

# **TeXRank: Texture Image Analysis and Machine Learning for Crystallizing Difficult Proteins**

**Carien Dekker**

*CPC/ Structural Biophysics, Novartis Institutes for Biomedical Research, Virchow-16-4.249.16, CH-4002 Basel, Switzerland*

While automation of both crystallization and drop imaging is now widespread, the flood of data remains largely unexploited. Algorithmic approaches have historically focused on per-drop scoring or else retrospective data mining; it remains unclear what impact the resulting computational tools have had on daily practice. We revisited drop analysis using modern texture methods, in particular the texton technique, with the ultimate goal of quantifying a protein's precipitation behaviour across the entire set of drops in a coarse screen, to serve as a fingerprint of its crystallizability that can be compared to the vast dataset of historical experiments at the SGC. The freely available program TeXRank comprises a suite of computational tools for day-to-day analysis of crystallization experiments, including: ranking of drops by likelihood of crystallinity; clear-drop analysis for discovery of better buffer solutions for the protein; and multi-drop analysis to provide a read-out of a protein's usefulness even in the absence of crystals along with pre-calculated optimisation screens. These tools will be described and results discussed, along with questions of uptake amongst users pre-conditioned to expect very little from a given experiment, and further possibilities opened by these approaches.