MULTIDIMENSIONAL ENGINEERING OF CHYMOSIN FOR EFFICIENT CHEESE PRODUCTION BY MACHINE-LEARNING GUIDED DIRECTED EVOLUTION

Christian Jäckel, Chr. Hansen A/S, Hørsholm, Denmark
DKCHJL@chr-hansen.com
Martin Lund, Chr. Hansen A/S, Hørsholm, Denmark
Hans van den Brink, Chr. Hansen A/S, Hørsholm, Denmark
Claes Gustafsson, ATUM, Newark, CA, USA
Sridhar Govindarajan, ATUM, Newark, CA, USA

Key Words: ProteinGPS, Multidimensional Engineering, Substrate Specificity, Design of Experiment

The global cheese market today exceeds $100B/year. Chymosin (a.k.a. rennin) is an aspartic endopeptidase produced by the stomach lining of new-born mammals. During cheese production chymosin is added to the milk where it cleaves the glycomacropeptide (GMP) from the surface of casein micelles to initiate milk coagulation. Current commercial recombinant chymosin enzymes derived from *Bos taurus* (cow) or *Camelus dromedarius* (camel) are limited in their proteolytic specificity leading to incomplete milk-to-cheese conversion. Increasing the chymosin specificity for GMP cleavage would significantly decrease the amount of milk needed for cheese production thereby reducing cost and decreasing environmental footprint of the dairy industry. Separate from milk coagulation, chymosin dependent release of N-terminal peptides from alphaS1 casein during cheese ripening leads to unwanted softening, accompanied with cheese loss during industrial processing such as slicing and shredding. Furthermore, chymosin dependent cleavage of the C-terminal end of beta casein contributes to unwanted bitterness of the cheese. Improvement of chymosin proteolytic specificity in both milk coagulation and cheese ripening is consequently of high commercial relevance.

We here present how we applied the ProteinGPS protein engineering platform to improve the specificity of camel chymosin for GMP cleavage, increase the milk clotting efficiency, while simultaneously reduce the off-target casein cleavage in cheese resulting in unwanted softening and bitter taste. A total of 108 amino acid substitutions found in naturally existing chymosin homologs or selected from structure-based design approaches were systematically introduced into the enzyme backbone using Design of Experiment principles for a total of ~300 chymosin variants distributed over four iterations (Fig. 1). The relative contribution and epistatic effect of each of the 108 substitutions in each functional dimension was modeled independently using modern machine learning algorithms. As little as 300 different chymosin variants covered the statistical sampling of a total sequence space of ~10^{32} illustrating the information efficiency of the method. This low number of chymosin variants to be produced and analyzed was small enough to allow for near-product grade quality in application relevant assays, including an LC-MS/MS-based mapping of casein degradation products in micro cheeses. The best performing engineered chymosin is 9 amino acid substitutions away from the parent camel chymosin, has increased GMP cleavage specificity by 30-fold, doubled milk clotting efficiency, and reduced unwanted proteolysis in cheese by 60-80%. This enzyme engineering work resulted in a commercial product to be released by Chr. Hansen in 2019.

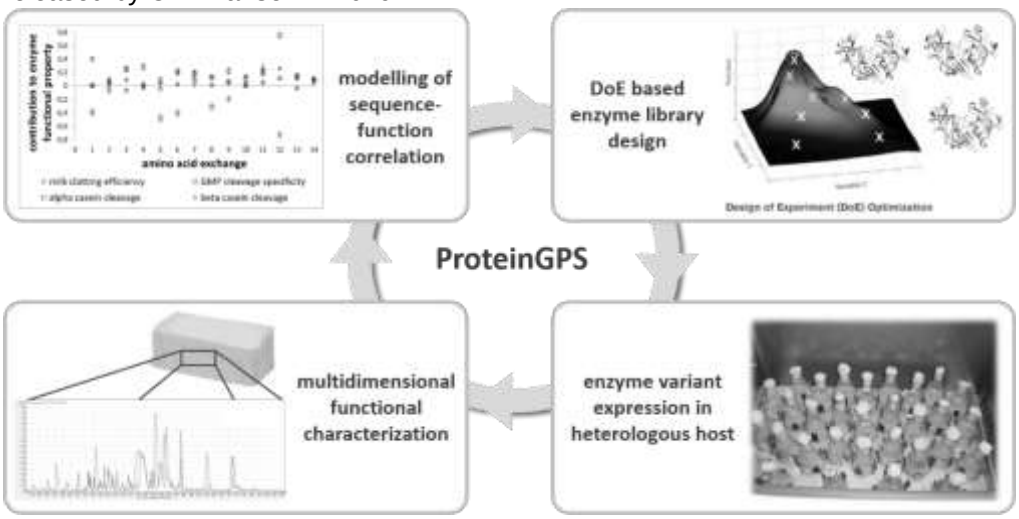

Figure 1: Multidimensional engineering of chymosin for cheese production by repetitive cycles of machine-learning guided directed evolution.