Predicting cryptic ligand binding sites based on normal modes guided conformational sampling

Wenjun Zheng¹

¹University at Buffalo

June 29, 2020

Abstract

To greatly expand the druggable genome, fast and accurate predictions of cryptic sites for small molecules binding in target proteins are in high demand. In this study, we have developed a fast and simple conformational sampling scheme guided by normal modes solved from the coarse-grained elastic models followed by atomistic backbone refinement and sidechain repacking. Despite the observations of complex and diverse conformational changes associated with ligand binding, we found that simply sampling along each of the lowest 30 modes is near optimal for adequately restructuring cryptic sites so they can be detected by existing pocket finding programs like fpocket and concavity. We further trained machine-learning protocols to optimize the combination of the sampling-enhanced pocket scores with other dynamic and conservation scores, which only slightly improved the performance. As assessed based on a training set of 84 known cryptic sites and a test set of 14 proteins, our method achieved high accuracy of prediction (with area under the receiver operating characteristic curve > 0.8) comparable to the CryptoSite server. Compared with CryptoSite and other methods based on extensive molecular dynamics simulation, our method is much faster (1-2 hours for an average-size protein) and simpler (using only pocket scores), so it is suitable for high-throughput processing of large datasets of protein structures at the genome scale.

Hosted file

Cryptic pocket prediction.pdf available at https://authorea.com/users/337624/articles/463257predicting-cryptic-ligand-binding-sites-based-on-normal-modes-guided-conformationalsampling