27O-ISMS25 **Machine Learning Techniques for Peptide Optimization from Sequence Information**

Mark R. HANSEN[1], Evgenia RAZUMOVSKAYA[1], Jason HODGES[1], Hugo O VILLAR[1], ○Fumiaki ARUGA[2]
[1]Altoris, Inc., [2]Patcore, Inc.

Traditionally peptides have been optimized in a series of sequential steps where activity is maximized, followed by selectivity and subsequently other pharmacokinetic and toxicological properties are adjusted. Gains in efficiency could be realized if multiple properties are optimized simultaneously via the use of machine learning techniques. Unfortunately, the most appropriate methods for each property can be different and require some expertise in their use. We have created a computer system that automatically examines the utility of different machine learning techniques for a dataset and selects the most predictive methods for different properties of clinical interest for peptides. The combination of these best in class algorithms provides an avenue to solve the multifactorial problem of peptide optimization. A wide range of machine learning techniques are needed to make accurate prediction for the full range of preclinical properties of interest. To that end, we analyzed large datasets that revealed that while properties such as half-life or affinity towards a target can be modeled using Bayesian regression techniques, in other cases techniques such as support vector machines are needed to predict amyloid aggregation, when using positional physicochemical descriptors. The models can be used to predict peptides most likely to have the desired activity from within the combinatorial expansion of amino acids, natural, unnatural including modified peptides. Our ultimate goal is to develop a decision support system that guides the optimization of peptides towards the definition of a clinical candidate minimizing the number of peptides that need to be evaluated and useful to non-experts.