

**MASARYKOVA UNIVERZITA**  
**PŘÍRODOVĚDECKÁ FAKULTA**  
**ÚSTAV BIOCHEMIE**

# **Bakalářská práce**

**BRNO 2019**

**RADKA SEDLÁKOVÁ**



**MASARYKOVA UNIVERZITA**  
**PŘÍRODOVĚDECKÁ FAKULTA**  
ÚSTAV BIOCHEMIE

---



# **Atomové typy v metodách pro výpočet parciálních atomových nábojů**

Bakalářská práce

**Radka Sedláková**

**Vedoucí práce: RNDr. Tomáš Raček**

**Brno 2019**

# Bibliografický záznam

<b>Autor:</b>	Radka Sedláková Přírodovědecká fakulta, Masarykova univerzita Ústav biochemie
<b>Název práce:</b>	Atomové typy v metodách pro výpočet parciálních atomových nábojů
<b>Studijní program:</b>	Biochemie
<b>Studijní obor:</b>	Chemoinformatika a bioinformatika
<b>Vedoucí práce:</b>	RNDr. Tomáš Raček
<b>Akademický rok:</b>	2018/2019
<b>Počet stran:</b>	počet stran od druhé stránky dokumentu po poslední stránku obsahu + počet stránek práce od první strany Úvodu
<b>Klíčová slova:</b>	atomy, magie

# Bibliographic Entry

**Author:** Radka Sedláková  
Faculty of Science, Masaryk University  
Department of Biochemistry

**Title of Thesis:** Atom types in methods for calculation of partial atomic charges

**Degree Programme:** Biochemistry

**Field of Study:** Chemoinformatics and Bioinformatics

**Supervisor:** RNDr. Tomáš Raček

**Academic Year:** 2018/2019

**Number of Pages:** počet stran od druhé stránky dokumentu po poslední stránku obsahu + počet stránek práce od první strany Úvodu

**Keywords:** atoms, magic

# Abstrakt

Parciální atomové náboje

# Abstract

In this thesis we study

**Místo tohoto listu vložte kopii oficiálního (podepsaného) zadání práce.**

# Poděkování

Na tomto místě bych chtěla poděkovat stránce LOTR University Memes za zlepšování nálady ve chvílích největší bezradnosti.

# Prohlášení

Prohlašuji, že jsem svoji bakalářskou práci vypracovala samostatně s využitím informačních zdrojů, které jsou v práci citovány.

Brno 15. května 2019

.....  
Radka Sedláková

# Obsah

<b>Přehled použitého značení</b> .....	<b>ix</b>
<b>Kapitola 1. Úvod</b> .....	<b>1</b>
<b>Kapitola 2. Teorie</b> .....	<b>2</b>
2.1 Parciální atomové náboje .....	2
2.2 Kvantově-mechanické metody .....	2
2.2.1 Základy kvantové mechaniky .....	3
2.2.2 Přehled kvantově-mechanických metod .....	3
2.3 Empirické metody .....	5
2.3.1 PEOE .....	5
2.3.2 EEM .....	6
2.3.3 Parametrizace .....	6
2.3.4 Atomové typy .....	6
2.4 Statistické pojmy .....	6
2.4.1 Průměrná a maximální absolutní odchylka .....	6
2.4.2 RMSD .....	6
2.4.3 Pearsonův korelační koeficient .....	6
<b>Kapitola 3. Metody</b> .....	<b>7</b>
3.1 Vstupní data (Formát souborů) .....	7
3.2 (SMILES/SMARTS - jestli je použiju) .....	7
3.3 RDKit .....	7
3.4 NumPy .....	7
3.5 Matplotlib? Co mi Ondra ukazoval jeho program, tak ten grafy taky tvoří ..	7
3.6 Ondrův program .....	7
<b>Kapitola 4. Implementace</b> .....	<b>8</b>
4.1 Klasifikátory (Co jiného?) .....	8
<b>Kapitola 5. Výsledky a diskuse</b> .....	<b>9</b>
5.1 Výsledky parametrizace .....	9
5.1.1 Klasifikátor1 .....	9
5.1.2 Klasifikátor2 .....	9
5.1.3 Klasifikátor3 .....	9



5.2 Srovnání navržených klasifikátorů . . . . .	9
<b>Kapitola 6. Závěr . . . . .</b>	<b>10</b>
<b>Příloha . . . . .</b>	<b>11</b>
<b>Seznam použité literatury . . . . .</b>	<b>12</b>

# Přehled použitého značení

Pro jednodušší orientaci čtenářů v textu bakalářské práce uvádím seznam zkratek, které jsou v textu použity.

