



Dr hab. Stanisław Ołdziej
Pracownia Struktury Biopolimerów
Międzyuczelniany Wydział Biotechnologii
Uniwersytetu Gdańskiego i Gdańskiego Uniwersytetu Medycznego
ul. Antoniego Abrahama 58
80-307 Gdańsk
e-mail: stan@biotech.ug.edu.pl

Gdańsk 19.02.2018

Recenzja rozprawy doktorskiej zatytułowanej: „*Molecular modeling of complexes of PrkC protein kinase with possible endogenous targets of phosphorylation activity*” przedstawianej przez mgra Marcina Augustyniaka. Promotorem rozprawy jest dr hab. Rajmund Kaźmierkiewicz prof. UG, a kopromotorem prof. dr hab. Michał Obuchowski.

Rozprawa napisana jest w języku angielskim i ma klasyczny dla tego rodzaju prac układ poszczególnych części: streszczenie (w języku polskim i angielskim); wstęp, cel pracy, materiały i metody, wyniki, dyskusja, konkluzje i bibliografia. Praca została wydrukowana na 123 stronach. Spis treści, spisy tabel i rysunków, spis skrótów zajmuje 10 stron, bibliografia obejmuje 239 pozycje i zajmuje 23 strony tekstu, na pozostałych 90 stronach Autor umieścił aż 50 rysunków i 20 tabel, co nie pozostawia zbyt wiele miejsca na tekst. Uwagi dotyczące strony edytorskiej rozprawy zostaną przedstawione w dalszej części recenzji (patrz: Pozostałe uwagi).

Przedstawiona do recenzji rozprawa dotyczy próby teoretycznego przewidywania miejsc fosforylacji białka CpgA przez kinazę PrkC, oba białka pochodzą z bakterii *Bacillus subtilis*). Aktywność kinazy PrkC obejmuje szerokie spektrum substratów (patrz Tabela 1.2 strona 23) i związanych z tymi substratami procesów biochemicznych. Drugie z badanych w rozprawie białek to białko CpgA. CpgA jest to GTP-aza kluczowa w procesie tworzenia rybosomu. Wyniki dotychczasowych badań eksperymentalnych wskazywały, że kinaza PrkC fosforyluje przynajmniej trzy reszty aminokwasowe białka CpgA (T166, T192, S226). Stosując szereg metod teoretycznych (dynamika molekularna, dokowanie molekularne) Autor starał się wskazać najbardziej prawdopodobne miejsca fosforylacji białka CpgA. Podsumowując wyniki swoich analiz, Autor konkluduje (strona 99), że najbardziej prawdopodobnymi resztami aminokwasowymi ulegającymi fosforylacji (katalizowanej przez kinazę PrkC) w białku CpgA są reszty S227 i T205 zaś reszta T166 raczej fosforylacji nie powinna ulegać. O ile w przypadku reszt S226 i S227 jest pewna zgodność danych eksperymentalnych z wynikami symulacji (wskazano jako miejsce działania kinazy dwie sąsiadujące reszty), to w przypadku pozostałych miejsc fosforylacji wyniki badań eksperymentalnych i teoretycznych są wyjątkowo

rozbieżne, szczególnie dotyczy to reszty T166. Uderzające jest to, że Autor nie stara się znaleźć wyjaśnienia dla zaobserwowanych sprzeczności pomiędzy wynikami swoich eksperymentów, a wynikami innych badań prezentowanymi w literaturze.

Uwaga ogólna:

W każdej dziedzinie wiedzy wypracowano pewne standardy techniczne/warsztatowe, które muszą być spełnione aby można było uznać wyniki prac badawczych za wiarygodne. Bardzo często czasopisma naukowe w wymaganiach dla autorów publikują podstawowe kryteria jakie muszą spełniać eksperymenty opisywane w manuskryptach prac naukowych. Pozwolę sobie przytoczyć fragment takich technicznych wymagań dla czasopisma *Journal of Molecular Modeling*, które specjalizuje się w publikacji wyników badań takich jak te przedstawione w recenzowanej rozprawie doktorskiej mgra Marcina Augustyniaka.

