

Christian Jelsch
Director of Research CNRS
CRM2 Université de Lorraine, Faculté des Sciences et Technologies
Laboratoire Cristallographie Résonance Magnétique et Modélisations
Vandoeuvre-les-Nancy France

Vandoeuvre lès Nancy, February, 27th 2019.

Review on the doctoral Thesis

Application of simplified models of electron density for estimation of the electrostatic properties of molecules

Author: **Sławomir A. Bojarowski**

Supervisor: Paulina M. Dominiak

Faculty of Chemistry, University of Warsaw.

The PhD manuscript comprises four parts after the summary:

- (1) A comprehensive introduction describing the state of the art
 - (2) A short text about objectives
 - (3) The reprint of the article “Bojarowski, S. A., Kumar, P., & Dominiak, P. M. (2017). “Interplay of point multipole moments and charge penetration for intermolecular electrostatic interaction energies from the University at Buffalo pseudoatom databank model of electron density.” *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials*, 73(4), 598-609.”
 - (4) The reprint of the article “, S. A., Kumar, P., & Dominiak, P. M. (2016). A Universal and Straightforward Approach to Include Penetration Effects in Electrostatic Interaction Energy Estimation. *ChemPhysChem*, 17(16), 2455-2460.”
 - (5) The reprint of article : “ Bojarowski, S. A., Kumar, P., Wandtke, C. M., Dittrich, B., & Dominiak, P. M. (2018). Universal Method for Electrostatic Interaction Energies Estimation with Charge Penetration and Easily Attainable Point Charges. *Journal of Chemical Theory and Computation*, 14(12), 6336-6345.”
 - (6) One page of conclusions.
 - (7) Supporting information which includes notably tables of energy values for different methods and molecular dimers.
- (1) The introduction is very instructive with a nice overview of the field of intermolecular energy calculations.

An important and recent reference is however missing in the introduction. This article is related to the aug-PROmol method developed as it describes another way to compute rapidly the electrostatic energy:

Nguyen, D., Kisiel, Z., & Volkov, A. (2018). Fast analytical evaluation of intermolecular electrostatic interaction energies using the pseudoatom representation of the electron density. I. The Löwdin α -function method. *Acta Crystallographica Section A: Foundations and Advances*, 74(5), 524-536.

The article proposes a rapid analytical calculation of the energy, without approximation, when the Hansen & Coppens multipolar atom model with Slater functions is used.

Concerning the PhD manuscript quality, the introduction has a non-negligible number of typographic errors and English language approximations.

(3) Paper Sławomir A. Bojarowski, Prashant Kumar and Paulina M. Dominiak*

Acta Cryst. (2017). B73, 598–609.

One major discovery in this article is that the electrostatic energy derived from fitted point charges agrees, surprisingly, relatively well with energy computed from the multipolar UBDB model, at longer distances. Molecules from the S66 benchmark of energy dimers are used and the contribution of the different multipole level components are analyzed. The penetration of the two electron densities has a large effect on the resulting electrostatic energy at the shortest interaction distances. This is not well modelled by the point charges.

The expression ‘‘pi-pi stacking’’ is found in the article. It is used very commonly in the literature, but shall be misleading as this interaction is more about stacking of flat cycles which display electrostatic complementarity. Heterocycles bearing electro-positive and negative atoms have indeed a higher propensity to be involved in parallel stacking.

See paper Martinez & Iverson, 2012, *Chemical Science*, ‘‘Rethinking the term pi-stacking’’,

(4) The article describes an empirical new method aug-PROmol to estimate rapidly the electrostatic energy (E_{es}) which is nearly of comparable quality to that obtained from the UBDB multipolar transfer and first principle calculations. The method is based on atomic RESP charges fitted against an electrostatic potential obtained from first principle calculations. It ignores the multipolar part, which has limited contribution to E_{es} . The procedure, based on empirical fitting of several coefficients to reproduce the electrostatic energy, avoids the lengthy integration procedure of EPMM. The energy deviation are computed and compared for UBDB, point RESP charges and aug-PROmol on the dimers found in the S66 database. The empirical fitting does model very well the penetration effect which occurs for short distance interactions. The article is relatively concise, easy to read and well presented.

