1067-92-471 **Ye (Alice) Tian*** (ytian@math.wsu.edu), Department of Mathematics, Washington State University, P.O.Box 643113, Pullman, WA 99164, and **Christopher Deutsch** and **Bala Krishnamoorthy**. *Optimized Scoring Function to Predict Solubility Mutagenesis.*

Background: Mutagenesis is commonly used to engineer proteins with desirable properties not present in the wild type (WT) protein, such as increased or decreased stability, reactivity, or solubility. Experimentalists often have to choose a small subset of mutations from a large number of candidates to obtain the desired change, and computational techniques are invaluable to make the choices. Results: We use concepts from computational geometry to define a three body scoring function that predicts the change in protein solubility due to mutations. The scoring function captures both sequence and structure information. By exploring the literature, we have assemble a substantial database of 137 single- and multiple-point solubility mutations. Our database is the largest such collection with structural information known so far. We optimize the scoring function using linear programming (LP) methods to derive weights based on training. We compare the LP method to the standard machine learning techniques of support vector machines (SVM) and the Lasso. Using statistics for leave-one-out (LOO), 10-fold, and 3-fold cross validations (CV) for training and prediction, we demonstrate that the LP method performs the best overall. For the LOOCV, the LP method has an overall accuracy of 81%. (Received September 05, 2010)