**The Molecular Graphics & Modeling Society:**
**Docking and Scoring in Structure Guided Drug Design,**
**University of Southampton, April 5th – 7th, 2006.**

The conference website can be found at: http://www.soton.ac.uk/~jwe1/index.html

The main focus of the conference was on protein-ligand and protein-protein docking, high-throughput virtual screening and scoring of molecular interactions. Several new docking methodologies – both algorithms and their implementations – and improvements on existing ones were presented, as well as actual results obtained with these methods.
The evaluation of molecular interactions (scoring) is one of the fundamental problems of molecular docking; Several methods for obtaining high quality absolute binding free energies (free energy calculations), which are mostly used during the lead optimization phase, were discussed, whereas less attention was directed at computationally less demanding scoring functions which are used during actual high- to mid-throughput docking runs. Pre- or post-docking processing of results, such as various filtering strategies, were also discussed by several speakers.

Another aspect of the docking procedure is the validation of results and, in particular, the quality of the used test-sets. This issue, as well as the feasibility of using e.g. root mean square deviation (RMSD) as a measure for similarity (the problems associated with the RMSD measure are quite well known, but still its use, and misuse, persists) when validating results, was addressed by several speakers.

Although nothing particularly ground-breaking was presented at the conference, it certainly is evident that the field is steadily evolving – and has been for the last two decades: there is increasing interest towards the field and significant effort is put into the development of new docking methods, more accurate scoring, etc. However, ligand docking and scoring are very complex problems to which there yet are no good solutions. William Jorgensen, who is one of the pioneers of computational drug design and free energy calculations, expressed this point quite clearly: “if you think you have found an easy solution, then you're wrong”.

Below are short descriptions of some of the talks and posters which I found quite interesting. Note that they contain comments and interpretations that may not represent the views of the authors of the original works. Abstracts of oral presentations and posters can be found in PDF format at the MGMS conference website (http://www.soton.ac.uk/~jwe1/abstracts.html).

Santeri Puranen
Structural Bioinformatics Laboratory
Åbo Akademi University

**Importance of Accurate Charges in Molecular Docking: Quantum Mechanical/Molecular Mechanical (QM/MM) Approach (Andrew Sparkes, Schrödinger Ltd.)**

The accuracy of ligand docking can be increased if the partial charges of the ligand atoms are calculated using QM methods and updated for each new ligand pose, i.e. the ligands are treated as electrostatically polarizable. This is implemented in Schrödingers QSite program which is basically a Python script that uses the GLIDE docking program and the Jaguar QM package. Schrödinger calls this Quantum Polarized Ligand Docking (QPLD). According to a Schrödinger representative the tool is good for finding decent binding modes, thus, it is suited for lead optimization (one may not obtain quantitatively accurate ligand scores, though).

**QM/MM-Poisson-Boltzmann scoring of protein/ligand binding affinities (Sonja Schwarzl et al., University of Heidelberg)**

Solvation effects have been estimated using the Poisson-Boltzmann/Surface Area method. The methodology allows for relatively fast and accurate calculation of absolute binding free energies (about 10 CPU-minutes per conformer and a 1.2kcal/mol RMS error for test-cases). One of the problems with using molecular mechanics methods is that interaction parameters are needed for all atom types, bonds, etc. These parameters are typically available for protein atoms, whereas ligands may contain types that are not accounted for in a particular force field. This problem can be overcome by using an ab initio or semi-empirical treatment of the ligand. In this case the calculations are based on a combination of semiempirical calculations (AM1) and molecular mechanics (CHARMM). Semi-empirical methods may also increase the accuracy of the free energy calculations by a more accurate treatment of electrostatics.

**PLANTS: Protein-Ligand ANT System (Oliver Korb and Thomas Exner, Universität Konstanz)**

PLANTS is a ligand docking tool that uses the Ant Colony Optimization (ACO) algorithm for finding the optimal fit for a ligand. This is interesting since ACO has not been used for ligand docking before. ACO is a fairly new type of optimization algorithm which is suitable for solving multi-variate optimization problems with multi-modal objective functions, as is the case for the docking of a ligand molecule to a receptor: the typical energy landscape found in ligand docking is highly rugged and noisy, making it a very challenging target for optimization. The results so far have been quite promising, although the current implementation only treats ligands as flexible, whereas the protein is rigid. However, the authors stated that the incorporation of protein side chain flexibility is one of the main priorities in future development.

**Structure-Based Virtual Screening, Validation and Application (Marcel Verdonk, Astex Therapeutics)**

One of the challenges with validating docking results is that existing test-sets may be ambiguous in terms of protonation states, tautomeric forms, water mediation and the test-sets may in some cases even include reference complexes where there is experimental uncertainty (such as poor electron density and clear errors even). To remedy this the authors have prepared “a new validation set [containing 85 diverse structures], each with a unique drug target, for which the ligand is a drug-like compound and has unambiguous electron density”. The validation set should be of high quality and will be made publicly available “soon” (when the work has been properly finished). Moreover, it is widely known that RMSD is not always a good measure for a fit, but still it is routinely used. The authors suggest that a similarity measure based on volume overlap or electron density should be preferred instead.

**Permuting input strings for more effective generation of 3D conformers (Giorgio Carta, et al., Trinity College Dublin)**

The authors had noticed that different permutations of SMILES strings, all describing the same molecule, resulted in different conformations when converted to 3D molecules and that all major tools that perform this conversions (such as Corina) seem to behave in this way (note that this is purely a property of the algorithms employed in these programs and that canonical SMILES strings always give the same geometry). This property (or bug, depending on how one looks at it) can be exploited as a fast and efficient way of making conformationally diverse sets of molecules: all one needs to do is to make systematic permutations to SMILES strings. One should also be aware of this property when doing ligand docking with molecules that are converted from SMILES strings into 3D structures: docking programs may not thoroughly sample the available conformational space of a ligand (it is e.g. common to only modify bond torsions), hence, the docking result may depends on the input geometry of the ligand and, ultimately, on how the SMILES string was formulated.