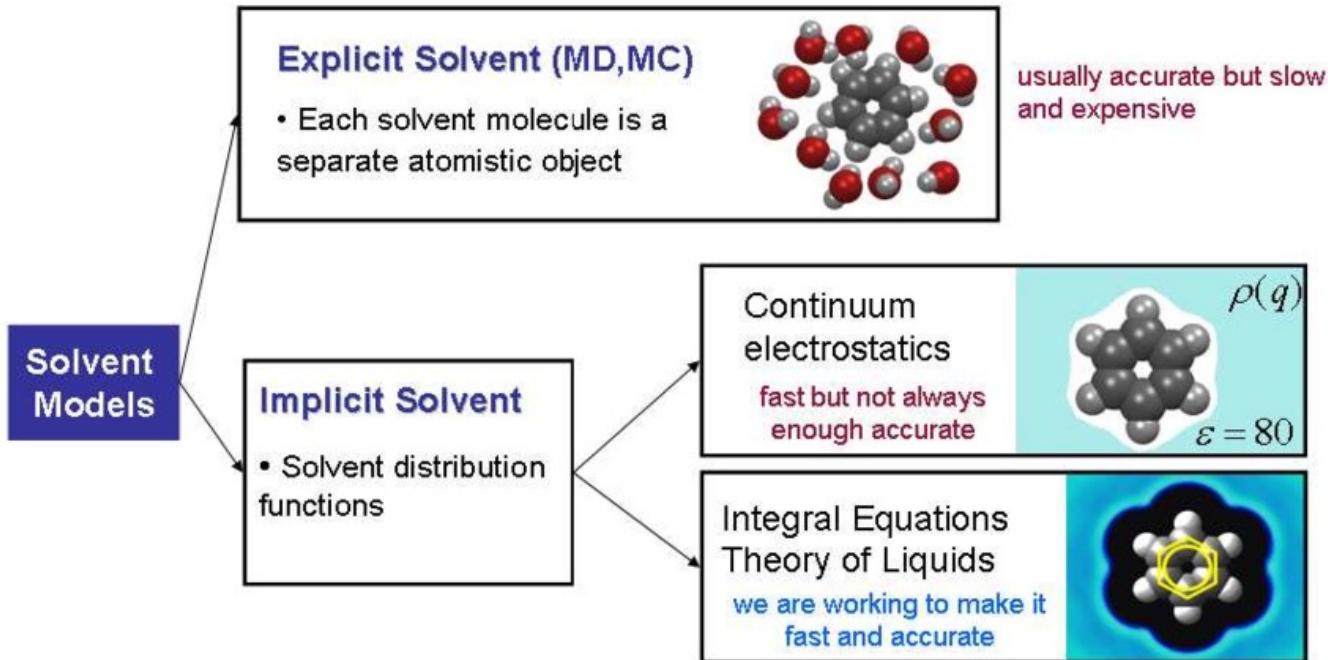


Fig. 2 Hypothetical energy landscape and pathways for the interaction

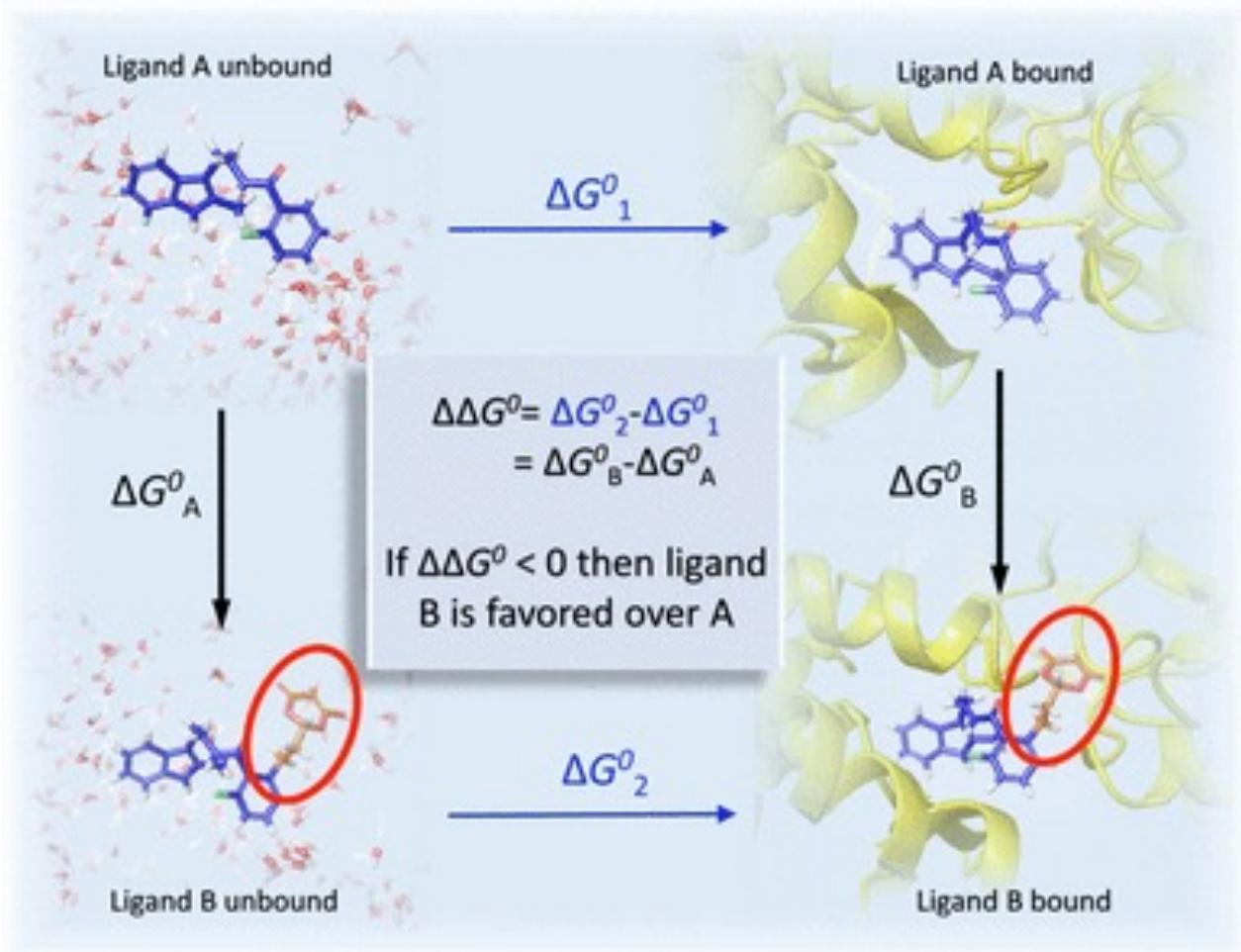
Comportamiento en disolvente

A medida que se insertan las nuevas cargas, el solvente se adapta o reorganiza, polarizándose por las nuevas cargas.

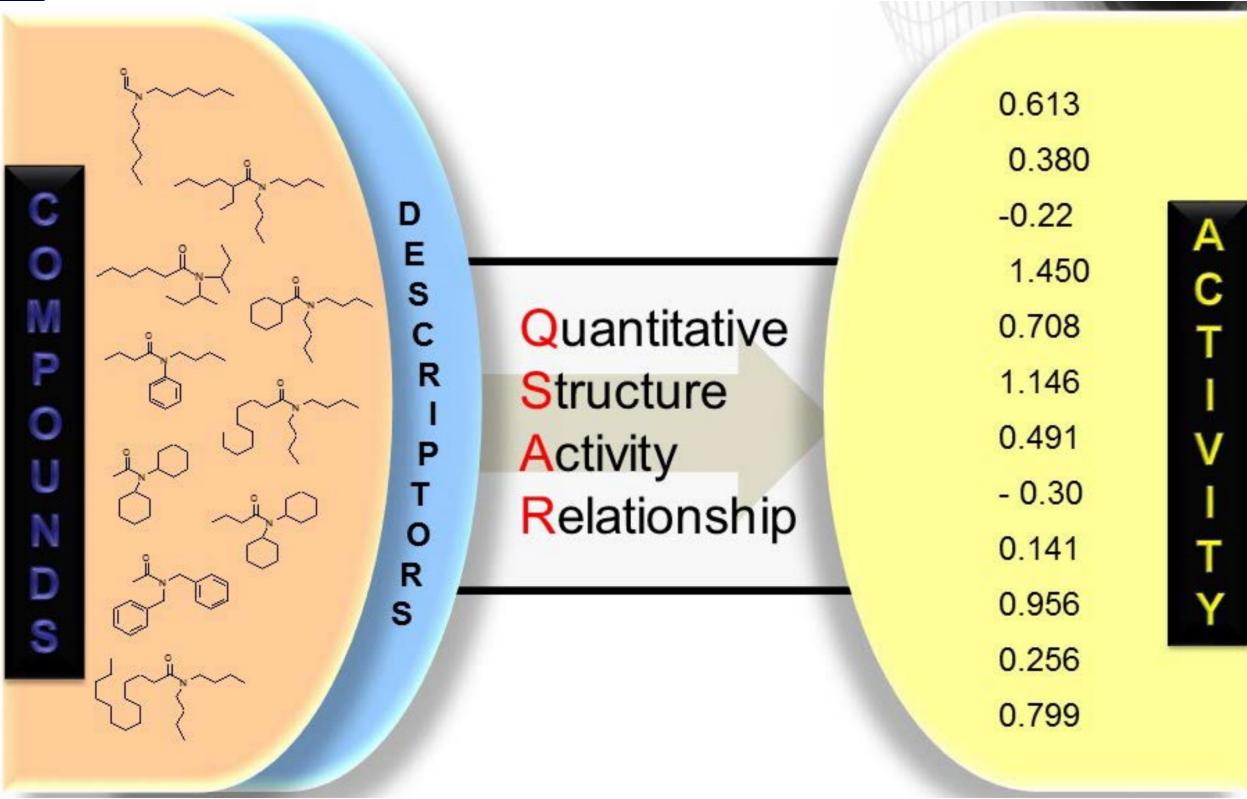


Ciclo termodinámico

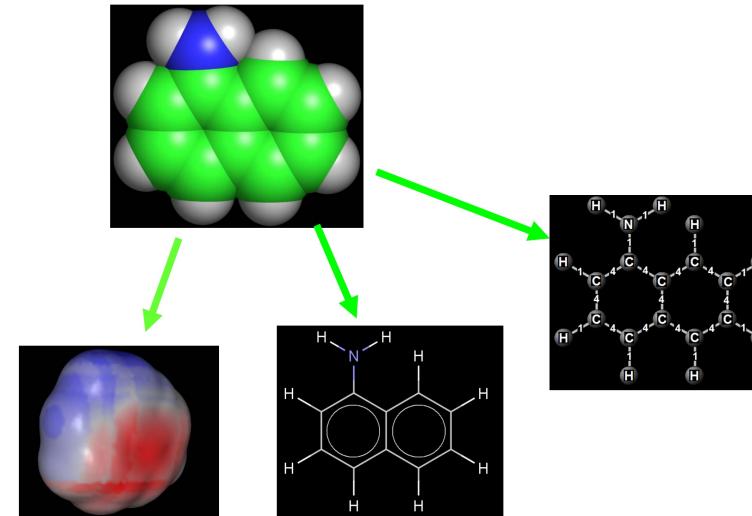
El cálculo de la energía libre de unión relativa del ligando B con respecto al ligando A, y modela el proceso de unión de cada ligando a la proteína. Ciclo termodinámico cerrado. En este ciclo, la diferencia en la energía libre entre la mutación del ligando A a B en el solvente y la misma mutación en el complejo, que es la energía libre de interés. La mutación del ligando A a B, que es la adición del sustituyente naranja, es factible computacionalmente. Si esta diferencia en la energía libre de unión es negativa, se espera que la unión del ligando B sea más favorable que la de A.



Métodos basados en relación estructura actividad (SAR y QSAR).



Representación de la estructura



Vertices
(atomic type,
coordinates etc.)

16	15	0	0	0	0	0	0	0.0999	V2000
-0.1958	-2.9667	0.0000	C	0	0	0	0		
0.5167	-2.5500	0.0000	C	0	0	0	0		
0.5125	-1.7250	0.0000	O	0	0	0	0		
1.2292	-2.9625	0.0000	N	0	0	0	0		
1.9417	-2.5458	0.0000	C	0	0	0	0		
2.6542	-2.5983	0.0000	C	0	0	0	0		
3.3667	-2.5417	0.0000	C	0	0	0	0		
4.0792	-2.9542	0.0000	C	0	0	0	0		
4.7917	-2.5375	0.0000	C	0	0	0	0		
5.5042	-2.9500	0.0000	C	0	0	0	0		
1.2250	-3.7875	0.0000	C	0	0	0	0		
0.8083	-4.5000	0.0000	C	0	0	0	0		
1.3917	-5.0833	0.0000	C	0	0	0	0		
0.9750	-5.7958	0.0000	C	0	0	0	0		
1.5583	-6.3792	0.0000	C	0	0	0	0		
0.9708	-6.9625	0.0000	C	0	0	0	0		
8	9	1	0	0	0	0			
4	5	1	0	0	0	0			
9	10	1	0	0	0	0			
2	3	2	0	0	0	0			
4	11	1	0	0	0	0			
5	6	1	0	0	0	0			
11	12	1	0	0	0	0			
1	2	1	0	0	0	0			
12	13	1	0	0	0	0			
6	7	1	0	0	0	0			
13	14	1	0	0	0	0			
2	4	1	0	0	0	0			
14	15	1	0	0	0	0			
7	8	1	0	0	0	0			
15	16	1	0	0	0	0			
M	END								

Vertices
(atomic type,
coordinates etc.)

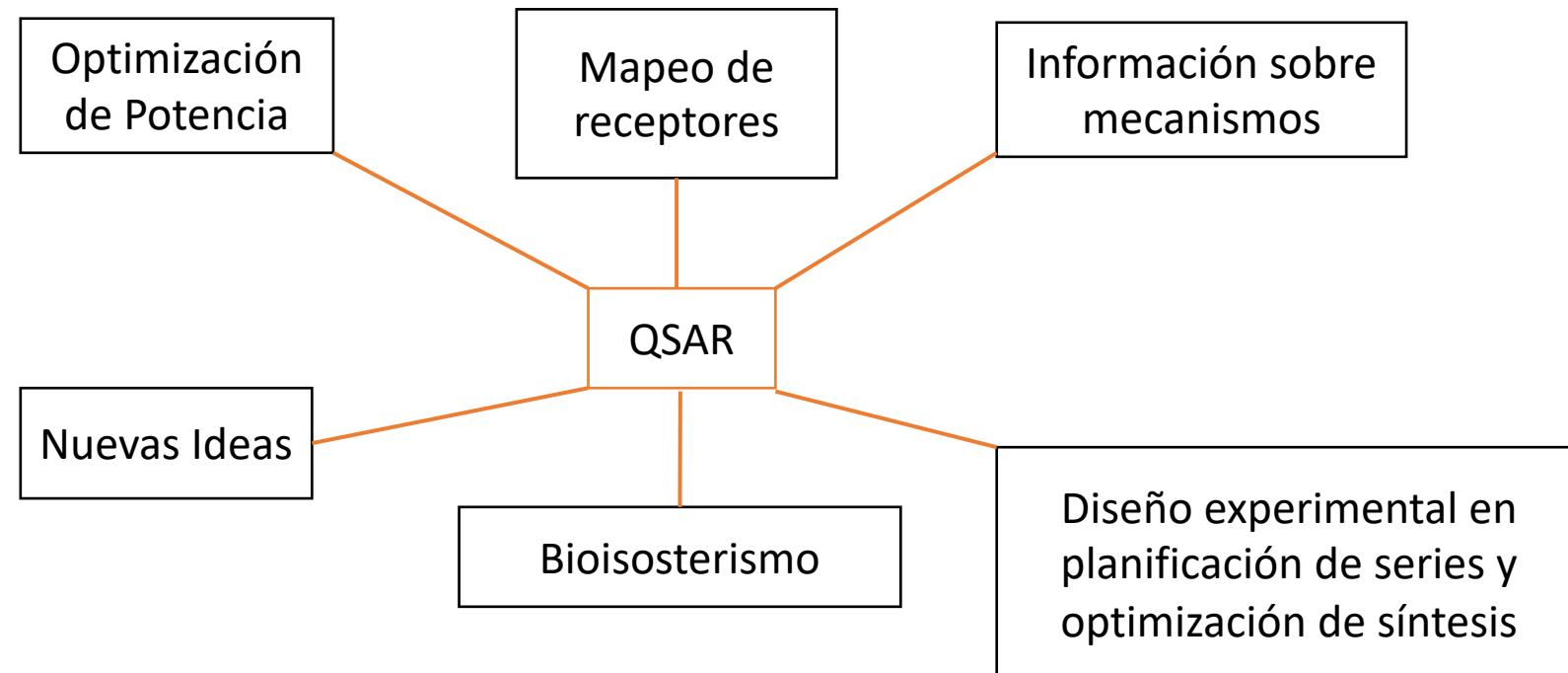
Edges
(connectivity table,
label-types of bonds)

Los gráficos se utilizan ampliamente para representar y diferenciar estructuras químicas, donde los átomos son vértices y los enlaces se expresan como aristas que conectan estos vértices.

