

DISCUSSION

Design of chimeric enzymes from naturally-occurring enzymes is a straightforward approach to creating a novel enzymatic function or altering catalytic efficiency. Our model system consists of two β -glucosidases, Glu1 and Dhr1, each with a strict specificity for its own physiological substrate, although they share 70% sequence identity and contain identical catalytic amino acids and glycone recognition and binding motifs (TFNEP and ITENG). Based on the modeled 3D structures of Glu1 and Dhr1, we have been able to combine two different substrate specificities in a single chimeric enzyme by replacing the C-terminal domain of Glu1 with the homologous domain from sorghum Dhr1. This strategy added novel substrate specificity (e.g., dhurrin hydrolysis) to the maize Glu1 isozyme and improved its catalytic efficiency on other substrates.

We produced five chimeric Glu1/Dhr1 β -glucosidases for the purpose of delineating the regions of primary structure that contain key amino acids and sequence motifs that determine the substrate (or aglycone) specificity. Each of the five chimeric enzymes exhibited the combined substrate specificities of the parental enzymes and, on average, 2-4 fold higher catalytic efficiency on certain substrates than the parental enzymes. The basis of this broadened substrate specificity and improved catalytic efficiency is thought to reside in amino acid substitutions that occur in the 53-amino-acid-long C-terminal domain of Dhr1 or its shorter fragments that were swapped with the homologous regions of Glu1. These results allow us to draw two conclusions. 1) The C-terminal domain of β -glucosidases includes residues that are necessary, but not sufficient, for aglycone recognition and binding. This is in agreement with the results of Singh and Hayahashi (20); they showed a chimeric enzyme obtained by replacing the C-terminal domain of a *C. gulvis* β -glucosidase with the homologous C-terminal domain of an *A. tumefaciens* β -glucosidase had the substrate specificity of the C-terminal region donor *A. tumefaciens* β -glucosidase. 2) The substrate specificity differences between maize and sorghum β -glucosidases are very likely due to amino acid substitutions in the extreme 47 to 53 amino acid long C-terminal domains of these enzymes. The N-terminal region, at least the first 41 amino acids of Glu1, does not appear to have a discernable role in substrate specificity. We replaced amino acids 1-41 of Dhr1 with the corresponding N-terminal region of Glu1. The resulting Dhr1/Glu1 chimeric enzyme hydrolyzed only dhurrin as Dhr1 does (data not shown). Thus, these data strongly suggest that the N-terminus of Glu1, and possibly of other β -glucosidases, is not involved in substrate (i.e., aglycone)

specificity. However, it is involved in catalysis because it contains a universally conserved amino acid (Q38), which is in the glycone-binding pocket of the active site (16, 17, 18).

We postulate that although Glu1 and Dhr1 differ by 151 amino acid substitutions, 5 deletions and 7 additions at 514 positions (~30% sequence divergence), only a small number of these changes are relevant to the substrate specificity differences between them. Indeed only 4 of the 22 variant amino acid sites map to or around the active site of the modeled enzymes (Fig. 5). We postulate further that more than 90% of the amino acid substitutions separating Glu1 and Dhr1 are adaptively and functionally neutral based on Khimura's theory of neutral evolution (35). There are well-documented examples in the literature supporting this postulate. For example, eubacterial and mitochondrial isocitrate dehydrogenases differ with respect to coenzyme specificity; the former is NADP-dependent while the latter is NAD-dependent. Moreover, both enzymes have essentially the same tertiary structure although they differ by 250 amino acid substitutions at 320 positions. Only 6 of these 250 amino acid substitutions determine coenzyme specificity, as shown elegantly by shifting NADP specificity to NAD specificity or vice versa by replacing these amino acids in the coenzyme binding pocket (36). Similarly, Wilks et al. (37) changed a lactate dehydrogenase to a malate dehydrogenase by replacing a single key amino acid although the two enzymes differed by 230 amino acid substitutions. Other examples of bringing about dramatic changes in substrate specificity and catalytic properties include: changing the substrate specificity and double-bond positional specificity of an acyl-carrier protein desaturase by five amino acid replacements (22); increasing catalytic efficiency and broadening substrate specificity in a chimeric protease constructed by recombining the N-terminal domain of coagulation factor X with the C-terminal domain of trypsin (23); and introducing the active site of nonheme iron superoxide dismutase into *E. coli* thioredoxin and changing it to a superoxide dismutase (38).

