# Supporting Information

# Improving Scoring-Docking-Screening Powers of Protein-Ligand Scoring Functions using Random Forest

*Cheng Wang*[1] *and Yingkai Zhang*[1,2,*]

[1]Department of Chemistry, New York University, New York, New York 10003

[2]NYU-ECNU Center for Computational Chemistry at NYU Shanghai, Shanghai 200062, China

*To whom correspondence should be addressed.

E-mail: yingkai.zhang@nyu.edu

## Table of Contents

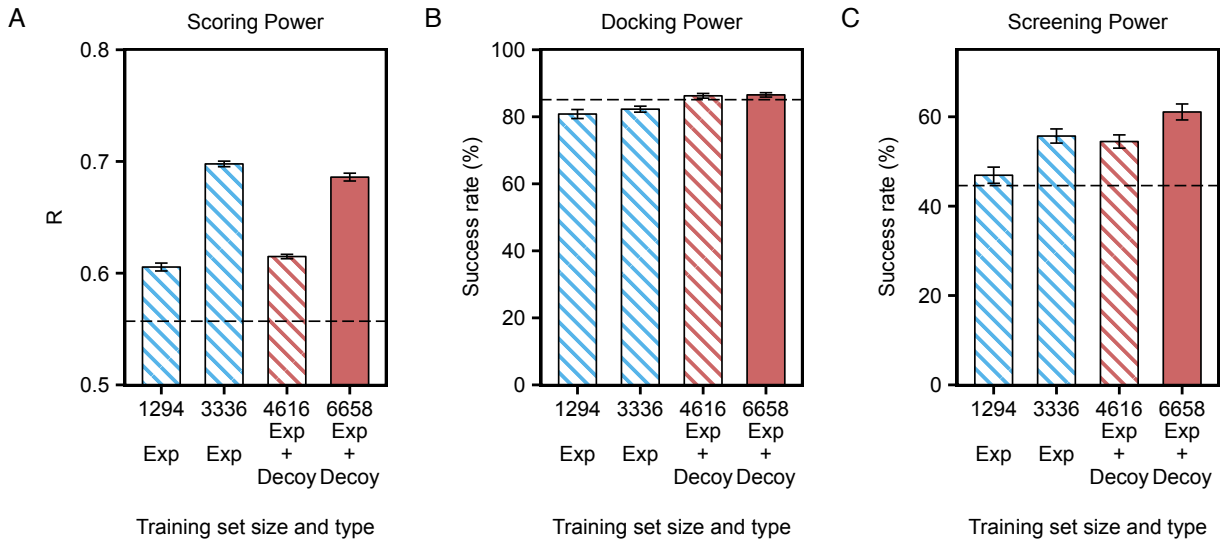


**Figure S1.** CASF-2013 benchmark test performance of scoring functions developed with the $\Delta_{vina}$RF approach using the same feature set (20 features as in $\Delta_{vina}RF_{20}$) but **different training sets**. (A) Scoring power. (B) Docking power; (C) Screening power. For each training set, parameterization is run 10 times with different random seed for random forest and calculated by averaging over 10 performances. Experimental only training set is colored in blue and mixed training set is colored in red. 1294 structures are from PDBbind Refined 2007 + Weak + Decoy Crystal – Test. 3336 structures are PDBbind Refined 2014 + Weak + Decoy Crystal – Test. 4616 structures are combined by 1294 crystal structures and 3322 decoys structures. 6658 structures are combined by 3336 crystal structures and 3322 decoys structures. The AutoDock Vina performance is indicated by dashed line.

