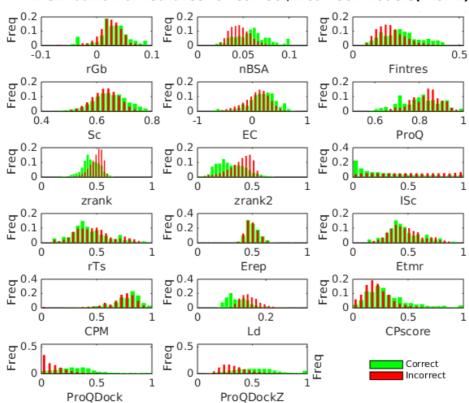
## Distribution of features for correct/incorrect models(MOAL)



**Fig. S1.** Distribution of training features on the MOAL set for correct (green) and incorrect (red) models. The features that are outside the -1,1 range are scaled using logistic scaling (see Methods) to put them in the range 0-1. For the methods that are energy terms, zrank, zrank2, Isc, rTs, Erep, and Etmr lower scores are better, for the others higher are better.

## Distribution of features for correct/incorrect models (CAPRI) 0.2 0.1 0.2 0.1 Fed 0.5 0 0 0 -0.02 0.05 0 0.02 0 0.1 0 0.5 rGb nBSA Fintres 0.2 0.1 0.2 0.1 0.4 0.2 0 0 0 0.5 1 0 0.6 0.8 0 -1 1 Sc EC ProQ 0.2 0.1 P.0 4 Freq 0.2 0.5 0 0 0 0 0.5 0 0.5 1 0 0.5 1 zrank zrank2 ISc Freq 0.5 Ped 0.1 1 Fed 0.5 0 0 0 0 0 0.5 1 0.5 1 0 0.5 1 rTs Erep Etmr P 0.2 P.0.4 0.2 P.0 4 0 0 0 0 0.5 0 0.2 0 0.5 1 CPM Ld CPscore P 0.2 0.2 0.1 Freq Correct

Fig. S2. Distribution of training features on the CAPRI set for correct (green) and incorrect (red) models. The features that are outside the -1,1 range are scaled using logistic scaling (see Methods) to put them in the range 0-1. For the methods that are energy terms, zrank, zrank2, Isc, rTs, Erep, and Etmr lower scores are better, for the others higher are better.

0.5

ProQDockZ

Incorrect

0

0

1

0

0

0.5

ProQDock