

## Recenzja

pracy doktorskiej pani mgr inż. Roksany Duszkiewicz pt. „*Ocena liganda jako potencjalnego leku na próbie wybranych baz wielkich danych*”, wykonanej w ramach studiów doktoranckich Instytutu Chemii Uniwersytetu Śląskiego w Katowicach w Zakładzie Chemii Organicznej, pod kierunkiem prof. dr hab. inż. Jarosława Polańskiego.

Poszukiwanie nowych leków to główna domena działań przemysłu farmaceutycznego wzbogacona także przez ośrodki akademickie czy instytuty naukowe. O ile celem badań firm farmaceutycznych jest uzyskanie leku jako produktu handlowego to celem badań pozostałych instytucji jest odkrywanie nowych cząsteczek bioaktywnych, ligandów jako potencjalnych leków, które mogą być przedmiotem dalszych prac, najczęściej we współpracy z firmami. Aktualnie droga do wprowadzenia nowego leku na rynek to wielokierunkowe, wieloletnie i kosztowne badania. Proces ten jednak stale jest analizowany pod kątem zmniejszenia ryzyka i kosztów prac, które są najwyższe na etapie badań klinicznych.

Przedstawiona do oceny praca doktorska pt. „*Ocena liganda jako potencjalnego leku na próbie wybranych baz wielkich danych*” wpisuje się w ten nurt badań, których wyniki mogą być przydatne do oceny cząsteczki związku chemicznego na etapie badań przedklinicznych. Praca jest kontynuacją ale także ciekawym rozwinięciem wcześniejszych zainteresowań profesora Polańskiego dotyczących wykorzystania dużych baz danych do analizy wybranych parametrów istotnych w procesie poszukiwania nowego leku.

Dysertacja napisana jest w oparciu o 5 opublikowanych prac, ale nie są to kolejne opisy poszczególnych publikacji tylko raczej łączna dyskusja uzyskanych wyników i analizowanych zagadnień. Celem pracy, który doktorantka sformułowała w pierwszej części, było zbadanie problemu oceny liganda jako potencjalnego leku poprzez analizę estymatora (parametru) wydajności liganda (*ligand efficiency*; LE) zawierającego w sobie komponentę aktywności biologicznej cząsteczki związku chemicznego. Do analiz wykorzystano dwie duże molekularne bazy danych tj. ChEMBL i PubChem oraz własną, mniejszą bazę danych literaturowych utworzoną z baz: *Binding Database*, *Psychoactive Drug Screening Program* i wybranych danych z publikacji. Dane te zostały dołączone do pracy doktorskiej w formie dysku pamięci.

Teoretyczny wstęp pracy opisuje krótko zagadnienia związane z badaniami doktorantki tj. proces projektowania leku, parametry wyrażające aktywność biologiczną cząsteczki związku chemicznego, deskryptory molekularne, właściwości cząsteczkowe, lekopodobieństwo, wydajność liganda, zagadnienia opisu aktywności biologicznej poprzez różne parametry, zagadnienia ekonomiczne oraz korzyści z zastosowań analiz wielkich baz

danych. Najwięcej uwagi poświęcone jest wskaźnikowi efektywności cząsteczki bioaktywnej czyli wydajności liganda, jego zdefiniowanie, opis zastosowań w projektowaniu leków oraz kontrowersji związanych z używaniem tej miary. Wydajność wiązania liganda szacuje wolną energię wiązania przez element strukturalny, który został zaproponowany w formie prostej funkcji jako iloraz (stosunek) miary aktywności (zazwyczaj stosuje się log z aktywności  $pIC_{50}$  lub  $pK_i$ ) i liczby ciężkich atomów w cząsteczce. Prostota wyliczenia wartości wydajności liganda dla cząsteczki bioaktywnej spowodowała zainteresowanie tym parametrem i stosowanie go w projektach poszukiwania leków w obrębie określonych serii związków, analizie zależności pomiędzy strukturą i aktywnością.

