

Modern visualization of partial atomic charges in Mol*

Bachelor's thesis defense

Dominik Tichý

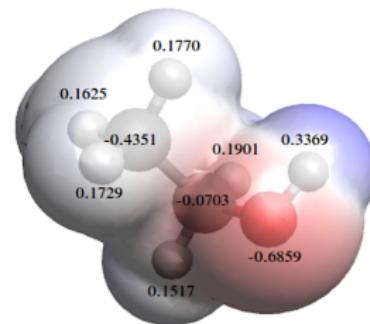
Supervisor: RNDr. Tomáš Raček, Ph.D.

Faculty of Informatics, Masaryk University

June 30, 2023

Partial atomic charges

- real numbers describing the distribution of electron density among atoms 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

- web applications for calculating partial atomic charges:
 - Atomic Charge Calculator II (ACC II)¹
 - α Charges (AlphaCharges)²
- developed by the Structural Bioinformatics research group at the National Centre for Biomolecular Research

¹ Tomáš Raček et al. “Atomic Charge Calculator II: web-based tool for the calculation of partial atomic charges”. In: *NAR* (2020).

² Ondřej Schindler et al. “ α Charges: partial atomic charges for AlphaFold structures in high quality”. In: *NAR* (2023).

Motivation & Thesis goals

Motivation:

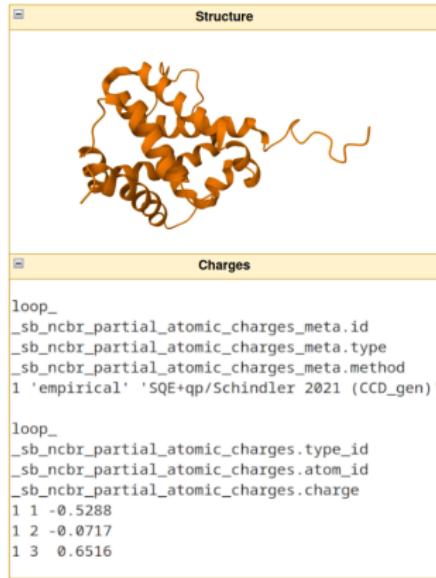
- both ACC II and α Charges use the LiteMol Viewer
 - no longer maintained → needs to be replaced
- the Mol* Viewer is the modern replacement for LiteMol
 - no support for partial atomic charges

Goals:

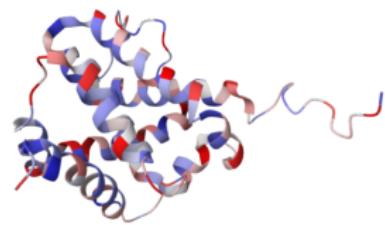
1. study the necessary theory on partial atomic charges
2. extend the Mol* Viewer to support charge visualization
3. integrate the updated Mol* Viewer into:
 - ACC II
 - α Charges

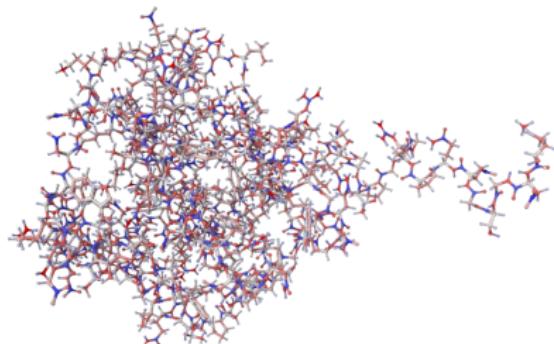
Mol* extension

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

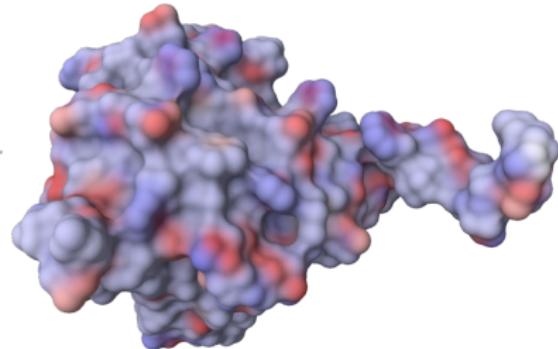


→ Mol*

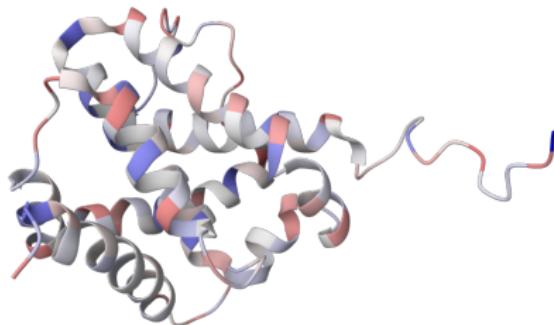




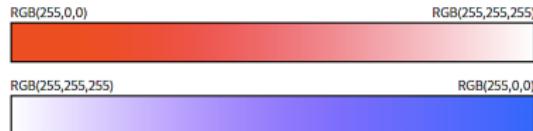
(a) Ball and stick



(b) Surface



(c) Cartoon



(d) Color gradients

Integration of Mol* viewer into ACC II

- replaced the LiteMol Viewer with the Mol* Viewer
- added support for multiple calculations on one request

