



dr hab. Aleksandra Bocian, prof. PRz
Katedra Biotechnologii i Bioinformatyki
Wydział Chemiczny
Politechnika Rzeszowska im. Ignacego Łukasiewicza w Rzeszowie

Rzeszów, 14.04.2023

**Recenzja rozprawy doktorskiej mgr inż. Anny Glodek pod tytułem
„Deisotoping methods in MALDI ToF Mass Spectrometry Imaging”**

W dzisiejszych czasach nikt chyba nie wyobraża sobie badań biologicznych, biomedycznych czy chemicznych bez spektrometrów mas. Pomiar stosunku masy do ładunku cząsteczek i ich fragmentów umożliwia identyfikację, określenie struktury a często także analizy ilościowe. Rozwój spektrometrów mas, w tym głównie opracowanie źródeł jonów umożliwiających łagodną jonizację i wysokorozdzielczych analizatorów pozwolił na rozkwit proteomiki, niezwykle ważnej i popularnej dziedziny badań biologicznych. Jednak analiza tak dużych cząsteczek jak białka jest trudna, nie tylko ze względu na rozbudowaną strukturę tych makrocząsteczek, ale także fakt, że w próbkach biologicznych tworzą skomplikowane mieszaniny składników różniących się właściwościami fizyko-chemicznymi i procentowym udziałem. W przypadku większości analiz mających na celu poznanie proteomu białka jako takie są zbyt duże, aby analizować je w postaci natywnej. Dlatego też niezbędny jest etap fragmentacji białka na peptydy, który w technice bottom-up wykonuje się poprzez trawienie enzymatyczne. Dzięki temu etapowi z białek powstają mniejsze peptydy, jednak powoduje to, że analizowana mieszanina staje się jeszcze bardziej skomplikowana, co pociąga za sobą znaczny wzrost złożoności widm MS i utrudnienie ich analizy.

Kolejnym problemem w analizach MS, który bezpośrednio wiąże się z tematyką recenzowanej rozprawy doktorskiej jest występowanie w naturalnych cząsteczkach izotopów o różnych masach. W przypadku białek problem ten dotyczy głównie atomów węgla, których około 1% to izotop C^{13} . Duża liczba atomów węgla w peptydach i białkach pochodzących z materiału biologicznego powoduje, że tak naprawdę każdy peptyd występuje w kilku wariantach. Jeden z nich, zwany w spektrometrii mas monoizotopowym, będzie zawierał tylko węgle C^{12} , ale znajdą się też takie zawierające jeden izotop C^{13} , dwa takie izotopy i więcej, w zależności od wielkości polipeptydu. Współczesne spektrometry mas mają wystarczającą rozdzielczość, aby te warianty od siebie odróżnić i dlatego też na widmach masowych obserwowane są zbiory sygnałów pochodzących od jednego peptydu, które nazywamy obwiednią izotopową. Jednak z biologicznego punktu widzenia każdy z tych sygnałów reprezentuje tę samą cząsteczkę,



czyli peptyd o określonej sekwencji, do identyfikacji i kwantyfikacji którego wystarczy tylko jeden z tych sygnałów na widmie MS. Powstaje więc problem jak wyłuskać z widma piki monoizotopowe, zwłaszcza jeżeli widmo MS jest bardzo skomplikowane a obwiednie izotopowe zachodzą na siebie. W wielu przypadkach rozwiązaniem jest dekompleksacja widma poprzez wcześniejsze frakcjonowanie próbki np. z użyciem elektroforezy lub chromatograficznego rozdzielania białek czy peptydów po trawieniu. Niestety zabieg ten jest niemożliwy w eksperymentach obrazowania MS takich, jak wykorzystany w rozprawie doktorskiej MALDI ToF Imaging. Wyniki uzyskane tą techniką dostarczają unikatowych informacji o przestrzennej dystrybucji poszczególnych cząsteczek w analizowanym materiale, niezwykle cennych między innymi z punktu widzenia proteomiki klinicznej.

