

Data: Toruń, 20 grudnia, 2017

L. Dz.

Prof. dr hab. Wiesław Nowak
Zespół Teoretycznej Biofizyki Molekularnej
Instytut Fizyki
Uniwersytet M.Kopernika w Toruniu
87-100 Toruń, ul. Grudziądzka 5
(Wiesiek@fizyka.umk.pl)

**RECENZJA rozprawy doktorskiej
Mgr. inż. Pawła P. Woźniaka pt.**

“Supporting Protein Tertiary Structure Modelling with Contact Sites Prediction Methods”

Jak słusznie zauważył doktorant w jednym z akapitów swojej rozprawy, nie tylko jest ważne „skąd przychodzimy?”, czy „dokąd idziemy?”, ale też „kim jesteśmy?”. To pytanie filozoficzne skierowane może być do naukowców badających człowieka. Jednym z podstawowych materiałów budulcowych naszych ciał, w tym mózgow, są białka. Te fascynujące polimery wchodzi w skład wszelkich znanych obecnie form życia. Fundamentalną zasadą w biologii jest wiara, oparta na licznych eksperymentach, w ścisły związek pomiędzy strukturą, a funkcją białek. Funkcje te to, m.in., utrzymywanie właściwego kształtu i rozmiarów komórek, transport substancji, komunikacja między komórkami organizmu, przepisywanie informacji genetycznej, udział w samopowielaniu się organizmu, utrzymywanie równowagi termodynamicznej, przekazywanie impulsów nerwowych i wiele innych. Rola białek w złożonych systemach biologicznych jest fundamentalna. Wiedza na temat ich struktury rośnie w zawrotnym tempie, ale jest wciąż niedostateczna i niepełna. Korzystają z niej medycyna, farmacja, biotechnologia, zielona chemia i inne dziedziny techniki oraz nauk praktycznych. Potrafimy stosunkowo łatwo określić sekwencje białek, tj. kolejność setek aminokwasów składających się w łańcuchy polipeptydowe, ale nie potrafimy łatwo przewidzieć jaka będzie ich budowa przestrzenna (struktura 3D, lub struktura trzeciorzędowa). Eksperymenty (techniki CryoEM, NMR, X-ray, itd) są możliwe tylko dla niewielkiej części białek interesujących przyrodników. Co do reszty struktur – pozostają białą plamą, jeśli chodzi o ich kształt (fold, zwój) zdani jesteśmy na domyśle albo na badania komputerowe.

Technologia komputerowa jest w XXI wieku wspaniale rozwinięta. Komputery są tanie i łatwo dostępne. Pożytek z wykorzystania cybernetyki (rozumianej tu jako pewnego rodzaju technika komputerowa) w badaniu białek krytycznie zależy od dostępności algorytmów (programów lub metod) prowadzenia takich badań. Póki co najmańdrzejszy nawet komputer sam się nie zaprogramuje i nie wykona badań przynoszących nową wiedzę o białkach. Nowe metody i algorytmy mogą za to być proponowane przez doktorantów.

Przedstawiona mi do recenzji rozprawa doktorska mgr. Inż. P.P. Woźniaka wpisuje się w ważny nurt badań komputerowych białek. Nie trzeba siedzieć w dusznych laboratoriach by zdobywać nowe dane o tych niezmiennie ważnych układach. Tematyka badań jest aktualna, istotna w środowisku międzynarodowym, wyniki mogą być potrzebne wielu grupom badawczym, zaś dziedzina, którą zajmuje się p. mgr Woźniak jest bardzo „gorąca”, w tym sensie, że badania z pogranicza biologii, informatyki, fizyki, chemii, statystyki czy metod sztucznej inteligencji i bioinformatyki prowadzone są w wielu czołowych uczelniach świata. Na pewno nie jest to tematyka niszowa, chociaż poziom formalizmu jest w niej wysoki i problemy, jakie będę za chwilę dyskutował są trudno zrozumiałe dla niespecjalistów.

