

Effect of Single Active-Site Cleft Mutation on Product Specificity in a Thermostable Bacterial Cellulase

TAUNA R. RIGNALL, JOHN O. BAKER,*
SUZANNE L. MCCARTER, WILLIAM S. ADNEY,
TODD B. VINZANT, STEPHEN R. DECKER,
AND MICHAEL E. HIMMEL

*Biotechnology for Fuels and Chemicals Division,
National Bioenergy Center, National Renewable Energy Laboratory,
1617 Cole Boulevard, Golden, CO 80401, E-mail: john_baker@nrel.gov*

Abstract

Mutation of a single active-site cleft tyrosyl residue to a glycyl residue significantly changes the mixture of products released from phosphoric acid-swollen cellulose (PSC) by EIcd, the catalytic domain of the endoglucanase-I from *Acidothermus cellulolyticus*. The percentage of glucose in the product stream is almost 40% greater for the Y245G mutant (and for an additional double mutant, Y245G/Q204A) than for the wild type enzyme. Comparisons of results for digestion PSC and of pretreated yellow poplar suggest that the observed shifts in product specificity are connected to the hydrolysis of a more easily digestible fraction of both substrates. A model is presented that relates the changes in product specificity to a mutation-driven shift in indexing of the polymeric substrate along the extended binding-site cleft.

Index Entries: *Acidothermus cellulolyticus*; endoglucanase; product specificity; saccharification; cellulose.

Introduction

Acidothermus cellulolyticus endoglucanase-I (EI, Cel5A) is a highly thermostable cellulase originating from the aforementioned hot springs organism (1). Both the complete protein molecule and its genetically truncated catalytic domain (rEIcd) rapidly reduce the viscosity of carboxymethylcellulose solutions (2), and both are highly synergistic with *Trichoderma reesei*

*Author to whom all correspondence and reprint requests should be addressed.

(*T. reesei*) cellobiohydrolase-I (CBHI) in the hydrolysis of cellulosic materials ranging from microcrystalline cellulose to dilute acid-pretreated yellow poplar (PYP) (3).

A high-resolution, X-ray crystallographic three-dimensional structure has been obtained for the catalytic domain of EI (4). Analysis of this structure suggested that removal of the aromatic/hydrophobic binding platform afforded by the binding-site cleft residue Tyr-245 might relieve inhibition by the product cellobiose. On construction and testing of the Y245G mutant of EIcd, this suggestion was in fact confirmed. These results and the relation of the relief of inhibition to the structural changes observed in the mutant are being published separately (5). During activity testing of the Y245G mutant, it was observed that in addition to the predicted relief of product inhibition, the mutant enzyme tended to produce a higher proportion of glucose (relative to cellobiose) than did the wild type enzyme, when acting alone on phosphoric acid-swollen cellulose (PSC). This article presents our first exploration of this apparent change in product specificity, including the construction and testing of an additional EI species having a second mutation in the same (+2) binding subsite where the original Y245G mutation is located (4,5).

Materials and Methods

Enzyme Production and Purification

rEIcd was purified either from the culture filtrate of an *Streptomyces lividans* strain containing the appropriate insert as detailed in Sakon et al. (4), or from a commercial cellulase product produced by a *T. reesei* strain. Mutants of EIcd were engineered and purified as detailed in McCarter et al. (6). *T. reesei* CBHI was purified from a commercial cellulase preparation following the general protocol described by Shoemaker et al. (7).

Assay Substrates

PSC was prepared from Avicel Type PH 101 (FMC, Philadelphia, PA). Five grams of Avicel was first moistened with deionized water, and then 150 mL of 85% phosphoric acid was added slowly over a period of 1 h with gentle stirring, with the slurry being maintained in an ice bath. After the addition of 100 mL of cold acetone, the slurry was centrifuged for 10 min at 5000 g. The pellet was washed three times by resuspension and centrifugation in deionized water and then thoroughly dispersed by blending with an IKA-Werk Ultra-Turrax (Tekmar, Cincinnati, OH). The swollen cellulose was stored at 4°C under acidic conditions (pH < 2.0).

For assays under the particular conditions of this series, a 1.25% (w/v) slurry was made and exchanged into assay buffer by successive centrifugations and resuspensions in 20 mM acetate, pH 5.0, 0.02% sodium azide, until the pH of the supernatant measured within 0.01 unit of 5.0. Dilute-acid-pretreated yellow poplar (PYP) was prepared as previously described (8,9).

