

ContaMiner and ContaBase: Automated identification of unwantedly crystallized protein contaminants

Stefan T Arold¹, Arnaud Hungler¹, Afaque A.I. Momin¹, Kay Diederichs²

¹Computational Bioscience Research Center (CBRC), KAUST, Thuwal, Saudi Arabia, ²Fachbereich Biologie, University Konstanz, Konstanz, Germany

E-mail: stefan.arold@kaust.edu.sa

Solving the phase problem in protein X-ray crystallography relies heavily on the identity of the crystallized protein, especially when molecular replacement (MR) methods are used. Yet, it is not uncommon that a contaminant crystallizes instead of the protein of interest. Such contaminants may be proteins from the expression host organism, protein fusion tags or proteins added during the purification steps. Many contaminants co-purify easily, crystallize and give good diffraction data. Identification of contaminant crystals may take time, since the presence of the contaminant is unexpected, and its identity unknown. We have established a webserver (ContaMiner) and a contaminant database (ContaBase), both available at strube.cbrc.kaust.edu.sa/contaminer/, to allow fast MR-based screening of crystallographic data against currently 71 known contaminants from seven different expression systems [1]. The program is also available in a github repository and can be installed locally. ContaMiner enables screening of novel crystals at synchrotron beamlines, and it would be valuable as a routine safety check for 'crystallization and preliminary X-ray analysis' publications. Thus, in addition to potentially saving X-ray crystallographers much time and effort, ContaMiner might considerably lower the risk of publishing erroneous data. We will present new features for enhancing speed, customization and visualization of results.

Hungler, A., Momin, A., Diederichs, K. & Arold, S.T. (2016). *J. Appl. Cryst.* (2016) 49, 2252-2258

Keywords: [molecular replacement](#), [contaminant crystals](#), [web server](#)