

Recenzja rozprawy doktorskiej magistra Piotra Urbaszka pt. „Przewidywanie sorpcji bromo- i chloro pochodnych trwałych zanieczyszczeń organicznych (TZO) na powierzchni fullerenu C₆₀”

Niniejszą recenzję sporządziłem na podstawie decyzji Rady Dyscypliny Nauki Chemiczne Uniwersytetu Gdańskiego, która w dniu 9 marca 2022 roku powołała mnie na recenzenta ww. rozprawy.

1. Przedmiot recenzji

Przedmiotem recenzji jest rozprawa doktorska mgr Piotra Urbaszka pt.: „Przewidywanie sorpcji bromo- i chloro pochodnych trwałych zanieczyszczeń organicznych (TZO) na powierzchni fullerenu C₆₀” ubiegającego się o stopień doktora nauk chemicznych przez Radą Dyscypliny Nauki Chemiczne Uniwersytetu Gdańskiego. Promotorem rozprawy doktorskiej jest Prof. Tomasz Puzyn a pracę wykonano na Wydziale Chemii UG, w pracowni Chemoinformatyki Środowiska. Problematyka rozprawy nawiązuje do głównego nurtu badawczego realizowanego w grupie badawczej, której przewodzi Prof. Tomasz Puzyn.

2. Ocena merytoryczna rozprawy doktorskiej

Przedstawiona do oceny rozprawa doktorska ma klasyczną formę monografii i zawiera 129 stron maszynopisu. Rozprawę kończy dodatek, w którym przedstawiono osiągnięcia naukowe i organizacyjne autora rozprawy. Po nim pojawia się streszczenie rozprawy w języku polskim i angielskim oraz spis literatury. W ocenie recenzenta zaproponowana struktura monografii nie odbiega znacząco od innych opracowań tego typu, lecz zasadniej byłoby umieścić spis literatury bezpośrednio po zakończeniu treści rozprawy zaś materiały dodatkowe przedstawić później. Rozprawę podzielono na 10 wyraźnie wyodrębnionych rozdziałów: Wstęp, Nanocząstki węglowe wśród nanomateriałów, Trwałe zanieczyszczenia organiczne, Wykorzystanie metod chemii komputerowej w badaniach nanocząstek, Oddziaływania między trwałymi związkami organicznymi a cząsteczką fullerenu – nadzieje i zagrożenia, Hipotezy i cele badawcze, Metodyka badań własnych, Wyniki badań własnych, Podsumowanie i Literatura. Spis literatury tworzy w sumie 187 pozycji, z czego dominująca większość to aktualne, anglojęzyczne prace naukowe opublikowane w ostatniej dekadzie w czasopismach o zasięgu ogólnosiwiatowym. Treść rozprawy wzbogacono 8 tabelami i 25 rysunkami, choć w przypadku tych drugich brakuje zupełnie rysunku 9a-d, na który autor powołuje się na stronie 117 (w rozprawie rysunek 9 nie występuje). W tekście najprawdopodobniej chodzi o rysunek 13a-d. Dodatkowo do pracy dołączono adres repozytorium, w którym znajdują się dane źródłowe i wyniki obliczeń, których, ze względu na ogromną ilość danych numerycznych, umieszczenie w monografii byłoby niemożliwe. Należy w tym miejscu docenić troskę autora o to aby wyniki tak obszernej pracy badawczej były dostępne dla zainteresowanych na zasadzie dostępu do „open-data-library”, co jest aktualnym trendem we współczesnej nauce. Treść pracy poprzedza spis ważniejszych oznaczeń, symboli i skrótów stosowanych w rozprawie choć wydaje się, że nie wykorzystano w pełni potencjału przyjętej tradycji stosowania tego typu rozwiązań. Po pierwsze zaprezentowany zestaw skrótów jest niepełny, a po wtóre

