From Substrate Specificity to Small Molecule Specificity

Birgit J. Waldner[1], Julian E. Fuchs[1], Michael Schauperl[1], Christian Kramer[1,†], Klaus R. Liedl[1]

1 Institute of General, Inorganic and Theoretical Chemistry, University of Innsbruck, Innrain 82, 6020 Innsbruck, Austria

† Present Address: F. Hoffmann- La Roche AG, Grenzacherstrasse 124, 4070 Basel, Switzerland

We present a way to use the rapidly growing amount of knowledge about protease peptide substrates as basis for a new virtual screening approach. We use the information on the specificity of the proteases and the physico-chemical features of the protease peptide substrates to find small molecule inhibitors. Modern database technology allows for easy access and sharing of the collected data on protease specificity and characteristics. The MEROPS database represents the biggest collection of known protease peptide substrates and is constantly improved and updated. The method represents a rapid and straightforward way of putting the MEROPS data on protease substrates to use for finding new small molecule inhibitors. We downloaded the peptide substrate sequences from the MEROPS database and used 3-4 substrate positions of each substrate to build the training set. Conversion of the 2D substrate sequences to 3D structures was carried out by mutating the residues in a template peptide for the corresponding protease taken from an X-ray structure of the protease-peptide complex. Considering the relative frequencies of substrate features, queries were created in ROCS. We show that the shape-based virtual screening gives good performance for four proteases, thrombin, factor Xa (fXa), factor VIIa (fVIIa) and caspase-3 (casp-3) with the DUD and DUD-E dataset. Thus, the method works for proteases with different specificity profiles as well as with different active site mechanisms and therefore should be applicable to any kind of protease.