Region-directed Mutagenesis of Residues Surrounding the Active Site Nucleophile in \(\beta\)-Glucosidase from \textit{Agrobacterium faecalis}*

Donald E. Trimbur\$, R. Antony J. Warren\$, and Stephen G. Withers\$

From the Departments of \$Chemistry and \$Microbiology, University of British Columbia, Vancouver, British Columbia V6T 1Y1, Canada

The active site nucleophile of the \(\beta\)-glucosidase of \textit{Agrobacterium faecalis} has recently been identified by the use of inhibitors. A combination of site-directed and \textit{in vitro} enzymatic mutagenesis was carried out on the \(\beta\)-glucosidase to probe the structure of the active site region. Forty-three point mutations were generated at 22 different residues in the region surrounding the active site nucleophile, Glu\(^{338}\). Only five positions were identified which affected enzyme activity indicating that only a few key residues are important to enzyme activity, thus the enzyme can tolerate a number of single residue changes and still function. The importance of Glu\(^{338}\) to enzymatic function has been confirmed and other residues important to enzyme structure or function have been identified.

\(\beta\)-Glucosidases (EC 3.2.1.21) are key enzymes in cellulose degradation, catalyzing the final step which converts cellulose to glucose (Shewale, 1982). We have been studying the \textit{Agrobacterium} \(\beta\)-glucosidase (Abg)\(^1\) because of its high specific activity and affinity for cellulose. Abg hydrolyzes glycosides with net retention of anomeric configuration (Day and Withers, 1986). The catalytic mechanism of \(\beta\)-glucosidases which retain anomeric configuration has been studied for many years. We have shown recently using the inhibitor, \(2',4'\)-dinitrophenyl 2-deoxy-2-fluoro-\(\beta\)-D-glucopyranoside that the mechanism involves the formation of a covalent intermediate. This inhibitor allowed the trapping of the covalent intermediate and allowed the identification of the active site nucleophile (Glu\(^{338}\) of Abg and the anomeric configuration of the glycosidic bond to the enzyme (Withers and Street, 1988; Withers et al., 1989).

Many \(\beta\)-glucosidases and \(\beta\)-galactosidases have been cloned and sequenced. These enzymes can be grouped into three families based on sequence alignments (Henrisat, 1991). Abg is a member of a large family containing enzymes from all three kingdoms. The catalytic mechanisms of these enzymes should be similar, and indeed alignments of the sequences of family members have shown that the active site nucleophile of Abg is conserved. Several other residues are also conserved in the region of the active site.

We have used the inhibitor studies and sequence alignments of related enzymes to target a region around the active site nucleophile for mutagenesis. We have generated a number of point mutations in this region and have measured their effect on activity. We have confirmed the importance of Glu\(^{338}\) and have identified other key residues. We have also shown that Abg can tolerate a number of mutations in this region without losing activity.

**MATERIALS AND METHODS**

**Strains, Culture Conditions—**\textit{Escherichia coli} strains JM101 (Viera and Messing, 1988) and RZ1032 (Kunkel, 1987) have been described. Plasmid pTZ18R:abg was constructed by isolating the SstI-SalI fragment containing the coding sequence of the \(\beta\)-glucosidase gene (abg) from pABG5 (Wakarchuk et al., 1986) and inserting it into pTZ18R (Mead et al., 1986). JM101 was maintained on M9 minimal medium (Viera and Messing, 1988). Plasmid-containing strains were grown in Luria broth (Miller, 1972) containing 100 \(\mu\)g/ml ampicillin.

