Supporting Information for

Improving docking results through re-ranking of diverse poses using MM-GBSA

P. A. Greenidge,[a,*] C. Kramer,[b] J.-C. Mozziconacci,[c] W. Sherman[d]

a) Novartis Institutes for Biomedical Research, Novartis Pharma AG, Forum 1, Novartis Campus, CH 4056 Basel, Switzerland

b) University of Innsbruck, Center for Chemistry and Biomedicine, Institute for General, Inorganic and Theoretical Chemistry, Innrain 82, 6020 Innsbruck, Austria

c) Schrödinger GmbH, Dynamostrasse 13, 68165 Mannheim, Germany

d) Schrödinger Inc., 120 West 45th Street, New York, NY 10036, USA

** Corresponding author: Email: Paulette.Greenidge@novartis.com*

The main scoring functions from GOLD produce lower correlations with the experimental data than do other separate components that contribute to the score (Table S1), notably vdW and PLP respectively from GoldScore and ChemPLP respectively. PLP models the steric complementarity between protein and ligand, while ChemPLP has additionally hydrogen bond and metal bonding terms from ChemScore. These extra terms improve the reproduction of binding mode accuracy.[1] GoldScore is a function that is based on molecular mechanics and is composed of four terms: protein-ligand hydrogen bonding, van der Waals interactions, ligand strain, and intra molecular hydrogen bonding. It is then not surprising that PLP and the vdW component of GoldScore give a very similar correlation with experimental data (R^2=0.57). GlideScore does not produce a high correlation with experimental binding affinity (R^2=0.36). The vdW component of the scoring function performs similarly to that from the GOLD scoring functions.

Table S1. Score in place results for the 855 complexes[2] and various components of the scoring functions with 95% confidence intervals indicated. Coefficients of determination (R^2) are for the negative log of experimental values (K_i or K_d) versus scoring function values. Simplex and no simplex refer to whether or not the ligand was locally minimized with respect to the protein.

R^2	ChemScore	GoldScore	ChemPLP	GlideScore SP
simplex	0.44 ± 0.05	0.38 ± 0.05	0.48 ± 0.05	0.36 ± 0.05
no simplex	0.41 ± 0.05	[a]n/a	0.49 ± 0.05	0.36 ± 0.05
[b]DG	0.47 ± 0.05	n/a	n/a	n/a
[c]vdW	n/a	0.57 ± 0.04	n/a	0.54 (0.49 – 0.58)
[d]PLP	n/a	n/a	0.57 ± 0.05	n/a

[a]Large steric clashes occurred without simplex minimization, therefore it was not appropriate to report a value for this option. n/a not applicable.[b] DG is a term in the Chemscore fitness function.[3] It was trained by regression against 82 complexes using binding data. [c]vdW is the van der Waals interaction energy. [d]PLP is a component of the ChemPLP scoring function that models the complementarity between the protein and ligand.[1]

1) Korb, O.; Stuetzle, T.; Exner, T. E. Empirical scoring functions for advanced protein-ligand docking with PLANTS. *J. Chem. Inf. Model*. **2009**, *49*, 84-96.
2) Greenidge, P. A.; Kramer, C.; Mozziconacci, J.-C.; Wolf, R. M. MM/GBSA Binding Energy Prediction on the PDBbind Data Set: Successes, Failures, and Directions for Further Improvement. *J. Chem. Inf. Model.* ***2013****, 53*, 201–209.
3) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical scoring functions. 1. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput.-Aided Mol. Des.***1997**, 11, 425–445.