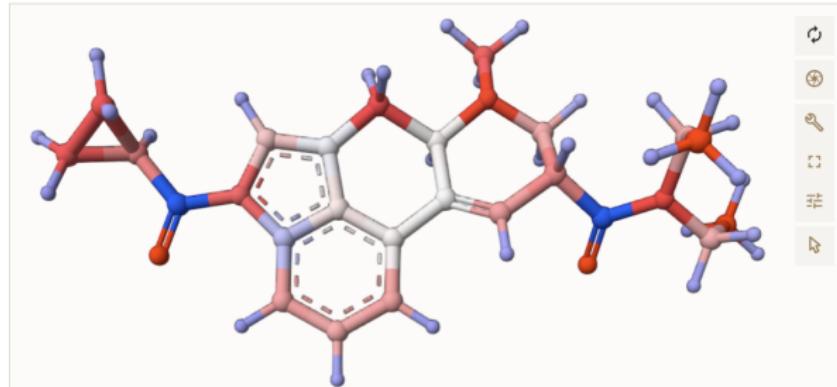
Structure
1CP_LSD

Charge set
SQE+qp/Schindler 2021 (CCD_gen)

View
 Cartoon Balls and sticks Surface

Coloring
 Structure Charges (relative) Charges (absolute)

Max value: 0.6381 Reset



Integration of Mol* viewer into α Charges

- the Mol* Viewer is used for displaying the results
- added a page for describing problematic atoms

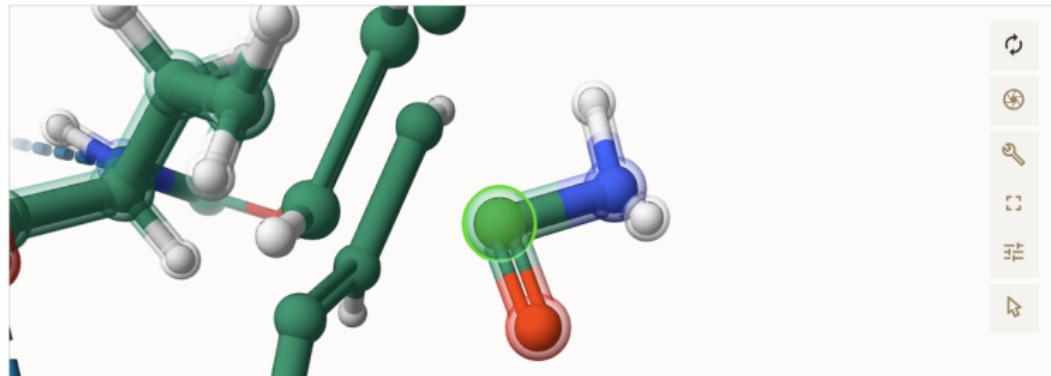
UniProt code: Q55GB6

AlphaFold2 prediction version: 4

pH: 7.2

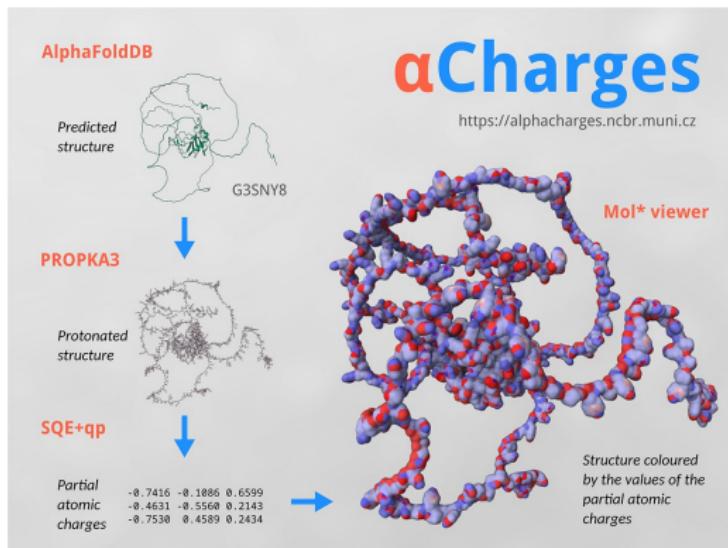
The bonding environment N,
O of this atom is non-
standard for proteins. This
atom is probably wrongly
predicted by AlphaFold2.

