

Modern visualization of partial atomic charges in Mol*

Bachelor's thesis defense

Dominik Tichý

Faculty of Informatics, Masaryk University

June 30, 2023

Partial atomic charges

- used to describe the distribution of electrons in a molecule
- applications:
 - molecular docking
 - pharmacophore modeling
 - molecular dynamics simulations
- calculation methods:
 - quantum mechanics - extremely slow
 - empirical methods - much faster with slight cost to accuracy

SB NCBR software tools

- developed at the Structural bioinformatics research group (SB NCBR) in Brno
- programs for calculating partial atomic charges:
 - Atomic Charge Calculator II (ACC II)
 - α Charges (AlphaCharges)

Motivation

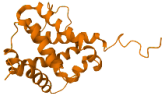
- ACC II uses the LiteMol Viewer
 - no longer maintained → needs to be replaced
- the Mol* Viewer is the modern replacement for LiteMol
 - no support for partial atomic charges

Thesis objectives

1. study the minimum necessary theory on partial atomic charges
2. extend the Mol* Viewer to support charge visualization
3. integrate updated Mol* Viewer into:
 - ACC II
 - α Charges

mmCIF file

Structure



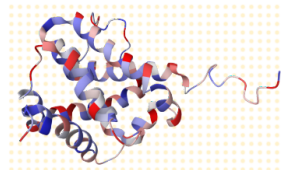
Charges

```

loop_
  sb_molr_partial_atomic_charges_meta_id
  sb_molr_partial_atomic_charges_meta_type
  sb_molr_partial_atomic_charges_meta_method
  1 empirical 'SQE+q5/Schindler 2021 (CCD_qm)'

loop_
  sb_molr_partial_atomic_charges_type_id
  sb_molr_partial_atomic_charges_atom_id
  sb_molr_partial_atomic_charges_charge
  1 1 -0.3266
  1 2 -0.4717
  1 3 0.6516
  1 4 -0.6844
    
```

⇒ Mol* ⇒



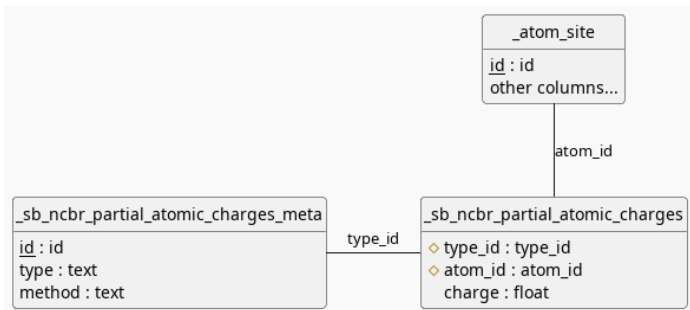
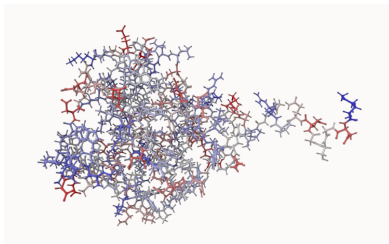
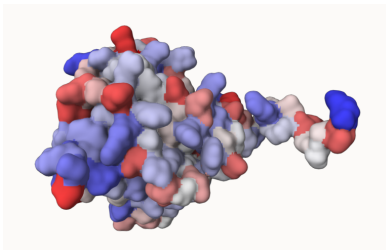


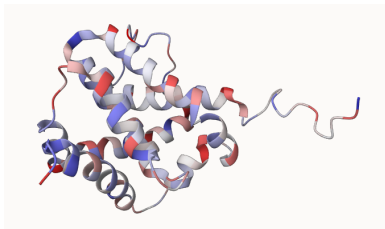
Figure: Diagram of the custom mmCIF categories



(a) Ball and stick



(b) Surface



(c) Cartoon

Integration of Mol* viewer into ACC II

- swapped the LiteMol Viewer for the Mol* Viewer
- added support for multiple calculations on one request
 - updated setup page to allow multiple calculations
 - added controls for switching between charge sets

Atomic Charge Calculator II

Computation results

Method: SQE+qp

Parameters: Schindler 2021 (CCD_gen)