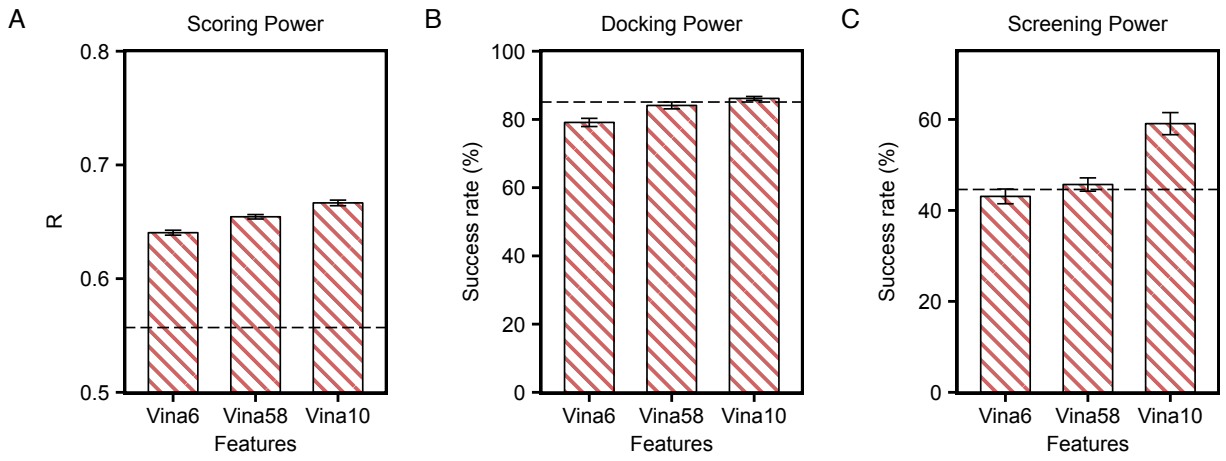


**Figure S2.** CASF-2013 benchmark test performance of scoring functions developed with the $\Delta_{vina}$RF approach using the combined experimental and decoy training set (6658) but with **different feature sets:** Vina6 refers the use of 6 features in the original Vina scoring function as the feature set, Vina58 means the use of all 58 features in the Vina source code as the feature set, and Vina 10 refers the use of top 10 features from the slection process as the featrure set. (A) Scoring power. (B) Docking power; (C) Screening power. For each feature set, parameterization has been run 10 times with different random seed for random forest and calculated by averaging over 10 performances. The AutoDock Vina performance is indicated by dashed line.

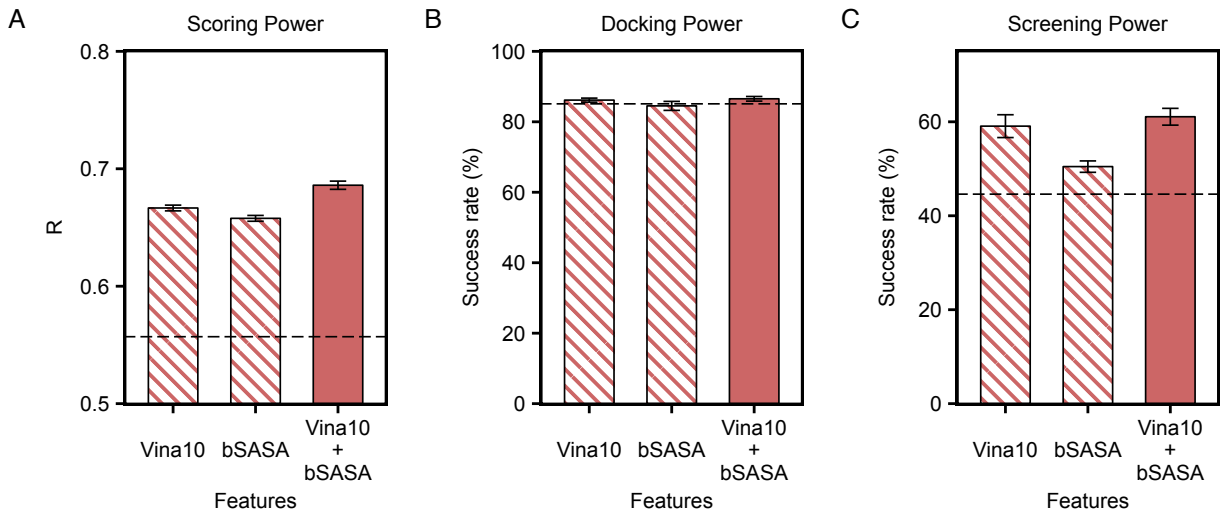


**Figure S3.** CASF-2013 benchmark test performance of scoring functions developed with the $\Delta_{vina}$RF approach using the combined experimental and decoy training set (6658) but with **different feature sets:** Vina10, bSASA, and Vina10 + bSASA features. (A) Scoring power. (B) Docking power; (C) Screening power. For each feature set, parameterization is run 10 times with different random seed for random forest and calculated by averaging over 10 performances. The AutoDock Vina performance is indicated by dashed line.