pomimo wyjaśnienia znaczenia niektórych skrótów na stronie 6 w samej treści pracy autor wielokrotnie, ponownie je wyjaśnia, przy czym wyjaśnienie niektórych z nich powtarza się w różnych fragmentach rozprawy w niezbyt spójnej formie. Przykładowo, w spisie akronimów pojawia się wyjaśnienie skrótu LVs – latent vectors (zmienne ukryte). W tekście na stronach odpowiednio 103 i 94 autor ponownie wyjaśnia ten skrót używając raz terminu zmienne utajone, w drugim zaś przypadku zmienne ukryte. Z jednej strony zrozumieliśmy jest dbanie o stosowanie terminów wieloznacznych w celu ograniczenia monotonii i wielokrotnych powtórzeń w treści, z drugiej zaś w tym konkretnym przypadku powinno się stosować konsekwentnie w całej pracy jednolita notację wyjaśnioną w skrócie akronimów na początku. W ocenie recenzenta nie ma potrzeby stosowania dwóch konwencji. Albo zamieszczamy skrót akronimów i nie wyjaśniamy w tekście pracy skrótów dodatkowo, albo stosujemy metodykę przyjętą w publikacjach naukowych, wyjaśniamy akronim w tekście w momencie jego pierwszego użycia a później już stosujemy skrót.

Wstęp w sposób syntetyczny wprowadza czytelnika w problematykę rozprawy nawiązując w formie raczej popularnonaukowej do zakresu stosowalności nanocząstek węglowych i omówienia zagrożeń związanych z obecnością Trwałych Związków Organicznych w środowisku. To połączenie informacji logicznie prowadzi czytelnika do konkluzji, że celem badań będzie prognozowanie właściwości chemicznych blisko dwóch milionów kongenerycznych halogenopochodnych wielopierścieniowych węglowodorów aromatycznych w aspekcie możliwego usuwania ich ze środowiska w efekcie oddziaływań z nanocząstkami węgla. Należy w tym miejscu podkreślić, że tematyka podjęta przez magistra Piotra Urbaszka jest tematyką ze wszech miar aktualną naukowo i dynamicznie się rozwijającą ze względu na współczesne możliwości zastosowania technik „data-science” do analizy zbiorów typ „big-data” metodami „in silico”. Efektywne połączenie metod chemii komputerowej, analizy wieloparametrycznej, technik numerycznych oraz wizualizacji pozwala z powodzeniem rozwijać obszar modelowania wielowymiarowego QSAR (Quantitative Structure-Activity Relationships). Metody QSAR w sposób najbardziej ogólny można zdefiniować jako odwzorowanie przestrzeni właściwości badanych obiektów w przestrzeni deskryptorów strukturalnych. Właściwości wybranych przedstawicieli badanych związków chemicznych wyznacza się eksperymentalnie zaś wartości deskryptorów wyznacza się obliczeniowo. Po uzyskaniu logicznego wyjaśnienia realnych właściwości chemicznych na podstawie zmienności wartości deskryptorów wnioskowanie rozszerza się na związki o podobnej strukturze chemicznej do mierzonych, związki zawierające analogiczne grupy funkcyjne, związki z szeregów homologicznych, etc. Stosowana metodyka rozwijana głównie w badaniach medycznych i środowiskowych pozwala w sposób obliczeniowy oszacować zachowanie ogromnych ilości związków bez potrzeby wykonywania kosztownych i czasochłonnych eksperymentów laboratoryjnych. W przypadku badań magistra Piotra Urbaszka opisaną metodykę wykorzystano do modelowania sorpcji TZO na powierzchni fullereny C₆₀, przy czym parametry sorpcji wyznaczono również w eksperymentach *in silico*. Powracając do analizy Wstępu autor przedstawił w nim ponadto krótki opis tematyki umieszczonej w każdej kolejnej części rozprawy.