Białka mające daną funkcję w konkretnych organizmach niekoniecznie mają taką samą sekwencję. W toku ewolucji (miliony lat) część aminokwasów ulega zamianie na inne w miarę jak powstają bardziej złożone organizmy, inne aminokwasy zachowują w zasadzie swoje położenia w łańcuchach białkowych. Mówimy, że są „silnie konserwowane” (taki slang). Strukturę przestrzenną (3D) białka można charakteryzować na wiele sposobów, jednym z nich jest tworzenie mapy kontaktów. Przez kontakty rozumie się niskie odległości (np. <8 Ang.) pomiędzy tzw. atomami C_{β} poszczególnych aminokwasów. Okazuje się, że z mapy kontaktów można próbować przewidywać ceną dla nas strukturę 3D białka. Przewidywania te jak dotąd nie są niestety 100 % skuteczne. Jeśli nawet (znając niektóre kontakty) algorytm coś „przewidzi”, to często (a nawet przeważnie) eksperymentalnie określona struktura jest odmienna a przewidywanie błędne czyli bezużyteczne. Celem wysiłków doktoranta było podanie sposobów poprawienia skuteczności takich przewidywań. Składające się na rozprawę 3 oryginalne publikacje w b. dobrych periodykach specjalistycznych, poprzedzone obszernym wstępem, dowodzą, że cel ten, nb. bardzo trudny, został osiągnięty. Było to możliwe dzięki głębokiej znajomości złożonych algorytmów, zrozumieniu ewolucji białek, opanowaniu współczesnych metod analizy kontaktów i pewnej pomysłowości w zaproponowaniu nowych rozwiązań. Uzyskanie wyników przedstawionych w rozprawie wymagało wiele wysiłku programistycznego, obliczeń, żmudnej analizy i wielu godzin prac optymalizujących kody. Widząc poziom publikacji nie mam najmniejszych wątpliwości co do opanowania przez mgr inż. Pawła P. Woźniaka warsztatu naukowego z dziedziny symulacji komputerowych białek czy innych złożonych systemów biologicznych.

Przedstawiona mi do oceny praca doktorska mgr inż. Pawła P. Woźniaka, napisana pod kierunkiem prof. dr hab. Małgorzaty Kotulskiej w Zakładzie Inżynierii Biomedycznej Wydziału Podstawowych Problemów Techniki Politechniki Wrocławskiej oraz (ko-promotor) prof. Gerta Vrienda z „Center for Molecular and Biomolecular Informatics”, Radboud Univeristy Medical Centre, z Nijmegen (Holandia) dotyczy metod wspierania komputerowego modelowania struktur trzeciorzędowych białek.

Doktorant w rozprawie przedstawił 3 ogólne tezy, które starał się „udowodnić”:

1. „Wykorzystanie charakterystyk predykcji miejsc kontaktowych w białkach może przyczynić się do poprawy skuteczności ich predykcji”.
2. „Przewidywane miejsca kontaktowe mogą rozróżnić strukturę prawidłową od fałszywego modelu”.
3. „Skuteczność predykcji miejsc kontaktowych dla konkretnego białka może być przewidziana”.

Można przyjąć, że wykazanie tych tez było celem rozprawy doktorskiej, zostały one opisane we wstępie do rozdziału wprowadzającego jak i w polskim streszczeniu.

Rozprawa, napisana w języku angielskim zawiera kolejno: Spis publikacji doktoranta (1,2), Rozdział wstępny składający się z Podsumowania rozprawy (3) Ogólnego wprowadzenia w tematykę (4) Wykazu tez (5) Omówienie przebiegu własnych badań i wyników rozprawy (6). Ten fragment ma 34 strony. Spis odnośników do tego rozdziału obejmuje ok 100 pozycji ułożonych w porządku alfabetycznym. Rozprawę zamykają streszczenie w języku polskim (2 strony) oraz kopie 3 artykułów wieloautorskich (l.a. 2-5) w czasopismach międzynarodowych:

