

Modelowanie komputerowe w badaniach nad lekiem – projektowanie i poszukiwanie cząstki aktywnej, ocena właściwości fizykochemicznych oraz aktywności biologicznej

Sebastian Polak, Barbara Wiśniowska

Pracownia Farmakoepidemiologii i Farmakoekonomiki Katedry Toksykologii CM UJ w Krakowie

Adres do korespondencji: Sebastian Polak, Pracownia Farmakoepidemiologii i Farmakoekonomiki, Katedra Toksykologii, Wydział Farmaceutyczny Collegium Medicum, Uniwersytet Jagielloński, ul. Medyczna 9, 30-688 Kraków

Computer modeling in drug investigation – projecting and research for active molecule, physico-chemical properties and biological activity evaluation

Pharmaceutical industry is one of the highly investing in research and development branch of industry. Such a situation is caused by both – their specificity (science depending) and world market competition. Being forced to fulfil growing expectations from all sides – patients who needs more effective and safe therapies, insurance companies and health care systems as they expect cost-effective drugs, pharmaceutical scientists are constantly looking for the new research methods and techniques. Fast development of the chemical and mathematical modelling techniques, effectively supported by the dramatic increase of the computer computational power, together with relatively low costs of implementation results in their wide usage in pharmaceutical science and industry. It begins at the level of new chemicals search and development, through activity, toxicity and safety assessment to the clinical studies modelling. Virtual pharmacy has become a fact.

Keywords: new research methods, in silico, computational pharmacy

© Farm Pol, 2009, 65(3): 214-223

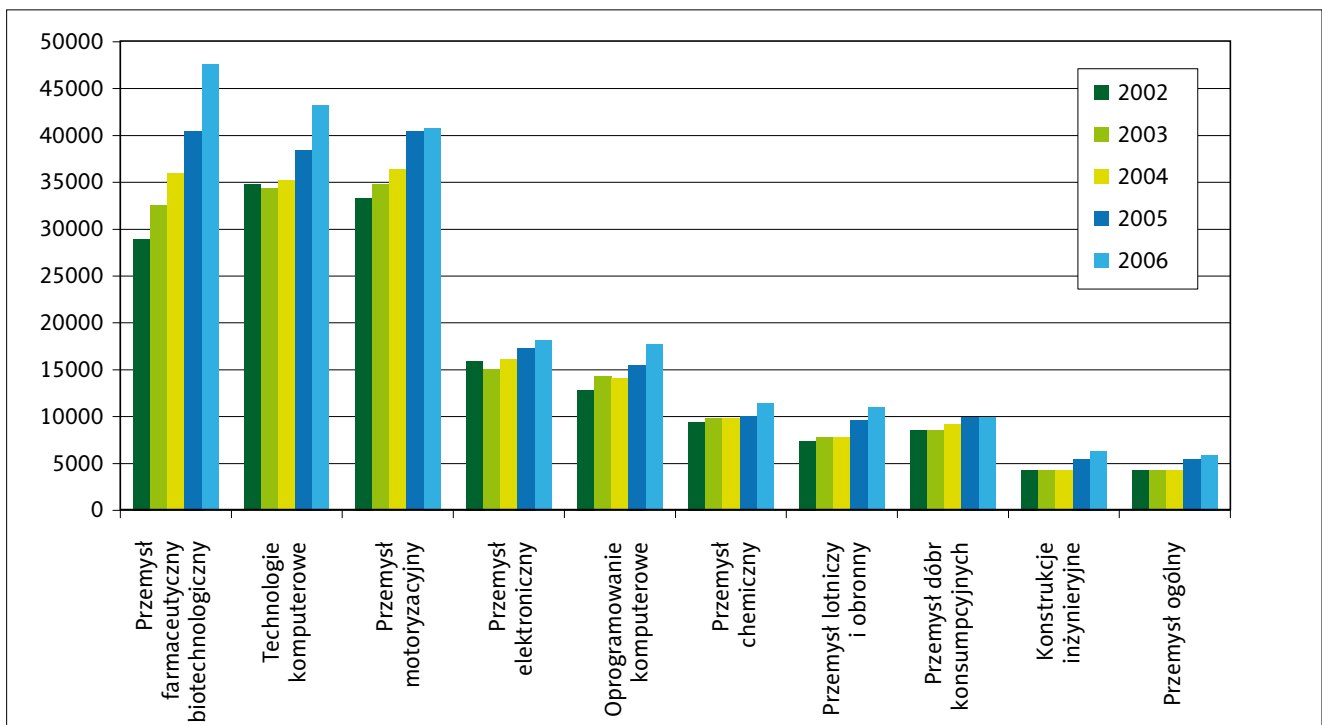
Metody poszukiwania nowych substancji o potencjalnym działaniu terapeutycznym i metodyka badań nowych produktów leczniczych przeszły w ciągu ostatnich dziesięcioleci ewolucję, podobną do wręcz rewolucyjnych zmian sposobów komunikacji, jaką możemy obserwować na co dzień, używając telefonów komórkowych czy też korzystając z możliwości oferowanych przez technologie internetowe. Praktyczne wykorzystanie technik informatycznych – podobnie jak w nowoczesnych systemach wymiany informacji – pozwoliło poszerzyć zakres poszukiwań i zmienić ich charakter, i choć zmiany w obu wymienionych dziedzinach zachodzą z inną dynamiką, m.in. ze względu na silne przyzwyczajenia i konserwatyzm środowisk naukowych, to wydaje się, że podobnie jak w przypadku świata telekomunikacji,

również dział farmacji zajmujący się badaniem i rozwojem leków wkracza nieodwracalnie na ścieżkę cyfrową.

W serii kilku artykułów, które chcielibyśmy Państwu zaprezentować, przedstawimy najważniejsze technologie informatyczne stosowane w trakcie prac nad nowymi substancjami chemicznymi o potencjalnym znaczeniu terapeutycznym. Nie opisując dokładnie długotrwałego, pracochłonnego i w swoim opisie wymagającego dużej staranności procesu badań nad lekiem, postaramy się przedstawić i odnieść do tych jego elementów, w których techniki komputerowe stały się narzędziami co najmniej równoważnymi, a niekiedy bardziej znaczącymi niż podejście tradycyjne. Całość cyklu jest dopełnieniem szkolenia, które zostało opublikowane w systemie szkoleń podyplomowych dla farmaceutów e-duk@cja (<http://www.e-dukacja.pl>) i wychodzi poza jego zakres.

Przemysł farmaceutyczny jest jedną z najintensywniej inwestujących w naukę i badania gałęzi gospodarki. Ocenia się, iż piątą część wśród firm przeznaczających najwięcej na rozwój stanowią firmy farmaceutyczne (ok. 15% zysków), które zdystansowały przemysł technologiczny oraz oprogramowania (ok. 10% zysków) (rycina 1).