The substrate specificity data from Chim 2 provided the first clue to the importance of the C-terminal 53-amino-acid-long domain of Dhr1 in dhurrin hydrolysis. These data suggested that this domain contains key substrate specificity determinants, but these determinants alone did not change the specificity of a Glu1/Dhr1 chimera entirely to that of the C-terminal region donor Dhr-1. The 53-amino-acid-long Dhr1 domain contains 22 amino acid substitutions distributed throughout and a 4-amino-acid-long add-on region at the C-terminus when compared to Glu1 (Fig. 2). However, from the Chim 2 data, it was not possible to identify specific amino acids or sequence motifs that are important in dhurrin binding and hydrolysis. Consequently, this Dhr1 domain was split into two

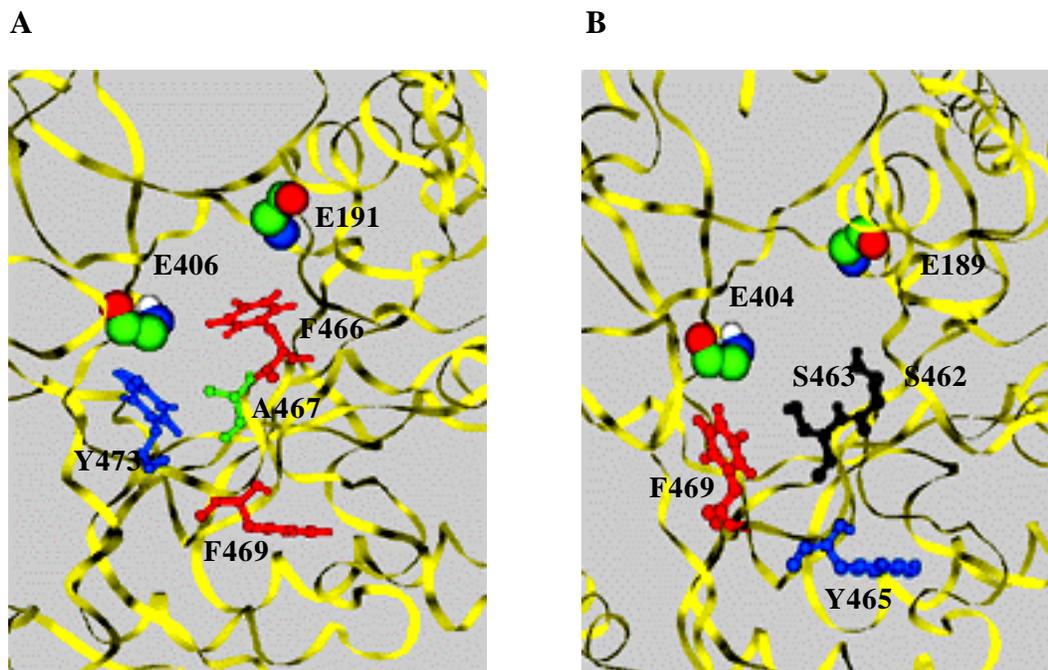


FIG. 5. Active sites of modeled Glu1(**A**) and Dhr1(**B**). The residues Y473, F469, A467, and F466 shown in **A** are postulated to be in the aglycone binding pocket of the active site and are necessary for DIMBOAGlc hydrolysis, while the residues F469, Y466, S463, and S462 in **B** are postulated to be necessary for dhurrin hydrolysis. The catalytic glutamic acids (E406 and E191, the nucleophile and the acid-base catalyst, respectively, in **A** and E404 and E189 in **B**) are shown in the space-filling form.

subdomains, which are represented in Chim 15 (23 amino acid-long C-terminal subdomain) and Chim16 (30-amino-acid-long N-terminal subdomain), to determine which subdomain was important in dhurrin specificity. The substrate specificity and kinetic data from these two chimeras unequivocally showed that the 30-amino acid long subdomain, which contains 10 amino acid substitutions, played a far greater role in dhurrin specificity than the 23 amino acid long subdomain. This is clearly evident in the fact that Chim 16 hydrolyzes dhurrin 12 times better than Chim 15 (Table 3). These data are also consistent with the modeling data in that none of the 23 amino acids from the C-terminus of Dhr1 or Glu1 maps to and around the active site of the modeled 3-D structures of these enzymes. Therefore, the 8 substitutions and 4-amino acid long addition that separate Glu1 from Dhr1 at the extreme C-terminus has a rather small and probably indirect effect on substrate specificity.