A) **Wozniak PP**, Kotulska M, Characteristics of protein residue-residue contacts and their application in contact prediction. *J Mol Model*. 2014 Nov; 20(11):2497. doi:

10.1007/s00894-014-2497-9. (IF₂₀₁₄: 1.736, 5-IF₂₀₁₄: 1.711, MNiSW₂₀₁₄: 25)

B) **Wozniak PP**, Vriend G, Kotulska M, Correlated mutations select misfolded from properly folded proteins, *Bioinformatics*. 2017 May 15;33(10):1497-1504. doi:

10.1093/bioinformatics/btx013. (IF₂₀₁₇: 7.307, 5-IF₂₀₁₆: 8.136, MNiSW₂₀₁₇: 45)

C) **Wozniak PP**, Konopka BM, Xu J, Vriend G, Kotulska M, Forecasting residue-residue

contact prediction accuracy, *Bioinformatics*. Epub 2017 June 26. doi: 10.1093/bioinformatics/btx416

(IF₂₀₁₇: 7.307, 5-IF₂₀₁₆: 8.136, MNiSW₂₀₁₇:45)

Prace A,B,C mają odpowiednio 17, 7 i 9 stron druku, IF czasopism wynosi 1.7, 7.3, 7.3 a mgr inż. P Woźniak jest w każdym artykule pierwszym autorem. Recenzent zapoznał się z oświadczeniami współautorów prac, dominujący wkład mgr. inż. P.P. Woźniaka do każdej z nich nie budzi wątpliwości.

Przedstawię kolejno ocenę rozdziału wprowadzającego, artykułów A-C i uwagi merytoryczne na ich temat. Uwagi formalne podam w dalszej części recenzji.

Rozdział wstępny napisany jest bardzo zwięźle, ale jest bardzo informatywny. Przedstawiona jest motywacja podjęcia tego typu badań. Wprowadzona jest używana dalej w rozprawie definicja kontaktu między aminokwasami ilustrowana przypadkiem krabiny. Wymienione są główne pozycje literaturowe związane z oceną kontaktów w białkach. Szczególnie ważny jest podrozdział o skorelowanych mutacjach. Opisany jest historyczny rozwój modeli przewidywania kontaktów w białkach, wprowadzone pojęcia wzajemnej informacji (MI) oraz informacji bezpośredniej (DI). Szczególnie cenna i nowa w tym przeglądzie jest moim zdaniem dyskusja problemu wyboru parametrów wejściowych do metod przewidywania kontaktów (4.3.3). Podrozdział ten ma bezpośredni związek z własnymi osiągnięciami doktoranta przedstawionymi w rozprawie. Pozytywnie oceniam niedługi, ale istotny fragment krytycznej analizy problemów na jakie napotykają metody oparte na wielokrotnym uliniowaniu (4.35). Omawiany rozdział przeglądowy dowodzi erudycji autora i świadczy o bardzo dobrej orientacji w aktualnej literaturze. Spis literatury sporządzony jest starannie, drobne niekonsekwencje zauważyłem na str. 35 (Bohr../ vs/ Bystroff; raz *J Mol Biol*, raz pełna nazwa czasopisma; NAR raz w skrócie raz pełna nazwa w Bateman et al..).

Końcowy fragment tego rozdziału (6) zawiera zwięzły opis historii prowadzonych badań, na tle rozwoju metod śledzonego w literaturze. Podane jest też krótkie omówienie najważniejszych wyników rozprawy, przedstawionych w publikacjach oraz autorski komentarz na temat ich

znaczenia. Wartościowe są odniesienia do najnowszych prac z tematyki doktoratu i pokazanie linii rozwojowej badań na temat przewidywania kontaktów w białkach.