⚠ There is a structural error with atoms **GLN 33 CD ?**, **GLN 33 CG ?** ! Calculation of partial atomic charges is not possible. ✖



Nucleic Acids Research paper

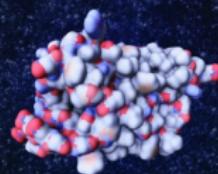
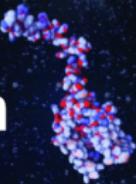
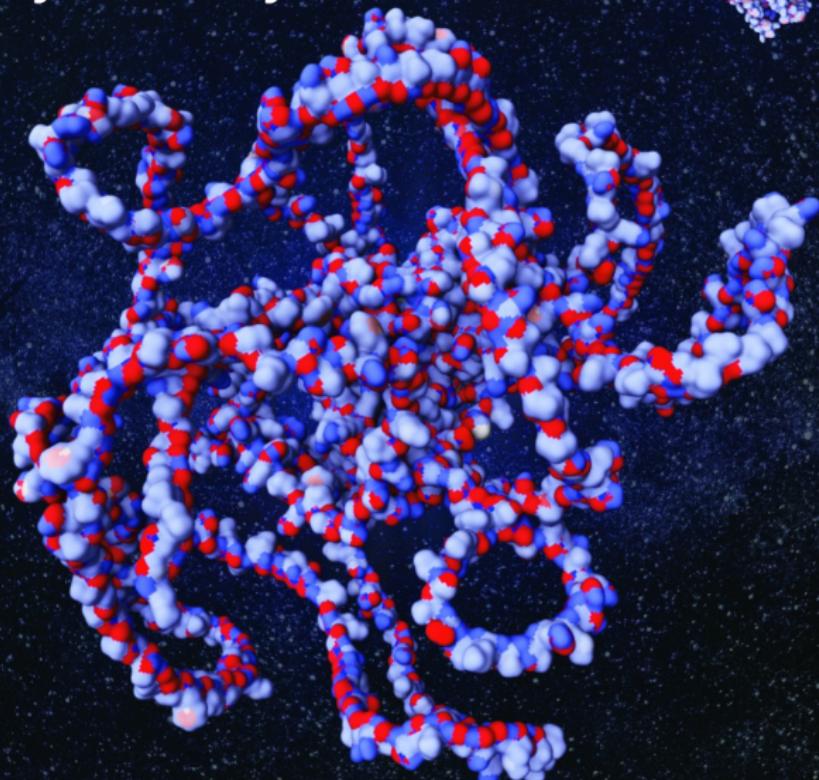
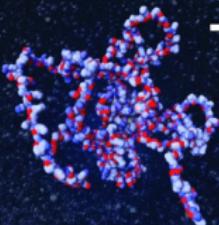
- the current impact factor of the journal is **19.160**
- Ondřej Schindler, Karel Berka, Alessio Cantara, Aleš Křenek, **Dominik Tichý et al.** “ α Charges: partial atomic charges for AlphaFold structures in high quality”. In: NAR (2023)