Los gráficos moleculares permiten el cálculo de numerosos índices para compararlos cuantitativamente.

Relación cuantitativa estructura-actividad (QSAR)

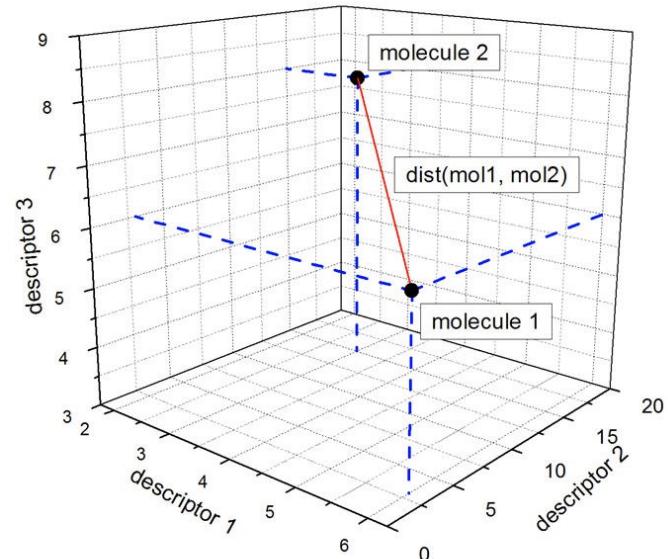
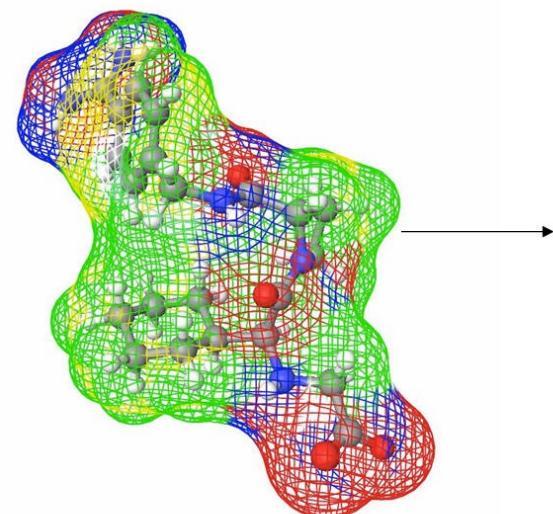
Métodos para correlacionar cuantitativamente características estructurales con actividad biológica. son herramientas importantes en la bio-quim-informática. Estas relaciones se pueden construir esencialmente basándose en datos generados a partir de modelación molecular y química computacional.



QSAR

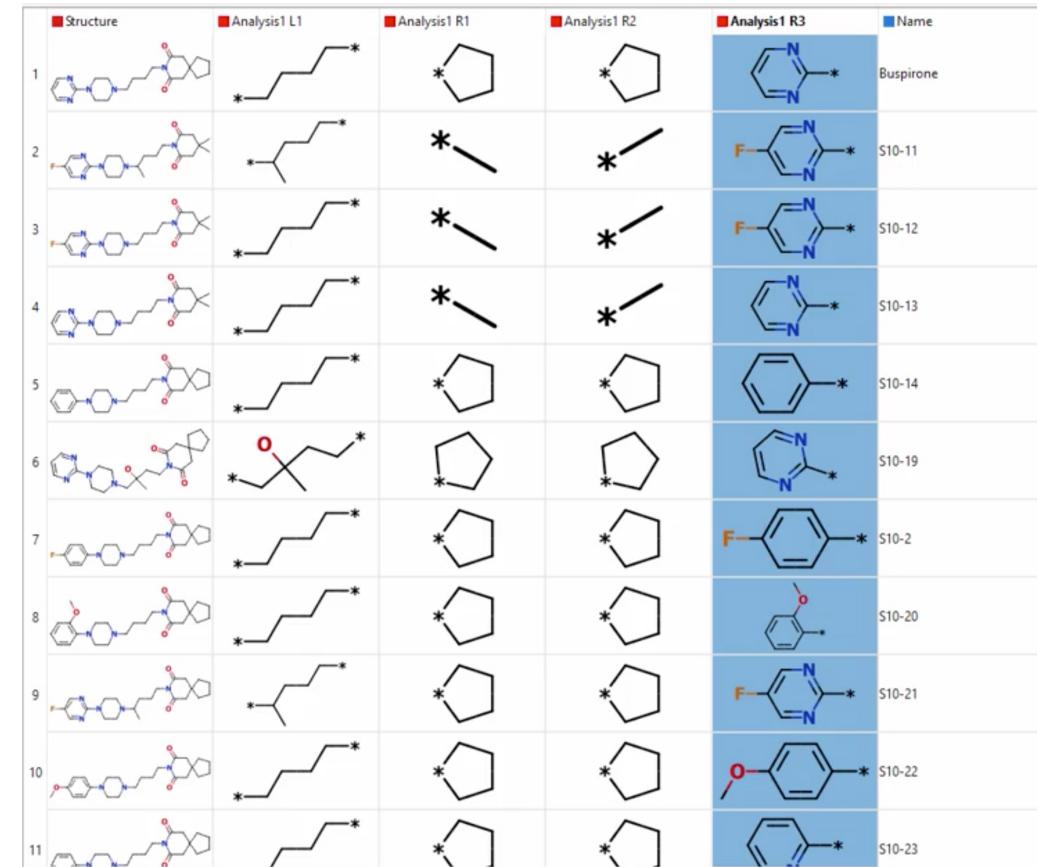
QSAR QSAR se puede definir como un conjunto de métodos que intentan encontrar una relación matemática entre un conjunto de moléculas y su actividad biológica.

Los descriptores se pueden derivar de datos experimentales o computacionales. Utilizando análisis de regresión, se puede obtener una relación matemática entre los descriptores químicos y la actividad.



La finalidad de QSAR es construir un modelo que permite predecir si un compuesto nuevo:

- será activo o inactivo
- tendrá una actividad x nM
- será selectivo
- proveerá nueva información
- tiene propiedades farmacocinéticas (ADME) y nivel de toxicidad Adecuados
- mejora nuestro conocimiento de las relaciones estructura actividad y permite mejorar nuestro modelo



Electrostáticos:

- ❖ Gasteiger-Marsili empirical atomic partial charges
- ❖ Zefirov's empirical atomic partial charges
- ❖ Mulliken atomic partial charges
- ❖ Minimum (most negative) and maximum (most positive) atomic partial charges
- ❖ Polarity parameters
- ❖ Dipole moment
- ❖ Molecular polarizability
- ❖ Molecular hyperpolarizability
- ❖ Average ionization energy
- ❖ Minimum electrostatic potential at the molecular surface
- ❖ Maximum electrostatic potential at the molecular surface
- ❖ Local polarity of molecule
- ❖ Total variance of the surface electrostatic potential
- ❖ Electrostatic balance parameter

Químico-cuánticos:

- ☒ Total energy of the molecule
- ☒ Total electronic energy of the molecule
- ☒ Standard heat of formation
- ☒ Electron-electron repulsion energy for a given atomic species
- ☒ Nuclear-electron attraction energy for a given atomic species
- ☒ Electron-electron repulsion between two given atoms
- ☒ Nuclear-electron attraction energy between two given atoms
- ☒ Nuclear repulsion energy between two given atoms
- ☒ Electronic exchange energy between two given atoms
- ☒ Resonance energy between given two atomic species
- ☒ Total electrostatic interaction energy between two given atomic species
- ☒ Total interaction energy between two given two atomic species
- ☒ Total molecular one-center electron-electron repulsion energy
- ☒ Total molecular one-center electron-nuclear attraction energy
- ☒ Total intramolecular electrostatic interaction energy
- ☒ Electron kinetic energy density

basados en DFT

Global Descriptors

^{DFT} Chemical Potential

^{DFT} Chemical Hardness

^{DFT} Softness

^{DFT} Electrophilicity Index

Local Descriptors

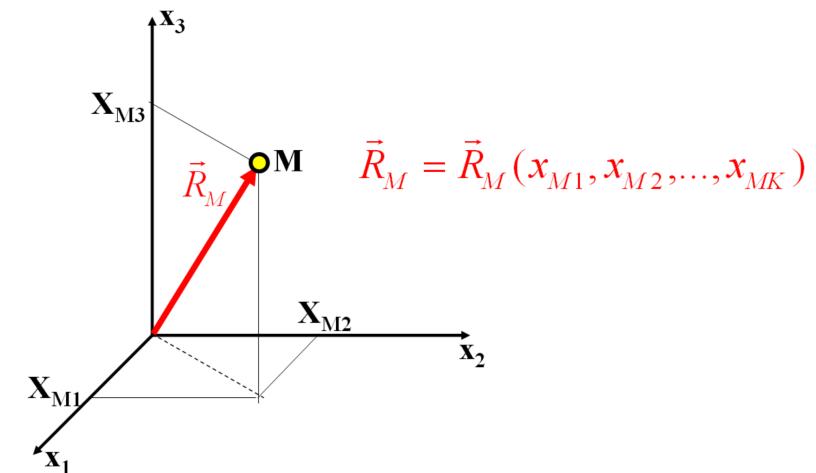
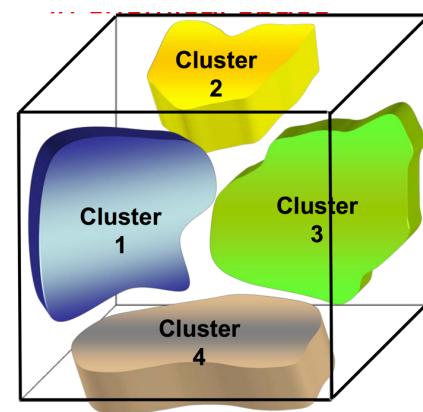
^{DFT} Condensed Fukui Function

^{DFT} Polarizability

^{DFT} Charge transfer

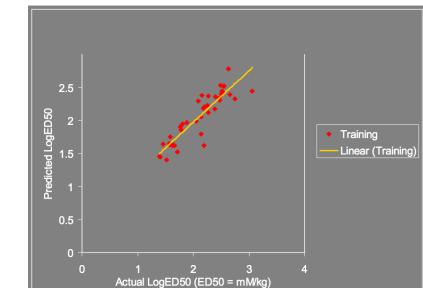
Modelado QSAR

Las moléculas se consideran vectores en el espacio de descriptores (espacio «químico»). Las dimensiones de este espacio corresponden al número de descriptores. Se emplean métodos de agrupación para analizar distancias entre compuestos e identificar agrupaciones.

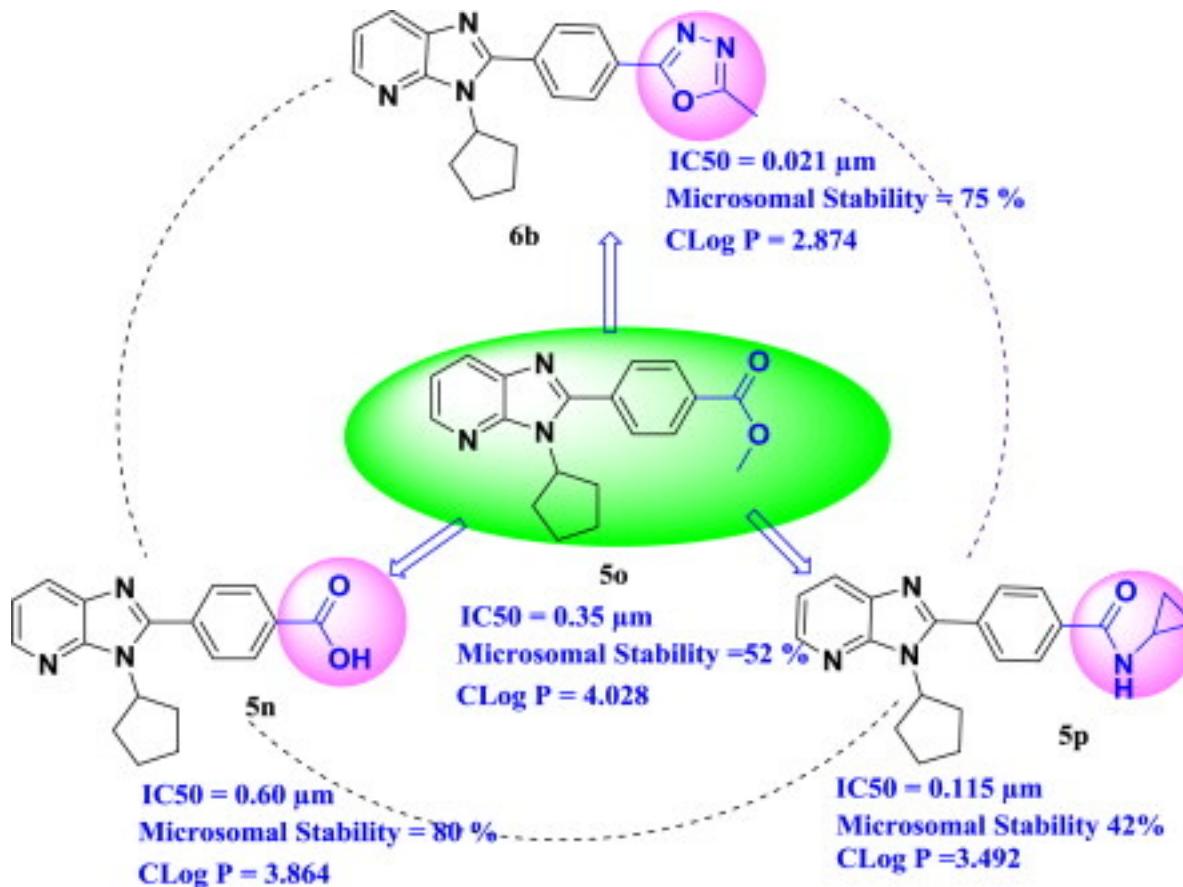


Chemistry	Bioactivity (IC50, Kd...)	Cheminformatics (Molecular Descriptors)
Comp.1	Value1	D ₁
Comp.2	Value2	D ₂
Comp.3	Value3	D ₃
Comp.N	ValueN	D _n

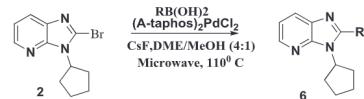
BA = F(D) (linear,
e.g., -LogIC50 = k₁D₁+k₂D₂+...+k_nD_n)
or non-linear, e.g. k nearest neighbors



Estudios SAR de isósteros de purina y su potencial como agentes citotóxicos



The IC₅₀ values (μ M) of compounds **6a–j** to the resistant breast cancer cell line (MDAMB-231).