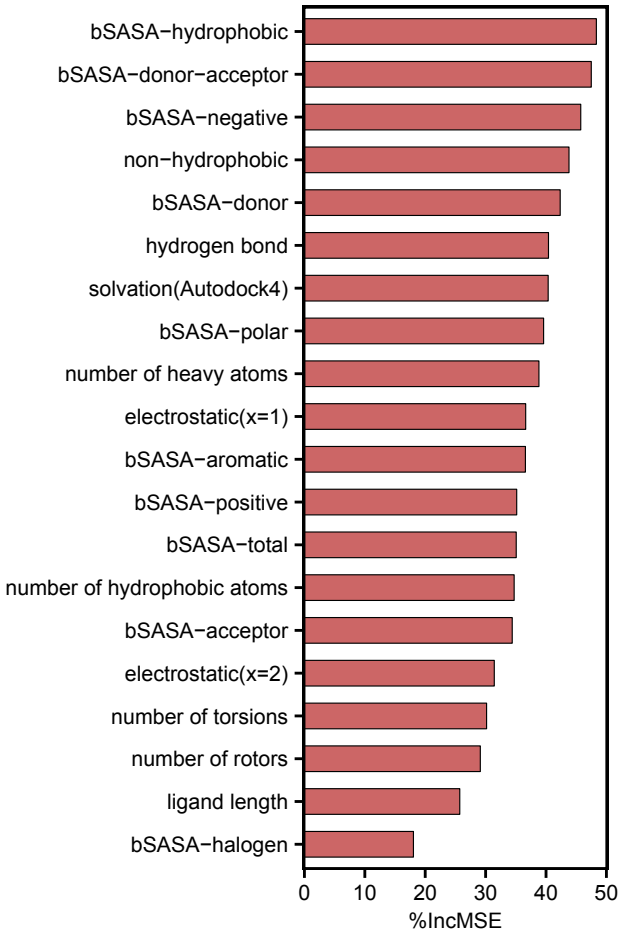


**Figure S4.** Importance of features measured by percentage of increased mean squared error (%IncMSE) for 20 features in $\Delta_{vina}RF_{20}$

**Table S1.** Training Set Description

| | number of structures | source |
|---|---|---|
| Experimental | 3336 | PDBbind v2014 refined set[1]<br>Selected structures from PDBbind v2014 general set[1]<br>Crystal structures from CSAR-NRC HiQ[2,3] |
| Decoy | 3322 | Selected decoy structures from CSAR-NRC HiQ[2,3] |

**Table S2.** AutoDock Vina Features in Source Code[a]

| type | terms | parameters | number of terms |
|---|---|---|---|
| Intermolecular Interaction[b] | ad4_desolvation | $q$ | 2 |
| | electrostatic | $x$ | 2 |
| | gauss | $o, w$ | 31 |
| | repulsion | $o$ | 8 |
| | hydrophobic | $b$ | 4 |
| | non_hydrophobic | | 1 |
| | vdW | | 1 |
| | hydrogen bond | $b$ | 3 |
| Conformation Independent | num_tors | | 1 |
| | num_rotors | | 1 |
| | num_heavy_atoms | | 1 |
| | num_hydrophobic_atoms | | 1 |
| | ligand_max_num_h_bonds | | 1 |
| | ligand_lengths_sum | | 1 |

[a]All terms are already defined with the AutoDock Vina source code. [b]The intermolecular interaction terms are the same as in smina.[4]

**Table S3.** Pharmacophore Type Definition

| pharmacophore type name | definition[a] |
|---|---|
| Positive | N.4 (4 *) |
| | N.2 (3 *) |
| | N.pl3 (C.cat) |
| Negative | O (C (2 O or S [*])) |
| | O (P (2 O or S [*])) |
| | O (S (3 O [*])) |
| | S (C (4 *)) [*] |
| | S (C (2 (O or S [*]))) |
| Donor-Acceptor | O (H) |
| | N.3 (H) |
| | N.2 (H) |
| | N.pl3 (H) |
| | S (H) |
| Donor | N.ar (H) |
| | N.am (H) |
| Acceptor | O Default |
| | N.3 |
| | N.1 |
| | N.ar (2 *) |
| | N.pl3 |
| | S [3 *] |
| Aromatic | N.ar |
| | C.ar |
| Hydrophobic | C [N] [O] [F] [P] [S] |
| Polar | N.am |
| | S (3 *) |
| | C (N) (O) (F) (P) (S) |
| | P |
| Halogen | F |
| | Cl |
| | Br |
| | I |
| Null | * |

[a]The syntax is the same as in DOCK.[5] The parentheses ( ) specify "atoms that must be bonded to the parent atom", while square brackets [ ] specify "atom that must not be bonded to the parent atom". The integer in the definition represents the number of atoms associated to the parent atom. For example, oxygen in carboxylate or thioacetate is denoted as O (C (2 O or S [*])). The syntax

means oxygen bonded to the carbon connected with at least 2 oxygen or sulfur atoms, which are not connected to other atoms. Pharmacophore types are only assigned to nine elements (C, N, O, F, S, P, Cl, Br, I), which is different from the pharmacophore types in DOCK. Only one pharmacophore type is assigned to each atom by the order of pharmacophore types in the table.