(5) The third article is a further development of the aug-PROmol methodology which used notably the tabulated atomic charges derived from INVARIOM electron density database developed by Birger Dittrich. These charges can be readily obtained by a transfer procedure and overcome the bottleneck of aug-PROmol which else requires atomic charges obtained from a fitting procedure and might require ab initio calculations. RESP charges (as in the precedent article) were obtained using the Antechamber toolkit of the Amber molecular modelling package and were also tested. Energy values are compared to the reference obtained from B3LYP calculations.

One interesting thing which could have been tested here in addition is the following: the use of atomic charges fitted against the electrostatic potential generated by the UBDB transfer.

Mr Bojarowski has conducted a PhD which has led to results which are very good both in terms of quantity and quality. The applicant has a total of 4 publications, 3 as first author, in good to top quality journals (JCTC).

The research project has led to original and important scientific results, which have been quite noticed by the Quantum Crystallography community. The rapidity and universal nature of the aug-PROmol methodology can be expected to lead in the future to further developments in the Molecular Modelling community.

As a conclusion, his scientific achievements allow clearly Mr Slawomir Bojarowski for the admission of the thesis to a public defense.

Typographic and English sentences errors and suggestions.

Abstract. Line N-2. "Error is 0.07 kcal/mol". It would be informative to give also the percentage of relative error, not just an absolute value.

page 17, line 14 "a deviations"

page 17, line 11. "The greater power of $1/R$, the range of interactions are shorter, and at larger distances will be giving only small contributions to the E_{int} ."

Sentence is not correct.

"The greater the power of $1/R$ is, the shorter the range of interactions are,"

Page 18, line 17 in 1.4.2 §. "treated as energy of a interaction" → "treated as interaction energy"

"since there are involved only excited states of both monomers" → "since only excited states of both monomers are involved".

"and it is the bigger," → "and it is the biggest"

Line N-6:; "causing lowering of the system energy"

Page 20. Line N-8 "When molecules are approaching each other the error is the greater, the closer molecules are".

"the closer molecules are" is probably redundant.

Example of correct sentence: "When molecules are approaching each other, the error becomes greater"

Page 20, line N-7: "At some distances the electron clouds cannot be described by dimensionless elements and the error becomes significant."

The phrase "cannot be described by dimensionless elements" is really not clear, what does it mean?

Page 22, line 9 "spherical harmonics are ranked"

Page 22, line 12: "The lower number of z the expansion will diverge easier." →

"The lower the number z is, the more easily the expansion will diverge."

Page 22, line 20. "It relies on partitioning the electron density"

Page 23, line 17, "IAM relies on a assumption" → "IAM relies on the assumption"

Page 23, N-8 "can be transferred"

Page 24, line 21: "to their equivalent"

Page 24, line N-2: "These studies"

Page 24, line N: "Thus, only these, which try to cope"

Page 25, line 17: "force field"

Page 25, line 19: Effective

Page 25, line 24: "Gordon et al. did a wide studies

Page 26, line 2 : "it relies on mutipole moments"

Equation 3.1.3. E values are not defined.

Page 27, line N-13: "Such a test"

Page 27, line N-1 "of the ~~the~~ AMOEBA force field"

Page 28 line 5. Give definition of Emtp.

Equation (3.1.5) Is it normal that there is both "r" and "rij" ?

If yes, define them.

There is a "{ " parenthesis in equation, but "}" is missing.

Equation 3.1.6. there are two summations over i,j. Is that equation correct ?

It would be better to add parentheses to better delimit summations.

Page 29, line 1. "The damping functions" ?

Page 29, line 10 "are at range" incorrect English

Page 29, line 19. "This studies".

Page 29, line 12: "by two Gaussian-type basis set, obtained at any theory."

Change to "basis sets"

"at any theory" is not clear/correct.

Page 29, line N-6. “Quantum chemical topology force field” “field”.

Page 30, line 8. “It uses a mathematical expressions”

Page 31, line 3: “author considers”

Page 31, line N-19: ‘pi-pi interactions’. Shall be replaced by “stacking interactions” for example.