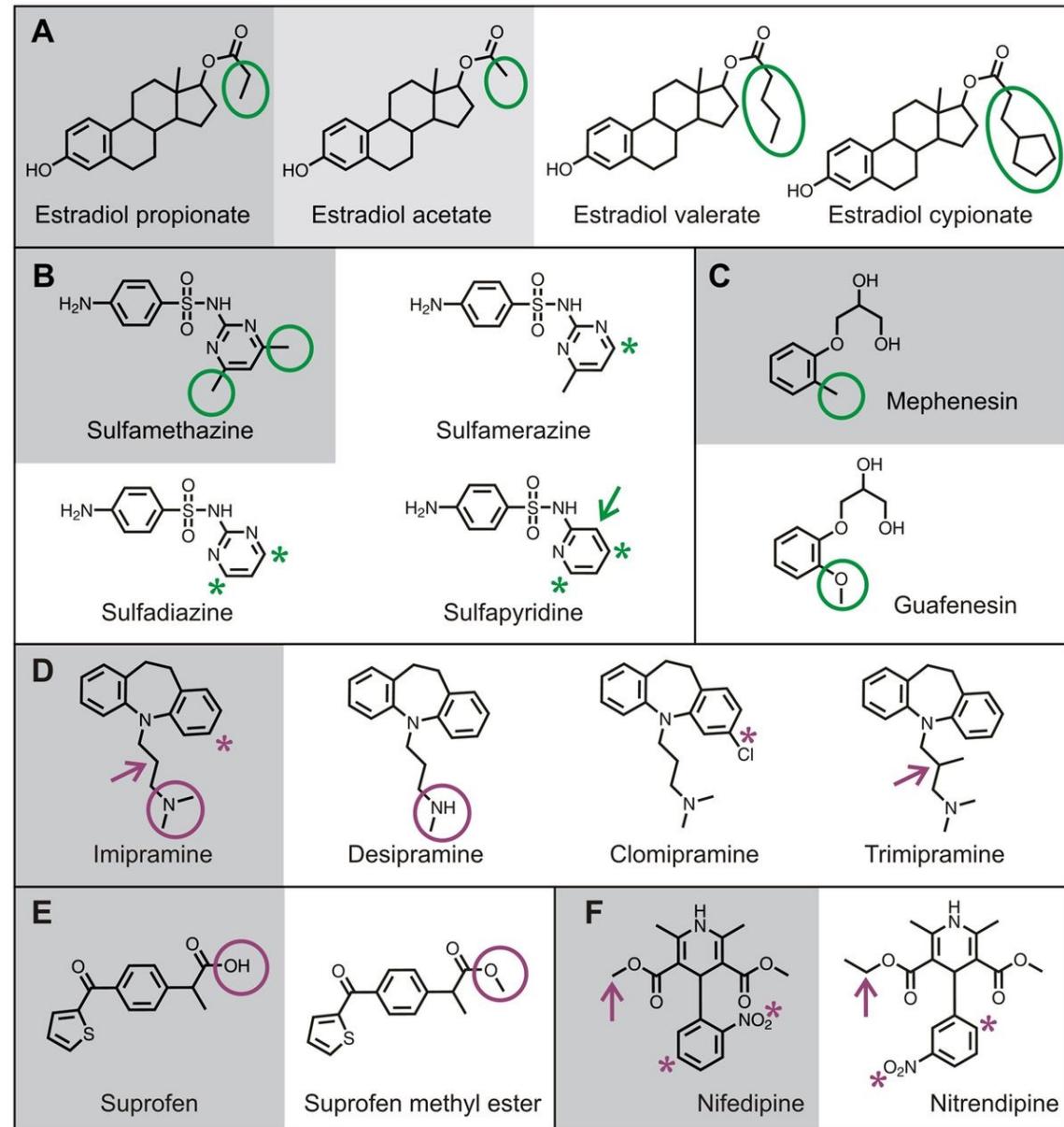
Compound	R ₂	Cytotoxicity IC ₅₀ (μ M) ^a	CLogP ^b
6a		>10	2.778
6b		0.021 ± 0.006	2.874
6c		0.045 ± 0.009	2.573
6d		0.138 ± 0.012	4.140
6e		>10	2.979
6f		0.348 ± 0.034	4.414
6g		0.07 ± 0.01	3.314
6h		0.219 ± 0.015	3.442
6i		0.513 ± 0.029	4.201
6j		0.09 ± 0.012	4.511
6k		0.765 ± 0.042	3.548
6l		>10	4.446
6m		0.087 ± 0.012	2.605
Taxol		0.457 ± 0.13	

^a IC₅₀ values were determined by triplicate experiments and expressed as mean.

^b CLogP values have been calculated using Organic Chemistry Portal software.

compuestos con posibles funciones anti-metástasis o potenciadoras de la cognición.

- SAR identifican características estructurales relativamente pequeñas pero biológicamente relevantes asociadas con los cambios dramáticos en la actividad
- similitudes con la sulfametazina de 0,98, 0,96 y 0,88.
- desipramina es una base más fuerte que la imipramina. Por lo tanto, es más difícil perder un protón para atravesar la membrana plasmática y entrar en la neurona. Esto sugiere que el acceso a una diana intracelular podría ser un factor que controle la actividad de los compuestos tricíclicos en el bioensayo de fascina.



Colección de datos



Software for calculating the descriptors and fingerprints.

Software	Descriptors	Type of descriptors	Web address	Status
ACD/labs	-	log P, log S, log D, pKa	www.acdlabs.com	Commercial
ADAPT	260	Topological, geometrical, electronic, physicochemical	www.research.chem.psu.edu	Freeware
ADAPT	260	Topological, geometrical, electronic, physicochemical	http://research.chem.psu.edu/pcjgroup/adapt.html	Freeware
ADMET predictor	297	Constitutional, functional group counts, topological, E-state, 3D descriptors, molecular patterns, acid-base ionization, empirical estimates of quantum	www.simulations-plus.com	Commercial
ADRIANA. Code	1244	Constitutional, functional group counts, topological, E-state, Moriguchi, Meylan flags, 3D descriptors, molecular patterns, etc.	www.molecularnetworks.com	Commercial
ALOGPS2.1	-	log P, log S	www.vcclab.org	Freeware
CDK	-	Topological, geometrical, electronic, constitutional	http://cdk.github.io	Freeware
ChemDes	-	Molecular descriptors	www.scbdd.com/chemdes	Webserver (Freeware)
CODESSA	1500	Constitutional, topological, geometrical, charge-related, semi-empirical, thermodynamical	www.codessa-pro.com	Commercial

DRAGON	4885	Constitutional, topological, 2D-autocorrelations, geometrical, WHIM, GETAWAY, RDF, functional groups, etc.	www.talete.mi.it	Commercial
E-DRAGON	–	Molecular descriptors	www.vcclab.org/lab/edragon/	Freeware
JOELib	40	Counting, topological, geometrical properties, etc.	www.ra.cs.uni-tuebingen.de	Freeware
MODEL	3778	Molecular descriptors	http://jing.cz3.nus.edu.sg/cgi-bin/model/model.cgi	Webserve (Freeware)
MOE	300	Topological, physical properties, structural keys, etc.	www.chemcomp.com	Commercial
MOLCONN-Z	40	Topological	www.edusoft-lc.com/molconn	Commercial
MOLD2	779	1D, 2D	www.fda.gov	Freeware
MOLGEN-QSPR	707	Constitutional, topological, geometrical, etc.	www.molgen.molgenqspr.html	Commercial
OEChem TK	–	166-bit MACCS, LINGO, Circular, Path (Daylight-like)	www.eyesopen.com	–
OpenBabel	–	MOLPRINT2D, 166-bit MACCS, Daylight fingerprint (FP2), structural key fingerprints	www.openbabel.org	Freeware
PADEL	1875	1D, 2D, 3D descriptors, molecular fingerprints	www.padel.nus.edu.sg	Freeware
PowerMV	1000	Constitutional, atom pairs, fingerprints, BCUT	www.niss.org/PowerMV	Freeware
PreADMET	955	Constitutional, topological, geometrical, physicochemical, etc.	http://preadmet.bmdrc.org	Freeware
Sarchitect	1084	Constitutional, 2D, 3D	www.strandls.com/sarchitect/index.html	Commercial

Metodologías QSAR-3D

El análisis de Hansch y los métodos de reconocimiento de tendencias usualmente no tienen en cuenta la estructura tridimensional de los compuestos de una manera explícita

Las metodologías que tienen en cuenta explicitamente la geometría 3D se denominan QSAR-3D. Los 2 métodos más usados son:

CoMFA: Comparative Molecular Field Analysis

hace un mapa 3D de interacciones electrostáticas y estéricas de los ligandos en estudio que luego se correlacionan con la actividad biológica

Cramer et al, J. Am . Chem. Soc., 110, 5959 (1988).

CoMSIA: Comparative Molecular Similarity Index Analysis

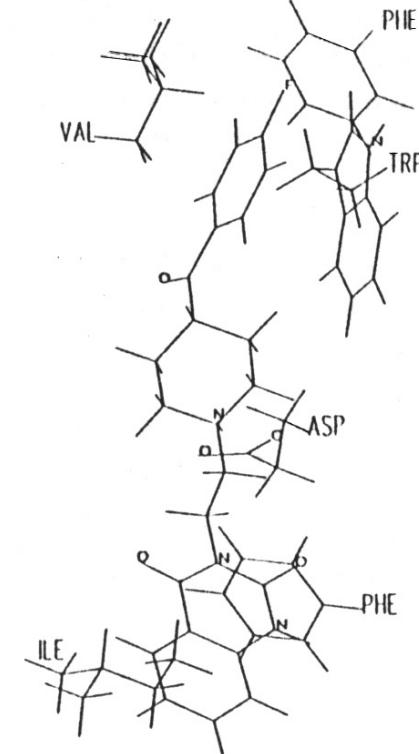
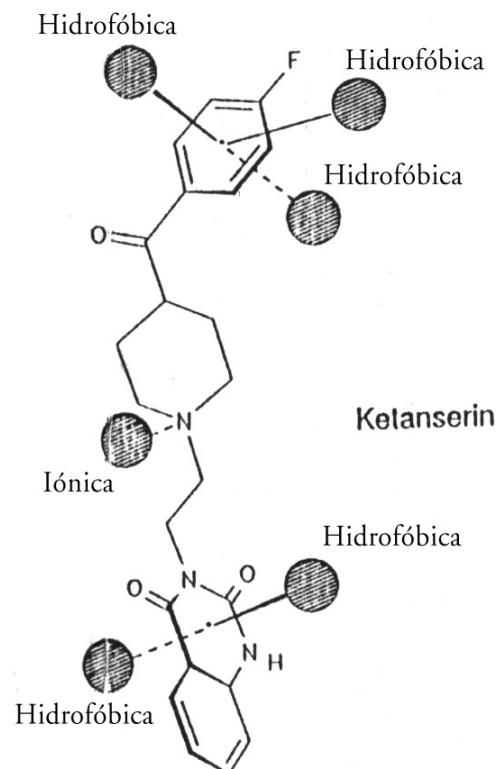
calcula índices de similitud para un conjunto de moléculas previamente alineadas con respecto a propiedades fisicoquímicas (interacciones de distintos tipos, puentes H, etc).

Esos índices luego se correlacionan con la actividad

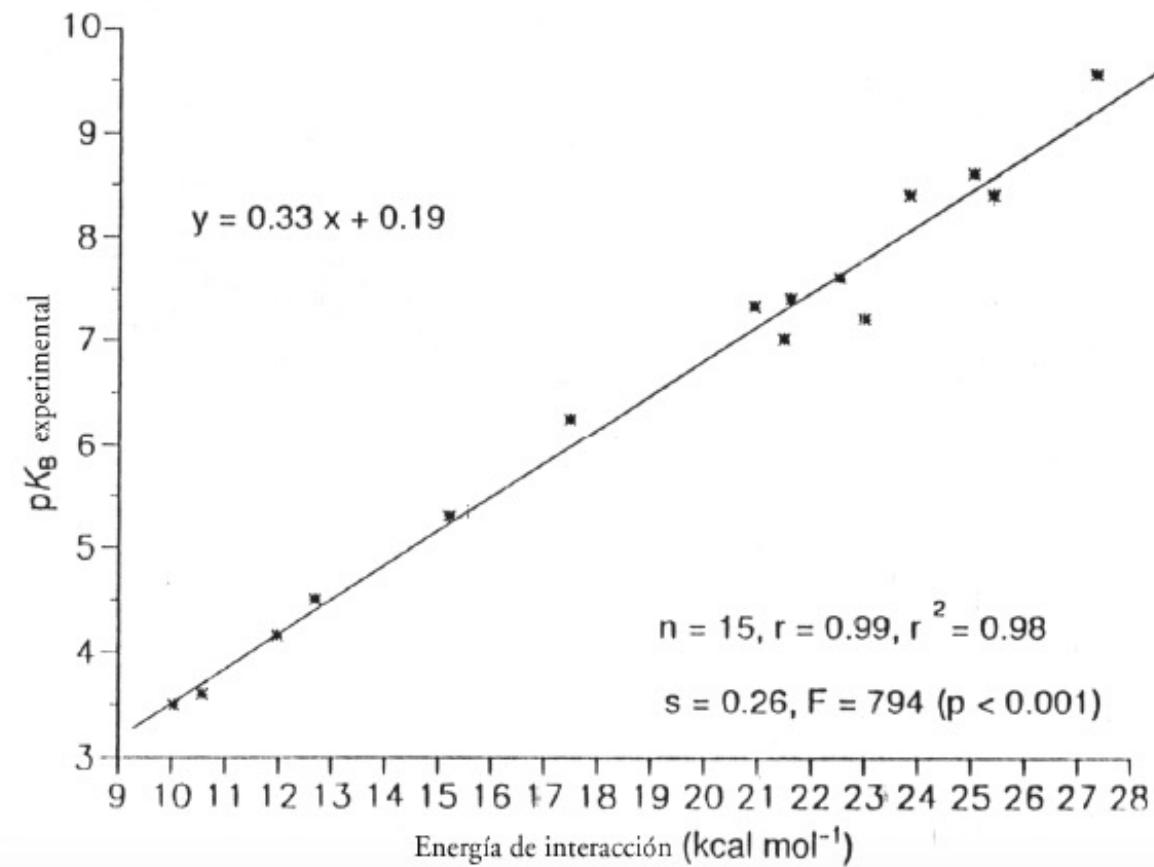
Klebe et al, J . Med. Chem., 37, 4130 (1994)

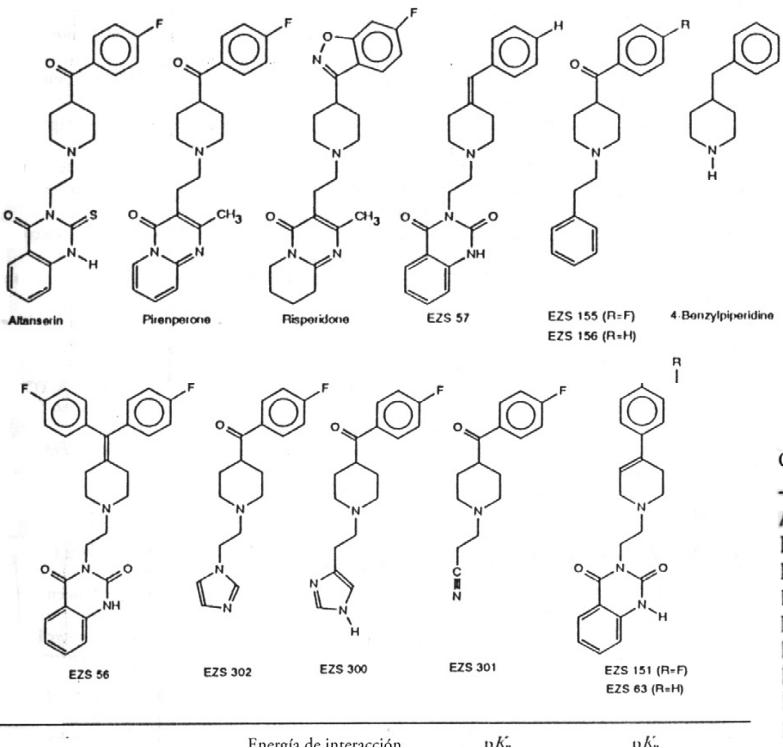
Deben ubicarse aminoácidos hidrofóbicos como PHE, TRI, VAL, LEU, ILE a los lado del sistema planar de anillos

