AEM Accepts, published online ahead of print on 15 July 2011
Appl. Environ. Microbiol. doi:10.1128/AEM.05532-11

1 **Title**:

2 Determination of the catalytic base in family 48 glycosyl hydrolases.

3 **Running title**:

4 Cel48 catalytic base.

5 **Authors**:

6 Maxim Kostylev and David B. Wilson*

7 **Address**:

8 Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853

9 **Corresponding author information**:

10 Cornell University, 458 Biotechnology Building, Ithaca, NY 14853; Tel: 607-255-5706; Fax:

11 607-255-2428; E-mail: dbw3@cornell.edu

Abstract: The catalytic base in family 48 glycosyl hydrolases has not been previously established experimentally. Based on structural and modeling data published to date, we used site directed mutagenesis and azide rescue activity assays to show definitively that the catalytic base in *Thermobifida fusca* Cel48A is aspartic acid 225. Of the tested mutants only Cel48A D225E retained partial activity on soluble and insoluble substrate. In azide rescue experiments only D225G, the smallest residue tested, showed an increase in activity with added azide.

Cellulosic biomass has the potential to form a significant component of a sustainable energy system (11). Carbohydrates obtained from the breakdown of cellulose can be utilized by various microorganisms to produce alcohols or combustible hydrocarbons for use as transportation fuels. One of the limiting steps in the development of an efficient and economically feasible system to produce biomass fuels is the enzymatic hydrolysis of cellulose (6, 16). It is important to generate an extensive knowledge base of the cellulases found in nature to develop better enzymatic cocktails for improved cellulose digestion in the industrial production of biofuels.

Family 48 cellulases form an important part of the bacterial cellulose degrading system. Despite their low specific activity on crystalline cellulose, they have high synergistic activity with other cellulases (1, 3, 7). The three most studied members of the family are Cel48F from *Clostridium cellulolyticum*, Cel48S from *Clostridium thermocellum*, and Cel48A from *Thermobifida fusca.* Cel48F and Cel48S are major components in the cellulose-degrading multienzyme complexes called cellulosomes (8, 17, 19), while Cel48A forms 34% of the total extracellular cellulases produced by *T. fusca* (15). The catalytic domain structures of Cel48F (8-10) and Cel48S (2) have been solved with various inhibitors and oligosaccharides, revealing that their active site is located at a junction between a tunnel and a cleft. The tunnel contains binding sites for the incoming cellulose chain and the cleft is the product binding site. The structure of Cel48A has not been solved, but a good model is available using the structure of Cel48F, with which Cel48A has 56% identity (3). Cel48 enzymes are believed to be processive, with cellobiose as their major product (3, 4, 12).

Family 48 cellulases employ concerted one-step general acid/base catalysis to hydrolyze the β-1,4 glycosidic bonds in cellulose (14). In this mechanism the catalytic acid serves as the

proton donor for the oxygen (O4) in the glycosidic bond undergoing hydrolysis, while the catalytic base activates the catalytic water molecule, which attacks the anomeric carbon (C1), resulting in the inversion of the original stereochemistry. Crystallographic data have made it possible to establish the catalytic acid in the Cel48 family. It is a glutamic acid residue (E55 in Cel48F, E87 in Cel48S, and E52 in Cel48A) located in the active site, which makes favorable hydrogen bonding to the O4 atom of the sugar unit in the +1 subsite (2, 8). The catalytic base, on the other hand, has not been experimentally established to date because there is no defined sugar residue in the -1 subsite in any of the obtained structures. Combined structural analyses (8, 9) and molecular simulations (13) suggest an aspartic acid residue (D230 in Cel48F, D255 in Cel48S, and D225 in Cel48A) to be the most likely candidate to carry out the catalytic base function in Cel48 enzymes. Here, we present strong experimental evidence that D225 is indeed the catalytic base in *T. fusca* Cel48A.

The catalytic domain of *T. fusca* Cel48A has been previously cloned and expressed in *E. coli* in our group (3). Using the QuickChange II-XL Site-directed Mutagenesis Kit (Agilent Technologies) we prepared four mutant versions of the protein: D225N, D225E, D225A, and D225G. The reasons for the chosen mutations are as follows. Asparagine is the closest structural analog of aspartate that is unable to catalyze hydrolysis. Glutamate is the closest structural and functional analog of aspartate that might retain some catalytic activity. Alanine and glycine are the two smallest amino acids, which would provide more flexibility within the active site, needed for the azide rescue experiments described below. All mutants were expressed and purified as previously described for the native enzyme (3). Activity of the enzymes was measured on insoluble swollen cellulose (SC) and on 1,4-β-D-celohexaose (G6), the longest soluble cellooligosaccharide. SC was prepared by treating crystalline cellulose

