

Szerkezet alapú virtuális szűrővizsgálatok a felfedező kutatásban

KESERŰ György Miklós*

CADD&HTS, Richter Gedeon Rt., Gyömrői út 19-21, H-1103 Budapest

1. Bevezetés

A mechanizmus alapú gyógyszerkutatás kémiai szempontból egyik legnagyobb kihívása a kutatási program alapjául szolgáló vezérmolekula azonosítása. A gyógyszerkutatás hőskorában a vezérmolekula azonosítását jórészt in vivo kísérletek segítségével végezték, amelyek általában megfelelő hatékonyságú és elfogadható farmakokinetikai tulajdonságokkal rendelkező vegyületeket eredményeztek. Ez a megközelítés jóllehet meglehetősen lassúnak és kiszámíthatatlannak bizonyult jelentős teret engedve a véletlen, sokszor mégis jelentős felfedezéseknek. Napjaink orvosi kihívásai, az emberi életminőség javításának programja, valamint a gyógyszeripari befektetők hozamvárásai ugyanakkor nem támogatják egy ilyen, nehezen tervezhető stratégia alkalmazását. Ez a felismerés vezetett a modern gyógyszerkutatás egyik legfontosabb felismeréséhez, a molekuláris célpontokon alapuló kutatási stratégia kidolgozásához. A megfelelő állatmodellek hiánya, valamint a hatásmechanizmus utólagos felderítésének nehézségei nagyban hozzájárultak ehhez a paradigmaváltáshoz, amelynek eredményeképpen az addigi in vivo modellek helyett in vitro eljárások kerültek a kutatási programok első vonalába. Az in vitro tesztek elterjedése jótékony hatással volt a racionális megközelítések elterjedésére is, ezek a módszerek voltak először képesek megfelelő minőségű adatokat szolgáltatni a szerkezet-hatás összefüggések számítógéppel segített vizsgálatához, elősegítve ezzel a molekulatervezési (CADD) módszerek elterjedését. Napjainkban az in vitro módszerek által szolgáltatott adatok, mint a gyógyszerkémiai erőfeszítések indikátorai, alapvető szerepet játszanak a vezérmolekula azonosítás és optimalizálás szakaszában. A molekuláris biológia, a biofizika, valamint az információs technológiák és a robotika területén bekövetkezett drámai változások a 90-es évek közepén újabb változást hoztak a vezérmolekula azonosítás gyakorlatában, a nagy áteresztőképességű tesztelés (High Throughput Screening, HTS) a korai fázisú gyógyszerkutatás meghatározó módszerévé vált. A HTS korai alkalmazásai egyértelműen a tömegmértű, többmillió kampányok irányába mutattak, jóllehet mára világossá vált, hogy a vezérmolekula azonosítás hatékonysága nem növelhető egyszerűen csak a vizsgált vegyületek, illetve a megvalósított kampányok számának növelésével. Egyfelől könnyen belátható, hogy még a manapság elérhetővé vált áteresztőképesség ($> 10^6$ vegyület/nap) mellett sem lehetséges a gyógyszereszerű vegyületeket magába foglaló kémiai tér (nagyságrendileg kb. 10^{40} molekula) teljes feltérképezése. Másfelől viszont, az áteresztőképesség növelésének ellenére, a célfehérje farmakológia validálásának nehézségei, a HTS-re alkalmas biológiai tesztek kidolgozásának problémái, valamint a tesztelésre kerülő vegyületek kezelésének és minőségbiztosításának

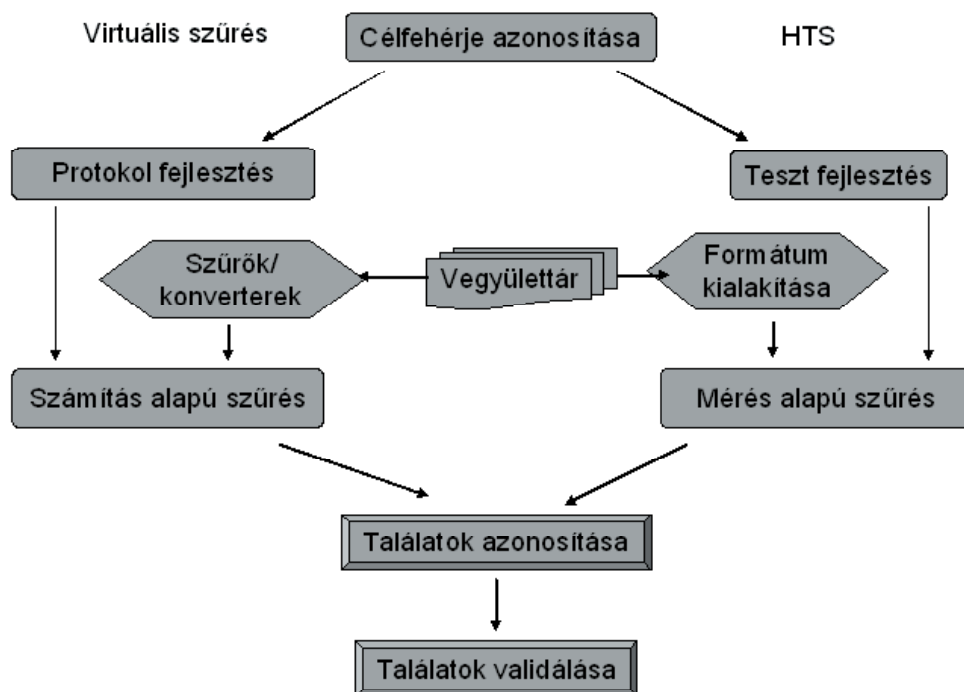
kérdései a megközelítés jelenlegi határait egyértelműen megmutatják. Tekintetbe véve a korábban általánosan alkalmazott tesztelési stratégia korlátait, a véletlen tesztelés olyan eljárásokkal kell felváltani, amelyek egyesítik a HTS áteresztőképességét a CADD módszerek racionális megközelítéseinek előnyeivel. A virtuális szűrővizsgálatok (virtual screening, VS) éppen ezt a koncepciót testesítik meg, amikor nagy áteresztőképességű molekulatervezési módszereket használunk fel nagy vegyületadatbázisok idő- és költségtakarékos vizsgálatára.

A HTS eljárások kidolgozása során elsőként a célfehérjét választjuk ki, ezt követi a vegyületek vizsgálatára alkalmas teszt kifejlesztése és a vizsgálandó vegyületek tesztelésre alkalmas formátumának kialakítása. A HTS kampányokat tipikusan automatizált környezetben valósítjuk meg, amelynek során a nagy számú vegyületnek köszönhetően nagy mennyiségű mérési adat keletkezik. A kampány kísérleti részének befejezését követi az adatok feldolgozásának, értelmezésének és kiértékelésének fázisa, amely jelentős számítógépes kapacitást és felkészültséget igényel.

A HTS és a virtuális szűrés közti rokonságot a megvalósítás főbb lépéseinek hasonlósága is jelzi (1. ábra). A célfehérje azonosítását követően ebben az esetben is az első feladat az aktivitás előrejelzésére szolgáló algoritmus kifejlesztése és optimalizálása, ezt tekinthetjük a HTS-ben alkalmazott in vitro teszt in silico megfelelőjének. A virtuális szűrővizsgálatok végrehajtása előfeltételezi a vizsgálandó vegyületeket tartalmazó adatbázis megfelelő előkészítését. Maga a tényleges szűrés a számítógép memóriájában történik, ennél fogva nagymértékben automatizált és a HTS-hez hasonlóan nagy mennyiségű adatot szolgáltat, amelyet tárolni, elemezni és értékelni szükséges.