Pomijając kwestie powtarzalności wyników, w technice MALDI ToF Imaging wielkim problemem jest ogromna ilość uzyskiwanych danych i ich duże skomplikowanie, co znacząco utrudnia analizę i interpretację wyników. Ponadto na złożoność widma MS wpływa także wspomniany wcześniej problem obwiedni izotopowych, znajdujących się na widmie w bezpośrednim sąsiedztwie lub wręcz zachodzących na siebie. W tym przypadku z pomocą mogą przyjść jedynie algorytmy upraszczające widma, wynajdujące sygnały monoizotopowe i przypisujące poszczególne piki do odpowiednich obwiedni izotopowych.

Stworzenie takiego algorytmu dedykowanego metodzie MALDI ToF Imaging Doktorantka postawiła sobie za cel, który z dużym sukcesem osiągnęła. W pracy wykorzystano system wnioskowania rozmytego Mandaniego-Assilana oraz ewaluacji przestrzennej dystrybucji pików do identyfikacji obwiedni izotopowych w widmach pochodzących z eksperymentu MALDI ToF Imaging a następnie model zweryfikowano za pomocą klasyfikatora. Cały system opiera się na porównywaniu par pików, które potencjalnie mogą należeć do jednej obwiedni izotopowej i wykluczeniu całej reszty. Zabieg ten, jak Doktorantka przedstawiła w tabeli 7, pozwala na zredukowanie liczby rozpatrywanych wariantów o co najmniej 96% co znacząco ułatwia i przyspiesza analizę. W kolejnym etapie wykorzystano przestrzenne mapy dystrybucji molekularnej do wyłonienia deskryptorów opartych na właściwościach pików oraz metrykach tekstury obrazu, aby przetestować działanie systemu. Działanie całej metody nazwanej DeisoLAB przetestowano następnie na zestawach danych pochodzących z eksperymentu MALDI ToF Imaging.

W tym miejscu chciałam zapytać Doktorantkę z czego wynikają tak drastyczne różnice uzyskane przy użyciu różnych klasyfikatorów, skoro wykorzystano te same dane wejściowe?





Recenzowana praca ma prawidłową strukturę, niemal typową dla rozpraw doktorskich. Przegląd literatury nie jest zbyt obszerny i pojawia się w nim kilka stwierdzeń, które nie doczekały się rozwinięcia, dlatego proszę o ich uzupełnienie w trakcie obrony:

1. Doktorantka napisała na stronie 9, że powstawanie jonów dodatnio naładowanych na końcu C jest korzystne z punktu widzenia spektrometrii mas. Dlaczego?
2. Jaka jest rola płytek pokrytych ITO w obrazowaniu MALDI?
3. Jakie jest wyjaśnienie faktu, że dla peptydów o masie do ok. 2 kDa najwyższy na widmie MALDI ToF jest pik monoizotopowy, a powyżej tej masy najwyższym jest jeden z kolejnych pików z obwiedni?
4. Które z algorytmów zestawionych w tabeli 4 są zaimplementowane w oprogramowaniu stosowanym przez użytkowników spektrometrów mas MALDI ToF, np. FlexAnalyser lub FlexImaging?

Sekcja bioinformatyczna nie jest napisana w przejrzysty sposób. Moim zdaniem łatwiejsze dla odbiorcy byłoby jasne rozdzielenie założeń i części literaturowej od opisu tego co faktycznie zostało przez Doktorantkę wykonane. W pracy brak również podstawowych informacji dotyczących konfiguracji systemu, na którym wykonywane były obliczenia. Czy zastosowany hardware to komputer klasy desktop, workstation czy może dedykowany serwer obliczeniowy? Pod jakim systemem operacyjnym działał? Nie przedstawiono również podstawowych informacji na temat wykorzystanego oprogramowania. Zaledwie w jednym punkcie wspomniane jest środowisko MATLAB brak jednak informacji o jego wersji czy zainstalowanych modułach. Co najważniejsze w rozprawie doktorskiej nie ma skryptu programu ani żadnej informacji czy został on ujawniony. Proszę o uzupełnienie tych informacji w trakcie obrony.