Structure

1C0Q

Charge set

SQE+qp/Schindler 2021 (CCD_gen)

View

☒ Cartoon ☐ Balls and sticks ☐ Surface

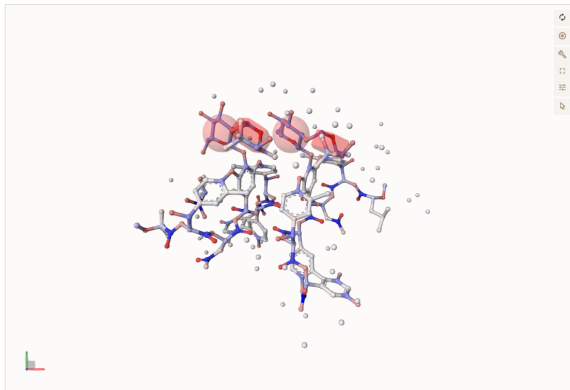
Coloring

☐ Structure

☒ Charges (relative) ☐ Charges (absolute)

Max value: 1.410

Reset



Download charges

Back to setup

Back to main page

Method

SQE+qp

Parameters

The most suitable parameters are shown first.

Schindler 2021 (CCD_gen)

[Add to calculation](#)**Full name**

Split-charge equilibration with parametrized initial charges

Publication

Schindler, O., Raček, T., Maršavelski, A., Koča, J., Berka, K., & Svobodová, R. (2021). Optimized SQE atomic charges for peptides accessible via a web application. Journal of cheminformatics, 13(1), 1-11. doi:10.1186/s13321-021-00528-w

Publication

Schindler, O., Raček, T., Maršavelski, A., Koča, J., Berka, K., & Svobodová, R. (2021). Optimized SQE atomic charges for peptides accessible via a web application. Journal of cheminformatics, 13(1), 1-11. doi:10.1186/s13321-021-00528-w

Will compute the following charges:

SQE+qp (Schindler 2021 (CCD_gen))



EEM (Racek 2016 (ccd2016_npa))



QEq (Rappe 1991)



EQeq (No parameters)

[Compute](#)[Back to main page](#)

αCharges – Calculation results

UniProt code: [Q9FFD0](#)

AlphaFold2 prediction version: 4

pH: 7.2

Number of atoms: 5508

View

- ☒ Cartoon
- ☐ Surface
- ☐ Ball & Stick

Coloring

- ☐ Structure
- ☐ Model confidence
- ☒ Charges (relative)
- ☐ Charges (absolute)

Max value: 0.9703

Reset



Download charges and protonated structure

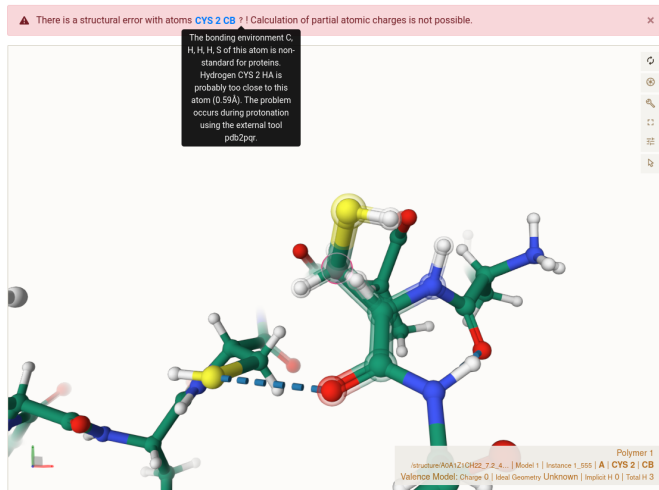
Back to main page

α Charges – Wrong structure

UniProt code: A0A1Z1CH22

AlphaFold2 prediction version: 4


pH: 7.2

[Download structure](#)[Back to main page](#)

Nucleic Acids Research paper

JOURNAL ARTICLE

α Charges: partial atomic charges for AlphaFold structures in high quality

Ondřej Schindler, Karel Berka, Alessio Cantara, Aleš Křenek, Dominik Tichý, Tomáš Raček ,
Radka Svobodová 

Nucleic Acids Research, gkad349, <https://doi.org/10.1093/nar/gkad349>

Published: 09 May 2023 **Article history** ▼

Figure: α Charges paper published in NAR [1]