Un aminoácido acídico (aspártico) debe ubicarse opuesto al N protonado



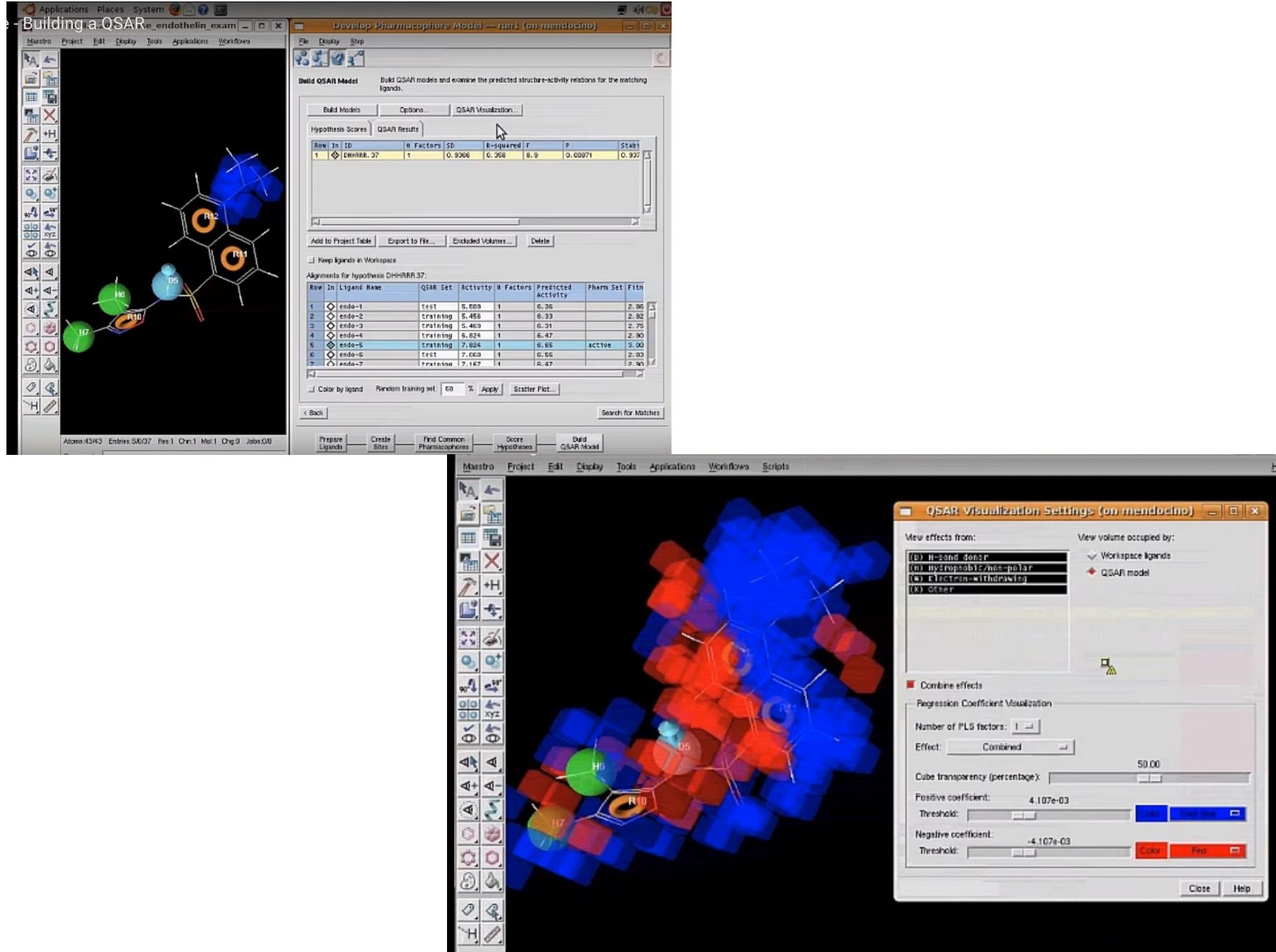
Las constantes de unión pueden correlacionarse con las energías de interacción





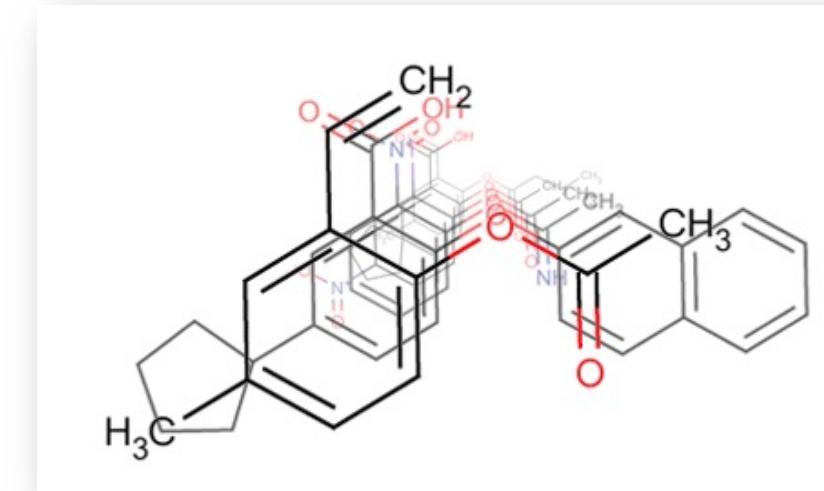
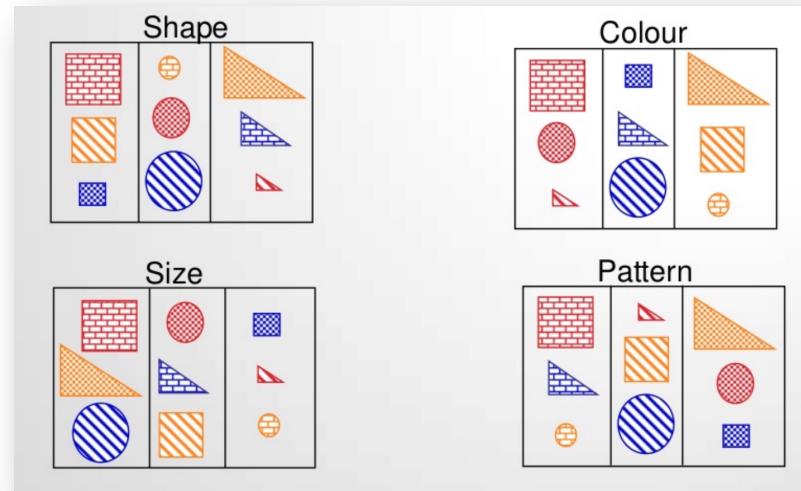
Con esa correlación pueden predecirse afinidades de nuevas estructuras

Compuesto	Energía de interacción (kcal mol ⁻¹)	K_a Calculado	K_a Experimental
Altanserin	27.62	9.31	9.80
Pirenperone	27.16	9.15	9.30
Risperidone	28.69	9.66	9.70
EZS-57	22.27	7.54	7.20
EZS-155	23.91	8.08	8.50
EZS-156	21.61	7.32	8.20
EZS-302	22.09	7.48	7.70
EZS-300	23.23	7.86	7.65
EZS-301	19.37	6.58	6.25
4-Benzylpiperidine	12.64	4.36	5.00
EZS-56	25.53	8.62	8.60
EZS-151	23.64	7.99	7.40
EZS-63	21.41	7.26	7.20



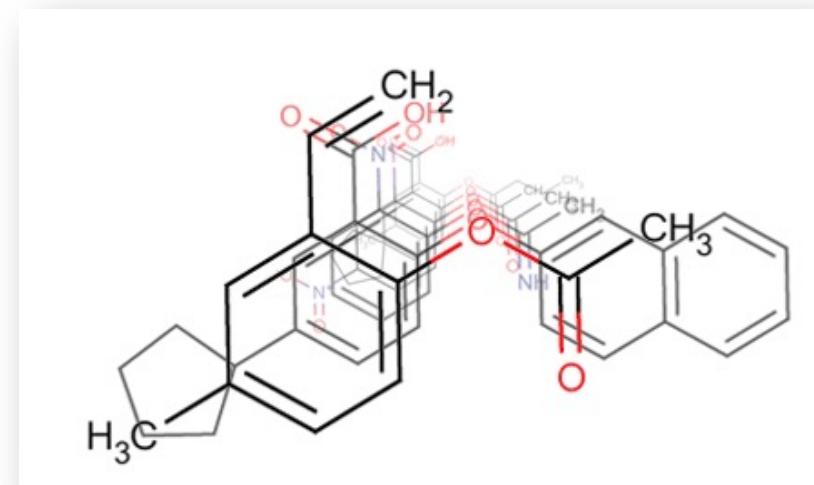
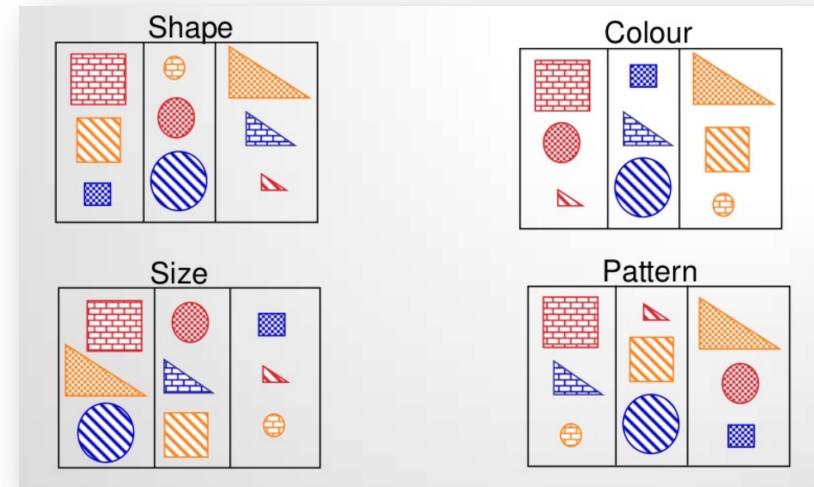
- <https://www.molsoft.com/gui/3Dqsar-tutorial.html>
- <https://www.molsoft.com/gui/cheminformatics.html>
- <https://www.perkinelmer.com/es/product/signals-lead-discovery-sld>
- <https://www.optibrium.com/stardrop/stardrop-torch3D.php>
- <https://www.optibrium.com/stardrop/stardrop-adme-qsar-models.php>
- <http://www.molsoft.com/chemistpro/sar-analysis.html>
- <http://www.openmolecules.org/datawarrior/>
- DOI: [10.1039/C1RA00924A](https://doi.org/10.1039/C1RA00924A) (Review Article) [RSC Adv.](https://doi.org/10.1039/C1RA00924A), 2012, **2**, 369-378

Análisis de similitud molecular (MSA).



Análisis de similitud molecular (MSA).

- El número de características que se pueden comparar entre dos moléculas es bastante grande y algunas pueden contener información redundante.
- Se utilizan una variedad de descriptores y algoritmos para capturar las características esenciales y cuantificar la coincidencia.
- Diferentes representaciones y sus correspondientes medidas de similitud pueden conducir a resultados significativamente diferentes.



Métodos de forma

- Esta categoría engloba un grupo de metodologías con complementariedad de forma y tamaño. Los métodos de forma, exigen implícitamente modelos 3D y la consideración de un conjunto diverso de confórmeros para moléculas inherentemente flexibles.

Method	Number of Variables	Advantages	Limitations	Reference
USR	3	Fast comparison; Descriptors generation	Chirality not recognized; Chemical incompatibilities	Ballester et al. (2009)
CSR	3	Chirality recognition	Chemical incompatibilities	Armstrong et al. (2009)
ElectroShape	4 (charge) or 5 (charge and ClogP)	Chirality recognition; Charge/lipophilic compatibility	Dependence on charge and ClogP calculations	Armstrong et al. (2010)
UFSRAT	3 and pharmacophoric sets	Pharmacophoric compatibility	Centroids' dependence on atom types	Hsin et al. (2011)
USRCAT	3 and pharmacophoric sets	Pharmacophoric compatibility with aromaticity; Real-time screening		Schreyer and Blundell (2012)

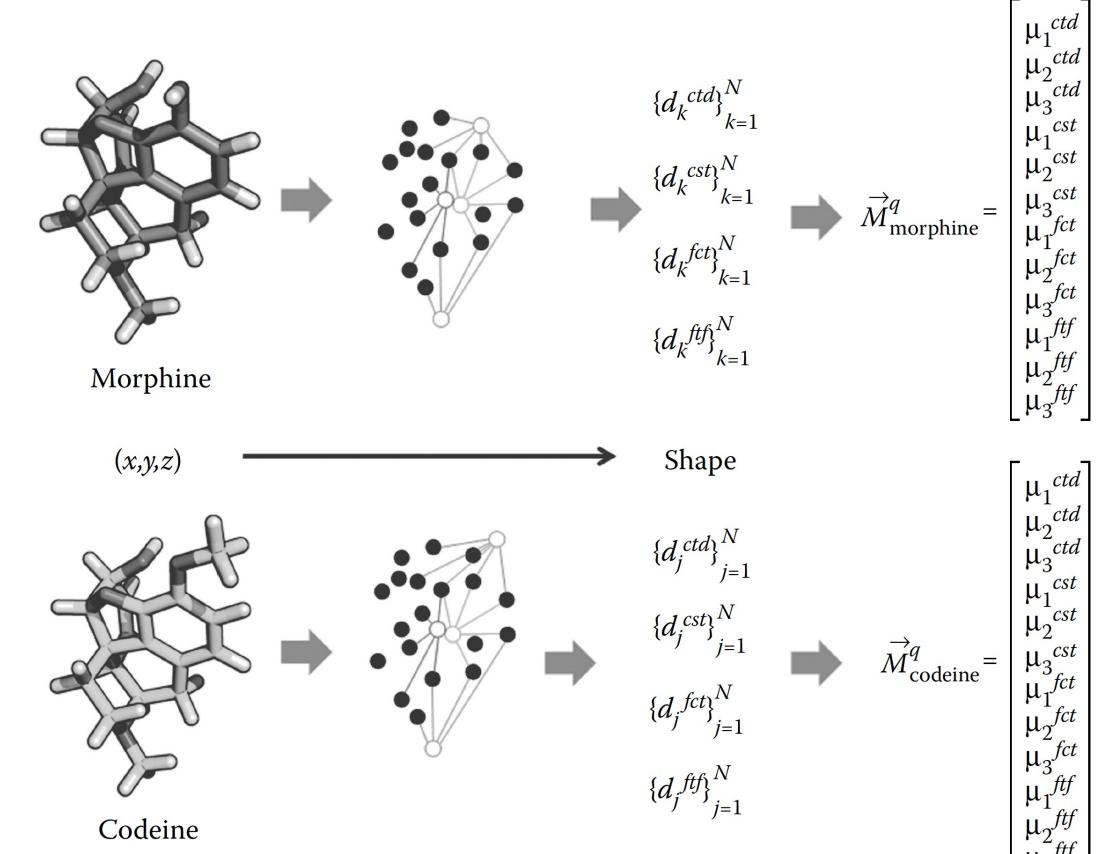
Los métodos de este tipo no tienen en cuenta explícitamente el volumen o la forma en sí mismos.

Ultrafast shape recognition-USR: Evaluating a new ligand based virtual screening technology

El reconocimiento de ultraformas (USR) se concibió bajo la premisa de que la forma de una molécula se puede determinar de forma única mediante un conjunto seleccionado de distancias medidas desde cada átomo hasta cuatro ubicaciones distintas.

Cada conformación molecular se describe mediante cuatro distribuciones de distancias atómicas, siendo el número de esta última proporcional al número de átomos.

