

**Figure S1.** Histograms and dependencies of several characteristics of the learning protein set (CC – correlation coefficient).