Enzyme Assays

Conventional Closed-System End-Point Assays

Assays were carried out in sealed 1.8-mL HPLC sample vials at 40°C and pH 5.0 in 20 mM acetate. Each 1.0-mL reaction contained 5 mg of acid-swollen cellulose and 28 µg (approx 70 nmol) of either native EIcd or Y245G mutant. Enzyme or enzyme mixture (total volume of 0.6 mL) was prepared beforehand in each reaction vial; reactions were then initiated by adding 0.4 mL of 1.25% (w/v) PSC in reaction buffer (pipetted from a well-stirred slurry). The reaction vials were then sealed with aluminum crimp-caps (silicone septum) and incubated for 4 h in a water bath at 40°C, with constant mixing by inversion at 10 rpm (axis of rotation 35° from horizontal, with long dimensions of vials tangential to the direction of rotation). At the end of the incubation period, the vials were removed from the water bath, cooled in an ice bath, and the reaction mixture was quickly removed by syringe needle through the septum and filtered (Acrodisc-13, 0.2-µm pore size, Pall-Gelman, Ann Arbor, MI) to remove the remaining substrate. Glucose and cellobiose were then quantified by chromatography on a HPX-87H column (Bio-Rad, Hercules, CA) at 65°C with 0.01 N sulfuric acid (0.6 mL/min) as the mobile phase in an automated chromatograph with refractive index detection (models 1100 and 1047, respectively; Agilent, Wilmington, DE).

Diafiltration Saccharification Assays

In the diafiltration saccharification assay (DSA) previously developed at National Renewable Energy Laboratory (9), cellulase enzymes carry out substantial conversion of an insoluble cellulosic substrate in a continuously buffer-swept, stirred-tank membrane reactor. The solubilized saccharification products are continuously removed by the buffer flux through an ultrafiltration membrane (Biomax-5; Millipore, Bedford, MA), while the insoluble substrate and macromolecular enzymes are retained in the reaction chamber. The course of the reaction is then followed by HPLC analysis of the soluble products in the effluent stream. All DSAs were carried out at 38°C in 20 mM, pH 5.0 sodium acetate buffer (plus 0.02% sodium azide to prevent microbial growth). Substrate loading for each assay was 104 mg (dry wt) of PYP, which amounted to a loading of 4.2% in terms of biomass and 3.2% in terms of cellulose. The substrate was finely ground, with the bulk of the material consisting of particles between 10 and 500 µ in maximum dimension.

For DSAs, the standard loading of the EIcd endoglucanase species involved was 1.54 nmol (61.9 µg) per assay (2.2-mL reaction volume). For assays with PYP as substrate, the substrate loading was 96 mg (dry wt) of biomass per assay, with a biomass cellulose content of 0.576 g/g resulting in a loading ratio of 27.7 nmol (1.12 mg) / g of biomass cellulose. For DSAs of endoglucanases vs PSC, substrate loading was 7.95 mg of cellulose per assay, for a loading ratio of 193 nmol (7.79 mg) of enzyme/g of cellulose. In all DSAs involving *T. reesei* CBHI, the CBHI loading was 534 nmol

(1.545 mg) per assay. For all DSAs, cellobiose and glucose contents of effluent fractions were determined by HPLC as described earlier for filtered closed-system assay samples.

Statistical Evaluation of Kinetically Determined Values

The statistical significances of differences observed between the means of triplicate determinations were evaluated by means of a one-tailed *t*-test (10), using the percentage points of the *t*-distribution tabulated by Pearson and Hartley (11).