A technikai megvalósítás mellett a HTS és VS módszerek koncepciójukban is számos hasonlóságot mutatnak. Mindkét eljárás nagy számú vegyület gyors, korlátozott pontosságú vizsgálatára alkalmas, ebben az értelemben tehát inkább klasszifikációra mintsem a biológiai aktivitás kvantitatív meghatározására szolgálnak. Ebből következően e módszerek valójában nem megerősített találatokat, hanem jelölteket azonosítanak, amelyek további, immár hagyományos, jól reprodukálható módszerekkel történő vizsgálata vezet a valódi találatokhoz. A nagy áteresztőképességű módszerek, jellegükönél fogva - a megvizsgált vegyületek számához képest csak kis mértékben - hamis eredményeket is szolgáltathatnak. Lesznek olyan vegyületek, amelyeket ezek a vizsgálatok aktívnak jeleznek, de a validáló vizsgálatokban ezt az aktivitást nem tudjuk megerősíteni, ezek a hamis

* Tel.: +36-1-4314605 ; fax: +36-1-4326002 ; e-mail: gy.keseru@richter.hu



1. Ábra. A virtuális illetve valós szűrővizsgálatok legfontosabb lépései

pozitívok. Ugyanakkor bizony számíthatunk olyan vegyületekre is, amelyeket tényleges aktivitásuk ellenére a nagy áteresztőképességű módszerekkel nem tudunk azonosítani, ezek a hamis negatívok. Bár a hamis eredmények elkerülése érdekében, mind a HTS, mind a VS vizsgálatok során a vizsgálati teszt gondos optimálásával mindent el kell követni, ezek valójában e módszerek közös működési filozófiáját és alkalmazását nem befolyásolják. Mindaddig, amíg az eljárások képesek az adott kutatási program kémiai kiindulópontjául szolgáló szerkezeti kör azonosítására, a hamis pozitív és hamis negatív találatokat mint zajt kezeljük és számukat minimálisra szorítva jelenlétüket elfogadjuk.

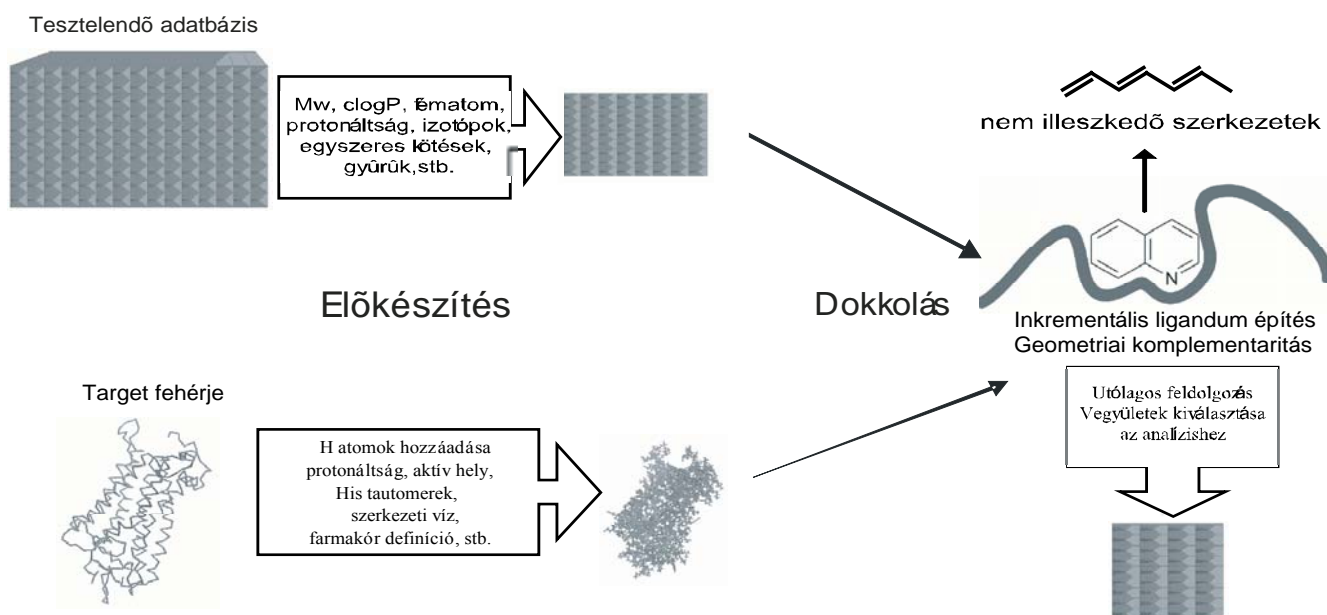
A HTS és VS módszerek közti leglényegesebb különbség az alkalmazott kísérleti illetve elméleti megközelítésben van. A VS elméleti megközelítéseinek köszönhetően 1) hatékonyabb költség-, idő-, és erőforrás-gazdálkodást eredményez; 2) a gyógyszereszerű vegyületek szélesebb – akár eddig elő nem állított – körének vizsgálatára alkalmas; 3) a random tesztelésnél nagyobb találati arányt eredményez. A közelmúlt összehasonlító vizsgálatai ugyanakkor nyilvánvalóvá tették, hogy a virtuális szűrés nem helyettesíti a HTS vizsgálatokat, ezek a módszerek valójában kiegészítik egymást¹⁻³. A HTS megközelítés a rendelkezésre álló vegyületek gyors vizsgálatára szolgál, míg a VS legfőbb alkalmazási területe az adott célfehérjére vagy célfehérje családra fókuszált vegyülettárak előállítására és a kereskedelmi forgalomban beszerezhető vegyületeket tartalmazó adatbázisok szűrése.

A VS módszerek, nagy áteresztőképességű molekulatervezési módszerek lévén, a CADD eljárásokhoz hasonlóan ligandum-, illetve szerkezet alapú megközelítés alapján dolgoznak. A ligandum-alapú módszerek a kismolekulás kémiai térben azonosított aktív illetve inaktív molekulák

alapján kísérlelnek meg új aktív vegyületeket azonosítani. A szerkezet alapú eljárások ezzel szemben a célfehérje 3D szerkezetének felhasználásával vizsgálják a potenciális ligandum és az aktív helyet alkotó aminosavak között kialakuló kölcsönhatásokat. Ebben az értelemben tehát a HTS módszerek in silico analógjai a szerkezet alapú VS eljárások, ezekkel foglalkozunk a továbbiakban.

2. Szerkezet alapú virtuális szűrővizsgálatok

Erlich szerint a gyógyszerhatás a gyógyszer és makromolekuláris célfehérje között kialakuló molekuláris felismerési folyamat eredménye. A racionális megközelítés ennek megfelelően igényli e folyamat szerkezet alapú kvalitatív és kvantitatív energetikai leírását⁴. A szerkezet alapú módszerek e kölcsönhatások vizsgálatát tűzik ki célul. A molekulatervezésben már régóta alkalmazott módszerek teljesítőképességét a korlátozott számítási teljesítmény mellett a jobbra csak alacsony felbontásban rendelkezésre álló fehérjeszerkezetek korlátozták. Mára mindkét területen jelentős előrelépés történt. A számítógépklaszterek és a Linux operációs rendszer olcsó számítási kapacitást kínál, míg a szerkezeti biológia a széles körben hozzáférhető szinkrotron forrásokkal a jobb minőségű szerkezeteket eredményez. Ebben a megváltozott környezetben a szerkezet alapú szűrővizsgálati módszerek gyógyszereszerű molekulák tömegét dokkolják a fehérje aktív helyére ahol a kötőmód predikcióját követően sor kerül a térbeli, elektrosztatikus és hidrofób illeszkedés értékelő-függvényeken alapuló kvantitatív vizsgálatára^{5,32-35}. Az így kapott pontszámok alkalmasak a vizsgált vegyületek affinitás alapú rangsorolására, amely alapját képezi a legkedvezőbb pontszámú vegyületek további, elsősorban kísérleti vizsgálatára. A szerkezet alapú virtuális szűrővizsgálatok folyamatábráját a 2. ábra mutatja.