W przedstawionych badaniach podjęta dyskusja nad fizycznym znaczeniem czy matematyczną poprawnością tego estymatora (parametru) pozwoliła na zaprezentowanie wyników analiz opartych na dużej puli danych oraz ograniczeń bezkrytycznego stosowania go do porównywania działania cząsteczek bioaktywnych. W swojej pracy doktorantka opisała zaproponowane nowe estymatory będące modyfikacjami wydajności liganda tj. – *product ligand efficiency* (PLE) oraz SCORE. Nie przedstawiono wersji polskiej nowego proponowanego estymatora *product ligand efficiency*, wypada zatem zadać pytanie jaka zdaniem doktorantki jest poprawna forma tego tłumaczenia - produktowa wydajność liganda czy też produkt wydajności liganda. PLE zdefiniowano jako iloczyn siła działania wyrażanej w formie logarytmicznej i liczby atomów ciężkich, natomiast SCORE jako funkcję sumy tych dwóch wartości w formie logarytmicznej wraz z odpowiednimi współczynnikami. Następnie wykonano szereg analiz porównujących zależności pomiędzy aktywnością biologiczną wyrażaną przez wartość  $pAC_{50}$ , wydajnością liganda (LE), wskaźnikiem wydajności wiązania (*binding efficiency index*; BEI) a liczbą ciężkich atomów (HAC) lub masą cząsteczkową. Analizy te wykonywane były na różnych, dużych zbiorach danych i przedstawiane w dwóch formach także uwzględniających kategoryzację zmiennych tj. proces binowania. Ponadto przeprowadzono dyskusję dotyczącą analizy właściwości związanych ze skalą molową, wagową lub skalą pojedynczej cząsteczki. Uzyskane wyniki wykazały zależność LE względem liczby ciężkich atomów w formie funkcji hiperbolicznej.

Kolejne prace obejmowały analizę zmian wartości LE dla projektów opartych o optymalizację fragmentów określanych jako F2L (*fragment to lead*), w których LE może się zmniejszać lub zwiększać, zatem delta-LE może być ujemna lub dodatnia. Określono to jako efekt F2L a wyniki analiz wykazały, że LE jest estymatorem, którego wartość w dużym stopniu jest zdominowana przez wartość HAC.

Dalsze analizy statystyczne uwzględniały także badanie zależności pomiędzy estymatorem PLE, funkcją SCORE dla cząsteczek leków oraz dla fragmentów. Ciekawym pomysłem było włączenie do analiz porównawczych zagadnień z obszaru ekonomii, gdzie badano zależności pomiędzy ceną związku chemicznego a jego masą cząsteczkową. Zaobserwowano podobne trendy zmienności przy analizie cen związków chemicznych w skali molowej i wagowej do tych obserwowanych dla aktywności biologicznej.

Następnie analizowano wydajność liganda względem temperatury wrzenia cząsteczki (*boiling point*; BP), która jest związana ze średnią energią cząsteczek układu i nie odnosi się ani do skali molowej czy wagowej. Dla porównania analizowano także zależności pomiędzy temperaturą wrzenia a masą cząsteczkową oraz liczbą ciężkich atomów. Zaobserwowano podobieństwo zależności dla analiz stosujących BP czy LE co wskazuje na powiązanie wielkości BP ze skalą molową, chociaż łączy się ona z modelem pojedynczej cząsteczki.

W podsumowaniu pracy doktorantka zebrała wnioski, które wysunięto w wyniku wykonania wielu powyżej krótko opisanych analiz. Hiperboliczny obraz zależności aktywności biologicznej wyrażonej w skali  $pAC_{50}$  czy też LE w funkcji  $1/HAC$  wskazuje, że wartość LE jest w dużym stopniu determinowana przez czynnik  $1/HAC$ . Przy czym zauważono, że czynnik  $1/HAC$  lub  $(1/MW)$  jest również operatorem zamieniającym wielkości wyrażone w skali molowej na skalę wagową. Konsekwencją tego dla wartości wydajności liganda jest preferencja ligandów o niższej masie cząsteczkowej i niższych wartościach ciężkich atomów (HAC). Dane te wskazują, że funkcja LE nie jest kontrolowana przez statystykę Avogadry charakterystyczną dla efektów w skali molekularnej.