Wprowadzenie na rynek nowego leku, który może trafić na półki apteczne, to kosztowny, wymagający wielu lat pracy proces. Ocena kosztów nie jest prosta, a końcowy wynik w dużej mierze zależy od metodyki przyjętej przez przeprowadzającego daną analizę badacza i przede wszystkim informacji, którymi dysponuje, a które nigdy nie są kompletne, choćby ze względu na konsekwentnie prowadzoną politykę ochrony danych finansowych firm farmaceutycznych (rycina 2). Bez względu jednak na uzyskaną kwotę końcową, wahającą się w granicach 300–900 mln dolarów, należy ją ocenić jako niebagatelną nawet

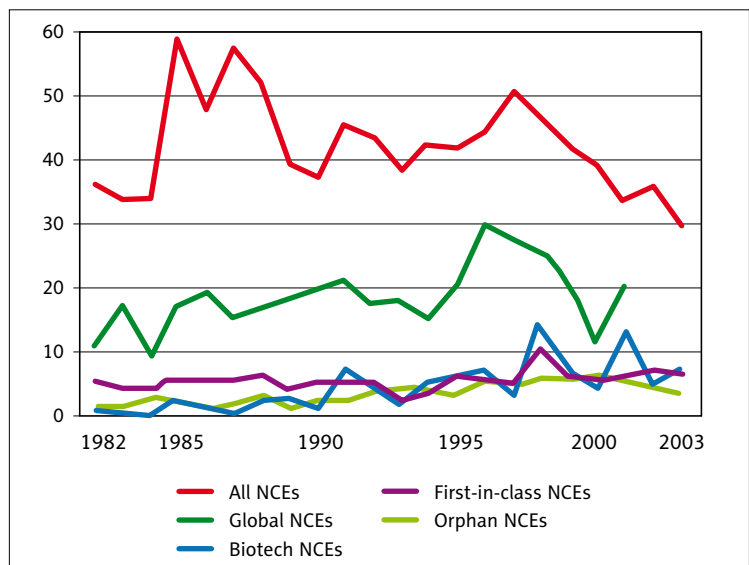


Rycina 1. Porównanie wydatków na badania i rozwój (R&D) dla różnych gałęzi przemysłu [7]

dla najzamożniejszych graczy rynku leków i to wraz z założeniem, iż jest ona rozłożona na cały 15–20-letni okres badań nad lekiem.

Koszty te nie rozkładają się równomiernie na poszczególne etapy badań potencjalnej substancji leczniczej i wdrażania leku, jako że większość pochłania przygotowanie i przeprowadzenie całości procedury badań klinicznych, zgodnie z obowiązującymi – coraz surowszymi zarówno pod kątem naukowym, jak i etycznym – procedurami i zasadami (rycina 3). Powyższe stwierdzenia są jednak prawdziwe tylko wtedy, gdy oceniany koszt jest przeliczany na jedną cząstkę aktywną. Należy pamiętać, że do fazy badań klinicznych dociera jedynie niewielki odsetek testowanych na wstępnym etapie molekuł, a jeszcze mniej jest wprowadzanych do obrotu (ocenia się, że w Stanach Zjednoczonych dziewięć z dziesięciu substancji poddawanych badaniom klinicznym nie jest wdrażane na rynek), tak więc jeśli kalkulacja zostanie wykonana zostanie pod kątem czasu przeznaczanego na poszczególne etapy, koszt poszczególnych części określanych jako przedkliniczne i kliniczne wyrównuje się.

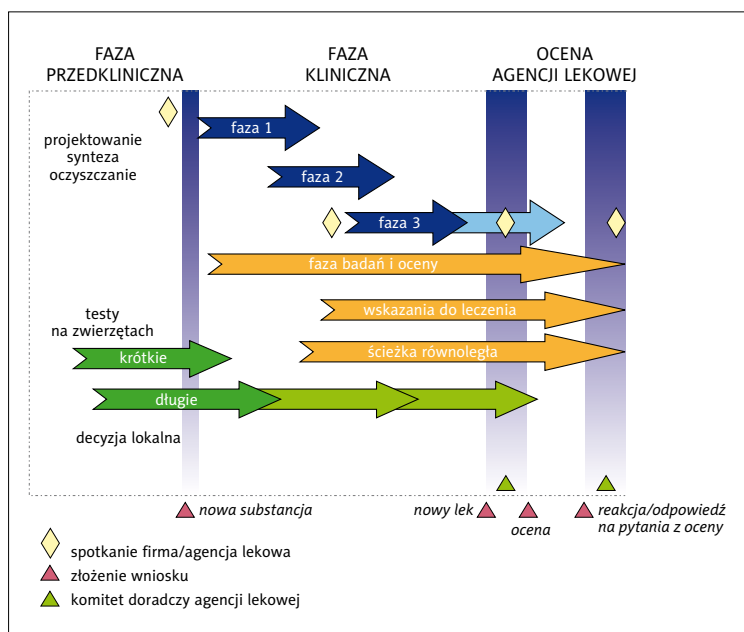
Powyższe informacje powinny zostać zestawione z wciąż malejącą liczbą rejestracji, rozumianych jako wprowadzenie na rynek nowych cząsteczek, a więc leków określanymi jako innowacyjne (w przeciwieństwie do wciąż rosnącego rynku leków generycznych) (rycina 4). Trend ten dotyczy również liczb bezwzględnych, jednak jest szczególnie wyraźnie widoczny w zestawieniu z wydatkami przeznaczanymi przez firmy farmaceutyczne na badania nad nowymi lekami (tzw. R&D – *Research and Development*).



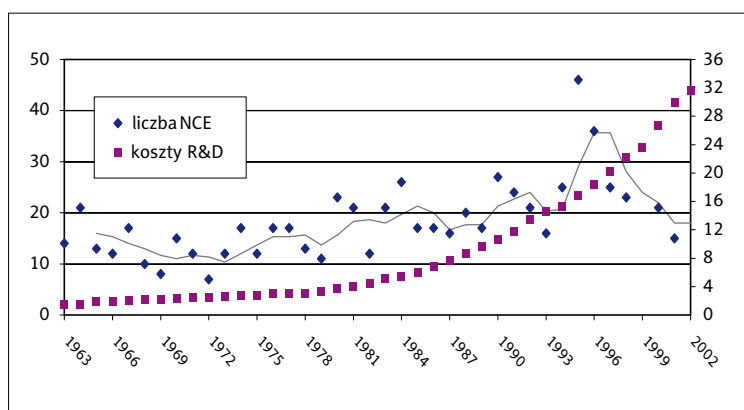
Rycina 2. Liczba nowych substancji leczniczych zarejestrowanych w latach 1982–2003 [1]

Jest to sytuacja, której przyczyn należy upatrywać m.in. w rosnących wymogach dotyczących bezpieczeństwa prowadzenia badań klinicznych, ich kosztach oraz sytuacji wyczerpania klasycznych źródeł nowych substancji oraz niezwykle kosztownych i długotrwałych badań związanych z makrocząsteczkami (m.in. białka – np. insulina i jej pochodne, hormony, szczepionki itd.).

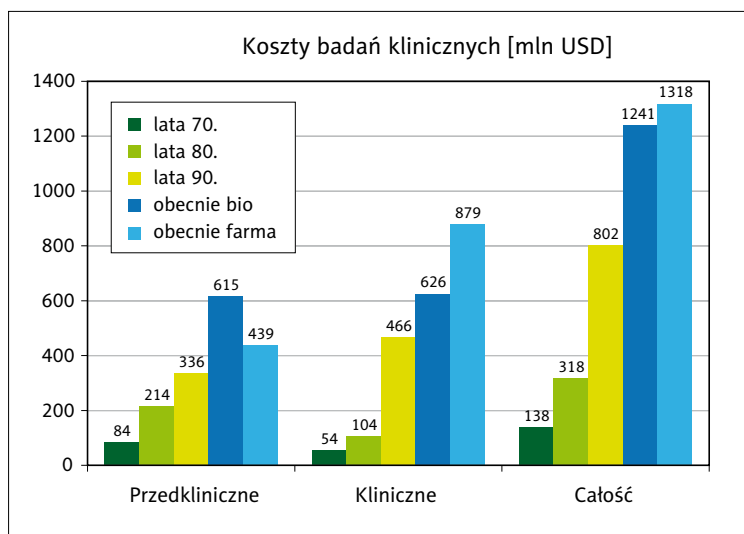
Najważniejszą, choć stosunkowo mało odkrywczą konkluzją wyciąganą z powyższych danych jest konieczność redukcji kosztów oraz – jak widać na rycinie 5, w sposób bezpośredni przeliczanego na



Rycina 3. Etapy wdrażania leku



Rycina 4. Zmiana kosztów wdrożenia leku w skali lat w zestawieniu z liczbą nowych rejestracji [3].



