

On conformational changes of proteins using collective motions in torsion angle space and L1 regularization

Ryabinina R. B.¹, Emtsev D. I.²

¹ ryabinina.rb@phystech.edu ² daniil.emcev.ru@yandex.ru

¹²MIPT

Abstract

Investigation of conformations in proteins is an important and well-studied problem in bioinformatics with applications ranged from drug design to understanding hidden effects in the data. This is an important and open question in computational structural bioinformatics - how to efficiently represent transitions between protein structures. Here we address the problem of how sparse subset of collective coordinates in the torsion subspace can describe functional conformational changes in proteins. The following strategy consists in determining the change of torsion angles through the fit of the linearized change of Cartesian coordinates. However, if the fit is not regularized the structures produced in this way demonstrate the deviation of several Angstrom from targets. Rescaled ridge regression (RRR) has been recently introduced to regularize multi-dimensional regressions with correlated explanatory variables. The resulting torsional conformational changes generate conformations that are much more similar to the target conformations, and they are better correlated with the thermal fluctuations of torsion angles and with the normal modes predicted by the TNM than the torsional conformational changes obtained through ordinary regression. Our goal is to find a solution of a ridge regression problem with an L1 regularization constraint using the LASSO formulation. Not much has been done in the torsional angle subspace (internal coordinates) and nearly nothing has been done using L1 regularization.bhbjh