**Table S4.** Summary of the Scoring Functions Evaluated in CASF-2007 Benchmark

| score function | number of descriptors | reference |
|---|---|---|
| LigScore1/LigScore2@DS | 3/3 | 6 |
| PLP1/PLP2@DS | 2/3 | 7,8 |
| PMF/PMF04@DS | 1/1 | 9-12 |
| Jain@DS | 5 | 13 |
| LUDI1/LUDI2/LUDI3@DS | 4/4/4 | 14,15 |
| GoldScore@GOLD | 4 | 16 |
| ChemScore@GOLD | 6 | 17,18 |
| ASP@GOLD | 3 | 19 |
| G-Score@SYBYL | 6 | 16 |
| PMF@SYBYL | 1 | 9-12 |
| D-Score@SYBYL | 2 | 20 |
| ChemScore@SYBYL | 4 | 17,18 |
| F-Score@SYBYL | 5 | |
| GlideScore-SP/XP@Schrödinger | 10/10 | 21-23 |
| $DrugScore^{PDB}$/$DrugScore^{CSD}$ | 2/2 | 24,25 |
| X-Score | 4 | 26 |
| AutoDock Vina | 6 | 27 |
| $\Delta_{vina}RF_{20}$ | 20 | |

**Table S5.** Summary of the Scoring Functions Evaluated in CASF-2013 Benchmark

| score function | number of descriptors | reference |
|---|---|---|
| LigScore1/LigScore2@DS | 3/3 | 6 |
| PLP1/PLP2@DS | 2/3 | 7,8 |
| PMF/PMF04@DS | 1/1 | 9-12 |
| Jain@DS | 5 | 13 |
| LUDI1/LUDI2/LUDI3@DS | 4/4/4 | 14,15 |
| GoldScore@GOLD | 4 | 16 |
| ChemScore@GOLD | 6 | 17,18 |
| ChemPLP@GOLD | 9 | 28 |
| ASP@GOLD | 3 | 19 |
| G-Score@SYBYL | 6 | 16 |
| PMF@SYBYL | 1 | 9-12 |
| D-Score@SYBYL | 2 | 20 |
| ChemScore@SYBYL | 4 | 17,18 |
| GlideScore-SP/XP@Schrödinger | 10/10 | 21-23 |
| London-dG@MOE | 4 | |
| ASE@MOE | 1 | 29 |
| Affinity@MOE | 6 | |
| Alpha-HB@MOE | 2 | |
| X-Score | 4 | 26 |
| AutoDock Vina | 6 | 27 |
| $\Delta_{vina}RF_{20}$ | 20 | |

**Table S6.** Performance of 18 Scoring Functions in the Scoring Power Test in CASF-2007[a]

| score function[b] | *N* | *R* | SD |
|---|---|---|---|
| $\Delta_{vina}RF_{20}$ | 195 | 0.732 | 1.62 |
| X-Score::HMScore | 195 | 0.644 | 1.83 |
| DrugScore$^{CSD}$ | 195 | 0.569 | 1.96 |
| AutoDock Vina | 195 | 0.566 | 1.96 |
| SYBYL::ChemScore | 195 | 0.555 | 1.98 |
| DS::PLP1 | 195 | 0.545 | 2.00 |
| GOLD::ASP | 193 | 0.534 | 2.02 |
| SYBYL::G-Score | 195 | 0.492 | 2.08 |
| DS::LUDI3 | 195 | 0.487 | 2.09 |
| DS::LigScore2 | 193 | 0.464 | 2.12 |
| GlideScore-XP | 178 | 0.457 | 2.14 |
| DS::PMF | 193 | 0.445 | 2.14 |
| GOLD::ChemScore | 178 | 0.441 | 2.15 |
| SYBYL::D-Score | 195 | 0.392 | 2.19 |
| DS::Jain | 189 | 0.316 | 2.24 |
| GOLD::GoldScore | 169 | 0.295 | 2.29 |
| SYBYL::PMF-Score | 190 | 0.268 | 2.29 |
| SYBYL::F-Score | 185 | 0.216 | 2.35 |

[a]The results except $\Delta_{vina}RF_{20}$ and AutoDock Vina are obtained from literature.[30] [b]Scoring functions are ranked by the Pearson correlation coefficients.