Rycina 5. Zmiana kosztów wdrożenia leku w skali lat – porównanie kosztów badań dla przemysłu farmaceutycznego i biotechnologicznego [4]

wartości monetarne – czasu, stąd coraz częstsze sięganie po komputerowe metody wspomaganie projektowania i wdrażania nowych leków. Modelowanie i symulacje są wykorzystywane na każdym etapie badań nad lekiem, a kolejne artykuły obejmują opis wybranych kroków wraz z przykładami stosowanych metod modelowania oraz realizującymi je programami.

Kolejnym argumentem przemawiającym za szeroką i wciąż rosnącą gamą modeli realizowanych w warunkach *in silico*, są względy humanitarno-etyczne wskazujące na konieczność uzyskania maksymalnej możliwej redukcji doświadczeń przeprowadzanych na zwierzętach oraz optymalizacja koniecznych badań prowadzonych na ludziach (zdrowych ochotnikach oraz osobach chorych, w warunkach klinicznych). Z pewnych przyczyn ten pierwszy element – badania na zwierzętach – rodzi silne i niekiedy skrajne emocje, od całkowitego zawieszenia wyników uzyskiwanym na ich podstawie, aż po całkowitą negację i krytykę. Temat ten stanowi zupełnie osobne zagadnienie, które – mimo dość dużej, choć sezonowej popularności – wymagałoby pełnego, rzetelnego i obiektywnego opracowania, niemniej jednak w dalszych częściach naszej pracy będziemy się do niego odnosić, stąd krótki akapit poświęcony temu problemowi.

Warto przypomnieć, że ścisłe standardy badań nad lekiem, w tym obowiązek przeprowadzania doświadczeń na zwierzętach zostały wprowadzone w roku 1968 (amerykański *Medicines Act*), jako wynik tragedii związanych ze stosowaniem talidomidu u ciężarnych kobiet. Jako lek uważany za bezpieczny, był on podawany w przypadkach porannych mdłości i wymiotów. Niestety niepożądanym efektem jego stosowania jest zaburzenie rozwoju zawiązków kończyn górnych płodu. Wyniki stosowanych obecnie testów na zwierzętach potrafią wskazać na możliwość działania teratogennego badanych związków chemicznych. Zwolennicy wykorzystywania zwierząt laboratoryjnych w procesie badań nad lekiem podają przykłady zagadnień, których dogłębna analiza – zgodnie z rosnącymi wymaganiami wdrażanymi przez organizacje zajmujące się nadzorem nad rynkiem leków – musi znaleźć się w składanym dossier rejestracyjnym, a których przeprowadzenie w innych warunkach jest albo niemożliwe, albo co najmniej niezwykle trudne. Kilka z nich to m.in. problem przenikania przez barierę krew-mózg i penetrowanie ośrodkowego układu nerwowego czy też analiza wchłaniania leków w przewodzie pokarmowym. Jednocześnie jednak, mimo malejącej liczby zwierząt poświęcanych w trakcie badań (w samej Wielkiej Brytanii, w której centra badawcze mają wszystkie największe firmy farmaceutyczne, ich liczba spadła z ok. 6 milionów rocznie w latach 70 do ok. 2,5 miliona na początku lat 2000), grono sceptyków ich powszechnego stosowania nie maleje. Pozostawiając na boku argumenty niemerytoryczne, wymieńmy najważniejsze, które mają mocną podstawę merytoryczną, wśród nich najważniejsze wydają się być dwa. Przede wszystkim

gwałtowny rozwój metod alternatywnych obejmujących badania przeprowadzane w warunkach:

- *in vitro*, a obejmujących badania z wykorzystaniem hodowli komórkowych czy kultur tkankowych; wzrost ich znaczenia jest ściśle powiązany z rozwojem metod oznaczeń substancji chemicznych w materiale biologicznym (wśród nich przede wszystkim wysokosprawną chromatografię cieczową z detektorem masowym)
- *in silico* – zawierających w sobie całą gamę modeli, które zostaną szczegółowo opisane poniżej oraz w kolejnych publikacjach
- *in vivo* – u ludzi i o ile nie są i nie będą one w stanie całkowicie zastąpić pewnych elementów badań m.in. z przyczyn wymienionych powyżej o tyle nowe techniki ich przeprowadzania (np. microdosing) pozwalają drastycznie zredukować liczbę zwierząt laboratoryjnych, bez zwiększenia ryzyka zaangażowanych ochotników;

Drugi argument, choć wydaje się być dość oczywisty, jest często niestudnie bagatelizowany, a podejmuje niemożliwe do usunięcia różnice międzygatunkowe, utrudniające, a niekiedy uniemożliwiające ekstrapolowanie wyników badań wykonanych na zwierzętach bezpośrednio na ludzi. Pewnym wyjściem jest wykorzystywanie tych gatunków, które genetycznie najmniej różnią się od genomu ludzkiego (świnie, małpy), co jednak nie rozwiązuje tego problemu całkowicie. Być może i tutaj rozwój technik biotechnologicznych pozwoli np. na humanizowanie pewnych narządów u zwierząt i obserwowanie skutków zastosowania substancji chemicznej w warunkach narządów ludzkich, niemniej jednak jest to wciąż stosunkowo odległa przyszłość.

Całość tego typu działań określa się jako *3Rs'* (ang. *Replace, Reduce, Refine*), a wśród nich coraz większą rolę odgrywają techniki modelowania komputerowego.

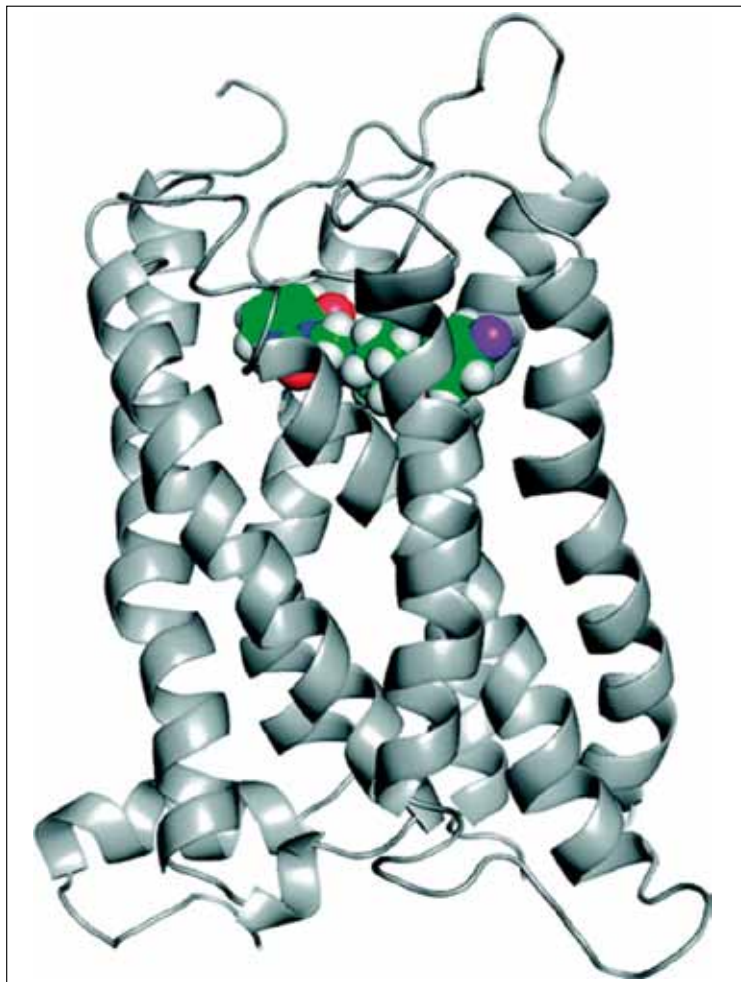
Projektowanie i poszukiwanie cząsteczki aktywnej

Początek pierwszego etapu poszukiwań nowej cząsteczki, mogącej mieć zastosowanie w leczeniu nie jest prosty do określenia. Punktem wyjścia może być zarówno choroba i dokładne poznanie mechanizmu patologii, jak i teoretyczna ocena już istniejących struktur i empiryczne poszukiwanie zmodyfikowanych cząstek. Nie należy również lekceważyć znaczenia – tak częstego w nauce – przypadku, choć przyznać trzeba, iż jego rola maleje.

