Warszawa, dn. 15.06.2016

Prof. dr hab. Sławomir Filipek,
Wydział Chemii, Centrum Nauk Biologiczno-Chemicznych,
Uniwersytet Warszawski,
ul. Pasteura 1, 02-093 Warszawa
Tel. 22-55-26405,
E-mail: sfilipek@chem.uw.edu.pl

Recenzja
rozprawy doktorskiej mgr. Stefana Mordalskiego, pt.

Computational studies on the structures of G protein-coupled receptors. Tools supporting homology modeling and the analysis of ligand-receptor interactions

Przedstawiona mi do oceny praca doktorska dotyczy badań teoretycznych przeprowadzonych dla receptora GPCR β_2-adrenergicznego z wykorzystaniem narzędzi wytworzonych przez doktoranta dla badania i opisu receptorów GPCR. Praca ma charakter klasycznej rozprawy i obejmuje 84 strony oraz dodatek zawierający dwie publikacje doktoranta: jedna w *Nucleic Acids Research*, a druga w *Journal of Cheminformatics*. Doktorant jest tam odpowiednio drugim i pierwszym autorem, co wskazuje na duży wkład pracy włożonej w otrzymanie i analizę wyników. Wstęp pracy doktorskiej liczy 19 stron, opis stosowanych Metod oraz skonstruowanych narzędzi obliczeniowych 16 stron, Wyniki 18 stron, Dyskusja 7 stron, oraz Podsumowanie 2 strony. W tekście pracy zamieszczono 35 figur, 12 tabel, oraz wykaz skrótów. Praca jest napisana bardzo dobrze pod względem językowym. Występują bardzo nieliczne błędy niewarte wymienienia. Bibliografia obejmuje 105 pozycji, w większości z ostatnich 10 lat. W pozycjach literaturowych podawanych jest tylko do 3 pierwszych autorów, a szkoda, bo najważniejszym jest często autor końcowy. Ponadto, w pozycjach 11 i 15 podany jest tylko jeden autor, a są to prace wieloautorskie.

We wprowadzeniu doktorant przedstawia ogólną charakterystykę receptorów GPCR, czyli jak wynika z nazwy, receptorów sprzężonych z białkami G, nazywanych również zgodnie z ich budową receptorami 7TM. Doktorant przedstawia klasyfikację tych 800 receptorów opartą na podobieństwie sekwencyjnym, tzw. podział GRAFS, jak również

starszy podział na klasy (lub rodziny) A-F. Przedstawiony jest również ogólny schemat budowy dla poszczególnych klas tych receptorów. Na rys. 3 przedstawiono podział helis receptora na te, które odpowiadają za sygnalizację poprzez białko G oraz na odpowiadające za, niedawno odkrytą, sygnalizację poprzez arestynę (helisa TM7). Niestety, nie podano odnośnika do tych informacji. Ponieważ do tej pory nieznane są wszystkie szczegóły aktywacji i sygnalizacji tych receptorów, a w szczególności sygnalizacji przez arestynę, wobec tego podawanie takich informacji uważam za przedwczesne. Należy raczej oczekiwać, że rozróżnienie tych obu typów sygnalizacji jest znacznie bardziej subtelne, szczególnie, że sam doktorant przedstawia na rys. 10 niemal identyczne struktury rodopsyny ze związanymi fragmentami białka G i arestyny.

We Wstępie przedstawiony jest jednolity sposób numeracji reszt helis transbłonowych receptorów GPCR zwany od nazwisk autorów numeracją Ballesterosa-Weinsteina oraz nowszy sposób, oparty na strukturach, a nie tylko sekwencji tych receptorów, zastosowany w serwisie GPCRdb. Doktorant skupia się na opisie zewnątrzkomórkowej części tych receptorów, która jest odpowiedzialna za wiązanie ligandów, jednak przedstawia także zachowane ewolucyjnie motywy strukturalne w całym receptorze, które związane są z tzw. przełącznikami molekularnymi pełniącymi istotną rolę w procesie aktywacji receptorów GPCR. Na rys. 12 przedstawione są zmiany położenia helis transbłonowych w czasie aktywacji tych receptorów. Na tym rysunku helisa TM6 podlega tylko obrotowi, brakuje natomiast przesunięcia cytoplazmatycznej części tej helisy na zewnątrz receptora. Wygląda więc na to, że przedstawione zmiany dotyczą tylko części zewnątrzkomórkowej, bowiem ruchy obu części receptora są w miarę niezależne. Na str. 14 napisano, że receptory klasy A mają jedynie krótkie N-końce bez struktury II-rzędowej - końce te są odcinane aby ułatwić krystalizację. Jednakże, nawet krótki N-koniec w rodopsynie, nadal jedynym receptorze GPCR, którego struktura jest znana w całości, posiada beta-kartkę, która ma funkcjonalne znaczenie. Na rys. 4 podane statystyki dotyczące znanych struktur receptorów są zamienione miejscami. Z kolei na str. 20 fragment (*accommodation of the orthosteric ligands in class B-F*) i odnośnik do rys. 3 prawdopodobnie dotyczy rys. 2.

