

Metody spektrometrii mas oparte o dane (DDA) oraz metody niezależne od danych (DIA) wykorzystywane w analizie materiału biologicznego

dr Marta Nolka-Szaszner^{1✉},

dr inż. Aleksander Strugała^{2✉},

dr Łukasz Marczał²

¹Zakład Proteomiki Biomedycznej, Instytut Chemii Bioorganicznej Polskiej Akademii Nauk, Poznań

²Pracownia Spektrometrii Mas, Instytut Chemii Bioorganicznej Polskiej Akademii Nauk, Poznań

https://doi.org/10.18388/pb.2021_535

✉ autor korespondujący: mnolka@ibch.poznan.pl, astruga@ibch.poznan.pl

Słowa kluczowe: spektrometria mas, DDA, DIA, analiza proteomiczna, analiza metabolomiczna

Wykaz skrótów: DDA – analiza zależna od danych (ang. *Data-Dependent Acquisition*); DDIA – połączona analiza zależna i niezależna od danych (ang. *Data Dependent-Independent Acquisition*); DE – dynamiczne wykluczenie (ang. *Dynamic Exclusion*); DIA – analiza niezależna od danych (ang. *Data-Independent Acquisition*); IE – wielkokrotne wykluczenie (ang. *Iterative Exclusion*); IL – lista włączenia (ang. *Inclusion List*); MRM – tryb monitorowania reakcji następczych (ang. *Multiple Reaction Monitoring*); PRM – tryb monitorowania reakcji równoległych (ang. *Parallel Reaction Monitoring*); RTW – okno czasu retencji (ang. *Retention Time Window*); SRM – tryb monitorowania pojedynczej reakcji (ang. *Single Reaction Monitoring*); XIC – chromatogram wyodrębnionych jonów (ang. *Extracted Ion Chromatogram*)

Finansowanie: Artykuł powstał podczas realizacji projektu OPUS nr 2020/39/B/NZ9/03336 sfinansowanego ze środków Narodowego Centrum Nauki.

STRESZCZENIE:

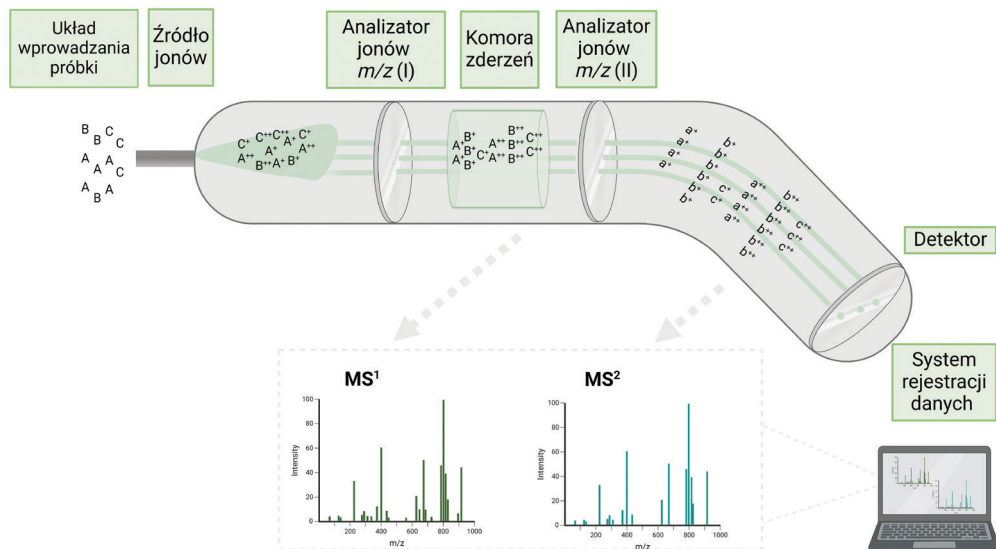
Spektrometria mas jest nieocenionym narzędziem w analizie proteomicznej, metabolomicznej i lipidomicznej. Dla pełnego wykorzystania jej potencjału kluczowym jest wybranie i skonfigurowanie odpowiedniego podejścia analitycznego. W przypadku badań niecelowanych dostępne są dwie główne strategie: analiza zależna od danych (DDA) oraz analiza niezależna od danych (DIA). Obie metody różnią się sposobem przeprowadzenia analizy i stopniem pokrycia uzyskanych danych, dlatego każda z nich może znaleźć zastosowanie w różnych typach badań. Metoda DDA opiera się na bieżącym skanowaniu analizowanych jonów, w efekcie czego prekursorzy o najwyższej intensywności poddawane są fragmentacji. Z kolei technika DIA, dzięki zastosowaniu łączonych zakresów izolacji jonów prekursorowych, pozwala na głębszą analizę badanych związków. Oba podejścia można modyfikować usprawniając ich działanie i umożliwiając uzyskanie większej ilości danych. Na horyzoncie pojawiają się także metody łączące obie techniki, jak np. DDIA, która wykorzystuje zalety obu tych metod otwierając nowe możliwości analityczne.

WPROWADZENIE

Ostatnie dwie dekady, w których badania genomu i transkryptomu wiodły zdecydowany prym, pozwoliły na głębsze poznanie wielu procesów biologicznych leżących u podstaw m.in. dziedziczenia, procesów patologicznych, czy efektów terapii medycznych. Jednakże zdobyta w ten sposób wiedza uwypukliła fakt, że bez informacji o pełnym profilu białkowym i metabolomicznym, pozostaje ona co najmniej niepełna. Stężenie białek i metabolitów w komórce, ich profil, czy modyfikacje, które stanowią o funkcji i fenotypie komórki, są na tyle zróżnicowane, że ich analiza nastęrcza wiele trudności. Obserwujemy jednak bardzo szybki rozwój technik umożliwiających analizę nie tylko pojedynczych białek i metabolitów, ale także, a może przede wszystkim, szerszego spektrum proteomu i metabolomu. Metodą z wyboru w ich analizie coraz częściej jest spektrometria mas. Umożliwia ona zarówno badania ilościowe jak i jakościowe, co wykorzystywane jest nie tylko w diagnostyce medycznej, projektowaniu i produkcji nowych leków [1] czy poszukiwaniu biomarkerów [2], ale także w tak na pozór odległych dziedzinach, jak archeologia i paleontologia [3,4].

Spektrometr mas zbudowany jest z trzech głównych komponentów: źródła jonów, analizatora oraz detektora (Ryc. 1). Powstałe w źródle jony przyśpieszane są w polu elektromagnetycznym i kierowane do pracującego pod wysoką próżnią analizatora mas. Rozdział analizowanych jonów na podstawie stosunku masy do ładunku (m/z) przedstawiany jest graficznie w postaci wykresu MS^1 – widma mas, na którym zaznaczana jest na osi odciętych wartość stosunku m/z , a na osi rzędnych intensywność sygnałów. Analizatory można łączyć uzyskując spektrometry tandemowe, które umożliwiają głębszą analizę badanych próbek. W pierwszym analizatorze odbywa się wybór jonu o określonej wartości m/z . Następnie wybrany jon ulega rozpadowi w komorze zderzeń, a masy powstałych fragmentów mierzone są za pomocą drugiego analizatora. Uzyskane w ten sposób wyniki rejestrowane są w postaci widma fragmentacyjnego – MS^2 .