Conclusion

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

Thank you for your attention



Otázky 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

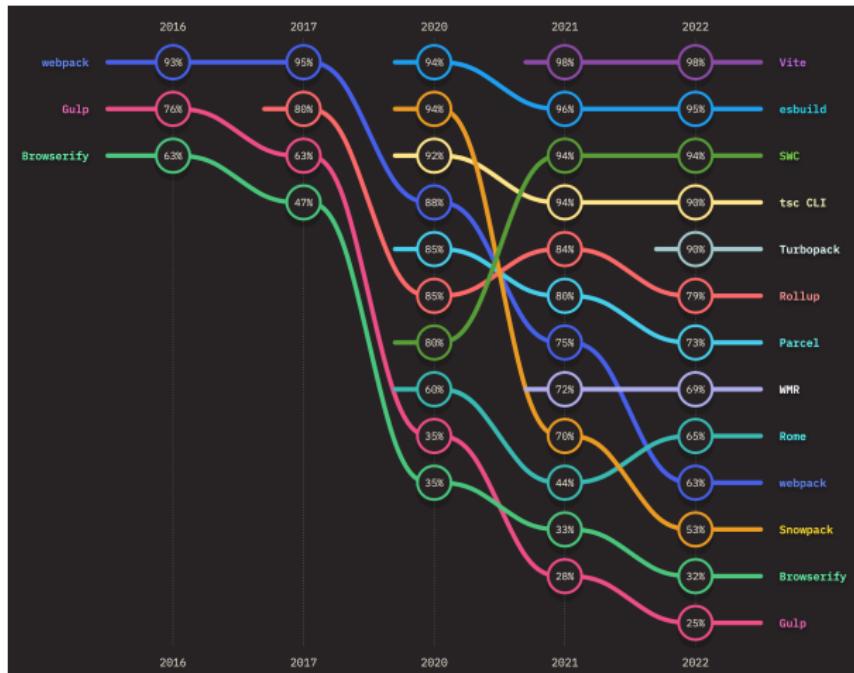


Figure: Retence build nástrojů v průzkumu State of JS 2022 ³

³ <https://2022.stateofjs.com>

Otázky oponenta 2

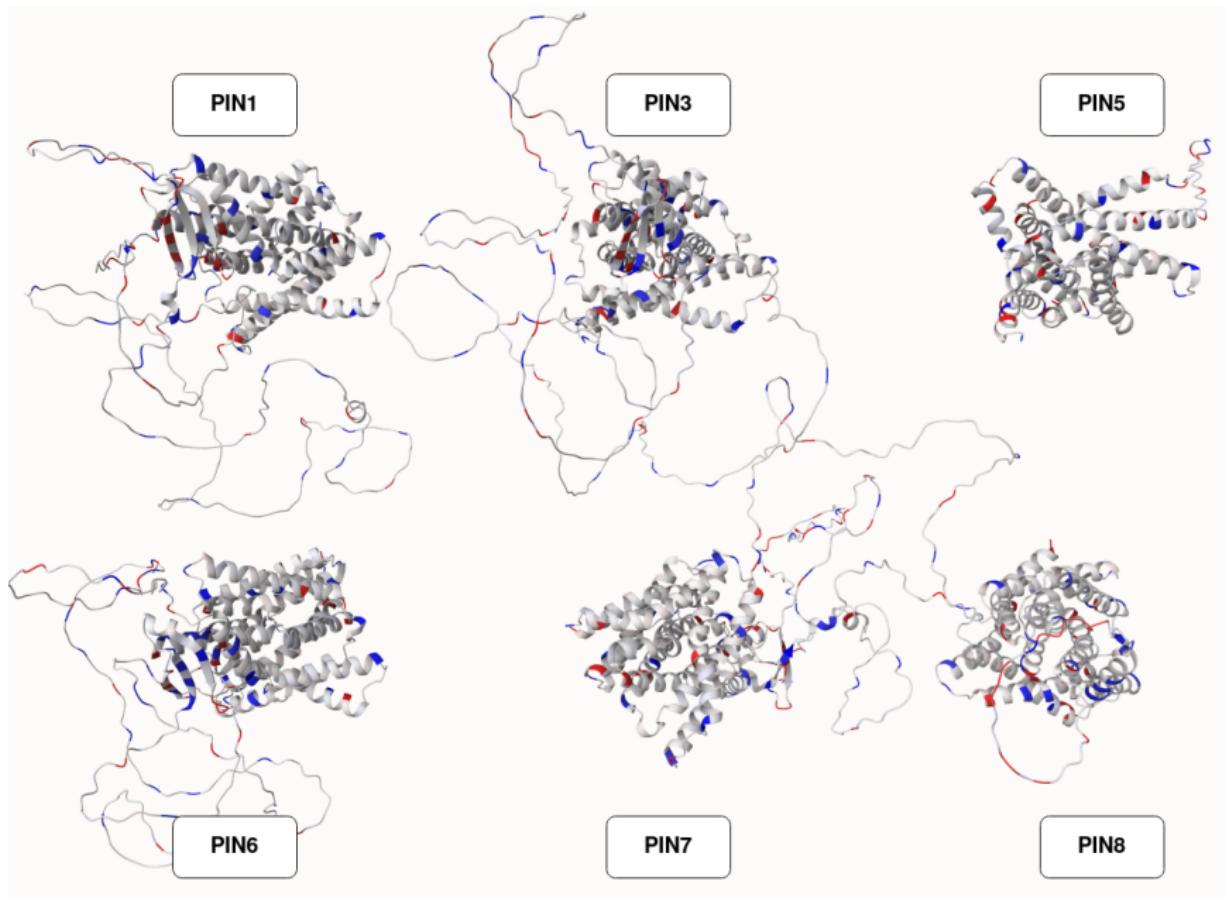
V jakých formátech je možno načíst náboje do Mol*?

- Mol* Viewer 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 nástroje na importování nábojů
 - tento přístup byl použit v rozšíření LiteMol Vieweru

Otázky oponenta 3

Jste schopen využít funkcionality 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?

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



Otázky 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?

