



prof. dr hab. n. farm. Łukasz Komsta
Uniwersytet Medyczny w Lublinie
Wydział Farmaceutyczny
Katedra i Zakład Chemii Leków
ul. Jaczewskiego 4, 20-090 Lublin, tel. 81 4487387, fax 81 4487381

Recenzja pracy doktorskiej mgr Agnieszki Kamedulskiej

*„Opracowanie bayesowskich modeli hierarchicznych
opisujących retencję w wysokosprawnej chromatografii ciekowej
w odwróconym układzie faz”*

PRZEWIDYWANIE retencji chromatograficznej związku na podstawie jego budowy chemicznej nie jest zadaniem łatwym. Interakcja cząsteczki z fazą stacjonarną oraz ruchomą jest procesem bardzo złożonym i zaangażowane są w nią różne rodzaje oddziaływań. Skuteczny model umożliwiający przewidywanie retencji wnosi istotny wkład zarówno w zrozumienie procesów zachodzących wewnątrz kolumny, jak również ma wyraźne zastosowanie praktyczne, skraca bowiem czas optymalizacji metod analitycznych.

Mimo upływu ponad stu lat od wynalezienia chromatografii, nikomu nie udało się stworzyć jednego modelu przewidującego retencję dla każdej cząsteczki, a biorąc pod uwagę kompleksowość procesów, wydaje się to mało prawdopodobne do zrealizowania w najbliższym czasie. Niemniej jednak można już konstruować modele przewidujące retencję w określonym układzie dla pewnych dość dużych obszarów różnorodności molekularnej.

Przewidywanie retencji jest jednym z przypadków przewidywania własności cząstek na podstawie budowy chemicznej QSPR (*Quantitative Structure-Property Relationships*). Na przestrzeni dziejów przewidywanie właściwości z deskryptorów molekularnych wzbudzało skrajnie różne emocje — od przekonania, że wkrótce można będzie modelować wszystko na podstawie wszystkiego, do całkowitego zanegowania przydatności takich metod w praktyce.

Overoptimism, jak określano w literaturze nadmierną fascynację nie do końca zwalidowanymi modelami, publikowanie modeli czysto przypadkowych, czy też wreszcie stosowanie do tworzenia modeli podejść szczególnie podatnych na przeuczenie (*overfitting*) przyniosło z kolei tej metodyce

szerokie grono przeciwników.

Jednakże modele dobrze skonstruowane, właściwie zwalidowane oraz interpretowalne mają pełną rację bytu i mogą być istotnie przydatnym narzędziem. Wszelka aktywność naukowa w tym zakresie jest bardzo wartościowa i przyczynia się do rozwoju wiedzy o procesie chromatograficznym, a przykładem tego jest przedłożona do recenzji rozprawa, wykonana pod kierunkiem prof. dr hab. n. farm. PAWŁA WICZLINGA.

Rozprawa składa się z cyklu trzech publikacji w wysoko punktowanych czasopismach wraz z dołączonym autoreferatem-komentarzem o objętości blisko 50 stron, zawierającym ponad 70 starannie dobranych pozycji literaturowych. Badania były finansowane z kilku projektów: SONATA BIS 5, POWR współfinansowanego przez Unię Europejską ze środków Europejskiego Funduszu Społecznego w ramach Programu Operacyjnego Wiedza Edukacja Rozwój, jak również korzystano z infrastruktury Centrum Informatycznego Trójmiejskiej Akademickiej Sieci Komputerowej do wykonywania obliczeń.

Do otrzymanej dokumentacji nie dołączono wprawdzie oświadczeń dotyczących wkładu współautorów w powstanie poszczególnych publikacji, jednakże na podstawie analizy treści autoreferatu oraz opisu wkładu autorów w publikacje można stwierdzić, że część badań analizowana przez Autorkę rozprawy jest wyraźnie wyodrębniona i wystarczająca do spełnienia wymagań Art. 187 Ustawy z dnia 20 lipca 2018 r. *Prawo o szkolnictwie wyższym i nauce*¹.

Doktorantka postawiła sobie za cel opracowanie modeli opisujących retencję szerokiego spektrum analitów w warunkach izokratycznych z wykorzystaniem masy molowej oraz grup funkcyjnych, bądź też $\log P$ i pK_a jako predyktorów. Ponadto powzięto udaną próbę opracowania bayesowskiego modelu hierarchicznego modelującego retencję analitów w warunkach gradientowych obejmujących różne wartości pH, rodzaje modyfikatora organicznego, temperatury i programu gradientowego. Do stworzenia tego modelu wykorzystano $\log P$, pK_a oraz liczebności poszczególnych grup funkcyjnych.