The substrate specificity and kinetic data clearly suggested that one or more of the 10 amino acid substitutions in the 30 amino acid long Dhr1 subdomain in Chim 16 plays a key role in dhurrin recognition and hydrolysis. Again, to bring further clarity to the specific site(s) responsible for dhurrin recognition and hydrolysis, the 30-amino-acid-long subdomain was divided into two segments, after leaving out the invariant region GIVYVDR separating them. The resulting two Dhr1 peptides ⁴⁶²SSGYTERF⁴⁶⁹ and ⁴⁷⁷ENGCERTMKR⁴⁸⁶ were used to replace their homologues in Glu1 by domain swapping, which yielded Chim 21 and 22, respectively (Table 1). The substrate specificity and kinetic data obtained with these two chimeras indicated that Chim 21 hydrolyzed dhurrin and had the best kinetic properties, having the lowest K_m and highest k_{cat} and efficiency coefficient among five Glu1/Dhr1 chimeras (Table 3). Thus, the Dhr1 peptide ⁴⁶²SSGYTERF⁴⁶⁹ alone is sufficient to convert Glu1 to Dhr1 specificity when it replaces the homologous peptide ⁴⁶⁶FAGFTERY⁴⁷³ of Glu1. This observation strongly suggests one or more of the 4 amino acid substitutions that occur within this peptide are in the aglycone binding pocket of the active site, and in this case they are necessary for the binding of dhurrin in correct angle and steric complementarity for hydrolysis. We postulate that S462 and Y465 are the key residues because they are unique to Dhr1 and they and their homologues map to the active site of the modeled β -glucosidases (Fig. 5). These two amino acids are ideal candidates to form hydrogen-bonding interactions through their side chain -OH group with the -OH group in the aglycone (*p*-hydroxy-(S)-mandelonitrile) moiety of dhurrin. Such interactions might be important in the attainment and stabilization of the transition state by the enzyme-dhurrin complex. Interestingly, these two sites are also different between the two maize β -

glucosidase isozymes Glu1 (F466 and F469) and Glu2 (Y466 and Y469), which differ in substrate specificity. The former hydrolyzes 6BNGlc and a variety of nitrophenyl β -glycosides while the latter does not hydrolyze 6BNGlc at all and hydrolyzes nitrophenyl β -glycosides 5-6 times less efficiently than Glu1 (24). Moreover, the 3-D structure of white mustard myrosinase provided information about key residues involved in substrate binding, where soaking enzyme crystals with the natural substrate sinigrin has identified the glucose-binding site, although the aglycone-binding site could not be identified. However, the docking of sinigrin into the active site suggested that the aglycone moiety is located in a hydrophobic pocket that is formed by residues F371, F473, I257 and Y330 (17). The F473 of myrosinase is homologous with Y473 of Glu1 and F469 of Dhr1, which are bolded in peptides ⁴⁶⁶FAGF**TER**^{Y473} and ⁴⁶²SSGY**TER**^{F469}, respectively, supporting our hypothesis these are key residues in aglycone recognition, and thus substrate specificity.

The results also indicate that the Dhr1 peptide ⁴⁷⁷ENG**CERT**MKR⁴⁸⁶ in Chim 22 substantially contributes to dhurrin hydrolysis. However, in this case the catalytic efficiency coefficient (k_{cat}/K_m) is less than one-third of Chim 21 (Table 3), and it is not possible to predict which of the five amino acid substitutions that are in this peptide imparts dhurrin hydrolysis capability to Chim 22. This is because none of these five sites maps to the active site of the modeled enzymes, nor were their homologues implicated in substrate binding and catalysis directly in the literature. Thus it is conceivable that one or more of these sites have indirect global effects on the structure of the active site for the binding of dhurrin in correct angle and steric complementarity for hydrolysis.

One of the significant findings of this study was to show that the physiological substrates DIMBOAGlc and dhurrin, respectively, of Glu1 and Dhr1, were potent competitive inhibitors of their heterologous enzymes. The K_i values of 0.076 mM for dhurrin and of 0.009 mM for DIMBOAGlc indicate that each substrate binds to the ground state of its heterologous enzyme with high affinity, but the hydrolysis of the β -glycosidic bond does not occur. It should be pointed out that the structures and substituents of the two aglycones, DIMBOA and p-hydroxy-(S)-mandelonitrile, are different (Fig. 1), the former being bulkier than the latter. The question of why DIMBOAGlc is not hydrolyzed by Dhr1 or of why dhurrin is not hydrolyzed by Glu1 in view of tight binding has only speculative answers at this time. The most plausible scenario is that the binding, although tight, does not position the glycosidic bond in correct angle and distance (~ 2.5 Å) with respect to the nucleophile (E406 in Glu1 and E404 in Dhr1 in the motif ITENG, Fig. 1) or it does not allow the enzyme-substrate complex to attain a transition state energetically favorable for