- ano, je to možné
- časový odhad: pár týdnů

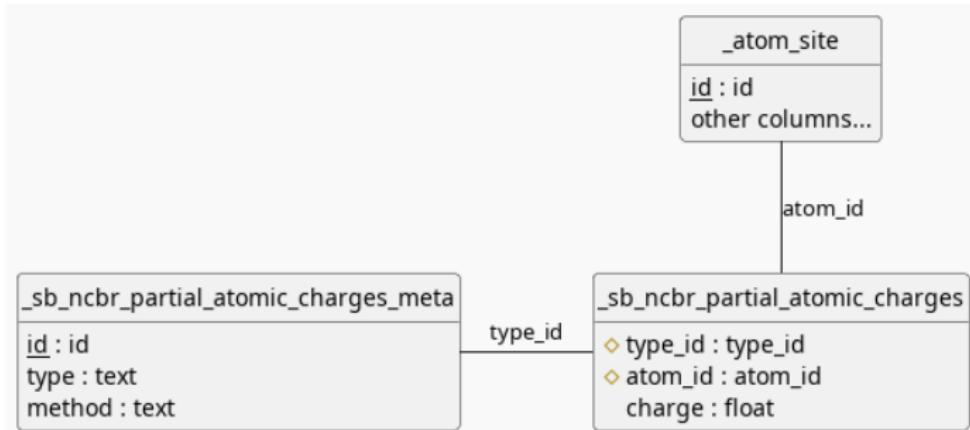
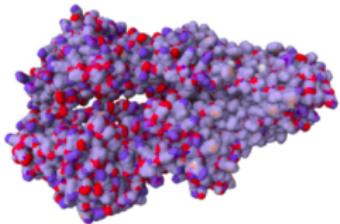


Figure: Diagram of the custom mmCIF categories

Examples

P-glycoprotein

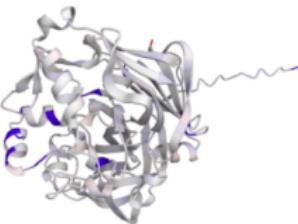


P-glycoprotein is part of the ABC transporter proteins that decrease drug accumulation in cancer cells ([Leslie2005](#)). It is a 170-kDa protein which consists of a nucleotide-binding domain and a transmembrane domain ([Ward2013](#)). Partial atomic charges calculated by [eCharges](#) demonstrate the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

P-glycoprotein

Pepsin

pH: 1

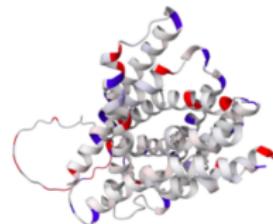


Pepsin is the enzyme that is majorly involved in protein digestion in the stomach. It is secreted as a zymogen and activated by the acidic pH created by the stomach parietal cells. Pepsin is the most effective at a pH of approximately 1.5 to 2, and it becomes inactive when the pH rises above 6 ([Heda2022](#)).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