Los momentos primero (distancia promedio), segundo (raíz cuadrada de la varianza) y tercero (raíz cúbica de la asimetría) de estas cuatro distribuciones constituyen un vector de 12 descriptores de forma que codifica de forma única la forma y el tamaño moleculares.

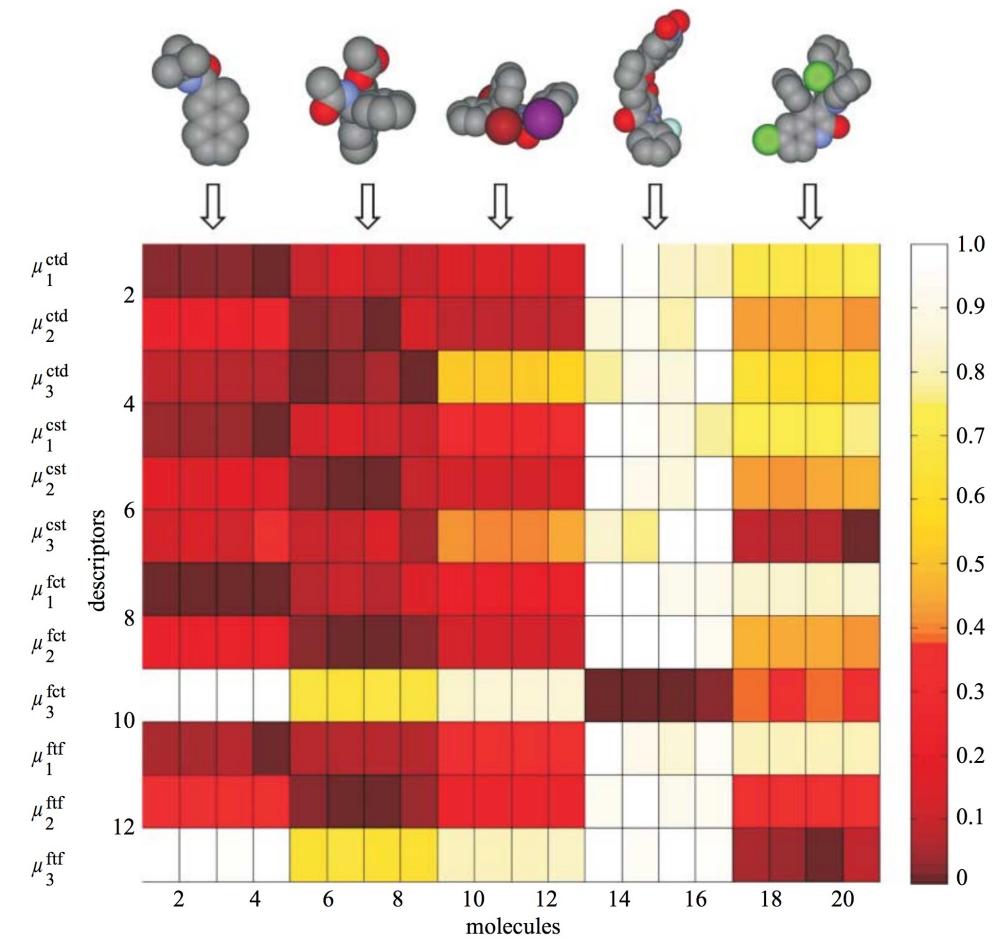


Ultrafast shape recognition-USR, CSR

Los vectores USR pueden almacenarse con requisitos mínimos de espacio y emplearse en diferentes proyectos para determinar qué moléculas en una base de datos tienen una forma más similar a la de una molécula de consulta determinada.

Introducción de un procedimiento de cálculo diferente conocido como reconocimiento de forma quiral (CSR).

Comprensión de que dos moléculas podían compartir una forma idéntica y aún así no poder unirse al mismo bolsillo debido a incompatibilidades polares, principalmente debido a repulsiones carga-carga y / o grandes costos de desolvatación.

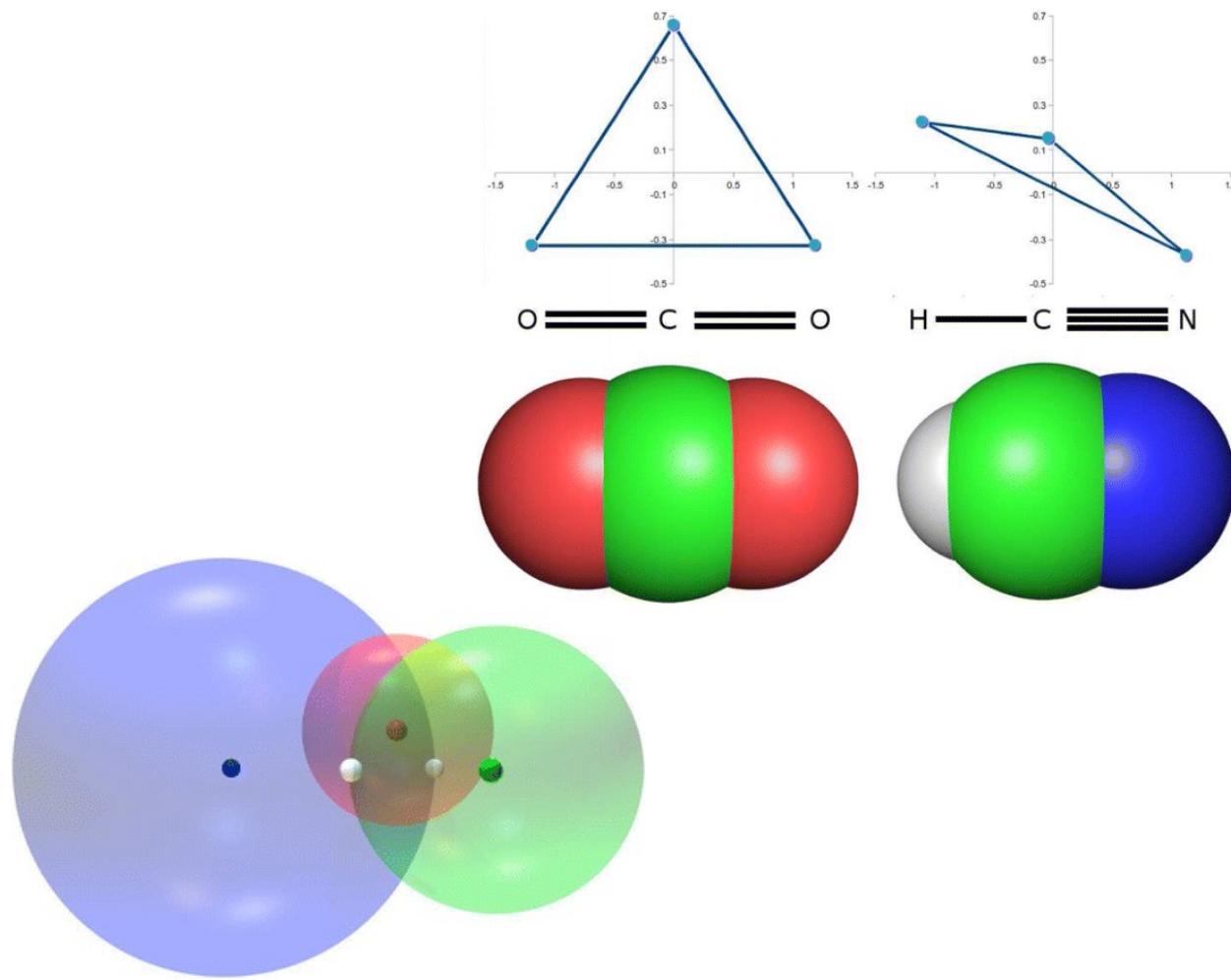


ElectroShape: fast molecular similarity calculations incorporating shape, chirality and electrostatics

Se introdujeron una cuarta (carga atómica parcial) y una quinta ($c\text{LogP}$) variable, dando origen al método ElectroShape.

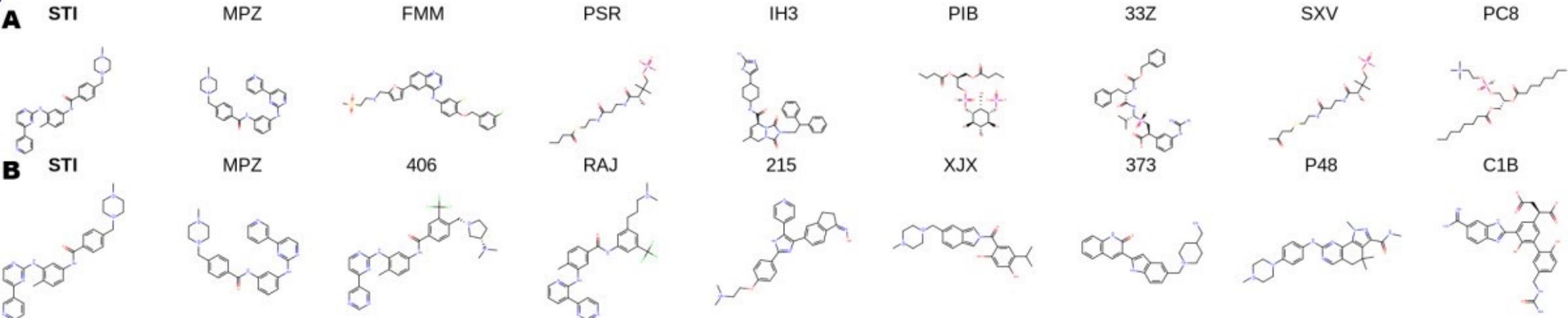
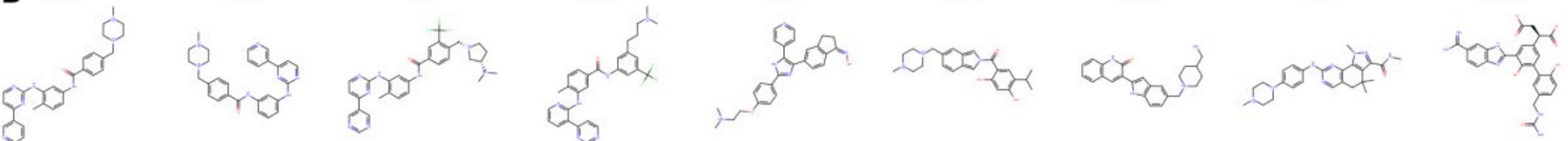
Combina la forma y la información electrostática en un marco único y unificado. Basándose en el enfoque de reconocimiento de forma USR para el cribado virtual rápido no superpuesto basado en formas, amplía el método al representar la información de carga parcial como una cuarta dimensión.

Estas nuevas variables se escalaron adecuadamente mediante un procedimiento de optimización entrenado con conjuntos de pruebas de VS comunes.



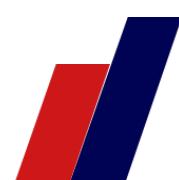
Agregar carga parcial como una dimensión adicional ayuda a discriminar entre moléculas de forma similar.

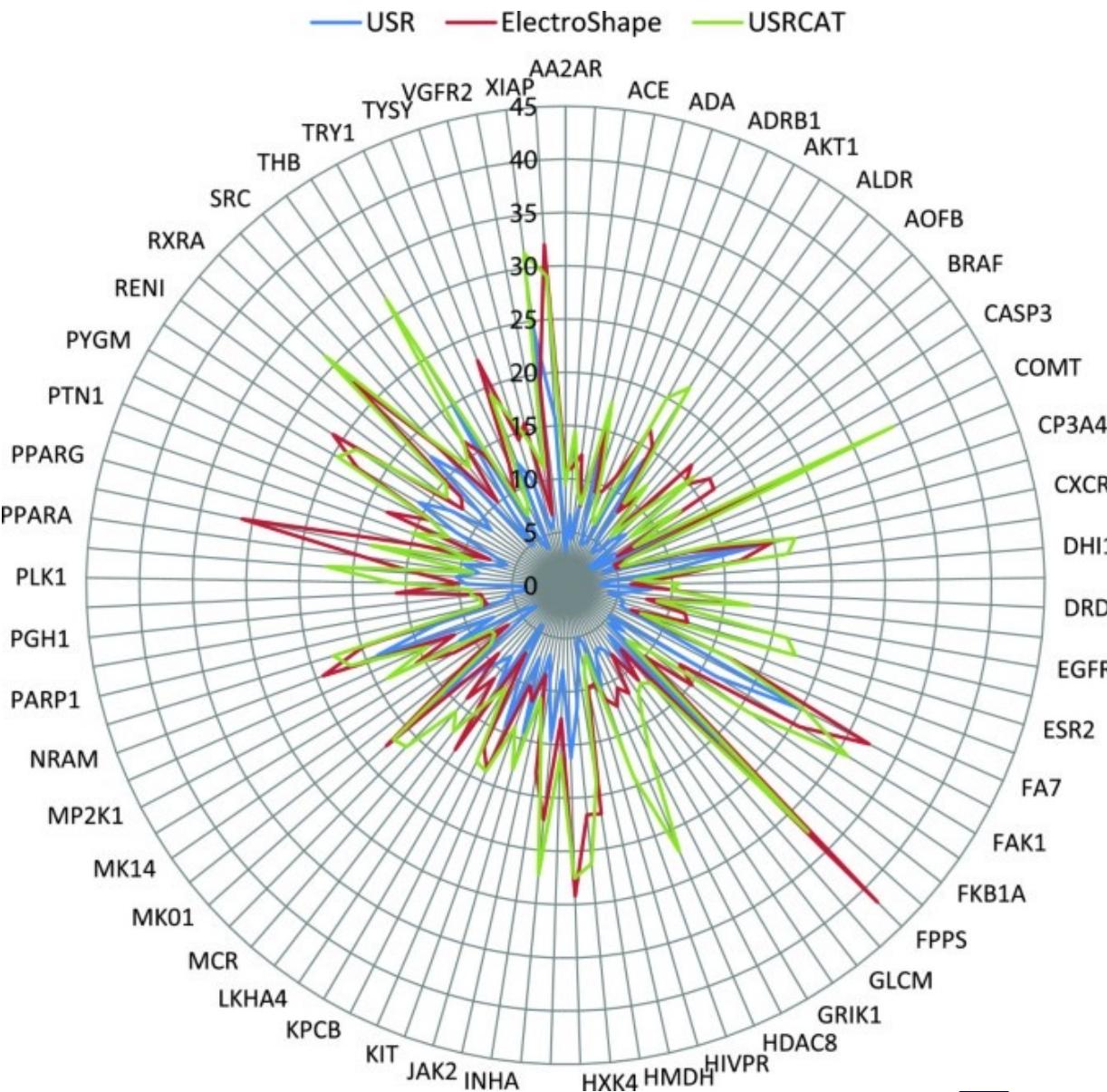
USRCAT: real-time ultrafast shape recognition with pharmacophoric constraints

**A****B**

Reconocimiento de formas ultrarrápido con tipos de átomos (UFSRAT).

