

Preface

Protein crystallization usually requires many experiments that check combinations of various factors such as pH, ionic strength, etc. for a successful crystalline outcome. Nevertheless, as crystalline outcome especially for proteins to difficult crystallize such as membrane proteins in the presence of lipids and detergents is rare, many trials have been set up. These protein crystallization trials are usually analyzed by an expert using a microscope. Going over thousands of unsuccessful trials for a few successful (but important) outcomes has been tedious. In recent years, automated robotic high-throughput systems are proposed to conduct many experiments and fast detection of crystalline conditions. Initially, these high-throughput systems were costly and accessible only by major research laboratories. The significant cost of these systems made these research systems available only big research laboratories. Recent advancements in computational aspects and analysis of protein crystallization and decreasing cost of hardware architectures make automated systems available to also small research laboratories. Moreover, new protein crystallization techniques such as trace fluorescence labeling do not only reduce the time for preparation and analysis of crystallization experiments but also help to develop fast and accurate computational methods for protein crystallization analysis. This book covers how to build low-cost but fairly accurate protein crystallization analysis system thus enabling small research groups to build their own robotic high-throughput systems and crystallization analysis systems.

This book covers various aspects of computational aspects of protein crystallization. Previously, the computational aspects were usually covered as supplementary information in major crystallization books or sometimes they were ignored. This book unites important aspects of data analytics for protein crystallization into a single book. The methods and programs were developed as a part of collaborative research by iXpressGenes, Inc. and the University of Alabama in Huntsville funded through NIH-STTR grants. These projects funded two Ph.D. students, six M.S. students, and two part-time students along with other ongoing contributors. We have developed a number of systems for analyzing protein crystallization process while comparing our work with the state-of-art techniques. This

book also covers relevant research that will help readers understand different dimensions of protein crystallization analysis.

This book is relevant to researchers who would like to know about computational aspects and data analytics components of protein crystallization. While the book is relevant to the community of structural biology, it also serves computer scientists who would like to get into the protein crystallization field.

This data analytics book on protein crystallization analysis covers the complete cycle of data analysis for protein crystallization. It starts from background information on protein crystallization, setting up screens by analyzing prior crystallization trials, building robotic setups, classifying crystallization trial images by effective feature extraction, analyzing crystal growth in time series images, segmenting crystal regions in images, providing focal stacking methods for crystallization images captures at varying fields of depth, and visualization of trials. The book is organized as follows:

“Chapter 1: Introduction to Protein Crystallization” gives information about how protein crystallization experiments are conducted in a wet lab. Besides traditional experiments, we also cover trace fluorescence labeling that helps data analytics.

“Chapter 2: Scoring and Phases of Crystallization” covers scoring and categorization of crystallization image trials. Researchers came up with their own way of categorization in the literature. This chapter presents a variety of ways for categorization.

“Chapter 3: Computational Methods for Protein Crystallization Screening” presents computational methods for determining cocktails to be tested based on the results of prior experiments and their scoring. While commercial screens enable setting up plates with many successful cocktails, the analysis of unsuccessful trials has been left to the experts. This chapter provides approaches for setting up plates for successful crystalline outcomes.

“Chapter 4: Robotic Image Acquisition” presents the hardware and software architectures for a basic high-throughput system.

“Chapter 5: Classification of Crystallization Trial Images” presents overview of features used in protein crystallization image classification. As feature extraction has been the bottleneck of high-throughput systems, this chapter categorizes features and analyzes their running-time for real-time systems.

“Chapter 6: Crystal Growth Analysis” presents spatiotemporal analysis of protein crystal growth. This chapter analyzes the formation of new crystals as well as the growth of crystals in size.

“Chapter 7: Focal Stacking for Crystallization Microscopy” presents how to generate in focus crystallization images from a set of images that are captured at varying depths of field of a microscope. Since crystals usually float in a 3D well, some crystals may be out of focus and focal stacking may be necessary for proper analysis.

“Chapter 8: Crystal Image Region Segmentation” presents how to extract regions of crystals as thresholding or binarization has been one of the challenging issues in image segmentation.

“Chapter 9: Visualization of Crystallization Trial Experiments” introduces how plates can be visualized before/after scoring, temporal visualization of wells under different lighting conditions, and enabling/updating scoring by experts.

“Chapter 10: Other Structure Determination Methods” provides alternate methods to obtain a 3D structure (neutron diffraction, cryogenic electron microscopy, nuclear magnetic resonance, and X-ray free electron laser diffraction) and methods suitable for more general structural information (chemical cross-linking, fluorescence resonance energy transfer, and circular dichroism).

“Chapter 11: Future of Computational Protein Crystallization” provides overview of methods in progress and future trends for protein crystallization.

Huntsville, AL, USA
August 2017

Marc L. Pusey
Ramazan Savaş Aygün

Data Analytics for Protein Crystallization

Pusey, M.L.; Aygün, R.S.

2017, XX, 231 p. 66 illus., 56 illus. in color., Hardcover

ISBN: 978-3-319-58936-7