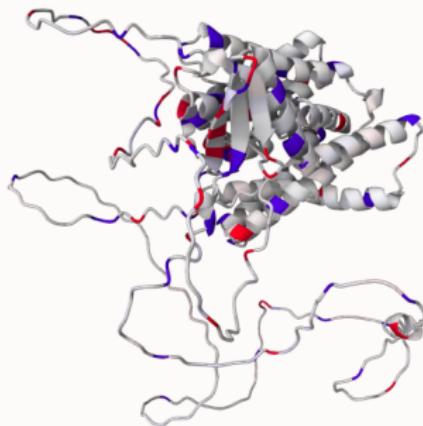
PIN7

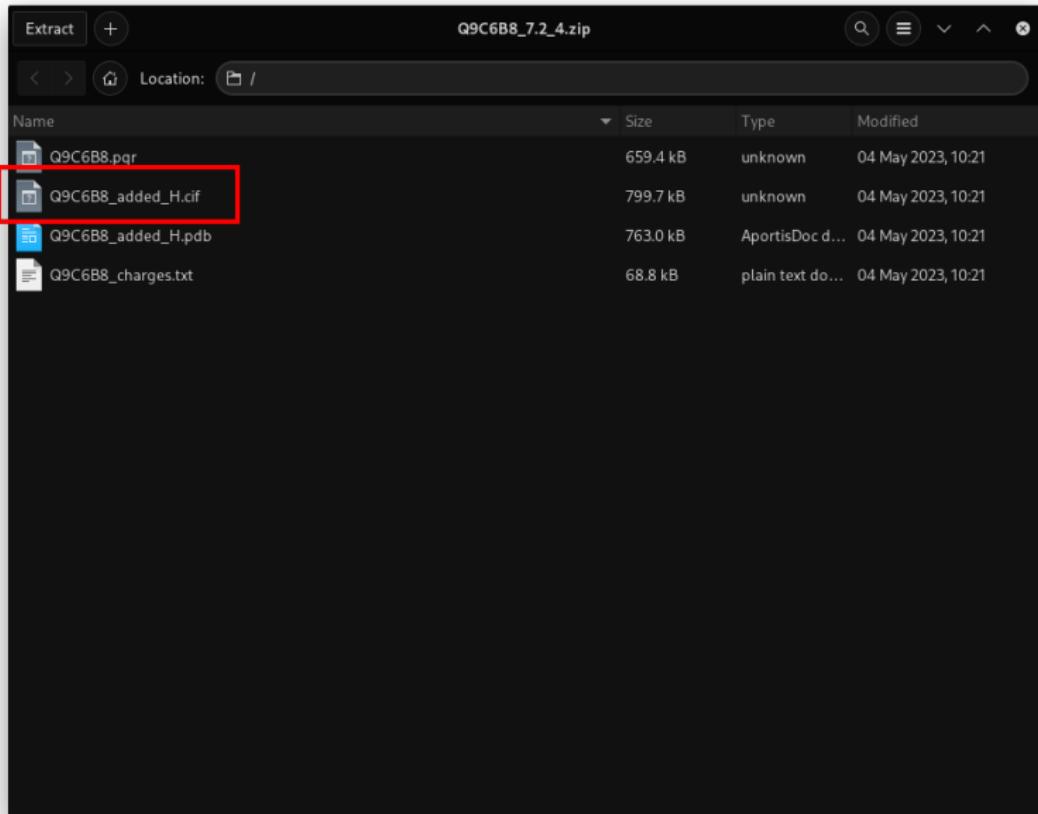
PIN8

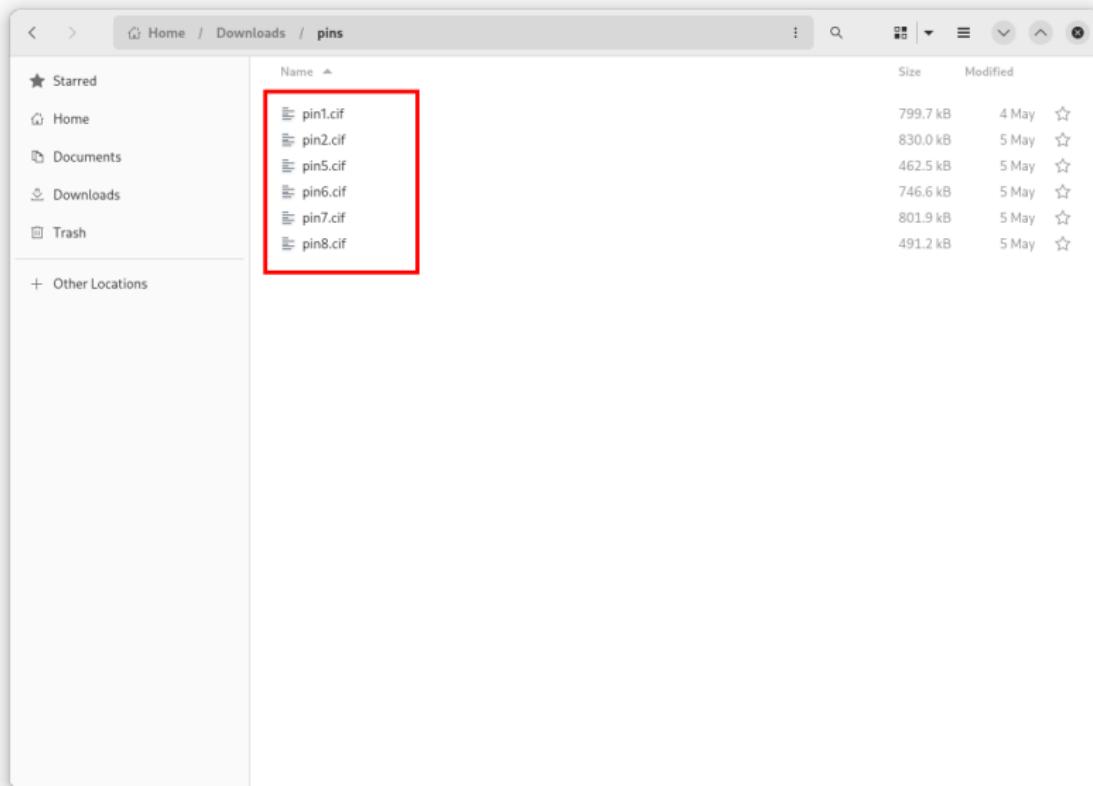
Max value:

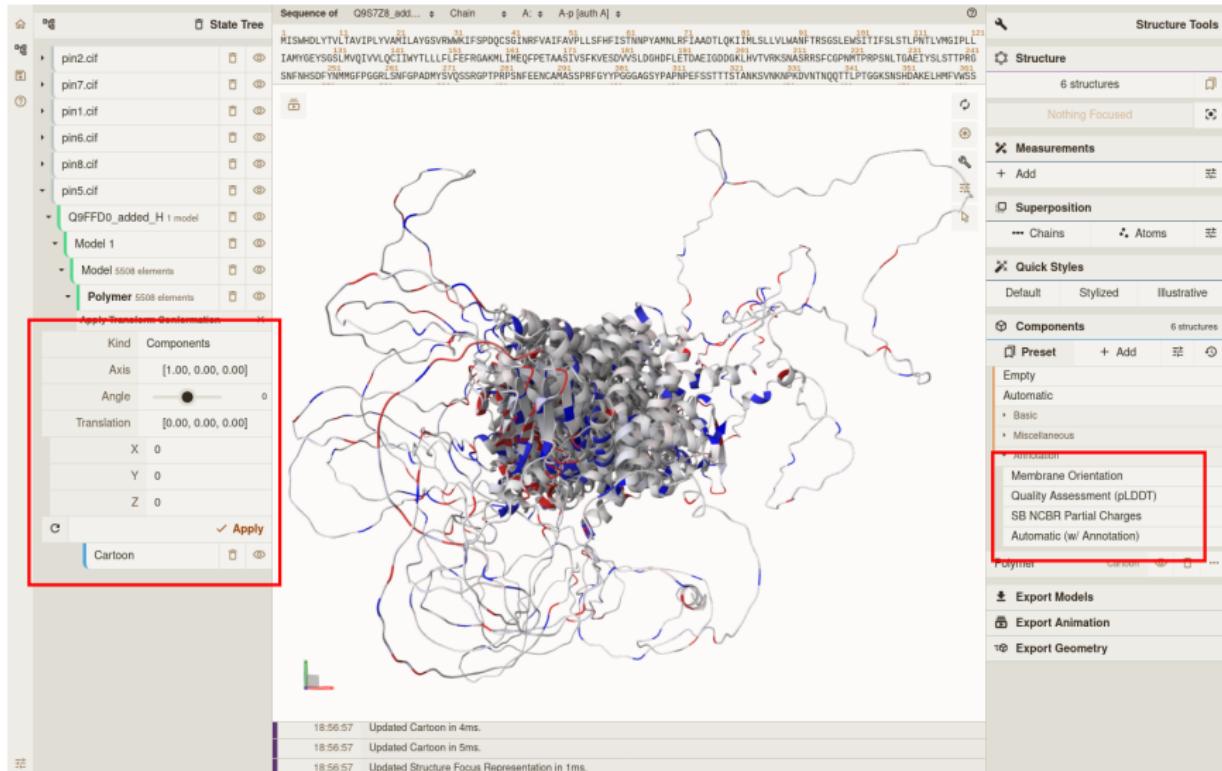
1.031

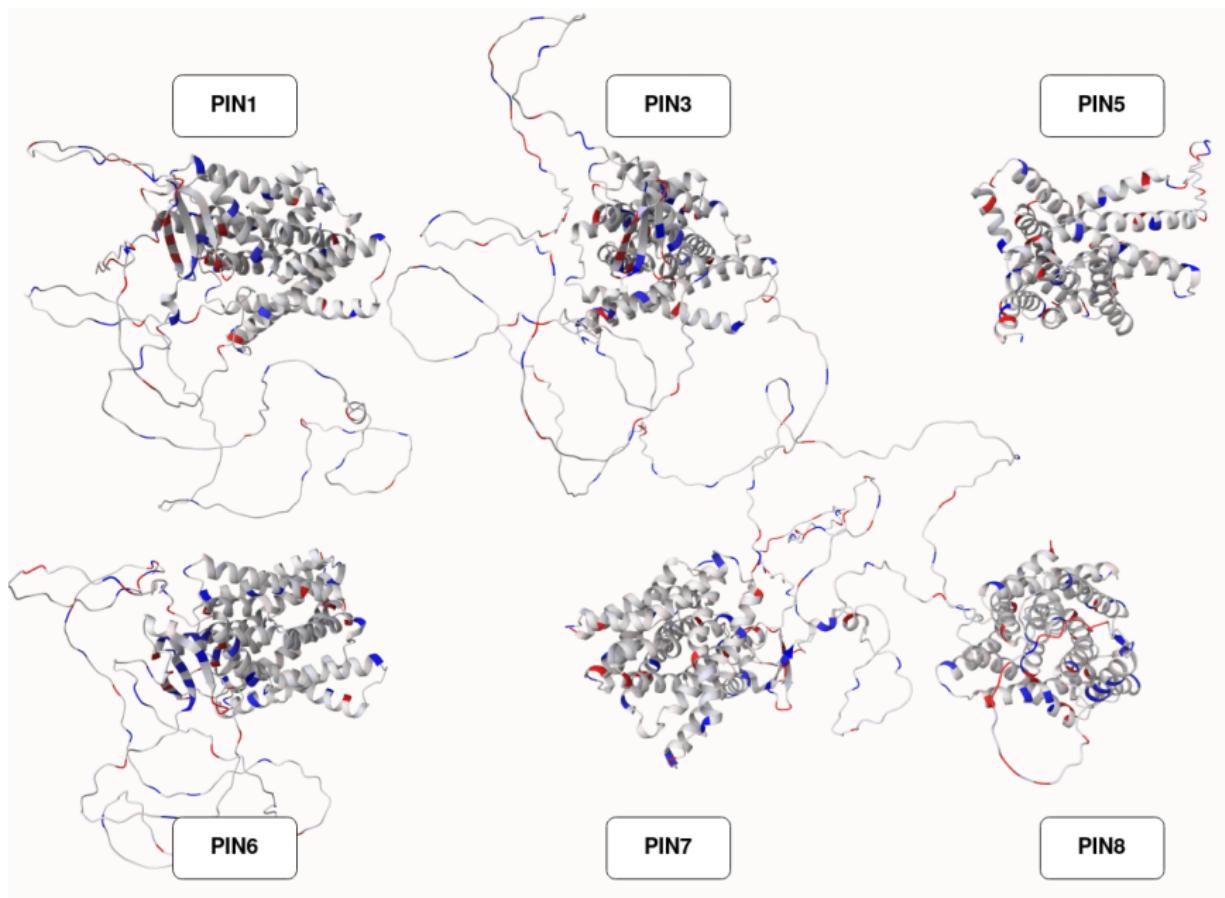
Reset

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









■ 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

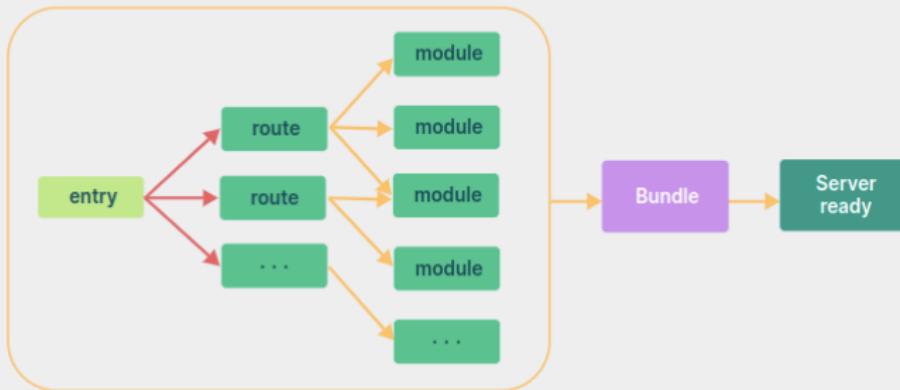


