

## **SAMPL: Statistical assessment of the modeling of proteins and ligands**

A. Skillman, G. Warren,  
P. Hawkins, A. Nicholls  
*OpenEye Scientific Software, Inc, Santa Fe, USA*

Opportunities for prospective or blinded validation of computational models in drug discovery are rare yet valuable. SAMPL provides the computational chemistry community an opportunity to evaluate a variety of methods on previously unpublished or difficult to discover data.

We will report on the second annual SAMPL evaluation, a blinded trial of a variety of computational tasks. We received three unpublished data collections for use in the study. Abbott Laboratories provided twenty-seven Urokinase inhibitors with measured affinities and co-crystal structures. Vertex Pharmaceuticals provided fifty-two JNK-3 Kinase inhibitors with measured affinities and co-crystal structures. Finally, Peter Guthrie provided sixty-three water-vacuum transfer energies, calculated from data collected from obscure sources.

We used these data to generate eleven different blinded experiments in which participants could make predictions. All experiments were open to the public for approximately four months. Four experiments examined virtual screening, four experiments examined pose-prediction, two experiments examined affinity prediction, and one experiment examined vacuum-water transfer energies. We encouraged professional modelers to participate using third-party software in addition to both academic and industrial software developers. SAMPL attracted approximately fifty participants from North America and Europe and we accepted over one hundred public and anonymous predicted data sets.

The virtual screening experiments were designed to address important questions in evaluation design as well as generate feedback for individual algorithms. These four experiments allowed evaluation of both ligand-based and structure-based design programs. Pairs of experiments with the same ligands were used to measure the change in performance with improved knowledge of the protein structure. They also included simultaneous comparison of six different decoy sets including DUD-like decoys (1), Drug-like decoys (2) and the Rognan decoys (3) allowing evaluation of each method's performance on these well-known decoy collections in the same system. In addition, the predictions were evaluated on all the ligands provided as well as on a subset of low-potency chemically independent ligands, to simulate a prospective approach to a new target.

Pose prediction experiments included an initial cross-docking phase followed by a self-docking phase. Each method is evaluated based how well they perform in each experiment and how much they improve going from cross-docking to self-docking experiment. We will discuss important issues regarding crystal structure quality and its impact on structure reproduction evaluation.

At the last stage, the co-crystal models were provided to participants to predict binding affinities. This portion of SAMPL saw the most varied methods, including QSAR models, Docking algorithms, implicit solvent molecular mechanics calculations, explicit solvent molecular mechanics as well as Monte Carlo perturbation approaches. We will present an analysis of each model's ability to predict the absolute affinity, the relative affinity, and the rank of the inhibitors, including an assessment of the experimental error.

The vacuum-water transfer energy prediction data set included highly flexible multifunctional molecules that are much more challenging than typical transfer energy data sets. This data set also included some large drug-like molecules. This valuable data set allowed new insights into the effects of partial charge models upon salvation energy.

We will present a summary of all the results. It will include an assessment of the field in general, the effects of increasing information content on our ability to make predictions, trends in virtual screening predictions with different decoys and a discussion of affinity prediction and transfer energy. Finally, we will briefly discuss plans for the next SAMPL challenge in 2009.

1. Haung, N; Shoichet, B.K.; Irwin, J.J. Benchmarking Sets for Molecular Docking. *J. Med. Chem.* **2006**, 49, 6789-6801.
2. Ajay; Walters, W.P.; Murcko, M.A. Can We Learn To Distinguish between “Drug-like” and “Nondrug-like” Molecules? *J. Med. Chem.* **1998**, 41, 3314-3324.
3. Bissantz, C.; Folkers, G.; Rognan, D. Protein-Based Virtual Screening of Chemical Databases. 1. Evaluation of Different Docking/Scoring Combinations. *J. Med. Chem.* **2000**, 43, 4759-4767.