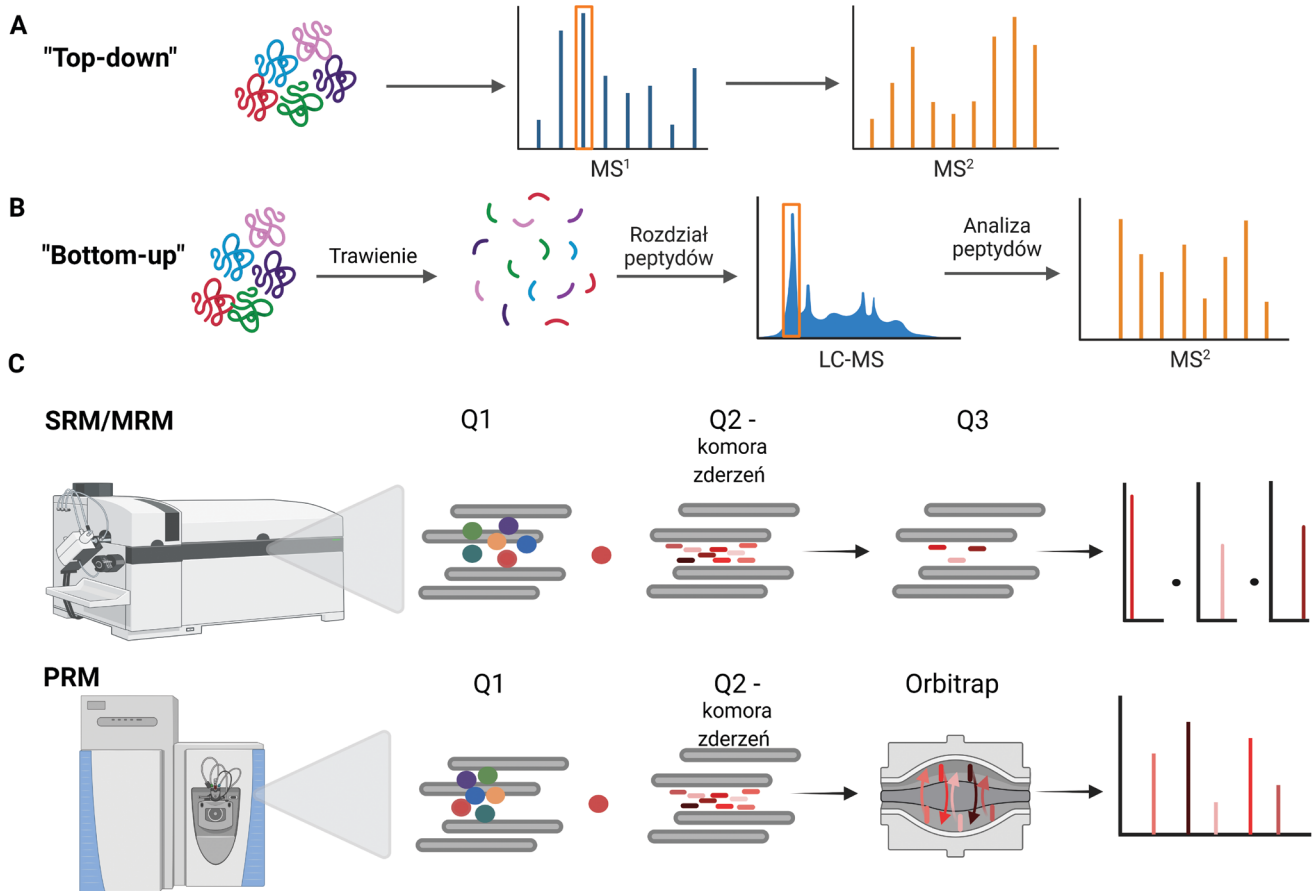
Zarówno w badaniach proteomicznych, metabolomicznych czy lipidomicznych, analiza interesujących nas obszarów z wykorzystaniem spektrometrii mas dostarcza ogromnych ilości danych. Kluczowym aspektem jest więc zastosowanie odpowiedniej strategii w przygotowaniu materiału badawczego, a także metod analizy instrumentalnej i chemometrycznej, służących uzyskaniu danych pozwalających na identyfikację badanych molekuł. W przypadku spektrometrii mas stosuje się dwie główne strategie analityczne, tzw. „top-down” oraz „bottom-up”. Na przykładzie proteomiki, strategia „top-down” zakłada, że całe, nietrawione białka zostają poddane jonizacji, a następnie fragmentacji w celu uzyskania widma MS^2 . Natomiast, strategia „bottom-up” polega na uprzedniej hydrolizie białek przy użyciu odpowiedniego enzymu (najczęściej trypsyny), a



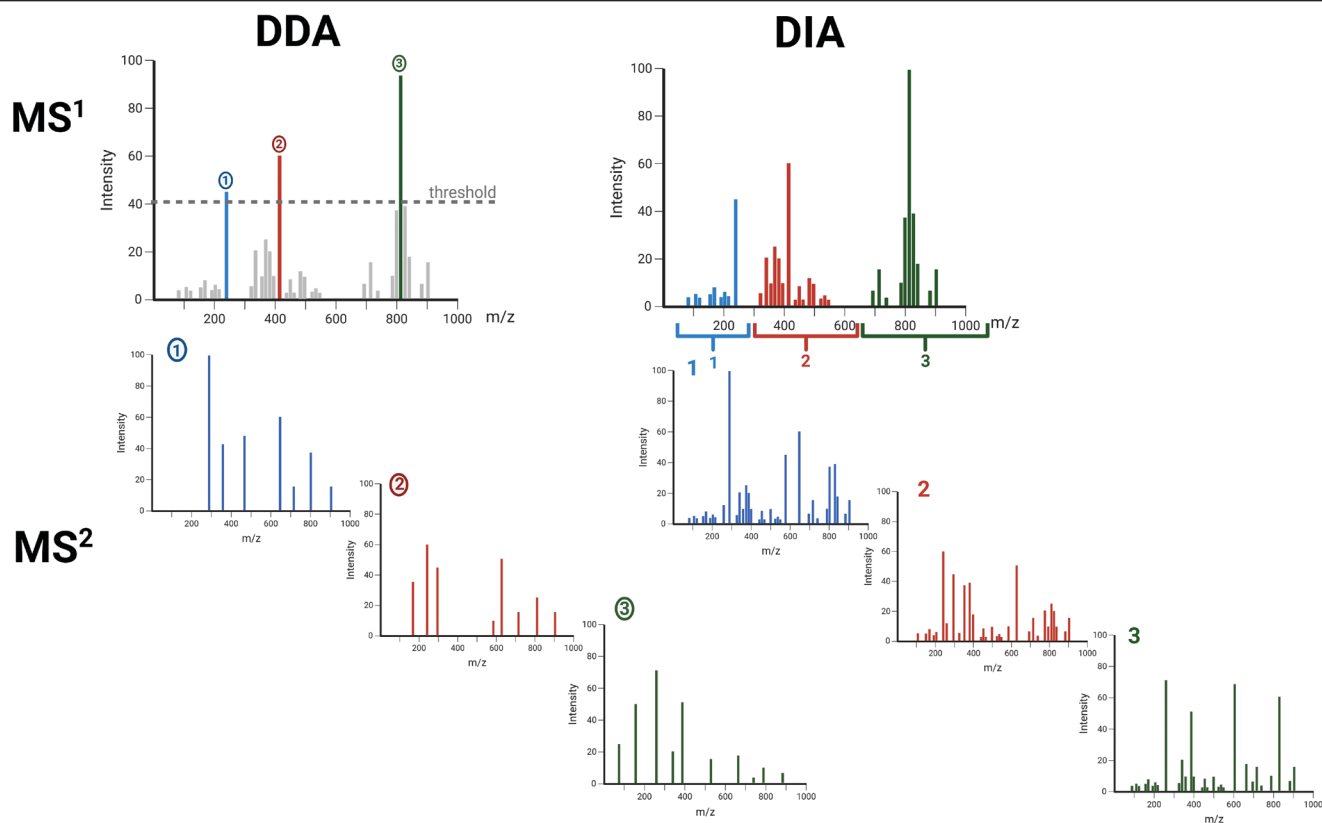
Rycina 1. Ogólny schemat blokowy spektrometru mas, który składa się ze źródła jonów, analizatora(ów) oraz detektora. Dużymi literami oznaczono jony prekursorowe, małymi literami jony fragmentacyjne.

następnie na separacji mieszaniny peptydów z użyciem metod chromatograficznych. Tak rozdzielone peptydy trafiają do spektrometru mas, gdzie po jonizacji następuje ich rozdział w oparciu o stosunek masy do ładunku (m/z), a w dalszej kolejności identyfikacja i kwantyfikacja. Podejścia analityczne oparte o MS można również podzielić na celowane

i niecelowane. Głównymi metodami analizy celowanej są podejścia: SRM/MRM (ang. *Selective/Multiple Reaction Monitoring*) i PRM (ang. *Parallel Reaction Monitoring*), a analizy niecelowanej: DDA (ang. *Data Dependent Acquisition*) i DIA (ang. *Data Independent Acquisition*) (Ryc. 2, 3).