Figure: Bundle-based development server ⁴

⁴<https://vitejs.dev/guide/why.html>

Native ESM based dev server

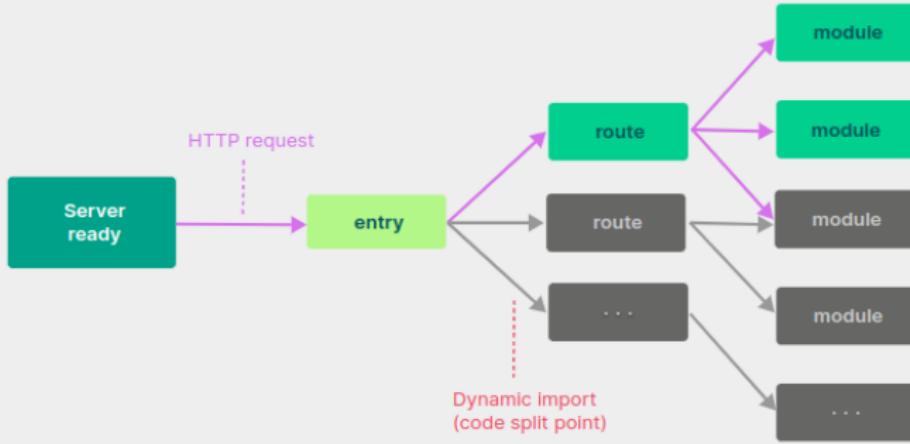


Figure: Native ESM module development server ⁵

⁵<https://vitejs.dev/guide/why.html>

MUNI
FACULTY
OF INFORMATICS