W rozdziale II (strony 11-26) autor opisał zakres stosowalności nanotechnologii oraz scharakteryzował nanocząstki węglowe, jako struktury o wzrastającym potencjale wykorzystania komercyjnego. W tej części autor opisał także budowę i właściwości fullereny C₆₀ (dalej w skrócie C₆₀), który wykorzystywał w badaniach. Scharakteryzował też główne mechanizmy potencjalnych efektów toksycznych związanych z obecnością nanocząstek w środowisku. Warto w tym miejscu podkreślić, że autor przedstawił również dowody na niespójność aktów prawnych, na podstawie których nanocząstki węglowe wprowadza się do obrotu komercyjnego. Dowodzi to ogromnej potrzeby uszczelnienia prawodawstwa, które powinno opierać się na wynikach badań naukowych, wykonywanych w oparciu o powszechnie zaakceptowane i przyjęte metodyki. Rozdział II to dobrze i ciekawie napisany fragment pracy, w którym w sposób zrównoważony przedstawiono wiele wątków opisanych powyżej. Drobną

uwaga związana z tym fragmentem dotyczy braku informacji o źródłach pochodzenia grafik wykorzystanych na rysunkach. Poza kilkoma wyjątkami nie wiadomo czy są to grafiki pobrane z innych źródeł czy też samodzielnie wykonane przez autora.

W rozdziale III (str. 27-44) autor przedstawił ogólną charakterystykę związków należących do grupy trwałych zanieczyszczeń organicznych z uwzględnieniem badanych w ramach rozprawy kongenerów. Autor opisał komercyjne zastosowania TZO, ich właściwości fizykochemiczne oraz mechanizm wywoływania przez nie efektów toksycznych. Podobnie jak w poprzedniej części, autor nawiązał do prawodawstwa związanego ze stosowaniem TZO w aspekcie bezpieczeństwa chemicznego. Rozdział III to ponownie precyzyjnie przygotowany tekst, poprawny pod względem gramatycznym i merytorycznie spójny. Jedyna uwaga dotyczy formy Tabeli 2. W opinii recenzenta można by było ją przygotować w nieco innej formie np. prezentując w pojedynczych, oddzielonych kolumnach dane dotyczące współczynników podziału badanych TZO, zakresu stosowalności czy też nawiązania do poszczególnych aktów prawnych czy rekomendacji w nawiązaniu do ich toksyczności.

W rozdziale IV (str. 45-57) autor przedstawił metody chemii obliczeniowej oraz chemoinformatyki wykorzystywane w badaniu oceny ryzyka dla nowych materiałów, ze szczególnym uwzględnieniem metod ilościowego modelowania zależności pomiędzy strukturą w właściwościami bądź aktywnością biologiczną. Z treści rozdziału dowiadujemy się przede wszystkim, że stosowanie metod QSAR do badań nad nanocząstkami nie należy do wyzwań trywialnych ze względu na małą dostępność i niską jakość danych eksperymentalnych, brak odpowiednich deskryptorów opisujących specyfikę nanostruktur węglowych, ograniczoną wiedzę o oddziaływaniu nanocząstek z układami biologicznymi, białkami, lub innymi związkami występującymi w środowisku jak również brak ujednoliconych wytycznych i procedur modelowania, uwzględniających zmienność strukturalną wynikającą z rozmiaru nanocząstek oraz zmiennych warunków zewnętrznych wpływających na agregację i aglomeracje nanocząstek. Biorąc chociażby pod uwagę jedynie fakty przytoczone powyżej wyniki badań przedstawione w pracy stanowią istotny wkład w rozwój nauki i znacząco poszerzają aktualny stan wiedzy w tematyce badań. W rozdziale IV autor omówił skrótowo etap zbierania i gromadzenia danych, podziału danych na zbiory treningowe, uczące i walidacyjne, podstawy najbardziej popularnych algorytmów wspierających ww. podziały jak również scharakteryzował rodzaje deskryptorów struktury chemicznej badanych związków z uwzględnieniem podziału na deskryptory konstytucyjne i strukturalne. Cenne, w aspekcie kolejnych rozdziałów było nawiązanie do zagadnień związanych z kalibracją modelu QSAR, definiowania dziedziny modelu oraz statystyk, na podstawie których można ocenić jakość dopasowania, elastyczność i zdolności predykcyjne modeli QSAR/QSPR. Treść rozdziału IV jest klarowna, przejrzysta i nie nasuwa uwag o charakterze dyskusyjnym. Jedyną uwagę o charakterze stylistycznym można mieć do pierwszego zdania w paragrafie dotyczącym zbierania danych. Wydaje się, że zdanie cyt...."Zbieranie i uporządkowanie danych jest pierwszym, niezwykle istotnym etapem modelowania QSAR/QSPR jest zbieranie dostępnych danych, ich zmierzenie lub też ich uporządkowanie dla modelowanej wielkości (ang. endpoint)...." zawiera pewne powtórzenie myśli, przez co staje się trudne w interpretacji. Należało także nieco bardziej precyzyjnie wyjaśnić znaczenie terminu endpoint.