hydrolysis. The fact that two (Chim 16 and 22) of our five chimeras hydrolyzed dhurrin when a Dhr1 peptide that does not contain an active site residue replaced its Glu1 homologue suggests that the problem is indeed one of positioning of the substrate for hydrolysis, which can be affected indirectly by amino acid substitutions that are outside the active site. The above question also has bearing on the evolution and existence of two distinct β -glucoside biosynthesis pathways and chemical defense compounds in two closely related plant genera, *Zea* and *Sorghum*, which are thought to have diverged from a common ancestor only 25-30 million years ago. For example, the dhurrin biosynthesis pathway starts with the amino acid tyrosine as precursor and produces a cyanogenic β -glucoside (dhurrin) as end product, while the DIMBOAGlc biosynthesis pathway starts with indole (a tryptophan analogue) as precursor and produces a hydroxamic acid glucoside (DIMBOAGlc) as end product. Which pathway did the common ancestor of sorghum and maize have and how and why was another β -glucoside pathway and defense compound 'invented' in one of the lineages are important questions for the evolutionary biologist to answer.

In conclusion, we were able to convert the maize Glu1 isozyme (a DIMBOA-glucosidase) to a dhurrinase by domain-swapping and all of the five chimeric enzymes produced hydrolyzed the artificial substrate *p*NPG with 1.5 to 4.4 fold relative efficiency and *o*NPG with 1.5 to 3.1 fold relative efficiency than Glu1. This was accomplished by replacing a 47-amino-acid-long C-terminal domain of Glu1 and its smaller segments with the homologous Dhr1 domain and its smaller segments. The shortest Dhr1 peptide to convert Glu1 to dhurrinase was 8-amino-acid-long, differing by 4 amino acid substitutions, two of which mapped to and around the active site of the modeled enzymes. Although all of the five Glu1/Dhr1 chimeric enzymes hydrolyzed both dhurrin and DIMBOAGlc, none of them either equaled or exceeded their parental enzymes in terms of catalytic efficiency for these natural substrates. However, with one exception (Chim 16) they were better DIMBOA-glucosidase than dhurrinase, having 65 to 88% catalytic efficiency of Glu1. In general, DIMBOAGlc hydrolysis and dhurrin hydrolysis efficiencies were negatively correlated. These data, taken together, show that the C-terminal domain of β -glucosidases contain sites that are necessary for aglycone recognition and binding, as well as positioning of the β -glycosidic bond for the catalytic nucleophile. In our specific model system, the substrate specificity differences between maize and sorghum β -glucosidases are very likely due to amino acid substitutions in the extreme 47 to 53 amino acid long C-terminal domains of these enzymes. However, these amino acid substitutions (specifically S/Y and Y/F) are

essential but not sufficient for determining substrate specificity because modeling and substrate docking suggest the contribution of the sites outside the C-terminal domain to the structure of the aglycone binding pocket within the active site. Precise definition of specific residues that make up the aglycone binding pocket and thus the substrate specificity will await the results of 3-D studies on enzyme-aglycone and enzyme-competitive inhibitor complexes by x-ray crystallography, as well as those of site-directed mutagenesis targeting candidate sites identified by domain swapping and modeling. In collaboration with B. Henrissat and his associates, we have ongoing 3-D studies on crystallized maize Glu1 isozyme and an inactive E191D mutant Glu1-dhurrin complex. Additional domain-swapping studies focusing on Dhr1/Glu1 chimeras and site-directed mutagenesis targeting the above-mentioned candidate sites will be underway in the near future.

LITERATURE CITED

1. **Bell, A.** (1981) *Annu. Rev. Plant Physiol.* **32**, 21.
2. **Conn, E. E.** (1981) *Cyanogenic glycosides*, eds Stumpf, P. K., & Conn, E. E. (Academic Press), Vol. **7**, pp. 479-500.
3. **Niemeyer, H. M.** (1988) *Phytochemistry.* **27**, 3349-3358.
4. **Poulton, J. E.** (1990) *Plant. Physiol.* **94**, 401-405.
5. **Brzobohaty, B., I. Moore, P. Christofferson, L. Bako, N. Campos, J. Schell, & K. Palme.** (1993) *Science* **262**, 1051-1054.
6. **Matsuzaki, T. & Koiwai, A.** (1986) *Agric. Biol. Chem.* **50**, 2193-2199.
7. **Schliemann, W.** (1984) *J. Plant Physiol.* **116**, 123-132.
8. **Smith, A. R. & van Staden J.** (1978) *J. Exp. Bot.* **29**, 1067-1073.
9. **Wiese, G. & Grambow H.** (1986) *Phytochemistry* **25**, 2451-2455.
10. **Dharmawardhana, D. P., Ellis, B. E., & Carlson, J. E.** (1995) *Plant Physiol.* **107**, 331-339.
11. **Leah, R., Kigel, J., Svendsen, I., & Mundy, J.** (1995) *J. Biol. Chem.* **270**, 15789-15797.
12. **Simos, G., Panagiotidis, C. A., Skoumbas, A., Choli, D., Ouzounis, C., & Georgatsos, J. G.** (1994) *Biochim. Biophys. Acta.* **1199**, 52-58.
13. **Henrissat, B.** (1991) *Biochem. J.* **280**, 309-316.