Results and Discussion

A previous X-ray crystallographic study of the complex of the catalytic domain of *A. cellulolyticus* EI with cellotetraose (4) has demonstrated that this enzyme has an extended binding site for the glucosyl residues of a cellulose chain. Hydrophobic bonding between as many as six aromatic amino acid side chains and the hydrophobic faces of cellulose glucosyl residues has been implicated in substrate binding (Fig. 1), with four such residues specifically identified as interacting with specific glucosyl residues of the bound cellotetraose. The two mutant proteins described in this study both involve mutations of residues associated with the binding site for Glc1, which is the reducing-terminal residue of the bound cellotetraose (4) and the second residue away from the cleavage site (represented by Glu-162 in Fig. 1). As shown by the ball-and-stick representation of Fig. 2, the side chain of Tyr-245 serves as the hydrophobic binding platform for Glc1. Removal of this platform has been shown to result in a substantial decrease in binding strength at this position, as shown by an increase in the K_i value for cellobiose (5), which might be expected to bind strongly at the Glc1 and Glc2 positions (4). The ϵ -oxygen of Gln-204, while somewhat distant in the native enzyme for strong hydrogen bonding to O2 of Glc1, present the possibility of collapsing toward Glc1 in the perhaps loosened structure of the Y245G mutant. Increased hydrogen bonding between O2-Glc1 and the Gln-204 oxygens might therefore partially offset the decrease in binding strength at this position resulting from the removal of the hydrophobic platform. Because of this possibility, and because it was reasonable to relate the initially observed changes in product mixture to changes in patterns of binding of polymeric substrate, we decided to investigate the possibility that performing an additional mutation (Gln-204 to Ala-204, or Q204A) on Y245G might increase the effect already shown by the Y245G mutation upon the product mixture for EIdc.

The closed-system assay results shown in Table 1 clearly demonstrate that the product mixture for the Y245G mutant is shifted toward higher glucose content. The small relative standard deviations (SDs) for the means of the quadruplicate determinations of the percentages of glucose, plus the fact that the difference between the means of the mutant/native ratios for percentages of glucose is significant well beyond the $p < 0.0001$ level (one-tailed *t*-test) (10,11), confirm that the observed effect is real.

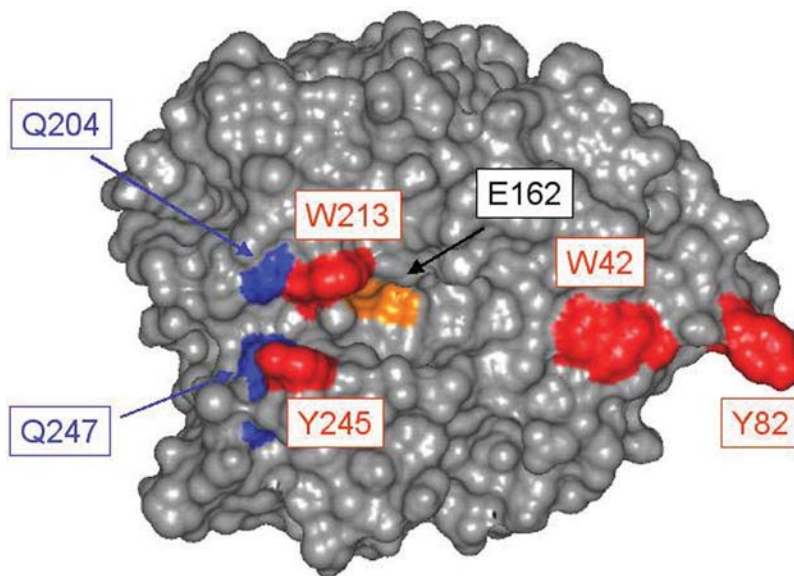


Fig. 1. Space-filling model of *A. cellulolyticus* EIcd. The highlighted W42, Y82, W213, and Y245 residues are hydrophobic binding platforms that help guide the substrate toward the active site, E162. The Q204 and Q247 residues provide hydrogen-bonding stabilization.