Rozdział V to dwustronicowy tekst, w którym autor opisał główne nadzieje i zagrożenia wynikające z potencjalnych zastosowań fullerenów do badań i aplikacji komercyjnych z TZO. Wydaje się, że treść tego rozdziału można było wkomponować w rozdział IV, tym bardziej, że wiele z opisywanych oddziaływań nawiązuje do tematyki deskryptorów struktury, o których była mowa wcześniej.

W rozdziale VI (str. 61-63) autor przedstawił 4 hipotezy badawcze, których weryfikacja uwzględniła konieczność realizacji 7 celów pracy. Za główne hipotezy przyjęto stwierdzenia, że:

- w oparciu o reprezentatywnie wybraną podgrupę związków oraz zestaw wyznaczonych metodami półempirycznymi deskryptorów struktury możliwe jest określenie podobieństwa halogeno- podstawionych kontenerów TZO w wielowymiarowej przestrzeni z wykorzystaniem metod analizy podobieństwa,
- pomimo pewnych różnic strukturalnych kongenery będą wykazywały podobieństwo związane z miejscem podstawienia i ilości halogenów w przestrzeni obliczonych deskryptorów struktury,
- dzięki zastosowaniu metod „*in silico*” możliwe jest przewidywanie wpływu podstawników halogenowych na wartości energii oddziaływań z C₆₀ a tym samym analizę procesu sorpcji bromowanych i/lub chlorowanych kongenerów dibenzo-p-dioksyny z C₆₀,
- możliwe jest wykorzystanie modeli QSAR/QSPR do oceny wartości energii oddziaływań pomiędzy fullerenem C₆₀ a wybranymi halogenowanymi TZO o różnej budowie szkieletu węglowego.

W rozprawie za główne cele badawcze uznano:

- wybór reprezentatywnej grupy związków ze zbioru ponad 1,8 mln halogenopochodnych (gdzie X – Cl, Br) kongenerów TZO różniących się budową szkieletu węglowego w celu ułatwienia wizualizacji otrzymanych wyników,
- zbadanie podobieństw w zbiorze reprezentatywnie wybranych TZO w przestrzeni 26 deskryptorów struktury obliczonych półempiryczną metodą PM6 z wykorzystaniem technik analizy podobieństwa,
- opracowanie metodyki prowadzenia obliczeń energii oddziaływań TZO a C₆₀ na przykładzie oddziaływań C₆₀ a 2,3,7,8-tetrachlorodibenzo-p-dioksyny
- wyznaczenie różnymi metodami energii oddziaływań C₆₀ a reprezentatywnym zbiorem TZO,
- kalibracji i walidacji modelu QSAR celem analizy jego użyteczności do przewidywania energii oddziaływań między C₆₀ a halogeno podstawionymi kontenerami dibenzo-p-dioksyny,
- zbadanie wpływu wielkości szkieletów węglowych 23 struktur niepodstawionych na oddziaływanie z cząsteczką TZO metodami chemii teoretycznej
- przewidywanie energii sorpcji za pomocą modelu QSAR dla pierwotnej grupy ponad 1,8 mln TZO.

