

# Modern visualization of partial atomic charges in Mol\*

Bachelor's thesis defense

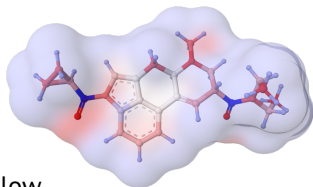
**Dominik Tichý**

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



# Empirical methods – implementations

- can contain parameters
- web applications developed at SB NCBR for calculating partial atomic charges:
  - Atomic Charge Calculator II (ACC II)<sup>1</sup>
  - $\alpha$ Charges (AlphaCharges)

---

<sup>1</sup>Tomáš Raček et al. “Atomic Charge Calculator II: web-based tool for the calculation of partial atomic charges”. In: *Nucleic Acids Research* 48.W1 (May 2020), W591–W596. DOI: 10.1093/nar/gkaa367.

# Motivation & Thesis goals

## Motivation:

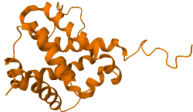
- ACC II uses the LiteMol Viewer
  - no longer maintained → needs to be replaced
- the Mol\* Viewer is the modern replacement for LiteMol
  - no support for partial atomic charges

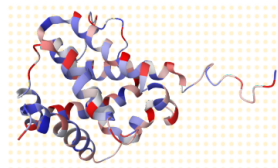
## Goals:

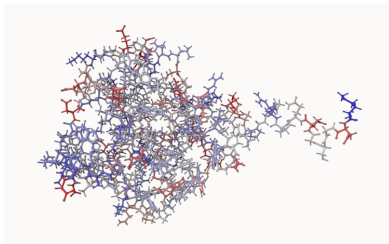
1. study the minimum necessary theory on partial atomic charges
2. extend the Mol\* Viewer to support charge visualization
3. integrate updated Mol\* Viewer into:
  - ACC II
  - $\alpha$ Charges

# Mol\* extension

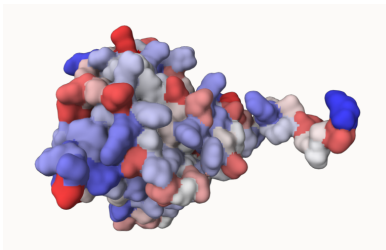
- created mmCIF categories for storing partial atomic charges
- extended Mol\* Viewer to support visualization of partial charges

Structure

Charges
<pre>loop_ _sb_ncbr_partial_atomic_charges_meta.id _sb_ncbr_partial_atomic_charges_meta.type _sb_ncbr_partial_atomic_charges_meta.method 1 'empirical' 'SQE+qp/Schindler 2021 (CCD_gen)'  loop_ _sb_ncbr_partial_atomic_charges.type_id _sb_ncbr_partial_atomic_charges.atom_id _sb_ncbr_partial_atomic_charges.charge 1 1 -0.5288 1 2 -0.0717 1 3 0.6516 1 4 -0.6044</pre>

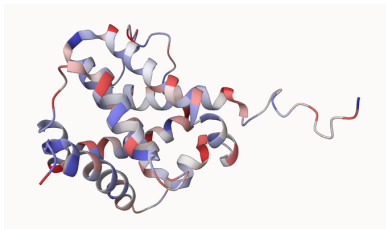




(a) Ball and stick



(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

Structure  
155884675

Charge set  
SQE+qp/Schindler 2021 (CCD\_gen)

View  
☐ Cartoon ☒ Balls and sticks ☐ Surface

Coloring  
☐ Structure ☒ Charges (relative) ☐ Charges (absolute)