2. Ábra. A szerkezet alapú szűrővizsgálatok legfontosabb lépései.

2.1. A célfehérje szerkezeti előkészítése

A szerkezet alapú szűrővizsgálatok első lépése a célfehérje szerkezetének előállítás. A nagy áteresztőképességű vizsgálatok általában nagy felbontású szerkezetet igényelnek, amelynek leginkább a röntgendiffrakcióval, illetve NMR-rel meghatározott szerkezetek felelnek meg leginkább. Számos esetben azonban, amikor a célfehérje nem, vagy csak nehezen kristályosítható, illetve mérete kívül esik az NMR mérésstartományon, homológiamodellek használatára is sor kerülhet⁶. Bár jelenleg az ismert fehérjék közel 56%-ról tudunk ilyen módon szerkezeti információt szerezni, ezek a modellek korlátozott felbontásuk miatt általában csak durva szűrőként alkalmazhatók. A közelmúltban, elsősorban GPCR fehérjékre⁷ bevezetett, ligandum-támogatott homológiamodellezés a modellek aktív helye környezetében elért minőségjavulás következtében reményeink szerint javíthat ezen a helyzeten⁸.

A nagy áteresztőképességű módszerek nem képesek a fehérje felszínének teljes feltérképezésére, így kezdeti feltételként igénylik az aktív hely meghatározását. Ezt a feladatot a szokásos kísérleti módszerek mellett (egykristály diffrakció, NMR, irányított mutációk) elméleti szinten leginkább a fehérjefelszín üregeit érzékelő algoritmusok alkalmazásával oldhatjuk meg^{9,10}, de azonosításuk fizikai-kémiai vagy geometriai leírók segítségével is lehetséges¹¹. Az azonosított aktív hely előkészítése alapvetően az alkalmazott dokkoló algoritmustól függ, általában magába foglalja a szükséges hidrogén atomok pótlását, a titrálható aminosavak protonálódási állapotának meghatározását, a megfelelő tautomer forma, valamint a kötésben szerepet játszó szerkezeti vízmolekulák azonosítását. Bár a nagy áteresztőképesség biztosítása érdekében a legtöbb dokkoló algoritmus az aktív helyet merev testként kezeli és a dokkolás során csak a ligandumok flexibilitásával számol, az aktív hely aminosav oldalláncainak flexibilitása a kötőkölsönhatások kialakulásában lényeges szerepet játszik. A napjainkban használatos módszerek a fehérje flexibilitását csak korlátok között képesek kezelni. Ennek következtében a megfelelő

célfehérje szerkezet kiválasztása elsődleges fontosságú, amennyiben a célfehérjéről több szerkezet is rendelkezésre áll, a kiválasztott szerkezetnek reprezentatívnak kell lennie. Több konformációs család esetén családonként egy szerkezet figyelembevétele indokolt.

2.2. Dokkoló algoritmusok

A leggyakrabban alkalmazott dokkoló algoritmusokat az 1. táblázat foglalja össze. Bár a dokkolás során csak a célfehérje aktív helyét vesszük figyelembe, az ezt kialakító aminosavak leírása az eredmény szempontjából egyáltalán nem közömbös. Az erre a célra általánosan alkalmazott térgeometriai deskriptorok magukban¹² vagy fizikai-kémiai jellemzőkkel kombinálva^{27,28}, illetve kölcsönhatási pontok vagy felszínek²⁹. A korai algoritmusok a ligandumokat is merev testként kezelték, manapság azonban a dokkoló programok többsége képes a potenciális ligandumok flexibilitásának figyelembevételére.

Erre a célra alapvetően két megközelítést használnak, ismertek a flexibilitást a teljes molekulán illetve annak fragmensein értelmező eljárások. A korai merev test dokkolások egyik tipikus kiterjesztése lehetővé teszi a konformációs flexibilitás figyelembevételét. A DOCK programban megvalósított merev test dokkolás során geometriai illesztési kritériumok segítségével illesztették a ligandumot az aktív hely „negatívjára”. Ezt a megoldást a későbbiek folyamán erőter alapú értékelő-függvénnyel, optimáló algoritmussal egészítették ki. Az ilyen módon továbbfejlesztett módszer képes kis energiájú kötőkonformációk azonosítására oly módon, hogy a ligandumok előre elkészített konformereit tartalmazó multi-konformer alkönyvtárakat dokkolnak az aktív helyre. Bár a multi-konformer dokkolás már képes a ligandumok flexibilitásának részleges figyelembevételére, ez a stratégia a ligandumok konformációs szabadsági fokainak növekedésével egyre körülményesebbé válik. Ezt a problémát kezelik hatékonyan a fragmens alapú módszerek,

amelyek a ligandumból képzett fragmenseket elkülönülten vagy inkrementálisan helyezik a fehérje aktív helyére.

Program	Módszer	Ref.
DOCK	Geometriai illesztés, inkrementális ligandépítés	12
FlexX	Geometriai illesztés, inkrementális ligandépítés	13,14
SLIDE	Geometriai illesztés, multikonformer ligandumkönyvtár	15
AutoDock	Energia alapú/sztochasztikus Monte Carlo szimulált hőkezelés	16
ICM	Monte Carlo minimálás, Pszeudo-Brown mintavételezés és lokális minimálás	17
QXP	Monte Carlo minimálás	18
MDD	Molekuladinamika	19
Glide	Szisztematikus keresés	20
GOLD	Genetikus algoritmus	21
PRO_LEADS	Tabu keresés	22
MOE-Dock	Tabu keresés vagy szimulált megeresztés	23
FRED	Szisztematikus keresés, multikonformer ligandumkönyvtárak	24,25
FLOG	Szisztematikus keresés	26

1. Táblázat. A legismertebb dokkoló algoritmusok

A szerkezet alapú virtuális szűrővizsgálatok egyik legnagyobb kihívása a fehérje flexibilitásának kezelése, amelyet több összefoglaló munka is tárgyal^{30,31}. A flexibilitás figyelembevétele a különböző megoldásokban igen változatos mértékben valósul meg. A jelenleg rendelkezésre álló módszerek a merev fehérjeszerkezet kismértékű változtatásán, az aktív hely fehérje oldalláncainak mozgatóján, illetve több fehérjeszerkezet egyidejű figyelembevételén alapulnak. A következőkben röviden ismertetjük a ligandum, illetve fehérje flexibilitás figyelembevételére alkalmas algoritmusokat.