Rycina 2. Schematyczne przedstawienie podejść „top-down” (A) i „bottom-up” (B) w analizie proteomicznej oraz podstawowych metod analiz celowanych opartych o MS (C). MS¹ – podstawowy skan MS; MS² – skan MS po fragmentacji wybranego jonu; LC-MS – chromatografia cieczowa z detekcją MS; Q1, Q2, Q3 – analizatory typu kwadrupol w tandemowym spektrometrze mas; na rysunku C kolorami oznaczono poszczególne jony o różnych wartościach m/z .



Rycina 3. Porównanie metody DDA oraz DIA. W podejściu DDA po uzyskaniu skanu podstawowego MS¹ wykonywane są skany fragmentacyjne w oparciu o ściśle określone w metodzie parametry, najczęściej na podstawie intensywności jonów prekursorowych. W podejściu DIA wszystkie prekursorowe z zadanego zakresu *m/z* są wspólnie fragmentowane dając uśrednione widmo MS². W ten sposób w analizie DIA wszystkie prekursor są poddawane fragmentacji.

CHARAKTERYSTYKA METODY MS ZALEŻNEJ OD DANYCH (DDA)

W podejściu DDA informacje uzyskane na podstawie skanu MS¹ służą do dalszej analizy w kolejnych skanach MS². W pierwszym analizatorze (najczęściej jest to kwadrupol) następuje pomiar jonów według ustalonego wcześniej zakresu *m/z*. Na podstawie uzyskanego widma MS¹ typowane są jony prekursorowe, które następnie ulegają fragmentacji w komorze zderzeń (ang. *Collision Induced Dissociation*, CID) i powstałe fragmenty są badane w drugim analizatorze [6]. Ze względu na drogę, którą musi pokonać jon (od pierwszego analizatora, przez komorę zderzeń, do drugiego analizatora), ważnym parametrem jest dobór odpowiedniego czasu skanowania zarówno w trybie MS¹ jak i MS² oraz samej liczby generowanych skanów MS² [5]. Wraz ze wzrostem czasu skanowania zwiększa się czułość pomiaru. Nie można jednak w nieskończoność rozciągać czasu pomiaru i najlepiej ustalić te wartości eksperymentalnie. W przypadku analiz metabolomicznych przyjmuje się wartości mieszczące się w zakresie od 20 do 800 ms czasu przeznaczanego na skan oraz 3 do 20 skanów MS² [5,7].

W przypadku prowadzenia analiz MS zależnych od danych należy zwrócić uwagę na dobór właściwego analizatora, który powinien charakteryzować się wysoką rozdzielczością i odpowiednią szybkością akwizycji uzyskiwanych danych. Chociaż w teorii metoda DDA może być implementowana na każdym tandemowym spektrometrze mas, najczęściej do tego typu analiz wykorzystuje się nowocze-

sne analizatory typu Q-TOF lub Orbitrap [5]. Spektrometry z analizatorami typu potrójny kwadrupol lub pułapka jonowa stosowane są stosunkowo rzadko do wykonywania wysokoprzepustowych analiz DDA, głównie ze względu na ich niską rozdzielczość i małą dokładność pomiaru masy.

Kolejnymi ważnymi parametrami w przypadku prowadzenia analiz metodami DDA są tzw. punkty odcięcia, według których spektrometr zakończy proces skanowania MS¹ i przejdzie do trybu MS². Mogą to być punkty odcięcia określane na podstawie parametrów takich jak: minimalna intensywność sygnału dla jonu lub minimalna, czy maksymalna dopuszczalna wartość *m/z*. W przypadku analiz proteomicznych parametrem odcinającym może być również krotność jonizacji badanego związku, gdzie w przypadku peptydów pomija się cząstki jednokrotnie zjonizowane. Parametr ten ma o tyle istotne znaczenie, że zbyt wysoko ustawiona linia odcięcia będzie powodowała ominięcie istotnych jonów, jednak ustawiona na zbyt niskim poziomie spowoduje obniżenie jakości widm MS² [5].

Wybór jonów do analizy MS² jest kluczowym elementem działania metody DDA. Odbywa się on w sposób semistochastyczny [8]. W efekcie, jony pochodzące od istotnych biologicznie peptydów lub metabolitów, które w badanym materiale są obecne w mniejszej ilości, mogą zostać pominięte w trakcie analizy. Aby jednak zniwelować ten efekt i uzyskać z danego eksperymentu jak najwięcej interesujących nas danych, można odpowiednio zoptymalizować ustawienia spektrometru [9].

Przy wyborze jonów prekursorowych najczęściej stosuje się tzw. podejście TopN. Polega ono na wyborze N liczby jonów, które zostaną poddane analizie MS². Wyboru tego dokonuje się na podstawie intensywności ich sygnału *m/z*. Jednym z narzędzi pozwalających na bardziej dogłębną analizę próbki jest statyczna lista wykluczenia (ang. *Static Exclusion List*, SEL), zawierająca jony, które mają być pomijane podczas analizy MS, co ma tym samym umożliwić detekcję jonów, o niższej abundancji. W zależności od zastosowanego analizatora lista może zawierać różną liczbę jonów. Nowsze układy MS pozwalają na usunięcie nawet kilkuset prekursorów korzystając z tej metody. Oprócz statycznej listy, stosuje się również dynamiczne listy wykluczenia. W metodzie TopN najczęściej używany jest jeden z dwóch głównych algorytmów: dynamiczne wykluczenie (ang. *Dynamic Exclusion*, DE) lub wielokrotne wykluczenie (ang. *Iterative Exclusion*, IE) [10]. W algorytmie DE po pierwszym skanie wytypowane zostają jony, które trafiają na „dynamiczną”, czyli czasową listę. Jony te nie będą brane pod uwagę w następnych skanowaniach. Procedura ta powtarzana jest wielokrotnie i dzięki temu można wytypować wiele potencjalnych prekursorów, jednak nadal są to jony o największej intensywności [10]. Warto zaznaczyć, że związki które trafiają na listę DE, są wykluczane z analizy tylko na pewien czas, dzięki czemu pomijane w badaniu jony o danych wartościach *m/z* będą mogły być analizowane później. Ma to szczególne znaczenie w przypadku możliwej obecności związków izomerycznych, czy izobarycznych w próbce. Drugie podejście (IE) opiera się na wielokrotnym analizowaniu danej próbki. Podobnie jak w przypadku DE, pierwszy skan pozwala wytypować najbardziej intensywne jony i umieścić je na liście, która w ponownej analizie posłuży do wykluczenia wybranych jonów z analizy fragmentacyjnej. Zastosowanie IE pozwala na wytypowanie jonów mających niższe intensywności niż w metodzie DE, wymaga jednak większych objętości próbki, gdyż jest ona analizowana wielokrotnie, co zmniejsza wydajność i przepustowość metody. Próbą odpowiedzi na niedoskonałości metody TopN jest tzw. metoda DiffN, zaprezentowana na przykładzie analizy lipidów [10]. Nie bazuje ona na absolutnej abundancji jonu, ale na porównaniu różnic ilości danego jonu w kolejnych analizowanych próbkach. W efekcie, do analizy MS² jesteśmy w stanie wybrać jony prekursorowe, które najbardziej różnicują w danym materiale. Metoda ta może być wykorzystywana m.in. przy poszukiwaniu nowych biomarkerów [10].

Oprócz listy wykluczającej, w metodzie DDA istnieje także opcja utworzenia „listy włączającej” jony do analizy (ang. *Inclusion List*, IL) [5]. IL może być utworzona przed analizą lub też po pierwszym skanowaniu próbki. W pierwszym przypadku odgórnie wybieramy jony prekursorowe, natomiast w drugim, analiza widma MS¹ dostarcza informacji o jonach, które zostały pominięte w badaniu. W dalszej kolejności, na podstawie wartości *m/z* oraz tzw. okna czasu retencji (ang. *Retention Time Window*, RTW) algorytm typuje jony do IL, a tym samym do ich dalszej analizy. Podejście takie zwiększa czułość analizy, a równocześnie liczbę związków, które można zidentyfikować [11]. Metoda DDA z zastosowaniem IL ma pewne cechy analizy celowa-

nej, dzięki czemu pozwala na częściowe zawężenie badania i jego udokładnienie.

