Pharmaceutical Bioinformatics: new approaches for well-established research

Erxleben A, Flemming S, Grüning B, Senger C, Günther S*

Department of Pharmaceutical Bioinformatics,
Institute for Pharmaceutical Sciences, University of Freiburg, Germany
**e-mail: stefan.guenther@pharmazie.uni-freiburg.de*

The discipline of Bioinformatics was formed to use computers for answering biological questions of modern research. In the last two decades, new experimental technologies in the Life Science area produced information about biological systems to an extent that we have never had before. The examination and interpretation of these data revealed fundamental new insights into the organisation of living cells and the functionality of the involved molecules. The increasing amount of information demands new techniques and approaches in data evaluation, simulation, and visualization. The variety of the work of the Pharmaceutical Bioinformatics Group includes the development of new algorithms and software, modelling of molecule interactions and biological systems, data management, and visualization. Here, we give a brief overview of four different projects that we are currently working on:

Automated genome annotation of Streptomyces spp.

With the increasing amount of available genome sequences over the last few years, there is a high demand for genome annotation. Genome annotation is the process of attaching biological information to sequences of identified elements on the genome. The structural annotation includes the identification of the gene structure and coding regions, open reading frame prediction, and the localization of regulatory motifs. By functional annotation biochemical and biological functions can be assigned to genes. Automatic annotation tools are able to perform these steps by computer analysis. We are using Galaxy[1] as framework for interactive large scale genome analysis which allows the integration of user-defined tools. In our current project we are working in cooperation with the Pharmaceutical Biology & Biotechnology and Pharmaceutical & Medicinal Chemistry on genome annotation of industrially relevant Streptomyces spp. Those bacteria are able to produce a variety of bioactive compounds, e.g. antibiotics[2]. Performing complete genome alignments (Fig. 1) will allow for a comparison of identified genes, products, and metabolites. The results will be used for constructing kinetic models to direct metabolic engineering.

References

1. Blankenberg D, *et al.* Galaxy: a web-based genome analysis tool for experimentalists. Curr Protoc Mol Biol, 2010. 19: 1-21.

2. Watve MG, *et al.* How many antibiotics are produced by the genus Streptomyces? Arch Microbiol, 2001. 176: 386-90.

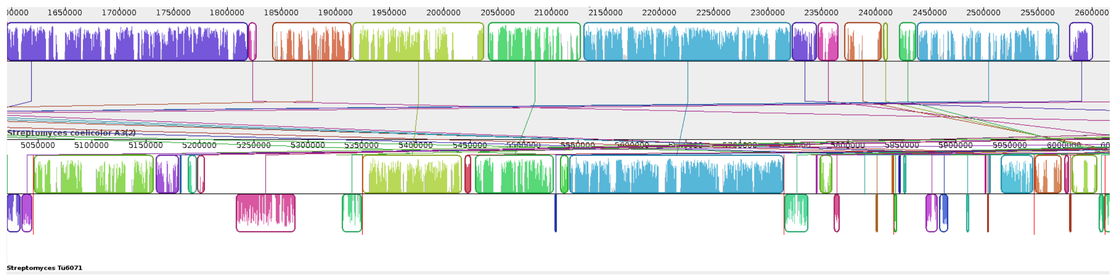

Fig. 1: Genome alignment of two different Streptomyces spp.

Prediction of drug-metabolizing enzymes by support vector machines

Machine learning approaches

A variety of methods exist to classify and predict biological properties of chemical compounds, e.g. principal component analysis, partial least squares, artificial neural networks, evolutionary algorithms, and support vector machines (SVMs).
SVMs are models for non-linear classification and regression. They find a hyperplane with the maximum margin separating samples of two classes of a training set. If samples are not separable linearly, they will be mapped to a high-dimensional "feature" space to find a hyperplane separating classes linearly in that space (Fig. 2). To form a set of meaningful descriptors for classification, different characteristics of chemical compounds like size, shape, surface, ring counts, etc. have to be computed.

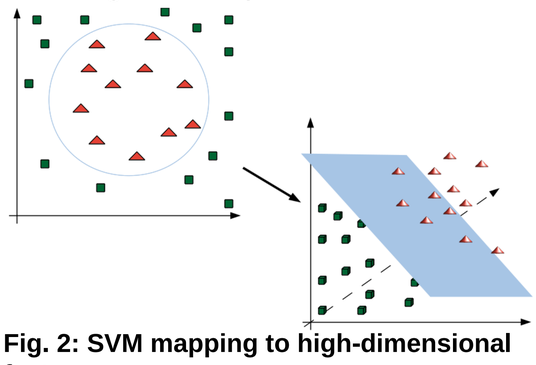

Fig. 2: SVM mapping to high-dimensional feature space.

Application

Cytochromes P450 (CYPs) account for ~75% of all drug-metabolizing enzymes. The prediction of affinities of enzymes and drugs is important for the prevention of drug-drug interactions, frequently caused by multi-medications in elder or intensive care unit patients[1,2].
Databases like SuperCyp[3] provide classified, comprehensive datasets on CYPs, metabolized drugs, and enzyme or drug structures. Utilizing this data as training and test sets allows for the calculation of descriptor sets and creation of prediction models for compounds whose metabolisms are not known, yet.

References

1. Bergk V, *et al.* Drug interactions in primary care: impact of a new algorithm on risk determination. Clin Pharmacol Ther, 2004. 76: 85-96.