<http://www.springer.com/chemistry/theoretical+and+computational+chemistry/journal/894>

„For docking studies, at least 3 different conformations of the target (from multiple crystal structures or generated using an MD simulation of at least 10ns) must be used to generate consensus poses. Docked poses must again be confirmed using MD simulations of at least 5-10 ns if there is no experimental validation.” Z przytoczonego tekstu wynika, że aby można było dyskutować na temat struktur kompleksów białko-białko (takie zagadnienie jest poruszane w tej rozprawie) przynajmniej trzy różne wersje (pozy) takiego kompleksu powinny zostać zaprezentowane, a prawdopodobieństwo występowania każdej z takich struktur powinno zostać zweryfikowane za pomocą odpowiednio długiej symulacji z zastosowaniem dynamiki molekularnej, albo za pomocą ukierunkowanych przewidywań teoretycznymi badaniami eksperymentalnymi. Autor rozprawy przedstawia znaczną liczbę modeli strukturalnych badanego układu białko-białko. Niestety tylko dla jednego z kompleksów przedstawiono wyniki weryfikacji struktury tego kompleksu za pomocą odpowiedniej i prawidłowo przeprowadzonej symulacji dynamiką molekularną (patrz Uwagi merytoryczne I i II). Można więc podsumować, że wyniki przedstawione przez Autora w rozprawie doktorskiej nie spełniają minimalnych wymogów technicznych/warsztatowych stawianych dla tego rodzaju badań przez specjalistyczne czasopismo naukowe. Innymi słowy, prezentowany materiał nie byłby nawet poddany recenzjom merytorycznym, a odrzucony na poziomie wstępnej decyzji redakcji. **Przedstawiona do recenzji praca doktorska, a właściwie badania tam opisywane są po prostu niedokończone.** Sam Autor przyznaje otwarcie w tekście rozprawy, że część planowanych badań została wykonana tylko w około 50% (strona 84, wiersz 1-2): *„At time of writing this text, rigid-body docking of PrkCc and CpgA conformations obtained from their aMD trajectories is complete at approximately 50% (...)”*. W innym miejscu (strona 97, wiersz 12-13) Autor stwierdza, że nie zostały przeanalizowane wszystkie wyniki: *„Although, at time of writing this, I have analyzed only 6 out of 979 combinations of PrkCs’s (10 plus the initial computational model) and CpgA’s (89) conformations (...)”*. W subiektywnej ocenie Recenzenta, Autor rozprawy bez dogłębnego rozeznania stopnia skomplikowania badanego problemu, podjął się zadania, które Go (Autora) przerosło pod względem skomplikowania technicznego, potrzebnych do jego realizacji zasobów, co w efekcie skutkowało niezrealizowaniem zamierzonych zadań badawczych i wyciągania wniosków z niekompletnych danych pochodzących z niedokończonych badań.

Uwagi merytoryczne:

I) Struktura białka CpgA zdeponowana w bazie PDB (identyfikator rekordu 1T9H) nie zawiera współrzędnych atomów reszt aminokwasowych od reszty P186 do reszty T206. W rozdziale 4.1 Autor poświęca opisowi rekonstrukcji tego fragmentu jedno zdanie zilustrowane obrazem zrekonstruowanego fragmentu (Rysunek 4.1). Z Rysunku 4.1 wynika, że nieokreślony w kryształach fragment białka CpgA ma w większości strukturę α -helikalną. Biorąc pod uwagę wysoką rozdzielczość pomiaru krystalograficznego wynoszącą 1,64 Å wydaje się niezwykle mało prawdopodobne, że tak dobrze zdefiniowany fragment struktury drugorzędowej jakim jest α -helisa nie został jednoznacznie zidentyfikowany na mapie gęstości elektronowej będącej wynikiem pomiarów dyfrakcyjnych, szczególnie przy tak wysokiej rozdzielczości pomiaru. Niestety Autor rozprawy zdaje się nie przejmować tymi przesłankami i przyjmuje, bez dodatkowych badań i analiz, że struktura zrekonstruowanego fragmentu jest poprawna. Fragment, o którym mowa jest dość kluczowy, bo większość wyników prezentowanych przez Autora dotyczy reszty T205, a więc reszty aminokwasowej zlokalizowanej w omawianym regionie sekwencji. **W opinii recenzenta rutynowe i bezrefleksyjne potraktowanie rekonstrukcji fragmentu 197-205 białka CpgA bez dodatkowych badań i analiz uwiarygadniających poprawność otrzymanego modelu jest poważnym błędem. Należy pamiętać, że jeśli został popełniony błąd przy budowie modelu białka CpgA, to rzutuje on na wszystkie wyniki otrzymane przez Autora rozprawy.**