Bardzo skromnie opisana jest sekcja metodyczna dotycząca eksperymentu MS. Proszę o informację na temat zastosowanej procedury trawienia, zastosowanej matrycy i sposobie jej aplikowania.

Na stronie 29 znajduje się opis zestawu danych z próbek świeżo mrożonych i jest to 45 tysięcy surowych widm. Wcześniej podano, że próbki pochodziły od czterech pacjentów. Czy jest to suma dla wszystkich czterech próbek? Pytanie nasuwa się ponieważ na stronie 30 Doktorantka podaje analogiczne informacje dla próbek utrwalonych w formalinie i tam podane są cztery liczby, tak jakby dla każdego z pacjentów osobno. Z czego wynikają różnice w opisie danych?

Analizowano dwa zbiory danych: pochodzący z tkanek świeżo mrożonych i utrwalanych w formalinie. Analizowano je tak samo, natomiast nie przedstawiono danych dla obu zestawów. Proszę pokazać tabelę dla FFPE analogiczną do tabeli 9 i dla FF analogiczną do tabeli 10. Proszę też wyjaśnić z czego wynikają tak duże różnice pomiędzy FF i FFPE.



Z formalnego punktu widzenia praca ma szereg niedociągnięć. Kropka kończąca zdanie powinna być zawsze po przypisie a nie przed. Skrót *m/z* zawsze pisany jest kursywą.

W pracy dwukrotnie powtarzają się identyczne paragrafy tekstu (odpowiednio strony 85 i 89 oraz 85 i 90), w których zmienione są tylko dane liczbowe. Na rysunku 77 są nieprawidłowo wprowadzone strzałki. Co najmniej 48 ze 123 pozycji literaturowych nie zawiera nazwy czasopisma, co samo w sobie byłoby być może akceptowalne gdyby zachowana była konsekwencja w całym spisie. W dwóch pozycjach nieprawidłowo podano nazwiska autorów a w jednej pozycji w ogóle nazwisk nie ma.

Pomimo licznych uwag przedstawionych powyżej recenzowana rozprawa doktorska rozwiązuje istotny problem badawczy, więc postawiony cel badawczy został osiągnięty. Opracowany system skutecznie ogranicza ilość danych podlegających analizie mającej na celu zidentyfikowanie obwiedni izotopowych przez co zwiększa przepustowość i pozwala na analizę większych zestawów danych. System radzi sobie z nakładającymi się obwiedniami izotopowymi, co daje mu przewagę nad większością istniejących algorytmów. Jak się okazuje porównywanie par pików było strzałem w dziesiątkę i czyni opracowany algorytm jedynym w swoim rodzaju. Na uwagę zasługuje także fakt, że metoda rozpoznaje więcej obwiedni izotopowych niż ekspert, uzyskując skuteczność rzędu 97%. Nasuwa się jednak pytanie jak należałoby przekształcić ten system, żeby mógł być wykorzystany do analizy danych pochodzących ze spektrometrów mas ze źródłem generującym jony wielokrotnie naładowane i czy jest to w ogóle możliwe?

Podsumowując przedstawioną rozprawę doktorską Pani mgr inż. Anny Glodek oceniam pozytywnie oraz stwierdzam, że spełnia ona wymogi określone dla osób ubiegających się o nadanie stopnia naukowego doktora w myśl ustawy z dnia 3 lipca 2018 r (przepisy wprowadzające ustawę Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2018 r., poz. 1669, z późn. zm.) oraz ustawa z 20 lipca 2018 r. — Prawo o szkolnictwie wyższym i nauce (Dz.U. z 2018 r., poz. 1668 z późn. zm)). Doktorantka udowodniła, że samodzielnie potrafi rozwiązywać sformułowany problem naukowy poprzez odpowiednio zaplanowane eksperymenty, ich interpretację, krytyczną dyskusję oraz wyważone wnioski. W związku z powyższym wnoszę do Rady Naukowej Dyscypliny Inżynieria Biomedyczna Politechniki Śląskiej o dopuszczenie Pani mgr inż. Anny Glodek do dalszych etapów postępowania w przewodzie doktorskim.

dr hab. Aleksandra Bocian, prof. PRz