New structure based substitution matrices and their applications.

Andreas Prlic, Francisco Domingues, Manfred Sippl
Center of Applied Molecular Engineering
University of Salzburg
Jakob-Haringerstr. 3
A-5020 Salzburg
Austria
email: andreas@came.sbg.ac.at

A set of 122 protein structure pairs, showing high structural and low sequence similarity was prepared. The ProSup [1] structure alignments of these pairs were used to derive an amino acid substitution matrix called Hellmax. Observed amino acid substitution frequencies were calculated from the structural equivalent residues in the ProSup alignment.

The matrix was applied in a jack-knife performance test for alignment accuracy. The accuracy of a sequence alignment was measured by calculating the percentage of correctly aligned residues relative to the structure alignment. Hellmax has one of the best performances comparing to other frequently used matrices.

A second matrix, Homomax, was derived out of the subset of the pairs, using homologous proteins only. It was applied in a scoring scheme. This score can be used to identify pairs of structures that share evolutionary relationship.

References

[1] Z.K. Feng and M.J. Sippl. Optimum superimposition of protein structures, ambiguities and implications. *Folding & Design*, 1:123–132, 1996.