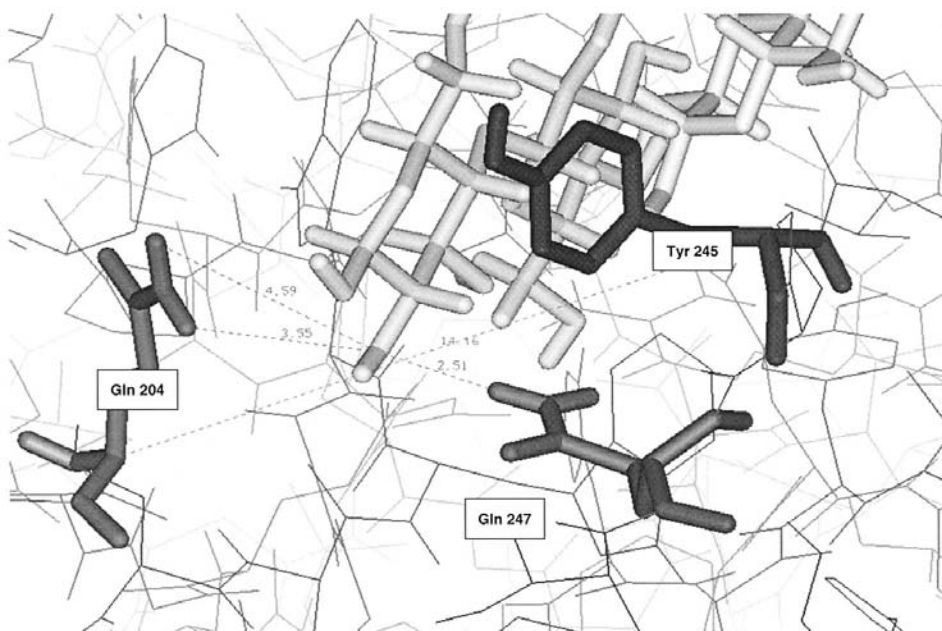


Fig. 2. Closer view of the leaving-group binding site that displays hydrogen bonding, based on structure determined in ref. 4.

Table 1
Product Percentage of Glucose for Wild Type and
Mutant Endoglucanases Acting on PSC,
as Measured by Closed-System Assays and DSA

	% <i>anhydro</i> -Glucose	Standard Deviations	Ratio
Closed System PSC Assays			
• Wild type	13.1	± 0.004	Ratio _{Y245G/Wild} = 1.398
• Y245G	18.2	± 0.002	
DSA PSC Assays			
• Wild type	11.4	± 0.005	Ratio _{Y245G/Wild} = 1.404
• Y245G	16.0	± 0.005	
• Y245G/Q204A	17.2	± 0.001	Ratio _{O204A/Wild} = 1.509

The closed-system assays, run as they are for a fixed time, give us one snapshot of the digestion process. A more detailed picture covering more of the time course of the reaction can be obtained by means of DSA, which provides continuous collection of product samples by sweeping soluble sugars out of the reaction cell by means of a buffer flux through the ultra-filtration membrane that retains the enzymes and the insoluble substrate. Table 1 displays the results of DSA-monitored duplicate digestions of PSC by the same two enzymes reported on by the closed-system assay. DSA effluent fractions have been selected to correspond to the same extents of conversion achieved by the closed-system assays by the end of the fixed (4 h) digestion period used in those experiments. For digestion of equivalent fractions of the substrate, the proportions of glucose produced by the wild type and Y245G mutant enzymes are seen to be roughly in the same proportions as measured by the two methods. Also shown are the results of duplicate DSA experiments employing the double mutant Y245G/Q204A. Like the single (Y245G) mutant, the double mutant produces a higher percentage of glucose than does the wild type enzyme. The excellent agreement between the duplicate assays (as shown by the small SDs) indicates that the difference between percentages of glucose for the single and double mutants may be real in the sense of being statistically significant. However, the question remains, is the difference, small as it is, important. It is essentially beyond question, though, that both mutants produce a higher percentage of glucose than does the wild type. The addition of the second (Q204A) mutation may not provide a great increase over the effect of Y245G alone, but it does not remove any of the effect of the first mutation.

The bar graph of Fig. 3 displays, as a function of digestion time, the percentages of glucose in the DSA products for all three enzymes acting on PSC. In general, the percentages of glucose for the two mutants are seen to be higher than those for the wild type in the rapid, early portion of the reaction. The percentage of glucose for the wild type increases, however,