We Wstępie przedstawiono także ogólny schemat modelowania przez homologię oraz opis serwisu GPCRdb wraz ze spisem nowych funkcji dodanych po 2013 roku, od kiedy opiekę na tym serwisem przejął dr David Gloriam.

W Metodach sporo miejsca zajmuje opis nowo-skonstruowanych narzędzi do analizy dokowania ligandów oraz rozróżniania funkcji liganda, tj. czy ligand aktywuje bądź nie

aktywuje dany receptor. Doktorant opisuje też użyte źródła ligandów, w tym bazy danych: DUDs oraz ZINC. W sumie do badań użyto 271 agonistów, 324 antagonistów i inwersyjnych agonistów, oraz po 2000 nieaktywnych związków z baz DUDs i ZINC. Dokowania zostały przeprowadzone za pomocą programu GLIDE z pakietu Schrödinger 2014. Opisana jest także procedura *machine learning*, która została wykorzystana do podziału ligandów na aktywne i nieaktywne. Opisano także strukturę profilu oddziaływań ligand-receptor za pomocą *Structural Interaction Fingerprints* (SIFt) oraz nowy typ deskryptora macierzowego 2D-SIFt, który został skonstruowany przez doktoranta i wykorzystany do analizy dokowania do receptora adrenergicznego β_2-AR, zarówno do jego struktur krystalicznych jak i modeli homologicznych.

Jak podaje doktorant były trzy równoległe cele pracy: 1) rozwój algorytmów i narzędzi do modelowania homologicznego receptorów GPCR; 2) nowa metoda opisu oddziaływań ligand-receptor; i 3) rozwój narzędzi do pracy nad sekwencjami i strukturami receptorów GPCR. Cel pierwszy był celem głównym jednakże, jak wyjaśnia później doktorant, rozumianym jako wykorzystanie tych narzędzi do tzw. wirtualnego skriningu, obejmującego przesianie baz danych ligandów pod kątem rozróżnienia aktywny/nieaktywny ligand. Ponieważ receptory GPCR występują w bardzo wielu stanach konformacyjnych, bardzo ważne jest opracowanie odpowiedniej procedury pozwalającej na powyższe rozróżnienie w celu znalezienia bądź skonstruowania nowych związków chemicznych mogących służyć jako nowe potencjalne leki. W tym celu doktorat użył nie jednej ale całego zestawu struktur danego receptora GPCR, w tym przypadku β_2AR, i porównał użyteczność wielu struktur krystalicznych tego receptora z różnymi ligandami względem zestawu modeli homologicznych otrzymanych z różnych wzorców, będących innymi receptorami GPCR z tej samej klasy A, tj. receptorów podobnych do rodopsyny. Ta procedura jest odzwierciedleniem faktu, że receptory GPCR podlegają znacznym zmianom swojej struktury po związaniu niektórych ligandów, w szczególności pełnych agonistów lub pełnych inwersyjnych agonistów. Ponadto, pełna zmiana struktury receptora w czasie jego aktywacji jest możliwa tylko w obecności białka G lub przeciwciał, wiążących się w tym samym miejscu receptora. To wszystko sprawia, że wirtualny skrining receptorów GPCR jest trudny i podatny na błędy.

Jak pokazuje doktorant, użycie wielu struktur tego samego receptora w procesie rozpoznawania ligandów podnosi dokładność tej metody o ok. 10 % przy 20 modelach homologicznych. Najlepsze wyniki dało modelowanie na wzorcu będącym receptorem muskarynowym M2, a najgorsze dla receptora dopaminowego D3 co wskazało, że

