

Warszawa, 2.10.2012

Katarzyna Jarzemska  
Pracownia Krystalochemii  
Wydział Chemii  
Uniwersytet Warszawski

Autoreferat rozprawy doktorskiej pt.:  
**„ROZSZERZENIE BANKU ASFERYCZNYCH PSEUDOATOMÓW W KIERUNKU  
KWASÓW NUKLEINOWYCH I JEGO ZASTOSOWANIE W BADANIACH STRUKTURY,  
ENERGII I ROZKŁADU GĘSTOŚCI ELEKTRONOWEJ CZĄSTECZEK.”**

**(„EXTENSION OF THE ASPHERICAL PSEUDOATOM DATABANK TOWARDS NUCLEIC  
ACIDS AND ITS APPLICATION IN STRUCTURAL, CHARGE DENSITY, AND ENERGY  
STUDIES.”)**

*Promotor pomocniczy: Dr Paulina M. Dominiak  
Promotor: Prof. dr hab. Krzysztof Woźniak*

Mój projekt doktorski składa się z dwóch wyraźnie zarysowanych części. Pierwsza z nich obejmuje modyfikację i rozszerzenie banku asferycznych pseudoatomów UBDB (University at Buffalo Databank) wraz ze sprzężonym programem *LSDB*, a także testy stosowalności opracowywanej metody do szacowania właściwości elektrostatycznych molekuł. Druga część mojej pracy opiera się na użyciu nowej wersji banku w trzech różnych celach: (1) w celu rekonstrukcji rozkładu gęstości elektronowej makromolekuł, (2) jako źródła asferycznych atomowych czynników rozpraszania, a w końcu, (3) w celu przygotowania wstępnego modelu multipolowego w eksperymentalnych badaniach gęstości elektronowej. Zastosowanie banku UBDB jest tu punktem wyjścia do dalszych, szeroko zakrojonych rozważań z obszarów biochemii i inżynierii krystalicznej. Poglądowy plan mojej pracy doktorskiej przedstawiony jest na poniższym schemacie.



**Schemat 1.** Ilustracja przebiegu projektu doktorskiego.

W wyniku realizacji projektu przygotowałam nową wersję banku UBDB, rozszerzoną o atomy występujące w kwasach nukleinowych, DNA i RNA, a także innych ważnych biologicznie molekułach. Zmodyfikowałam również istniejące definicje atomowe. Nowy bank UBDB2011 zawiera obecnie ponad 200 różnych typów atomowych wymodelowanych na podstawie ponad 600 cząsteczek organicznych. Dodatkowo, równoległe do pracy nad bankiem, aktualizowałam i poprawiałam program *LSDB* odpowiedzialny za przypisywanie atomom lokalnych układów współrzędnych i nakładanie parametrów multipolowych przechowywanych w banku na wybraną molekułę o zadanej geometrii. Do bieżącej wersji kodu włączone są również zaktualizowane 'średnie neutronowe' długości wiązań z wodorem, X-H (X – atom niewodorowy), zgodnie z najnowszą publikacją Allen *et al.* 2010. Oczywiście program jest w pełni kompatybilny z istniejącą wersją banku. Nowa wersja *LSDB*, oprócz plików wejściowych do pakietu *XD*, automatycznie produkuje także pliki w formacie coraz powszechniej używanego pakietu *MOPRO*.

Mając rozszerzoną wersję banku, UBDB2011, przeprowadziłam testy jego stosowalności do szacowania właściwości elektrostatycznych cząsteczek. Weryfikację przeprowadziłam na podstawie analizy elektrostatycznej energii oddziaływania zestawu kompleksów zawierających zasady azotowe lub reszty aminokwasowe. Okazało się, że metoda UBDB+EPMM (EPMM, *ang.* Exact Potential Multipole Method) w sposób satysfakcjonujący odzwierciedla wartości energii badanych układów w odniesieniu do obliczeń teoretycznych (*ab initio* i DFT). Korelacja między metodami była wysoka ( $R^2 > 0.9$ , RMSD = 3.7 kcal·mol<sup>-1</sup>), natomiast zmienność energii zachowana. Ważną cechą metody bazującej na banku jest również uwzględnianie kierunkowości oddziaływań, co jest niezwykle ważne, a zaniedbywane przez metody oparte na ładunkach punktowych.

Podjęcie UBDB+EPMM ma jednak kilka ograniczeń. Parametry przechowywane w banku nie odwzorowują efektów pola krystalicznego czy zmienności konformacyjnej molekuł. Dodatkowym ograniczeniem banku jest adekwatność definicji typów atomowych oraz metodologia zastosowana do obliczania parametrów gęstościowych (użyta baza funkcyjna, metoda DFT, obcięcie transformaty Fouriera, ograniczenia formalizmu Hansena-Coppensa itd.). W konsekwencji dokładność metody UBDB+EPMM wynosi około 5 kcal·mol<sup>-1</sup>. Ponadto bank UBDB2011 nie nadaje się do opisu niestandardowych molekuł, czy układów zawierających szereg sprzężonych wiązań podwójnych, co pokazałam na przykładzie cząsteczki amfoterycyny B.