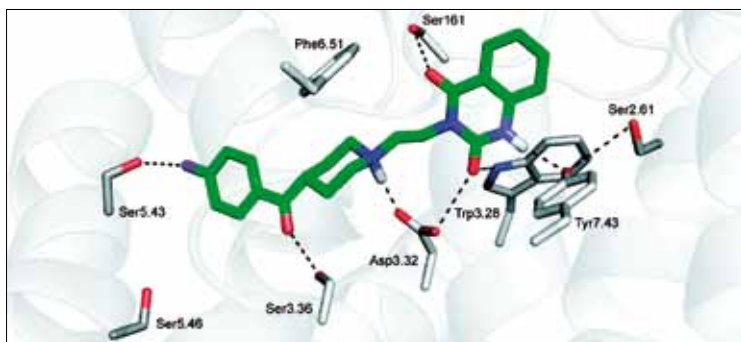
Wszystkie opisywane w artykule techniki zaliczane są do narzędzi modelowania molekularnego. Jego komponentami są między innymi grafika molekularna i zagadnienia związane z realistyczną wizualizacją badanych cząstek, narzędzia z zakresu chemii obliczeniowej, obszerne bazy danych, niekiedy

Tabela. Etapy badań klinicznych

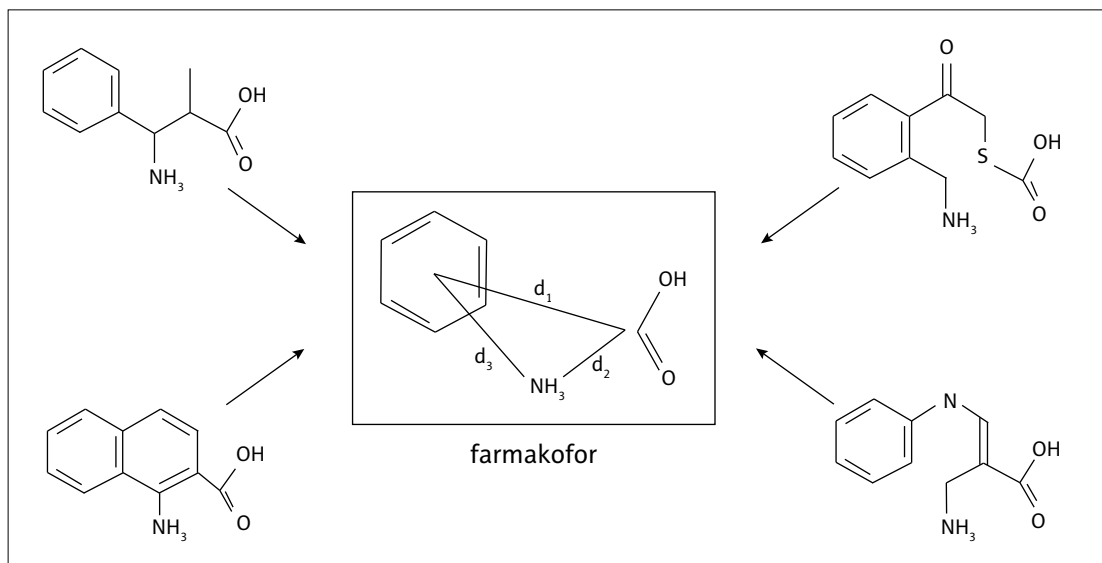
Badania przedkliniczne (<i>in vitro</i> i na zwierzętach)
faza I – na niewielkiej grupie (20–80) zdrowych ochotników; ma na celu zbadanie bezpieczeństwa, toksyczności, farmakokinetyki i farmakodynamiki terapii. W badaniach nad niektórymi lekami (np. przeciwnowotworowymi czy anti-HIV) w tej fazie badań biorą udział pacjenci w zaawansowanym stadium choroby
faza II – na większej grupie (20–300); ma na celu zbadanie klinicznej skuteczności terapii
faza III – randomizowane badania na dużej grupie (300–3000 lub więcej) pacjentów; ma na celu pełną ocenę skuteczności nowej terapii



Rycina 6. Jedna z zamodelowanych struktur receptora serotoninowego 5-HT2A wraz ze związaną z nim cząsteczką ketanseryny [12]



Rycina 7. Miejsce wiązania receptora serotoninowego 5-HT2A wraz ze związaną z nim cząsteczką ketanseryny [12]



Rycina 8. Przykład farmakoforu [9]

skomplikowanych cząstek chemicznych czy w końcu określanie właściwości fizyko-chemicznych na podstawie struktury (ang. QSAR – *Quantitative Structure-Activity Relationship*). Jak łatwo sobie wyobrazić, wszystkie wymienione metody wymagają dużych mocy obliczeniowych, a niewątpliwym postępem w tej dziedzinie jest związany bezpośrednio z rewolucją informatyczną.

Ze względu na zakres możliwych oddziaływań leku na organizm, często współwystępujących, klasyfikacje pozwalające na zdefiniowanie mechanizmów działania muszą siłą rzeczy mieć charakter ogólny. Jedną z istniejących możliwości jest określenie ich jako receptorowe oraz niereceptorowe. W pierwszym przypadku modelowanie na ekranie monitora można rozpocząć od odtworzenia struktury przestrzennej receptora (ang. *receptor based design*), co jest realizowane przy użyciu specjalistycznych programów, pozwalających nie tylko zobrazować, ale także ocenić prawdopodobieństwo stabilności nawet bardzo dużych cząstek białkowych, jaki-

mi są receptory. Każdy model jest mniej lub bardziej udanym oddaniem rzeczywistości, a więc z samej swojej definicji niesie ze sobą pewien błąd jej odtworzenia, toteż najlepszą możliwą sytuacją jest projektowanie na podstawie znanej struktury krystalograficznej białka receptorowego lub miejsca wiązania z lekiem. Większość ze wspomnianych programów dodatkowo oferuje możliwość analizy prawdopodobieństwa łączenia się receptora z badaną cząstką chemiczną, właśnie w miejscu receptorowym. Poznane sekwencje aminokwasów są gromadzone oraz udostępniane za pośrednictwem ogólnodostępnych i wyspecjalizowanych baz danych takich

jak ExpASY (*Expert Protein Analysis System*) – <http://www.expasy.org>, TMPRED – <http://www.ch.embnet.org>, czy też Protein Data Bank – <http://www.pdb.org>, niemniej jednak badacze zajmujący się tą problematyką po pobraniu struktury bazowej mogą ją dowolnie, wirtualnie modyfikować oddając najbardziej możliwą strukturę statyczną, ale także modelując dynamikę tego niekiedy niezwykle skomplikowanego związku. Dodatkową komplikacją i kolejnym krokiem może być umieszczenie stworzonego modelu receptora w dynamicznym środowisku błony komórkowej w przypadku receptorów błonowych. Na podstawie kształtu i właściwości miejsc wiązania receptora planowana jest struktura chemiczna potencjalnego leku oraz zestawiana wirtualnie z odtworzonym cyfrowo komputerowym modelem receptora. Aby było to możliwe konieczne jest określenie i dokładne zdefiniowanie miejsca wiążącego, a więc tego elementu struktury, który odpowiedzialny jest za łączenie się z lekiem. Wśród stosowanych metod najważniejsze są oparte na homologii czyli porównaniu ze znanymi strukturami miejsc wiążących.