- $\mathbb{C}$  množina všech komplexních čísel
- $\mathbb{R}$  množina všech reálných čísel
- $\mathbb{Z}$  množina všech celých čísel
- $\mathbb{N}$  množina všech přirozených čísel

# Kapitola 1

## Úvod

Distribuce elektronů v m

# Kapitola 2

## Teorie

### 2.1 Parciální atomové náboje

Parciální atomové náboje [1] jsou reálná čísla, která popisují asymetrické rozložení elektronové hustoty na chemické vazbě. Vznikají v důsledku rozdílných elektronegativit vazebných partnerů. Pokud v chemické vazbě figuruje vysoce elektronegativní atom, pak tento k sobě přitahuje vazebný elektronový pár, čímž se zvyšuje elektronová hustota v jeho okolí a dochází ke vzniku parciálního záporného náboje ( $\delta^-$ ). V okolí elektro-pozitivnějšího vazebného partnera se elektronová hustota naopak snižuje a na atomu dochází ke vzniku parciálního kladného náboje ( $\delta^+$ ).

Koncept parciálních atomových nábojů je pouze teoretický, hodnoty nábojů proto nelze získat pomocí experimentu [2]. Jelikož jsou ale významným faktorem pro predikci fyzikálních, chemických a biologických vlastností molekul, bylo pro jejich stanovení vyvinuto množství výpočetních metod. Tyto se dělí na metody kvantově-mechanické a metody empirické. Kvantově-mechanické metody poskytují přesnější výsledky, ovšem za cenu vysoké časové náročnosti. Empirické metody dosahují v porovnání s QM metodami velmi dobrých výsledků, a to ve výrazně kratším čase. Žádná z vyvinutých empirických metod však není uznána za všeobecně platnou a použitelnost konkrétních metod se hodnotí na základě reprodukovatelnosti výsledků [3].

Aplikaci parciálních atomových nábojů lze nalézt ve výpočetní chemii a chemoinformaticce, kde slouží k predikci elektrostatických a termodynamických vlastností popisujících reaktivitu molekul. Uplatňují se v molekulových simulacích [4], ve virtuálním screeningu [5], při hledání vazebných míst proteinů nebo při návrhu farmakoforů. Prokázaly se jako platné deskriptory v QSAR a QSPR modelech [6, 7]. V anorganické chemii se uplatňují při popisu toku elektronů v bateriích a katalyzátorech [8].

### 2.2 Kvantově-mechanické metody

Kvantově-mechanické metody pro výpočet parciálních atomových nábojů jsou založeny na poznatcích kvantové mechaniky. Dělí se na tři hlavní skupiny, a to metody semi-empirické, metody odvozené od teorie funkcionálu hustoty a metody *ab initio*. *Ab initio* metody (lat. *ab initio* - od počátku) staví výpočty na teoretickém aparátu a k

řešení Schrödingerovy rovnice přistupují numericky, z čehož vyplývá jejich velká výpočetní náročnost. Metody semi-empirické jsou stejně jako metody *ab initio* založeny na řešení SE, část výpočtů však parametrizují nebo aproximují na základě experimentálních dat.

Limitujícím faktorem pro použití QM metod je jejich složitost, konkrétně pro *ab initio* metody až  $O(B^4)$ , kde  $B$  je číslo rovno počtu elektronů v molekule nebo větší.

### 2.2.1 Základy kvantové mechaniky

K rozvoji kvantové mechaniky došlo ve 20. letech 20. století v reakci na newtonovskou mechaniku, jejíž aparát nepostačoval pro popis mikrosvěta. Základním principem QM je vlnově-korpuskulární dualismus. Vlna je v kvantové mechanice reprezentována matematickou funkcí  $\Psi$ , tzv. vlnovou funkcí, která popisuje dynamický stav částice a nese veškeré informace, které lze o částici získat [9].

Základním úkolem kvantové mechaniky je výpočet vlnové funkce systému. Vlnová funkce je řešením Schrödingerovy rovnice

$$H\Psi = E\Psi$$

kde  $H$  je operátor Hamiltonián a  $E$  je energie systému. Hamiltonián bere na vstup vlnovou funkci  $\Psi$  a transformuje ji na funkci jinou. Řešením Schrödingerovy rovnice je soubor funkcí, které lze po aplikaci Hamiltoniánu zapsat jako součin původní funkce a skaláru  $E$ . Takovéto funkce označujeme jako *vlastní funkce* a odpovídající skaláry jako *vlastní hodnoty* operátoru [10].