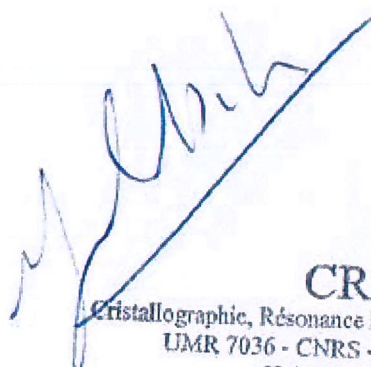
See paper Martinez & Iverson, 2012, Chemical Science, “Rethinking the term pi-stacking”, Which suggests this expression is misleading and inappropriate.

Page 32, line 4: “were used to obtain”

Reference 48 Popelier: year is missing.

Page 32, line N-3 “The studies were performed on a small organic molecules”

Page 33, line 3 “Finally, for a net dipole moments, the RMSE”



CRM2
Cristallographie, Résonance Magnétique et Modélisations
UMR 7036 - CNRS - Institut Jean Barriol
Université de Lorraine
BP 70239 - 54506 VANDŒUVRE LES NANCY CEDEX - France

Christian Jelsch

Director of Research CNRS

CRM2 Université de Lorraine, Faculté des Sciences et Technologies Laboratoire
Cristallographie Résonance Magnétique et Modélisations Vandoeuvre-les-
Nancy France

Vandoeuvre lès Nancy, February, 27th 2019.

Recenzja pracy doktorskiej

Application of simplified models of electron density for estimation of the electrostatic properties of molecules

Autor: **Sławomir A. Bojarowski**

Promotor: Paulina M. Dominiak

Wydział Chemii, Uniwersytet Warszawski.

Praca doktorska składa się z czterech części po streszczeniu:

- (1) Wszechstronne wprowadzenie opisujące aktualny stan wiedzy
- (2) Krótki tekst dotyczący celów
- (3) Reprint artykułu „Bojarowski S. A., Kumar P. & Dominiak P. M. (2017). „Interplay of point multipole moments and charge penetration for intermolecular electrostatic interaction energies from the University at Buffalo pseudoatom databank model of electron density”, *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials*, 73(4), 598-609.”
- (4) Reprint artykułu: „Bojarowski S. A., Kumar P., and Dominiak P. M., A universal and straightforward approach to include penetration effects in electrostatic interaction energy estimation, *ChemPhysChem*, 17(16), 2455-2460.”
- (5) Reprint artykułu: Bojarowski S. A., Kumar P., Wandtke C. M., Dittrich B. and Dominiak P. M. Universal Method for Electrostatic Interaction Energies Estimation with Charge Penetration and Easily Attainable Point Charges, *Journal of Chemical Theory and Computation*, 14(12), 6336-6345. *J. Chem. Theory Comput.* 2018, 14, 6336- 6345.”
- (6) Jedna strona wniosków.
- (7) Supplement, który obejmuje w szczególności tabele wartości energii dla różnych metod i różnych dimerów cząsteczkowych.

(1) Wprowadzenie jest bardzo pouczające, z ładnym przeglądem metod obliczeń energii międzycząsteczkowej.

Brakuje jednakże we wstępie cytowania ważnego i nowo opublikowanego źródła. Ten artykuł jest związany z opracowaną metodą aug-PROmol, ponieważ opisuje inny sposób szybkiego obliczania energii elektrostatycznej: Nguyen, D., Kisiel, Z., & Volkov, A. (2018). Fast analytical

evaluation of intermolecular electrostatic interaction energies using the pseudoatom representation of the electron density. I. The Löwdin α -function method. *Acta Crystallographica Section A: Foundations and Advances*, 74(5), 524-536.

Artykuł proponuje szybkie analityczne obliczenia energii, bez przybliżeń, przy wykorzystaniu modelu multipolowego atomu Hansena i Coppensa z funkcjami Slatera. Jeśli chodzi o jakość pracy doktorskiej, wprowadzenie ma nieistotną liczbę błędów typograficznych i uproszczeń w języku angielskim.