Kolejnym ważnym czynnikiem w przypadku analizy DDA jest energia kolizji jonów użyta do ich fragmentacji. Przy zbyt małej energii jon albo nie ulegnie rozbiciu, albo jego dysocjacja będzie tylko częściowa, co spowoduje niewystarczającą intensywność sygnału jonów fragmentacyjnych [12]. Odwrotnie, przy zbyt wysokiej energii kolizyjnej, obserwuje się efekt nadmiernej fragmentacji, co z kolei prowadzi do otrzymania widm z niewspółmiernie reprezentowanymi jonami o niskich wartościach *m/z*. Ogranicza to istotnie przydatność takich widm do analiz strukturalnych i ilościowych.

Programy sterujące spektrometrami mas pozwalają również ustawić filtry ułatwiające wyłapywanie właściwych jonów prekursorowych. Najczęściej stosowane są dwa typy takich filtrów. Pierwszym jest funkcja wyłączania izotopów (ang. *Monoisotopic Precursor Selection*, MPS). Dzięki temu filtrowi, spektrometr nie będzie analizował izotopologów danego jonu i tym samym zaoszczędzi czas na analizę tylko jonów monoizotopowych, zwiększając tym samym całkowitą sumę przebadanych jonów różnego pochodzenia. Drugim, często stosowanym filtrem jest filtr separujący jony o większej liczbie ładunków od jonów pojedynczo naładowanych, co ma znaczenie głównie w analizach metabolomicznych czy lipidomicznych [5].

NARZĘDZIA BIOINFORMATYCZNE WYKORZYSTYWANE W METODZIE DDA

Kluczowym elementem analizy danych uzyskanych w efekcie zastosowania tandemowej spektrometrii mas jest dobór narzędzi bioinformatycznych służących do analizy widm MS². W przypadku analiz proteomicznych istnieje wiele algorytmów wspierających opracowanie widm, przeszukiwanie baz danych oraz bibliotek widm masowych. Wśród nich najbardziej znanymi i zarazem najczęściej wykorzystywanymi są: SEQUEST, Mascot oraz MaxQuant (Andromeda). Podstawową różnicą między nimi jest funkcja oceny wyniku analizy, czyli tzw. „score”, która odpowiada za dopasowanie zarejestrowanego widma fragmentacyjnego do widm pochodzących z repozytorium. SEQUEST był jednym z pierwszych, komercyjnie dostępnych programów służących do analizy widm MS². Jego funkcja „score” opiera się na korelacji krzyżowej i szybkiej transformacji Fouriera (ang. *Fast Fourier Transform*, FFT) [13]. SEQUEST jest bardzo czuły, jednakże ze względu na użytą FFT jego działanie jest bardzo wolne. Istnieje modyfikacja algorytmu używanego przez SEQUEST, w której pomija się liczenie FFT [13]. Algorytm SequestHT pozwala na przyspieszenie procesu analizy danych przy jednoczesnym poprawieniu stopnia identyfikacji.

Innym przykładem narzędzia służącego do identyfikacji i kwantyfikacji białek w spektrometrii mas jest Mascot. Algorytm użyty w programie Mascot w analizie widm peptydowych stosuje podobne podejście jak wspomniany wyżej SEQUEST, jednak operacje matematyczne leżące u podstaw obu metod są różne [14]. Mascot wykorzystuje obliczenia probabilistyczne do porównania danych uzyskanych z wid-

ma fragmentacyjnego i trawienia białka *in silico*. Co więcej oba algorytmy dają nieco inne wyniki w zależności od użycia w trakcie analizy układu MS [14].

Jednym z najczęściej stosowanych programów bioinformatycznych w analizie danych proteomicznych jest MaxQuant, który swoją premierę miał w roku 2008 [15]. Program ten wyposażony jest we własny algorytm do przeszukiwania baz danych peptydów – Andromedę. Jest on w stanie zidentyfikować więcej niż jeden peptyd z widma MS². Wynika to z tego, że MaxQuant prowadzi drugie skanowanie widma MS² w poszukiwaniu kofragmentujących jonów [15,16]. MaxQuant do analizy danych wykorzystuje tzw. strategię „target – decoy”, która polega na wyszukiwaniu błędnych identyfikacji białek, co w przypadku analizy bardzo dużej liczby peptydów ma ogromne znaczenie [15]. Warto dodać, że MaxQuant jest programem niekomercyjnym i stale rozwijanym, co zdecydowanie zwiększa jego aplikacyjność w badaniach.

Oprócz wyżej omówionych narzędzi, istnieją inne algorytmy i programy, które mogą być wykorzystane do analizy widm masowych mieszaniny peptydów uzyskanych przy pomocy metody DDA. Wśród nich są m.in.: X!Tandem, Comet, czy OMSSA. Warto jednak zwrócić uwagę, że wszystkie te narzędzia wykorzystują nieco inne podejście do analizy widm, a zatem wyniki, które uzyskamy z ich użyciem mogą się między sobą różnić. Jednym z ciekawszych przykładów podnoszących tą kwestię, jest praca dotycząca analizy spektrometrycznej białek jadu kobry [17]. W badaniu tym użyto 2 różnych algorytmów stosowanych do analizy ilościowej peptydów, NSAF+ zaimplementowa-

nego w programie PeptideShaker oraz algorytmu iBAQ zaimplementowanego w programie MaxQuant. Wnioski z tej analizy jasno wskazują, że w zależności od zastosowanego narzędzia, liczba identyfikowanych białek może się różnić. Praca ta podkreśla, że należy być ostrożnym w interpretacji danych oraz sugeruje, że dobór narzędzi bioinformatycznych powinien być starannie przemyślany i dostosowany do rodzaju badanych próbek [17]. Podsumowanie narzędzi bioinformatycznych przedstawiono w Tabeli 1.

WYKORZYSTANIE SPEKTROMETRII MOBILNOŚCI W METODZIE DDA (DDA – PASEF)

W chromatografii cieczowej sprzężonej ze spektrometrią mas wiele związków stanowiących potencjalne prekursorzy opuszcza kolumnę chromatograficzną równocześnie. W metodzie DDA w danym czasie, z całej mieszaniny jonów, tylko jeden jon prekursorowy może zostać poddany fragmentacji, a reszta zostaje utracona. W odpowiedzi na ten problem opracowana została metoda PASEF (ang. *Parallel Accumulation – Serial Fragmentation*) wykorzystująca dodatkowy wymiar analizy jakim jest spektrometria mobilności jonowej TIMS (ang. *Trapped Ion Mobility Spectrometry*) w spektrometrach firmy Bruker Daltonics [18]. Analiza ta składa się z 2 etapów. W pierwszym następuje akumulacja jonów, która odbywa się w module TIMS. Następnie, poprzez obniżanie napięcia pola elektrycznego, zatrzymane jony podlegają selektywnemu uwolnieniu do dalszej analizy. Dzięki takiemu podejściu znacznie więcej jonów może ulec fragmentacji. Zastosowanie metody PASEF pozwala na uzyskanie 10 krotnie większej szybkości analizowanych próbek przy zachowaniu tej samej czułości metody [18].

