

## Application of neural networks for small angle scattering data analysis

A Kikhney<sup>1</sup>, D Molodenskiy<sup>2</sup>, D Svergun<sup>3</sup>

<sup>1</sup>*EMBL Hamburg, 22607*, <sup>2</sup>*EMBL Hamburg, Hamburg, NON-US*, <sup>3</sup>*EMBL Hamburg, Hamburg, Germany*

*a.kikhney@embl-hamburg.de*

Small angle X-ray scattering (SAXS) is a powerful technique for analyzing dilute solutions of proteins, nucleic acids and other macromolecules in solution under a wide range of conditions. The primary SAXS data analysis steps include estimation of the overall particle parameters (the radius of gyration  $R_g$ , maximum intraparticle distance  $D_{max}$ , molecular weight  $MW$ ) and computation of the pair distance distribution function  $p(r)$ . The latter represents the histogram of distances between pairs of points in the particle, weighted by the product of their scattering contrasts. Mathematically, the  $p(r)$  function is a Fourier transform of the scattering intensities versus the momentum transfer (i.e. scattering angle). The limited angular range of recorded experimental data, as well as the presence of experimental noise, makes the evaluation of  $p(r)$  an ill-posed problem. This problem is typically solved by an indirect Fourier transformation (IFT) approach, e.g. Glatter (1977), Svergun (1992), Hansen (2000). We propose a new approach for computing  $p(r)$  from the experimental data based on the state-of-the-art machine learning principles. We employed neural networks trained on a manifold of simulated and experimental SAXS data to perform primary data analysis. For a given experimental data set from protein, RNA or DNA our stack of networks evaluates  $R_g$ ,  $D_{max}$ ,  $MW$ ,  $p(r)$  and produces a denoised scattering curve. This approach has proved to be robust against experimental errors, applicable to data from particles of various nature, size and shape and does not require human input. The implementation of this approach as a publicly available web service with a graphical interface was developed, providing the possibility to inspect and download the results. This work is supported by the SAS-BSOFT grant: "Curation and development of databases for X-ray and neutron small angle control in biology and soft matter".