Bardzo przydatną funkcją oferowaną przez producentów podobnego oprogramowania są zaawansowane opcje wizualizacji, co ułatwia ocenę otrzymanej struktury, w tym miejsca wiążącego.

Drugie możliwe podejście to wirtualna modyfikacja istniejących cząstek w poszukiwaniu pochodnych, cechujących się korzystniejszym profilem farmakokinetyczno-farmakodynamicznym. Dzięki znajomości struktury ligandów uproszczone zostaje poszukiwanie pochodnych oryginalnej cząsteczki lub struktur o podobnym mechanizmie działania, dokonywane tradycyjnie na drodze syntezy. Cząstki wykazujące aktywność w modelach komputerowych podlegają dalszym badaniom, często już po syntezie chemicznej.

Kolejny krok, a w zasadzie jego punkt końcowy, jest podobny dla obu dróg, choć sposób jego

Przemysł farmaceutyczny jest jedną z najintensywniej inwestujących w naukę i badania gałęzi gospodarki. Ocenia się, iż piątą część wśród firm przeznaczających najwięcej na rozwój stanowią firmy farmaceutyczne (ok. 15% zysków), które zdystansowały przemysł technologiczny oraz oprogramowania (ok. 10% zysków).

osiągnięcia inny. W zależności od wybranej drogi lub danych, którymi dysponuje się na początku procesu poszukujemy miejsca wiążącego oraz konstrukcji wiązań farmakoforowych (pierwsze podejście) lub farmakoforu, a więc wspólnej części badanych struktur, w znaczeniu relacji przestrzennych atomów budujących cząsteczkę.

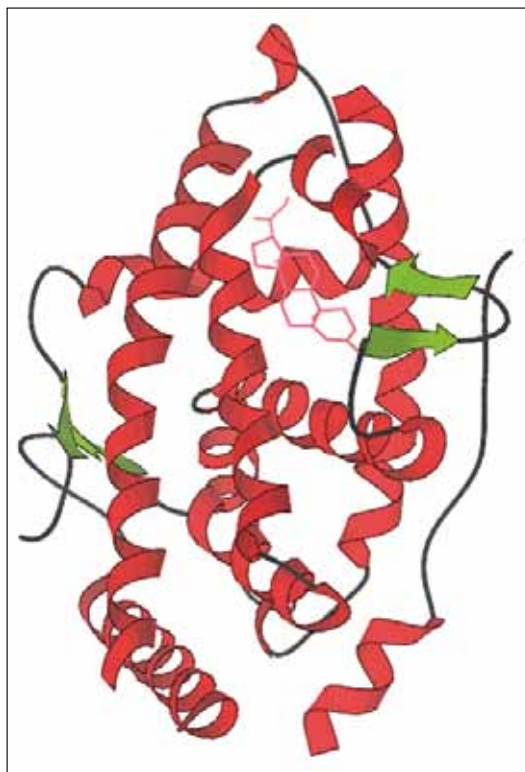
Wśród algorytmów wykorzystywanych w trakcie poszukiwania farmakoforów znajdują się oprócz stosunkowo prostych, zaawansowane matematycznie metody oparte m.in. na inteligencji obliczeniowej czy też algorytmy genetyczne, bazujące na odkryciach dotyczących biologicznego dziedziczenia cech.

Ocena właściwości fizykochemicznych

Jednym z obowiązkowych kroków w trakcie opracowywania nowej, potencjalnej struktury chemicznej jest określenie jej właściwości fizykochemicznych, które warunkują losy badanej cząstki już po podaniu w postaci leku. Podstawowe informacje obejmują masę molową, rozpuszczalność (w różnych warunkach – rozpuszczalnik, pH, dodatek substancji zwiększających rozpuszczalność np. kwasów żółciowych), współczynnik podziału n-oktanol/woda ($\log P$; $\log D$), stała dysocjacji (pK_a), PSA (ang. *polar surface area*), liczba wiązań będących donorami (ang. HBD – *Hydrogen Bond Donor*) czy też akceptorami atomów wodoru (ang. HBA – *Hydrogen Bond Acceptor*). Na ich podstawie można wyciągać wnioski zarówno na temat aktywności biologicznej, jak i potencjalnego mechanizmu wiązania z receptorem. Są one wyznaczane najczęściej na wczesnym etapie badań nad lekiem, niemniej jednak liczba syntetyzowanych lub badanych wirtualnie cząstek praktycznie uniemożliwia wykonywanie w warunkach laboratoryjnych dokładnych pomiarów dla każdej z nich, co jest powodem wykorzystywania metod opartych na technikach obliczeniowych, które pozwalają przewidzieć wszystkie wymienione powyżej właściwości oraz wiele innych. Wśród programów wykorzystywanych w tym celu można między innymi wymienić:

- ACDLabs – komercyjny program kanadyjskiej firmy ACD, posiadający swoją okrojoną darmową wersję zawierającą kalkulator umożliwiający określenie wartości $\log P$ na podstawie struktury,
- Marvin – zestaw doskonałych programów zawierający m.in. narzędzia umożliwiające określenie wartości $\log P$, pK_a i PSA; darmowy do zastosowań niekomercyjnych,
- MMPPro.

Te programy są to jedynie przykładami całej gamy podobnych narzędzi. Na szczególną uwagę zarówno ze względu na jakość, jak i sposób wykorzystania modeli oraz prezentacji danych, zasługuje system sieciowy zbudowany przez doktora Igora Tetko. Jego wirtualne laboratorium w całości dostępne w Internecie



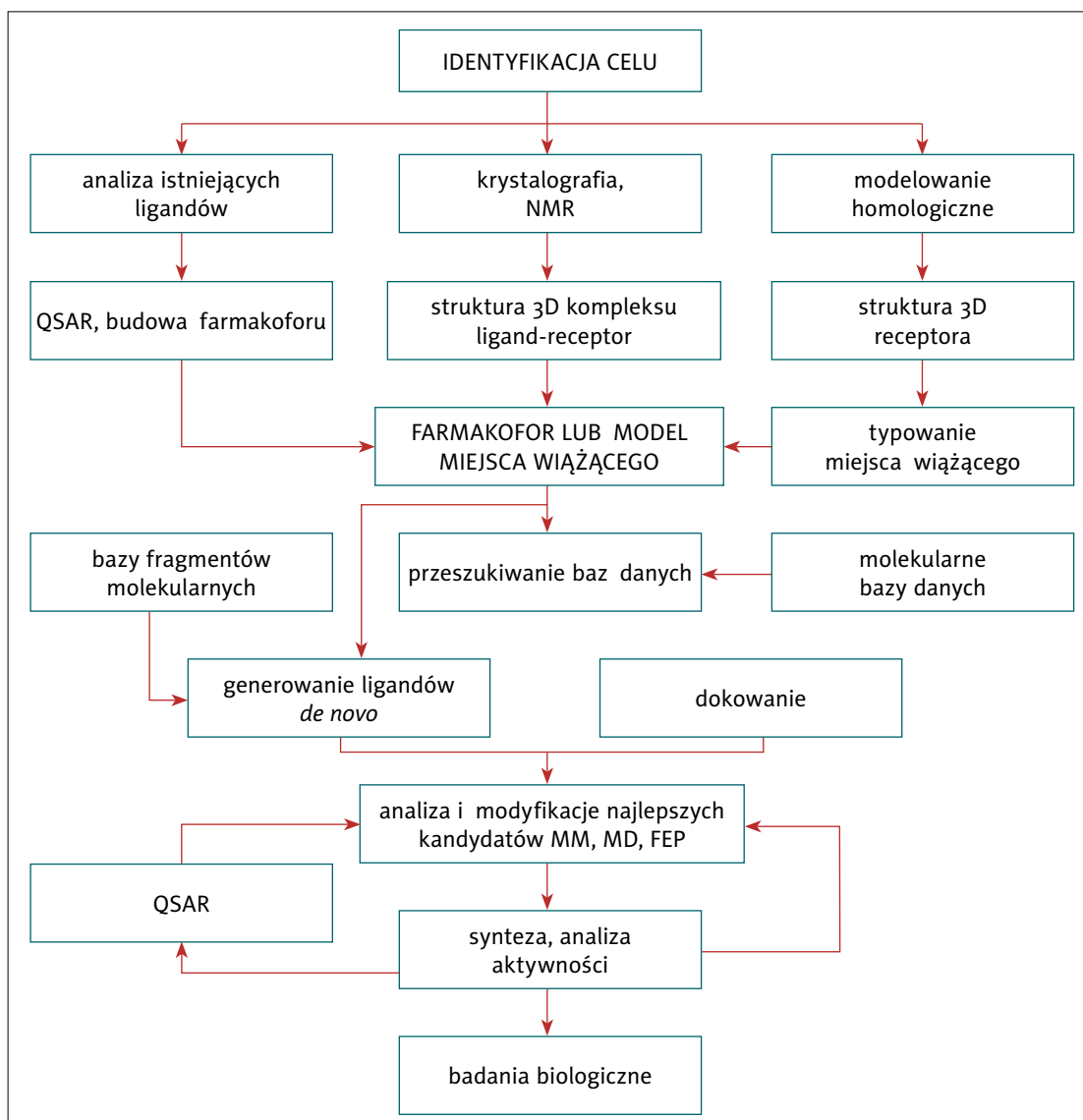
Rycina 9. Struktura krystaliczna domeny wiążącej ligand (deoksykortykosteron) ludzkiego receptora mineralokortykoidowego [9]