Tabela 1. Narzędzia bioinformatyczne używane w analizach DDA i DIA

PROGRAM	ANALIZA		OPROGRAMOWNIE			DARMOWY DOSTĘP	UWAGI	LINK
	DDA	DIA	Windows	Linux	Mac			
SEQUEST	✓	✗	✓	✓	✗	✗	nie jest już wspierany	https://proteomicsresource.washington.edu/protocols06/sequet.php
Mascot	✓	✓	✓	✓	✗	✗	DIA możliwe tylko dla analiz wykonywanych na spektrometrach ThermoScientific i w wąskich zakresach okna	https://www.matrixscience.com/
MaxQuant	✓	✓	✓	✓	✗	✓	DIA możliwe po aktualizacji 2.0 i z użyciem biblioteki MaxDIA	https://www.maxquant.org/
XITandem	✓	✗	✓	✓	✓	✓	-	https://www.thegpm.org/tandem/
Comet	✓	✗	✓	✓	✗	✓	-	https://comet-ms.sourceforge.net/
OMSSA	✓	✗	✓	✓	✓	✓	-	https://bioinformaticshome.com/tools/proteomics/descriptions/OMSSA.html#gsc.tab=0
PeptideShaker	✓	✗	✓	✓	✓	✓	-	http://compomics.github.io/projects/peptide-shaker
Skyline	✗	✓	✓	✗	✗	✓	-	https://skyline.ms/project/home/begin.view?/
DIA-NN	✗	✓	✓	✓	✗	✓	-	https://github.com/vdemichev/DiaNN
Spectronaut	✗	✓	✓	✓	✗	✗	-	https://biognosys.com/software/spectronaut/
DIA-Umpire	✗	✓	✓	✓	✗	✓	-	https://diaumpire.nesvilab.org/
PECAN	✗	✓	✓	✓	✗	✓	-	https://bitbucket.org/maccosslab/pecan/src/master/
FragPipe	✓	✓	✓	✓	✓	✓	-	https://fragpipe.nesvilab.org/
PEAKS Studio	✓	✓	✓	✓	✗	✗	-	https://www.bioinfor.com/peaks-studio/
EncyclopeDIA	✓	✓	✓	✓	✓	✓	-	https://bio.tools/encyclopedia
Scaffold DIA	✓	✓	✓	✓	✓	✓	jest również wersja dla analizy DDA: Scaffold DDA	https://www.proteomesoftware.com/products/scaffold-dia

Dzięki możliwości analize kilku jonów w jednym skanie, metoda PASEF pozwala również na wykorzystanie mniejszej ilości materiału. Jest to szczególnie ważne w przypadku analizy próbek takich jak np. tkanki z biopsji, gdzie sam proces jej pozyskania warunkuje małą ilość materiału, czy w analizach pojedynczej komórki (ang. *Single Cell Proteomics*), w których takie podejście istotnie poprawia wydajność analityczną [19].

Metoda DDA, mimo, że powoli wypierana przez DIA, nadal wykorzystywana jest w badaniach biomedycznych. Kompleksowe analizy metabolomiczne i proteomiczne oparte o podejście DDA umożliwiły na przykład identyfikację potencjalnych biomarkerów związanych z odpowiedzią raka odbytnicy na przedoperacyjną radioterapię [20].

CHARAKTERYSTYKA METODY MS NIEZALEŻNEJ OD DANYCH (DIA)

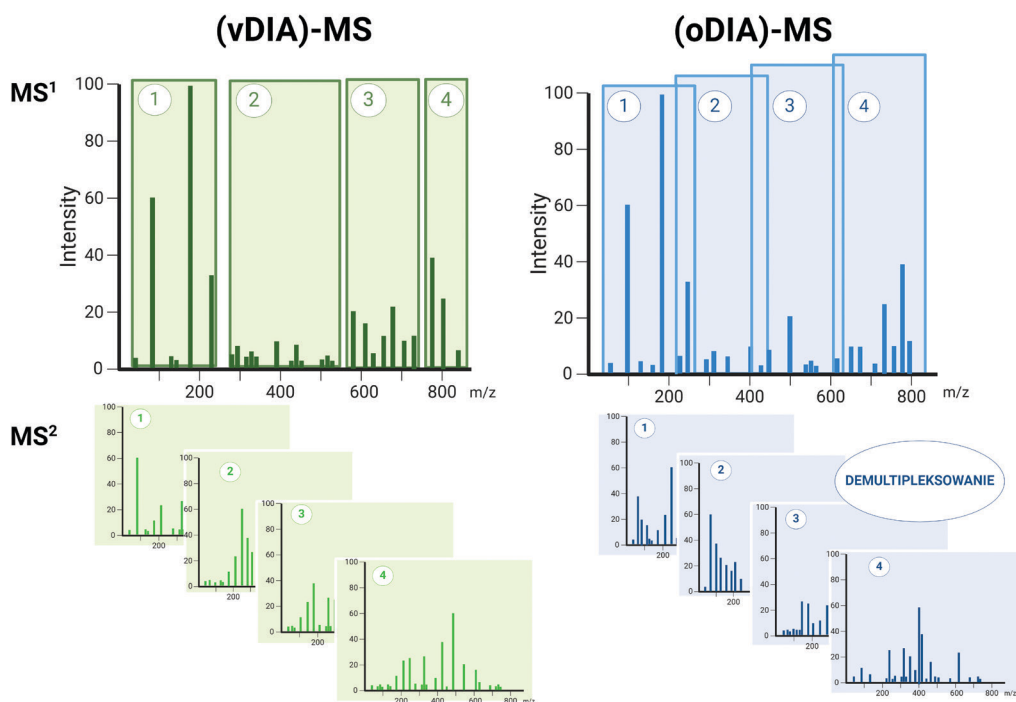
Analizy DDA pomimo swoich zalet: łatwości aplikacyjnej i powszechnego zastosowania, mają jedną zasadniczą wadę jaką jest utrata informacji dotyczącej jonów, które nie zostały wskazane do fragmentacji i w konsekwencji nie były jej poddane. Fakt ten powoduje, że analizy DDA nie umożliwiają uzyskania pełnej informacji o składzie analizowanej próbki, a otrzymane wyniki często nie są powtarzalne. Rozwój narzędzi bioinformatycznych oraz technologii w konstrukcji spektrometrów mas sprzyjał opracowaniu podejścia analizy niezależnej od danych DIA, omijającego powyższe niedogodności. Największą zaletą DIA jest możliwość uzyskania pełnej informacji o jonie na podstawie skanowania w trybach MS^1 oraz MS^2 , a także precyzja i powtarzalność będąca unikalną cechą metod SRM/PRM. DIA polega na równoczesnym wykonywaniu fragmentacji

(MS^2) dla wszystkich związków ulegających w danym momencie jonizacji. W metodzie tej analizowane są wszystkie jony obecne w próbce jednocześnie. Możliwa jest również analiza partiami, w szybko następujących po sobie skanach MS^2 , w komplementarnych zakresach mierzonych mas. W praktyce oznacza to, że żaden związek nie jest pomijany, ani wykluczany z analizy.

Najczęściej w technice DIA stosuje się podejście pozwalające na sekwencyjne skanowanie, w którym szeroki zakres mas jest dzielony na mniejsze zakresy m/z . Następnie, w każdym oknie wszystkie jony są równocześnie izolowane i poddawane fragmentacji. Dzięki temu można uzyskać łączone widma MS^2 dla wszystkich związków w badanej próbce, co umożliwia bardziej kompleksową analizę jej składu.

Metoda DIA jest szczególnie przydatna w badaniach biologicznych i medycznych, w których próbki są często bardzo złożone pod względem składu molekularnego, a analiza wszystkich obecnych związków jest istotna dla zrozumienia procesów biologicznych, czy diagnostyki choroby.

Podstawowe podejście DIA z selekcją jonów w pełnym zakresie zostało wykorzystane w technice MS^E , w której równocześnie uzyskuje się zarówno widma jonów prekursorowych jak i fragmentacyjne w czasie jednego rozdzielu chromatograficznego, poprzez skanowanie naprzemienne przy wysokich i niskich wartościach energii kolizji [21]. Inne podejście, stosujące wiele zakresów izolacji po raz pierwszy wykorzystano w pracy Venable i wsp. Zastosowano w niej liniową pułpkę jonową (ang. *Linear Ion Trap*) i przeprowadzono sekwencyjną izolację i fragmentację dla 20 okien (każde o szerokości 10 m/z). Pozwoliło to na po-



Rycina 4. Schematyczne przedstawienie metod wyboru szerokości okna: elastyczne (ang. *variable-window DIA*, (vDIA)-MS) oraz *overlapping-window DIA* (oDIA)-MS. W pierwszej metodzie szerokość okna dostosowana jest do częstości występowania peptydów natomiast, w metodzie drugiej, okna izolacji zachodzą na siebie i są demultiplikowane, co redukuje złożoność widm MS^2 .

krycie zakresu 900–1100 m/z oraz kilkukrotnie zwiększyło stosunek sygnału do szumu i umożliwiło porównanie poziomów akumulacji białek dla dwóch stanów rozwojowych *Caenorhabditis elegans* [22]. Jednak pomimo tych zalet, strategia predefiniowanych okien nie była szeroko stosowana ze względu na ograniczenia wynikające z szybkości skanowania oraz niską rozdzielczość spektrometrów mas. Wraz z wprowadzeniem spektrometrów wysokorozdzielczych zaproponowano nową technikę "SWATH MS" (Sciex), która wykorzystuje stałą szerokość okna Q1 (25 Da). Taka szerokość została uznana za równoważną z wielokrotnymi (pięć lub więcej) przejściami dla peptydu w metodzie SRM, bez zakłóceń i dokładnością pomiaru do 10 ppm [23].