14. **Trimbur, D. E., Warren, R. A. J. & Withers, S. G.** (1992) *J Biol Chem.* **267**, 10248-10211.
15. **Trimbur, D., Warren, R. A. J. & Withers, S. G.** (1993) β -Glucosidases: *Biochemistry and Molecular Biology*, eds. Esen, A., pp. 42-55.
16. **Barrett, T., Suresh, C. G., Tolley, S. P., Dodson, E. J., & Hughes, M. A.** (1995) *Structure* **3**, 951-960.
17. **Burmeister, W. P., Cottaz, S., Driquez, H., Iori, R., Palmieri S., & Henrissat, B.** (1997) *Structure* **5**, 663-675.
18. **Sanz-Aparicio, J., Hormoso, J. A., Martinez-Ripoll, M., Lequerica, J. L. & Polaina, J.** (1998) *J. Mol. Biol.* **275**, 491-502.
19. **Davies, G., & Henrissat, B.** (1995) Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853-859.
20. **Singh, A., Hayashi, K., Hoa, T. T., Kashiwagi, Y., & Tokuyasu, K.** (1995) *Biochem. J.* **305**, 715-719.
21. **Terashima, M., Hosono, M., & Katoh, S.** (1997) *Appl. Microbiol. Biotechnol.* **47**, 364-367.
22. **Cahoon, B. E., Lindqvist, Y., Schneider, G., & Shanklin, J.** (1997) *Proc. Natl. Acad. Sci. USA* **94**, 4872-4877.
23. **Hopfner, K. P., Kopetzki, E., Kresse, G. B., Bode, W., Huber, R., & Engh, R. A.** (1998) *Proc. Natl. Acad. Sci. USA* **95**, 9813-9818.
24. **Cicek, M. & Esen, A.** (1999) *Biotech. & Bioeng.* **63**, 392-400.
25. **Horton, R., Cai, M. Z., Ho, S. N. & Pease, L. R.** (1990) *Biotechniques* **8**, 528-535.

26. Esen, A., & Cokmus, C. (1990) *Biochemical Genetics* **28**, 319-336.
27. Martin, T., Wohner, R. V., Hummel, S., Willmitzer, L., & Frommer, G. (1992) *GUS protocols*, eds. Gallagher, S. R., pp 23-42.
28. Raabo, E. & Terkildsen, T. C. (1960) *Scand. J. Clin. Lab Invest.* **12**, 402
29. Cicek, M., & Esen, A. (1998) *Plant Physiol.* **116**, 1469-1478.
30. Robyt, F. J., & White, B. J. (1990) *Biochemical Techniques, Theory and Practice*, pp. 107-108.
31. Sali, A., & Blundell, T.L. (1993) *J. Mol. Biol.* **234**, 779-815.
32. Pearlman, D.A., Case, D.A., Caldwell, J.W., Ross, W.S., Cheatham, T.E. III, Ferguson, D.M., Seibel, G.L., Singh, U.C., Weiner, P.K., & Kollman, P.A. (1995) AMBER4.1, University of California, San Francisco.
33. Hosel, W., Tober, I., Eklund, S. H., & Conn, E. E. (1987) *Arch. Biochem. Biophys.* **252**, 152-162.
34. Babcock, G. W., & Esen, A. (1994) *Plant. Sci.* **101**, 31-39.
35. Kimura, M. (1983) *The Neutral Theory of Molecular Evolution*. (Cambridge University Press), Cambridge, England.
36. Dean, A. M. (1998) The Molecular anatomy of an ancient adaptive event. *Amer. Sci.* **86**, 26-37.
37. Wilks, H. M., Hart, K. W., Feeney, R., Dunn, C. R., Muirhead, H., Chia, W. N., Barstow, D. A., Atkinson, T., Clarke, A. R. & Holbrook, J. J. (1988) *Science* **242**, 1541-1544.

38. Pinto, A. L., Hellings, H. W. & Caradonna, J. P. (1997) Proc. Natl. Acad. Sci. USA 94, 5562-5567.