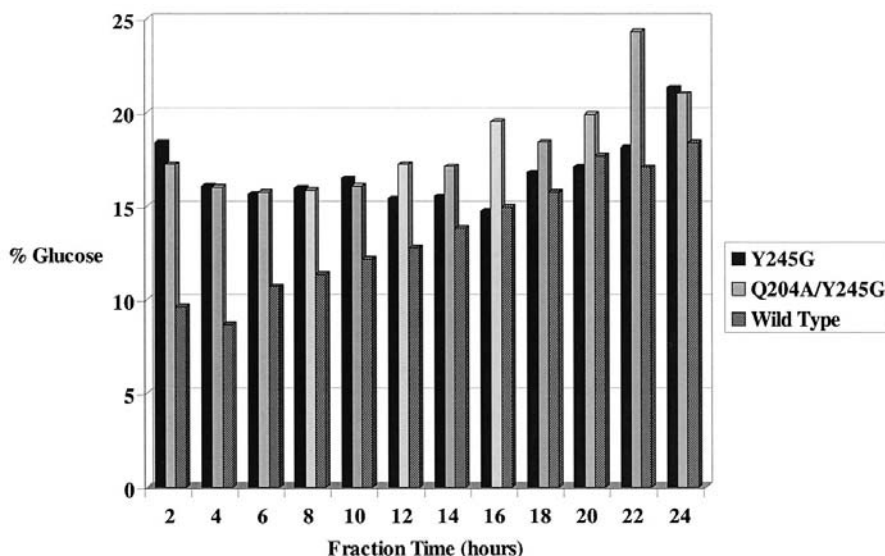


Fig. 3. Comparison of percentages of glucose by fraction time in DSA using PSC as substrate.

over the course of the reaction, and by 20–24 h has nearly equaled those for the two mutants. Figure 4 represents a corresponding bar graph of DSA percentage of glucose for these three enzymes acting on a different substrate, dilute acid PYP. Here a quite different trend is seen. As in the case of assays with PSC as substrate, the early fractions show both mutants producing higher percentages of glucose than that produced by the wild type. Over the first 24 h, however, there is not nearly as clear an indication that the wild type is catching up to the mutants in percentage of glucose; for the 21 and 24 h fractions the percentage of glucose produced by the wild type enzyme still only averages 0.63 times what the mutants display. (Percentages of glucose were calculated for all fractions yielding HPLC peaks for glucose that were large enough to yield reliable ratios. After 24 h, although the cellobiose peaks were still substantial, the glucose peaks were so small that the calculated percentages became so noisy as not to be useful.)

Some understanding of the data on percentage of glucose in Figs. 3 and 4 may be had from Figs. 5 and 6, which display standard DSA progress curves for these three enzymes acting against the two different substrates. These curves represent the overall conversion of cellulose to soluble sugar (glucose *plus* cellobiose), with all curves being averages of duplicates. The enzyme performances vs PSC (Fig. 5) are clearly biphasic, with a rapid phase being essentially complete by approx 20 h of digestion, followed by a phase of much slower hydrolysis. This biphasic nature is a commonly observed feature of endoglucanases acting alone on PSC and is probably related to the fact that most PSC preparations are only partially “swollen” and consist of a fraction that is relatively “amorphous,” mobile, and hydrated, plus another fraction that is still essentially crystalline. PYP (Fig. 5) is an even

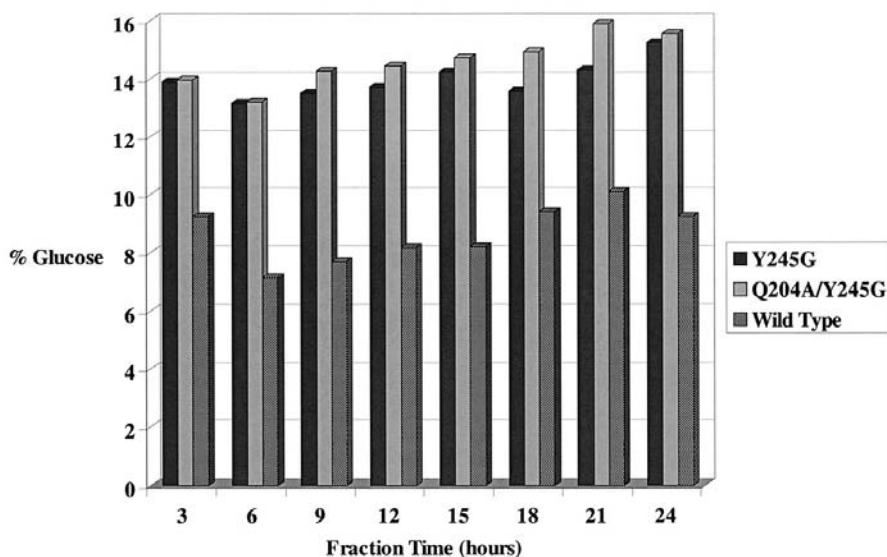


Fig. 4. Comparison percentages of glucose by fraction time in DSA using PYP as substrate.

more complex material that contains, in addition to cellulose, significant quantities of lignin (approx 40% by weight for the present material) in varying degrees of association with the cellulose and also retains a significant amount of the original wood structure. Accumulated kinetic evidence, such as that in Fig. 5, suggests that in this more complex material, the gradations of cellulose accessibility and digestibility may be more subtle than in PSC and may involve, effectively, a larger number of fractions defined in terms of accessibility.