Przy doborze metody MS należy zrównoważyć kilka istotnych parametrów, m.in: wpływ złożoności próbki, szerokość okna izolacji oraz pokrycie jak najszerszego zakresu mas. Wynika to z faktu, że poprzez zmniejszanie szerokości okna izolacji prekursorów m/z zostaje zwiększona czułość i precyzja metody, równocześnie zmniejszeniu ulega liczba peptydów reprezentowanych w każdym widmie MS^2 , co w konsekwencji zmniejsza pokrycie prekursorów w całym zakresie m/z . Opracowano dwa podejścia ułatwiające wybór szerokości okna: elastyczne (ang. *variable-window DIA*, (vDIA)-MS) [24] i *overlapping-window DIA* (oDIA)-MS [25]. Pierwsze dostosowuje szerokość okna do częstości występowania peptydów: w obszarach, w których sygnały MS są bogato wykrywane, zmniejsza szerokość okna izolacji i zwiększa szerokość w obszarach gdzie sygnały są mniej intensywne. Natomiast w metodzie oDIA-MS okna izolacji zachodzą na siebie i są demultipleksowane, co redukuje złożoność widm MS^2 . Podejście to poprawia prawie dwukrotnie selektywność wyboru prekursorów bez utraty rozdzielczości w analizie masy, rozdzielczości chromatograficznej, szybkości skanowania i innych kluczowych parametrów akwizycji (Ryc. 4).

NARZĘDZIA BIOINFORMATYCZNE WYKORZYSTYWANE W METODZIE DIA

Dane uzyskane w wyniku prowadzenia analiz DIA, ze względu na kompleksowość i brak przypisania widm MS^2 do konkretnych prekursorów, cechują się znacznie większą złożonością niż dane uzyskane w eksperymencie DDA. Stąd oczywistym staje się konieczność korzystania z zaawansowanych metod bioinformatycznych w celu dekonwolucji i interpretacji uzyskiwanych widm. Identyfikacja białek obecnych w próbce analizowanej metodą DIA odbywa się przede wszystkim na podstawie złożonych analiz opartych o biblioteki widm peptydowych. Mogą one być tworzone bądź to w oparciu o dostępne atlasy peptydowe, bądź być kompilowane na podstawie badań własnych próbek, wykorzystując wyniki wcześniejszej analizy DDA. Biblioteki te katalogują czasy retencji, masy prekursorów, wzory fragmentacji peptydów i znane zakłócenia. Należy tutaj wyraźnie zaznaczyć, że jakość dekonwolucji widm i identyfikacji oraz badania ilościowego w oparciu o DIA zależy w oczywisty sposób od jakości dostępnych bibliotek. Tak więc dobór tych ostatnich i sposób ich przygotowania jest kluczowym aspektem całego eksperymentu.

Zaawansowane programy informatyczne dedykowane analizom DIA, takie jak DIA-Umpire [26], PECAN [27] oraz EncyclopeDIA [28], umożliwiają poprawioną identyfikację oraz analizę ilościową białek i peptydów, w odniesieniu do opisywanych wcześniej analiz DDA. Analizę danych można przeprowadzać również przy wykorzystaniu komercyjnie dostępnych aplikacji takich jak: Scaffold DIA (Proteome Software Inc., Portland OR, USA), Spectronaut, a także PEAKS Studio. Dekonwolucja danych może odbywać się na kilka sposobów np. stosując metodę najmniejszych kwadratów w połączeniu z techniką MSX (ang. *multiplexing*). Wykorzystanie jej jest możliwe jednak jedynie w przypadku spektrometrów mas typu kwadrupol-orbitrap, gdzie izolacja, aktywacja i akumulacja prekursorów może odbywać się równolegle z akwizycją widma MS^2 [25]. Inny algorytm wykorzystywany w metodzie oDIA-MS, pozwala przypisać każdy sygnał z okna o szerokości 20 m/z do właściwego podokna 10 m/z , uzyskując wysokiej jakości piki chromatograficzne prezentujące charakterystyczny kształt wynikający z naprzemiennego skanowania w dwóch podoknach [25].

Biblioteki DIA można tworzyć z pomocą dedykowanego oprogramowania, a wśród dostępnych aplikacji są zarówno rozwiązania komercyjne jak i te z otwartym dostępem. Polecić należy zwłaszcza darmowy program Skyline (obecnie v2.5, Seattle, WA) [29], który nie tylko usprawnia proces przygotowania metod do analiz MS, ale również wspomaga opracowanie bibliotek MS^2 . Umożliwia on przeprowadzenie trawienia białek *in silico*, wspierając dobór peptydów i przejść fragmentacyjnych, a także automatyczne wybieranie filtrów do bibliotek spektralnych. Skyline współpracuje z układami MS większości producentów, takich jak Agilent, Sciex, Thermo Fisher Scientific, Bruker czy Waters [29]. Z innych programów warto wymienić też oprogramowanie DIA-NN [30], czy komercyjny Spectronaut [31]. Wymienione narzędzia bioinformatyczne podsumowano w Tabeli 1.

WYKORZYSTANIE SPEKTROMETRII MOBILNOŚCI W METODZIE DIA (DIA - PASEF)

Analizy DIA można również łączyć z analizami mobilności jonów (ang. *ion mobility spectrometry*, IMS), tak jak to opisano w przypadku analiz DDA (rozdział o DDA-PASEF). Metoda ta, nazwana analogicznie DIA-PASEF, wykorzystuje mobilność jonową, która zwiększa czułość analizy, przyspiesza ją oraz pozwala na zmniejszenie ilości próbki niezbędnej do przeprowadzenia badania. Uzyskiwane w czasie analizy widma są „czystsze”, gdyż metoda ta pozwala na rozróżnienie sygnałów pochodzących od różnych peptydów, przez co nie dochodzi do współfragmentacji. Co ważne, w przypadku analizy niewielkich ilości próbek, obserwowany jest 2–5-krotny wzrost czułości. Zależy on od schematu akwizycji i doboru odpowiedniego ułożenia okien izolacji jonów prekursorowych w wymiarze ruchliwości jonów, co tym samym zwiększa efektywność cyklu pracy [32]. Dodatkowo biorąc pod uwagę fakt, że wiele prekursorów jest dobieranych i fragmentowanych podczas pojedynczego skanowania TIMS, PASEF osiąga ponad dziesięciokrotny wzrost prędkości niż w przypadku DDA, bez utraty czułości, która jest charakterystyczna dla bardzo

Tabela 2. Istotne czynniki przy analizach DIA i DDA.