Conclusion

- created an Mol* extension for visualizing partial atomic charges
- integrated the updated Mol* viewer into the ACC II and α Charges web applications
- extended the capabilities of ACC II to support multiple calculations on one request
- contributed with the work to the α Charges NAR paper

Oponents' questions

1. Proč jste pro vývoj Mol* pluginu zvolil Vite build tool? Zvažoval jste i jiné alternativy?
2. V jakých formátech je možno načíst náboje do Mol*?
3. Jste schopen využít funkcionalitu pro vizualizaci nábojů v Mol* i pro více přiložených molekul? Pokud ano, mohl byste to ukázat na příkladu? Např. podle nábojů obarvit a přiložit PIN proteiny, které jsou využity jako use case na webu α Charges?
4. Bylo by možno do budoucna podle nábojů obarvit i jiné objekty než atomy, elementy sekundární struktury nebo povrchy atomů - např. póry a kanály v proteinech? A pokud ano, jak pracné by to bylo?

Oponents' question 1

Proč jste pro vývoj Mol* pluginu zvolil Vite build tool? Zvažoval jste i jiné alternativy?

- Mol* repository uses Webpack for bundling
- Vite provides a much faster development server
 - **Fast start**
 - prebundles dependencies using *esbuild*
 - serves source code with native ESM modules
 - **Fast updates**
 - HMR is performed over native ESM
 - leverages HTTP headers to speed up full page reloads
- Vite provides optimized production builds with Rollup

Bundle based dev server

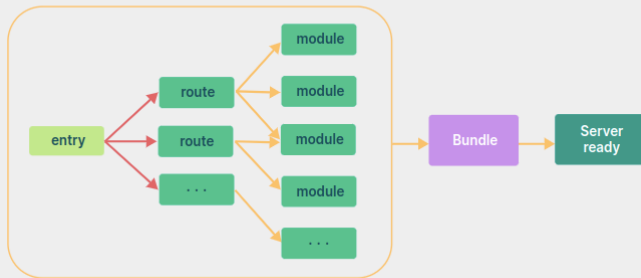


Figure: Bundle-based development server¹

¹<https://vitejs.dev/guide/why.html>

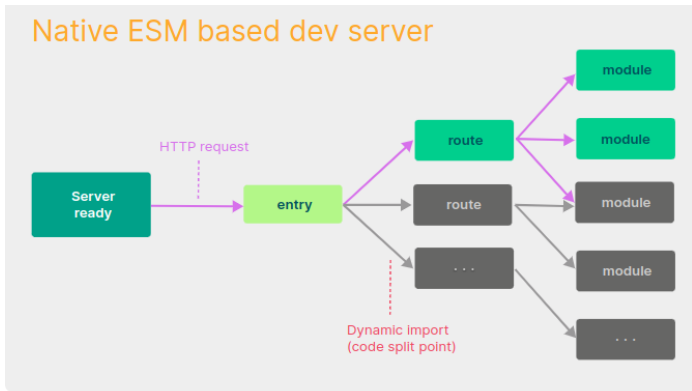


Figure: Native ESM module development server²

²<https://vitejs.dev/guide/why.html>

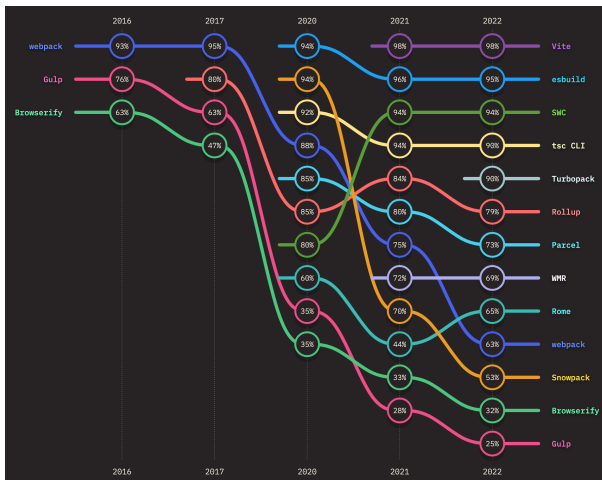


Figure: Build tools retention ranking ³

³<https://2022.stateofjs.com>

Oponents' question 2

V jakých formátech je možno načíst náboje do Mol*?