According to this picture, the ability of both mutant endoglucanases to generate larger percentages of glucose is tied to the hydrolysis of a certain, more accessible fraction of each substrate. In the case of the DSA experiments with PSC as substrate, the disappearance of the mutant "advantage" in generating glucose (Fig. 3) roughly coincides with the apparent exhaustion of the "easily digestible" fraction of the material (Fig. 5). In the case of the more complex PYP, the combination of a more refractory substrate and a lighter enzyme loading have resulted in a lesser extent of hydrolysis (Fig. 6). In this case, the cellulosic material from which the mutants produce elevated levels of glucose has not been exhausted at 24 h into the digestion.

Given what is known about the structure of the wild type Elcd and the Y245G mutant (for both of which high-resolution X-ray crystallographic structures have been obtained) (4,5), one can visualize at least one possible explanation for the difference between the product specificities of the mutant and wild type enzymes. This scheme is illustrated in Fig. 7, in which the four glucosyl residue binding sites identified from the crystal structure of the Elcd-cellotetraose complex (4) are shown as white circles, which may or may not be occupied by dark, filled circles representing glucosyl residues of substrate or product. In the wild type Elcd, the polymeric substrate

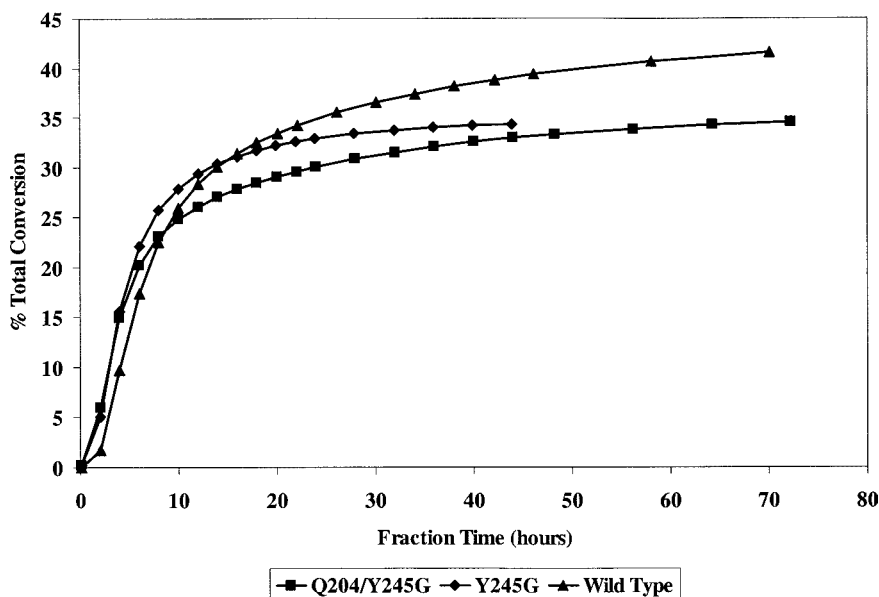


Fig. 5. DSA/PSC progress curves comparing the three endoglucanases.

