COMPUTATIONAL PREDICTION OF EXPRESSION AND SOLUBILITY OF RECOMBINANT BIOPHARMACEUICALS

Alan Dickson, Manchester Institute of Biotechnology, UK
alan.dickson@manchester.ac.uk
Swee Jiun Fu, Faculty of Life Sciences, University of Manchester, UK
Manuel Carballo Amador, Manchester Institute of Biotechnology, UK
Jim Warwicker, Manchester Institute of Biotechnology, UK

Protein based therapeutics have emerged as a successful class of pharmaceutical. However, it is well known that much of the current therapeutic protein discovery and development processes is based around existing molecular frameworks and that novel formats offer significant challenges for expression. Efficient production of protein is required to meet the growing demands and increasing expectation of the patients and healthcare providers. Major obstacles during biopharmaceutical production are linked to the efficiency of the protein expression system and the biophysical properties of proteinbased products that can lead to aggregation and subsequent problems for purification, quality and effectiveness. Computational tools have been developed to aid prediction of protein solubility and aggregation propensity to support enhanced certainty of optimal generation of product with desirable properties. In this study, we have used an in-house computational tool for prediction of soluble protein expression (developed around protein structure and surface electrostatic properties of human erythropoietin, HuEPO) was developed in E. coli and has been validated experimentally with several bacteriallyexpressed model protein variants. The application of the computational approach has been extended to mammalian expression platforms. A significant inverse correlation was observed between positive surface patches and the expressability of HuEPO in transient mammalian cells (HEK and CHO cell lines). Mechanistically the differential expression operates at a level post-transcriptionally, associated with ribosome-secretory complex engagement, protein stability or secretory processes. The results demonstrate the potential of application of a predictive computational algorithm as a design tool in rational protein engineering to improve expression of novel protein formats in mammalian systems as well as E.coli. In summary, optimization of molecular patches on the surface of proteins may be a viable strategy to enhance protein soluble expression and therefore a potential solution for development of novel proteins that might otherwise fit into the category of "difficult-toexpress" proteins.