Zagadnienia podjęte w pracy doktorskiej pani mgr Roksany Duszkiewicz są bardzo ciekawe i wpisują się w nurt poszukiwania zależności i prawidłowości pomiędzy aktywnością cząsteczki bioaktywnej a jej strukturą, z której dalej wynikają inne właściwości, które można wyrazić w formie różnych parametrów czy deskryptorów. Wyniki takich analiz mogą być przydatne w procesie projektowania i poszukiwania nowych cząsteczek bioaktywnych i dalek leków. Najwięcej uwagi doktorantka poświęciła w swojej pracy wydajności liganda, przedstawiając kontrowersje wokół tej wartości na podstawie analizy wielu czynników i dużych baz danych. Wydaje się jednak, że opisane krytyczne spojrzenie na tą wartość nie dyskredytuje jej ale wskazywać może na pewne ograniczenia w stosowaniu. Należy przykładowo wziąć pod uwagę, że zaproponowana prawie ćwierć wieku temu Reguła 5 Lipinskiego wyklucza związki pochodzenia naturalnego oraz cząsteczki, które są substratami dla transporterów biologicznych. Ponadto wykazano, że 16% stosowanych leków doustnych narusza przynajmniej 1 kryterium a 6% zawodzi w 2 lub więcej tej Reguły. Nie dyskredytuje to jej jednak i standardowo stosuje się jako filtr dla optymalizacji właściwości lekopodobnych w procesie projektowania leku.

Zaproponowane nowe estymatory uwzględniające wydajność liganda są także interesujące, jednak stosowanie ich i rozpowszechnienie w środowisku naukowym nie wydaje się proste. Należy zauważyć, o czym zresztą pisze doktorantka w pracy, że wartość wydajności liganda wymaga znajomości aktywności biologicznej badanej cząsteczki. W projektach poszukiwania leku zastosowanie LE może być zatem przydatne w analizie zależności struktura – aktywność i dalszej modyfikacji strukturalnej cząsteczki wiodącej. W takich projektach zazwyczaj ilość zbadanych cząsteczek jest ograniczona a ich wielkość wyrażona masą cząsteczkową i dalej ilością ciężkich atomów nie różni się znacząco co wynika z klasycznej drogi optymalizacji farmakoforu i konieczności odpowiedniego dopasowania do


celu molekularnego. Należy ponadto zaznaczyć, że parametr wydajności liganda odnosi się do aktywności biologicznej wykonanej w warunkach *in vitro*, to jest do wczesnego etapu badań przedklinicznych, w długiej drodze do nowego leku.

Podsumowując, praca doktorska pani mgr Roksany Duszkiewicz jest wartościowa, napisana poprawnie w oparciu o opublikowane już prace, w czasopismach o zasięgu międzynarodowym. Są to prace zazwyczaj z udziałem kilku autorów, w których pierwszym i korespondencyjnym autorem jest promotor pracy, w 2 pracach doktorantka jest drugim a w 3 trzecim autorem. Ważna dla doktorantki praca opublikowana w 2020 w *Chemometrics and Intelligent Laboratory Systems* (IF = 2,895) z 2020 jest opublikowana tylko z profesorem Polańskim. Dołączone oświadczenia współautorów prac nie budzą wątpliwości co do najbardziej znaczącego udziału doktorantki w wykonanych badaniach.

Przedstawiona do oceny praca odpowiada w pełni wymogom stawianym pracom naukowym na stopień doktora, dlatego stawiam wniosek Radzie Naukowej Instytutu Chemii Uniwersytetu Śląskiego w Katowicach o dopuszczenie pani mgr Roksany Duszkiewicz do dalszych etapów przewodu doktorskiego.

Biorąc pod uwagę ciekawą, oryginalną tematykę wykonanych analiz z wykorzystaniem dużych baz danych, zaproponowane nowe estymatory do opisu wydajności liganda, które mogą być wykorzystane w ocenie właściwości cząsteczek bioaktywnych oraz opracowanie wyników w formie już opublikowanych 5 prac skład także wniosek o wyróżnienie tej pracy doktorskiej.

Prof. dr. hab. Barbara Malawska

Katedra Chemii Farmaceutycznej UJ CM  
Zakład Fizykochemicznej Analizy Leku  
  
prof. dr hab. Barbara Malawska

Zakład Fizykochemicznej Analizy Leku  
Katedry Chemii Farmaceutycznej  
Wydział Farmaceutyczny  
Uniwersytet Jagielloński  
Collegium Medicum