W ocenie recenzenta cele badawcze wybrano adekwatnie do postawionych hipotez aczkolwiek dwie pierwsze hipotezy wydają się już zweryfikowanymi faktami naukowymi dla milionów innych związków chemicznych i nie stanowią wobec tego hipotez, które należy w dalszym ciągu weryfikować. Czy w opinii magistra Piotra Urbaszka należałoby się w tym konkretnym przypadku spodziewać tego, że podobieństwo budowy strukturalnej nie znajdzie tym razem odzwierciedlenia w wynikach analizy podobieństwa? Jeśli tak, jakie mogłyby być tego przyczyny?

Kolejny rozdział rozprawy to opis metodyki badań własnych (str. 64-78). W pierwszym etapie metodami kombinatorycznymi wygenerowano wszystkie możliwe Cl- i Br- podstawione kongenery podstawowych 23 struktur chemicznych wybranych do badań. Do stworzenia listy możliwych kongenerów zastosowano oprogramowanie ConGENER. Następnie każdy z kongenerów scharakteryzowano za pomocą 26 deskryptorów przy czym deskryptory 3D wyznaczono w oparciu o półempiryczną metodę PM6. Jakim kryterium stosowano przy wyborze konkretnej listy 26 deskryptorów? Z globalnej biblioteki kongenerów z zastosowaniem algorytmu Kennarda-Stonea wybrano 2% związków reprezentatywnych z każdej z 23 podstawowych rodzin. W efekcie uzyskano macierz danych o wymiarach 36811 x 26. W związku z tym, że w opisie metodyki badań własnych pojawia się wiele informacji dotyczących redukcji zbioru danych zdecydowanie brakuje schematu, który ułatwiłby czytelnikowi śledzenie poszczególnych kroków analizy a przede wszystkim stosowanych metod zmniejszania ilości struktur chemicznych analizowanych danym momencie. 26 deskryptorów przetworzono za pomocą metody analizy głównych składowych aby finalnie uzyskać z nich 5 głównych