(Sigmacell® 101, Sigma-Aldrich) with phosphoric acid, as described in (18), to reduce its crystallinity for easier digestion. Pure G6 was obtained from Megazyme International. All reactions were prepared by combining Cel48A with 1.7 mg/ml SC or 17 µM G6 in 600 µl of 50 mM sodium acetate buffer, pH 5.5. Enzyme concentration was 1 µM on SC and 0.17 µM on G6. Reactions were incubated at 50 °C for four (G6) or 18 (SC) hours and soluble products were analyzed via HPLC using Aminex HPX-87P column (Bio-Rad) and water as a mobile phase. As expected, the main products observed were cellobiose and cellotriose. Trace amounts of cellotetraose were also present in some cases. For SC percent digestion was calculated from the total measure amount of glucose units released from the bulk substrate. For G6 percent digestion was calculated from the amount of G6 left in the reaction buffer at the end of the reaction. Presented data are averages and standard deviation of triplicate measurements.

Figure 1 shows the activity of the wild type (WT) and mutant enzymes on SC and G6. On both soluble and insoluble substrates the substitution of aspartate with glutamate (D225E) resulted in significant partial retention of WT activity. All three other mutants showed only residual activity, which is consistent with the presumed catalytic base role of D225. It is interesting to note that the activity of D225E relative to WT was much lower on G6 than on SC. For cellulases acting on insoluble substrates it is generally considered that access to substrate, not hydrolysis, is the rate limiting step. This is evidenced by the much higher cellulase activities on soluble cellooligosaccharides than on insoluble cellulose. As a mutation of the catalytic base, D225E should only increase the activation barrier of the hydrolysis step without an effect on the ability of Cel48A to bind cellulose. Hence, it is likely that the effect of the glutamate mutation is partially masked on SC by the high activation energy required to bind an individual cellulose

chain in the active site, but is fully revealed on a soluble substrate, for which hydrolysis is the rate limiting step.

To further confirm that Cel48A D225 is the catalytic base and not just an important residue in the catalytic site, we tested the ability of azide to rescue the activity of the mutated versions of the protein (Figure 2) (5). For this test, activity assays were carried out as described above, but with the indicated amount of sodium azide added to the reactions. Of the four mutant enzymes only the D225G protein showed an increase in activity on SC with added azide. The position of the activated water molecule that attacks the anomeric carbon during hydrolysis is critical. We thus presume that the reason for restored activity in the D225G mutant only is due to the greater flexibility of its active site in comparison to the other mutants, which all have more sterically-hindering functional groups. The ability of azide to partially restore enzyme activity when D225 is mutated is consistent with its presumed role as the catalytic base in Cel48A.

In summary, using site-directed mutagenesis and azide rescue hydrolysis assays we have shown that the *T. fusca* Cel48A catalytic base is D225. This residue is equivalent to D230 of *C. cellulolyticum* Cel48F and D255 of *C. thermocellum* Cel48S. To our knowledge, this is the first time that the catalytic base in family 48 glycosyl hydrolases has been experimentally established.

This work was supported by the BioEnergy Science Center, a U.S. Department of Energy research center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

105 **References**

106 1. **Berger, E., D. Zhang, V. V. Zverlov, and W. H. Schwarz.** 2007. Two noncellulosomal
107 cellulases of Clostridium thermocellum, Cel9I and Cel48Y, hydrolyse crystalline
108 cellulose synergistically. FEMS Microbiol Lett **268:**194-201.
109 2. **Guimaraes, B. G., H. Souchon, B. L. Lytle, J. H. David Wu, and P. M. Alzari.** 2002.
110 The crystal structure and catalytic mechanism of cellobiohydrolase CelS, the major
111 enzymatic component of the Clostridium thermocellum Cellulosome. J Mol Biol
112 **320:**587-596.
113 3. **Irwin, D. C., S. Zhang, and D. B. Wilson.** 2000. Cloning, expression and
114 characterization of a family 48 exocellulase, Cel48A, from Thermobifida fusca. Eur J
115 Biochem **267:**4988-4997.
116 4. **Kruus, K., W. K. Wang, J. Ching, and J. H. Wu.** 1995. Exoglucanase activities of the
117 recombinant Clostridium thermocellum CelS, a major cellulosome component. J
118 Bacteriol **177:**1641-1644.
119 5. **MacLeod, A. M., T. Lindhorst, S. G. Withers, and R. A. Warren.** 1994. The acid/base
120 catalyst in the exoglucanase/xylanase from Cellulomonas fimi is glutamic acid 127:
121 evidence from detailed kinetic studies of mutants. Biochemistry **33:**6371-6376.
122 6. **Merino, S. T., and J. Cherry.** 2007. Progress and challenges in enzyme development for
123 biomass utilization. Adv Biochem Eng Biotechnol **108:**95-120.
124 7. **Olson, D. G., S. A. Tripathi, R. J. Giannone, J. Lo, N. C. Caiazza, D. A. Hogsett, R.**
125 **L. Hettich, A. M. Guss, G. Dubrovsky, and L. R. Lynd.** 2010. Deletion of the Cel48S
126 cellulase from Clostridium thermocellum. Proc Natl Acad Sci U S A **107:**17727-17732.