2.2.1. Szisztematikus módszerek

Ezek a módszerek a ténylegesen meglévő konformációs szabadsági fokok teljes feltérképezését tűzik ki célul, azonban a szabadsági fokok számának növekedésével teljesítőképességük erősen korlátozottá válik:

$$N_{\text{konformer}} = \prod_{i=1}^N \prod_{j=1}^{n_{\text{inkrementum}}} \frac{360}{\theta_{i,j}}$$

ahol $\theta_{i,j}$ az i -ik kötéshez tartozó j forgási inkrementum, N a forgatható kötések száma. E probléma elkerülésére vezették be a ligandumok aktív helyen történő inkrementális felépítésének módszerét, amelyet a FlexX algoritmusban implementáltak. A FlexX algoritmus alapját az a felismerés képezi, amely szerint a ligandum fizikai-kémiai tulajdonságai alapvetően meghatározzák majdani elhelyezkedését a fehérje aktív helyén. Az eljárás a ligandum alapfragszékének

dokkolásával kezdődik, majd ennek kedvező orientációjában a teljes ligandum inkrementális felépítésével folytatódik. Az inkrementális építkezés alapját a gráfelmélet fakereső algoritmus szolgáltatja. Természetesen a fa kiterjedésének növekedésével a teljes keresés nem ésszerű, ezért az egyes csomópontokban a kötődési szabadentalpia értékelő-függvényekkel történő becslése alapján irányított keresést valósítunk meg.

A DOCK algoritmusban a ligandumokat egy merev alapfragszékre és flexibilis oldalláncokra bontják. A dokkolás során a szterikus komplementaritás optimalizálásával kapunk alacsony energiájú kötőkonformációkat.

A ligandumok előre elkészített konformációinak merev dokkolása szintén a szisztematikus módszerek közé sorolható, ez valósul meg a FLOG programban²⁶. Jóllehet ebben az esetben a ligandumok konformációs terének feltérképezése valójában nem az aktív helyen történik, így a generált konformációk relevanciája kérdéses lehet.

2.2.2. Sztochasztikus módszerek

A legismertebb sztochasztikus módszerek általában Monte Carlo keresést, genetikus algoritmusok alkalmazását vagy a „tabu search” eljárást használják fel. A Monte Carlo algoritmusok véletlenszerű kezdőkonformációkat állítanak elő a ligandumokra a fehérje aktív helyén, majd a kialakult kölcsönhatások pontozásával meghatározzák az adott konformáció jószágát. Ezt követően a konformer generálási és pontozási lépésekből álló ciklusok által előállított konfigurációkat a Metropolis kritérium alapján osztályozzák. Ennek megfelelően egy konfigurációt újnak fogadunk el, ha a hozzá tartozó pontérték kedvezőbb, mint a korábbi konfigurációra kapott érték. Amennyiben az új konfiguráció ezt a feltételt nem teljesíti, úgy az új konfigurációt csak abban az esetben tartjuk meg, ha a konfiguráció megtartásának valószínűsége (P) az alábbi egyenletben leírt Boltzmann-eloszlást követi:

$$P = \exp \left[\frac{-(E_{\text{új}} - E_{\text{régi}})}{kT} \right]$$

ahol P az új konfiguráció elfogadásának valószínűsége.

Ezt a koncepciót használja az ICM algoritmus¹⁷, amely egy globális Monte-Carlo minimáláson alapuló dokkolást valósít meg. Az algoritmus további gyorsítása érdekében a kölcsönhatási energiákat egy előre számított rács segítségével határozzák meg.

A genetikus algoritmusok (GA-k) az evolúciós verseny és a populációs dinamika alapelveit alkalmazzák számítási problémák megoldására. A GA-k valójában sztochasztikus optimalizációs módszerekahol az optimalizációs probléma általánosságban genetikai leírásban kerül tárgyalásra. A kezdeti állapotban az algoritmusok egy véletlenszerű populációt állítanak elő, amely az adott probléma lehetséges megoldásait reprezentálja. A populáció tagjait kromoszómák írják le, ahol a gének az optimalizációs probléma változóinak felelnek meg. Egy dokkolási feladat esetében a gének kódolják a molekula külső (transzlációs és rotációs), valamint belső konformációs szabadsági fokait (torziók).

A keresztezés és mutációk alapján kialakuló új populáció egyedeit ezt követően dekódoljuk és számítjuk az adott konformációra a kötődési szabadentalpiát. Ezt a folyamatot addig folytatjuk, amíg egy előre rögzített generációs számot elérünk. A GA alapú virtuális szűrés legjellemzőbb példája a GOLD algoritmusahol a különböző, hidrogénkötésben résztvevő szerkezeti részleteket a génekben direkt módon kódolják. A szintén népszerű AUTODOCK programban alkalmazott GA az energiaminimálást lokális kereséssel oldja meg.

A Monte Carlo és GA eljárások mellett a PRO_LEADS programban egy további heurisztikus algoritmust, a tabu search módszert találjuk. Ebben az eljárásban a véletlenszerűen generált orientációkat egy tabu lista alapján vizsgálják. Ez a lista előre elkészített megoldásokat tartalmaz és peremfeltételeket szolgáltat a keresés számára. Például a ligandum véletlenszerű elmozdulása akkor tiltott, ha ez olyan megoldáshoz vezet, amely magasabb energiájú és nem különbözik számottevően a korábban kapott, már tárolt megoldásoktól.

2.3. Értékelő-függvények

A ligandum és a fehérje aktív helye között kialakuló kölcsönhatások a virtuális szűrővizsgálatok során reverzibilis egyensúlyt eredményeznek, amelyet kvantitatív módon a kötődési szabadentalpia számításával írhatunk le. Sajnálatos módon a kötődési szabadentalpia még jelentős számítási kapacitást igénylő módszerek alkalmazásával is csak korlátozott pontossággal számítható. A virtuális szűrővizsgálatok során a vizsgált ligandumok nagy tömegére számítandó kötődési szabadentalpia racionális idő alatt csak becsléssel határozható meg, ezt a célt szolgálják a különböző értékelő-függvények. A leggyakrabban használt értékelő-függvényeket a 2. táblázatban foglaltuk össze. Valamennyi függvény közös jellegzetessége, hogy a kötődési szabadentalpia becsléséhez azt a közelítést használják fel, amely szerint a teljes entalpia különböző entalpiájárulékok összegeként írható fel. Fontos megjegyeznünk, hogy a kötődési szabadentalpia állapotfüggvény és mint ilyenre az additivitás szigorú értelemben nem teljesül.

A kötődési szabadentalpia becslésére szolgáló különböző értékelő-függvényeket három csoportra oszthatjuk, megkülönböztetünk erőtér alapú, empirikus, és tudásbázis alapú függvényeket.

Valamennyi ma ismert értékelő-függvény alkalmas a kötődési szabadentalpia becslésére, ugyanakkor mindegyikük számos korláttal bír. A pontatlanságok jórésze abból származik, hogy minden értékelő-függvényt kísérleti adatokhoz illesztettek, ezért értelemszerűen magukban foglalják a mérési hibát. Mivel az illesztésbe bevont kísérleti adatok mennyisége mindenképpen véges, elmondhatjuk, hogy univerzálisan alkalmazható értékelő-függvény nem áll rendelkezésre. Ugyanakkor az illesztésbe bevont röntgenszerkezetek csak kísérletileg mért kedvező kölcsönhatásokat tartalmaznak és nem hordoznak információt a kedvezőtlen kölcsönhatásokról, ami a értékelő-függvények használati értékét tovább korlátozza. Hasonló korlátokat támaszt a bizonytalan protonálódási állapot, valamint a szerkezeti vizek által közvetített kölcsönhatások figyelembevétele is.