Są i drobne mankamenty. W rozdziale tym nie ma niestety definicji tego, co autor uważa za „przewidywanie kontaktów”. Jest to pojęcie nieostre i wcale nie oczywiste. Czy od metody oczekujemy by w oparciu o sekwencje przepowiedziała wszystkie kontakty jakie wystąpią w białku? Czy niektóre? Czy „najważniejsze”? W oparciu o jakie dane i czy dla jednej sekwencji czy dla klas białek o wspólnym zwoju (foldzie) spodziewamy się, że metoda będzie skuteczna? Według mnie zbyt skrótowo jest potraktowany wątek wprowadzający pojęcie mutacji skorelowanych. Przykład na rysunku 2 nie jest wcale oczywisty, i myślę, że dopiero podzielenie MSA na podgrupy (MM), (GL), (FH), (AS) pozwoliłoby Czytelnikowi lepiej zrozumieć o co tu chodzi. Kolory na rys. 1 nie są właściwie dobrane, w mojej kopii niebieskie obszary zlewały się z czarnymi. Pewien niedosyt budzi pominięcie w tym rozdziale chociaż prostego opisu modelu Potta, nie każdy biocybernetyk interesuje się fizyką statystyczną i uogólnieniami modelu Isinga.

Kolejny rozdział rozprawy to obszerna **publikacja (A)** (Wozniak PP, Kotulska M, *Characteristics of protein residue-residue contacts and their application in contact prediction*. J Mol Model. 2014) zamieszczona w dobrze rozpoznawalnym czasopiśmie Journal of Molecular Modeling wydawanym przez Springer. Autorzy wprowadzili dwa istotnie nowe ulepszenia do problemu oceny kontaktów w białkach (1) poddali analizie ujednoliconą bazę 5872 struktur białkowych, (2) zaproponowali serie parametrów wspomagających analizę właściwości zbiorów białek, w tym przewidywania nowych kontaktów. Rozwój metod predykcji struktur opiera się zwykle na właściwościach statystycznych obserwowanych w pewnych zbiorach. Łączy się cechy białek z ich klasyfikacją. Niestety, najważniejsze dwie bazy strukturalne foldów białek – CATCH i SCOP stosują nieco inne zasady klasyfikacji i oparcie się na danych strukturalnych wybieranych „losowo” może być problematyczne. W omawianej pracy zaproponowano spójny zestaw struktur, zawierający 29 różnych topologii, każda „topologia” ma przynajmniej 20 reprezentatywnych domen białkowych, co ważne, podjęto starania by usunąć redundancje (zbyt duże podobieństwo sekwencyjne elementów zbioru) i dokonano nieco arbitralnego podziału na klasy z dominującą strukturą α lub β lub bez dominacji ($\alpha+\beta$). Baza posłużyła do obliczania różnych parametrów związanych z częstością kontaktów, normalizowanych lub nie. Parametry te (W_c , W_t , f_p , f_{pn} , S_t) są w pewnym sensie „intuicyjne” i ich analiza wnosi ciekawe informacje o cechach strukturalnych białek. Nie są to informacje zupełnie nowe, wcześniej dokonywano podobnych analiz, jednak wartość tej pracy upatruje w systematycznym przebadaniu największego (wg mojej wiedzy) zbioru struktur oraz w porządnej analizie wyników obliczonych parametrów. Obliczenia wykonano własnymi programami, wyniki porównano, tam gdzie to możliwe ze znacznie skromniejszymi danymi literaturowymi. Okazało się, na przykład, że wartości parametrów f_p i f_{pn} pozwalają określić, które aminokwasy najczęściej tworzą kontakty. Wyniki (Cys, Ile, Leu, Met, Phe, Trp, Tyr, Val) nie są zaskakujące, są to aminokwasy hydrofobowe lokujące się zwykle w gęsto upakowanych wnętrzach białek. Dla mnie jednak ciekawa była klasyfikacja liczbowa tej „skłonności do oddziaływań” z innymi aminokwasami, bez porządnej analizy statystycznej dokonanej przez doktoranta takiej wiedzy byśmy nie mieli. Inny bardzo ciekawy wynik pojawił się przy analizie indeksu S_t , pokazującego „indywidualizm” danej grupy topologicznej. Jedna z grup wykazała wartość tego parametru 5-10 razy wyższą niż pozostałe. Autorzy zgrabnie wyjaśnili przyczynę tego efektu (topologia 12 odpowiada