(3) Artykuł Sławomir A. Bojarowski, Prashant Kumar and Paulina M. Dominiak* *Acta Cryst.* (2017). B73, 598–609.

Jednym z głównych odkryć w tym artykule jest to, że energia elektrostatyczna pochodząca z dopasowanego ładunków punktowych, na dłuższych dystansach, zgadza się, zaskakująco dobrze z energią obliczoną z multipolowego modelu UBDB. Wykorzystywane są cząsteczki z *benchmarku* dimerów energetycznych S66 i analizowany jest udział różnych składników multipolowych. Penetracja dwóch gęstości elektronowych ma duży wpływ na szacowaną energię elektrostatyczną dla niewielki odległości. Nie jest to zjawisko dobrze modelowane przez ładunki punktowe.

W artykule znajduje się wyrażenie "pi-pi stacking". Stosowane jest bardzo często w literaturze, ale wprowadza w błąd, ponieważ to oddziaływanie dotyczy raczej ułożenia płaskich pierścieni, które wykazują komplementarność elektrostatyczną. Heteropierścienie z atomami elektrododatnimi i ujemnymi mają istotnie większą skłonność do brania udziału w ułożeniu równoległym. Warto zobaczyć artykuł Martinez & Iverson, 2012, *Chemical Science*, "Rethinking the term pi-stacking".

(4) Artykuł opisuje nową metodę empiryczną aug-PROmol do szybkiego oszacowania energii elektrostatycznej (Ees), która jest prawie porównywalnej jakości do tej uzyskanej z transferu multipolowego UBDB i obliczeń *ab initio*. Metoda oparta jest na atomowych ładunkach RESP dopasowanych do potencjału elektrostatycznego uzyskanego z obliczeń *ab initio*. Metoda pomija część multipolową, która ma ograniczony wkład do Ees. Procedura, oparta na empirycznym dopasowaniu kilku współczynników w celu odtworzenia energii elektrostatycznej, pozwala uniknąć długiej procedury całkowania EPMM. Błędy energii, na dimerach z bazy danych S66, są obliczane i porównywane dla UBDB, ładunków punktowych RESP i aug-PROmol. Empiryczne dopasowanie bardzo dobrze modeluje efekt penetracji, który występuje w przypadku interakcji na krótkich dystansach. Artykuł jest stosunkowo zwięzły, czytelny i dobrze przedstawiony.

(5) Trzeci artykuł jest dalszym rozwinięciem metodologii aug-PROmol, która wykorzystuje w szczególności tabulowane ładunki atomowe pochodzące z bazy gęstości elektronowej INVARIOM opracowanej przez Birgera Dittricha. Ładunki te można łatwo uzyskać za pomocą procedury przenoszenia i w ten sposób ominąć wąskie gardło aug-PROmol, czyli ominąć dodatkowy wymóg dopasowania ładunków atomowych do potencjałów uzyskanych z obliczeń *ab initio*. Ładunki RESP (jak w poprzednim artykule) zostały uzyskane przy użyciu programu Antechamber z pakietu Amber i również zostały przetestowane. Wartości energii są porównywane z wartościami referencyjnymi uzyskanymi z obliczeń B3LYP.

Interesującą rzeczą, którą można tutaj przetestować, jest ponadto: wykorzystanie ładunków atomowych dopasowanych do potencjału elektrostatycznego generowanego przez UBDB.

Pan Bojarowski przeprowadził doktorat, którego wyniki są bardzo dobre zarówno pod względem ilościowym jak i jakościowym. Doktorant ma w sumie 4 publikacje, 3 jako pierwszy autor, w dobrych lub on najwyższej jakości czasopismach (JCTC).

Projekt badawczy doprowadził do oryginalnych i ważnych wyników naukowych, które zostały zauważone przez społeczność kwantowej krystalografii. Można oczekiwać, że szybkość i uniwersalność metodologii aug-PRomol doprowadzą w przyszłości do jej dalszego rozwoju w społeczności modelowania molekularnego.

