Welcome to the fifth edition of the BioSolveIT newsletter!

We will be publishing information on our new developments, events, milestones and industry facts on a quarterly basis.

FlexX Version 2.0 – beta phase started

A selected set of customers now evaluate the latest version of FlexX. The beta-phase will take approximately 3 months, after which the general release is planned. With this a major step in BioSolveIT's efforts to supply customers with the decisive advantage of the latest developments in docking technology coupled with convenient user-access is achieved.

Among the major features of 2.0 are speedup of up to 30% on the PC platform, completely re-worked static data files for improved chemistry, compatibility of the major modules PHARM, C, and FlexE, greatly enhanced FlexX-PHARM, a reworked Python-interface, SMARTS-support and the ability to read now all major file formats (MOL2, SDF, SLN, SMILES, CSLN, and PDB).

However, this is only the top-level functionality. A lot of improvement has been done behind the scenes to make FlexX even more reliable and robust. You’ll find the detailed list of features here.

Gene/Protein network visualisation – ToPNet Release 1.3 Beta available

A first release of ToPNet was distributed in 2003. Now a new version of our tool for the joint analysis of gene networks and expression data is available for evaluation.

ToPNet release 1.3 offers the following new features:

- **New Gene Expression GUI:**
  A completely re-designed GUI for loading gene expression data now makes it very easy to integrate protein or gene networks with your expression measurements.

- **GO-Annotation:**
  Extracted sub-networks can be analyzed according to GO-terms. You get a selection of related GO-terms that are most specific for your network. This may be useful, e.g. for relating biological functions with selected Significant Areas to see which biological processes or molecular functions are affected by your expression experiments.

- **Pathway Queries:**
  This new feature lets you search for pathways that are interesting with respect to certain criteria. For example, you may look for ‘druggable’ targets like kinases that should phosphorylate a transcription factor regulating genes with a significant expression change in a certain experiment. To allow for such complex queries, ToPNet contains an XML-based query language that is used to formulate pathway templates by posing restrictions on network nodes and their connections.

Permute – combinatorial compound variation on-the-fly

Permute is a powerful new FlexX-module for on-the-fly generation and docking of compound variations. Based on the widely used SMARTS substructure-definition, libraries for protonation-states, tautomers or close chemical analogues can be defined.

Permute generates a small combinatorial library for each ligand (containing e.g. the different protonation states) and these libraries are docked using the FlexX²-technology. Therefore Permute preserves the speed and comfort of single ligand docking.

The frequent issue of either generating multiple copies of the compound with the different protonation states, or a priori selection of a particular protonation state is a past issue with Permute.
Last year we first published results on the basis of this new module at the 3rd Joint Sheffield Conference on Chemoinformatics.

Please find more details on the official Permute page.

WebSite relaunch introduces new contents and functionality

As announced in Newsletter 04, our internet appearance has recently been revised. While some graphical elements had already been re-designed last year in order to implement the company's new corporate ID, the whole content has now been technically revised, and textually restructured and expanded. One major goal was to facilitate conciseness and clarity and to simplify navigation. This will enable visitors to straightforwardly reach their desired piece of information, service or file. Also, we implemented a registration mechanism which grants our users automatic access to all services and at the same time helps us to protect our site against unauthorised access. Find out all about the new structure, the comfort of fly-out menus, and other interesting details in a separate outline.

Tips and Tricks from the world of FlexX

In order to keep up with good traditions, in this section we focus on aspects of FlexX that are either unknown or sometimes misunderstood, yet we see them as important points to bring to your attention.

In this issue we take a closer look at the deepsite option. It is a well know problem that FlexX has a tendency to place ligands at the rim of the pocket so that oftentimes the binding pocket is not completely occupied. The user can actually do a lot to prevent such artifacts of the algorithm. A couple of tips and tricks to mitigate this problem at least to some extent can be found on the tips & tricks page.

If you have any questions or know of any tips and tricks yourself which you would like to share with the FlexX user community, we would appreciate your input at flexx-info@biosolveit.de.

