Poster

Machine guess of scattering factors - SALTED-NoSpherA2

F. Kleemiss[1], L. Seiffert[1], D. Brüx[1]

[1]Institute of Inorganic Chemistry, RWTH Aachen University, Landoltweg 1a, 52074 Aachen, Germany

florian.kleemiss@ac.rwth-aachen.de

Calculations of scattering factors from wavefunctions have shown remarkable improvements in the quality of structural models, including accurate hydrogen atom positions from X-ray diffraction data, better precision of disorder models in more straightforward cases, and the treatment of almost any type of crystal structures, and improvement of electron diffraction data interpretation. [1–4] One of the most time-consuming steps in this approach is the calculation and analysis of the electronic structure of the structural model. However, unlike wavefunction fitting, the molecular orbital coefficients are not required during the non-spherical atom refinement; solely the electron density or the resulting electrostatic potential is needed.

It was shown that the electron density of many chemical entities is transferrable using multipolar models trained on tailor-made experimental or theoretical structures.[5–7] The present approach uses symmetry-adapted learned three-dimensional electron densities (SALTED).[8] These densities are generated without a need for identification of atom types or species, despite their element symbol. The model can be trained on arbitrary basis set expansions and produces results comparable to complete DFT calculations in a fraction of the required time for the calculation. A direct consequence of this approach is the prediction of atom-centered electron densities without the need for assumptions like the Hirshfeld Stockholder partitioning scheme, which are directly usable in Fourier transform, as is possible for Multipole Models.

A question to be addressed is the choice of suitable training structures. Therefore, a set of structures from geometry-optimized ab initio molecular dynamics simulations is compared to the results obtained by a generalized set of training conformations, using a metaheuristic approach to sample the configuration space of training molecules to the highest possible extent. The results obtained using these two different training sets on a new set of atomic positions, not present in either training set, are compared, as exemplified in Figure 1. The obtained densities of both models can be used to refine high-quality data suitable for method benchmarking. The resulting statistics are compared to HAR results from a classical DFT calculation during non-spherical refinements.

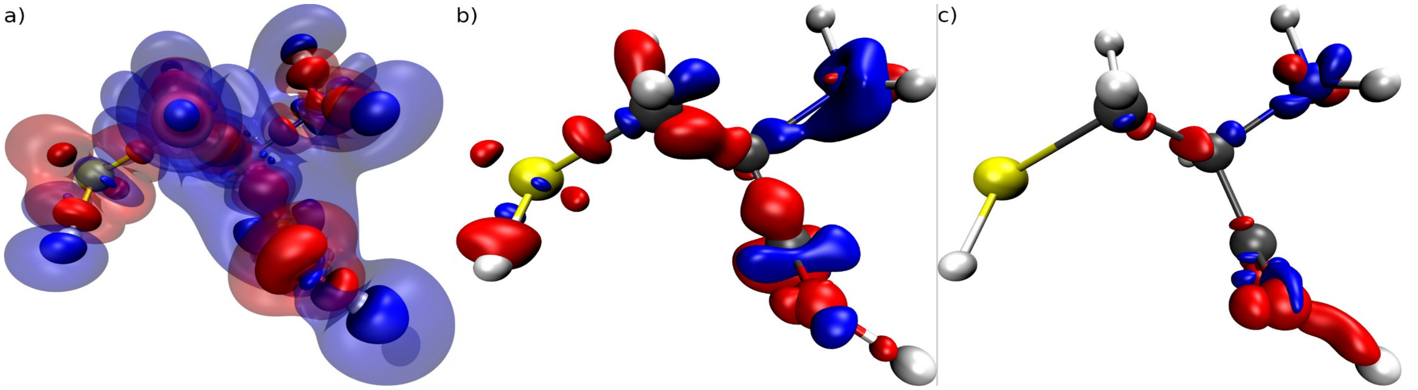

Figure 1. Static deformation densities from a) DFT calculation using cc-pVTZ basis set (isovalues: transp. ±0.005 a.u., opaque ±0.025 a.u.); b) Difference from a) to SALTED-model trained on DFT-optimized geometries (isovalue ±0.005 a.u.); c) using SALTED-model trained on generalized geometries from metaheuristics (isovalue: ±0.005 a.u.).

[1] S. C. Capelli, H.-B. Bürgi, B. Dittrich, S. Grabowsky, D. Jayatilaka, *IUCrJ* **2014**, *1*, 361–379.

[2] F. Kleemiss, O. V. Dolomanov, M. Bodensteiner, N. Peyerimhoff, L. Midgley, L. J. Bourhis, A. Genoni, L. A. Malaspina, D. Jayatilaka, J. L. Spencer, F. White, B. Grundkötter-Stock, S. Steinhauer, D. Lentz, H. Puschmann, S. Grabowsky, *Chem. Sci.* **2021**, *12*, 1675–1692.

[3] P. N. Ruth, R. Herbst-Irmer, D. Stalke, *IUCrJ* **2022**, *9*, 286–297.

[4] B. Gruza, M. L. Chodkiewicz, J. Krzeszczakowska, P. M. Dominiak, *Acta Cryst.* **2020**, *A76*, 92–109.

[5] P. M. Rybicka, M. Kulik, M. L. Chodkiewicz, P. M. Dominiak, *J. Chem. Inf. Model.* **2022**, *62*, 3766–3783.

[6] S. Domagała, B. Fournier, D. Liebschner, B. Guillot, C. Jelsch, *Acta Cryst.* **2012**, *A68*, 337–351.

[7] B. Dittrich, C. B. Hübschle, K. Pröpper, F. Dietrich, T. Stolper, J. J. Holstein, *Acta Cryst.* **2013**, *B69*, 91–104.

[8] A. Grisafi, A. M. Lewis, M. Rossi, M. Ceriotti, *J. Chem. Theory Comput.* **2023**, *19*, 4451–4460.