**DNA Techniques and in Vitro Mutagenesis—**Isolation of plasmid DNA was according to Maniatis et al. (1989). Single-stranded DNA was isolated by the following method. Cultures were grown on TYP (16 g/liter tryptone, 16 g/liter yeast extract, 5 g/liter NaCl, 2.5 g/liter KH\(_2\)PO\(_4\)) medium containing 100 \( \mu \)g/ml ampicillin and 10\(^5\) pfu/ml helper phage M13K07 (Viera and Messing, 1987). Kanamycin (50 \( \mu \)g/ml) was added 1 h after inoculation, and the culture was grown 6–10 h at 37 °C. Phagemid were precipitated with 1.7 M ammonium acetate and 12% (w/v) PEG-6000. Single-stranded DNA was isolated from the phagemid by the method of Kristensen et al. (1987). The region of the active site nucleophile was mutagenized by the random enzymatic mutagenesis method of Lehtovaara et al. (1988). Uracil-containing template was generated by growing the plasmid in strain RZ1032 (da" ura" ung"). TT DNA polymerase was used instead of the Klenow fragment of DNA polymerase I for the extension reactions. A 17-base pair oligonucleotide primer (pGGTTTACGCTCCGGCGC) 60 nucleotides away from the active site nucleophile was used for the mutagenesis. dNTP concentrations in the limited extension reaction were adjusted to generate mutations in a region covering the next 150 nucleotides. Site-directed mutants were generated by the method of Kunkel (1987) with modifications for phagemid vectors (McClary et al., 1989). Specific mutations of the active site nucleophile (Glu\(^{338}\)) were carried out with oligonucleotide primers: pTACATCACCZ(Ar$77, pGACCA GCCGAHACCTCGTATTATAC, Tyr$76, pCGCTCGATTTAC GCCGAAAC (n = A, C, G, or T; B = C, G, or T; H = A, C, or T). After in vitro mutagenesis, the plasmid DNA was transformed into JM101. Transformants were selected on LB agar containing 2% X-Glc, 1 mM isopropyl-\(\beta\)-D-thiogalactoside, and 100 \( \mu \)g/ml ampicillin. Possible mutants were screened by single-track sequencing and confirmed by complete sequencing reactions. DNA sequencing was performed by the procedure of Tabor and Richardson (1987) with the following modifications: the reaction temperature was increased to 45 °C and deaza-7-dGTP replaced dGTP.

**Enzyme Assays—**Cells were grown overnight in 2 ml of LB ampicillin at 37 °C. One ml of a culture was used to inoculate 10 ml of MM6 salts (Gerhardt et al., 1981) supplemented with 0.6% (v/v) glycerol, 1 mM isopropyl-\(\beta\)-D-thiogalactoside, and 100 \( \mu \)g/ml ampicillin. Cultures were grown for 6 h. Cells were harvested by centrifugation at 6,000 \( \times \)
Fig. 1. Multiple sequence alignment of proteins in the same family as Abg. BGLS$SULSO, $-galactosidase from Sulfolobus solfataricus (Cubellis et al., 1990); BGL$ALSULSO, $-galactosidase from S. solfataricus (Little et al., 1988); BGLBECOLI, $-phospho-$-glucosidase from E. coli (Schnetz et al., 1987); LACGS$TAU, $-phospho-$-galactosidase from Streptococcus lactis (Withers et al., 1988); LACGS$LACCA, $-phospho-$-galactosidase from Lactobacillus casei (Porter and Chassy, 1988); BGL$LACCHL, $-glucosidase from Clostridium thermocellum (Grabitz et al., 1991); BGL$SALCSA, $-glucosidase from Caldocellum saccharolyticum (Love et al., 1988); BGL$AGRSP, $-glucosidase from A. faecalis (Wakarchuk et al., 1988); BGL$BACPO, $-galactosidase from Bacillus polmyxma; BGLB$BACPO, $-galactosidase from B. polymyxa (Gonzalez-Candelas et al., 1990); LPHD$SHU, human lactase-phlorizin hydrolase domain 4; LPHD$SHA, rabbit lactase-phlorizin hydrolase domain 4; LPHD$SHU, human lactase-phlorizin hydrolase domain 3; LPHD$SHU, rabbit lactase-phlorizin hydrolase domain 3 (Mantei et al., 1988). The alignment was performed with the CLUSTER program (Higgins and Sharp, 1989) from PC GENE by Intelligenetic, Mountain View, CA.

Consensus: $-GL...L...Y...$-YITENG G-D...D...-RI-I-Y...-N..-AI...D...

Active Mutants

| F | S | G |
|---|---|---|
| L | V | P |

Inactive Mutants

| D | S | Y | N | P |
|---|---|---|---|---|
| C | A | |

RESULTS AND DISCUSSION

Mutant Design—Mutants that changed the nucleophile were generated by site-directed mutagenesis using specific oligonucleotide primers. In addition, we wanted to investigate the importance of the residues in the surrounding region. Point mutations distributed throughout this region of the protein were generated by the enzymatic misincorporation method (Lehtovaira et al., 1988). Alignment of $-glucosidases and $-galactosidases in the same family as Abg indicated 6 conserved residues in this region (Fig. 1). Site-directed mutants were generated at these residues to supplement the randomly generated mutations.