El método USR tradicional es capaz de recuperar resultados muy similares, pero tiene problemas con compuestos que parecen tener distribuciones de coordenadas atómicas similares pero que no tienen ninguna relación en términos de sus perfiles farmacofóricos



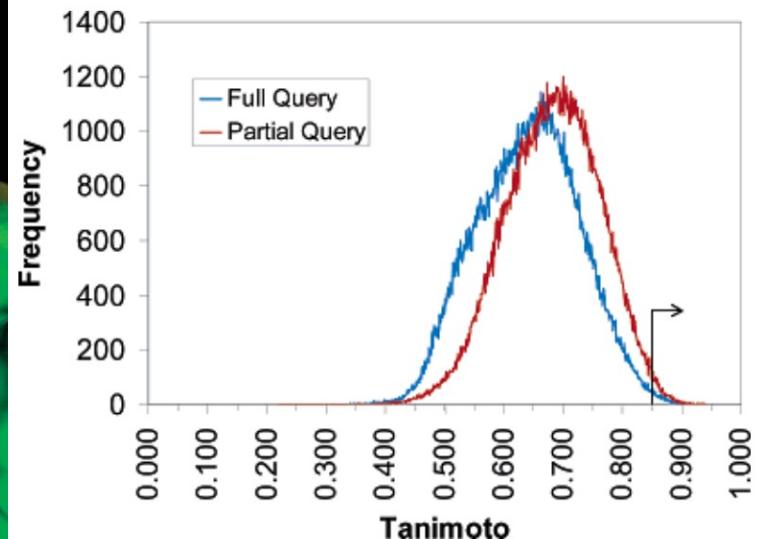
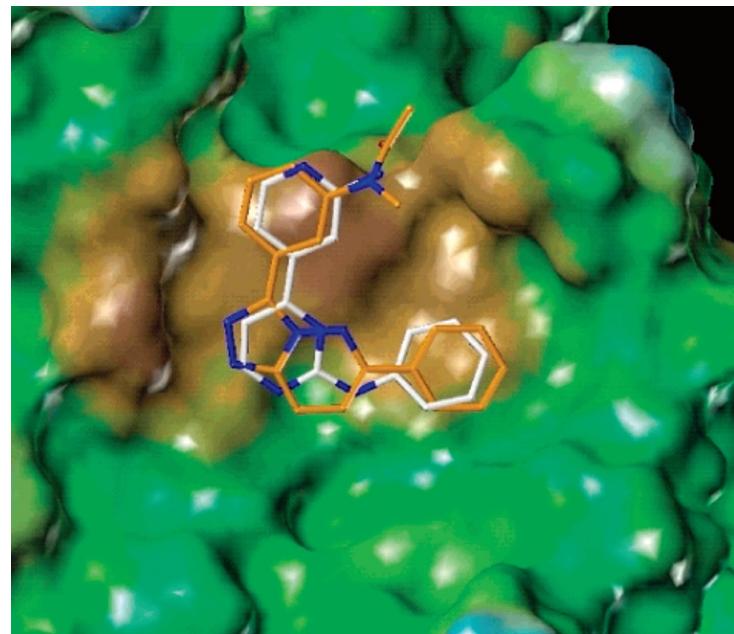
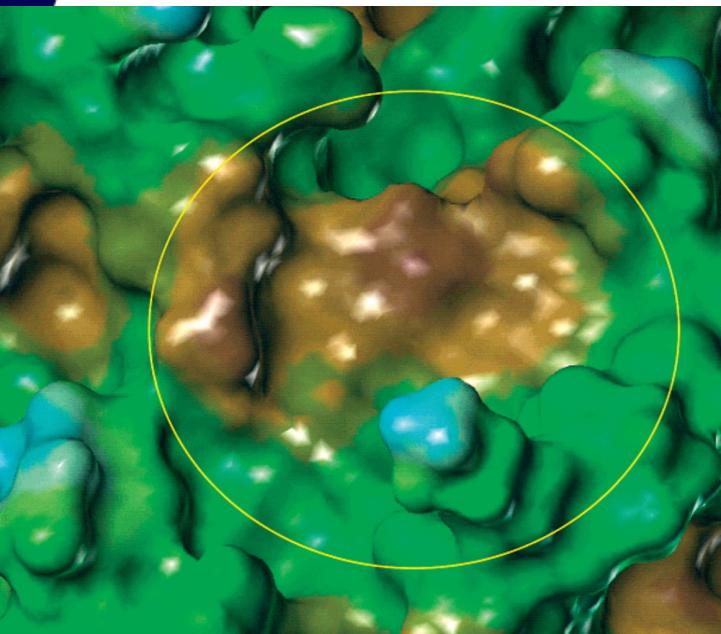


Métodos de forma gaussiana

Reemplazar la colección de esferas duras que se cruzan de diferentes radios utilizadas para representar una molécula con un conjunto de funciones gaussianas superpuestas centradas en el átomo. Esta forma se utilizó para comparar dos moléculas optimizando analíticamente su intersección de volumen.

Method	Description	Availability	References
ROCS	Fast Gaussian overlay based shape comparison. Widely used shape based virtual screening tool. GPU version also available.	Developed by OpenEye Scientific Software (https://www.eyesopen.com). Commercial.	Rush et al., 2005; Hawkins et al., 2007
PAPER	Accelerates large scale virtual screening experiments. Parallel implementation on NVIDIA GPUs.	Developed by Stanford University. Open source. Available from SimTK at https://simtk.org/projects/paper	Haque and Pande, 2010
MolShaCS	Uses Gaussian description of shape and charge. Hodgkin like similarity metric. Molecules are considered rigid.	Developed by University of Sao Paolo, Brazil. Open source tool available at https://code.google.com/archive/p/molshacs/downloads	Vaz de Lima and Nascimento, 2013
SHAFTS	It combines shape similarity with pharmacophoric features. Employs a hybrid similarity metric combining shape and chemical similarity. Suitable for large scale virtual screening.	Developed by Shanghai Key Laboratory of New Drug Design, East China University of Science & Technology, Shanghai, China. Available for download from http://lilab.ecust.edu.cn/home/resource.html	Liu et al., 2011
Phase Shape	Uses atom triplets to generate initial alignments which are refined by Gaussian overlay.	Developed by Schrodinger. (https://www.schrodinger.com). Commercial.	Sastray et al., 2011
ShaEP	Generate consensus shape pattern based on structural features of known ligands.	Developed by Abo Akademi University, Finland. Free for Academics. Available from the Abo Akademi University at http://users.abo.fi/mivainio/shaep/index.php	Vainio et al., 2009
SimG	Uses downhill simplex method to evaluate shape and chemical similarity between two molecules. Comparison of ligand and binding pocket shape or chemical similarity is also possible.	Developed by Shanghai Key Laboratory of New Drug Design, East China University of Science & Technology, Shanghai, China. Available for download from http://lilab.ecust.edu.cn/home/resource.html	Cai et al., 2013
SABRE	Uses consensus shapes to generate initial alignments which are later refined by rigid-body rotations and translations.	Academic license is available on request	Hamza et al., 2012, 2013
WEGA	Uses a weighted Gaussian function to improve the accuracy of first order approximation. A GPU implementation (gWEWA) is also available for large scale virtual screenings.	Developed by Research Center for Drug Discovery, Sun Yat-sen University, China. Academic license is available on request at http://www.rcdd.org.cn/home/program.html .	Yan et al., 2013

Superposición rápida de estructuras químicas

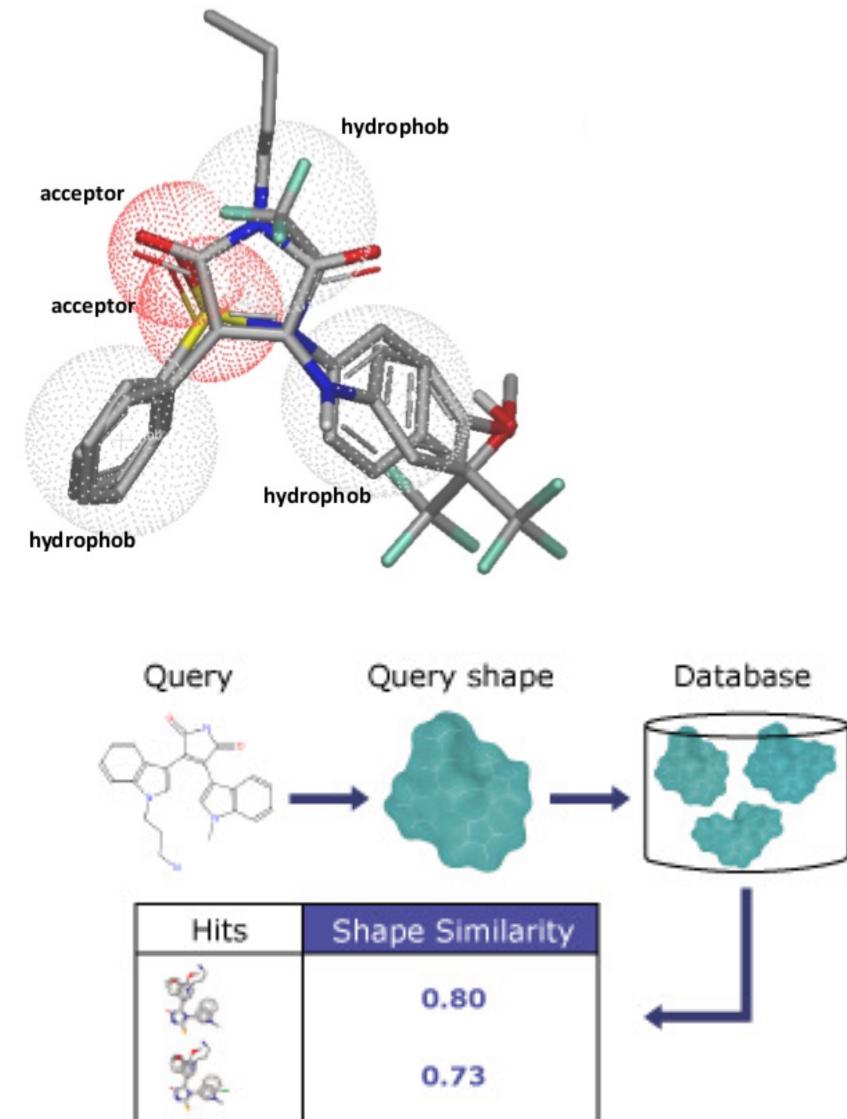


ROCS no usa directamente la información estructural de la proteína, las moléculas identificadas por el procedimiento pueden potencialmente tener un gran volumen de superposición con regiones del sitio de unión de ZipA.

El algoritmo primero usa ROCS para crear un volumen gaussiano tridimensional para la molécula de consulta y para buscar rápidamente moléculas con formas similares en una base de datos de múltiples conformeros . Luego, se realiza una evaluación simple de las interacciones de van der Waals de los impactos de ROCS con la proteína para eliminar moléculas con choques estéricos obvios con la proteína

hay que tener en cuenta que los procedimientos de optimización son muy sensibles a la posición y orientación inicial de las moléculas que se comparan, lo que hace que en algunos casos no se encuentre la respuesta correcta. se dan diferentes orientaciones iniciales como entrada a un algoritmo de optimización como SIMPLEX o simulado de Monte-Carlo todos los cuales han demostrado ser exitosos para el alineamiento molecular en VS.

En comparación con los métodos similares a USR, dos deficiencias son evidentes: (i) son mucho más lentos, y (ii) los resultados no se pueden almacenar y reutilizar. Sin embargo, una clara ventaja es que su aplicabilidad se extiende a otras técnicas de modelado más allá del VS porque los resultados se pueden verificar visualmente.

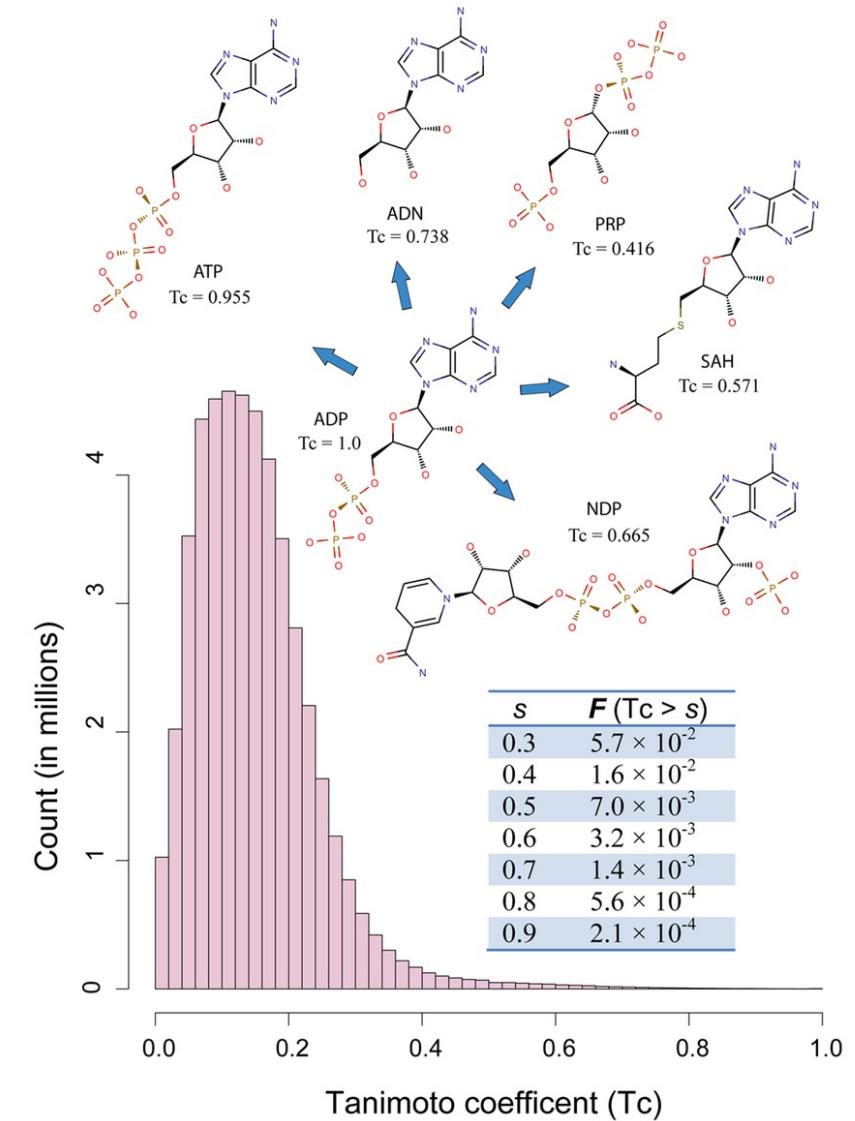


Huellas digitales y claves estructurales.

- La métrica normalmente utilizada para calcular la superposición entre dos moléculas

$$T_c = \frac{\text{overlap}_{ab}}{\text{overlap}_{ab} - \text{overlap}_{aa} - \text{overlap}_{bb}}$$

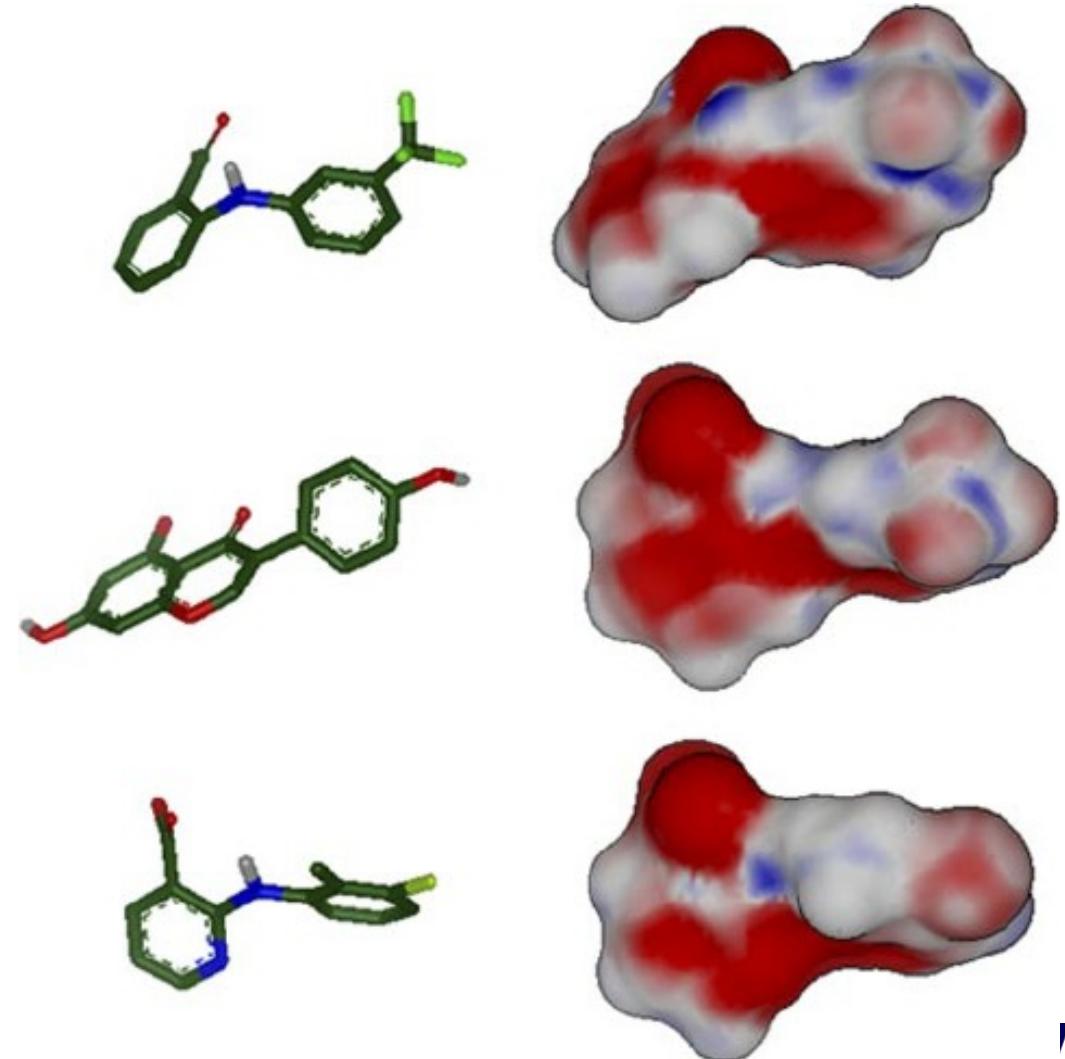
Donde la superposición ab, la superposición aa y la superposición bb son las superposiciones de volumen de la molécula a con la molécula b, de la molécula a consigo misma y de la molécula b consigo misma, respectivamente. Los valores de Tc oscilan entre 0 (sin similitud) y 1 (una coincidencia perfecta).