Podsumowując, jego dorobek naukowy wyraźnie pozwala panu Sławomirowi Bojarowskiemu na dopuszczenie pracy do publicznej obrony.

Błędy typograficzne i językowe (angielskie) zdań oraz sugestie .

Abstrakcyjny. Linia N-2. "Error is 0.07 kcal/mol". Byłoby pouczające podać również procent błędu względnego, a nie tylko wartość bezwzględną.

strona 17, wiersz 14 " a deviations "

strona 17, wiersz 11. " The greater power of $1/R$, the range of interactions are shorter, and at larger distances will be giving only small contributions to the Eint. " Zdanie nie jest poprawne. " The greater the power of $1/R$ is, the shorter the range of interactions are, "

Strona 18, wiersz 17 w 1.4.2 §. " treated as energy of a interaction" na "treated as interaction energy"

"since there are involved only excited states of both monomers" na "since only excited states of both monomers are involved".

"and it is the bigger," na "and it is the biggest"

wiersz N-6; "causing lowering of the system energy"

Strona 20. wiersz N-8 "When molecules are approaching each other the error is the greater, the closer molecules are".

"the closer molecules are" jest prawdopodobnie powtórzeniem

Przykładowe poprawne zdanie: "When molecules are approaching each other, the error becomes greater"

Strona 20, wiersz N-7: "At some distances the electron clouds cannot be described by dimensionless elements and the error becomes significant."

Fraza: "cannot be described by dimensionless elements", nie jest do końca jasna.

Strona 22, wiersz 9 "spherical harmonics are ranked"

Strona 22, wiersz 12: "The lower number of z the expansion will diverge easier." na "The lower the number z is, the more easily the expansion will diverge."

Strona 22, wiersz 20. "It relies on partitioning the electron density"

Strona 23, wiersz 17, "IAM relies on a assumption" na "IAM relies on the assumption"

Strona 23, N-8 "can be transferred"

Strona 24, wiersz 21: "to their equivalent"

Strona 24, wiersz N-2: "These studies"

Strona 24, wiersz N: "Thus, only these, which try to cope"

Strona 25, wiersz 17: "force field"

Strona 25, wiersz 19: Effective

Strona 25, wiersz 24: "Gordon et al. did a wide studies

Strona 26, wiersz 2 : "it relies on mutipole moments" równanie 3.1.3. wartości E nie są zdefiniowane.

Strona 27, wiersz N-13: "Such a test"

Strona 27, wiersz N-1 "of the ~~the~~ AMOEBA force field"

Strona 28 wiersz 5 dać definicję E_{mp}.

równanie (3.1.5) czy to normalne, że jest "r" i "rij" ? Jeśli tak to trzeba je zdefiniować.

W równaniu występuje nawias "{", ale brakuje "}".

Równanie 3.1.6. są dwa podsumowania nad i, j. Czy to równanie jest poprawne? Byłoby lepiej dodać nawiasy, aby lepiej ograniczyć liczbę podsumowań.

Strona 29, wiersz 1. "The damping functions" ?

Strona 29, wiersz 10 "are at range" niepoprawny angielski

Strona 29, wiersz 19. "This studies".

Strona 29, wiersz 12: "by two Gaussian-type basis set, obtained at any theory." Change to "basis sets" "at any theory" jest niejasne/niepoprawne

Strona 29, wiersz N-6. "Quantum chemical topology force filed" "field".

Strona 30, wiersz 8. "It uses a mathematical expressions"

Strona 31, wiersz 3: "author considers"

Strona 31, wiersz N-19: "pi-pi interactions" Powinno być zmienione na na przykład "stacking interactions".

Zobacz Martinez & Iverson, 2012, Chemical Science, "Rethinking the term pi-stacking", Co sugeruje, że to wyrażenie jest mylące i niewłaściwe.

Strona 32, wiersz 4: "were used to obtain"

pozycja 48 Popelier: brakuje roku.

Strona 32, wiersz N-3 "The studies were performed on a small organic molecules"

Strona 33, wiersz 3 "Finally, for a net dipole moments, the RMSE"



dr hab. Paulina Maria Dominiak, prof. UW
Promotor