Analysis of Mutations—Mutants were initially characterized by plating on X-Glc plates and active mutants were classified by their ability to cleave X-Glc and form blue colonies. In the region of 60 residues surrounding the nucleophile, 43 point mutations were generated at 22 different residues (Table I). Only five positions were identified at which mutations inactivated the enzyme (white colonies on X-Glc plates). Four of the five positions were conserved residues.

Role of Glu$358$—Inhibitor studies indicate that Glu$358$ is the nucleophile (Withers et al., 1990). Multiple sequence alignments of family members show it is absolutely conserved in all catalytically competent members of the group, and the 5 residues around Glu$358$ (YITENG) are also highly conserved. Nine different mutations were generated at Glu$358$, all of which resulted in a decrease in activity of at least 10,000-fold. Indeed, only the most conservatively modified mutant (E358D) retained any measurable activity under our assay conditions. These severe effects on enzyme activity of mutations at Glu$358$ confirm the importance of this residue and are consistent with its presumed role. Since Glu$358$ must form a covalent intermediate during catalysis, the spacing and the ability to act as a nucleophile would be expected to be critical to catalysis. The change of Glu$358$ to aspartic acid retains the same chemical group but shortens the side chain and would be expected to strain any covalent bond formed during catalysis.
play a role in, for example, positioning the nucleophile. In type activity. Although this is a very conservative mutation, the equivalent residue in all members of the family is smaller to allow alignment beyond. Secondary structure predictions are completely conserved in all family members with the sequence D-R-Y, corresponding to residues 374-380 of Abg. One possible reason for the small effects of mutations at Arg377 and Tyr380 had only very small effects on enzyme activity, with only the highly disruptive change of R377P severely decreasing activity. This is most surprising since Arg377 and Tyr380 are conserved in all family members, thus would be expected to be important for enzymatic function. The conservative changes of R377K and Y380F reduce activity only 2-fold, which may not even be significant in our assay system. While other changes at these residues decrease activity 10-fold it is not at all clear why these residues are so highly conserved, since they clearly do not play an important role in catalysis. One possible reason for the small effects of mutations at Arg377 and Tyr380 may have been that these residues are involved in binding the second sugar moiety of cellobiose, thus use of the artificial substrate pNPG would not detect these changes. Assays with cellobiose were therefore performed to test this hypothesis and essentially identical levels of activity to those with pNPG were found, indicating that these residues do not affect the second glucose-binding site.

**Conclusions**—This study has further confirmed the essential role of Glu358 in catalysis and has identified Asp374 as a candidate for the role of acid-base catalyst. Four other residues in this region of the protein were identified as being of potential importance based upon sequence homologies. Any modification of one of these (Gly360) resulted in severe reductions in activity while changes in the others resulted in smaller (2-10) fold reductions in activity. In contrast, mutations of most other, non-conserved, residues did not significantly affect activity, even when drastic changes such as glutamate to lysine were effected. These results therefore strongly support the use of sequence alignments to indicate important regions of a protein as long as the number of sequences compared is large enough. However, they do point out also that sequence alignment alone is not sufficient to identify residues since mutation of several fully conserved residues resulted in only relatively modest activity decreases.