Przygotowany i przetestowany bank pseudoatomów otworzył mi drogę do drugiej części doktoratu. Jak już wspominałam we wstępie, w tej części mojej pracy można wyodrębnić trzy oddzielne analizy. Pierwsza z nich jest poświęcona badaniu elektrostatyki oddziaływań w kompleksach neuraminidazy wirusa grypy z inhibitorami z wykorzystaniem UBDB2011, ze szczególnym uwzględnieniem udziału cząsteczek wody. W kolejnej zajęłam się tematyką serii pochodnych uracylu. Tu bank służył, jako źródło asferycznych czynników rozpraszania atomowego i został użyty w stosunkowo nowym typie udokładnienia struktury tzw. udokładnieniu TAAM (*ang.* Transferable Aspherical Atom Model). Ostatnie badanie polegało zaś na zastosowaniu banku do przygotowania startowego modelu gęstości elektronowej. Tu zajęłam się dwoma

układami, dla których uzyskałam dane wysokorozdzielcze, tj. 6-metylo-2-tiouracylem i ko-kryształem komplementarnej pary zasad DNA 1-metylotyminy z 9-metyloadeniną.

Badania układu neuraminidaza-inhibitor pokazały potencjał banku w analizie kompleksów biochemicznych oraz aktywności leków. Dzięki użyciu metody UBDB+EPMM możliwe było oszacowanie wkładów do całkowitej elektrostatycznej energii oddziaływania pochodzących od pojedynczych grup funkcyjnych ligandu, czy też poszczególnych aminokwasów części białkowej. W ten sposób łatwo jest zidentyfikować miejsce aktywne, czy kluczowe grupy funkcyjne limitujące, bądź wspomagające wiązanie inhibitora. Moim zadaniem było zbadanie wpływu cząsteczek wody na trwałość układu neuraminidaza-inhibitor. Udało mi się ustalić poglądowy udział rozpuszczalnika w stabilizacji kompleksu. To przyczyniło się po części do wytłumaczenia różnic w aktywności inhibicyjnej wybranych ligandów, która w niektórych przypadkach nie korelowała bezpośrednio z elektrostatyczną energią wiązania samego ligandu z białkiem. Takie badania mają większe znaczenie dla układów naładowanych, gdzie elektrostatyka przeważa w zdecydowany sposób nad innymi rodzajami oddziaływań. Takim właśnie układem był badany kompleks.

Analiza serii 10 pochodnych uracylowych (uracylu, 1-metylouracylu, 1,5-dimetylouracylu, 2-tiouracylu, 4-tiouracylu, 2,4-ditiouracylu, 1-metylo-4-tiouracylu, 2-tio-5-metylouracylu, 2-tio-6-metylouracylu, 5-fluorouracylu) pokazała zaś wyższość udokładnienia typu TAAM nad standardowym udokładnieniem struktury w sferycznym modelu niezależnych atomów IAM (*ang.* Independent Atom Model). Badania energetyczne i geometryczne wykazały, że uzyskane przeze mnie geometrie cząsteczek z wykorzystaniem banku są bardzo zbliżone zarówno do tych uzyskanych z badań gęstościowych czy neutronowych, jak również do struktur okresowo zoptymalizowanych w pakiecie *CRYSTAL*. Również w przypadku energii kohezji, czy oddziaływania motywów strukturalnych, trendy są zgodne z danymi teoretycznymi, podczas gdy wyniki dla struktur uzyskanych w metodzie IAM silnie od nich odbiegają. Moja analiza zaowocowała siedmioma dobrej jakości strukturami badanych układów, w tym jednej zupełnie nowej dla 4-tiouracylu. Wszystkie związki, poza 2,4-ditiouracylem, tworzą warstwy molekularne utrzymywane przez wiązania wodorowe i wzajemnie oddziaływujące za pomocą kontaktów  $\pi$ -elektronowych i z przewagą oddziaływań dyspersyjnych. Różnica energii kohezji między najstabilniejszym termodynamicznie 5-fluorouracylem a najmniej korzystną siecią krystaliczną 1-metylo-4-tiouracylu wynosi około  $40 \text{ kJ}\cdot\text{mol}^{-1}$ . Analiza energii motywów strukturalnych badanych związków pokazała wpływ grupy metylowej na stabilizację oddziaływań międzypłaszczyznowych oraz podstawnika siarkowego w pozycji 4 na jakość struktury. Również te dwa czynniki istotnie wpływają na wartości indeksów aromatyczności, *NICS* i *HOMA*, które jednak nie są wzajemnie zgodne. Wielkość energii deformacji molekuly przy przeniesieniu jej z próżni do sieci krystalicznej zależy natomiast od siły i liczby oddziaływań, w które zaangażowana jest dana cząsteczka. Dodatkowo analizowanym problemem okazał się szczególny motyw tworzony przez cząsteczki w strukturze 5-fluorouracylu. Tu atomy fluoru grupują się po cztery w tetramerycznym motywie. Na podstawie banku została odtworzona gęstość

deformacyjna wskazująca na potencjalny udział stabilizujących kontaktów F...F typu 'lump to hole'.

