

Structural analysis of *Thermobacillus xylanilyticus* GH-51 arabinofuranosidase : from the sequence to the crystal structure

Hélène Débat^{*§}, Claire Dumon^{*}, Gabriel Pažs[€], Lars K. Skov[£], Caroline Rémond[€], Jette S. Kastrup[¶], Michael Gajhede[¶] Osman Mirza[¶] and Michael J. O'Donohue^{€*}

^{*}INRA-INSA UMR 792, Toulouse, France ; [§] Université Versailles Saint-Quentin, France ; [€]INRA UMR FARE 614, Reims, France ; [£]Novozymes A/S, Bagsværd, Denmark ; [¶]University of Copenhagen, Copenhagen, Denmark.

Hemicelluloses represent 20-35% of plant biomass and are the second most abundant source of renewable carbon after cellulose. Heteroxylans are the most common hemicelluloses and complex polysaccharides composed of a -D-xylose backbone that can be substituted with various sugars (arabinose, galactose and glucuronic acid), as well as with acetyl groups and phenolic acids. Due to the high complexity and structural variability of heteroxylans, enzymatic hydrolysis is achieved using a vast array of enzymes that include endoxylanases (EC 3.2.1.1) and debranching enzymes such as -L-arabinofuranosidases (Abf) (EC 3.2.1.55). According to the glycoside-hydrolase classification system (CAZy database), GH-51 and GH-54 are the two major families containing almost exclusively Abfs. Family GH 51 currently contains 128 sequences of bacterial origin with very variable sequence similarity. GH-51 α -L-Arabinofuranosidase of *Thermobacillus xylanilyticus* (*Tx*-Abf) is a retaining enzyme that catalyzes the hydrolysis of glycosidic bonds through a double displacement mechanism with Glu176 as the acid/base and Glu298 as the nucleophile.

Here we present the crystal structure of the native *Tx*-Abf and an inactive mutant Glu176Gln in complex with a branched pentasaccharide, a fragment of its natural substrate xylan. The overall structure shows the two characteristic GH-51 domains: a catalytic domain that is folded into a ($/$)₈-barrel and a C-terminal domain that displays jelly-roll architecture. The pentasaccharide is bound in a groove on the surface of the enzyme, with the arabinosyl entering a tight pocket harbouring the catalytic dyad. Interestingly, the protein bound to its substrate shows an open and closed conformation of the β 2 α 2 loop. However, superposition of the structures of *Geobacillus stearothermophilus* *Gs*-Abf and *Clostridium thermocellum* *Ct*-Abf onto the open conformation of *Tx*-Abf reveals that *Gs*-Abf and *Ct*-Abf probably cannot adopt this open conformation. This β 2 α 2 loop movement in *Tx*-Abf has been emphasized by molecular dynamic (MD) calculation. MD also revealed a movement of β 6 α 6 loop in *Tx*-Abf, which is longer than in *Gs*-Abf and *Ct*-Abf. This difference could contribute to *Tx*-Abf substrate specificity.

The results of this structural study has been used to guide site-directed mutagenesis experiments designed that aimed at the identification of the key substrate specificity determinants.