Decoding hidden structural information in solution wide-angle X-ray scattering (WAXS): from ensemble modeling to machine learning

Y. L. Chen[1], L. Pollack[2]
[1]Cornell University [2]Cornell University
yc2253@cornell.edu

Solution X-ray scattering is a powerful method for determining in vitro macromolecular structures. Small-angle X-ray scattering (SAXS) has been used to investigate global shape and size of a myriad of molecules at 10s-100

Å resolution. Its extension, wide-angle X-ray scattering (WAXS), sharpens the resolution of solution scattering to below 10 Å, enhancing information about macromolecular structures. These WAXS profiles, although informative, can be challenging to interpret, both as a result of contributions of the solvent, and the unknown structures related to the finer features of the solute. We applied WAXS to study the salt dependent structures of a short (12-base pair long) RNA duplexes. With a structural pool containing a variety of conformations, we leveraged ensemble modeling to understand the RNA's solution conformations [1]. Furthermore, we applied trained supervised machine learning (ML) models, mapping WAXS profiles to structural features, to interpret the noisy, yet information rich experimental WAXS profiles [2]. These well-trained ML models faithfully captured the WAXS fingerprints of RNAs. Importantly, these models identified specific scattering angles, or regions of scattering angles, which reflect (and can be used to predict) distinct structural parameters (WAXS feature importance). Therefore, we demonstrated that the ensemble and ML methods can provide novel frameworks for extracting highly relevant structural information from solution experiments on biological macromolecules.

1. Chen, Y.L. and Pollack, L. Salt dependence of A-form RNA duplexes: structures and implications, J. Phys. Chem. B 2019, 123, 46, 9773–9785
2. Chen, Y.L. and Pollack, L. Machine learning deciphers structural features of RNA duplexes measured with solution X-ray scattering, IUCrJ 2020, 7, 5,