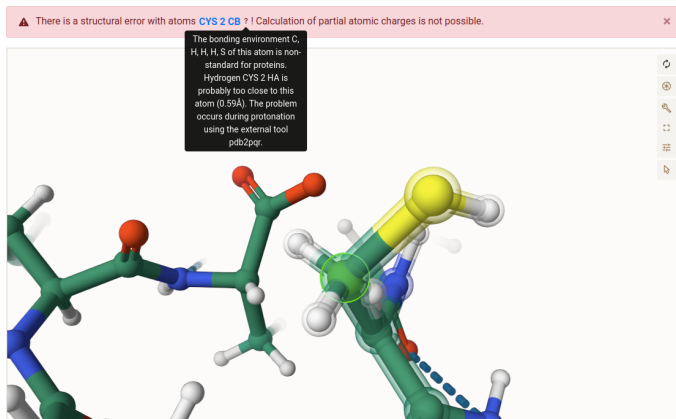
Max value: 0.6333 Reset



The image displays the Mol\* viewer interface. At the top, there are controls for 'Structure' (ID: 155884675) and 'Charge set' (SQE+qp/Schindler 2021 (CCD\_gen)). Below these are 'View' options (Cartoon, Balls and sticks, Surface) and 'Coloring' options (Structure, Charges (relative), Charges (absolute)). A 'Max value' field shows 0.6333 with a 'Reset' button. The main area shows a 3D ball-and-stick model of a complex organic molecule with partial atomic charges represented by a color gradient from blue (negative) to red (positive). The molecule features a central benzene ring fused to a five-membered ring, with various side chains including amide and ester groups. A vertical toolbar on the right contains icons for zooming, rotating, and other standard molecular viewer functions.

# Integration of Mol\* viewer into $\alpha$ Charges

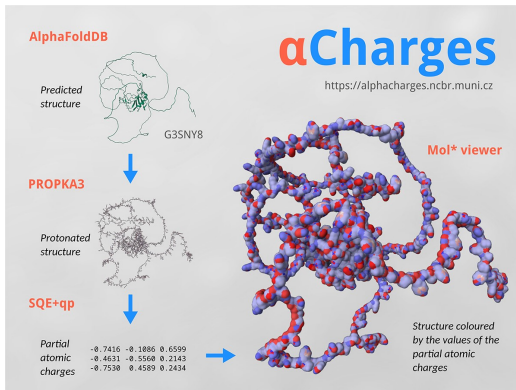
- Mol\* Viewer is used for displaying the results
- new coloring by pLDDT confidence score
- added a page for describing problematic structures





# Nucleic Acids Research paper

- impact factor: 19.160 (2021)
- Ondřej Schindler, Karel Berka, Alessio Cantara, Aleš Křenek, Dominik Tichý et al. “ $\alpha$ Charges: partial atomic charges for AlphaFold structures in high quality”. In: *Nucleic Acids Research* (2023). DOI: 10.1093/nar/gkad349



# Conclusion

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

## Otázka oponenta 1

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

- Mol\* používá Webpack
- Vite poskytuje rychlejší development server a optimalizované produkční buildy

