AIST RESEARCH HOT LINE

UPDATES FROM THE CUTTING EDGE

(July - Sep. 2002)

The abstracts of the recent research information appeared on the Vol.2 No.7-No.9 of "AIST Today" are introduced and classified by research area.
For inquiry about the full article, please contact the author directly.

Life Science & Technology

Systematic Clustering of a Large Superfamily Based on Sequence, Structure and Function

- Application to the TIM Barrel Glycosidase Superfamily -

Nozomi NAGANO
Computational Biology Research Center
e-mail:
n.nagano@aist.go.jp
AIST Today Vol. 2, No. 7 (2002) 16

To untangle the distant relationships between enzymes in a large superfamily, comprehensive sequence and structure analyses can be performed. By applying the analyses to the TIM barrel glycosidase superfamily, including 30 distinct sequence families, we generated four structural clusters (S1, S2, S3, and S4), as in Figure 1. By this analysis, we could find the evolutionary relationships between two distinct enzyme groups, β-amylase and so-called 4/7 superfamily enzymes, including endoglucanase. Moreover, by combining the multiple-alignment of the tertiary structures, the local structure analyses and sequence analyses such as PSI-BLAST, we could cluster these enzymes hierarchically, and suggested that the two distinct structural clusters, S1 and S2 are distantly related, as are S3 and S4, which have the common chemical functions, *N*-acetylated substrates (Figure 2).

Furthermore, this combined method is expected to be applicable to other large superfamilies and even to folds, which are at higher level of protein structures.

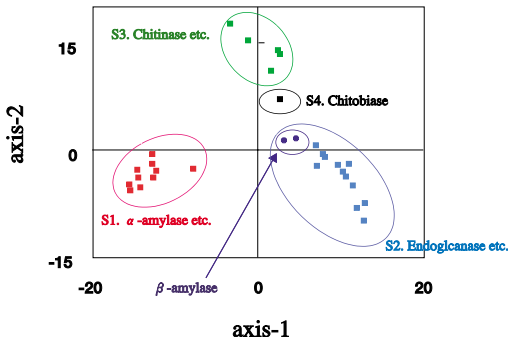
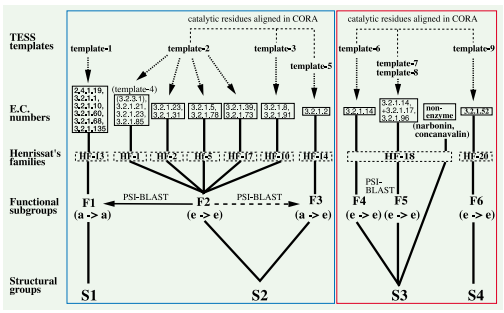


Fig.1 Clustering based on sequence identities: Multi-dimensional Scaling (MDS) plot of maximum sequence identities between 30 sequence families. The clusters are labeled with the names of the structural groups, α-amylase (S1), endoglucanase (S2), chitinase (S3) and chitobiase (S4). In the S2 group, the two β-amylases are indicated in circles, whilst the remainder is in squares.



N-Acatylated substrastes

Fig.2 Hierarchic clustering of glycosidase superfamily