II) Białko CpgA zawiera motyw wiążący jon cynku (reszty C252, C257, C265, H259), a struktura krystalograficzna (1T9H) potwierdza obecność jonu cynku związanego w tym regionie. W rozdziale opisującym Metody (Rozdział 3) nie ma wzmianki o tym, czy, a jeśli tak to w jaki sposób, jon cynku był uwzględniany w symulacjach dynamiki molekularnej zarówno kompleksów białek PrkC – CpgA, jak i symulacji samego białka CpgA. Znaczna dynamika regionu 250-260 białka CpgA pokazana na Rysunku 4.24 (strona 67) może sugerować, że jon cynku raczej nie został uwzględniony w symulacjach. Jeśli jednak obecność jonu cynku uwzględniono, a jedynie pominięto ten fakt w rozdziale Metody (Rozdział 3), to zastanawiające jest dlaczego omawiany fragment strukturalny jest tak dynamiczny? Związany jon cynku powinien znacząco usztywniać strukturę w tym regionie sekwencji. **Autor niestety w rozprawie nie przedstawił odpowiednich opisów czy analiz, które wyjaśniałyby dlaczego region obejmujący reszty 250-260 białka CpgA jest tak bardzo dynamiczny, pomimo tego, że wszelkie inne przesłanki wskazują, że tak nie powinno być. Podobnie jak w przypadku poprzedniego zarzutu jakkolwiek błąd czy niedopatrzenie przy przeprowadzeniu badań rzutuje na prawie wszystkie wyniki otrzymane przez Autora.**

III) W trakcie realizacji badań Autorowi duży problem sprawiała ocena generowanych w czasie dokowania kompleksów pod względem ich potencjalnej funkcjonalności. Przez funkcjonalność w tym przypadku należy rozumieć ułożenie reszt seryny, treoniny lub tyrozyny względem cząsteczki ATP, które może w prosty sposób implikować reakcję fosforylacji. Autor jako jedno z kryteriów funkcjonalności stosuje odległość pomiędzy atomem fosforu gamma z cząsteczki ATP, a tlenem hydroksylowym reszt seryny, treoniny lub tyrozyny. Drugim kryterium używanym przez Autora jest zdefiniowany na Rysunku 3.2 (strona 40) kąt θ pomiędzy wektorem zakotwiczonym na atomie fosforu gamma cząsteczki ATP i prostopadłym do płaszczyzny wyznaczonej przez atomy tlenu O1G, O2G i O3G cząsteczki ATP, a wektorem łączącym atom fosforu gamma cząsteczki ATP i atom tlenu grupy hydroksylowej reszty seryny, treoniny lub tyrozyny. W rozdziale 3.1.3. (strona 39) Autor stwierdził, że jako najlepsze struktury będą wybierane takie, których miara kąta θ będzie jak najbliższa zera stopni. Nie wiem, dlaczego Autor przyjął takie założenie? Oczywiście w trakcie reakcji fosforylacji w stanie/prодукcie przejściowym reakcji, można oczekiwać, że miara kąta θ powinna być równa zero stopni, ale Autor badał kompleksy molekularne, a nie stany przejściowe reakcji chemicznej. W znanych strukturach krystalograficznych kinaz w kompleksach z inhibitorami tzw. „analogami stanu przejściowego” miara kąta θ wynosi około 28 stopni, np. struktura zdeponowana w bazie PDB: rekord 1GY3 Cook, A. i inni (2002) Biochemistry 41: 7301. Jednak w kompleksach kinaz z

ATP i substratem peptydowym miara kąta θ wynosi około 75 stopni. Np. struktury zdeponowane w bazie PDB rekordy 1B38, 1B39 Brown, NR. (1999) J. Biol. Chem. 273: 8746. **W większości analiz prezentowanych w pracy Autor odrzucał modele dla których miara kąta θ była większa niż 45 stopni (patrz Rozdział 4.6), a więc takie struktury, które według dostępnych informacji literaturowych powinny być jak najbardziej pożądane !!!** W świetle dostępnych danych literaturowych przyjęte przez Autora kryteria funkcjonalności badanych kompleksów, szczególnie te dotyczące wzajemnej orientacji przestrzennej cząsteczki ATP względem fosforylowanej grupy hydroksylowej, nie znajdują żadnego uzasadnienia. W trakcie analizy funkcjonalności generowanych kompleksów Autor całkowicie pominął ważny element centrum katalitycznego kinaz. Integralnym elementem centrum katalitycznego kinaz jest grupa karboksylowa łańcucha bocznego kwasu asparaginowego (bardzo rzadko glutaminowego), która pełni rolę zasady odbierającej proton z fosforylowanej grupy hydroksylowej (Brown, NR. (1999) J.Biol.Chem. 273: 8746: Sith, GK (2011) J. Phys. Chem. B, 115: 13713; Leigh, KN i Webster, CE (2014) Dalton Trans. 443: 3039). Autor w swoich badaniach nie starał się zidentyfikować kwasowej reszty aminokwasowej kinazy PrkC i nie użył tej reszty do weryfikacji funkcjonalności generowanych w trakcie badań modeli strukturalnych.