Schrödingerova rovnice je exaktně řešitelná pouze pro vybrané problémy. Jedním z nich je atom vodíku. Pro víceelektronové systémy je nutno do výpočtu zavádět velké množství aproximací, z nichž nejznámější je Born-Oppenheimerova aproximace. Jejím základním konceptem je oddělení pohybu jader atomů od pohybu elektronů, vycházející z předpokladu, že jádra, mnohonásobně těžší než elektrony, se pohybují výrazně pomaleji než elektrony samotné. Řešení Schrödingerovy rovnice

$$H\Psi(r, R) = E\Psi(r, R)$$

se tak rozkládá na řešení popisující elektrony v souboru fixních jader, po němž následuje řešení rovnice zahrnující kinetickou a potenciální energii jader obklopených polem elektronů.  $\Psi(r, R)$  je vlnová funkce systému, závislá jak na souřadnicích elektronů ( $r$ ), tak na souřadnicích jader ( $R$ ).

### 2.2.2 Přehled kvantově-mechanických metod

Důležitým krokem *ab initio* a semi-empirických metod je výběr báze sady. Báze sady je soubor vlnových funkcí reprezentujících atomové orbitály, jejichž vhodnou lineární kombinací (LCAO) lze následně vyjádřit vlnovou funkci molekuly. Pro popis funkcí reprezentujících atomové orbitály se používají orbitály Gaussova typu (GTO). Kombinace několika Gaussových orbitalů přibližuje tzv. Slaterův orbital (STO), který je pro výpočet vlnové funkce molekuly méně vhodný z důvodu složitosti výpočtů. Báze sady

existuje nepřeberné množství, např. báze sady STO-3G, STO-4G či obecně STO- $n$ G, kde  $n$  je počet orbitalů Gaussova typu reprezentujících jeden atomový orbital.

Krokem vedoucím k výpočtu hodnot parciálních nábojů je provedení populační analýzy, která popisuje rozložení elektronové hustoty v molekule. Příkladem je Mullikenova populační analýza (*Mulliken Population Analysis*, MPA), která elektronovou hustotu určuje dle obsazenosti atomových orbitalů elektrony. V rámci chemické vazby je elektronová hustota rovnoměrně rozdělena mezi vazebné partnery, není tedy brána v potaz možná rozdílnost elektronegativit. Výsledky MPA jsou také silně závislé na použitém kvantově-mechanickém přístupu a na velikosti báze sady. Nevýhody MPA, zejména nepřesnost výsledků související s rozšiřováním báze sady, řeší přirozená populační analýza (*Natural Population Analysis*, NPA), pracující s přirozenými atomovými orbitaly. Přirozené atomové orbitály jsou nejprve vypočteny z báze sady a jsou následně použity pro výpočet ortonormálních přirozených vazebných orbitalů (*Natural bonding orbitals*, NBO). Na základě NBO se poté provádí populační analýza.

Odlišný přístup finálního výpočtu parciálních atomových nábojů představuje metoda *Atoms-in-Molecules* (AIM), která přiřazuje náboje atomům na základě integrace elektronové hustoty přes prostor příslušící danému atomu. Dalším možným přístupem je výpočet parciálních atomových nábojů na základě elektrostatických potenciálů molekuly (metody založené na *Molecular Electrostatic Potential-derived charges*, MEP).

### Hartree-Fockova metoda

Problém řešení víceelektronových systémů nastává při zahrnutí elektronových interakcí do výpočtu. Hartree-Fockova metoda rozkládá původní problém  $n$ -elektronové Schrödingerovy rovnice na řešení  $n$  jednoelektronových rovnic. Využívá přiblížení pomocí metody nezávislých částic (*Self-Consistent Field*, SCF), která pracuje s modelem elektronu pohybujícím se v průměrném poli ostatních elektronů. HF metoda tedy nezahrnuje korelaci pohybu elektronů. Jednoelektronové rovnice jsou určeny předpisem

$$\hat{F}\chi_i = \varepsilon_i \chi_i \quad (2.1)$$

kde Fockův operátor  $\hat{F}$  je Hamiltonián aplikovaný na jednoelektronový (atomový nebo molekulový) orbital a  $\varepsilon_i$  je odpovídající Langrangeův multiplikátor. Metoda pracuje iterativně a konečné řešení rovnic určuje na základě ustálení výsledků jednotlivých iterací výpočtu.