### Table I

| Strain    | Mutation | pNPG activity     |
|-----------|----------|-------------------|
| ABG (wild-type) |          | 1.9               |
| 1068AB78 | H 339 L  | 1.8               |
| 1068AB59 | H 339 P  | 2.0               |
| 1068AB104| E 343 V  | 1.7               |
| 1068AW101-1 | T 344 S | 2.1               |
| 1068AB100| Y 346 F  | 1.4               |
| 1068AB9 | D 350 V  | 3.0               |
| 1068AB174| Y 355 F  | 3.4               |
| 1091GP3-1| T 357 P  | 2.2               |
| E356D-1 | E 358 D  | 0.0008            |
| E356N-1 | E 358 N  | 0.0002            |
| E356G-1 | E 358 E  | 0.0002            |
| E356C-8 | E 358 C  | 0.0001            |
| E358H-4 | E 358 H  | 0.0001            |
| 1068AW116| E 358 V  | 0.0001            |
| 1068AW88 | E 358 A  | 0.0002            |
| 1091GW5 | E 358 K  | 0.0001            |
| 1068AB200| N 359 S  | 0.1               |
| 1068GW69 | G 360 C  | 0.0001            |
| 1091GW2 | G 360 S  | 0.002             |
| 1068AW62 | C 362 Y  | 0.0001            |
| 1068AW92 | Y 363 S  | 3.2               |
| 1068AB184| Y 363 F  | 3.8               |
| 1068GB43| M 365 I  | 1.5               |
| 1086GB8 | V 367 I  | 2.8               |
| 1068AB192| E 368 D  | 1.6               |
| 1068GB9 | G 370 C  | 1.9               |
| 1068GB52| G 370 S  | 1.9               |
| 1068GB1 | E 371 K  | 1.4               |
| D374-2  | D 374 V  | 0.001             |
| D374-6  | D 374 G  | 0.003             |
| 1091GW6 | D 374 N  | 0.006             |
| 1068AW117| D 374 A  | 0.007             |
| D374-5  | D 374 E  | 0.02              |
| 1068AW6 | R 377 P  | 0.002             |
| R377I-19| R 377 I  | 0.1               |
| R377T-13| R 377 T  | 0.3               |
| R377-42 | R 377 K  | 0.8               |
| Y380-31 | Y 380 C  | 0.1               |
| Y380-30 | Y 380 S  | 0.3               |
| Y380-42 | Y 380 F  | 0.8               |

* μmol of pNPG released/min/mg of total cell protein.

**Importance of Asn358 and Gly360**—These residues are conserved in all members of the family, suggesting that they may play a role in, for example, positioning the nucleophile. In contrast, the region following Glu358 (Ala359, Asp374) shows very little sequence conservation and requires the insertion of gaps to allow alignment beyond. Secondary structure predictions for this region of Abg suggest the presence of short β-sheets and β-turns. Interestingly, the only mutant obtained involving a change in residue 359 (N359S) retained essentially wild-type activity. Although this is a very conservative mutation, it suggests that this residue is not absolutely critical to activity. In contrast, replacement of Gly360 by serine or cysteine results in severe reductions in activity. The small size of glycine may well be required to allow a turn, thus mutations which insert a large side chain may not be tolerated. A similar situation may obtain for Cys362. Even though this residue is not conserved, the mutation C362Y destroys all activity. Since the equivalent residue in all members of the family is smaller it is likely that the increase in bulk of the side chain is sufficient to affect secondary structure.

**Conserved D-R-Y Sequence**—There are 3 residues which are completely conserved in all family members with the sequence D-R-Y, corresponding to residues 374-380 of Abg. Point mutations generated at these residues decreased the activity of Abg by differing amounts. Mutations at Asp374 had the greatest effect on activity with the highly conservative change D374E reducing activity 100-fold and other changes decreasing activity by up to 2,000-fold. However, all of the mutations at Asp374 did retain some measurable activity, in contrast to mutations at Glu358, suggesting a very important, but not critical role in catalysis. Several possibilities arise concerning the potential role of this residue. It could serve a structural role in maintaining the required conformation, or it might have a role in binding the substrate, or more critically in binding and stabilizing the transition state. It is, however, tempting to suggest another role for this residue, that of acid-base catalyst which protonates the leaving group and subsequently deprotonates the water as it attacks. In other glycosidases, in particular the lysozymes (Sinnott, 1990; Grutter et al., 1983; Inoto et al., 1972) this acid-base catalyst has been suggested to be a carboxylic acid. There are only 5 conserved aspartate or glutamate residues in Abg, excluding the nucleophile. Residue Asp374 is the closest of these to Glu358, and the linear separation of these 2 residues is tantalizingly similar to the separation between the 2 equivalent residues in hen lysozyme (13 residues). It is therefore tempting to suggest that Asp374 serves as the acid-base catalyst. Indeed, the 100-2,000-fold reduction in activity upon mutation is consistent with this assignment and is also consistent with the reduction of 1,000-fold or greater observed for the hen and T4 lysozymes when the acid catalyst was mutated (Malcolm et al., 1989; Anand et al., 1988). However, further studies will be required to confirm or refute this suggestion.
REFERENCES

