

Modelowanie molekularne kompleksów kinazy białkowej PrkC z domniemanymi endogennymi celami aktywności fosforylacyjnej.

Marcin Augustyniak

2017

Streszczenie

Doniesienia z eksperymentów przeprowadzonych *in vivo* lub *in vitro* wskazują, że CpgA (YolQ), cyklicznie permutowana GTPaza, jest endogennym substratem serynowo-treoninowej kinazy białkowej PrkC z *B. subtilis* [1, 4, 13]. PrkC jest białkiem przezbłonowym, z cytoplazmatyczną domeną katalityczną i zewnątrzkomórkową domeną receptorową, uczestniczącym w procesach sporulacji i formowania biofilmu. Struktura i funkcja PrkC wyazuje homologię do receptora TGF- β [10, 11].

Analiza moich wstępnych modeli kompleksów PrkCc \cdot Mg²⁺ \cdot ATP \cdot CpgA wskazała, że CpgA może być fosforylowane przez PrkC na innych resztach aminokwasowych niż dotychczas opublikowane w literaturze, tzn. innych niż Thr166 [13], Thr192 lub Ser226 [1]. Dominującym miejscem phosphorylacji na podstawie moich wstępnych modeli miała być Thr205. Powyższa rozbieżność wyzwoliła oczywistą ciekawość dlaczego i w jaki sposób wyniki molekularnego modelowania obliczeniowego mogą tak odbiegać od doniesień z eksperymentów “mokrych”. Analiza pierwszych modeli wskazała również na kilka reszt tyrozynowych CpgA, zwłaszcza Tyr240 lub Tyr298, jako potencjalne miejsca fosforylacji, również przez PrkC.

Klasyczne symulacje dynamiki molekularnej (cMD - ang. classical molecular dynamics) kompleksów PrkCc \cdot ATP \cdot CpgA lub PrkCc \cdot Mg²⁺ \cdot ATP \cdot CpgA pozwoliły na wgląd w dynamikę obu protomerów, jak również zakwestionowały osiągalność uzyskania przekonywujących modeli PrkCc \cdot Mg²⁺ \cdot ATP \cdot CpgA w początkowo zakładany sposób, tzn. wykorzystując sztywne dokowanie z następującą obróbką wyselekcjonowanych, najbardziej obiecujących modeli za pomocą klasycznych symulacji dynamiki molekularnej. W trakcie obu symulacji cMD, o długości 5 ns każda, relacje przestrzenne hydroksylogowego atomu tlenu Thr205 w CpgA i atomu γ -fosforu ATP oraz kluczowych reszt aminokwasowych PrkC nie ewoluowały w kierunku aranżacji umożliwiającej transfer grupy γ -fosforanowej w zakładanym, typowym dla większości eukariotycznych kinaz białkowych, bezpośrednim mechanizmie o liniowym uszeregowaniu atomu γ -fosforu oraz atomów - docelowego i opuszczającego - tlenu, z odwróceniem chiralności atomu fosforu. Obie symulacje były jednak zgodne co do niezbędności kationu Mg²⁺ dla aktywności enzymatycznej kinaz białkowych.

W trakcie symulacji kompleksu PrkCc · ATP · CpgA, tzn. bez kationu Mg^{2+} tworzącego kompleks z ATP, relacje przestrzenne w obrębie szczeliny katalitycznej PrkC uległy istotnej reorganizacji. Rozpoznałem również możliwość zastosowania giętkiego dokowania białek wykorzystującego LMOD [6–8] do modelowanego systemu, z uwagi jednak na wielkość systemu oraz koszt obliczeniowy opartej na LMOD procedury giętkiego dokowania kontynuowałem poszukiwania bardziej wydajnych metod, rezerwując LMOD, podobnie jak classical molecular dynamics (cMD), do końcowej obróbki wyselekcjonowanych struktur.

Większość dostępnych programów lub algorytmów służących do dokowania białek przypomina, o ile nie naśladuje model ‘indukowanego dopasowania’ [9].

Teoria krajobrazu energetycznego (ang. energy landscape theory) [5, 15] i oparty na niej model selekcji konformacyjnej i przesunięć populacyjnych rozpoznania molekularnego i oddziaływań międzybiałkowych postulują, że wchodzące w interakcje cząsteczki białek wywodzą się z już obecnych w roztworze populacji, nie koniecznie z tych najkorzystniejszych energetycznie, z natępnym, po ewentualnym utworzeniu kompleksu, przesunięciem między populacjami (ponownym osiągnięciem równowagi) [2].

Zastosowanie przyspieszonej dynamiki molekularnej (aMD, ang. accelerated molecular dynamics) uczyniło próbkowanie przestrzeni konformacyjnej cząsteczek białek osiągalnym, w niemalże wyczerpującym stopniu, za wyjściową przyjmując konformację niskoenergetyczną, bliską naturalnie występującej w roztworze, np. strukturę uzyskaną w oparciu o dane krystalograficzne [12]. Postęp w projektowaniu układów GPU oraz oprogramowania wykorzystującego możliwości obliczeniowe takich układów sprawił, że przeszukiwanie przestrzeni konformacyjnej makrocząsteczek stało się również dostępne na komputerach klasy stacji roboczych wyposażonych w nowoczesne karty GPU [14].

W pracy tej proponuję protokół przewidywania i modelowania kompleksów białkowych wykorzystujący model selekcji konformacyjnej i przesunięć populacyjnych rozpoznania molekularnego i oddziaływań międzybiałkowych [2] poprzez (i) próbkowanie przestrzeni konformacyjnej protomerów oddzielnie, (ii) wyznaczenie ich rozróżnialnych konformacji, (iii) dokowanie kombinacji tak uzyskanych konformacji, (iv) szczegółową obróbkę i analizę wybranych struktur.

Zastosowałem powyższe podejście do PrkCc i CpgA poprzez próbkowanie ich przestrzeni konformacyjnej oddzielnie za pomocą symulacji accelerated molecular dynamics (aMD), które przeprowadziłem wykorzystując implementację programu pmemd dla układów GPU (pmemd.cuda) wchodzącego w skład pakietu Amber16 [3]. Konformacje obu protomerów wyznaczyłem poprzez analizę klastrową ich trajektorii. Utworzyłem struktury kompleksów poprzez sztywne dokowanie uzyskanych konformacji w możliwych kombinacjach, przeszukując dane w poszukiwaniu możliwie optymalnych relacji przestrzennych substratu, kompleksu $Mg^{2+} \cdot ATP$, oraz kluczowych reszt enzymu. Jako fundamentalne kryterium oceny i selekcji kompleksów przyjąłem ograniczenia geometryczne podyktowane typowym dla większości eukariotycznych kinaz białkowych mechanizmem transferu grupy fosforanowej, tzn. dysocjacyjnym, bezpośrednim, o liniowym uszeregowaniu atomów fosforu oraz obu atomów tlenu (opuszczającego i wchodzącego), z odwróceniem chiralności atomu fosforu przenoszonego ugrupowania fosforanowego.