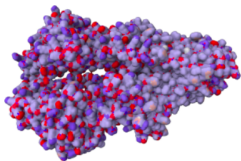
- the Mol* extension only supports the mmCIF format
- it is possible to add support for other formats which have data fields for charges
 - MOL2
 - PQR
- separate structure and charges files would require custom import controls
 - this approach was used in the LiteMol extension
 - would make the integration into Mol* more difficult

Oponents' question 3

Jste schopen využít funkcionalitu pro vizualizaci nábojů v Mol* i pro více přiložených molekul? Pokud ano, mohl byste to ukázat na příkladu? Např. podle nábojů obarvit a přiložit PIN proteiny, které jsou využity jako use case na webu α Charges?

Examples

P-glycoprotein

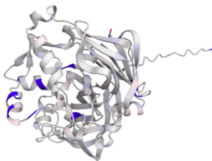


P-glycoprotein is part of the ABC transporter proteins that decrease drug accumulation in cancer cells ([Leslie2005](#)). It is a 170-kDa protein which consists of a nucleotide-binding domain and a transmembrane domain ([Ward2013](#)). Partial atomic charges calculated by [aCharges](#) demonstrate the differences in charge distribution between transmembrane parts and the extracellular/intracellular one.

P-glycoprotein

Pepsin

pH: 1

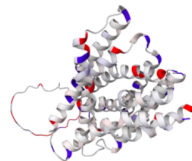


Pepsin is the enzyme that is majorly involved in protein digestion in the stomach. It is secreted as a zymogen and activated by the acidic pH created by the stomach parietal cells. Pepsin is the most effective at a pH of approximately 1.5 to 2, and it becomes inactive when the pH rises above 6 ([Heda2022](#)). Pepsin remains structurally stable until at least a pH of 8. Differences in a charge distribution between its active form (pH 2) and inactive form (pH 8) can be seen in this use case. The alkaline environment causes an increase in negative charges, which contributes to the structural instability of the pepsin, causing the shift to an inactive form of the protein ([Tanaka2001](#), [Grahame2021](#)).

Active

Inactive

PIN proteins



The PIN family proteins control plant growth by regulating auxin export from the cytosol to the extracellular space. Eight types of PIN proteins are known (PIN1-PIN8), and last year, structures of three of them were discovered and published in Nature (i.e., articles about [PIN1](#), [PIN3](#), [PIN8](#)). Partial atomic charges play an important role in PINs functionality. The PIN protein part inside the cytosol (containing the cytosolic loop) is charged more than the part outside the cytosol. Questionable is the charge distribution of PIN5, which structure differs from other PINs ([Ung2022](#)) and was not experimentally determined yet. In the use case, you can compare the charge distribution of AlphaFold2 predicted PINs from *A. thaliana*.

