Are Scoring Functions in Protein−Protein Docking Ready to Predict Interactomes? Clues from a Novel Binding Affinity Benchmark *J. Proteome Res.* **2010,** *9,* 2216–2225. Panagiotis L. Kastritis and Alexandre M. J. J. Bonvin*

We have recently reported a binding affinity benchmark consisting of binding constants (K_d's) for 81 complexes. This benchmark was used to assess the performance of nine commonly used scoring algorithms along with a free energy prediction algorithm in their ability to predicting binding affinities.

In a collaborative effort with the groups of Prof. Janin (Université Paris-Sud), Dr. P. Bates (Cancer Research UK, London) and Prof. Z. Weng (University of Massachusetts Medical School, Worcester) to extend this benchmark, we have discovered a number of discrepancies in some of the reported values and PDB entries in our previously published benchmark. Accordingly, only 46 of the 81 reported binding affinity data can be considered fully accurate at this time. The other entries have various levels of inaccuracies and should preferably be discarded in any work using those data for the development of new scoring or binding affinity functions for example. This subset includes complexes with PDB IDs 7CEI, 1DFJ, 1BVN, 1IQD, 1MAH, 1EZU, 1JPS, 1IBR, 1R0R, 1T6B, 1KXP, 2FD6, 2I25, 2B42, 2JEL, 1ML0, 1BJ1, 1KXQ, 1OPH, 1M10, 2AJF, 1IJK, 1H1V, 1E6J, 2HLE, 1A2K, 2C0L, 1RLB, 1GRN, 1E6E, 1J2J, 2BTF, 1HE8, 1B6C, 1I4D, 1GHQ, 2MTA, 1E96, 1Z0K, 1QA9, 1AK4, 1GCQ, 1WQ1, 2OOB, 1AKJ, and 1S1Q.

We have reanalyzed the accuracy of the various scoring functions on this subset and, although correlations are slightly improved ($0 < R^2 < 0.3$), current functions still do not hold any predictive capacity (Figure 1). The updated trusted benchmark is available from http://haddock.chem.uu.nl/services/affinity.

10.1021/pr101118t

Published on Web 11/17/2010

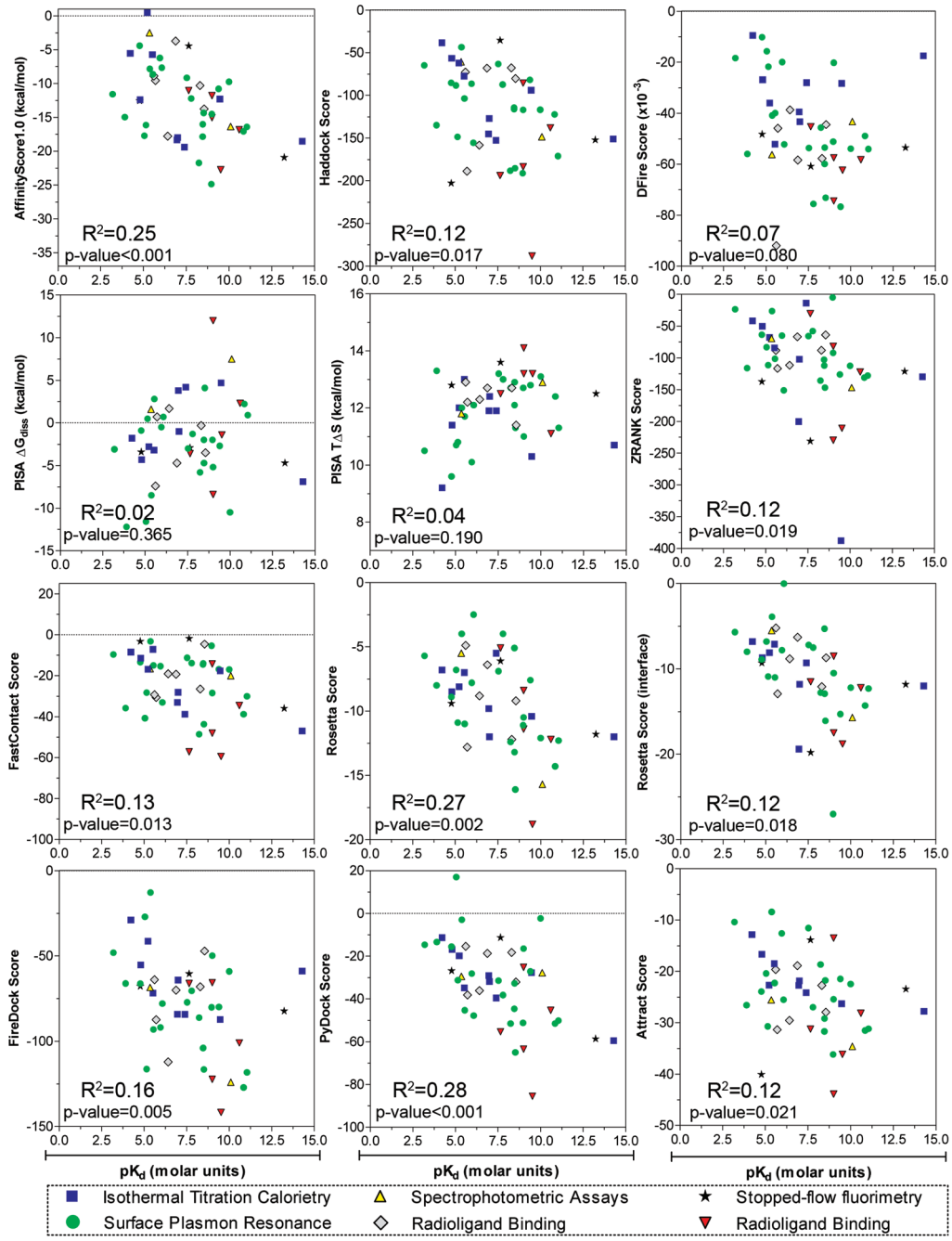

Figure 1. Scatter plots of binding affinity ($-\log(K_d)$) (*x*-axis) versus score (*y*-axis) for the 46 water refined complexes from our corrected binding affinity benchmark. The scores were calculated using nine different scoring functions and a binding affinity prediction algorithm (panels A−L, see text). r^2 and *p*-values are indicated in each plot. Different colors and shapes of the data points correspond to different methodologies followed to experimentally determine the binding affinity.

PR101118T