2. Táblázat. A leggyakrabban használt értékelő-függvények

Típus	Függvény	Ref.
Erőtér-alapú	CHARMM	
	AMBER	
	LIE	
	OWFEG	36, 37
	MM PB/SA	
	GOLD/Gold(G)	
Empirikus	DOCK/Dock(D)	
	PLP	
	FlexX	
	ChemScore	14,38,39
	ScreenScore	
Tudásbázis-alapú	HINT	
	DrugScore	
	PMF	40,41,42
	BLEEP	

A függvények által szolgáltatott, a kötődési szabadentalpiával arányos pontszámok gyakran mutatnak összefüggést a molekulamérettel, jóllehet ezt a jellegzetességet a biológiai vizsgálatok nem erősítik meg. Mindezen hiányosságok részben csökkenthetők a különböző értékelő-függvények kombinációjának alkalmazásával. A konszenzus pontozásnak nevezett megoldás egyesíti a komponensként alkalmazott függvények egyes kölcsönhatástípusok leírásában mutatkozó előnyös sajátosságait és így általában jobb eredményt szolgáltat mint amilyen a komponensek egyedi alkalmazásával elérhető.

2.3.1. Erőtér-alapú értékelő-függvények

Az erőtérre alapuló függvények a rendszer energiáját a molekulamechanikából ismert energiatagokkal fejezik ki, ilyen például a DOCK és a GOLD pontozás. A fehérje-ligandum kölcsönhatás kvantitatív jellemzésére a molekulamechanikában használatos energiatagok közül elsőként a nemkötő kölcsönhatások leírására alkalmas tagokat használták fel. A ligandum és kötőhelye között fellépő kötési kölcsönhatási energia számítását leggyakrabban az elektrosztatikus és van der Waals tagok összegzésével becslik. Az ilyen elven működő erőtér alapú pontozási módszerek legnagyobb hiányossága, hogy a fehérje-ligandum kölcsönhatásokat eredetileg gázfázisú számításokra kifejlesztett összefüggésekkel írják le, amelyek elhanyagolják a biológiai rendszerekben meghatározó jelentőségű solvatációs- és entrópia-effektusokat. Szintén gyakori hiba a dokkolás során kialakuló sztérikus ütközések kezelése. Az erőtér alapú értékelő-függvények a van der Waals kölcsönhatások számítására általánosan alkalmazott Lennard-Jones 12-6 potenciál kis távolságoknál jelentkező nagy meredeksége miatt a közeli kölcsönhatásokra jelentős taszítással válaszolnak (pl. DOCK pontozás). Az ilyen félrevezető effektusok csökkentése érdekében a Lennard-

Jones tagot a korszerűbb értékelő-függvényekben skálázzák (pl. GOLD pontozás).

A GOLD értékelő-függvény különlegessége a H-hidas kölcsönhatások számítása. A pontozás figyelembe veszi a donorok és akceptorok deszolvatációját, a bevezetett diszperziós tag felelős a hidrofób kölcsönhatások leírásáért, míg hagyományos molekulamechanika energiátagek szolgálnak a ligandum belső energiájának meghatározására. Ennek következtében ez a értékelő-függvény elsősorban azokban az esetekben ad megbízható eredményt a ligandum és az aktív hely kölcsönhatásában a poláris kölcsönhatások meghatározóak.

A GOLD értékelő-függvény³⁶ a ligandum-fehérje kölcsönhatásra:

$$E_{vdW} + E_{H-bond} = \sum_{protein} \sum_{ligand} \left[\left(\frac{A_{ij}}{d_{ij}^8} - \frac{B_{ij}}{d_{ij}^4} \right) + (E_{da} + E_{ww}) - (E_{dw} + E_{aw}) \right]$$

a ligandum belső energiája pedig:

$$E_{vdW} + E_{torsion} = \sum_{ligand} \left(\frac{C_{ij}}{d_{ij}^{12}} - \frac{D_{ij}}{d_{ij}^6} \right) + \sum_{ligand} \frac{1}{2} V \left[1 + \frac{n}{|m|} \cos(n|\omega|) \right]$$

ahol A_{ij} , B_{ij} , C_{ij} és D_{ij} az i és j atomok van der Waals paraméterei, d_{ij} az i és j atom távolsága.

A DOCK függvény mind elektrosztatikus, mind pedig hidrofób kölcsönhatási tagot tartalmaz, azonban az entrópia-effektusokat nem veszi figyelembe³⁷. A poláris kölcsönhatásokat távolságfüggő dielektromos állandó segítségével vissza lehet skálázni.

$$E_{vdW} + E_{electrostatic} = \sum_{protein} \sum_{ligand} \left[\left(\frac{A_{ij}}{d_{ij}^{12}} - \frac{B_{ij}}{d_{ij}^6} \right) + 332 \frac{q_i q_j}{\epsilon(d_{ij}) d_{ij}} \right]$$

ahol $\epsilon(d_{ij})$ a távolságfüggő dielektromos állandó.

2.3.2. Empirikus értékelő-függvények

Az empirikus értékelő-függvények a kötődési szabadentalpia becslését a fehérje-ligandum kölcsönhatásban résztvevő explicit hidrogénkötések és hidrofób kontaktok súlyozott összegzésével valósítják meg, ilyen függvény például a LUDI, a FlexX és a ChemScore. A különböző megoldások nemcsak az adott kölcsönhatások számítására használt formalizmusban térnek el, hanem a különféle értékelő-függvényekben számos további kiegészítő tag bukkanhat fel. Így egyes függvények a hidrogén kötések leírására távolságfüggvényt használnak, amely az ideális geometriától való eltérést bünteti, míg más esetekben ezeket a kölcsönhatásokat szögfüggőként írják le. A kötődési szabadentalpia számításához feltesszük, hogy $\Delta G_{bind} \cong \sum \Delta G_i f_i$ ahol ΔG_i a megfelelő súlyfaktor, f_i pedig az i -ik kölcsönhatáshoz tartozó kölcsönhatási tag. A súlyfaktorokat kísérletileg meghatározott kötődési szabadentalpia értékek és 3D szerkezeti információk alapján regressziós analízissel állítják elő.

Az empirikus értékelő-függvények kétségtelenül legnagyobb előnye, hogy gyors és hatékony lehetőséget teremtenek a kötődési szabadentalpia becslésére. Ugyanakkor fontos megjegyezni, hogy ezek a értékelő-függvények az erőter alapú függvényekhez hasonló elhanyagolásokkal élnek, így a gyakorlati alkalmazás korlátai is hasonlóak. Az empirikus függvények teljesítőképessége természetesen függ a súlyfaktorok meghatározására felhasznált kísérleti adatkészlettől.

Az empirikus értékelő-függvényeket az entrópia-effektus becslése érdekében forgási taggal is kiegészíthetik, de a leggyakrabban használt függvények a szolvatációs effektusokat továbbra is csak részben képesek kezelni.

A LUDI pontozás a hidrogénkötő kölcsönhatásokat semleges és ionos jellegű kölcsönhatásokra bontja, míg a hidrofób kölcsönhatásokat a molekulafelszín számításával veszi figyelembe²⁹.

$$\Delta G_{köt} = \Delta G_{H-köt} \sum_{H-köt} f(\Delta R, \Delta \alpha) + \Delta G_{ionos} \sum_{ionos} f(\Delta R, \Delta \alpha) + \Delta G_{hidrofob} \sum_{hidrofob} |A_{hidrofob}| + \Delta G_{rotor} N_{rotor} + \Delta G_0$$

ahol $A_{hidrofob}$ a molekula hidrofób felszíne.