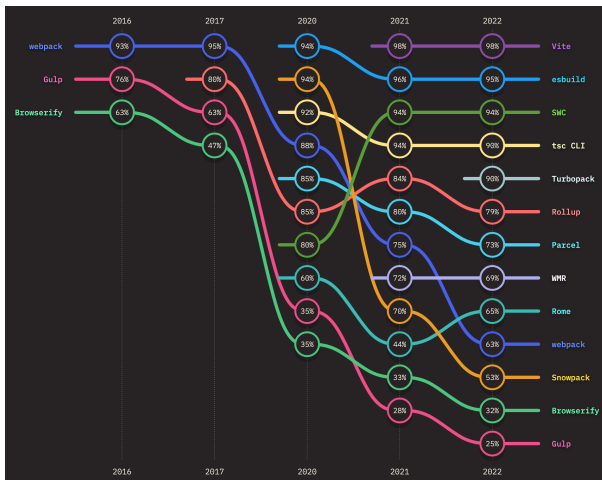


Figure: Retence build nástrojů v průzkumu State of JS 2022 <sup>2</sup>

<sup>2</sup><https://2022.stateofjs.com>

## Otázka oponenta 2

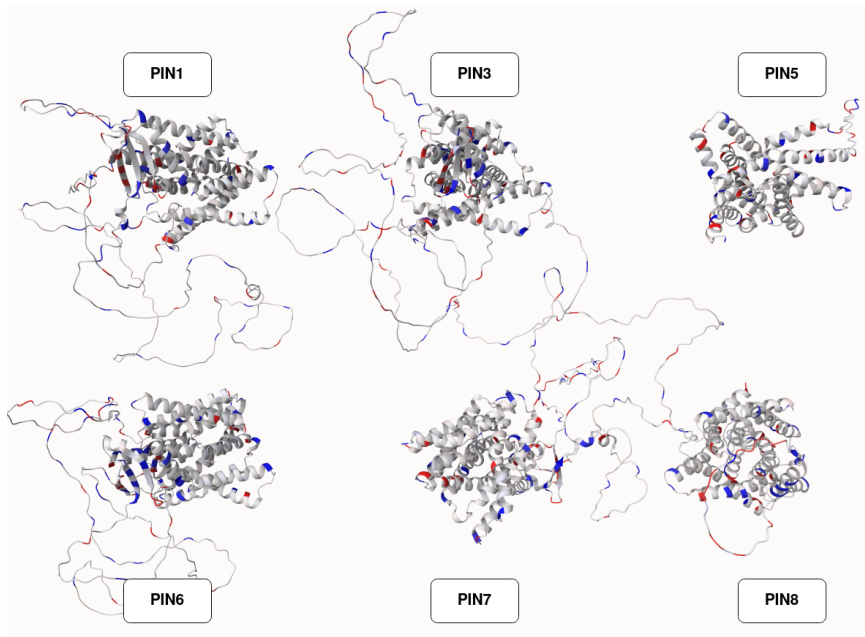
V jakých formátech je možno načíst náboje do Mol\*?

- Mol\* umožňuje načíst náboje pouze v mmCIF formátu
- je možné přidat podporu pro další formáty, které mají datová pole pro náboje
  - MOL2
  - PQR
- lze také oddělit strukturu od nábojů a načítat je zvlášť
  - vyžadovalo by to vytvořit importy pro náboje
  - tento přístup byl použit v rozšíření LiteMol

## Otázka oponenta 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  $\alpha$ Charges?

- ano, struktury s vypočítanými náboji z  $\alpha$ Charges lze nahrát do Mol\* a v něm provést přiložení



## Otázka oponenta 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?

- obarvování používá Location objekt, podle kterého se určí jakou barvu obarvit daný prvek reprezentace (např. atom, residue, surface)
- bylo by nutné vytvořit novou 3D reprezentaci (např. pro póry, kanály)
  - tato reprezentace by musela poskytnout Location objekty pro každou část póru/kanálu



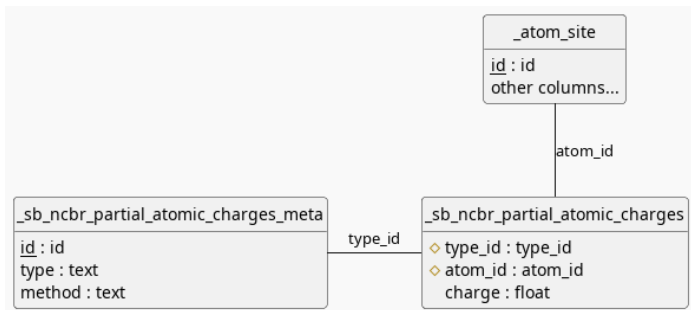
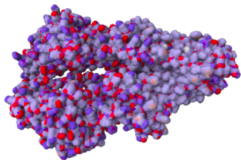


Figure: Diagram of the custom mmCIF categories

## Examples

### P-glycoprotein

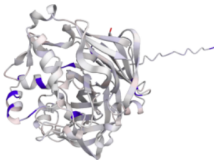


P-glycoprotein is part of the ABC transporter proteins that decrease drug accumulation in cancer cells ([Leslie2005](#)). It is a 170-kDa protein which consists of a nucleotide-binding domain and a transmembrane domain ([Ward2013](#)). Partial atomic charges calculated by [aCharges](#) demonstrate the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

P-glycoprotein

### Pepsin

pH: 1

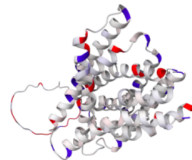


Pepsin is the enzyme that is majorly involved in protein digestion in the stomach. It is secreted as a zymogen and activated by the acidic pH created by the stomach parietal cells. Pepsin is the most effective at a pH of approximately 1.5 to 2, and it becomes inactive when the pH rises above 6 ([Heda2022](#)). 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