składowych, które sumarycznie wyjaśniały 82,9% wariacji. W tym miejscu chciałbym zapytać czy wraz z wyborem 26 deskryptorów dokonano analizy tego, czy część z nich z definicji nie jest skorelowana ze sobą? Dlaczego nie zbadano kryterium stosowalności techniki PCA, np. za pomocą testu Bartletta? Dlaczego nie przedstawiono wartości własnych głównych składowych i dlaczego nie przedstawiono wariacji wyjaśnionej przez każdą z nich. Na jakiej podstawie dokonano wyboru 5-ciu głównych składowych? Jakie kryterium wyboru liczby składowych zastosowano na tym etapie? W kolejnych etapach analizy wykonywano optymalizację metodyki prowadzenia obliczeń energii oddziaływań pomiędzy wybranymi strukturami TZO a C₆₀, między innymi z metodą półempiryczną PM6, oraz metodami opartymi o teorię funkcjonały gęstości z wykorzystaniem funkcjonałów B3LYP i M06-2X. Ostatni funkcjonał wykorzystano do wyznaczenia energii oddziaływań kompleksów PXDD@C60. Funkcjonał M06-2X w bazie People'a 6-31 uzupełniono funkcjami polaryzacyjnymi i dyfuzyjnymi, a dla przyspieszenia obliczeń zastosowano metodę ONIOM. W kolejnych etapach kalibrowano, budowano i walidowano odpowiednie modele QSPR z zastosowaniem metody PLS z algorytmem genetycznym oraz podziałem zbioru danych na zbiory: modelowy, treningowy i testowy. W rozdziale VII zdecydowanie brakuje schematu procedury prowadzonych analiz np. z uwzględnieniem odmiennej kolorystyki dla etapu PCA oraz QSPR. Lekturę rozdziału znacząco utrudnia przemieszanie metodyki badań własnych z opisem teoretycznym wykorzystanych w nim metod. Wydaje się, że opis teoretyczny metod powinien stanowić odrębny rozdział np. „Materiały i metody”, w którym więcej uwagi poświęcono by metodom skomplikowanym i unikalnym, zaś zdecydowanie mniej metodom standardowo stosowanym z analizie wielowymiarowej, jak np. analiza składowych głównych. Pomimo podania precyzyjnych odnośników literaturowych i częściowego wyjaśnienia przyczyn zastosowania niektórych metod czytelnik może czuć się nieco zagubiony. Przykładowo, nie wyjaśniono, choćby jednym zdaniem na czym polegają metody AM1, RM1, PM3, PM6 i wreszcie PM7 zaś bardzo szczegółowo przedstawiono podłoże merytoryczne analizy głównych składowych [str. 74]. Korzystając z okazji prowadzenia dyskusji z doktorantem chciałbym zapytać czy schemat przedstawiony na stronie 74 w rzeczywistości nie odnosi się do techniki analizy czynnikowej, w której jedną z metod wyodrębniania czynników może być metoda głównych składowych? Wydaje się bowiem, że w pierwotnej formie analizy PCA nie zakłada się uzyskania macierzy reszt od modelu. Poczucie lekkiego zagubienia wzmocniają zdania takie jak: „...PCA została wykorzystana do analizy zmienności zbioru 1 840 951 kongenerów ...[str. 75]”. Jest to przecież skrót myślowy, gdyż z opisu wynika, że do badań wyselekcjonowano 36811 związków. Bardziej precyzyjne byłoby określenie, że wyniki z analizy PCA uogólniono na zbiór ponad 1,8 kongenerów.

Początkowo, bardzo pozytywne odczucie dotyczące jakości „przeptywu myśli” w rozprawie znacząco osłabło w trakcie lektury rozdziału VIII (Wyniki badań własnych, str. 80-124). W rozdziale dochodzi do wielokrotnych powtórzeń prezentowanych treści (np. ponowne stwierdzenie o zastosowaniu oprogramowania ConGENER do typowania wszystkich kombinacji kongenerów badanych związków, wcześniej str. 65). Opis redukcji liczby analizowanych kontenerów za pomocą algorytmu Kennarda-Stonea stanowi kolejny przykład takiego powtórzenia [str. 81]. Autor rozprawy nie koncentruje się często na wynikach i ich analizie ale powraca do metodyki. To znacząco utrudnia lekturę rozprawy. Wracając do propozycji tworzenia schematu prowadzonych analiz można by było przedstawić na nim kolejny etap redukcji danych do PCA, czego efektem było uzyskanie macierzy 184 x 5 z początkowej 36811 x 26. Dopiero w tym fragmencie tekstu czytelnik dowiaduje się np. o zastosowanej metodzie Malinowskiego do szacowania bezwzględnej wartości ładunków czynnikowych zmiennych [str. 82], które będą brane pod uwagę w trakcie oceny merytorycznej głównych składowych oraz o procentowym udziale wariacji wyjaśnianej przez każdą z nich [str. 81]. Przywołany już wcześniej rysunek 13a-d przedstawiony w wersji czarno-białej jest bardzo trudny w interpretacji. Na rysunkach 13b-d nie widać rozdzielenia poszczególnych Cl- i Br- pochodnych a śledzenie „rozdzielania” kongenerów na podstawie interpretacji znaczenia głównych składowych opisanych w tekście jest bardzo trudne. Zdecydowanie polecałbym zastosowanie innych wizualizacji w których za pomocą