A széles körben alkalmazott FlexX pontozás az ideáltól eltérő hidrogén-hidas, ionos, aromás és egyéb kölcsönhatásokat geometria változókat tartalmazó büntetőpotenciálokkal írja le¹⁴.

$$\Delta G_{köt} = \Delta G_{H-köt} \sum_{H-köt} f(\Delta R, \Delta \alpha) + \Delta G_{ionos} \sum_{ionos} f(\Delta R, \Delta \alpha) + \Delta G_{aromás} \sum_{aromás} f(\Delta R, \Delta \alpha) + \Delta G_{kontakti} \sum_{kontakti} f(\Delta R, \Delta \alpha) + \Delta G_{rotor} N_{rotor} + \Delta G_0$$

A ChemScore tanító adatkészlete összesen 82 különböző típusú fehérje-ligandum komplexet tartalmazott³⁸. A függvény a hidrogénkötő és hidrofób kölcsönhatások mellett a fehérjékben gyakorta előforduló fémekkel kialakuló kölcsönhatásokat, valamint a ligandum flexibilitáscsökkenése következtében fellépő entrópia-effektust is figyelembe veszi.

$$\Delta G_{köt} = \Delta G_{H-köt} \sum_{H-köt} f(\Delta R, \Delta \alpha) + \Delta G_{fém} \sum_{fém} f(\Delta R, \Delta \alpha) + \Delta G_{lipo} \sum_{lipo} f(\Delta R) + \Delta G_{rotor} \sum_{rotor} f(P_{nl}, P'_{nl}) + \Delta G_0$$

2.3.3. Tudásbázis-alapú értékelő-függvények

A tudásbázis-alapú értékelő-függvények a fehérje-ligandum komplexekben implicit módon kódolt tudás hasznosítására jöttek létre. Az ilyen típusú függvények a Boltzmann törvény inverz megfogalmazására épülnek, amely szerint:

$$E_{ij} = -kT \ln(p_{ijk}) - kT \ln(Z)$$

ahol E_{ij} az átlagos potenciál, p_{ijk} a valószínűségi sűrűség, Z az állapotösszeg

Ezt a megközelítést arra lehet felhasználni, hogy fehérjeszerkezeteket tartalmazó adatbázisok felhasználásával olyan értékelő-függvényt állítsunk

elő, amely a fehérjeszerkezetek kvantitatív elemzését a szerkezetpredikciók aspektusából valósítja meg. Ennek érdekében az i , j , és k változókat a fehérje és ligandum atomtípusainak illetve a köztük lévő távolságoknak feleltetik meg. Az így definiált kölcsönhatások előfordulási gyakorisága pedig meghatározza az adott kölcsönhatás hozzájárulásának mértékét a kötődési szabadentalpiához. A széles körben használt DrugScore és a PMF (Potential of Mean Force) értékelő-függvények ezt a koncepciót valósítják meg:

DrugScore:

$$\Delta W = \gamma \sum_{\text{fehérje ligand}} \Delta W_{ij}(r) + (1-\gamma) \times \left[\sum_{\text{ligand}} \Delta W_i(SAS, SAS_0) + \sum_{\text{fehérje}} \Delta W_j(SAS, SAS_0) \right]$$

ahol SAS az oldószer által elérhető molekulafelszín, W_{ij} a távolságfüggő párpotenciál.

$$\text{PMF: } PMF = \sum_{\text{fehérje ligand}} A_{ij}(d_{ij})$$

$$A_{ij}(d_{ij}) = -k_B T \ln \left[f_{\text{terfogatkorr}}^j(r) \frac{\rho_{\text{seg}}^{ij}(r)}{\rho_{\text{tomb}}^{ij}} \right]$$

ahol k_B a Boltzmann állandó.

2.4. Utólagos feldolgozás

A jelenleg elterjedt értékelő-függvények az előzőekben ismertetett korlátok miatt nem alkalmasak a kötődési szabadentalpia megbízható, pontos számítására. Ezt a hiányosságot további, utólagos számításokkal kísérhetjük meg csökkenteni. E megközelítés szerint az egyszerű és gyors értékelő-függvények segítségével a részletesebb vizsgálatok számára egy racionális módon szűkített, a rendelkezésre álló számítási kapacitással feldolgozható dimenziójú problémát állítunk elő. Ezekben a számításokban már lehetőség nyílik a szolvatációs- és entrópia-effektusok megfelelő szintű figyelembevételére is. A szűrt ligandumokra elvégzett szabadenergia-perturbációs számítások, illetve a manapság népszerű lineáris válasz-elmélet alkalmazásával a kötődési szabadentalpiát szerencsés esetben 1-2 kcal/mol pontossággal becsülhetjük.

Az utólagos vizsgálatok egyik leggyakrabban alkalmazott módszere az újrapontozás, illetve a konszenzus pontozás⁴³. Az újrapontozás során annak lehetőségét vizsgáljuk meg, hogy az elsődleges értékelő-függvény segítségével kiválasztott kötőkonformációkat egy, vagy több további függvénnyel pontozva javítani tudjuk –e a kötődési szabadentalpia becslését. A konszenzus pontozás során olyan kötőkonformációkat választunk ki, amelyek nem egyetlen értékelő-függvénnyel mutatkoznak optimálisnak, hanem olyan konformációkat keresünk, amelyeket több különböző értékelő-függvény is kedvezőnek ítél. Ebben a megoldásban a különböző értékelő-függvényeket súlyfaktoruk függvényében azonos, illetve különböző súllyal vehetjük figyelembe. A konszenzus pontozást külön is alkalmazhatjuk a kötőkonformáció kiválasztására, illetve a ligandumok affinitás szerinti sorrendjének kialakítására is.

Az utólagos vizsgálatok egy speciális típusát az aktívnak jószolt vegyületek mesterséges dúsulásának elkerülésére

irányuló kísérletek. Ezek során a értékelő-függvények azon hiányosságát próbáljuk meg javítani, hogy általában a nagyobb méretű molekulák tényleges aktivitásuktól függetlenül is kedvező pontszámot kaphatnak. Szokásos eljárásként a vizsgálandó ligandumok fizikai kémiai paraméterek alapján történő szűrés jön szóba, gyakori azonban a pontszámok molekulatömeggel, illetve lipofilicitási paraméterekkel történő skálázása is⁴⁴.

Az utólagos vizsgálatok közül nem hagyhatjuk ki a leghatékonyabbnak becsült ligandumok kötőkonformációinak vizuális vizsgálatát sem. Ez a nehezen automatizálható lépés jelentős mértékben hozzájárul az eredmények értelmezéséhez, az esetleges szisztematikus hibák felismeréséhez és ahhoz, hogy a vezérmolekulakeresés fázisát követően értékes szerkezeti információkkal segítsük az optimalizálást.

2.5. Sikeres alkalmazások

A dokkoló algoritmusok és értékelő-függvények folyamatos fejlődése eredményeképpen mára számos a szerkezet alapú virtuális szűrővizsgálatoknak számos sikeres alkalmazása ismert (lásd még ref. 45-49). A 3. táblázatban ezeket igyekeztünk összegyűjteni megadva célfehérjét, az alkalmazott módszert, a tesztelt ligandumok forrását, valamint a szűrés hatékonyságára jellemző dúsítási faktorok (EF) ideális és ténylegesen elért mértékét. A dúsítási faktorokat a szűrt, rangsorolt adatbázis adott százalékának megfelelő vegyületre számítják, ezt az adott faktorok mellett minden esetben megadjuk.

