Deep learning approach for solving crystal structures from low-resolution data

T. Rekis[1,2], D. M. Melgalvis [1]

1 Faculty of Medicine and Life Sciences, University of Latvia, Jelgavas iela 1, Riga, LV1004, Latvia, 2 Institute for Inorganic and Analytical Chemistry, Goethe-University Frankfurt, Max-von-Laue Straße 7, Frankfurt am Main, 60438, Germany

rekis@chemie.uni-frankfurt.de

Crystal structure determination from single-crystal diffraction data is a well-established procedure. Nevertheless, there are experimental scenarios where the obtained diffraction data are not sufficiently good. The crystallographic phase problem does not have a general solution for incomplete and low-resolution data. This is often the case in high-pressure and electron diffraction where completeness and resolution are limited by the experimental setups and the nature of sample and radiation interaction. Furthermore, samples of some compounds inherently have low diffracting powers resulting in poor diffraction data. In a wider context, lack of general-purpose methods for phasing low-resolution data hinders structure solution from powder X-ray diffraction data. Even though high-resolution data can be easily measured, there is a significant peak overlap in higher 2θ ranges leading to ambiguous intensity assignment of these high-angle reflections. Consequently, powder patterns are rarely phased directly to get the starting crystal structure model.

Recently, we showed that deep learning can be successful to solve the phase problem [1]. The neural network PhAI was designed to solve structures in space group P21/c from data with down to 2 Å resolution, but it is limited for small unit-cell structures only. Here, we present further work. We have used the power of graphs to develop GraPhAI neural networks for solving centrosymmetric structures from low-resolution data [2]. The new models are trained on carefully designed training sets of hundreds of millions of artificial crystal structures [3].

Currently, our GraPhAI models can solve centrosymmetric crystal structures without strict restrictions on the space group symmetry, as well as for considerably larger unit cells than the original PhAI neural network. The results are summarized in Figure 1. Testing 10 000 randomly selected structures from the structure databases it was revealed that majority of them can be solved from diffraction data with 1 Å and 2 Å resolution. We also demonstrated that powder X-ray diffraction patterns can in principle be directly phased with our methods, but we also note that indexing and proper intensity extraction is still challenging.

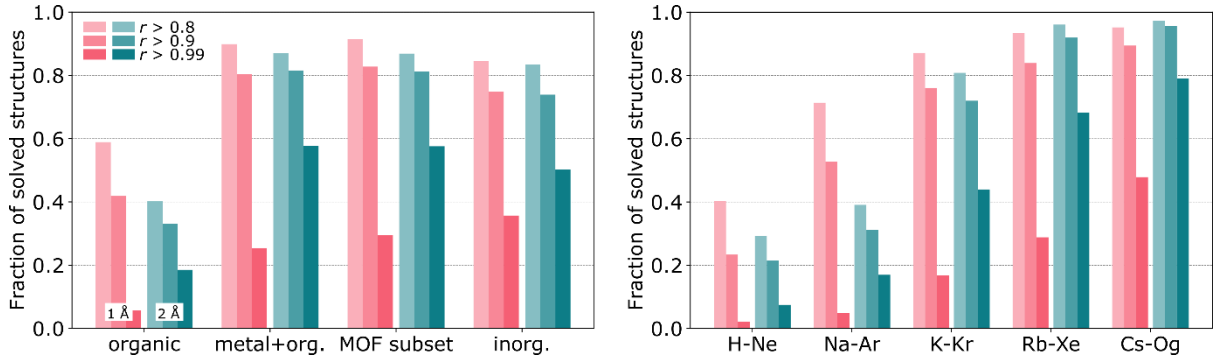


Figure 1. *Results of phasing 10 000 centrosymmetric crystal structures with unit cell dimensions up to 20 Å. Left: structures partitioned by compound class. Right: structures partitioned by heaviest element present. The red bars correspond to 1 Å resolution data, the green bars correspond to 2 Å resolution data. Darker shade corresponds to a stricter criterion for assuming the structures to be solved.*

[1] Larsen, A. S.; Rekis, T.; Madsen, A. Ø., Science (2024), 385, 522–528.
[2] Melgalvis, D. M.; Rekis, T., ChemRxiv (2026), https://doi.org/10.26434/chemrxiv.15000845/v1.
[3] Melgalvis, D. M.; Rekis, T., Acta Cryst. A (2026), 82, 32–40.