(<http://vclab.org/>), to bezpłatny zestaw oprogramowania, umożliwiający określenie właściwości fizykochemicznych jedynie na podstawie podanej struktury (istnieje możliwość rysowania wzoru substancji online w specjalnie przygotowanym edytorze). Warto wspomnieć, że wysoka jakość tego systemu, oceniana jako różnica między wartościami przewidzianymi i uzyskanymi w warunkach laboratoryjnych, została osiągnięta między innymi dzięki wykorzystaniu metod inteligencji obliczeniowej (w tym przypadku są to sztuczne sieci neuronowe), choć jednocześnie należy pamiętać, że oznaczenie laboratoryjne ma zawsze wyższą wartość niż ocena najlepszego nawet modelu.

Wszystkie podane wyżej przykłady deskryptorów fizykochemicznych nie oddają w pełni charakteru cząsteczki, ze względu na brak lub niepełne wykorzystanie informacji o strukturze przestrzennej. Opracowywane obecnie metody m.in. analiza właściwości pól molekularnych (ang. CoMFA – *Comparative Molecular Field Analysis*) umożliwiają kwantyfikację informacji opisujących strukturę trójwymiarową.

Określenie aktywności biologicznej

Dalszy etap, a więc poszukiwanie korelacji między strukturą chemiczną cząsteczek a ich właściwościami biologicznymi jest określane jako QSAR lub



Rycina 10. Schemat wykorzystania podstawowych metod komputerowych we wstępnych etapach projektowania leków [8]

QSPR. Przykładami programów komputerowych wykorzystywanych do realizacji opisanych zadań są: bezpłatne – BALLView, Ghemical i MMTK oraz komercyjne – Gaussian, Ceriusz, InsightII, Molsoft ICM, PyMOL, VMD, GROMOS, Sirius, NOCH, Sybyl, MOE, Agile Molekule, SPARTAN i Millsian.

Znając strukturę i charakterystykę zarówno miejsca wiązania, jak i badanej cząsteczki lub grup cząstek przeprowadza się ich wirtualne łączenie. Wśród ocenianych parametrów znajduje się prawdopodobieństwo ułożenia ligandu w miejscu wiążącym, jak i siła wiązania. W trakcie badań należy pamiętać o tym, że wiązanie obu molekuł jest zależne od ich dynamicznej struktury chemicznej ale także, że samo wiązanie spowoduje odkształcenie cząstek. Przeszukanie i ocena wszystkich możliwych układów jest – ze względu na ich liczbę – praktycznie niemożliwe, dlatego wykorzystuje się algorytmy półempiryczne. Jest to etap prac nad wirtualnym lekiem, który w dużej mierze

wymaga doświadczenia i intuicji badacza. Generowane olbrzymie ilości danych wymagają niezwykle starannej analizy, ponieważ istnieje prawdopodobieństwo przeoczenia najlepszej możliwej konfiguracji.

Rycina 10 podsumowuje schemat postępowania w trakcie wirtualnego poszukiwania nowych, obiecujących cząstek chemicznych o potencjalnej aktywności farmakodynamicznej. Na uwagę zasługuje fakt konieczności gromadzenia rozległych zbiorów informacji chemicznej oraz związanych z tym efektywnych narzędzi przeszukiwania baz danych i wnioskowanie na podstawie uzyskanych rezultatów. Są to niekiedy bazy danych zawierające miliony rekordów, z których każdy jest opisany wieloma tysiącami parametrów. System przeszukiwania takiego zbioru informacji musi być nie tylko efektywny, ale i szybki, oferując możliwość uzyskania rezultatów w czasie rzeczywistym. Całość technik matematycznych, realizowanych w praktyce w warunkach *in silico*, a wykorzystywanych w tym

właśnie celu określa się angielskojęzycznym mianem data mining, co w wolnym przekładzie może znaczyć zgłębianie danych.

Ze względu na bardzo duże zapotrzebowanie na komputerową moc obliczeniową wykorzystywaną przy projektach wirtualnego poszukiwania nowych struktur chemicznych, potencjalnych leków wykorzystuje się zaawansowane techniki rozpraszania obliczeń na dużą ilość stosunkowo słabych komputerów (tzw. gridy obliczeniowe). Jednym z przykładów ciekawego rozwiązania jest wykorzystywanie wolnej mocy domowych komputerów. Po zainstalowaniu niewielkiego programu, który analizuje wykorzystanie naszego komputera osobistego i włącza własne, drobne fragmenty większych zadań obliczeniowych, po czym wysyła wyniki do komputera centralnego, możemy pomóc projektować nowe leki przeciwnowotworowe (<http://boinc.bakerlab.org/rosetta/> lub <http://www.boincatpoland.org/>).

Choć mogłoby się wydawać, że lek z komputera to wciąż przyszłość jednak istnieją już dostępne na rynku leki, podczas projektowania których wykorzystano techniki modelowania molekularnego. Najbardziej spektakularnym przykładem są blokery receptorów dla angiotensyny – sartany.

Ocena właściwości biologicznych

Uzbrojeni w zdobyte do tej pory dane, możemy przejść do kolejnego etapu badań nad naszym lekiem – oceny właściwości biologicznych. Najczęstszym do tychczas stosowanym modelem były oczywiście – wciąż wykorzystywane – modele zwierzęce (myszy, szczury, świnki morskie, króliki, psy, małpy). Niemniej jednak coraz wyraźniej widać tendencję do minimalizowania wykorzystania zwierząt laboratoryjnych. Jest to podyktowane zarówno względami humanitarnymi, jak i praktycznymi – skalowanie allometryczne, a więc przenoszenie obserwacji ze zwierząt na ludzi jest zawsze obciążone błędem. Wynika to z innej fizjologii nawet najbardziej zbliżonych genetycznie do człowieka gatunków zwierząt (np. szczury nie mają woreczka żółciowego – wydzielanie żółci ma charakter ciągły, a nie wzbudzany pokarmem, jak u człowieka). Oprócz tego postuluje się wykorzystywanie danych uzyskiwanych w trakcie doświadczeń na ludziach, izolowanych komórkach lub organellach komórkowych (np. HLM – *Human Liver Microsomes*). Zastosowanie takiej metodyki jest kosztowne i obciążone wymogiem spełnienia bardzo wyśrubowanych wymogów związanych z projektowaniem badań, choć dostarcza niezwykle dokładnych i reprezentatywnych danych.