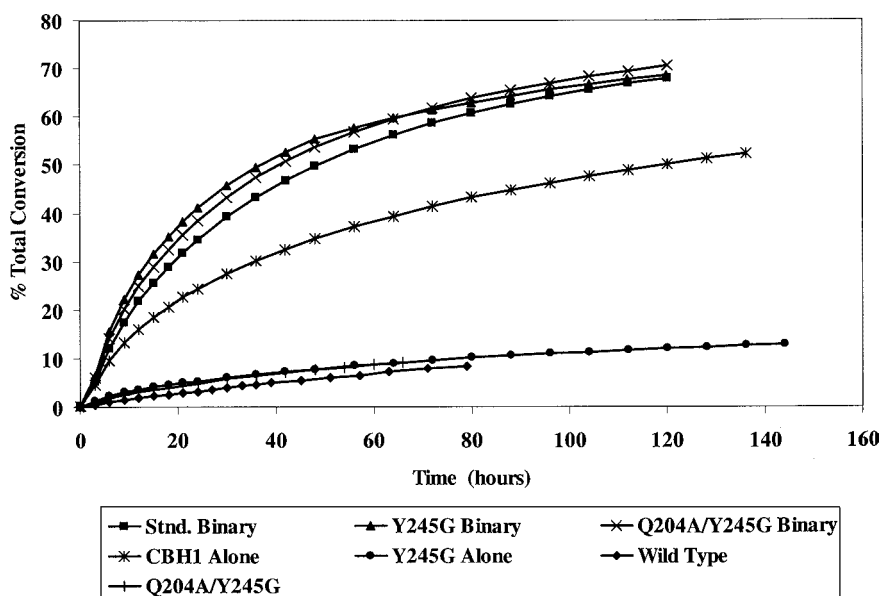


Fig. 6. DSA/PYP progress curves comparing the three endoglucanases, both assayed run alone and in binary assay. The binary assay utilizes a 1:19 ratio of the E1cd to CBHI.

is shown as having bound to the enzyme so as to occupy all of the glucosyl binding sites shown, and then having been cleaved into an $(n - 2)$ polymer plus a molecule of cellobiose. Because of the strong glucosyl-binding potential of subsites +1 and +2 (subsite +2 being the subsite on the extreme left,

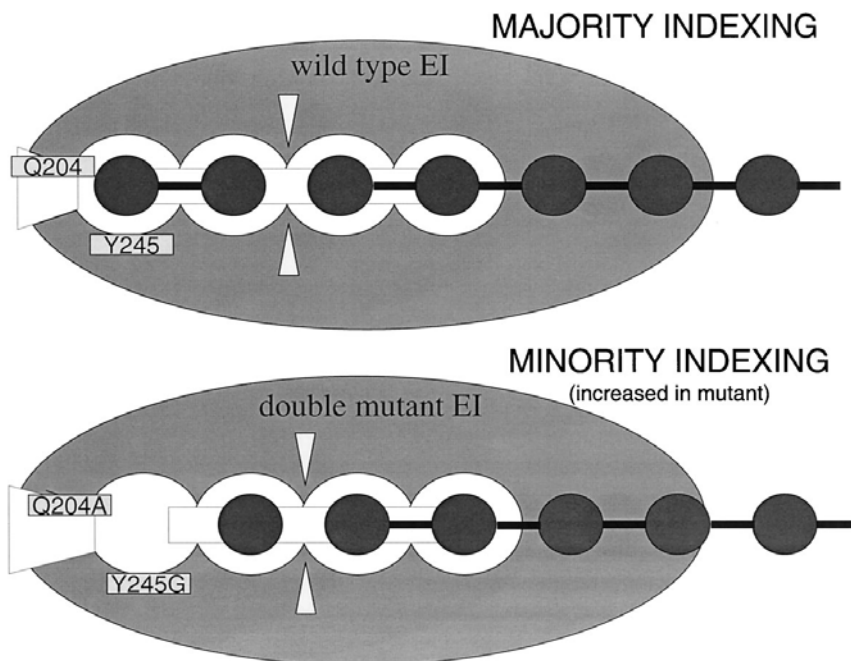


Fig. 7. Schematic representation of proposed explanation of mutant effect on composition of product mixture.

the one associated with Tyr-245 and Gln-204), the wild type enzyme has a strong preference for the pattern of binding that is shown as "majority indexing," which produces cellobiose as soluble product from the end of a relatively accessible "amorphous" portion of a cellulose chain. In both mutants, the removal of the binding platform provided by Tyr-245 has greatly decreased the glucosyl-binding capability of subsite +2 and, therefore, has increased the relative probability of binding such as that shown as "minority indexing." This binding alignment produces glucose as soluble product. The Y245G mutation, and perhaps to some extent the Q204A mutation, can be seen as allowing the endoglucanase to "cut closer to the end."

Of interest from an applications viewpoint is the additional observation that the Y245G mutant effect on glucose output expresses itself even in the assay of a binary mixture of enzymes in which the endoglucanase is a small minority component. When a 1:19 molar ratio of EICd (or Y245G) and *T. reesei* CBHI is assayed vs PYP at a loading of 0.028 mg of total cellulase /g of cellulose, and the percentage of glucose is measured at time points corresponding to the same (22%) extent of cellulose conversion for each mixture, the mean percentage of glucose in the DSA output stream is $5.34 \pm 0.09\%$ for the wild type enzyme, whereas it is $6.09 \pm 0.19\%$ for the Y245G mutant. The difference between the means, while small, appears to be real in that it is statistically significant at the level of $p < 0.0025$. These means result in a ratio (mutant/wild type) of 1.14, a value that is small compared

with the ratio of 1.4 found when the wild type and Y245G mutant endoglucanases were used alone against PSC; however, it is worth noting that the mutant effect does show up under such circumstances.