odpowiednio dobranych skali kolorystycznych punktów przedstawionych na wykresie różnicowaniu podlegałyby np. związki o wzrastającej masie cząsteczkowej czy liczbie atomów wodoru w cząsteczce (odnośnie interpretacji PC1). Bardzo proszę o kilka dodatkowych słów wyjaśnienia, które dotyczą tego, że brak zróżnicowania rodzin kontenerów jest pozytywny i potwierdza słuszność obranego w dalszej części kierunku prac badawczych. Czy można to interpretować w taki sposób, że np. antraceny będą zachowywały się w stosunku do C₆₀ analogicznie jak tioksanteny? Czy w takim razie czynnikiem decydującym o ogólnych relacjach TZO i C₆₀ nie będzie mniej lub bardziej rozbudowana struktura aromatyczna?

W kolejnych fragmentach autor opisuje wyniki badania wpływu podstawienia atomami bromu i/lub chloru na energię oddziaływań dibenzo-p-dioksyn z powierzchnią fullerenu C₆₀. Autor ponownie podaje szereg danych metodologicznych, np. wizualizację pozycji startowych aby finalnie stwierdzić, że optymalna odległość wynosi 3Å, którą stosuje w kolejnych obliczeniach. Brakuje informacji o tym dlaczego uznano tę odległość za optymalną. Opis wyników uzyskanych w efekcie kalibracji i budowy modelu QSPR dowodzi, że doktorant opanował warsztat prezentacji wyników eksperymentów dotyczących QSPR lecz precyzyjne oddzielenie wyników od metodologii znacząco podniosłoby ocenę końcową. W wielu fragmentach autor podaje wartości uzyskane w jednym z badań, które wpływają na możliwości interpretacji kolejnych. Ponieważ w ramach testowania modelu QSPR wykorzystano w PLS w miejsce zmiennych pierwotnych wyodrębnione wcześniej główne składowe zdecydowanie zalecam początkowe sprawdzenie kryterium stosowalności PCA, o którym wspominałem powyżej. W kilku fragmentach zestawienie treści prezentowanych w opisie z wizualizacją stanowi ogromne wyzwanie dla czytelnika. Do przykładów tego typu należy analiza danych przedstawionych na rysunku 17a-c. Dotyczy on projekcji rozmieszczenia kongenerów ze zbioru testowego i walidacyjnego w przestrzeni zmiennych ukrytych z PLS. W tekście autor wyjaśnia znaczenie merytoryczne „latent variables” lecz odpowiednie punkty na wykresie okraszono jedynie etykietą niebieskiego kwadratu lub czerwonego trójkąta. Nie wiadomo np. które kongenery z tabeli przedstawionej na stronie 92 zestawionej na stronie charakteryzują się wysokimi, dodatnimi wartościami czynnikowymi dla LV1. Taki opis i przedstawienie danych sprawia, że czytelnik musi niejako wierzyć autorowi, że odnaleziono opisywane zależności. Analogiczne przemyślenia dotyczą analizy wyników walidacji modelu nano-QSPR, który stworzono do przewidywania energii oddziaływania między Br- i Cl- kongenerami TZO a powierzchnią fullerenu C₆₀ (w szczególności rysunku 24 w zestawieniu z tekstem). Poza wizualną oceną oddalenia niektórych punktów od głównej chmury czytelnik nie może ocenić, dla których konkretnych kongenerów ta zależność występuje. W ujęciu ogólnym treść rozdziału VIII mogłaby być nieco bardziej klarowna, ponieważ bardzo pozytywne odczucie dotyczące warsztatu metodologicznego stosowanego przez doktoranta blaknie wraz z lekturą kolejnych paragrafów rozdziału VIII, w których większość informacji należy przyjąć na wiarę a mniej można zweryfikować na wykresach.