Czynnik	DIA	DDA
Przygotowanie próbek	Procedura przygotowania próbki jest zazwyczaj kompleksowa, z naciskiem na skuteczną ekstrakcję składowych. Metoda nie ogranicza wyboru analizowanego rodzaju próbki w procesie jej przygotowania.	Bardziej ukierunkowane podejście do przygotowania próbki. Kryteria selekcji jonów prekursorowych są z góry określone, co pomaga w ekstrakcji i przygotowaniu określonych składników molekularnych.
Jonizacja i analiza MS	Spektrometr mas systematycznie fragmentuje jony w predefiniowanych oknach masowych, zapewniając kompleksowe pokrycie widma masowego. Ta systematyczność i powtarzalność jest kluczową cechą wyróżniającą DIA.	DDA obejmuje selektywną jonizację. Spektrometr izoluje i fragmentuje jony prekursorowe w oparciu o określone kryteria, na które często wpływa intensywność lub liczebność jonów. To ukierunkowane podejście pozwala na dogłębną charakterystykę określonych jonów.
Akwizycja danych	Jednoczesna akwizycja w całym zakresie mas. Powstały zestaw danych jest kompleksowy i bezstronny. Podejście to minimalizuje ryzyko pominięcia nieoczekiwanych peptydów lub peptydów o niskiej liczebności.	W DDA proces akwizycji danych jest bardziej selektywny. Prowadzi to do zbioru danych skupiającego się na określonych składnikach molekularnych. Chociaż to ukierunkowane podejście zapewnia szczegółowe informacje na temat wybranych jonów, może powodować utratę danych o mało licznych jonach.
Analiza danych	Generowany jest kompleksowy zestaw danych, który wymaga zaawansowanych narzędzi do analizy. Narzędzia i algorytmy bioinformatyczne zaprojektowane do dekonwolucji i niecelowanej analizy danych są niezbędne do wnioskowania z eksperymentów DIA.	Analiza danych DDA jest zwykle prostsza i koncentruje się na interpretacji danych pochodzących od wyselekcjonowanych jonów. Wyzwania mogą pojawić się w sytuacjach, w których kluczowe znaczenie mają jony o niskiej intensywności, ponieważ mogą zostać przeoczone w procesie selektywnego gromadzenia danych.

szybkich cykli fragmentacji [18]. Dzięki zastosowaniu DIA-PASEF udało się np. zidentyfikować 9 000 białek w próbce komórek HeLa zawierającej 200 ng peptydów, stosując rozdział chromatograficzny trwający zaledwie 93 minuty [33].

Analiza wykorzystująca podejście DIA gwarantuje zatem wysokie pokrycie uzyskiwanych danych, zapewnia wysoką czułość jak również pozwala na względną analizę ilościową badanych białek oraz jest wysoce powtarzalna. Metoda ta umożliwia przeprowadzenie analiz proteomicznych, metabolomicznych i lipidomicznych. Jako przykład wszechstronnej analizy z wykorzystaniem DIA można podać badania Zhang i wsp. którzy przeprowadzili analizę proteomiczną komórek dendrytycznych wyizolowanych z krwi obwodowej oraz analizę metabolomiczną moczu pozyskanego od zdrowych ochotników [24]. Analiza DIA umożliwiła także badanie kinaz i czynników transkrypcyjnych w próbkach tkanek mózgu myszy [34]. Niewątpliwą zaletą DIA jaką jest znacząca liczba identyfikacji białek (>10 000) uzyskana

przy długim rozdziale na kolumnie chromatograficznej (6h gradient) przedstawiono w pracy Muntel i wsp. [35].

PODSUMOWANIE

W przeszłości metody DIA były wykorzystywane głównie do jakościowych badań przesiewowych, jednakże rozwój spektrometrii mas pozwolił na ogromny postęp w tej dziedzinie znacznie poprawiając wydajność tej metody. Oczywiście rozwój nie ominął również technik opartych o podejście bardziej klasyczne jakim jest DDA. Spektrometry są coraz dokładniejsze, bardziej czułe, a przede wszystkim oferują większą częstotliwość skanowania. Przekłada się to na poprawę skuteczności analiz opartych o dane, co przybliża tą metodę do skuteczności DIA. DIA wykazuje większą czułość niż DDA, szczególnie w przypadku, kiedy w analizowanej próbce znajduje się niewielka liczba jonów ulegających jednoczesnej elucji z kolumny. DIA pozwala również na detekcję większej liczby jonów oraz z powodzeniem jest

Tabela 3. Podsumowanie zalet i wad metod DDA i DIA.

	DDA	DIA
Zalety	Łatwiejsza konfiguracja metody i tańsza analiza Mniejsze zapotrzebowanie na zasoby obliczeniowe Możliwa względna kwantyfikacja peptydów między próbkami przy użyciu różnych metod znakowania chemicznego (np. SILAC lub iTRAQ/TMT)	Nadaje się doskonale do analiz niecelowanych Uwzględnia wszystkie peptydy Umożliwia ilościowe oznaczanie białek w złożonych mieszaninach w szerokim zakresie dynamicznym Zapewnia wyższą precyzję i lepszą odtwarzalność Uzyskane dane mogą być analizowane retrospektywnie za pomocą ulepszonych algorytmu
Wady	Dobór prekursorów wybranych do fragmentacji odbywa się w czasie rzeczywistym Biblioteki widm DDA mogą zawierać „luki”, gdy peptydy zostaną zidentyfikowane tylko w niektórych próbkach Niższa precyzja i odtwarzalność Metoda niewystarczająca do analizy peptydów o niskiej abundancji	Wysokie zapotrzebowanie na zasoby obliczeniowe wynikające z ilości generowanych danych Wymagająca analiza danych wynikająca z multipleksowego charakteru widm MS ² Konieczność ulepszenia narzędzi i oprogramowania wykorzystywanego do dekonwolucji uzyskanych widm Brak możliwości przesłania jonów fragmentacyjnych w widmach MS ² do ich prekursorów Droższa analiza

używana do badania modyfikacji potranslacyjnych białek. DDA z kolei częściej stosuje się w analizie immunopeptydów. Istotnymi czynnikami wpływającymi na dobór podejścia analitycznego będą potencjalne sposoby przygotowania próbki, a także sposób przeprowadzenia analizy oraz tryb opracowania uzyskanych danych (Tabela 2.) Wady i zalety obu podejść analitycznych zostały zebrane w Tabeli 3.

Wielu badaczy uważa, że ze względu na ciągle udoskonalanie algorytmów i oprogramowania do dekonwolucji złożonych danych DIA, obie metody ostatecznie połączą się w jedną, hybrydową. Teza ta wydaje się powoli urzeczywistniać biorąc pod uwagę propozycje zastosowania metod takich jak DDIA [36], czy DaDIA [37]. Łączą one oba podejścia wykorzystując jednocześnie narzędzia głębokiego uczenia (ang. *deep learning*), w celu usprawnienia analizy danych.

Podsumowując, ze względu na łatwość konfiguracji metod i analizy, podejście DDA nadal jest lepszym rozwiązaniem, szczególnie jeśli eksperymentator nie ma doświadczenia w tandemowej spektrometrii mas oraz biegłości w prowadzeniu przesiewowych analiz biologicznych. Z drugiej strony DIA daje lepsze rezultaty, wymaga jednak większego doświadczenia, bezstronnego i głębszego spojrzenia na profil badanych próbek, szczególnie, gdy pochodzą one ze słabo zbadanego organizmu lub typu komórek. Tak więc, DIA powinno być wyborem w przypadku dysponowania odpowiednio zaawansowanym spektrometrem przez doświadczonego operatora oraz zawsze wtedy gdy dysponujemy możliwością korzystania z bogatej i wszechstronnej biblioteki widm pozwalającej na skuteczną dekonwolucję uzyskiwanych danych.

PIŚMIENNICTWO

1. Meissner F, Geddes-McAlister J, Mann M, Bantscheff M (2022) The emerging role of mass spectrometry-based proteomics in drug discovery. *Nat Rev Drug Discov* 9(9):637-654
2. Niu L, Thiele M, Geyer PE, i in. (2022) Noninvasive proteomic biomarkers for alcohol-related liver disease. *Nat Med* 28(6):1277-1287
3. Wang H, Lim KP, Kong W, Gao H, Wong BJH, Phua SX, Guo T, Goh WWB (2023) MultiPro: DDA-PASEF and diaPASEF acquired cell line proteomic datasets with deliberate batch effects. *Sci Data* 10(1):1-11
4. Warinner C, Korzow Richter K, Collins MJ (2022) Paleoproteomics. *Chem Rev* 122(16):13401-13446
5. Defossez E, Bourquin J, von Reuss S, Rasmann S, Glauser G (2023) Eight key rules for successful data-dependent acquisition in mass spectrometry-based metabolomics. *Mass Spectrom Rev* 42(1):131-143
6. Li J, Smith LS, Zhu HJ (2021) Data-independent acquisition (DIA): An emerging proteomics technology for analysis of drug-metabolizing enzymes and transporters. *Drug Discov Today Technol* 39:49-56
7. Pezzatti J, González-Ruiz V, Boccard J, Guillaume D, Rudaz S (2020) Evaluation of different tandem ms acquisition modes to support metabolite annotation in human plasma using ultra high-performance liquid chromatography high-resolution mass spectrometry for untargeted metabolomics. *Metabolites* 10(11):1-17
8. Hilaire PB Saint, Rousseau K, Seyer A, Dechaumet S, Damont A, Junot C, Fenaille F (2020) Comparative evaluation of data dependent and data independent acquisition workflows implemented on an orbitrap fusion for untargeted metabolomics. *Metabolites* 10(4):158
9. Kalli A, Smith GT, Sweredoski MJ, Hess S (2013) Evaluation and optimization of mass spectrometric settings during data-dependent acquisition mode: Focus on LTQ-orbitrap mass analyzers. *J Proteome Res* 12(7):3071-3086

