3 months report, 28.10.2004

TArget Specific Scoring FUNction (tassfun)

Trypsin-like serine proteases are important targets for treating thrombo-embolic diseases, so the development of new specific inhibitors against trypsin, thrombin, urokinase type plasminogen activator (uPA), factor Xa (fXa) and related proteases is one of the major goals in structure-based drug design. The aim of this project is to apply and develop Comparative Binding Energy (COMBINE) analysis (1-2) to derive target specific scoring functions for predicting new potent inhibitors. The advantage of using COMBINE analysis instead of a classical three-dimensional quantitative structure activity relationship (3D-QSAR) is the possibility to correlate physico-chemical properties of protein-ligand-interactions based on crystal structures to experimental determined inhibition constants.

In the first three months of this project, the literature for a lead compound, investigated by Dr. B. A. Katz and colleagues (Axys Pharmaceuticals Corporation, South San Francisco) with dozens of published crystal structures (3-7), were inspected. The common structure element of these small molecule inhibitors is a 2-(2-phenol)indole or 2-(2-phenol)benzimidazole moiety, which is forming short and very short hydrogen bonds directly to the catalytic residues or mediated by water molecules.
Under supervision of Dr. Niklas Blomberg (GSI CompChem AZ Mölndal) the data of the published compounds were put into datasheets together with their K_i values according to thrombin, trypsin, uPA and fXa. In the group of Dr. Rebecca Wade (EML Research, Heidelberg) some of these ligands interacting with bovine trypsin were selected as a small test set for use in a first COMBINE analysis. Before the crystal structures, downloaded form the Protein Data Bank, could be used, preparation of the files was required. In view of running COMBINE analysis with large data sets of several hundreds of compounds, computer programs were written to prepare the files in a semi-automatic procedure. Subsequently, parameters for the ligands were assigned in AMBER8 by ANTECHAMBER and PARMCHK. Because of misassignments of atom types in the generated parameter files, some entries in the forcefields used had to be changed. After a short energy minimization in SANDER of AMBER8 with different restraints. These minimized structures were read into GOLPE 4.5 (8) for building Principal Component Analysis (PCA) and Partial Least Square (PLS) models. The first results are promising, so in next experiments, water molecules will take into consideration and the qualtiy of correlation between different components will be improved by selecting some variables.
The main work for the next few months will be running further analysis with larger testsets. With the help of these test sets the parameters of PCA and PLS will be optimized for reducing outliers and difficicult variables. For setting up larger test sets, the preparation of structure files should be done in a more automatic procedure by improving the written programs and by creating a database, which is linked to the existing datasheets.

COMBINE:
1. Ortiz, A.R. et al. (1995) Prediction of drug binding affinities by comparative binding energy analysis. *J. Med. Chem.* 38:2681-2691.
2. Wade, R.C. et al. (1998) Comparative Binding Energy Analysis. *Persp. Drug Disc. and Des.* 9:19-34**.**

Structures:
3. B.A. Katz *et al.* (2003) Elaborate manifold of short hydrogen bond arrays mediating binding of active site-directed serine protease inhibitors. *J. Mol. Biol.* 329:93-120.
4. B.A. Katz *et al.* (2001) A novel serine protease inhibition motif involving a multi-centered short hydrogen bonding network at the active site. *J. Mol. Biol.* 307:1451-1486.
5. E. Verner *et al.* (2001) Development of serine protease inhibitors displaying a multicentered

short (<2.3 A) hydrogen bond binding mode: Inhibitors of urokinase-type plasminogen activator and factor Xa. *J. Med. Chem.* 44:2753-2771.
6. B.A. Katz *et al.* (2001) Engineering inhibitors highly selective for the S1 sites of Ser190 trypsin-like serine protease drug targets. *Chemistry & Biology* 8:1107-1121.
7. B.A. Katz *et al.* (1998) Design of potent selective zinc-mediated serine protease inhibitors. *Nature* 391:608-612.

GOLPE 4.5:
8. Multivariate Infometric Analysis Srl., Viale dei Castagni 16, Perugia, Italy (1999).