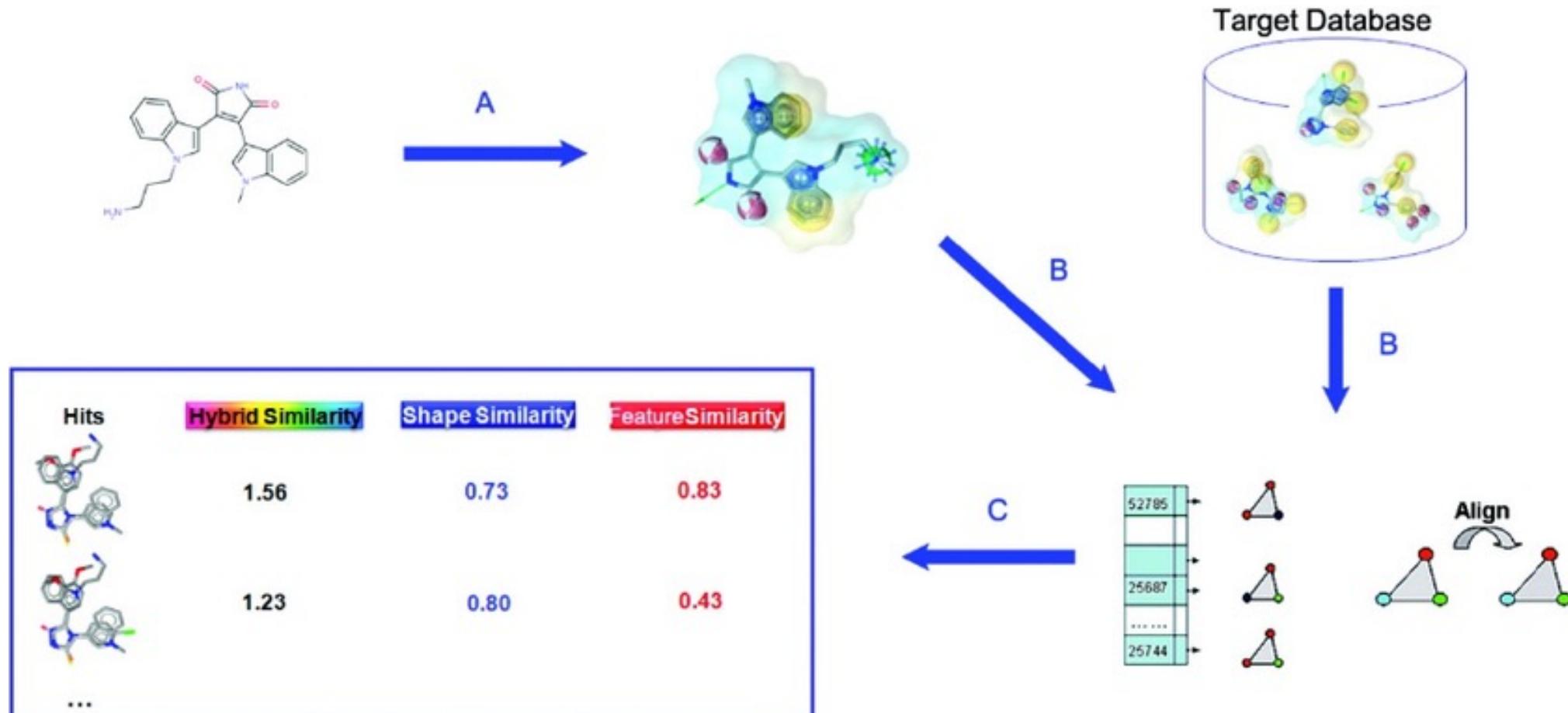
ShaEP (Shape and Electrostatic Potential)

Se desarrolló con el objetivo de capturar las fortalezas de los enfoques volumétricos y de campo. Esta superposición se optimiza maximizando la superposición de sus volúmenes moleculares.

La alineación de campo es capaz de reconocer esas coincidencias parciales, pero no es tan rápido como el método de superposición de volumen y generalmente no es práctico desde el punto de vista computacional.



SHAFTS (SHApE FeaTure Similarity)

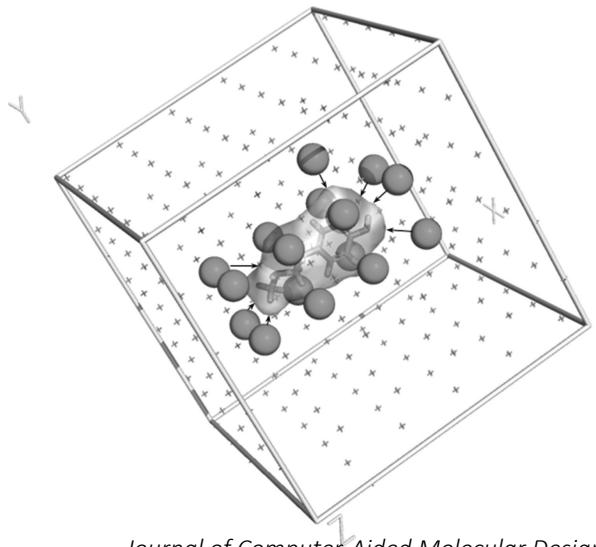


SHAFTS. (A) Genera puntos característicos del farmacóforo de la molécula activa seleccionada; (B) Busque en una base de datos superponiendo triplete de características; (C) Clasifique los compuestos por la puntuación de similitud

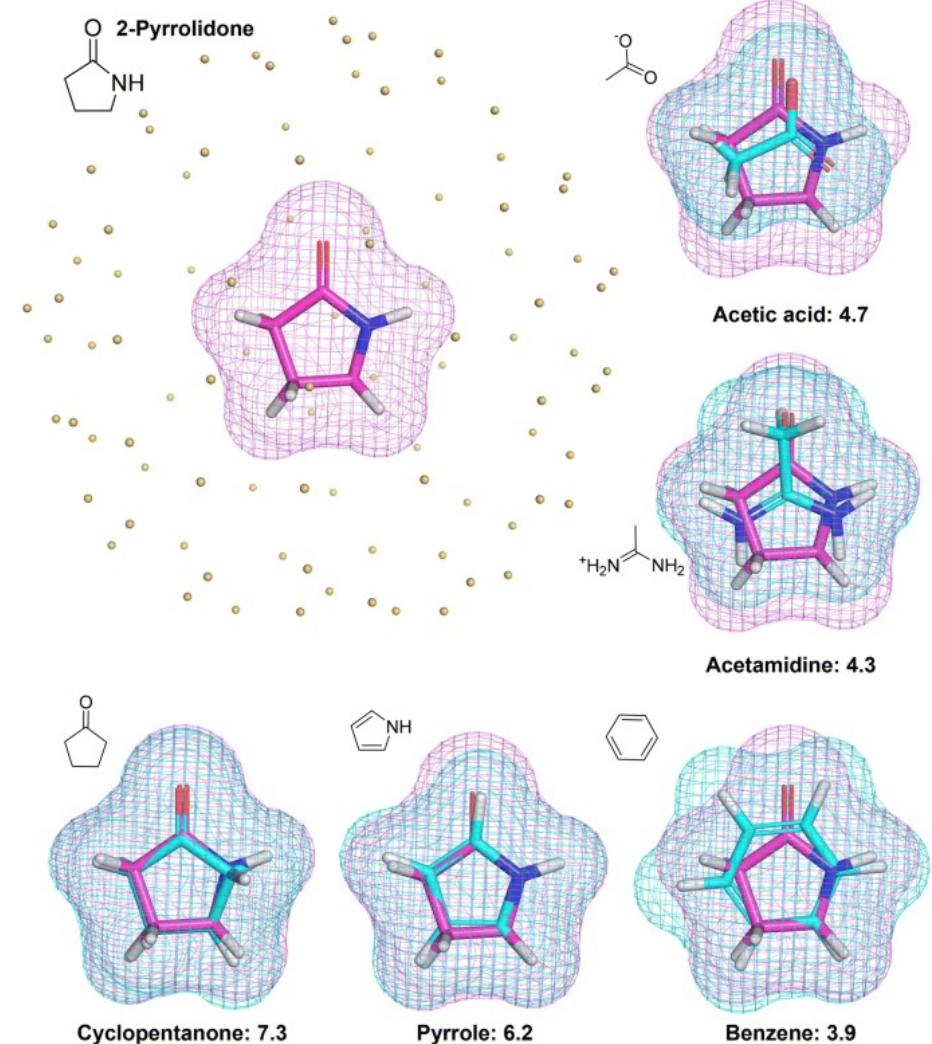
Métodos de superficie local

Utilizan una función de similitud morfológica basada en superficies moleculares. En un primer paso, se define un conjunto de puntos alrededor de la molécula.

Estos "puntos de observación" se utilizan para medir varios tipos de distancias a la superficie molecular. Se han utilizado métodos de superficie local en la superposición molecular para estudios QSAR y predicciones fuera del target.



Journal of Computer-Aided Molecular Design volume 33, pages 865–886(2019)



Huellas digitales y claves estructurales

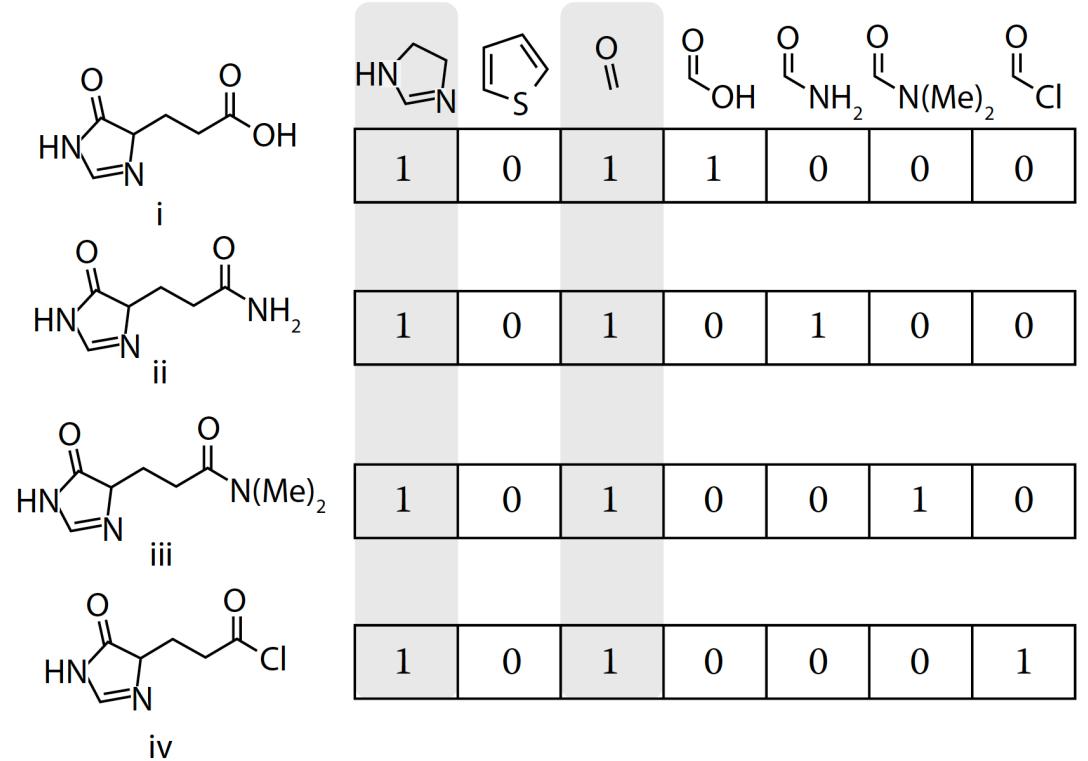
Esta categoría agrupa una serie de métodos que tienen en común la codificación de información molecular. La mayoría de ellos se basan en una representación molecular 2D reducida, es decir, un gráfico, pero también son aplicables a estructuras moleculares 3D.

Type	Examples	Reference
Structural keys	MDL MACCS	Durant et al. (2002)
Extended	ECFP, FCFP	Rogers and Hahn (2010)
Atom distances	TAD (typed atom distances)/TGT (typed graph triangles)	Bender et al. (2009)
Path-based	Daylight	Bender et al. (2009)
Pharmacophoric	CATS (2D), piDAPH3 (pharmacophore atom triangle)	Schneider et al. (1999)

MACCS

"huella molecular" fue necesario desarrollar algoritmos rápidos para comparar y encontrar similitudes entre moléculas. diccionario más o menos amplio de entidades relevantes como átomos, enlaces y fragmentos que se clasifican en diferentes tipos.

La estructura se codifica como una serie de indicadores binarios que indican la presencia (1) o ausencia (0) de estas unidades particulares, los 0 y los 1 se almacenan en cadenas comúnmente denominadas "claves estructurales". Las cadenas de bits son extremadamente rápidas para comparar computacionalmente debido a su naturaleza binaria.



codificación de claves estructurales en huellas dactilares para la molécula i y tres análogos

Molecule A:

00010100010101000101010011110100

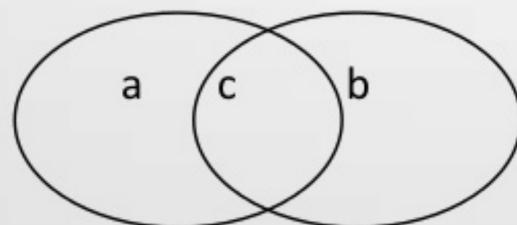
Molecule B:

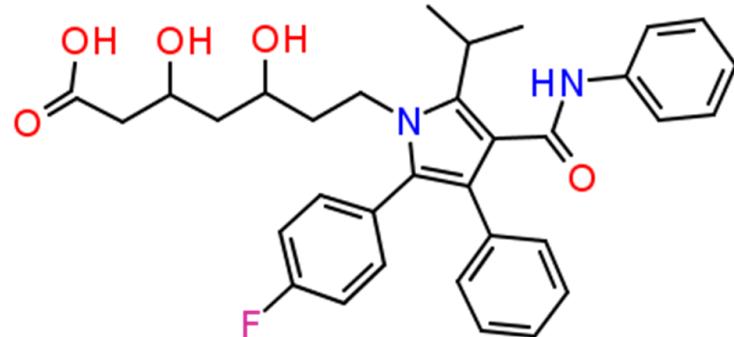
00000000100101001001000011100000

$$\text{Tanimoto coefficient} = \frac{c}{(a + b) - c}$$

Where $c = A \text{ AND } B$

$$\text{Tanimoto} = 6 / (13 + 8) - 6 = 0.4$$





C N O F

... 1 1 1 1 0 ...