### DFT

Metody založené na teorii funkcionálu hustoty (*Density Functional Theory*, DFT) nevycházejí z řešení vlnové funkce, ale poznatky staví na rozložení elektronové hustoty v molekule, ze které následně odvozují energii systému a další vlastnosti molekuly. Do výpočtů DFT metod je narozdíl od Hartree-Fockovy metody zahrnuta i korelační energie elektronů. Ve výpočtech figurují pouze tři neznámé (souřadnice  $x$ ,  $y$ ,  $z$ ), zatímco řešení Schrödingerovy rovnice obsahuje  $4n$  neznámých, kde  $n$  představuje počet elektronů systému. Metody odvozené od DFT jsou tak výpočetně méně náročné a poskytují přesnější výsledky.

Jedním z cílů DFT metod je výpočet celkové energie elektronů na základě elektronové hustoty. *Funkcionál* je v rámci DFT metod chápán jako zobrazení, které zobrazuje funkci, představující elektronovu hustotu, do množiny reálných čísel popisujících energii elektronů. Říkáme, že energie elektronů je funkcionálem elektronové hustoty.

## Semiempirické metody

## 2.3 Empirické metody

### 2.3.1 PEOE

Metoda PEOE (*Partial Equalization of Orbital Electronegativity*), známá také pod jménem autorů jako metoda Gasteiger-Marsili, byla poprvé publikována v roce 1980 [c]. Metodu lze aplikovat pouze na systémy obsahující  $\sigma$  vazby a nekonjugované  $\pi$  vazby, později však byla autory rozšířena i o výpočet systémů konjugovaných  $\pi$  vazeb. V rámci metody není uvažována 3D struktura molekuly, pracuje se pouze s její topologií.

Koncept elektronegativity atomových orbitalů, na němž je metoda založena, vychází z Mullikenovy definice elektronegativity  $\chi_A$  atomu  $A$

$$\chi_A = \frac{1}{2}(I_A + E_A) \quad (2.2)$$

Dle Mullikena je elektronegativita atomu určena hodnotami elektronových afinit  $E_A$  a ionizačních potenciálů  $I_A$  jeho valenčních stavů. PEOE připisuje na základě hodnot  $I_A$  a  $E_A$  elektronegativitu každému orbitalu valenčního stavu atomu. Elektronegativita  $\chi_{iv}$  orbitalu  $iv$  na atomu  $i$

$$\chi_{iv} = a_{iv} + b_{iv}Q_i + c_{iv}Q_i^2 \quad (2.3)$$

je ovlivněna náboji ostatních orbitalů a tedy i celkovým nábojem příslušného atomu  $Q_i$ . Koeficienty  $a_{iv}$ ,  $b_{iv}$  a  $c_{iv}$  jsou empirické parametry vypočtené z ionizačních potenciálů a elektronových afinit neutrálního, kationtového a aniontového stavu příslušného orbitalu. (za PEOE vložit tabulku parametrů abc, viz Gasteiger s. 330)

Při vzniku vazby dochází na základě elektronegativity atomů k přesunu elektronů od elektropozitivnějšího atomu směrem k elektronegativnějšímu. Interagují spolu příslušné atomové orbitály a dochází k částečné ekvalizaci (vyrovnání) jejich nábojů. Množství přeneseného náboje mezi atomy  $A$  a  $B$  v  $k$ -té iteraci výpočtu, kde atom  $B$  má vyšší elektronegativitu, je definováno jako

$$Q^{(k)} = \frac{\chi_B^{(k)} - \chi_A^{(k)}}{\chi_A^+} \cdot \alpha^k \quad (2.4)$$

kde  $\chi_A^+$  označuje elektronegativitu kationtu atomu  $A$ . Iniciální výpočet elektronegativity orbitalu (2.3) pracuje s formálním nábojem atomu. Po výpočtu příspěvků přenesených nábojů (2.4) všech vazebných partnerů atomu je náboj daného atomu přepočítán a použit v další iteraci. Množství přeneseného náboje mezi dvěma atomy se v každé iteraci výpočtu snižuje a vypočtené hodnoty nábojů atomů postupně konvergují. Přibližně po šesté iteraci dochází k ustálení výpočtů.

Díky přesnosti a rychlosti výpočtů byla PEOE implementována do většiny programů pro molekulové modelování jako základní metoda výpočtu atomových nábojů. Reziduální elektronegativita atomů, získaná na základě PEOE, se prokázala vhodnou pro popis indukčního efektu v molekulách.