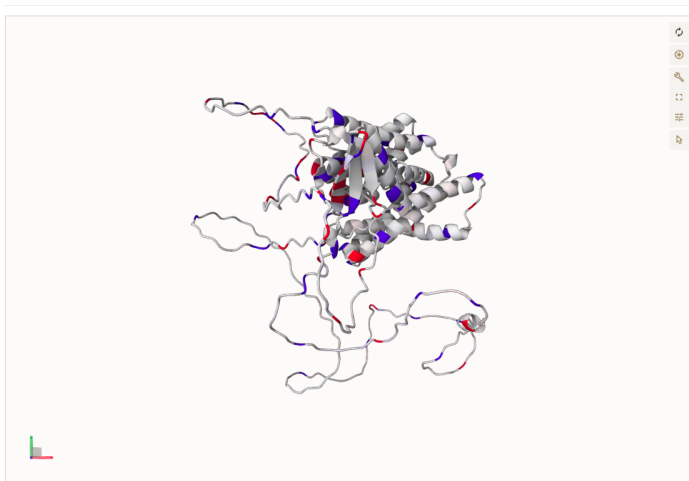
PIN6

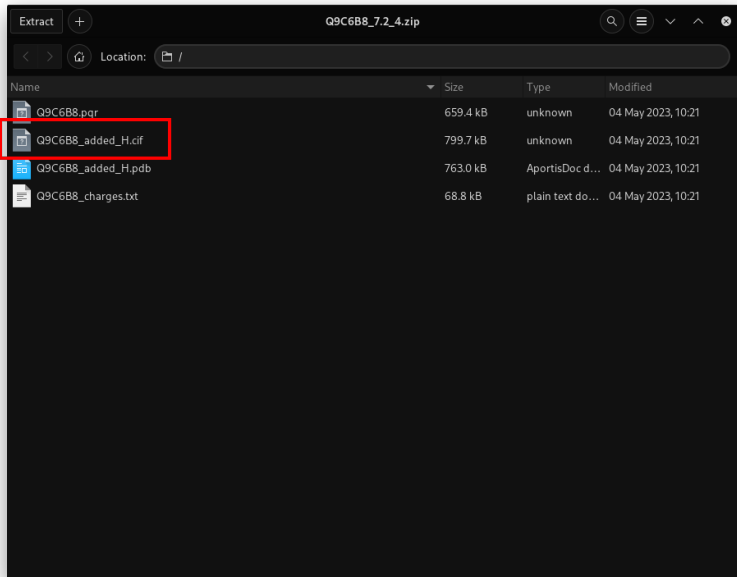
PIN7

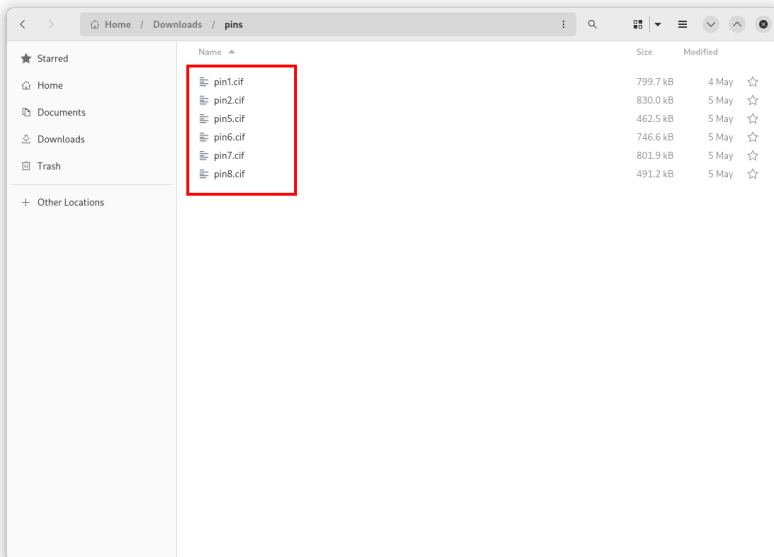
PIN8

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 Q9SZ8\_add... Chan A: A-p [auth A]