3. Következtetések

Napjainkra a szerkezet alapú virtuális szűrés a nagy áteresztőképességű szűrővizsgálatok családjának sikeres tagjává vált. Ezek a módszerek az irodalom és saját tapasztalataink alapján (3, 54-57) a vezérmolekula azonosításának stádiumában jól kiegészítik a kísérleti megközelítéseket, különösen pedig a HTS vizsgálatokat. Ezek a tapasztalatok arra utalnak, hogy a virtuális szűrés, a rendelkezésre álló vegyülettáron alkalmazva, az adott célfehérjére fókuszált alkonyvtárak kialakításával jelentős mértékben javíthatja a HTS hatékonyságát. Másfelől a virtuális szűrés alkalmas az adott kutatóhelyen meg nem található, de kereskedelmi forrásból beszerezhető vegyületek vizsgálatára is. Ez utóbbi alkalmazás lehetővé teszi, hogy a vezérmolekula keresés ne csupán a rendelkezésre álló vegyületek által lefedett kémiai térre korlátozódjon.

Bár a sikeres alkalmazások a metodika hasznosságát egyértelműen jelzik, a módszer korlátairól is fontos megemlékeznünk. A jelenleg rendelkezésre álló dokkoló algoritmusok legnagyobb hiányossága a fehérje flexibilitásának hiányos kezelése. A értékelő-függvények teljesítőképessége a kívánalmaktól elmarad, a szolvatációs- és entrópia-effektusok figyelembevétele, valamint a mért kötődési affinitásértékkel való korreláció javítása változatlanul a kutatások középpontjában áll. Végezetül nem elhanyagolható, hogy a szerkezet alapú virtuális szűrővizsgálatok továbbra is jelentős számítási kapacitást, szoftver és hardver oldalról jelentős ráfordítást igényelnek.

3. Táblázat. A szerkezet alapú virtuális szűrővizsgálatok sikeres alkalmazása (A: az aktív vegyületek száma, EF: a dúsítási faktor)

Célfehérje	Program	Jellemzők				Ref.	
		Szűrt adatbázis	A	EF	Ideális EF		EF/Ideális EF
Timidine kináz	DOCK	ACD: 990; (szűrők: M_w reagensek, inorganic), véletlen válogatás, Gasteiger-Marsili töltések	10	7(10)	10	0.7	50
	FlexX			8(10)	10	0.8	
	GOLD			10(10)	10	1	
Estrogen receptor alfa	DOCK		10	7(10)	10	0.7	
	FlexX			8(10)	10	0.8	
	GOLD			10(10)	10	1	
Protein Tirozin Foszfátáz-1B	DOCK	ACD: 179805; (szűrők: fémek, izotópok, clogP), véletlen válogatás, Gasteiger-Marsili töltések, ionizált	1327	7(10)	10	0.7	51
	FlexX			9(10)	10	0.9	
	Glide			9(10)	10	0.9	
Protein Kináz B	DOCK		266	0.1(10)	10	0.01	
	FlexX			1(10)	10	0.1	
	Glide			5(10)	10	0.5	
Ciklooxygenáz-2	FRED	WDI: 7528; (szűrők: M_w hasonlóság, forgatható kötések, makrociklus és szteroid kizárása, farmakofor pontok, ionizált pH 7)	128	5(10)	10	0.5	52
	FlexX			6.2(10)	10	0.62	
	Glide			6.3(10)	10	0.63	
Estrogen Receptor béta	FRED		55	7.9(10)	10	0.79	
	FlexX			5.8(10)	10	0.58	
	Glide			7.5(10)	10	0.75	
p38 MAP kináz	FRED		72	1.9(10)	10	0.19	
	FlexX			7.9(10)	10	0.79	
	Glide			4.1(10)	10	0.41	
Giráz B	FRED		36	4.1(10)	10	0.41	
	FlexX			7(10)	10	0.7	
	Glide			5.8(10)	10	0.58	
Trombin	FRED		67	7.9(10)	10	0.79	
	FlexX			8.1(10)	10	0.81	
	Glide			8.3(10)	10	0.83	
Gelatináz A	FRED		43	4.1(10)	10	0.41	
	FlexX			6.3(10)	10	0.63	
	Glide			6.1(10)	10	0.61	
Neuraminidáz	FRED	51	8.3(10)	10	0.83		
	FlexX		6.2(10)	10	0.62		
	Glide		9.8(10)	10	0.98		
Ciklooxygenáz-2	FlexX	128	20(1)	68	0.29	53	
Estrogen Receptor beta	FlexX	55	39(1)	100	0.39		
p38 MAP kináz	FlexX	72	18(1)	100	0.18		
Giráz B	FlexX	36	21(1)	100	0.21		
Trombin	FlexX	67	25(1)	100	0.25		
Gelatináz A	FlexX	43	16(1)	100	0.16		
Neuraminidáz	FlexX	51	22(1)	100	0.22		
Neuraminidáz	GOLD	ATLAS (vállalati adatbázis): ~3.1 millió	15	43(1)	100		0.43
PTF-1B			25	20(1)	100	0.2	
CDK-2 kináz $M_w < 250$			41	10(1)	100	0.1	
CDK-2 kináz $M_w > 250$			23	23(1)	100	0.23	
Estrogen Receptor béta (agonisták)			20	60(1)	100	0.6	
Estrogen Receptor béta (antagonisták)?			17	58(1)	100	0.58	

Mindezek alapján megállapíthatjuk, hogy a szerkezet alapú virtuális szűrés a vezérmolekula azonosításának kétségkívül hasznos, de távolról sem kiváltságos módszere. Meggyőződésünk, hogy legnagyobb hatékonyságát a már alkalmazásban lévő metodikákkal, a HTS-sel, a kémiai analóg kutatással és a de novo tervezéssel kombinálva érheti el.

Köszönetnyilvánítás

Köszönettel tartozom Polgár Tímeának a kézirat elkészítése során nyújtott segítségéért.