**Table S7.** Performance of 18 Scoring Functions in Ranking Power Test in CASF-2007[a]

| score function[b] | success rates (%) |
|---|---|
| X-Score::HSScore | 58.5 |
| $\Delta_{vina}RF_{20}$ | 56.9 |
| AutoDock Vina | 53.8 |
| DS::PLP2 | 53.8 |
| $DrugScore^{CSD}$ | 52.3 |
| SYBYL::ChemScore | 47.7 |
| SYBYL::D-Score | 46.2 |
| SYBYL::G-Score | 46.2 |
| GOLD::ASP | 43.1 |
| DS::LUDI3 | 43.1 |
| DS::Jain | 41.5 |
| DS::PMF | 41.5 |
| SYBYL::PMF-Score | 38.5 |
| GOLD::ChemScore | 36.9 |
| DS::LigScore2 | 35.4 |
| GlideScore-XP | 33.8 |
| SYBYL::F-Score | 29.2 |
| GOLD::GoldScore | 23.1 |

[a]The results except $\Delta_{vina}RF_{20}$ and AutoDock Vina are obtained from literature.[30] [b]Scoring functions are ranked by their success rates.

**Table S8.** Success rates of 18 Scoring Functions in the Docking Power Test in CASF-2007[a]

| | Docking (%) | | |
|---|---|---|---|
| scoring function[b] | The top pose | Top two poses | Top three poses |
| GOLD::ASP | 82.5 | 90.2 | 92.3 |
| $\Delta_{vina}RF_{20}$ | 80.5 | 87.7 | 89.2 |
| AutoDock Vina | 77.9 | 86.2 | 88.2 |
| DS::PLP1 | 75.4 | 86.9 | 90.2 |
| DrugScorePDB::PairSurf | 74.3 | 89.1 | 91.8 |
| GlideScore::SP | 73.2 | 83.1 | 86.9 |
| DS::LigScore2 | 71.6 | 85.8 | 88.0 |
| GOLD::ChemScore | 70.5 | 78.7 | 82.0 |
| GOLD::GoldScore | 68.9 | 79.8 | 85.2 |
| X-Score1.2::HMScore | 68.3 | 82.0 | 84.2 |
| SYBYL::F-Score | 64.5 | 74.3 | 82.0 |
| SYBYL::ChemScore | 60.1 | 72.1 | 78.7 |
| DS::LUDI2 | 57.4 | 70.5 | 75.4 |
| SYBYL::PMF-Score | 48.1 | 61.7 | 63.4 |
| DS::Jain | 44.8 | 63.9 | 70.5 |
| DS::PMF | 43.7 | 53.0 | 56.8 |
| SYBYL::G-Score | 41.5 | 55.7 | 67.2 |
| SYBYL::D-Score | 30.6 | 50.3 | 59.0 |

[a]The results except $\Delta_{vina}RF_{20}$ and AutoDock Vina are obtained from literature.[30] [b]Scoring functions are ranked by success rates when the top one poses are considered.

**Table S9.** Performance of 22 Scoring Functions in the Scoring Power Test in CASF-2013[a]

| score function[b] | *N* | *R* | SD |
|---|---|---|---|
| $\Delta_{vina}RF_{20}$ | 195 | 0.686 | 1.64 |
| X-Score$^{HM}$ | 195 | 0.614 | 1.78 |
| ChemScore@SYBYL | 195 | 0.592 | 1.82 |
| ChemPLP@GOLD | 195 | 0.579 | 1.84 |
| PLP1@DS | 195 | 0.568 | 1.86 |
| G-Score@SYBYL | 195 | 0.558 | 1.87 |
| AutoDock Vina | 195 | 0.557 | 1.85 |
| ASP@GOLD | 195 | 0.556 | 1.88 |
| ASE@MOE | 195 | 0.544 | 1.89 |
| ChemScore@GOLD | 189 | 0.536 | 1.90 |
| D-Score@SYBYL | 195 | 0.526 | 1.92 |
| Alpha-HB@MOE | 195 | 0.511 | 1.94 |
| LUDI3@DS | 195 | 0.487 | 1.97 |
| GoldScore@GOLD | 189 | 0.483 | 1.97 |
| Affinity-dG@MOE | 195 | 0.482 | 1.98 |
| LigScore2@DS | 190 | 0.456 | 2.02 |
| GlideScore-SP | 169 | 0.452 | 2.03 |
| Jain@DS | 191 | 0.408 | 2.05 |
| PMF@DS | 194 | 0.364 | 2.11 |
| GlideScore-XP | 164 | 0.277 | 2.18 |
| London-dG@MOE | 195 | 0.242 | 2.19 |
| PMF@SYBYL | 191 | 0.221 | 2.20 |

[a]The results except $\Delta_{vina}RF_{20}$ and AutoDock Vina are obtained from literature. [b]Scoring functions are ranked by the Pearson correlation coefficients.