...	1	1	1	1	1	1	...
-----	---	---	---	---	---	---	-----

C-C C-N C-O N-S

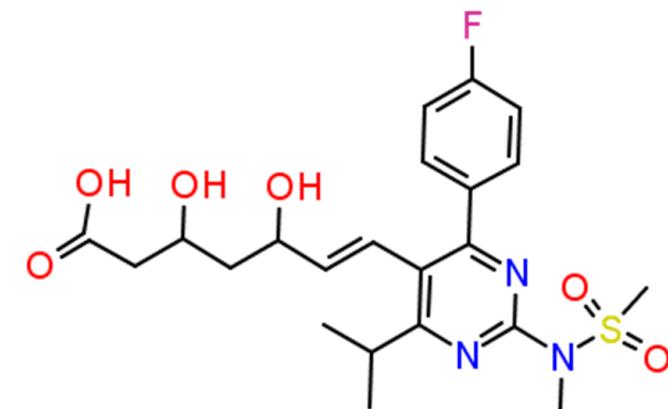
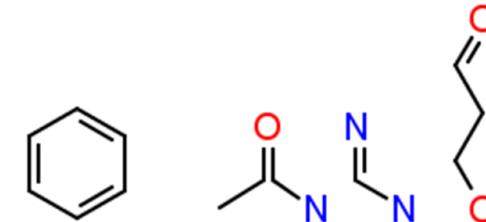
1 1 1 0

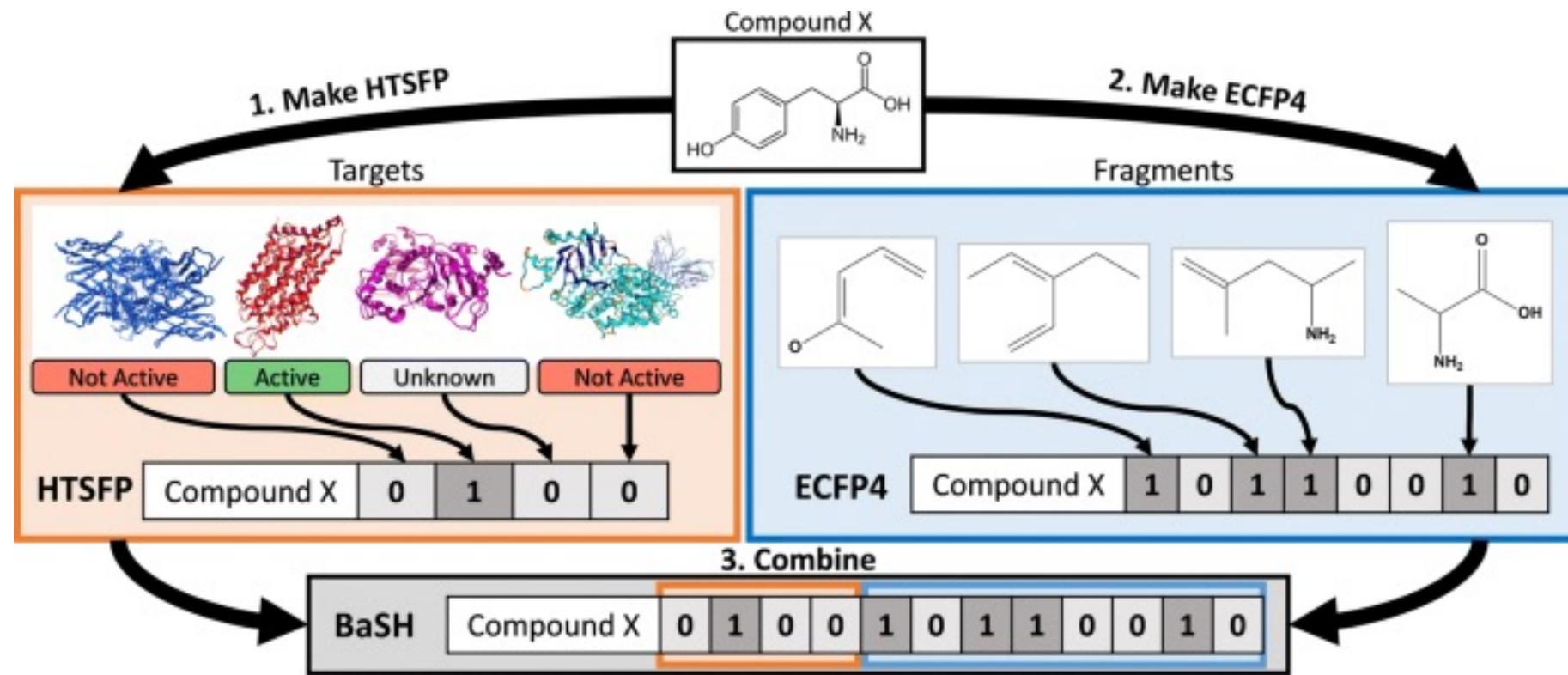
1 1 1 1

C=C C=N C=O

1 1 1

1 1 1



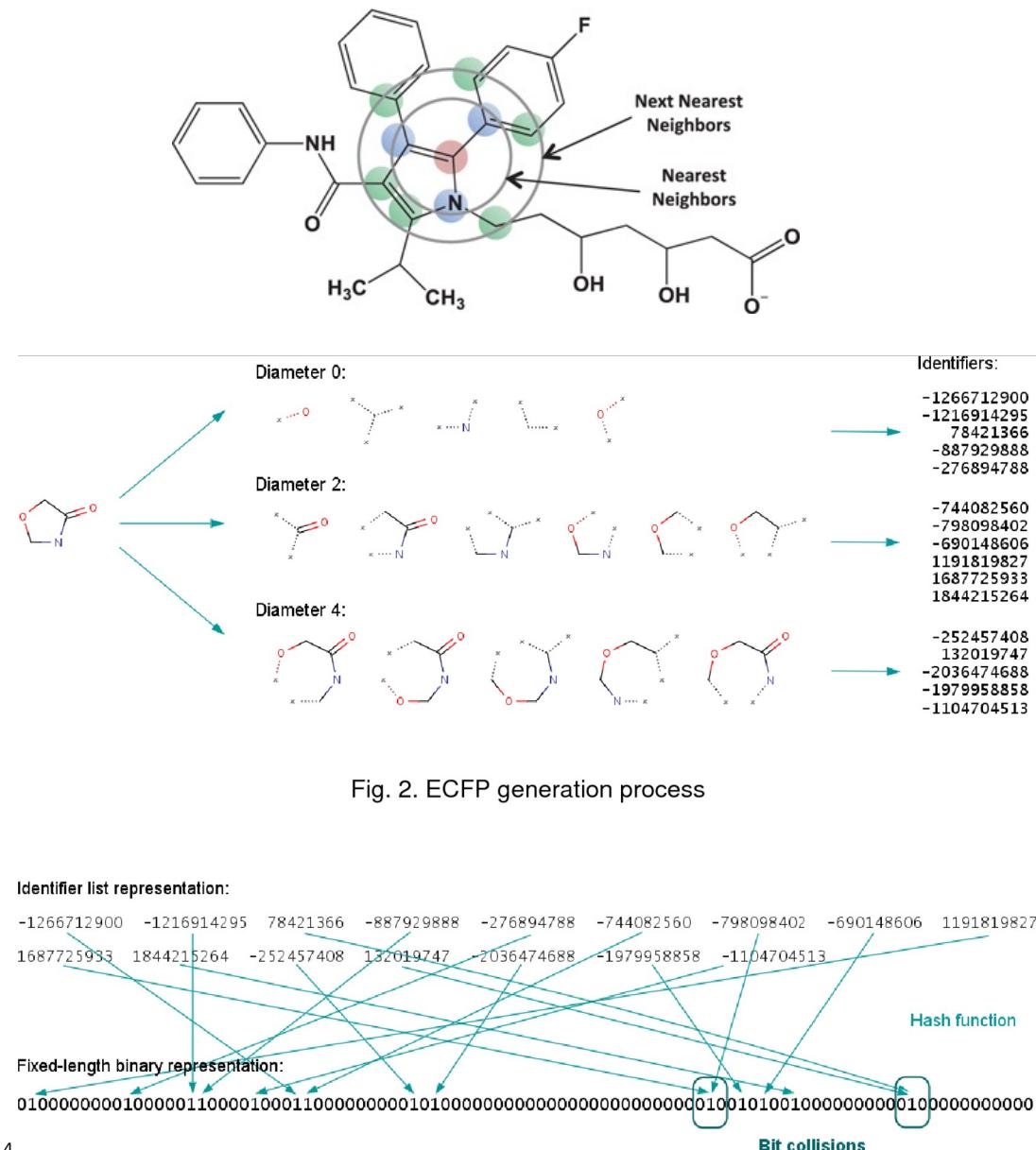


- Las huellas dactilares más comunes incluyen patrones químicos generales o fragmentos como múltiples átomos enlazados (típicamente hasta 7), ramas y ciclos, distancias topológicas por pares (es decir, distancias medidas como el número de enlaces entre dos puntos seleccionados en el gráfico), etc. , a cada uno de los cuales se le asigna un valor entero.

Huellas digitales de conectividad extendida- ECFP

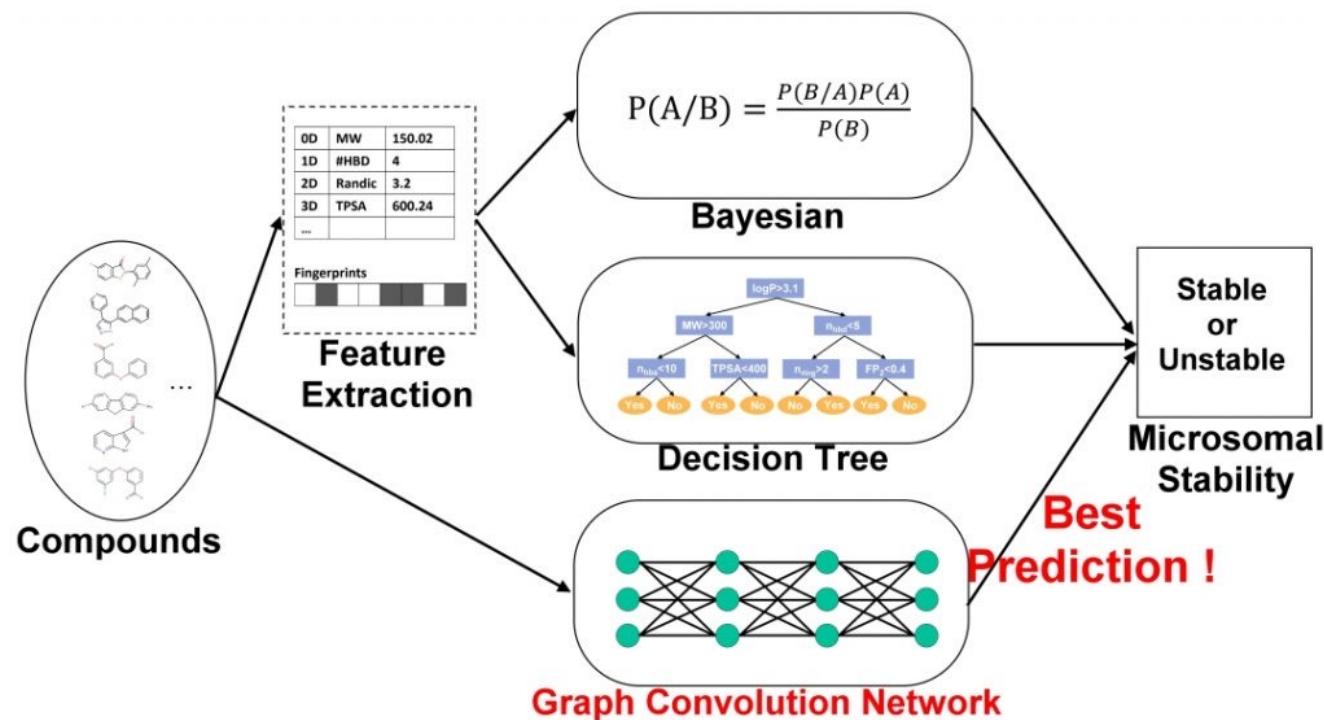
Versión modificada de algoritmo que etiqueta inequívocamente átomos en moléculas. Estas etiquetas forman la base de la huella digital y se generan iterativamente sobre la base de las conectividades y vecinos.

- (i) ECFP, que utilizan seis valores que no dependen del orden de numeración (número atómico, masa atómica, carga atómica parcial, conectividad, número de enlaces a átomos que no son de hidrógeno y número de átomos de hidrógeno unidos),
- (ii) Huellas dactilares de clase funcional (FCFP), en este caso, los seis valores iniciales se activan o desactivan dependiendo de los siguientes tipos de átomos: aceptor de enlace de hidrógeno, donante de enlace de hidrógeno, con carga positiva y negativa, aromático y halógeno.



Metodos alternativos

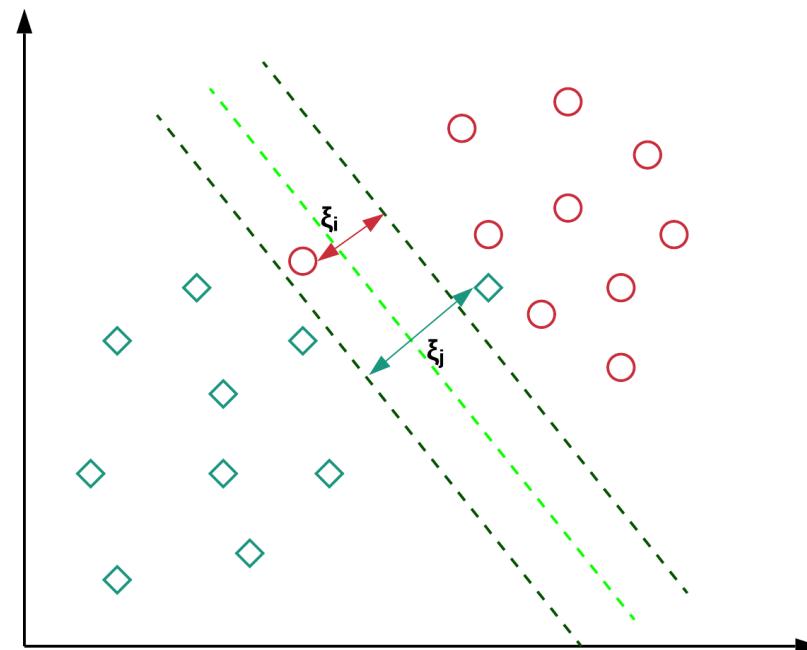
- Los métodos descritos hasta ahora básicamente estudian la similitud de una molécula con respecto a una estructura de referencia que muestra cierta actividad. Otros métodos en el descubrimiento de fármacos basados en ligandos abordan la tarea de identificar nuevas moléculas activas como si fuera un problema de clasificación.



Otros métodos

Máquinas de vectores de soporte

Los métodos que utilizan SVM requieren la disponibilidad de información vinculante sobre moléculas positivas ("activas") y negativas ("baja actividad" o "inactivas"). Conjuntos de entrenamiento y de prueba, y se calcula una serie de descriptores que representan adecuadamente las propiedades más significativas físicamente responsables de la vinculación al objetivo de interés.



Otros métodos

Métodos bayesianos

Los métodos basados en el teorema de Bayes, ampliamente utilizados en la teoría de la probabilidad y la estadística, tienen como objetivo cuantificar la probabilidad de que un compuesto determinado sea activo contra un objetivo seleccionado.

$$P(\text{class}_n \mid V) = \frac{P(V \mid \text{class}_n)P(\text{class}_n)}{P(V)}$$