Dwa pierwsze modele zostały stworzone na podstawie ogólnodostępnych danych retencyjnych 1026 związków chemicznych, natomiast model gradientowy uzyskano na podstawie licznych własnych doświadczeń Zespołu Badawczego na 187 związkach. Wszystkie trzy publikacje przeszły pozytywnie rygorystyczny proces wydawniczy, a ich uważna lektura nie nasuwa żadnych wątpliwości wymagających rozwiania.

¹Dz. U. z 2023 r. poz. 742, 1088, 1234, 1672

Nowość naukowa zaprezentowanych modeli jest wyraźna i oczywista:

- (I) Podejście hierarchiczne wykazało szereg zalet w stosunku do najczęściej przedstawianego w dotychczasowej literaturze podejścia niehierarchicznego: głównie możliwość włączenia wcześniejszej wiedzy *a priori* łącznie z aplikowaniem tej wiedzy jedynie do określonych grup związków, jak również regularyzację parametrów.
- (II) Połączenie równania NEUE z podejściem bayesowskim pozwoliło na nowatorskie potraktowanie problemu, kwantyfikację niepewności predykcji i otrzymanie chromatogramu niepewności, zupełnie zmieniając spojrzenie na modele o szerokiej przestrzeni predykcyjnej.
- (III) Zaobserwowano bardzo istotne w praktycznym zastosowaniu zjawisko względnie stałej różnicy logarytmu współczynnika retencji pomiędzy formą zjonizowaną i niezjonizowaną, która ma związek z energią transferu międzyfazowego analitu.
- (IV) Stwierdzono większą liniowość (mniejszą krzywiznę) zależności logarytmu współczynnika retencji od zawartości modyfikatora organicznego dla molekuł zjonizowanych, a wyraźnie większą dla niezjonizowanych.
- (V) Wykryto istotną różnicę między rozkładem *a priori* i *a posteriori* dla parametrów dS_1a , S_2 i τ , co sugeruje różnicę w retencji pomiędzy kwasami i zasadami nieco inną przy takim modelowaniu, niż można byłoby oczekiwać.
- (VI) Zidentyfikowano parametry chromatograficzne w przybliżeniu niezależne od fazy stacjonarnej, a ich rozkłady mogą być użyte jako wiedza *a priori* w kolejnych badaniach.

Poniższe przemyślenia i pytania, jakie nasunęły mi się podczas lektury, nie stanowią o jakichkolwiek uchybieniach, a są jedynie zachętą do ewentualnej dyskusji na obronie pracy:

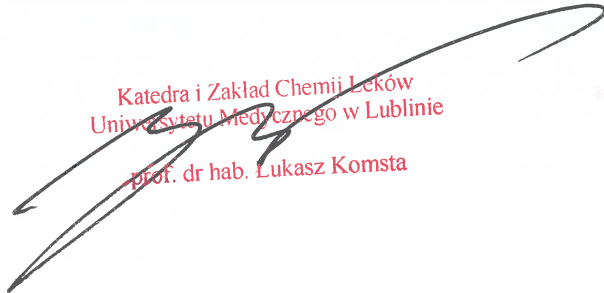
1. Doktorantka pisze (str. 36), że *nawet pojedynczy eksperyment zwykle daje dużo informacji, i tym samym pozwala zmniejszyć niepewność predykcji*. Jak w praktyce wyglądałoby modelowanie retencji w przypadku posiadania jakiejś informacji z pojedynczego eksperymentu?
2. W pracy stosowano różne oprogramowanie: R, Stan, Matlab itd. Jakie argumenty przemawiały za wyborem takiego, a nie innego oprogramowania? Jakie istniały alternatywy z grupy *Open Source* lub programów komercyjnych?
3. Która część obliczeń wymagała najwięcej czasu procesora i zmuszała do korzystania z infra-

struktury obliczeniowej?

4. Jakie alternatywne dla równania NEUE modele można by przyjąć w tworzeniu tych modeli i co przesądziło o wyborze akurat tego równania?
5. Na ile mogłyby się zmienić wyniki obliczeń, gdyby zamiast rozkładu półnormalnego dla parametrów skali został użyty rozkład χ^2 o pewnej (jakiej?) liczbie stopni swobody?

Podsumowując, praca doktorska mgr Agnieszki Kamedulskiej spełnia wszystkie wymagania stawiane pracom doktorskim, zawiera istotne elementy nowości naukowej i wnioskuję o dopuszczenie Doktorantki do dalszych etapów przewodu doktorskiego.

Dodatkowo, biorąc pod uwagę wyraźny i znaczny wkład badań w rozwój teorii chromatografii, chemometrię i przestrzeń zastosowań modeli bayesowskich, z przyjemnością wnioskuję o wyróżnienie przedłożonej mi do recenzji pracy.



Katedra i Zakład Chemii Leków
Uniwersytetu Medycznego w Lublinie
prof. dr hab. Łukasz Komsta

Lublin, 11 września 2023.