**Table S10.** Performance of 22 Scoring Functions in Ranking Power Test in CASF-2013[a]

| | Ranking success rates (%) | |
|---|---|---|
| score function[b] | high-level | low-level |
| X-Score$^{HM}$ | 58.5 | 72.3 |
| ChemPLP@GOLD | 58.5 | 72.3 |
| $\Delta_{vina}RF_{20}$ | 55.4 | 73.8 |
| PLP2@DS | 55.4 | 72.3 |
| GoldScore@GOLD | 55.4 | 76.9 |
| ChemScore@SYBYL | 53.8 | 67.7 |
| Affinity-dG@MOE | 53.8 | 66.2 |
| LigScore1@DS | 52.3 | 61.5 |
| Alpha-HB@MOE | 52.3 | 66.2 |
| G-Score@SYBYL | 52.3 | 72.3 |
| LUDI1@DS | 52.3 | 69.2 |
| AutoDock Vina | 49.2 | 72.3 |
| D-Score@SYBYL | 49.2 | 63.1 |
| PMF@DS | 49.2 | 66.2 |
| ASP@GOLD | 47.7 | 72.3 |
| ChemScore@GOLD | 46.2 | 63.1 |
| London-dG@MOE | 43.1 | 60.0 |
| PMF@SYBYL | 43.1 | 61.5 |
| GlideScore-SP | 43.1 | 56.9 |
| Jain@DS | 41.5 | 58.5 |
| ASE@MOE | 40.0 | 64.6 |
| GlideScore-XP | 35.4 | 47.7 |

[a]The results except $\Delta_{vina}RF_{20}$ and AutoDock Vina are obtained from literature.[31] [b]Scoring functions are ranked by high-level success rates.

**Table S11.** Success rates of 22 Scoring Functions in the Docking Power Test in CASF-2013[a]

| | docking (%) | | |
|---|---|---|---|
| score function[b] | The top pose | Top two poses | Top three poses |
| $\Delta_{vina}RF_{20}$ | 86.7 | 93.3 | 94.4 |
| AutoDock Vina | 85.1 | 89.7 | 92.3 |
| ChemPLP@GOLD | 81.0 | 86.7 | 89.7 |
| ChemScore@GOLD | 77.9 | 83.1 | 88.2 |
| GlideScore-SP | 78.5 | 85.6 | 87.7 |
| LigScore2@DS | 76.9 | 84.1 | 86.7 |
| PLP1@DS | 76.9 | 84.1 | 86.2 |
| Alpha-HB@MOE | 75.4 | 82.6 | 86.2 |
| GlideScore-XP | 74.4 | 82.6 | 85.6 |
| ASP@GOLD | 71.8 | 81.5 | 87.2 |
| GoldScore@GOLD | 71.3 | 81.0 | 85.6 |
| LUDI1@DS | 59.0 | 75.4 | 83.1 |
| Affinity-dG@MOE | 63.1 | 74.9 | 81.0 |
| ChemScore@SYBYL | 59.5 | 69.2 | 75.4 |
| London-dG@MOE | 59.5 | 73.8 | 78.5 |
| X-Score$^{HM}$ | 61.0 | 73.3 | 77.9 |
| G-Score@SYBYL | 45.1 | 61.5 | 72.3 |
| Jain@DS | 48.2 | 62.1 | 70.8 |
| PMF@SYBYL | 51.8 | 60.0 | 66.7 |
| PMF04@DS | 51.8 | 62.6 | 66.2 |
| ASE@MOE | 51.3 | 60.0 | 63.6 |
| D-Score@SYBYL | 18.5 | 29.7 | 42.6 |

[a]The results except $\Delta_{vina}RF_{20}$ and AutoDock Vina are obtained from literature.[31] [b]Scoring functions are ranked by success rates when the top one poses are considered.