Pozostałe uwagi (w tym dotyczące strony edytorskiej pracy):

- 1)** Jak określono w Rozdziale 2 (strona 36) celem pracy było zaproponowanie struktur kompleksów tworzonych przez kinazę PrkC z *B. subtilis* z endogennym substratem jakim jest białko CpgA. Zwraca uwagę niezgodność tytułu pracy, który wyraźnie sugeruje, że badane będą różne substraty kinazy PrkC, a celem pracy wskazującym tylko jeden taki substrat.
- 2)** Numeracja stron podana w spisie treści, spisie rysunków i tabel nie odpowiada rzeczywistej lokalizacji w tekście poszczególnych rozdziałów, tabel i rysunków.
- 3)** W tekście całej rozprawy zabrakło informacji o planie badań, jak Autor zamierza osiągnąć cel nakreślony w Rozdziale 2, jakie badania są planowane i w jakiej kolejności.
- 4)** Spis literatury jest ponumerowany i obejmuje 239 pozycji. Odnośniki literaturowe nie są jednorodne. Autor w większości przypadków używa pełnych pierwszych imion autorów i inicjału w przypadku kolejnych imion, ale w wielu pozycjach literaturowych są tylko same inicjały przykładowo pozycje 9, 13, 15. Autor stosuje dość nieortodoksyjną metodę cytowania literatury. Jak napisałem, odnośniki są ponumerowane, co sugerowałoby, że w takiej kolejności jak pojawiają się w spisie literatury powinny się pojawiać w tekście. Niestety pierwsze cytowania jakie pojawiają się w Abstrakcie są to cytowania o numerach 3, 32, 177, 133, 134. Zdecydowanie nie jest to przyjęty w literaturze naukowej sposób prezentacji i organizacji odnośników literaturowych.
- 5)** Na stronie 41 są odwołania do nieistniejącego w tekście rysunku/rysunków.
- 6)** Autor raz używa skrótów jednoliterowych aminokwasów, a w innych miejscach skrótów trójliterowych.
- 7)** Praca napisana jest w języku angielskim. Niestety w wielu przypadkach tekst jest albo niezrozumiały albo tak niejednoznaczny, że utrudnia albo wręcz uniemożliwia rzetelną ocenę. Autor bardzo często używa bardzo długich zdań, których sensu można się tylko domyślać. Przykładowo zdanie ze strony 36: „*I have assumed that modeling of an enzyme in complex with its substrate must be focused on function, so I have concentrated my efforts on the enzymatic activity of phosphoryl group transfer by PrkC to one of serine, threonine, or, subsequently also taken into consideration, tyrosine residues of CpgA.*”

8) Autor używa zwrotu „*numeric properties*” w opisie rysunków 4.43 (strona 88); 4.44 (strona 89), chyba nie zdając sobie sprawy, co ten zwrot oznacza.

9) Praca pełna jest rysunków przedstawiających struktury cząsteczek lub wybrane fragmenty strukturalne. Graficzna prezentacja wyników jest niezwykle sugestywna i pomocna w zrozumieniu przekazu Autora. Jednak graficzne reprezentacje wyników wymagają, aby były to obrazy przemyślane, wykonane poprawnie technicznie, oraz były wyczerpująco opisane. Niestety zdecydowana większość z zamieszczonych rysunków strukturalnych nie spełnia podanych powyżej kryteriów. Przykładowo Rysunek 4.17 str. 59, opis rysunku brzmi: „*Of the best 21 initial complexes targeted to T205, 8 – the best per cluster – are rendered. PrkCc rendered with green cartoon. CpgA rendered with ribbon, and Mg²⁺ with a silver sphere. Clustering and selection criteria are described (...)*” (dalsza część opisu nie jest istotna). Opis nie zawiera informacji, co ten rysunek przedstawia (struktury kompleksów białek PrkCc – CpgA – dopisek recenzenta), w cytowanym opisie wymieniony jest kolor zielony (rzeczywiście jedna ze struktur jest pokolorowana w taki sposób), ale są też struktury pokolorowane na żółto, niebiesko, jasnobrązowo (szaroczerwono) jednak opis rysunku nie wyjaśnia, co te kolory oznaczają. Objasnienie dotyczące pojawiających się na rysunku kolorów znajduje się w tekście na stronie 52 (siedem stron od omawianego rysunku). Struktury są narysowane dość grubą kreską, co powoduje, że jest to plątanina linii niemożliwa do analizy. Rysunki 4.18, 4.27, 4.29, 4.30, 4.31 przedstawiają tak wiele szczegółów, nie opisanych dobrze w podpisach, że są kompletnie nieczytelne. Użycie kolorów w Rysunkach 4.37, 4.38, 4.39 (nie opisanych co oznaczają) i źle dobranych grubości linii pozwala je oceniać jedynie w wymiarze artystycznego, a nie informacyjnego przekazu.