H T S M D L Y T V L T A Y I P L Y V A R I L A Y S V P W A K I F S P Q C G I N R F V A I F A P L L S F H F I S T N P Y A N H L R P L A A D T L Q K I I R L S L V L M A N F T R S G L E M S I Y I F S L S T U R N T L V N S T P L L  
 I A M Y G E Y S G L M V Q I V L V Q C I I M Y T L L L C L F E F R G A K N L I E Q F P E T A S I V S F K V E S V S L D G H D F L E T D A E I G D G K L H V T Y R K S N A S R S F C G P W T P R P S N L T G A E I Y S L T T P R D  
 S N F N H S D F Y M M G F P G G R L S N F G P A D M Y S V Q S S R G P T R P R S N F E E M C A N A S S P R F G Y P G G A G S Y P A P R P E F S T T T T S T A N K S V N K N P K D Y N T N Q I T L P T G G K S N S H D A K E L H M F V S S

Apply Transform Configuration

Kind	Components
Axis	[1.00, 0.00, 0.00]
Angle	<input type="range"/> 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)

Export Models

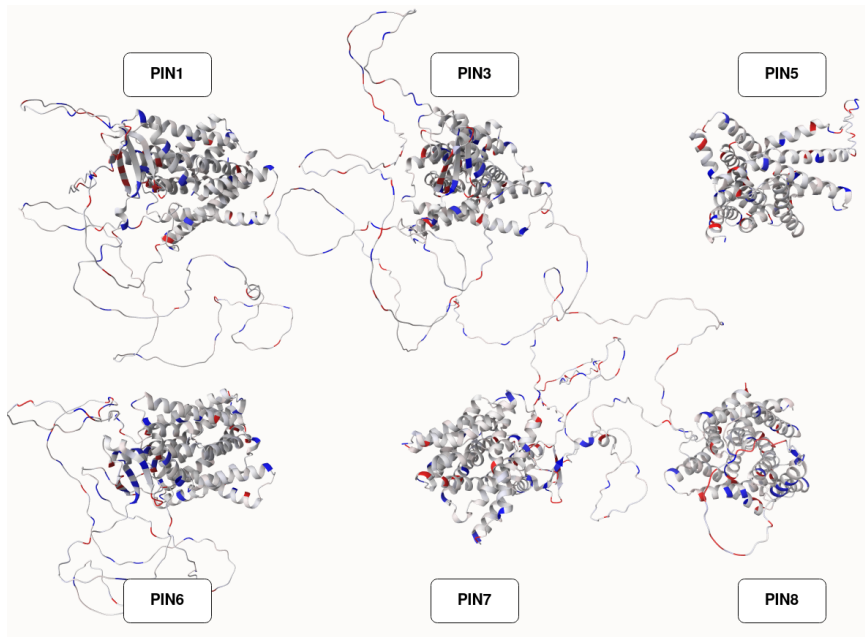
Export Animation

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.



## ■ 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



## Bundle based dev server

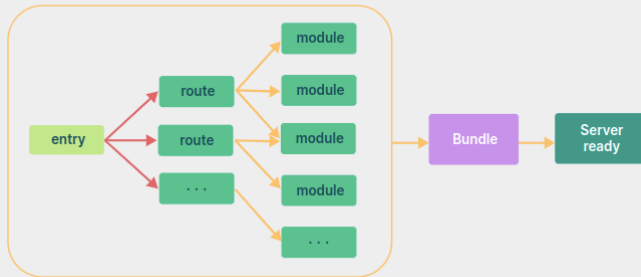


Figure: Bundle-based development server <sup>3</sup>

<sup>3</sup><https://vitejs.dev/guide/why.html>

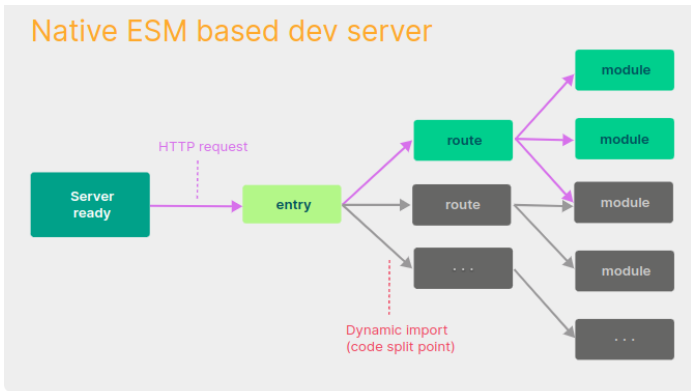


Figure: Native ESM module development server<sup>4</sup>

<sup>4</sup><https://vitejs.dev/guide/why.html>

**MUNI**

FACULTY

OF INFORMATICS