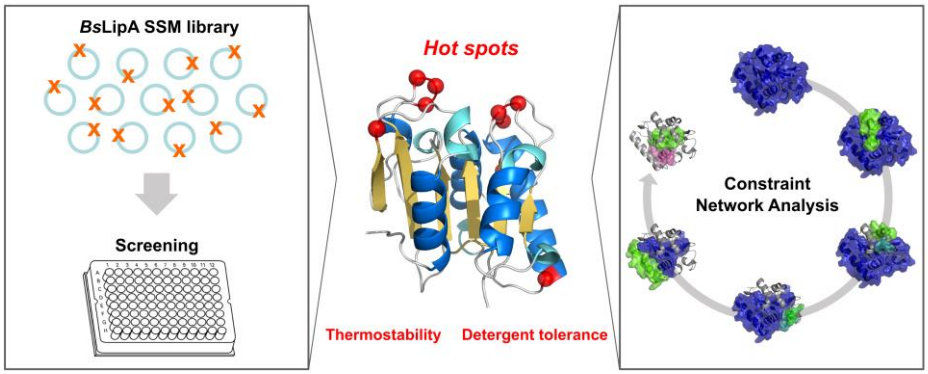
Large-scale analysis of protein thermostability and detergent tolerance

Christina Nutschel[1,2], Alexander Fulton[3,7], Olav Zimmermann[2], Ulrich Schwaneberg[4,5], Karl-Erich Jaeger[3], Holger Gohlke[1,2,6,]

[1] *John von Neumann Institute for Computing (NIC) and Institute for Complex Systems - Structural Biochemistry (ICS6), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany*
[2] *Juelich Supercomputing Centre (JSC), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany*
[3] *Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany*
[4] *Institute of Biotechnology, RWTH Aachen University, 52074 Aachen, Germany*
[5] *DWI-Leibniz-Institute for Interactive Materials, 52056 Aachen, Germany*
[6] *Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany*
[7] *Present Address: Novozymes A/S, Krogshoejvej 36, 2880 Bagsvaerd, Denmark*

Understanding how structure and activity of proteins are related to their (thermo-)stability [1-2] or tolerance against solvents [3] and detergents [4] is of utmost importance in protein engineering. However, until now, the role and impact of the location and type of substitution on protein properties has been predominantly studied from analyses of only a few protein variants or large-scale studies over data sets of many proteins, only contributing a few variants each. Furthermore, most analyses considered only one protein property at a time. Here, for the first time, we exhaustively characterize changes in thermostability *and* detergent tolerance for a complete site-saturation mutagenesis (SSM) library containing 3439 single variants of the *Bacillus subtilis* lipase A (*Bs*LipA). To establish a set of generally applicable guidelines regarding improved protein thermostability *and* / *or* detergent tolerance, we identified *hot spot* residues, i.e. those with a high likelihood to yield stabilized variants, and probed if they can be predicted based on protein structural characteristics or via rigidity theory-based Constraint Network Analysis (CNA) [5]. The main outcome is that CNA can precisely predict *hot spots* for rational protein design aiming at improved protein thermostability *and* / *or* detergent tolerance.



[1] P.C. Rathi, K.-E. Jaeger, H. Gohlke, *PLOS One* **2015**, *10*, e0130289.
[2] P.C. Rathi, A. Fulton, K.-E. Jaeger, H. Gohlke, *PLOS Comp. Biol*. **2016**, *12*, e1004754.
[3] V.J. Frauenkron-Machedjou, A. Fulton, L. Zhu, C. Anker, M. Bocola, K.-E. Jaeger, U. Schwaneberg, *ChemBioChem* **2015**, *16*, 937-945.
[4] A. Fulton, V.J. Frauenkron-Machedjou, P. Skoczinski, S. Wilhelm, L. Zhu, U. Schwaneberg, K.-E. Jaeger, *Chembiochem* **2015**, *16*, 930-936.
[5] S.M.A. Hermans, C. Pfleger, C. Nutschel, C. Hanke, C.A. Hanke, H. Gohlke, *WIREs Comput. Mol. Sci.* **2017**, e1311.