podobieństwo sekwencyjne receptorów nie przekłada się automatycznie na najbardziej trafne dokowania ligandów. Najciekawszy jednak wynik osiągnięto porównując rezultaty wirtualnego skriningu ligandów opartego na strukturach krystalicznych i na modelach homologicznych. Użycie modeli homologicznych okazało się być dużo lepsze a dodawanie nowych struktur krystalicznych nie poprawiało wyników, jak to było przy modelach homologicznych. Do skriningu użyto zarówno jednowymiarowych opisów oddziaływań ligand-receptor SIFt jak również wersji 2D-SIFt, w postaci macierzy, rozwijanej przez doktoranta. Do rozróżnienia aktywny/nieaktywny ligand wykorzystano metody typu *Machine learning* a w szczególności algorytm SVM (*Support Vector Machines*). Stosując opis 2D-SIFt doktorant osiągnął rozróżnienie lepsze nie tylko od 1D-SIFt (0.81 w porównaniu do 0.74 parametru MCC, który przyjmuje wartości w zakresie 0-1) ale także od innych powszechnie używanych deskryptorów. W pracy stosowano modele receptorów bez pętli między helisami, w szczególności ECL2, bowiem źle wymodelowana pętla zagłębia się w miejsce wiążące receptora i ma ogromny wpływ na uzyskane wyniki. Tym niemniej doktorant stwierdza, że dokładniejsze wyniki dokowania, jak również skriningu, można uzyskać stosując te pętle. Jest to związane z oddziaływaniem aminokwasów z tej pętli z niektórymi ligandami związanymi w głównym miejscu wiążącym receptora, czyli w miejscu ortosterycznym. Nie brano pod uwagę ligandów allosterycznych – wynika to z braku odpowiednich struktur krystalicznych. Ligandy allosteryczne wiążą się głównie do N-końca receptorów GPCR, który jest jednak odcinany aby ułatwić krystalizację. Tym niemniej ligandy allosteryczne są coraz częściej w centrum zainteresowania badaczy z powodu ogromnych problemów z selektywnością ligandów ortosterycznych, bowiem główne miejsce wiążące jest zwykle mocno konserwowane ewolucyjnie i jest niemal identyczne w różnych podtypach tych receptorów.

Wiele struktur tego samego receptora może być również wygenerowanych przez dynamikę molekularną, i jak stwierdzono dla ligandów receptorów opioidowych mju i kappa, parametr SASA (*solvent accessible surface area*) może również być wykorzystany do rozróżniania typu liganda aktywny/nieaktywny. Chociaż stosowanie dynamiki molekularnej jest dużo bardziej czasochłonne niż generowanie struktur przez modelowanie homologiczne, tym niemniej daje wgląd w szczegóły procesu aktywacji receptora.

W czasie pracy nad serwisem GPCRdb doktorant zbudował różnorodne narzędzia ułatwiające korzystanie z tego serwisu - w szczególności parametr 2D-SIFt jest wykorzystywany do szczegółowego opisu oddziaływań pomiędzy ligandem a receptorem. Pozostałe narzędzia to: 1) automatyczne przypisywanie generycznych numerów reszt,

wg. schematu GPCRdb, dla aminokwasów z helis transbłonowych. Do tej pory było to robione ręcznie przez wszystkich badaczy receptorów GPCR; 2) automatyczne nakładanie struktur receptorów GPCR biorąc pod uwagę generyczne numery reszt, co automatycznie eliminuje złe nakładania, związane z wybrzuszeniami (*bulges*) bądź zwężeniami (*constrictions*) w helisach transbłonowych; oraz 3) generowanie modeli farmakoforowych opartych na 250 fragmentach ligandów ze wszystkich struktur receptorów GPCR.

Podsumowując, rozprawa jest bardzo zwarta, bez żadnych zbędnych fragmentów, i zawiera wiele cennych, opublikowanych wyników. Ponadto, opisane metody są już zaimplementowane w serwisie internetowym GPCRdb. Praca jest napisana po angielsku poprawnym językiem naukowym, co będzie sprzyjać jej rozpowszechnieniu i zapewne również cytowaniu w publikacjach naukowych i doniesieniach konferencyjnych. Wybór metod i krytyczna interpretacja uzyskanych wyników świadczą o dojrzałości naukowej Autora a wyżej wymienione drobne niedociągnięcia nie przesłaniają jednak znaczącej wartości naukowej dysertacji oraz znakomitego przygotowania merytorycznego doktoranta.

Przedstawiona mi do oceny rozprawa doktorska spełnia warunki określone w art. 13 Ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. nr 65, poz. 595 z późn. zm.), dlatego przedkładam wniosek do Rady Naukowej Wydziału Biochemii, Biofizyki i Biotechnologii Uniwersytetu Jagiellońskiego o dopuszczenie mgr. Stefana Mordalskiego do dalszych etapów przewodu doktorskiego. Wnoszę także o wyróżnienie tej rozprawy z powodu uzyskania wielu ciekawych i wartościowych wyników z zarówno poznawczego oraz także praktycznego punktu widzenia.

S. Filipek

Sławomir Filipek