pojedynczej alfa helisie) i słusznie zauważyli, że S_t może mieć duże znaczenie w diagnostyce modeli białek. Spore oczekiwania wiązano zapewne z wprowadzeniem informacji o wyznaczonych parametrach do procesu przewidywania nowych kontaktów. Zastosowano je w sprytnym algorytmie, próbując poprawić przewidywania uznaną metodą DCA (Morcos i in. PNAS, 2011). Efekt uzyskano, jednak w mojej ocenie jest on dość mizerny: poprawa jest wyraźna dla tzw. „top scores”, ale tylko w klasie białek ze strukturą α . Nie chcę tą uwagą deprecjonować znaczenia pracy (A): dokonanie tej analizy było konieczne, podejście jest właściwe, a wynik świadczy tylko o tym, jak trudno jest naturze wyrwać jej tajemnice. Poza tym sądzę, że zastosowany algorytm dorzucania informacji o parametrach statystycznych może wzbogacić jeszcze inne metody predykcji kontaktów, więc wysiłek doktoranta poszedł we właściwym kierunku i zasługuje na uznanie.

Praca (A) jest dość trudna w odbiorze, ale napisana z troską o jasność prezentacji. Są w niej drobne niedoskonałości techniczne: wg mnie rys. 4 i 5 są mało czytelne i dane tam zgromadzone powinny być rozdzielone na osobne diagramy dla separacji 7 i separacji 15 (nb. dlaczego odchylenia standardowe są tak wysokie?); pisownia słowa „amino acid” nie jest konsekwentna w podpisach do rysunków (Fig. 1,6,7 mają inaczej niż 5, 10, 12). Nie znalazłem też informacji czy zakodowany algorytm rozwinięty w tej pracy jest dostępny dla innych badaczy, byłoby to cenne, zważywszy na możliwe różne pomysły klasyfikowania białek. Również ważne jest moim zdaniem udostępnienie samej bazy struktur (klas α , β i $\alpha+\beta$). W samej treści artykułu nie ma informacji o jednej ważnej kwestii: Ile z drugiego (stosowanego do oceny predykcji) zbioru danych Pfam (562 domen) było wspólnych w pierwotnym zbiorze 5872 domen użytym do wyznaczania parametrów)?

Praca (B), opublikowana w tym roku w czołowym czasopiśmie bioinformatycznym [Wozniak PP, Vriend G, Kotulska M, *Correlated mutations select misfolded from properly folded proteins*, Bioinformatics. 2017 May 15;33(10):1497-1504], dotyczy ważnego problemu wspierania procesu wyboru właściwego modelu 3D białka w oparciu o przewidywanie kontaktów. Przedstawione wyniki opierają się na znanym z literatury założeniu, że jeśli robiąc wielokrotne uliniowanie (MSA) zauważymy w serii sekwencji homologicznych białek korelacje w występowaniu mutacji, to możemy wnioskować, że albo te skorelowane aminokwasy są powiązane funkcjonalnie albo przynajmniej występują w bliskim kontakcie. Autorzy wykorzystali nowoczesny wariant metody DCA (*Direct Coupling Analysis*) - gplmDCA do oceny kontaktów w dwóch starannie i pomysłowo dobranych grupach białek: (a) serii struktur białkowych z PDB, które są już poprawione (błędna-poprawna, 152 pary) oraz (b) serii białek z ich celowo „zepsutymi” odpowiednikami strukturalnymi (zastosowano bazę „Decoys ‘R’ Us, 26 par). Tym razem kontakty zdefiniowano poprzez atomy C α . Porównano kontakty przewidywane poprawnie dla białek błędnych i tych mających prawidłowy zwój w obu grupach. Obliczone dane wskazują, że mediana liczby poprawnych kontaktów dla struktur błędnie określonych i już nieaktualnych w PDB („obsolete”) oraz mediana dla struktur „dobrych” są istotnie różne, zwłaszcza dla danych uzyskanych metodą X-ray. Spektakularny wynik uzyskano porównując dane dla grupy (b) – tutaj metoda opracowana przez P.P. Wozniaka i współ. wykazuje 100% skuteczność w określaniu, do której klasy należy dana struktura, czyli czy fold jest poprawny. Autorzy w publikacji dyskutują konkretne przypadki analizy białek i pokazują jak proponowane podejście może (lub nie może) pomóc w ustaleniu właściwej struktury białka. Podoba mi się, że w publikacji nie