8. **Parsiegla, G., M. Juy, C. Reverbel-Leroy, C. Tardif, J. P. Belaich, H. Driguez, and R. Haser.** 1998. The crystal structure of the processive endocellulase CelF of Clostridium cellulolyticum in complex with a thiooligosaccharide inhibitor at 2.0 A resolution. Embo J **17:**5551-5562.

9. **Parsiegla, G., C. Reverbel-Leroy, C. Tardif, J. P. Belaich, H. Driguez, and R. Haser.** 2000. Crystal structures of the cellulase Cel48F in complex with inhibitors and substrates give insights into its processive action. Biochemistry **39:**11238-11246.

10. **Parsiegla, G., C. Reverbel, C. Tardif, H. Driguez, and R. Haser.** 2008. Structures of mutants of cellulase Cel48F of Clostridium cellulolyticum in complex with long hemithiocellooligosaccharides give rise to a new view of the substrate pathway during processive action. J Mol Biol **375:**499-510.

11. **Perlack, R. D., L. L. Wright, A. F. Turhollow, R. Graham, L., B. Stokes, J., and D. C. Erbach.** 2005. Biomass as Feedstock for a Bioenergy and Bioproducts Industry: the Technical Feasibility of a Billion-ton Annual Supply. Oak Ridge National Laboratory.

12. **Reverbel-Leroy, C., S. Pages, A. Belaich, J. P. Belaich, and C. Tardif.** 1997. The processive endocellulase CelF, a major component of the Clostridium cellulolyticum cellulosome: purification and characterization of the recombinant form. J Bacteriol **179:**46-52.

13. **Saharay, M., H. Guo, and J. C. Smith.** 2010. Catalytic mechanism of cellulose degradation by a cellobiohydrolase, CelS. PLoS One **5:**e12947.

14. **Shen, H., P. Tomme, A. Meinke, N. R. Gilkes, D. G. Kilburn, R. A. Warren, and R. C. Miller, Jr.** 1994. Stereochemical course of hydrolysis catalysed by Cellulomonas fimi

CenE, a member of a new family of beta-1,4-glucanases. Biochem Biophys Res Commun **199:**1223-1228.

15. **Spiridonov, N. A., and D. B. Wilson.** 1998. Regulation of biosynthesis of individual cellulases in Thermomonospora fusca. J Bacteriol **180:**3529-3532.

16. **Wen, F., N. U. Nair, and H. Zhao.** 2009. Protein engineering in designing tailored enzymes and microorganisms for biofuels production. Curr Opin Biotechnol **20:**412-419.

17. **Wu, J. H. D., W. H. Orme-Johnson, and A. L. Demain.** 1988. Two components of an extracellular protein aggregate of Clostridium thermocellum together degrade crystalline cellulose. Biochemistry **27:**1703-1709.

18. **Zhang, Y. H., J. Cui, L. R. Lynd, and L. R. Kuang.** 2006. A transition from cellulose swelling to cellulose dissolution by o-phosphoric acid: evidence from enzymatic hydrolysis and supramolecular structure. Biomacromolecules **7:**644-648.

19. **Zverlov, V. V., J. Kellermann, and W. H. Schwarz.** 2005. Functional subgenomics of Clostridium thermocellum cellulosomal genes: identification of the major catalytic components in the extracellular complex and detection of three new enzymes. Proteomics **5:**3646-3653.

166 **Figure 1**: Activity of the wild type Cel48A catalytic domain and Cel48A D225 mutants on G6
167 (soluble) and SC (insoluble); n = 3.

168 **Figure 2**: Effect of azide addition on the activity of wild type Cel48A catalytic domain and
169 Cel48A D225 mutants on SC; n = 3

170