Rozwiązaniem jest więc wykorzystanie modeli realizowanych w warunkach *in silico*. Na tym etapie badań kilkadziesiąt-kilkaset z tysięcy wcześniej badanych związków jest monitorowanych pod kątem właściwości biologicznych. Oznacza to w praktyce ocenę ADME/Tox, a więc przewidywanie zachowania leku w organizmie

(tox) oraz wpływu organizmu na lek (wchłanianie, dystrybucja, metabolizm, wydalanie – ADME). Należy zaznaczyć, że poszczególne etapy są traktowane bardzo szeroko i tak wchłanianie obejmuje nie tylko podanie doustne, ale i na skórę, wziewne czy doodbytnicze. Dystrybucja to także modelowanie przenikania przez barierę krew-mózg, a metabolizm obejmuje zarówno fazę I (najczęściej przez cytochrom P450 w jelitach i wątrobie), ale także fazę II (np. glukuronizacja za pośrednictwem enzymów z rodziny UGT). Liczba i różnorodność wykorzystywanych w licznych dostępnych programach komputerowych metod jest ogromna. Obejmuje zarówno metody analizy strukturalnej (fragmenty cząstki odpowiadające za jej charakter) jak i algorytmy matematyczne, metody statystyczne czy też oparte na sztucznej inteligencji. Przykładem podobnego oprogramowania są systemy ADME Boxes, ADMEnsa Interactive, q-Tox/q-ADME i wiele innych. Na szczególną uwagę i krótkie omówienie zasługują metody przewidywania metabolizmu cząsteczek z udziałem cytochromów z rodziny P450. Programy takie jak MetaSite czy MetaDrug oferują właśnie taką funkcjonalność, a dzięki zastosowanym algorytmom przewidują one nie tylko jakościowe (który enzym), ale i ilościowe (jakie powinowactwo) parametry metabolizmu danej substancji.

Do kolejnego etapu naszego wirtualnego doświadczenia przechodzimy z kilkoma najbardziej obiecującymi cząsteczkami, ponieważ pozostałe udało nam się odrzucić na tym etapie – nie spełniały wymogów biologicznych. Jedna z nich – mamy nadzieję – zostanie pełnoprawnym lekiem, co pozwoli nam na odzyskanie zainwestowanych do tej pory kwot.

Słownik używanych pojęć

Słownik stanowi integralną część całego cyklu i zawiera pojęcia, które w opinii autorów mogą okazać się przydatne w trakcie studiowania treści artykułów.

In silico – czyli dosłownie *w krzemie*, a więc w domyśle – wewnątrz krzemowej płytki procesora komputera; pojęcie to – będące modnym słowem kluczem – oznacza w zasadzie wszystkie możliwe sytuacje nadsładowania rzeczywistości za pomocą modeli obliczeniowych realizowanych komputerowo (i oczywiście niekoniecznie dotyczy to jedynie farmacji!); sama nazwa jest nieco myląca, jako że – co podkreślają umiarkowani zwolennicy podobnych metod – zdecydowaną większość podobnych kalkulacji można przeprowadzić na papierze, jedyną różnicą jest czas przeznaczony na realizację tychże zadań.

Kolejnym argumentem przemawiającym za szeroką i wciąż rosnącą gamą modeli realizowanych w warunkach *in silico*, są względy humanitarne wskazujące na konieczność uzyskania maksymalnej możliwej redukcji doświadczeń przeprowadzanych na zwierzętach oraz optymalizacja koniecznych badań prowadzonych na ludziach (zdrowych ochotnikach oraz osobach chorych, w warunkach klinicznych).

In vitro – bezpośrednie tłumaczenie łacińskiego wyrażenia *w szkle* może być nieco mylące dla szeroko pojmowanych badań farmaceutycznych, ponieważ postęp w preparatyce laboratoryjnej powoduje, że niekoniecznie musi to być szkło; w praktyce pojęcie to oznacza odtwarzanie (naśladowanie) w warunkach laboratoryjnych reakcji biologicznych zachodzących w żywych organizmach, a co warte podkreślenia obejmuje również doświadczenia z żywymi komórkami, wyizolowanymi z organizmu macierzystego (zwierząt i ludzi) i umieszczonymi w warunkach laboratoryjnych, umożliwiającym podtrzymywanie ich życia; przykładem podobnych komórek mogą być ludzkie mikrosomy (HLM); stało się to na tyle powszechne, że określenie badania *in vitro* oznaczają prowadzenie badań na żywych, wyizolowanych z organizmu komórkach, organellach lub substancjach; rodzi to jednak kolejny problem, a mianowicie „tłumaczenie” (korelowanie) wyników badań *in vitro* z warunkami *in vivo*.

In vivo – a więc wewnątrz funkcjonującego organizmu, obejmującego wszystkie tkanki i narządy składające się na działającą w określonych warunkach środowiska jednostkę; pojęcie to obejmuje zarówno wyniki badań prowadzonych na ludziach (ochotnicy) jak i na modelach zwierzęcych (w naukach farmaceutycznych m.in. myszy, szczury, świnki morskie, psy, mały); wyniki uzyskiwane w warunkach

in vivo stanowią najcenniejsze źródło informacji na temat procesów biologicznych, choć oczywiście możliwość ich uzyskania jest ograniczona ze względu na czynniki etyczne, finansowe i praktyczne (np. realizacja badania przenikalności bariery krew-mózg w warunkach *in vivo*), stąd niekiedy zastępowane przez badania określane mianem *ex vivo*.

Ex vivo – a więc przeprowadzane poza organizmem, ale na materiale pobranym z żywego, funkcjonującego organu; przykładem podobnych badań jest ocena składu żółci (i jej wpływu na proces rozpuszczania leków w jelitach), pobranej przy użyciu endoskopu.

IVIVC – (ang. *in vitro – in vivo correlation*) – niezwykle szerokie pojęcie oznaczające całość technik (matematycznych, statystycznych, obliczeniowych) wykorzystywanych w trakcie przenoszenia i wykorzystywania wyników badań *in vitro* na warunki *in vivo*.

HLM (ang. *human liver microsomes*) – ludzkie komórki wątroby (mikrosomy) to jeden z najczęściej

stosowanych modeli *in vitro* umożliwiających odtwarzanie i badanie funkcji metabolicznej wątroby w warunkach laboratoryjnych; obecnie wyniki badań z zastosowaniem HLM ze względu na oczywiste różnice między wyizolowanymi komórkami oraz całym narządem funkcjonującym w warunkach fizjologicznych, są wykorzystywane jako dane startowe do dalszego symulowania aktywności wątroby już w warunkach *in silico*.

QSAR (ang. *Quantitative structure-activity relationship*) – pojęcie obejmuje zakres technik stosowanych w trakcie poszukiwania korelacji między strukturą chemiczną a zdefiniowanym procesem zachodzącym z jej udziałem (np. proces biologiczny lub reakcja chemiczna); jedną z odmian jest 3D-QSAR, a więc ocena opisanej powyżej korelacji na podstawie właściwości cząstki obliczonych z wykorzystaniem jej struktury przestrzennej (trójwymiarowej); liczba stosowanych technik obliczeniowych jest ogromna, warto również pamiętać, że bardzo często wymagają one również specjalistycznego oprogramowania z zaimplementowanymi algorytmami oraz potężnych mocy obliczeniowych, daleko wykraczających poza te oferowane przez klasyczne komputery biurkowe.

QSPR (ang. *quantitative structure-property relationships*) – techniki stosowane przy poszukiwaniu korelacji między strukturą i właściwościami cząstek.