Anand, N. N., Stephen, E. R., and Narang, S. A. (1988) Biochem. Biophys. Res. Commun. 153, 862-868
Boizet, B., Villeval, D., Slos, P., Novel, M., Novel, G., and Mercenier, A. (1988) Gene (Amst.) 62, 249-261
Bradford, M. M. (1976) Anal. Biochem. 72, 248-254
Breidt, F. Jr., and Stewart, G. C. (1987) Appl. Environ. Microbiol. 53, 969-973
Brei!!!is, M. V., Rozzo, C., Montecucchi, P., and Rossi, M. (1990) Gene (Amst.) 94, 89-94
Day, A. G., and Withers, S. G. (1986) Biochem. Cell. Biol. 64, 914-922
Gerhardt, P., Murray, R. G. E., Costilow, R. N., Nester, E. W., Wood, W. A., Kreig, N. R., and Phillips, G. B. (1981) Manual of Methods for General Bacteriology, American Society for Microbiology, Washington, D. C.
Gilkes, N. R., Warren, R. A. J., Miller, R. C., Jr., and Kilburn, D. C. (1988) J. Biol. Chem. 263, 10401-10407
Gonzalez-Candelas, L., Romon, D., and Polaina, J. (1990) Gene (Amst.) 95, 31-38
Grabnitz, F., Seiss, M., Rucknagel, K. P., and Staudenbauer, W. L. (1991) Eur. J. Biochem. 206, 301-309
Gutter, M. G., Weaver, L. H., and Matthews, B. W. (1983) Nature 303, 828-830
Henrisatt, B. (1991) Biochem. J. 280, 309-316
Higgins, D. G., and Sharp, P. M. (1989) Comput. Appl. Biosci. 5, 151-153
Imoto, T., Johnson, L. N., North, A. C. T., Phillips, D. C., and Rupley, J. A. (1972) Enzymes 7, 665
Kristensen, T., Vees, H., and Ansorge, W. (1987) Nucleic Acid Res. 15, 5507-5516
Kunkel, T. A., Roberts, J. D., and Zakour, R. A. (1987) Methods Enzymol. 154, 367-382
Lammine, K. (1979) Nature 277, 680-685
Lehtovaara, P. M., Koivula, A. K., Bamford, J., and Knowles, J. K. C. (1988) Protein Eng. 2, 63-68
Little, S., Cartwright, P., Campbell, C., Prenneta, A., McChesney, J., Mountain, A., and Robinson, M. (1989) Nucleic Acids Res. 17, 7080
Love, D. R., Fisher, R., and Bergquist, P. L. (1988) Mol. Gen. Genet. 213, 84-92
Malcolm, B. A., Rosenberg, S., Corey, M. J., Allen, J. S., De Baetselier, A., and Kirsch, J. F. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 133-137
Maniatis, T., Futsch, E. F., and Sambrook, J. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
Mantzi, N., Villa, M., Enzler, T., Wacker, H., Boll, W., James, P., Hunziker, W., and Semenza, G. (1988) EMBO J. 7, 2705-2713
McCly, J. A., Wity, F., and Geiessboder, J. (1989) Biotechniques 7, 282-289
Mead, D. A., Sza!heska-Skorupa, E., and Kemper, B. (1986) Protein Eng. 1, 67-74
Miller, J. H. (1972) in Experiments in Molecular Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
Porter, E. V., and Chassy, B. M. (1988) Gene (Amst.) 62, 263-276
Schnetz, K., Toloczyki, C., and Kek, B. (1987) J. Bacteriol. 169, 2679-2690
Shewale, J. G. (1982) J. Biochem. (Tokyo) 14, 435-443
Sinnott, M. L. (1990) Chem. Rev. 90, 1173
Tabor, S., and Richardson, C. C. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 4767
Viera, J., and Messing, J. (1987) Methods Enzymol. 153, 3-11
Viera, J., and Messing, J. (1988) Gene (Amst.) 19, 259-268
Wakarchuk, W. W., Kilburn, D. C., Miller, R. C., and Warren, R. A. J. (1986) Mol. Gen. Genet. 205, 146-152
Wakarchuk, W. W., Greenberg, N. M., Kilburn, D. C., Miller, R. C., and Warren, R. A. J. (1988) J. Bacteriol. 170, 301-307
Withers, S. G., and Street, I. P. (1988) J. Am. Chem. Soc. 110, 655; Withers, S. G., Warren, R. A. J., Street, I. P., Rupitz, K., Kempton, J. B., and Aebbersold, R. (1990) J. Am. Chem. Soc. 112, 9887-9889