

Enzymy proteolityczne, czyli proteazy, odgrywają kluczową rolę w regulacji istotnych procesów biologicznych poprzez selektywne cięcie wiązań peptydowych w substratach białkowych i peptydowych. Są one ważne w procesach aktywacji i degradacji białek, trawieniu, apoptozie i replikacji patogenów. Złe funkcjonowanie proteaz jest związane z wieloma chorobami, co czyni je ważnymi celami dla nowych leków. Dokładna identyfikacja miejsc cięcia substratów jest kluczowa dla zrozumienia selektywności enzymu oraz projektowania nowych inhibitorów i rozwoju nowych strategii terapeutycznych.

Stosowane obecnie metody obliczeniowe do identyfikacji miejsc cięcia opierają się głównie na danych sekwencyjnych. Chociaż metody te mogą być skuteczne, mają też poważne ograniczenia, nie uwzględniają strukturalnego i dynamicznego kontekstu interakcji enzym-substrat. W rezultacie nasza wiedza na temat mechanizmów działania i selektywności enzymów pozostaje niekompletna, a projektowanie terapii ukierunkowanych na proteazy jest utrudnione.

Podczas realizacji projektu powstanie nowa metoda do identyfikacji miejsc cięcia peptydów przez enzymy, integrująca modelowanie strukturalne kompleksów enzym-substrat z oceną struktury wspartą przez algorytmy uczenia maszynowego. Wykorzystamy zaawansowane narzędzia do przewidywania struktury białek (np. AlphaFold-Multimer, AlphaFold3) i symulacje dokowania peptydów przy użyciu gruboziarnistego modelu CABS. Modele kompleksów enzym-substrat zostaną następnie zoptymalizowane z wykorzystaniem metod mechaniki i dynamiki molekularnej w celu poprawy ich geometrii.

Opracowana zostanie funkcja oceny przy użyciu metod uczenia maszynowego, która będzie mogła rozróżniać odmienne sposoby wiązania, charakterystyczne dla ciętych lub nieciętych peptydów. Do identyfikacji miejsc cięcia zostaną wykorzystane dwie uzupełniające się strategie: (1) podejście oparte na strukturze kompleksu, wykorzystujące kryteria geometryczne opisujące wzajemne położenie atomów wiązanego peptydu i aminokwasów z miejsca aktywnego enzymu; oraz (2) funkcja oceny wytrenowana w celu rozpoznawania cech wspólnych dla oddziaływania enzym-peptyd, charakterystycznych dla ciętych substratów.

Wstępne badania potwierdziły skuteczność tego podejścia dla układów enzym-substrat obejmujących pepsynę, reninę i proteazę HIV-1. W przypadku tych enzymów miejsca cięcia łańcucha peptydowego wyznaczone z wykorzystaniem metod dokowania gruboziarnistego i analizy strukturalnej, ściśle odpowiadały danym z eksperymentów. Ponadto nasze wyniki wykazały również, iż cięte i niecięte peptydy posiadają odmienne sposoby wiązania się do enzymu, co potwierdza możliwość wykrzesania metod opartych na uczeniu maszynowym w tym kontekście.

Proponowane podejście jest nowatorskie i odmienne od dotychczas stosowanych metod opartych wyłącznie na sekwencji. Metoda umożliwi przewidywanie miejsc cięcia w bardziej złożonych układach molekularnych i dostarczy wglądu w mechanizmy rozpoznawania substratu przez enzym. Ponadto, umożliwi identyfikację nowych motywów strukturalnych substratów, wyjaśni efekty kooperacji między wieloma resztami aminokwasowymi, oraz ujawni czynniki konformacyjne wpływające na specyficzność enzymatyczną. Poprzez integrację modelowania strukturalnego, symulacji dokowania, i metod uczenia maszynowego, projekt ten przyczyni się do postępu w dziedzinie enzymologii obliczeniowej. W rezultacie powstanie wydajne narzędzie dla badaczy akademickich i przemysłu, do identyfikacji miejsc cięcia substratów peptydowych, wspomagające projektowanie nowych inhibitorów, co będzie również prowadziło do lepszego zrozumienia biologii proteaz.