ADME lub LADME lub LADMET lub ADME/Tox lub LADME/Tox (ang. [*Liberation*], *Absorption, Distribution, Metabolism, Excretion, [Toxicology]*) – akronim obejmujący podstawowe etapy interakcji lek-organizm począwszy od uwalniania, poprzez wchłanianie, dystrybucję, metabolizm, wydalanie, aż do działania farmakologicznego/toksycznego na organizm; nazwa ta stała się tak popularna, że określenie „modelowanie ADME” oznacza bezpośrednio określenie dowolnych modeli wykorzystywanych w celu opisanie powyższych zjawisk.

Sztuczne sieci neuronowe – narzędzia analizy danych wykorzystujące podstawowe mechanizmy adaptacyjne znane z neurobiologii do samouczenia na dostępnych danych; z reguły mają nieliniowy charakter i pozwalają na identyfikację skomplikowanych i wielowymiarowych zależności w danych.

Modelowanie mechanistyczne – określenie obejmujące techniki modelowania matematycznego, dla którego punktem wyjścia jest zestaw założeń dotyczących charakteru tworzonego modelu.

ExpPASy (ang. *Expert Protein Analysis System*) – system magazynowania oraz udostępniania struktur białkowych możliwych do pobrania i dalszej analizy, rozprowadzanych w ogólnie przyjętym formacie komputerowym (pliki pdb); dodatkowo oferuje możliwość wizualizacji pobranej cząsteczki (<http://www.expasy.org>- <http://swissmodel.expasy.org/repository/>).

Ze względu na zakres możliwych oddziaływań leku na organizm, często współwystępujących, klasyfikacje pozwalające na zdefiniowanie mechanizmów działania muszą siłą rzeczy mieć charakter ogólny.

Jedną z istniejących możliwości jest określenie ich jako receptorowe oraz niereceptorowe.

W pierwszym przypadku modelowanie na ekranie monitora można rozpocząć od odtworzenia struktury przestrzennej receptora, co jest realizowane przy użyciu specjalistycznych programów, pozwalających nie tylko zobrazować, ale także ocenić prawdopodobieństwo stabilności nawet bardzo dużych cząstek białkowych, jakimi są receptory.

Grid computing – określenie obejmujące zarówno sprzętową, jak i logiczną strukturę informatyczną, umożliwiającą zarządzanie obliczeniami rozproszonymi – zespół komputerów zlokalizowanych w różnych miejscach pracujących nad jednym problemem obliczeniowym dzięki zaawansowanym technologicznie systemom rozpraszania obliczeń; struktury gridowe są niekiedy określane – nie do końca prawidłowo – jako wirtualne superkomputery

Skalowanie allometryczne – całość technik matematycznych pozwalających na przeniesienie rezultatów badań z jednego medium na inne z wykorzystaniem proporcjonalności w masie narządów i/lub budowy ciała; technika ta jest też stosowana do skalowania wyników badań uzyskanych z wykorzystaniem modeli zwierzęcych na analogiczne dane możliwe do zaobserwowania u ludzi.

NCE (ang. *New Chemical Entity*) – definicja podawana za firmą IMSHealth – pierwsze na świecie wdrożenie nowej substancji aktywnej, włączając w to produkty biotechnologiczne (np. rekombinowane proteiny, szczepionki); z założenia definicja ta wyklucza radioterapeutyki, testy diagnostyczne, wyciągi roślinne, szczepionki złożone, przeciwciała poliklonalne, w odróżnieniu od stosowanego przez FDA pojęcia NME (ang. *New Molecular Entity*), które oznacza każdy nowy produkt leczniczy wdrożony na rynek; proszę zwrócić uwagę, że definicje te nie dotyczą leków generycznych.

R&D (ang. *Research and Development*, dosłownie – badania i rozwój) – ogólne określenie obejmujące całość technik, metod, narzędzi i związanych z tym kosztów, systemów organizacji pracy i analizy wyników zaangażowany w proces poszukiwania i badania cząstek o potencjalnej aktywności farmakologicznej; ogólnie – każdy proces badawczy, bez względu na typ zaangażowanego przemysłu.

Microdosing – technika oznaczania profilu farmakokinetycznego leku w warunkach *in vivo* na ludziach, z zastosowaniem minimalnych dawek substancji aktywnej, nie wywołujących żadnego efektu farmakodynamicznego, a więc także niepożądanego; w praktyce stosuje się jedną setną część najmniejszej dawki wywołującej efekt farmakodynamiczny lub 100 µg, w zależności od tego, która z nich jest mniejsza; ze względu na bardzo niskie stężenia oznaczanych substancji oraz ich metabolitów konieczne jest stosowanie niezwykle czułych metod analitycznych jak chromatografia cieczowa z podwójnym detektorem masowym – LC-MS-MS, spektrometria mas – AMS (ang. *Accelerated Mass Spectrometry*) czy po-

zytronowa tomografia emisyjna – PET (ang. *Positron Emission Tomography*).

3Rs – ang. *Replace, Reduce, Refine*, co mogłoby być przetłumaczone na język polski jako 3Z – Zastąpić, Zredukować, Zdefiniować; określenie definiujące współczesne podejście do problemu wykorzystywania zwierząt do badań naukowych, którego celem jest właśnie redukcja ich liczby m.in. poprzez udoskonalenie i optymalizację procedur laboratoryjnych, jak i zastąpienie doświadczeń *in vivo* na zwierzętach innymi technikami, pozwalającymi osiągnąć podobne rezultaty.

Farmakofor – model opisujący relacje przestrzenne między elementami wspólnymi dla ligandów wiążących się lub inaczej oddziałujących na badany receptor.

Data mining (eksploracja danych, zgłębianie danych) – zbiorcze określenie technik odkrywania wiedzy w bazach i hurtowniach danych; koncepcja eksploracji danych opiera się na wykorzystaniu statystycznych, empirycznych i półempirycznych metod wyszukiwania powiązań, ukrytych dla człowieka między innymi ze względu na ilość informacji.

Jednym z obowiązkowych kroków w trakcie opracowywania nowej, potencjalnej struktury chemicznej jest określenie jej właściwości fizykochemicznych, które warunkują losy badanej cząstki już po podaniu w postaci leku. Podstawowe informacje obejmują masę molową, rozpuszczalność, współczynnik podziału n-oktanol/woda, stała dysocjacji, PSA (ang. polar surface area), liczba wiązań będących donorami czy też akceptorami atomów wodoru.

Przyjęto: 2009.01.18 · Zaakceptowano: 2009.01.25

Piśmiennictwo

- Grabowski H.G., Wang Y.R.: The Quantity And Quality Of Worldwide New Drug Introductions, 1982–2003. *Health Affairs* 2006, 25 (2), 452–460.
- Grabowski H.G.: Are the Economics of Pharmaceutical R&D Changing? Productivity, Patents and Political Pressures. *PharmacoEconomics* 2004, 22, Supp. 2, 15–24.
- DiMasi J.A., Hansen R.W., Grabowski H.G.: The price of innovation: new estimates of drug development costs. *Journal of Health Economics* 2003, 22, 151–185.
- DiMasi J.A., Grabowski H.G.: The Cost of Biopharmaceutical R&D: Is Biotech Different? *Manage. Decis. Econ.* 2007, 28, 469–479.
- <http://www.nc3rs.org.uk/>
- OECD Health Policy Studies. Pharmaceutical Pricing Policies in a Global Market. OECD 2008-12-27
- The 2007 R&D Scoreboard. Department for Innovation, Universities & Skills UK.
- Piotr Setny, Projektowanie leków, artykuł dostępny pod adresem http://www.icm.edu.pl/kdm/Projektowanie_lekow
- <http://www.pdb.org>
- <http://boinc.bakerlab.org/rosetta/>
- <http://www.boincatpoland.org/>
- Cristina Dezi, Jose Brea, Mario Alvarado i wsp.: Multistructure 3D-QSAR Studies on a Series of Conformationally Constrained Butyrophenones Docked into a New Homology Model of the 5-HT_{2A} Receptor. *J. Med. Chem.* 2007, 50, 3242–3255.