2. Bertsche T, *et al.* Prevention of adverse drug reactions in intensive care patients by a personal intervention based on an electronic clinical decision support system. Intens Care Med, 2010. 36: 665-72.

3. Preissner S, *et al.* SuperCYP: a comprehensive database on Cytochrome P450 enzymes including a tool for analysis of CYP-drug interactions. Nucleic Acids Res, 2010. 38: D237-43.

Regulatory network analysis based on miRNA and mRNA microarray data

Background

Knowledge about gene relationships and the influence of miRNA on gene regulation is very important for in-depth analyses of the causes and the development of complex diseases. In addition to traditional statistical evaluations of microarray data which aim at the detection of mis-regulated genes among different probes, we focus on a network-based approach. Thus, we rather consider groups of interacting genes to be important for the detection of conditional dependencies between individually regulated genes.

Application

In cooperation with the department of Rheumatology at the Charité, Berlin, we are using approaches like Bayesian networks to derive a network of interacting genes. Based on mRNA and miRNA array datasets related to different purified immune cells, the aim of our research is the identification of unknown dependencies to extend the knowledge of regulatory networks. Obtained results will support the detection of disease stages of different rheumatic disorders, the development of tools for the diagnosis of other complex diseases, and the identification of potential drug targets.

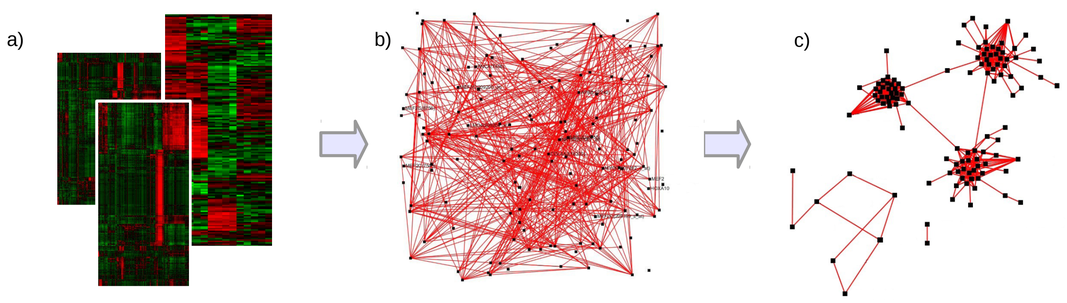

Fig. 3: Results of multiple microarray analyses (a) will be used to derive and augment a regulatory network (b) which will lead to clusters of conditionally dependent regulated genes (c).

In silico screening for new drugs interfering with protein-protein interactions

Motivation

Disease-relevant intracellular protein-protein interactions occurring at defined cellular sites have great potential as drug targets. They allow for highly-specific pharmacological interference with defined cellular functions. Drugs targeting such interactions are likely to act with fewer side effects than conventional medication influencing whole cell functions.

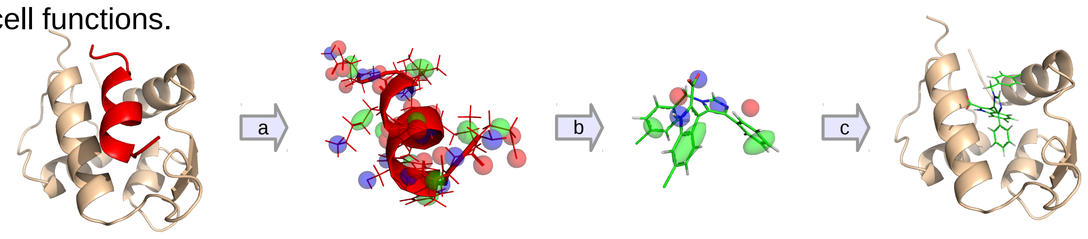

Fig. 4: Interfering with p53 and Mdm2 a) molecular features of a protein-protein interface are determined b) features are transferred to a pharmacophore that is subjected to *in silico* screening c) compounds matching the pharmacophore are docked onto the target protein.

Methods

Solved structures of protein-protein complexes give fundamental insights into protein function and molecular recognition. The number of structurally solved and diverse protein-protein complexes is limited (~1500)[1], but can be greatly extended by models generated by homology modeling and protein-protein docking[2]. The interface areas of complexes can be systematically screened for target pockets that are suitable for the binding of small molecules. Binding pockets will be transferred into a pharmacophore that will be applied for *in silico* screening of compound libraries. After a refinement step (compound docking), high-scoring drug candidates will be identified. The capability to inhibit specific protein-protein interactions will be tested by surface plasmon resonance (SPR) based technology.

References

1. Günther *et al.* ISEARCH--prediction of interactions using known interfaces. Proteins, 2007. 69: 839-44.
2. Günther *et al.* JAIL: a structure-based interface library for macromolecules. Nucleic Acids Res, 2009. 37: D38-41.

Pharmaceutical Bioinformatics
University of Freiburg

The working group of Pharmceutical Bioinformatics at the Institute for Pharmaceutical Sciences develops algorithms and software for pharmaceutical research. Our fields of research include the modeling of molecular interactions, prediction of biological effects of molecules, and identification of potential new drug agents. The working group is part of the University of Freiburg's Research Group Program of the Excellence Initiative of the federal and state governments.

Bundesministerium für Bildung und Forschung

UNI FREIBURG