ograniczono się do jednego wariantu metody DCA, ale sprawdzono inny wariant (mfDCA) – wyniki były podobne. Jestem ciekaw opinii doktoranta, jakie warunki musi spełniać każda inna metoda przewidywania kontaktów, by można było ją wykorzystać w proponowanym schemacie analizy.

W pracy zarysowano możliwość budowy serwera webowego służącego jako narzędzie do wspomaganie wyboru modeli strukturalnych (foldów) nowych białek. Dane analizowane w pracy są dostępne w sieci dla innych badaczy. Omawiany artykuł zawiera nową ideę – przedstawia propozycję nowej metody analizy struktur złożonych układów biologicznych. Dowodzi biegłości doktoranta w prowadzeniu zaawansowanych analiz przy użyciu środków bioinformatycznych. Wykazuje orientację w problemach analizy struktur białkowych i zdolność do dyskusji konkretnych problemów z dziedziny biologii molekularnej. Praca jest napisana jasno i czytałem ją z dużą przyjemnością.

W pracy (C) [Wozniak PP, Konopka BM, Xu J, Vriend G, Kotulska M, *Forecasting residue-residue contact prediction accuracy*, Bioinformatics. Epub 2017 June 26. doi: 10.1093/bioinformatics/btx416] autorzy zmagają się z jeszcze trudniejszym wyzwaniem: próbą przewidywania „a priori” jakości (czy skuteczności) predykcji kontaktów w białkach. Wiadomo z coraz liczniejszych studiów, że dobre metody przewidywania kontaktów, np. DCA, oparte na skorelowanych mutacjach, mają 40 procentową skuteczność (średnio). Wiemy, że 40 na 100 struktur w analizowanym zbiorze ma kontakty „poprawnie” określone. Problem w tym, że nie wiemy, które z tych 100 są właśnie poprawne, a które nie. Metoda komputerowa opracowana przez autorów, oparta na wykorzystaniu modeli regresyjnych, pozwala przewidzieć dokładność, czy skuteczność (*accuracy*) przewidywania kontaktów w poszczególnych strukturach na poziomie sięgającym 7 %. Kontakty były przewidywane metodami z klasy DCA (gplmDCA oraz PSICONV). Modele regresyjne wykorzystywały ok. 20 parametrów związanych z białkami (m.in. opisujące uliniowanie MSA, aminokwasy dostępne rozpuszczalnikowi, oraz oczywiście przewidywane kontakty) Badano jakie kombinacje parametrów najlepiej użyć, aby jak najlepiej przewidywać kontakty. Do konstrukcji modeli stosowano metody LASSO, TREE, SVM (oparte na metodach sztucznej inteligencji) oraz regresję liniową (wieloparametrową); testowana też metamodele będące kombinacjami prostszych metod. Algorytmy uczyły się na zbiorze 1128 białek z bazy SCOP, zaś przewidywania robiono dla zbioru testowego 446 białek. Analizowano pewien indeks tzw. PPV (*positive predictive value*) w zależności od użytego modelu (łącznie 9 modeli). Najlepszy okazał się model avLSF (kombinacja modeli LASSO, SVM i FOREST), w połączeniu z metodami gplmDCA i PSICONV dawał on niski błąd RMSE na poziomie 7.3 i 6.5 %. Oznacza to, że 200 najsilniejszych kontaktów może być przewidzianych z taką mniej więcej dokładnością. Testy numeryczne pokazały, że modele powinny być optymalizowane osobno dla gplmDCA i dla PSICONV. Zastosowanie zaproponowanego podejścia do jeszcze innego, odmiennego w swojej filozofii narzędzia przewidywania kontaktów – RaptorX, dało nieco gorsze wyniki (RMSE ok. 10%) jednak wykazano, że taka linia oceny jakości przewidywań jest ogólna i ma sens. Nowe wyniki cząstkowe przedstawiono w obszernym materiale dodatkowym towarzyszącym tej publikacji. Nie zauważyłem w pracy (C) opisu jak badano znaczenie poszczególnych parametrów (str. 3410) tj. wpływ na zachowanie modeli regresyjnych.

