

Odkrycie fragmentacji indukowanej wychwyceniem elektronu (ETD - z ang. Electron Transfer Dissociation) znacząco poszerzyło repertuar technik analitycznych dostępnych w proteomice. Celem niniejszego projektu jest poznanie dynamiki tych reakcji w spektrometrze mas poprzez sporządzenie ich uproszczonego statystycznego opisu. Nasze badania oprzyły o dane do wiadczalne pozyskane przez specjalistów z Centre For Proteomics, z Antwerp w Belgii. Zdobyte wyniki przedstawimy w formie praktycznej wiedzy dla chemików zajmujących się badaniami proteomicznymi.

W naszych badaniach skupimy się głównie na białkach, jednak prezentowana metodologia może uogólnić się na inne struktury. Białka pełnią przeto rolę w każdej komórce: katalizują reakcje metaboliczne, pełnią dozór nad prawidłową replikacją DNA oraz przekazują informacje w ramach sieci sygnałowych. Poznanie zestawu białek aktywnych w komórkach wirusów jest bezpo- rednio ze zrozumieniem ich działania, a przede wszystkim - z określeniem czynników patogennych, gdy przestają działać. Dlatego identyfikacja białek oraz pomiar ich stężenia w komórkach jest ważnym narzędziem w badaniach podstawowych biologii oraz w praktyce klinicznej. Spektrometria mas w pełni przyczyniła się do zwiększenia przepustowości badań proteomicznych poprzez umożliwienie jednoczesnego pomiaru mas i liczebności tysięcy molekuł tworzących analizowane próbki. Proste ważenie molekuł nie zawsze skutkuje jednoznacznym rozpoznaniem biomolekuł. Problem ten rozwiązuje się przez fragmentację biomolekuł zawartych w próbce. Jedną z nowszych, acz już rozpowszechnionych technik umożliwiających fragmentację wewnątrz instrumentu (tzw. podejście top-down) jest ETD. Szczegółowy mechanizm rozpadu nadal jest przedmiotem żarliwych debat.

Nasze badanie nie rozstrzygnie tego problemu, nie da się bowiem tego uczynić tylko na podstawie wyników badań spektrometrycznych. Dlatego też skupimy się na innym problemie: wiadomo, że fragmentacja nie jest jedynym zjawiskiem wynikającym z transferu elektronu. Mogą być inne reakcje, które skutkują obserwowalnymi wynikami. Do tej pory nie potrafiono stwierdzić, które reakcje zachodzą częściej od innych. Naszym zadaniem jest zrozumienie dynamiki tych zjawisk na podstawie eksperymentalnie pozyskanych danych.

Sporządzenie statystycznego opisu reakcji dostarczy praktycznych wskazówek dla chemików chcących optymalizować swoje protokoły badań oraz umożliwi pełniejsze porównanie maszyn. Upatrywane korzyści wynikają z lepszego zrozumienia następujących czasowych reakcji oraz możliwości monitorowania tylko niewielkiej liczby parametrów przy porównaniach różnych eksperymentów.

Obrana przez nas strategia modelowania zakłada pominięcie skomplikowanych kulomboskich interakcji przestrzennych oraz uzależnienie ewolucji układu jako całości od liczebności podgrup czy steczek go tworzących. Naturalne jest rozpatrzenie podgrup czy stek o tym samym ładunku. Zakładamy, że w ramach badanego zjawiska statystycznie najpierw powinny odreagować stężenie wy- szym ładunku, oraz że w danym momencie zachodzi co najwyżej jeden rodzaj reakcji, a sam rodzaj reakcji, której zostanie poddana molekuła jest losowy. Na początku naszej współpracy z Centre For Proteomics, specjaliści chemicy z tej grupy sporządzili list trzech reakcji, które mogłyby dostatecznie dobrze wytłumaczyć obserwowane dane. Po stworzeniu odpowiednich narzędzi identyfikujących fragmenty (projekt MassTodon), i przebadaniu ponad setki spektr masowych jesteśmy pewni, że list ten uzupełnić trzeba co najmniej o czwartą reakcję.

Chcemy teraz dalej ulepszyć narzędzia o wykorzystanie opracowanego przez nas metody wyliczenia zbioru krytycznego dokładnej wiązki izotopowej - algorytmu IsoStar. Wiązka izotopowa to zbiór danych zaobserwowanych w spektrometrze mas, który może być utworzony z jednym rodzajem czy steczki. W przyrodzie spotykamy izotopy, czyli atomy gnie dące w jądrze różną liczbę neutronów. To wpływa na ich masę, a więc także i na masę interesujących nas czy steczek chemicznych. Tym samym, przy pomiarze spektrometrycznym, kilka pików (zaobserwowanych par masa-intensywność) może odpowiadać jednemu i temu samemu rodzajowi czy steczki. Pików tych może być bardzo dużo, ale nie wszystkie są równie prawdopodobne. IsoStar wykorzystuje własność gromadzenia się prawdopodobieństwa tych rozkładów w pewnych obszarach i je eksploruje. Narzędzie to pozwoli nam wyczytać z danych informacje na temat liczebności obserwowanych produktów rozważanych przez nas reakcji.

Uzyskane tym sposobem oszacowania chcemy wykorzystać w bayesowskiej analizie rozkładu a posteriori parametrów opisujących dynamikę reakcji. Zakładamy bowiem, że stery naszego układu te są do pewnego stopnia losowe. Wyznaczenie rozkładu a posteriori udzieli odpowiedzi na pytanie, jak bardzo losowe. Aby ten rozkład poznać, przeprowadzimy ekstensywne symulacje komputerowe.

Zastosowana metoda badawcza ma na celu uchwycenie kwintesencji zjawisk spowodowanych wychwyceniem elektronu, zachodzących wewnątrz spektrometru mas. Dostarczy to informacji niezbędnych chemikom, którzy chcą efektywnie pozyskiwać informacje ze swoich maszyn. Nadto, nasz projekt doprowadzi do rozwoju szybkich algorytmów służących identyfikacji biomolekuł w danych z wysokoprzepustowych instrumentów, zarzucających nas danymi. W tym upatrujemy przyszłości metod obliczeniowych w proteomice i tym chcemy się zająć.