BioSolveIT news in Brief

DDB
BioSolveIT’s Docking database (DDB) is now ready for selected beta testers. With its powerful graphical user interface (GUI), you can sweep through massive amounts of docking data, view molecular sketches of your docked compounds and let the program highlight atoms/groups which form interactions to the protein. Moreover, the GUI assists you in defining complex queries for all sorts of mining your in silico data.

NovoBench
BioSolveIT and Partners (Molecular Networks, ZBH, CCC, Eli Lilly, 4SC, and Altana Pharma) have started a joint project with the aim of an integrated suite for fragment based 'de novo' design of lead compounds. The ‘integration’ of the software extends from a ‘synthesisability’ module to the disappearance of fuss with different data formats. The project with a volume of € 2.000.000 won a three-year grant by the German Federal Ministry of Education and Research (BmBF).

ACS National Meeting in San Diego
BioSolveIT will be at this year’s ACS National Meeting with two oral contributions:

1. Mining docking space (COMP 310) – Christian Lemmen et al.
   The talk will report the idea of our docking database (DDB) which is currently being interfaced to a generic toolbox for machine learning techniques. The graphical user interface makes the extraction of important information from hundreds of thousands of docked compounds a lot easier.

2. Boosting descriptors for similarity searches: feature trees trained by machine learning (CINF 83) – Marcus Gastreich et al.
   This talk focuses on an application of the feature trees descriptor (FTrees) to large, real life datasets. The combination of FTrees with machine learning techniques leads to improved enrichments. The design of both the descriptor and the methodology deliver information on what the important parts of actives are.
BioSolveIT will not maintain a booth at the conference exhibition, however, if you are interested in meeting the participating BioSolveIT members please just drop a note at BSI-at-ACS@biosolveit.de so that we can schedule an individual appointment.

**Literature Corner**

*FlexX-Scan: Fast Structure-Based Virtual Screening*
I. Schellhammer and M. Rarey  
PROTEINS: Structure, Function, and Genetics 2004 vol57, p504-517  
[details here](#)

*Combination of a Naive Bayes Classifier with Consensus Scoring Improves Enrichment of High-Throughput Docking Results*
A.E. Klon, M. Glick, and J.W. Davies  
J. Med. Chem.; (Letter); 2004; 47(18); 4356-4359  
[details here](#)

*Virtual Screening: finding needles in a haystack on a shoestring*
D.E. Clark, N.V. Harris, A.G. Roach, and A.D. Baxter  
Drug Discovery World, 2:37-41, 2004  
[details here](#)

**Upcoming Articles**

- **2Ddraw**
  Computationally it is a hard problem to draw compounds subject to certain constraints. An approach developed by Fricker et al. (at the ZBH Hamburg) covers applications ranging from the trivial alignment of cores in a combinatorial library to a complete 2D-alignment of molecules. Read more about this in the next BioSolveIT newsletter.

- **Distributor**
  The Distributor is an all-purpose “parallelisator” for any application that deals with large input lists. This can be a docking job with hundred-thousands of molecules to be processed but also any other computation that iterates over a list of items (e.g., parameters) and can be started on the command line. We have recently finished the coding of the Distributor and will present it in detail in the next issue.

- **FlexX-Scan**
  A special high-throughput screening module for FlexX: FlexX-Scan is a new development by Ingo Schellhammer and Matthias Rarey from Centre for Bioinformatics at Hamburg University (see also the Literature Corner). This FlexX extension module has shown to shorten the docking runtime to a mere 6 seconds per compound on a modern Linux PC. This means, with a cluster of 32 PCs you’d just require 2 days to dock your entire 1-million-inhouse-compound-library, with FlexX-Scan.

**Contact**

For further information please contact:  

**BioSolveIT GmbH**  
An der Ziegelei 75  
53757 Sankt Augustin  
Germany  
eMail: Newsletter@BioSolveIT.de  
www: www.BioSolveIT.de  
phone: +49-2241-25 25 0  
fax: +49-2241-25 25 525