W mojej ocenie publikacja (C) ma wysoką wartość ze względu na swój analityczny charakter. W bioinformatyce często poszukuje się polepszania wyników czy nowych danych poprzez tworzenie modeli „automatycznych”, ale opartych na przesłankach fizycznych. Tutaj w zaawansowany sposób wykorzystano ogólnie znane algorytmy do rozwiązania trudnego problemu – jak ocenić dokładność indywidualnego przewidywania kontaktów w białkach. Zaproponowane podejście jest oryginalne i pomysłowe. Realizacja badań i wykonanie analiz wymagały wielu żmudnych obliczeń i porównań. Zastosowano różnorodne metody numeryczne, wyniki poddano krytycznej ocenie, wnioski są ostrożne, ale optymistyczne.

Resumując, do obu ostatnich rozdziałów nie mam w zasadzie uwag krytycznych, sądzę, że proces recenzowania w „Bioinformatics” pozwolił wychwycić ewentualne usterki. Strona edytorska rozprawy doktorskiej jest bez zarzutu, przydałby się może wykaz skrótów, których jest niemało.

Po lekturze rozprawy mogę zgodzić się z mgr inż. Pawłem P. Woźniakiem, iż podał on solidne argumenty na słuszność tez postawionych w pracy doktorskiej. Zauważone drobne błędy nie umniejszają mojej bardzo wysokiej oceny całości badań.

KONKLUZJA

Uważam, że mgr inż. Paweł Woźniak realizując swoją rozprawę doktorską wykonał szereg trudnych badań, obliczeń i analiz. Przyczynił się do zaproponowania metod poprawiających jakość przewidywania kontaktów w białkach. Uzyskał tym samym wartościowe i oryginalne wyniki, które zostały opublikowane w czołowym światowym czasopiśmie bioinformatycznym oraz specjalistycznym czasopiśmie poświęconym modelowaniu komputerowemu.

Stwierdzam zatem, że przedłożona mi do oceny rozprawa doktorska mgr. Inż. Pawła P. Woźniaka

pt.

“Supporting Protein Tertiary Structure Modelling with Contact Sites Prediction Methods”

stanowi oryginalne rozwiązanie problemu naukowego i wykazuje ogólną wiedzę teoretyczną doktoranta w naukach technicznych w zakresie dyscypliny biocybernetyki i inżynierii biomedycznej oraz wykazuje umiejętność samodzielnego prowadzenia pracy naukowej, tj. spełnia w pełni wymagania ustawowe [Ustawa z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki ze zmianami późniejszymi (Dz. U. Nr 65, poz. 595, ze zm. w Dz. U. z 2005 r. Nr 164, poz. 1365 oraz w Dz. U. z 2011 r. Nr 84, poz. 455] i zwyczajowe stawiane rozprawom doktorskim. Wnoszę zatem o dopuszczenie Pana magistra inż. Pawła P. Woźniaka do dalszych etapów przewodu doktorskiego.

Ze względu na wyróżniający nakład pracy badawczej związanej z doktoratem, pomysłowość metod, duże znaczenie wyników dla dziedziny modelowania białek, opublikowanie dwóch artykułów w czołowym czasopiśmie „Bioinformatics” (45 pkt) wnoszę o wyróżnienie rozprawy zgodnie z panującymi na Wydziale zasadami.



Wiesław Nowak, prof. zw.