Hivatkozások

- Doman, T. N.; McGovern, S. L.; Witherbee, B. J.; Kasten, T. P.; Kurumbail, R.; Stallings, W. C.; Connolly, D. T.; Shoichet, B. K. *J. Med. Chem.* **2002**, *45*, 2213-21.
- Paiva, A. M.; Vanderwall, D. E.; Blanchard, J. S.; Kozarich, J. W.; Williamson, J. M.; Kelly, T. M. *Biochim. Biophys. Acta.* **2001**, *1545*, 67-77.
- Polgár, T.; Baki, A.; Szendrei-Ignác, Gy.; Keserű, G. M. *J. Med. Chem.* beküldve
- Böhm, H.-J.; Klebe, G. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2588 – 2614.
- Paul D. Lyne. *Drug Discov. Today* **2002**, *7*, 1047-1055.
- Hillish, A.; Pineda, L.F.; Hilgenfeld, R. *Drug Discov. Today* **2004**, *9*, 659-669.
- Becker, O. M.; Y.Marantz, Y.; Shacham, S.; Inbal, B.; Heifetz, A.; Kalid, O.; Bar-Haim, S.; Warshaviak, D.; Fichman, M.; Noiman, S. **2004**, *PNAS*, *101*, 11304 – 11309.
- Evers, A.; Klebe, G. *Angew. Chem. Int. Ed.* **2004**, *43*, 248-251.
- Hendlich, M.; Rippmann, F.; Barnickel, G. *J. Mol. Graph. Model.* **1997**, *15*, 359-363.
- Brady, G.P.; Stouten, P.F. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 383-401.
- Schmitt, S.; Hendlich, M.; Klebe, G. *Homology.* **2001**, *40*, 3141 – 3144.
- Kuntz, I.D.; Blaney, J.M.; Oatley, S.J.; Langridge R.; Ferrin T.E. *J. Mol. Biol.* **1982**, *161*, 269-288
- Hindle, S. A.; Rarey, M.; Buning, C.; Lengau, T. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 129-149.
- Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. *J. Mol. Biol.* **1996**, *261*, 470-489.
- Schnecke, V.; Swanson, C.A.; Getzoff, E.D.; Tainer, J.A.; Kuhn, L.A. *Proteins* **1998**, *33*, 74-87.
- Goodsell, D.S.; Olson A.J. *Proteins* **1990**, *8*, 195-202.
- Abagyan, R.; Totrov, M.; Kuznetsov, D. *J. Comput. Chem.* **2004**, *15*, 488 – 506.
- McMartin, C.; Bohacek, R.S. *J. Comput.-Aided Mol. Design.* **1997**, *11*, 333 – 344.
- Di Nola, A.; Roccatano, D.; Berendsen, H.J.C. *Proteins* **1994**, *19*, 174-182.
- Halgren, T.A.; Murphy, R.B.; Friesner, R.A.; Beard, H.S.; Frye, L.L.; Pollard, W.T.; Banks, J.L. *J. Med. Chem.* **2004**, *47*, 1750-1759.
- Friesner, R.A.; Banks, J.L.; Murphy, R.B.; Halgren, T.A.; Klicic, J.J.; Mainz, D.T.; Repasky, M.P.; Knoll, E.H.; Shaw, D.E.; Shelley, M.; Perry, J.K.; Sander, L.C.; Shenkin, P.S. *J. Med. Chem.* **2004**, *47*, 1739-1749.
- Jones, G.; Willett, P.; Glen, R.C. *J. Mol. Biol.* **1995**, *245*, 43-53.
- MOE (Molecular Operating Environment), Chemical Computing Group Inc., 1010 Sherbrooke St. West, Suite 910, Montreal, Quebec, H3A 2R7, Canada.
- Baxter, C.A.; Murray, C.W.; Clark, D.E.; Westhead, D.R.; Eldridge M.D. *Proteins* **1998**, *33*, 367-382.
- FRED, OpenEye Scientific Software, 3600 Cerrillos Rd., Suite 1107, Santa Fe, NM 87507
- Miller, M.D.; Kearsley, S.K.; Underwood, D.J.; Sheridan, R.P. *J. Comput.-Aided Mol. Design* **1994**, *8*, 153-174.
- Jianga, F.; Kimb, S.-H. *J. Mol. Biol.* **1991**, *219*, 79-102.
- Burkharda, P.; Taylora, P.; Walkinshawa, M.D. *J. Mol. Biol.* **1998**, *277*, 449-466.
- Böhm, H.-J. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 61-78.
- Carlson, H.A.; McCammon, J.A. *Mol. Pharmacol.* **2000**, *57*, 213-218.
- Claussen, H.; Buning, C.; Rarey, M.; Lengauer, T. *J. Mol. Biol.* **2001**, *308*, 377-395.
- Taylor, R.D.; Jewsbury, P.J.; Essex, J.W. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 151-166.
- Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R. *Proteins* **2002**, *47*, 409-443.
- Schneider G.; Böhm, H.-J. *Drug Discov. Today.* **2002**, *7*, 64-70.
- Kitchen, D.B.; Decornez, H.; Furr, J.R.; Bajorath, J. *Nature Rev. Drug Discov.* **2004**, *3*, 935-949.
- Vieth, M.; Hirst, J.D.; Kolinski, A.; Brooks, C.L. *J. Comput. Chem.* **1998**, *19*, 1612-1622.
- Shoichet, B.K.; Leach, A.R.; Kuntz, I.D. *Proteins* **1999**, *34*, 4-16.
- Eldridge, M.D.; Murray, C.W.; Auton, T.R.; Paolini, G.W.; Mee, R.P. *J. Comput.-Aided Mol. Design.* **1997**, *11*, 425-445.
- Stahl, M.; Rarey, M. *J. Med. Chem.* **2001**, *44*, 1035-1042.
- Gohlke, H.; Hendlich, M.; Klebe, G. *J. Mol. Biol.* **2000**, *295*, 337-356.
- Muegge, I.; Martin, Y. C. *J. Med. Chem.* **1999**, *42*; 791-804.
- Mitchell, J.B.O.; Laskowski, R.A.; Alex, A.; Thornton, J.M. *J. Comput. Chem.* **1999**, *20*, 1165-1176.
- Clark, R.D.; Strizhev, A.; Leonard, J.M.; Blake, J.F.; Matthew, J.B. *J. Mol. Graph. Mod.* **2000**, *20*, 281-295.
- Verdonk, M.L.; Berdini, V.; Hartshorn, M.J.; Mooij, W.T.M.; Murray, C.W.; Taylor, R.D.; Watson, P. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 793-806.
- Muegge, I.; Enyedy, J. *Curr. Med. Chem.* **2004**, *11*, 693-707.
- Davis, T.G.; Pratt, D.J.; Endicott, J.A.; Johnson, L.N.; Noble, M.E.M. *Pharm. Therap.* **2002**, *93*, 125-133.
- Furet, P.; Caravatti, G.; Lydon, N.; Priestle, J.P.; Sowadski, J.M.; Trinks, U.; Traxler, P. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 465-472.
- Lyne, P. D.; Kenny, P. W.; Cosgrove, D. A.; Deng, C.; Zabudoff, S.; Wendoloski, J. J.; Ashwell *J. Med. Chem.* **2004**, *47*, 1962-1968.
- Barril, X.; Hubbard, R.E.; Morley, S.D. *Mini Rev. Med. Chem.* **2004**, *4*, 779-791.
- Bissantz, C.; Folkers, G.; Rognan, D. *J. Med. Chem.* **2000**, *43*, 4759-4767.
- Klon, A.E.; Glick, M.; Thoma, M.; Acklin, P.; Davies, J.W. *J. Med. Chem.* **2004**, *47*, 2743-2749.
- Schulz-Gasch, T.; Stahl, M. *J. Mol. Mod.* **2003**, *9*, 47-57.
- Stahl, M.; Rarey, M. *J. Med. Chem.* **2001**, *44*, 1035-1042.
- Bauer, A.; Kovári, Z.; Keserű, G. M. *J. Mol. Struct. THEOCHEM* **2004**, *676*, 1-7.
- Virág, I.; Polgár, T.; Keserű, G. M. *J. Mol. Struct. THEOCHEM* **2005**, *725*, 239-242.
- Polgár, T.; Keserű, G. M. *J. Med. Chem.* **2005**, *48*, 3749-3755.
- Polgár, T.; Baki, A.; Szendrei, Gy.; Keserű, G. M. *J. Med. Chem.* **2005**, <http://dx.doi.org/10.1021/jm050504d>.

Structure-based virtual screening

Recent improvements in genomics yielded a substantial increase of potential, novel targets for drug discovery. Whereas combinatorial chemistry and high-throughput screening were supposed to produce a vast number of new drugs, neither of them has met fully the expectations. Virtual screening applies rule-based, property-based, pattern recognition and docking methods to screen large databases of virtual compounds. Hereby the number of compounds to be

assayed is drastically decreased and virtual screening can speed up the lead discovery process. This review summarizes the recent developments in structure-based virtual screening focusing the problems of docking algorithms and scoring functions. A collection of success stories is enclosed for supporting further applications in the lead discovery field.