**Table S12.** Enrichment Factors of 22 Scoring Functions in the Screening Power Test in CASF-2013[a]

| | enrichment factor | | |
|---|---|---|---|
| score function[b] | top 1% | top 5% | top 10% |
| $\Delta_{vina}RF_{20}$ | 20.95 | 6.44 | 3.89 |
| GlideScore-SP | 19.54 | 6.27 | 4.14 |
| ChemScore@GOLD | 18.90 | 6.83 | 4.08 |
| GlideScore-XP | 16.81 | 6.02 | 4.07 |
| LigScore2@DS | 15.90 | 6.23 | 3.51 |
| AutoDock Vina | 15.57 | 5.55 | 3.62 |
| ChemPLP@GOLD | 14.28 | 5.88 | 4.31 |
| LUDI1@DS | 12.53 | 4.28 | 2.80 |
| ASP@GOLD | 12.36 | 6.23 | 3.79 |
| Affinity-dG@MOE | 8.21 | 4.15 | 3.19 |
| London-dG@MOE | 8.08 | 3.36 | 2.51 |
| GoldScore@GOLD | 7.95 | 4.52 | 3.16 |
| PLP1@DS | 6.92 | 4.28 | 3.04 |
| Jain@DS | 5.90 | 2.51 | 1.80 |
| PMF@SYBYL | 5.38 | 2.21 | 1.90 |
| ChemScore@SYBYL | 5.26 | 2.38 | 2.18 |
| Alpha-HB@MOE | 4.87 | 3.23 | 1.32 |
| PMF04@DS | 4.87 | 2.87 | 2.63 |
| ASE@MOE | 4.36 | 2.35 | 1.59 |
| X-Score$^{HM}$ | 2.31 | 2.14 | 1.41 |
| D-Score@SYBYL | 2.31 | 1.79 | 1.46 |
| G-Score@SYBYL | 1.92 | 1.26 | 1.44 |

[a]The results except $\Delta_{vina}RF_{20}$ and AutoDock Vina are obtained from literature.[31] [b]Scoring functions are ranked by their average enrichment factor obtained at the top 1% level.

**Table S13.** Success Rates of Finding the Best Ligand Molecule of 22 Scoring Functions in the Screening Power Test in CASF-2013[a]

| | success rates of finding the best ligand molecule among (%)[c] | | |
|---|---|---|---|
| score function[b] | top 1% | top 5% | top 10% |
| $\Delta_{vina}RF_{20}$ | 60.0 (39) | 80.0 (52) | 86.2 (56) |
| GlideScore-SP | 60.0 (39) | 72.3 (47) | 76.9 (50) |
| GlideScore-XP | 52.3 (34) | 69.2 (45) | 73.8 (48) |
| ChemScore@GOLD | 49.2 (32) | 78.5 (51) | 83.1 (54) |
| LigScore2@DS | 47.7 (31) | 75.4 (49) | 83.1 (54) |
| AutoDock Vina | 44.6 (29) | 60.0 (39) | 69.2 (45) |
| ChemPLP@GOLD | 41.5 (27) | 70.8 (46) | 84.6 (55) |
| LUDI2@DS | 38.5 (25) | 53.8 (35) | 66.2 (43) |
| ASP@GOLD | 36.9 (24) | 75.4 (49) | 81.5 (53) |
| Affinity-dG@MOE | 23.1 (15) | 50.8 (33) | 66.2 (43) |
| PLP1@DS | 21.5 (14) | 52.3 (34) | 70.8 (46) |
| GoldScore@GOLD | 21.5 (14) | 52.3 (34) | 66.2 (43) |
| London-dG@MOE | 21.5 (14) | 36.9 (24) | 49.2 (32) |
| Jain@DS | 16.9 (11) | 29.2 (19) | 40.0 (26) |
| ChemScore@SYBYL | 15.4 (10) | 33.8 (22) | 50.8 (33) |
| Alpha-HB@MOE | 13.8 (9) | 36.9 (24) | 63.1 (41) |
| PMF@SYBYL | 13.8 (9) | 23.1 (15) | 38.5 (25) |
| PMF04@DS | 12.3 (8) | 30.8 (20) | 47.7 (31) |
| ASE@MOE | 12.3 (8) | 30.8 (20) | 38.5 (25) |
| X-Score$^{HM}$ | 9.23 (6) | 21.5 (14) | 32.3 (21) |
| D-Score@SYBYL | 6.15 (4) | 20.0 (13) | 24.6 (16) |
| G-Score@SYBYL | 4.62 (3) | 16.9 (11) | 30.8 (20) |

[a]The results except $\Delta_{vina}RF_{20}$ and AutoDock Vina are obtained from literature.[31] [b]Scoring functions are ranked by their success rates obtained at the top 1% level. [c]Numbers in brackets are the number of successful cases over total 65 cases.