PIN1

PIN3

PIN5

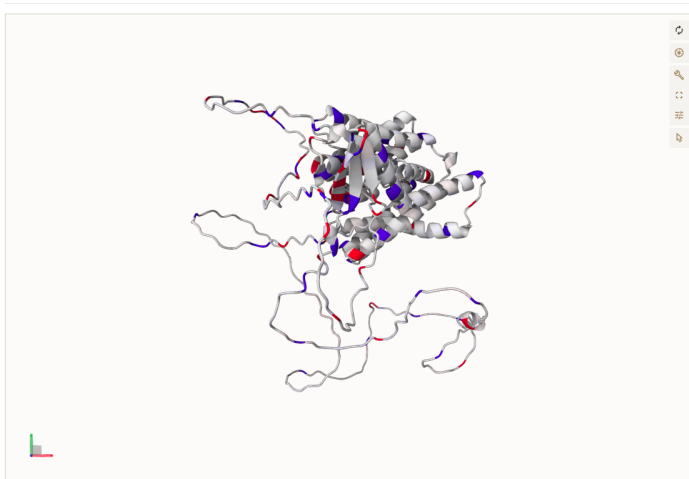
PIN6

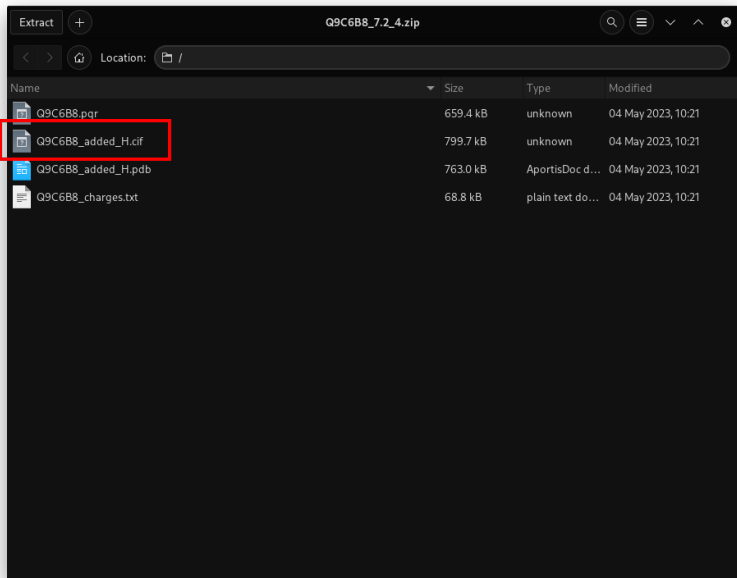
PIN7

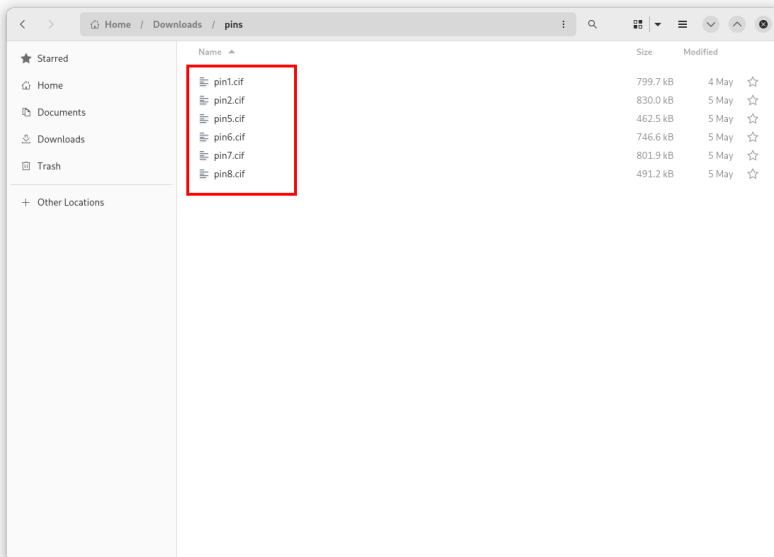
PIN8

Max value: 1.031

Reset

[Download charges and protonated structure](#)[Back to main page](#)





pin2.cif

pin7.cif

pin1.cif

pin6.cif

pin8.cif

pin5.cif

Q9FFD0_added_H1 model

Model 1

Model 5508 elements

Polymer 5508 elements

Sequence of Q9S7Z8_add... Chan A: A-p [auth A]

RISMDLYTVLTAVIPLYVAR¹²¹ILAYSVVRAKIFSPDQCG²¹¹INRFVAIFAV³¹¹LLSFHFTSTNNPYANNIL⁴¹¹LAADTLQK⁵¹¹IIRLSLLVMA⁶¹¹TRSGSELS⁷¹¹YIFSLSTL⁸¹¹NTLVNS⁹¹¹TPLL¹⁰¹¹

IAHYGEYSGLMVQIVVLQ¹¹¹CIIMYTL¹²¹LLCFEFGRAK¹³¹LINEQFETA¹⁴¹SVSVKVESV¹⁵¹SLDGDH¹⁶¹FLETD¹⁷¹AEI¹⁸¹GGGL¹⁹¹LVTVYKSN²⁰¹ASRRSFCGPW²¹¹TPRPS²²¹NLTGA²³¹EIYSL²⁴¹STTP²⁵¹RO