### **2.3.2 EEM**

### **2.3.3 Parametrizace**

### **2.3.4 Atomové typy**

## **2.4 Statistické pojmy**

### **2.4.1 Průměrná a maximální absolutní odchylka**

### **2.4.2 RMSD**

### **2.4.3 Pearsonův korelační koeficient**



# **Kapitola 3**

## **Metody**

**3.1 Vstupní data (Formát souborů)**

**3.2 (SMILES/SMARTS - jestli je použiju)**

**3.3 RDKit**

**3.4 NumPy**

**3.5 Matplotlib? Co mi Ondra ukazoval jeho program, tak ten grafy taky tvoří**

**3.6 Ondrův program**

## **Kapitola 4**

### **Implementace**

#### **4.1 Klasifikátory (Co jiného?)**

# **Kapitola 5**

## **Výsledky a diskuse**

### **5.1 Výsledky parametrizace**

#### **5.1.1 Klasifikátor1**

#### **5.1.2 Klasifikátor2**

#### **5.1.3 Klasifikátor3**

### **5.2 Srovnání navržených klasifikátorů**

## Kapitola 6

## Závěr

[illegible][illegible]

# Příloha

**Sem můžete přidat přílohu. Pokud chcete “Přílohy”, tak upravte definici záhlaví v souboru sci.muni.thesis.sty, viz příkaz \HlavickaPriloha.**

# Seznam použité literatury

- [1] ATKINS, P. W. & DE PAULA, J. *Atkins' physical chemistry*. 9th ed. Oxford: Oxford University Press, c2010. ISBN 978-0-19-954337-3.
- [2] LEACH, A. R. *Molecular modelling: principles and applications*. 2nd ed. New York: Prentice Hall, 2001. ISBN 0-582-38210-6.
- [3] GASTEIGER, J. & ENGEL, T. *Chemoinformatics: A Textbook*. Weinheim: Wiley-VCH, c2003. ISBN 978-3-527-30681-7.
- [4] RAPPE, A. K. & GODDARD, W. A. „Charge equilibration for molecular dynamics simulations“. *The Journal of Physical Chemistry*. 1991, **95**(8), 3358-3363. DOI: 10.1021/j100161a070. ISSN: 0022-3654. URL: <http://pubs.acs.org/doi/abs/10.1021/j100161a070>
- [5] LYNE, P. D. „Structure-based virtual screening: an overview“. *Drug Discovery Today*. 2002, **7**(20), 1047-1055. DOI: 10.1016/S1359-6446(02)02483-2. ISSN: 13596446. URL: <http://linkinghub.elsevier.com/retrieve/pii/S1359644602024832>
- [6] GHAFOURIAN, T., DEARDEN, J. C. & KATRITZKY, A. R. „The Use of Atomic Charges and Orbital Energies as Hydrogen-bonding-donor Parameters for QSAR Studies: Comparison of MNDO, AM1 and PM3 Methods“. *Journal of Pharmacy and Pharmacology*. 2000, **52**(6), 1027-1044. DOI: 10.1211/0022357001774435. ISSN: 00223573. URL: <http://doi.wiley.com/10.1211/0022357001774435>
- [7] KARELSON, M., LOBANOV, V. S. & KATRITZKY, A. R. „Quantum-Chemical Descriptors in QSAR/QSPR studies“. *Chemical Reviews*. 1996, **96**(3), 1027-1044. DOI: 10.1021/cr950202r. ISSN: 0009-2665. URL: <https://pubs.acs.org/doi/10.1021/cr950202r>
- [8] WANG, B., LI, S. L. & TRUHLAR, D. G. „Modeling the Partial Atomic Charges in Inorganometallic Molecules and Solids and Charge Redistribution in Lithium-Ion Cathodes“. *Journal of Chemical Theory and Computation*. 2014, **10**(12), 5640-5650. DOI: 10.1021/ct500790p. ISSN 1549-9618. URL: <http://pubs.acs.org/doi/10.1021/ct500790p>
- [9] CELÝ, J. *Základy kvantové mechaniky pro chemiky: I. Principy*. Brno: Rektorát UJEP Brno, 1986

- [10] JEAN, Y., VOLATRON, F. & BURDETT, J. K. *An introduction to molecular orbitals*. New York: Oxford University Press, 1993. ISBN 0-19-506918-8.