**References**


[1] Y. Li, Z. H. Liu, J. Li, L. Han, J. Liu, Z. X. Zhao, R. X. Wang, *J. Chem. Inf. Model.* **2014**, *54*, 1700.

[2] J. B. Dunbar, Jr., R. D. Smith, C. Y. Yang, P. M. Ung, K. W. Lexa, N. A. Khazanov, J. A. Stuckey, S. Wang, H. A. Carlson, *J. Chem. Inf. Model.* **2011**, *51*, 2036.

[3] S. Y. Huang, X. Q. Zou, *J. Chem. Inf. Model.* **2011**, *51*, 2107.

[4] D. R. Koes, M. P. Baumgartner, C. J. Camacho, *J. Chem. Inf. Model.* **2013**, *53*, 1893.

[5] L. L. Jiang, R. C. Rizzo, *J. Phys. Chem. B* **2015**, *119*, 1083.

[6] A. Krammer, P. D. Kirchhoff, X. Jiang, C. M. Venkatachalam, M. Waldman, *J. Mol. Graphics Modell.* **2005**, *23*, 395.

[7] G. Verkhivker, K. Appelt, S. T. Freer, J. E. Villafranca, *Protein Eng.* **1995**, *8*, 677.

[8] G. M. Verkhivker, D. Bouzida, D. K. Gehlhaar, P. A. Rejto, S. Arthurs, A. B. Colson, S. T. Freer, V. Larson, B. A. Luty, T. Marrone, P. W. Rose, *J. Comput.-Aided Mol. Des.* **2000**, *14*, 731.

[9] I. Muegge, Y. C. Martin, *J. Med. Chem.* **1999**, *42*, 791.

[10] I. Muegge, *Perspect. Drug Discovery Des.* **2000**, *20*, 99.

[11] I. Muegge, *J. Comput. Chem.* **2001**, *22*, 418.

[12] I. Muegge, *J. Med. Chem.* **2006**, *49*, 5895.

[13] A. N. Jain, *J. Comput.-Aided Mol. Des.* **1996**, *10*, 427.

[14] H. J. Bohm, *J. Comput.-Aided Mol. Des.* **1994**, *8*, 243.

[15] H. J. Bohm, *J. Comput.-Aided Mol. Des.* **1998**, *12*, 309.

[16] G. Jones, P. Willett, R. C. Glen, A. R. Leach, R. Taylor, *J. Mol. Biol.* **1997**, *267*, 727.

[17] M. D. Eldridge, C. W. Murray, T. R. Auton, G. V. Paolini, R. P. Mee, *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425.

[18] C. A. Baxter, C. W. Murray, D. E. Clark, D. R. Westhead, M. D. Eldridge, *Proteins: Struct., Funct., Genet.* **1998**, *33*, 367.

[19] W. T. M. Mooij, M. L. Verdonk, *Proteins: Struct., Funct., Bioinf.* **2005**, *61*, 272.

[20] T. J. A. Ewing, S. Makino, A. G. Skillman, I. D. Kuntz, *J. Comput.-Aided Mol. Des.* **2001**, *15*, 411.

[21] R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, P. S. Shenkin, *J. Med. Chem.* **2004**, *47*, 1739.

[22] T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard, J. L. Banks, *J. Med. Chem.* **2004**, *47*, 1750.

[23] R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin, D. T. Mainz, *J. Med. Chem.* **2006**, *49*, 6177.

[24] H. Gohlke, M. Hendlich, G. Klebe, *J. Mol. Biol.* **2000**, *295*, 337.

[25] H. F. G. Velec, H. Gohlke, G. Klebe, *J. Med. Chem.* **2005**, *48*, 6296.

[26] R. X. Wang, L. H. Lai, S. M. Wang, *J. Comput.-Aided Mol. Des.* **2002**, *16*, 11.

[27] O. Trott, A. J. Olson, *J. Comput. Chem.* **2010**, *31*, 455.

[28] O. Korb, T. Stutzle, T. E. Exner, *J. Chem. Inf. Model.* **2009**, *49*, 84.

[29] J. Goto, R. Kataoka, H. Muta, N. Hirayama, *J. Chem. Inf. Model.* **2008**, *48*, 583.

[30] T. J. Cheng, X. Li, Y. Li, Z. H. Liu, R. X. Wang, *J. Chem. Inf. Model.* **2009**, *49*, 1079.

[31] Y. Li, L. Han, Z. H. Liu, R. X. Wang, *J. Chem. Inf. Model.* **2014**, *54*, 1717.