SNFNHSD²⁶¹YNNMGFP²⁷¹GGRL²⁸¹SNFGPADMY²⁹¹SVQSSRG³⁰¹TPRPS³¹¹NFEENCA³²¹ASSPRF³³¹GYPGGAG³⁴¹SYPA³⁵¹NP³⁶¹EFSS³⁷¹TTT³⁸¹TANKSV³⁹¹NK⁴⁰¹PKD⁴¹¹YNT⁴²¹Q⁴³¹IT⁴⁴¹LPT⁴⁵¹GGKSN⁴⁶¹SHDA⁴⁷¹KEL⁴⁸¹HP⁴⁹¹W⁵⁰¹SS

Apply Transform Configuration

Kind	Components
Axis	[1.00, 0.00, 0.00]
Angle	<input type="text" value="0"/>
Translation	[0.00, 0.00, 0.00]
X	0
Y	0
Z	0

✓ Apply

Cartoon

Structure Tools

Structure

6 structures

Nothing Focused

Measurements

+ Add

Superposition

Chains Atoms

Quick Styles

Default Stylized Illustrative

Components 6 structures

Presets + Add

Empty

Automatic

Basic

Miscellaneous

Membrane Orientation

Quality Assessment (pLDDT)

SB NCBP Partial Charges

Automatic (w/ Annotation)

Polymer

Export Models

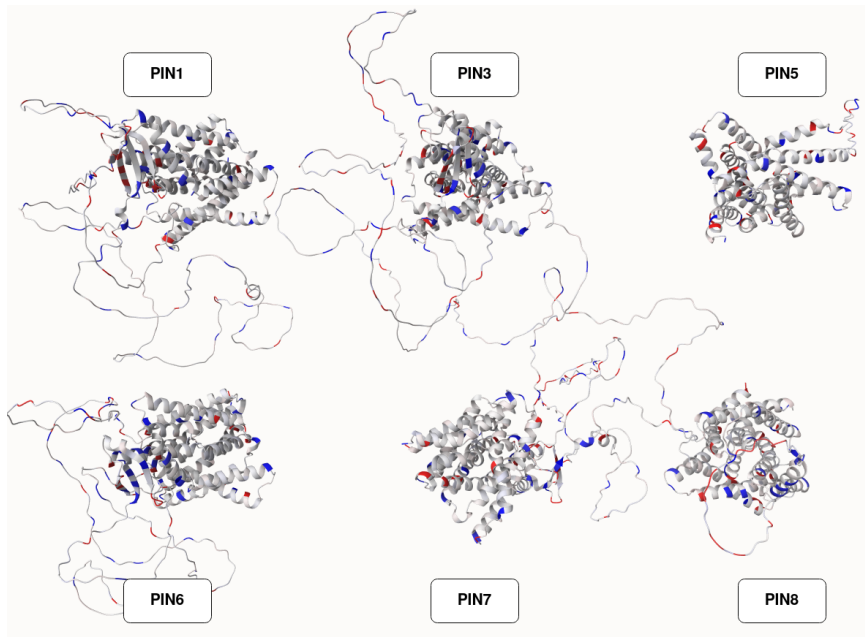
Export Animation

Export Geometry

18:56:57 Updated Cartoon in 4ms.

18:56:57 Updated Cartoon in 5ms.

18:56:57 Updated Structure Focus Representation in 1ms.



Oponents' question 4

Bylo by možno do budoucna podle nábojů obarvit i jiné objekty než atomy, elementy sekundární struktury nebo povrchy atomů - např. póry a kanály v proteinech? A pokud ano, jak pracné by to bylo?

- the coloring uses a Location object to determine which color to assign to the given representation element (e.g. atom, residue, surface)
- it would be necessary to create a custom representation (e.g. for pores, channels)
 - the representation would need to provide a Location object for each of its elements

Bibliography I

- [1] Ondřej Schindler et al. “ α Charges: partial atomic charges for AlphaFold structures in high quality”. In: *Nucleic Acids Research* (May 2023). gkad349. ISSN: 0305-1048. DOI: 10.1093/nar/gkad349. eprint: <https://academic.oup.com/nar/advance-article-pdf/doi/10.1093/nar/gkad349/50252244/gkad349.pdf>. URL: <https://doi.org/10.1093/nar/gkad349>.

MUNI

FACULTY

OF INFORMATICS