Po zakończeniu lektury rozprawy doktorskiej muszę stwierdzić, że mgr Piotr Urbaszek zrealizował bardzo ambitny i skomplikowany program badawczy. Wysoko oceniam poziom merytoryczny zastosowanej metodologii i wieloaspektowość badań, choć jak autor sam podkreśla, budowanie i kalibrowanie odpowiednich modeli QSPR przewidujących szukane wartości deskryptorów dla milionów związków na podstawie kalibracji modeli z wykorzystaniem kilkudziesięciu z nich musi charakteryzować się pewnymi niedoskonałościami. Pod względem merytorycznym zastosowany warsztat badawczy nie budzi zastrzeżeń a jego opis, w większości przypadków, dowodzi, że autor dokonał świadomego wyboru poprawnych narzędzi i technik. Jedyne fragment, w którym nie podzielam stwierdzenia autora dotyczy opisu przedstawionego na stronie 116, w którym autor pisze, że przy budowie modelu nano-QSPR wykorzystał maksymalną możliwą do interpretacji ilość zmiennych ukrytych. Autor stwierdził, że wybrane wyjaśniają 82% zmienności. A co z pozostałymi 18%?

3. Wnioski końcowe

Podsumowując ocenę merytoryczną z całym przekonaniem pragnę stwierdzić, że autor w pełni i konsekwentnie zrealizował sprecyzowane cele badawcze a poprzez to, wniósł wiele elementów nowości naukowej do badań QSPR. Do szczególnie wartościowych zaliczam parametryzację modelu nano-QSPR TZO@C60 i wykazanie możliwości jego zastosowania do modelowania wartości energii oddziaływań kongenerów TZO z powierzchnią fullereny C60 o różnej budowie szkieletowej. W mojej ocenie wyniki uzyskane w tym wątku mogą doprowadzić do świadomego podjęcia decyzji o potencjalnym zastosowaniu nanocząstek węglowych w aplikacjach środowiskowych, szczególnie gdy uwzględnimy fakt, że wpływ nanostruktur węglowych samych w sobie na środowisko nie jest do końca jednoznacznie poznany. Lektura rozprawy dowodzi także, że mgr Piotr Urbaszek posiadał biegłą znajomość stosowania metody QSPR oraz prawidłowo stosuje metodykę charakterystyczną w tego typu badaniach. Na uwagę zasługuje także dobra znajomość metod powszechnie stosowanych w analizie wielowymiarowej, jak PLS, PCA, regresja wieloraka co stanowi dodatkowy pozytywny aspekt, gdyż sprawia, że mgr Urbaszek może być cennym źródłem specjalistycznej wiedzy także w innych projektach z zakresu „data mining”.

Wysoką ocenę rozprawy doktorskiej samej w sobie należy uzupełnić wysoką oceną aktywności publikacyjnej doktoranta, który już w chwili obecnej jest autorem lub współautorem wielu publikacji docenionych w międzynarodowym obiegu naukowym a przekładających się na wskaźnik Hirscha wynoszący 6-7 (w zależności od bazy danych) i łączną liczbę cytowań bez autocytowań oscylującą w rejonie 300.

Z całym przekonaniem uważam, że przedstawiona mi do recenzji rozprawa doktorska pt.: „Przewidywanie sorpcji bromo- i chloro pochodnych trwałych zanieczyszczeń organicznych (TZO) na powierzchni fullereny C₆₀” spełnia wymogi stawiane rozprawom doktorskim przez Ustawę o Stopniach i Tytule Naukowym. Pojawia się w niej istotny element nowości naukowej, a w dalszej perspektywie nawet szansa wykorzystania wyników w praktyce przez podmioty rozważające zastosowanie technologii nanocząstek węglowych do usuwania TZO ze środowiska. W związku z powyższym wnoszę do Rady Dyscypliny Nauki Chemicznej Uniwersytetu Gdańskiego o dopuszczenie pana mgra Piotra Urbaszka do dalszych etapów przewodu doktorskiego.

Aleksander Ostel

3 maja 2022 r.