Ostatnim tematem poruszonym w mojej pracy doktorskiej jest analiza rozkładu gęstości w kryształach zasad azotowych. Badania 6-metylo-2-tiouracylu pokazały, że bank UBDB może być wykorzystany nie tylko, jako startowy model gęstości elektronowej, ale może również być pomocny w przypadku analizy nieporządku w sieci. W moim badaniu UBDB2011 posłużył do wymodelowania problematycznego atomu siarki, oraz przyczynił się do wysnucia hipotezy o obecności okso-tiolowej formy tautomerycznej w badanej strukturze, tłumaczącej systematyczny błąd rejestrowany podczas analizy danych gęstościowych i zagadkowe piki resztowe wokół siarki. Badanie pokazało również, jak ostrożnym należy być przy interpretacji danych i dobieraniu modelu. Przypadek 6-metylo-2-tiouracylu i regularne kryształy, które tworzy, umożliwił znalezienie zależności między strukturą wewnętrzną kryształu a jego cechami makroskopowymi. Badanie pokazało także zgodność energii oddziaływań wodorowych uzyskanych na podstawie analizy topologicznej danych gęstościowych z tymi uzyskanymi z obliczeń teoretycznych.

Podobną analizę przeprowadziłam dla ko-kryształu 9-metyloadeniny z 1-metylotyminą. Para zasad w kryształce występuje w nietypowym dla struktur kwasów nukleinowych kontakcie typu Hoogsteen-Watson-Crick (HW). Okazuje się, że pochodne adeniny wiążąc się w kryształce z modyfikacjami uracyli wykazują tendencję do tworzenia preferencyjnie właśnie tego motywu, podczas gdy nawet w cząsteczce RNA, o wiele bardziej labilnej niż DNA, tylko niewielki procent par zasad typu A:U występuje w takiej konfiguracji. Jednakże z obliczeń energetycznych wynika, że orientacja zasad typu HW jest nieznacznie stabilniejsza niż standardowe ustawienie typu Watsona-Cricka. Ponadto charakterystyka topologiczna rozkładu gęstości elektronowej układu jest zgodna z wynikami uzyskanymi dla analogicznych związków, np. 1-metylotymina tylko nieznacznie różni się od 6-metylouracylu przy porównaniu wspólnych fragmentów molekularnych. Podobnie, jak w 6-metylo-2-tiouracylu, również i w przypadku ko-kryształu energia wiązań wodorowych jest dobrze estymowana na podstawie parametrów topologicznych. Co więcej, model uzyskany bezpośrednio z banku, mimo że w detalach różny od eksperymentalnego rozkładu gęstości uzyskanego dla ko-kryształu, odzwierciedla satysfakcjonująco wielkość tych oddziaływań.

Podsumowując, moja praca doktorska dostarczyła pełniejszej wersji narzędzi do opisu gęstości elektronowej cząsteczek, które z powodzeniem mogą być wykorzystywane zarówno do badania elektrostatyki układów makromolekularnych, jak i do poprawiania geometrii udokładnianych struktur czy w analizie gęstości elektronowej. Ponadto przeprowadzone badania uwypuklają zależności między geometrią układów, motywami strukturalnymi, energetyką sieci a morfologią. Większość wyników została już opublikowana w liczących się czasopismach (w sumie do tej pory jestem współautorką 11 publikacji), a ostatnie manuskrypty są w przygotowaniu. Kierunkami rozwoju wskazanymi przez wyniki moich badań może być analiza kompleksów makromolekularnych ważnych biologicznie, rozpowszechnianie użycia banku w udokładnianiu strukturalnym i jego dalszy rozwój

lub też poszukiwanie lepszych modeli gęstości elektronowej w celu wyeliminowania przytoczonych wcześniej ograniczeń metody. Dodatkowo, ciekawą byłaby szersza analiza energetyczna, geometryczna i środowiskowa różnych połączeń zasad azotowych, a także ich oddziaływań z wodą, jonami metali i resztami cukrowymi. Przyczyniłoby się to do lepszego zrozumienia mechanizmów powstawania konkretnych kontaktów międzycząsteczkowych, także tych ważnych farmaceutycznie.