

Białka są podstawowym budulcem żywych organizmów. Nazwa tych molekuł pochodzi z greckiego πρωτεος, mającego znaczenie "pierwotny". Białka biorą udział w większości procesów w żywych komórkach. Funkcjonowanie wszystkich żywych organizmów w dużej mierze zależy od tego, że budujące je białka przyjmują podczas procesu fałdowania w warunkach fizjologicznych ściśle określoną strukturę trójwymiarową. Białka są zbudowane z długich łańcuchów reszt aminokwasowych, poszczególne białka różnią się sekwencją następujących po sobie aminokwasów. Każde białko posiada na ogół dobrze zdefiniowaną strukturę przestrzenną która, wraz z jego mobilnością, określa funkcję. Białka można zatem uważać za nanomaszyny. Wiedza o strukturze i ruchach białka jest niezbędna do zrozumienia jego funkcji. Błędy w sekwencji białek, znane jako mutacje, mogą spowodować błędy w funkcjonowaniu (złe zwinięcie lub agregacja) i mogą drastycznie zaburzyć funkcjonowanie całego organizmu, prowadząc do poważnych schorzeń takich, jak choroby nowotworowe czy amyloidozy (np. choroba Alzheimera).

Obecnie jedynie ok. 7% struktur białek o znanej sekwencji jest rozwiązanych metodami eksperymentalnymi. Z uwagi na ogromny przyrost liczby nowo poznawanych sekwencji białek różnica między liczbą sekwencji a liczbą znanych struktur pogłębia się. Powszechnie wykorzystywane eksperymentalne metody określania struktur białek, przede wszystkim krystalografia rentgenowska czy magnetyczny rezonans jądrowy wymagają długiego czasu oraz ogromnego nakładu pracy i kosztów. Ponadto, metody eksperymentalne są w stanie dostarczyć jedynie fragmentarycznych informacji o dynamice białek. Dlatego metody teoretyczne stały się powszechnym narzędziem przewidywania struktury i symulacji dynamiki białek. Obecnie najskuteczniejsze są tzw. metody bioinformatyczne, oparte na wykorzystaniu informacji ze strukturalnych baz danych. Przewidywana struktura jest konstruowana na podstawie podobieństwa do struktur z baz danych. Takie podejście jest uzasadnione ponieważ białka o podobnej sekwencji są związane ewolucyjnie i pełniły podobne funkcje w praorganizmach. Dlatego ich struktura powinna być podobna. To podejście można porównać do porównywania wyrazów podobnie brzmiących lub zapisanych w językach tej samej rodziny, np. 'dom' po polsku, 'dum' po czesku. Ta metoda zawodzi jednak gdy sekwencje są zbyt słabo do siebie podobne. Można po porównać do wyrazu 'pysk' brzmiącego tak samo ale mającego zdecydowanie inne znaczenie w języku polskim i czeskim. Z drugiej strony, metody oparte na fizyce oddziaływań są innym podejściem do przewidywania struktury białek i są one niezależne od baz danych. W metodach tych tworzy się model matematyczny oddziaływań fizycznych (np. odpychania i przyciągania między ładunkami elektrycznymi) zwany *polem siłowym*. Model ten jest używany do symulacji dynamiki, w których ruch białka jest opisany prawami Newtona. Jednakowoż metody te wymagają znacznie dłuższego czasu obliczeń niż metody bioinformatyczne i są niezwykle czułe na niedokładności modelowania oddziaływań.

Najlepsza metoda przewidywania struktury i przewidywania dynamiki białek prawdopodobnie powinna łączyć dwa wymienione powyżej podejścia. Część białka o sekwencji podobnej do białek o znanej strukturze powinna być modelowana metodami bioinformatycznymi a pozostała metodami fizycznymi. Takie podejście zaczęliśmy opracowywać w ramach poprzedniego projektu, wprowadzając informację z modelowania bioinformatycznego do modelu UNRES. W przeciwieństwie do pól siłowych, które używają reprezentacji pełnoatomowej, model UNRES jest wysoce uproszczony i każdy aminokwas jest reprezentowany tylko przez dwa centra. Pomimo tego uproszczenia pole UNRES jest ściśle oparte na fizyce oddziaływań. Uproszczona reprezentacja umożliwi nam prowadzenie symulacji w czasie ok. 1000-10000 krótszym niż symulacje pełnoatomowe. W ramach wnioskowanego projektu planujemy poprawę dokładności przewidywania struktury białek i rozszerzenie metody o dodatkowe funkcje: badanie przejść konformacyjnych w dużych białkach, w oparciu o informacje pomocnicze z eksperymentu lub modelowania struktury w oparciu o bazy danych oraz o modelowanie struktur białek i ich kompleksów w oparciu o niewystarczające do określenia struktury dane doświadczalne. Przewidujemy, że utworzony przez nas pakiet oprogramowania umożliwi badanie struktury i dynamiki elementów maszynierii komórkowej (enzymy oddechowe, białka motoryczne, itp.) i znajdzie zastosowanie w naukach biofizycznych i naukach o zdrowiu. Pakiet będzie udostępniony dla środowiska naukowego ze strony programu UNRES ([www.unres.pl](http://www.unres.pl)) oraz poprzez infrastrukturę PL-GRID.