10. Larson TS, Worthington CD, Verber MD, Keating JE, Lockett MR, Glish GL (2023) DiffN Selection of Tandem Mass Spectrometry Precursors. *Anal Chem* 95(25):9581-9588
11. Hoopmann MR, Merrihew GE, Von Haller PD, MacCoss MJ (2009) Post analysis data acquisition for the Iterative MS/MS sampling of proteomics mixtures. *J Proteome Res* 8(4):1870-1875
12. Révész Á, Hevér H, Steckel A, Schlosser G, Szabó D, Vékey K, Drahos L (2023) Collision energies: Optimization strategies for bottom-up proteomics. *Mass Spectrom Rev* 42(4):1261-1299
13. Eng JK, Fischer B, Grossmann J, MacCoss MJ (2008) A fast SEQUEST cross correlation algorithm. *J Proteome Res* 7(10):4598-4602
14. Elias JE, Haas W, Faherty BK, Gygi SP (2005) Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. *Nat Methods* 2(9):667-675
15. Tyanova S, Temu T, Cox J (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* 11(12):2301-2319
16. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen J V, Mann M (2011) Andromeda: A peptide search engine integrated into the MaxQuant environment. *J Proteome Res* 10(4):1794-1805
17. Wolters DA, Washburn MP, Yates JR (2001) An automated multidimensional protein identification technology for shotgun proteomics. *Anal Chem* 73(23):5683-5690
18. Meier F, Beck S, Grassl N, Lubeck M, Park MA, Raether O, Mann M (2015) Parallel Accumulation-Serial Fragmentation (PASEF): Multiplying Sequencing Speed and Sensitivity by Synchronized Scans in a Trapped Ion Mobility Device. *J Proteome Res* 14(12):5378–5387
19. Wang Y, Guan ZY, Shi SW, i in. (2024) Pick-up single-cell proteomic analysis for quantifying up to 3000 proteins in a Mammalian cell. *Nat Commun* 15(1):1279
20. Wojakowska A, Marczak L, Zeman M, Chekan M, Zembala-Nożyńska E, Polanski K, Strugała A, Widlak P, Pietrowska M (2024) Proteomic and metabolomic signatures of rectal tumor discriminate patients with different responses to preoperative radiotherapy. *Front Oncol* 14(February):1-11
21. Bilbao A, Varesio E, Luban J, Strambio-De-Castillia C, Hopfgartner G, Müller M, Lisacek F (2015) Processing strategies and software solutions for data-independent acquisition in mass spectrometry. *Proteomics* 15(5-6):964-980
22. Venable JD, Dong MQ, Wohlschlegel J, Dillin A, Yates JR (2004) Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nat Methods* 1(1):39-45
23. Gillet LC, Navarro P, Tate S, Röst H, Selevsek N, Reiter L, Bonner R, Aebersold R (2012) Targeted Data Extraction of the MS/MS Spectra Generated by Data-independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis. *Mol Cell Proteomics* 11(6):O111.016717
24. Zhang Y, Bilbao A, Bruderer T, Luban J, Strambio-De-Castillia C, Lisacek F, Hopfgartner G, Varesio E (2015) The Use of Variable Q1 Isolation Windows Improves Selectivity in LC-SWATH-MS Acquisition. *J Proteome Res* 14(10):4359-4371
25. Amodei D, Egertson J, MacLean BX, Johnson R, Merrihew GE, Keller A, Marsh D, Vitek O, Mallick P, MacCoss MJ (2019) Improving Precursor Selectivity in Data-Independent Acquisition Using Overlapping Windows. *J Am Soc Mass Spectrom* 30(4):669-684
26. Tsou CC, Avtonomov D, Larsen B, Tucholska M, Choi H, Gingras AC, Nesvizhskii AI (2015) DIA-Umpire: comprehensive computational framework for data-independent acquisition proteomics. *Nat Methods* 12(3):258-264
27. Ting YS, Egertson JD, Bollinger JG, Searle BC, Payne SH, Noble WS, MacCoss MJ (2017) PECAN: library-free peptide detection for data-independent acquisition tandem mass spectrometry data. *Nat Methods* 14(9):903-908
28. Searle BC, Pino LK, Egertson JD, Ting YS, Lawrence RT, MacLean BX, Villén J, MacCoss MJ (2018) Chromatogram libraries improve peptide

- de detection and quantification by data independent acquisition mass spectrometry. *Nat Commun* 9(1):5128
29. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* 26(7):966-968
 30. Demichev V, Messner CB, Vernardis SI, Lilley KS, Ralser M (2020) DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods* 17(1):41-44
 31. Barkovits K, Pacharra S, Pfeiffer K, Steinbach S, Eisenacher M, Marcus K, Uszkoreit J (2020) Reproducibility, Specificity and Accuracy of Relative Quantification Using Spectral Library-based Data-independent Acquisition. *Mol Cell Proteomics* 19(1):181-197
 32. Meier F, Brunner AD, Frank M, i in. (2020) diaPASEF: parallel accumulation–serial fragmentation combined with data-independent acquisition. *Nat Methods* 17(12):1229-1236
 33. Demichev V, Szyrwiel L, Yu F, i in. (2022) dia-PASEF data analysis using FragPipe and DIA-NN for deep proteomics of low sample amounts. *Nat Commun* 13(1):3944
 34. Kawashima Y, Watanabe E, Umeyama T, Nakajima D, Hattori M, Honda K, Ohara O (2019) Optimization of Data-Independent Acquisition Mass Spectrometry for Deep and Highly Sensitive Proteomic Analysis. *Int J Mol Sci* 20(23):5932
 35. Muntel J, Gandhi T, Verbeke L, Bernhardt OM, Treiber T, Bruderer R, Reiter L. (2019) Surpassing 10 000 identified and quantified proteins in a single run by optimizing current LC-MS instrumentation and data analysis strategy. *Mol Omi* 15(5):348-360
 36. Guan S, Taylor PP, Han Z, Moran MF, Ma B (2020) Data Dependent-Independent Acquisition (DDIA) Proteomics. *J Proteome Res* 19(8):3230-3237
 37. Guo J, Shen S, Xing S, Huan T (2021) DaDIA: Hybridizing Data-Dependent and Data-Independent Acquisition Modes for Generating High-Quality Metabolomic Data. *Anal Chem* 93(4):2669-2677

Data-driven mass spectrometry methods (DDA) and data-independent methods (DIA) used in the analysis of biological material

Marta Nólka-Szaszner¹✉, Aleksander Strugała²✉, Łukasz Marczak²

¹Department of Biomedical Proteomics, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań

²Laboratory of Mass Spectrometry, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań

✉corresponding author: mnolka@ibch.poznan.pl, astrugala@ibch.poznan.pl

Keywords: mass spectrometry, DDA, DIA, proteomic analysis, metabolomic analysis

ABSTRACT

Mass spectrometry is an important tool in proteomic, metabolomic and lipidomic analysis. To fully use its potential, it is crucial to select and configure the appropriate analytical approach. For untargeted research, there are two main strategies available: data-dependent analysis (DDA) and data-independent analysis (DIA). Both methods differ in the way the analysis is carried out and in the degree of coverage of the obtained data, which is why each of them can be used in various types of research. The DDA method is based on continuous scanning of the analyzed ions, as a result of which the precursors with the highest intensity are fragmented in the MS² mode. On the other hand, DIA, due to the use of combined ranges of precursor ion isolation, allows for a deeper analysis of the analyzed compounds. Both approaches also have modifications that improve their operation and enable obtaining more valuable data. Methods combining both techniques are also appearing on the horizon, such as DDIA, which uses the advantages of both methods, opening new analytical possibilities.