One must still be cautious about drawing mechanistic conclusions from these latter results (in the presence of CBH), because there is at least one other explanation (in addition to that of Fig. 7) for the mutant effect in the presence of CBH. Specifically, the mutant endoglucanase is simply more active than the wild type during the early stages of a DSA, cleaving more internal glycosidic bonds. This may simply provide the processive CBH with shorter cellulose chains on which to act, will by itself result in the production of higher percentages of glucose. The mutant effect is well established for the experiments in which the wild type and mutant endoglucanases act alone against cellulosic materials. We are then left with the finding that, by one mechanism or another, under a variety of reaction conditions, the Y245G mutation in EIcd increases the percentage of glucose in the product mixture.

Conclusion

Mutation of the active-site cleft residue Tyr-245 to glycine has been shown to change the product specificity of *A. cellulolyticus* EIcd, in that the percentage of glucose in the sugar stream released from cellulose is increased for the mutant, relative to that for the wild type enzyme. A simple preliminary scheme has been presented to present a possible explanation of this finding in terms of primary interaction between the enzyme and substrate.

Although the glucose streams for both wild type and mutant are small, and the increase is by a modest factor in the mutant, these results establish the principle that product specificity of cellulases may be modified by structure-guided, site-directed mutation. Given that many common industrial fermentative organisms are capable of fermenting glucose, but not cellobiose, this principle would appear to be worthy of further consideration.

Acknowledgment

This work was funded by the Biochemical Conversion Element of the Biofuels Program of the US Department of Energy.

References

1. Himmel, M. E., Adney, W. W., Grohmann, K., and Tucker, M. P. (1994), US patent no. 5,275,944.
2. Thomas, S. R., Laymon, R. W., Chou, Y-C., Tucker, M. P., Vinzant, T. B., Adney, W. S., Baker, J. O., Nieves, R. A., Mielenz, J. R., and Himmel, M. E. (1995), in *Enzymatic Degradation of Insoluble Polysaccharides*, ACS Series 618, Saddler, J. N. and Penner, M. H., eds., American Chemical Society, Washington, DC, pp.208–236.
3. Baker, J. O., Adney, W. S., Thomas, S. R., Nieves, R. A., Chou, Y-C., Vinzant, T. B., Tucker, M. P., Laymon, R. A., and Himmel, M. E. (1995), in *Enzymatic Degradation of*

- Insoluble Polysaccharides*, ACS Series 618, Saddler, J. N. and Penner, M. H., eds., American Chemical Society, Washington, DC, pp.113–141.
4. Sakon, J., Adney, W. S., Himmel, M. E., Thomas, S. R., and Karplus, A. P. (1996), *Biochemistry* **35**, 10,648–10,660.
 5. Baker, J. O., McCarley, J. R., Lovett, R., Yu, C-H., Adney, W. S., Rignall, T. R., Vinzant, T. B. Decker, S. R., Sakon, J., and Himmel, M. E. (2001), submitted.
 6. McCarter, S. L., Adney, W. S., Vinzant, T. B., Eddy, F. P., Decker, S. R., Baker, J. O., Sakon, J., and Himmel, M. E. (2002), *Appl. Biochem. Biotechnol.* **98–100**, 273–288.
 7. Shoemaker, S., Watt, K., Tsitowsky, G., and Cox, R. (1983), *Bio/Technology* **1**, 687–690.
 8. Vinzant, T. B., Ponfick, L., Nagle, N. J., Ehrman, C. I., Reynolds, J. B., and Himmel, M. E. (1994), *Appl. Biochem. Biotechnol.* **45/46**, 611–626.
 9. Baker, J. O., Vinzant, T. B., Ehrman, C. I., Adney, W. S., and Himmel, M. E. (1997), *Appl. Biochem. Biotechnol.* **63/65**, 585–595.
 10. Snedecor, G. W. and Cochran, W. G. (1967), *Statistical Methods*, 6th ed., Iowa State University Press, Ames, IA, pp.103–106.
 11. Pearson, E. S. and Hartley, H. O. (1976), in *Biometrika Tables for Statisticians*, vol.1, Pearson, E. S. and Hartley, H. O., eds., Biometrika Trust, University College, London, pp. 139,146.