m02.p03

Increasing Crystallization Trials Productivity through Imaging Automation

Conn Mallett, Pierre Le Magueres[1], Eric Hnath[1], Jian Xu[2]

[1] Rigaku Americas, 9009 New Trails Dr., The Woodlands, TX 77381. [2] Rigaku Automation Group, 5999 Avenida Encinas, Suite 150, Carlsbad, CA 92008

Keywords: crystal imaging, crystallization productivity, phase separation

The Minstrel I crystal imaging platform from Rigaku, consisting of a Minstrel I imager, a plate hotel and the database CrystalTrak, is designed to automate the process of recording, tracking and optimizing a vast number of protein crystallization trials.
The Minstrel I captures high quality images of crystallization drops from most commercial crystallization plates due to a flexible and upgradeable plate type library. The possibility to manually select different light patterns and polarized light filters further aids the identification of such critical features as crystalline precipitate, phase separation and crystals down to a size of a few microns.
Linked to a plate hotel with up to 160 SBS-type plates, and combined with the CrystalTrak database, the Minstrel I imaging system allows crystallographers to frequently image crystallization trials and electronically record relevant of data. Thanks to the option in CrystalTrak to automatically optimize crystallization conditions, the platform (Minstrel I + plate hotel + CrystalTrak) represents a solution for increasing crystallization trial productivity, and thus reduce the amount of time in obtaining crystals suitable for X-ray diffraction.

m02.p04

Machine Learning Strategies for Protein Crystallization

Mónica Pérez-Priede[a], Santiago García-Granda[a]

[a]Department of Physical and Analytical Chemistry, University of Oviedo, Spain. E-mail: mpp@fq.uniovi.es.

Keywords: macromolecular crystallography, macromolecular crystallization, bioinformatics

Despite general pessimism on applying techniques based on Artificial Intelligence, algorithms are able to learn from general crystallization data, improving protein crystallization success. Using data available in the BMCD Biological Macromolecule Crystallization Database [1] and in the PDB Protein Data Bank [2], two new crystallization local databases, which are the core of the process of knowledge-based predictions, were implemented. Machine Learning techniques were applied, not only to predict crystallization conditions for new problems, but also to optimise them in a particular case. Some of the results obtained will be shown. The figure shows the best mean absolute error when predicting crystallization temperature range (Tmin, Tmax) for enzymes and metalloproteins from BMCD. The table presents the percent of success when predicting precipitant and a crystallization method for metalloproteins from BMCD using four different learning algorithms which will be detailed.

	(1)	*(2)*	*(3)*	*(4)*
Precipitant	*27.14±2.78*	*39.71±12.26*	*35.92±9.78*	*39.07±12.4*
Method	*58.85±1.43*	*63.29±7.14*	*59.95±5.05*	*62.74±6.97*

Table: Percent of success on predicting precipitant and a crystallization method within those most used on crystallizing metalloproteins from BMCD. Columns match the results together with standard deviations, obtained using four different learning algorithms

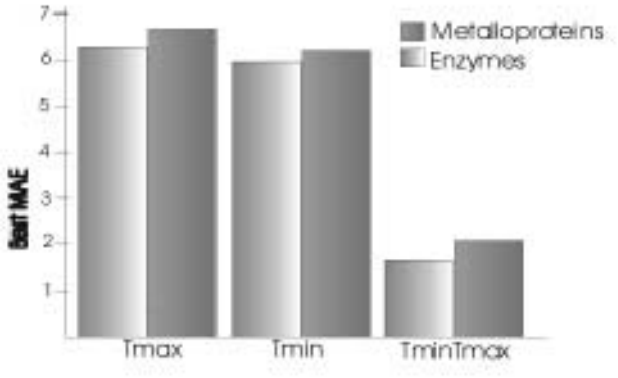

Fig. Mean absolute error on predicting temperature range for crystallizing any metalloprotein or enzyme from BMCD. In bars five and six TminTmax label means *Tmin* was used to learn and predict *Tmax*.

[1] Gilliland, G. L. & Tung, M. & Blakeslee, D.M. & Ladner, J. (1994). *Acta Crystallogr.* D50, 408-413.
[2] Berman, H.M.& Westbrook, J. & Feng, Z. & Gilliland, G. & Bhat, T.N. & Weissig, H. & Shindyalov, I.N. & Bourne, P.E. (2000). *Nucleic Acids Research* 28, 235-242.