

Identification and molecular modeling of a family 5 endocellulase from *Thermus caldophilus* GK24, a cellulolytic strain of *Thermus thermophilus*

Dooil Kim^{a,b,*}, Bo Hyun Park^a, Bo-Won Jung^a, Mi-kyung Kim^a, Suk-In Hong^{b,c}, Jungdon Bae^a, Dae-Sil Lee^{a,*}

^a*Systems Microbiology Research Center, KRIBB, Daejeon 305-806, Korea*

^b*Department of Biomicrosystem Technology, Korea University, Seoul 136-701, Korea*

^c*Department of Chemical and Biological Engineering, Korea University, Seoul 136-701, Korea*

The genome of *T. caldophilus* GK24 was recently sequenced and annotated as 14 contigs, equivalent to 2.3 mega basepairs (Mbp) of DNA. In the current study, we identified a unique 13.7 kbp DNA sequence, which included the endocellulase gene of *T. caldophilus* GK24, which did not appear to be present in the complete genomic sequence of the closely related species *T. thermophilus* HB27 and HB8. Congo-red staining revealed a unique phenotype of cellulose degradation by strain GK24 that was distinct from other closely related *Thermus* strains. The results showed that strain GK24 is an aerobic, thermophilic, cellulolytic eubacterium which belongs to the group *T. thermophilus*. In order to understand the mechanism of production of cellobiose in *T. caldophilus* GK24, a three-dimensional model of the endocellulase, *TcCel5A*, was generated based on known crystal structures. Using this model, we carried out a flexible cellotetraose docking study.

Keywords: Endocellulase; *Thermus caldophilus* GK24; Glycosyl hydrolase family 5; Homology modeling; Congo-red staining.