10) Znamienne dla całej rozprawy jest umieszczanie rysunków, wykresów czy tabel, które są jedynie zaanonsowane w pracy w jednym zdaniu. Ekstremalny przykład to Rozdział 4.4.2 (strony 72-74). Na trzech stronach rozprawy są trzy rysunki, jedna tabela i **cztery zdania opisu**. „*Table 4.9 summarizes the results of LMOD [106-108] (as implemented in NAB [132]) flexible docking of three structures of PrkCc · Mg²⁺ · ATP · CpgA complex targeted to Thr192, Thr205 and Ser227. The conformations obtained are visualized in Figures 4.30, 4.31 and 4.32. Each of the computations lasted approximately 36h. The structure targeted to Thr205 (ID 1.14179) is the same as that used for both cMD simulations.*” Znaczna część rysunków i tabel mogłaby być usunięta z tekstu (albo przesunięta do załączników), co w dużej mierze poprawiłoby czytelność samego tekstu. Ponadto w Rozdziale 5 - Dyskusja przydałyby się rysunki wybranych modeli kompleksów ilustrujących opisywane i dyskutowane wyniki.

11) Strona 41, wiersz 6-8: „*I have simply forgotten to reinsert the Mg²⁺ cation after rigid-body docking with ZDOCK (see Section 3.1) prior to starting the former of the simulations.*” Jak napisano na stronie 64 ta pomyłka kosztowała Autora 13 dni obliczeń (nie jest podane, jak dużo zasobów obliczeniowych procesoro/godzin (Servis Units SUs) poświęcono na tę pomyłkę). Ludzką rzeczą jest się mylić, należy z takich pomyłek wyciągać wnioski, ale rozprawa doktorska nie jest miejscem na to, aby wylewnie opisywać tego rodzaju zdarzenia.

12) W tytule rozprawy badana kinaza nazywana jest PrkC, jednak już w Abstrakcie pojawia się nazwa PrkCc. W dalszej części tekstu częściej, ale nie wszędzie stosowana jest nazwa PrkCc. Jak należy rozumieć stosowanie tych dwu podobnych ale różnych nazw? Czy nazwy te dotyczą tego samego enzymu?

13) W Rozdziale 1.5 (strona 24) opisywana jest struktura białka CpgA. Autor nie przedstawia rysunku obrazującego strukturę przestrzenną opisywanego białka. Brak jest wyjaśnienia czym są opisywane w tekście obiekty nazwane G1, G2, G3, G4?

Podsumowanie:

Jak napisano w części recenzji pt. **Uwagi ogólne** badania zaprezentowane w rozprawie doktorskiej mgra Marcina Augustyniaka są niedokończone. Ze względu na brak jasnego planu pracy (patrz **Pozostałe uwagi p.3**) trudno jest oszacować, jaka część pracy została wykonana i w jakim stopniu. Budowanie wniosków (jak to ma miejsce w tej rozprawie) na podstawie fragmentarycznych wyników ma raczej znamiona hipotezy badawczej, a nie kompletnej i dojrzałej pracy naukowej jaką powinna być rozprawa doktorska. Niestety nawet te fragmentaryczne wyniki przedstawione w rozprawie są obarczone poważnymi błędami i uchybieniami (patrz **Uwagi merytoryczne p. I-III**), co rzutuje na ich (wyników) wiarygodność. Przedstawione powyżej zarzuty niestety dyskredytują tę rozprawę jako opracowanie naukowe. Biorąc pod uwagę podsumowane powyżej zarzuty i uwagi stwierdzam, że przedstawiona do recenzji rozprawa doktorska zatytułowana: *„Molecular modeling of complexes of PrkC protein kinase with possible endogenous targets of phosphorylation activity”* autorstwa mgra Marcina Augustyniaka **nie spełnia wymogów ustawowych** zawartych w ustawie „O stopniach i tytułach naukowych” z dnia 14 marca 2003 r. Niniejszym wnioskuję do Rady Międzyuczelnianego Wydziału Biotechnologii Uniwersytetu Gdańskiego i Gdańskiego Uniwersytetu Medycznego o **niedopuszczenie** mgra Marcina Augustyniaka do dalszych etapów przewodu doktorskiego.

St. Oty