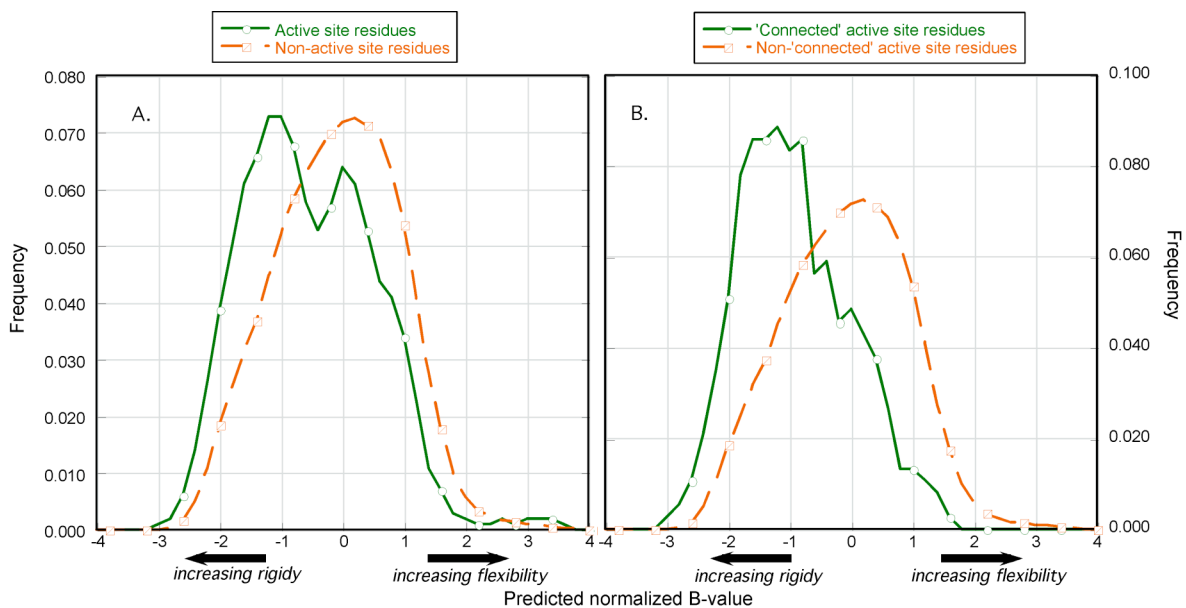


## Supplementary material

We validated our prediction method by trying to solve simple biological tasks. One example was the characterization of active sites in enzymes. The distributions of experimental normalized B-values differ between active and non-active site residues (active sites more rigid) (Yuan, 2003). Our prediction method almost replicated the experimental results (Fig. S1A). Pietrokovski and colleagues have shown that active sites in enzymes have a higher probability of being more connected internally than other residues (SARIG (Amitai, et al., 2004)). Here, we show that PROFbval can distinguish the residues that are in active sites and are strongly ‘connected’ to the network of residues from all other residues (Fig. S1B). The fact that the separation was found to be rather good when combining two properties that are only somewhat correlated (the Pearson correlation coefficient with the experimental and predicted normalized B-values was -0.52 and -0.49 respectively) may be the basis for developing an effective enzyme active site prediction method.

### Figures



**Fig S1: Predicted normalized B-values for active sites vs. non-active sites residues.** The distributions are given for residues of a set of 69 non-redundant enzymes (Yuan, 2003). (A) Active sites were identified by their PDB (Berman, et al., 2002) annotations as ‘SITE’ (marked in black). All residues without this annotation were classified as ‘non-active site’

(marked in gray). (B) Connected active site residues are the residues that are both annotated in PDB (Berman, et al., 2002) and predicted to be functional by SARIG (Amitai, et al., 2004): the separation is better when combining both methods than when using any single one of the two.

Amitai, G., Shemesh, A., Sitbon, E., Shklar, M., Netanel, D., Venger, I. and Pietrokovski, S. (2004) Network analysis of protein structures identifies functional residues, *J. Mol. Biol.*, **344**, 1135-1146.

Berman, H.M., Battistuz, T., Bhat, T.N., Bluhm, W.F., Bourne, P.E., Burkhardt, K., Feng, Z., Gilliland, G.L., Iype, L., Jain, S., Fagan, P., Marvin, J., Padilla, D., Ravichandran, V., Schneider, B., Thanki, N., Weissig, H., Westbrook, J.D. and Zardecki, C. (2002) The Protein Data Bank, *Acta Crystallogr D Biol Crystallogr*, **58**, 899-907.

Yuan, Z., Zhao, J., Wang, Z.X. (2003) Flexibility analysis of enzyme active sites by crystallographic temperature factors., *Protein Eng.*, **16**, 109-114.