Investigation of the functional relevance of the catalytically important Glu^{28} in family 51 arabinosidases

Mounir Ferchichi, Caroline Rémond, Roselyne Simo, Michael J. O'Donohue*

Institut National de la Recherche Agronomique, UMR FARE, 8, rue Gabriel Voisin, P.O. Box 316, 51688 Reims Cedex 2, France

Received 15 July 2003; revised 14 September 2003; accepted 15 September 2003

First published online 26 September 2003

Edited by Ulf-Ingo Flügge

Abstract **The α-L-arabinofuranosidase (AbfD3) from *Thermobacillus xylanilyticus* is a family 51 glycosyl hydrolase. According to classification hierarchy, family 51 belongs to clan GH-A. While the major GH-A motifs, the catalytic acid-base and nucleophile, are conserved in AbfD3, a third catalytically important residue (Glu^{28}) does not appear to be analogous to any known GH-A motif. To evaluate the importance of Glu^{28}, bioinformatics analyses and site-saturation mutagenesis were performed. The results indicate that Glu^{28} forms part of a family 51 arabinosidase motif which might be functionally homologous to a conserved N-terminal motif found in exo-acting enzymes from families 1 and 5. Importantly, the data reveal that Glu^{28} is a key determinant of substrate recognition in the −1 subsite, where it may also play an important role in water-mediated deglycosylation of the glycosyl–enzyme covalent intermediate.**

Key words: α-L-Arabinofuranosidase; Glycosyl hydrolase; Family 51; Clan GH-A; Transglycosylation

1. Introduction

Arabinofuranosidases (EC 3.2.1.55), which are found in soil bacteria, fungi and plants, are accessory enzymes which participate in the degradation of plant hemicelluloses. By catalysing the hydrolysis of α-1,2, α-1,3 or α-1,5 L-arabinofuranoside bonds in polymers such as arabinoxylan or arabinan, arabinofuranosidases complement the action of the major depolymerising enzymes, such as xylanases, and allow the complete degradation of plant biomass. Recently, this hitherto understudied class of enzymes has been the subject of renewed interest because new potential industrial uses for arabinofuranosidases have been identified. Importantly, it is now considered that robust arabinofuranosidases will probably be required for enzymatic bioconversion strategies aimed at the complete saccharification of lignocellulose residues [1]. In addition, arabinofuranosidases have been shown to be useful as flavour enhancers for the wine industry and as nutritional improvers in animal feed [2]. Finally, arabinofuranosidases which display transglycosylation ability constitute potentially interesting tools for chemoenzymatic synthesis of arabinose-containing compounds that are difficult to access via organic synthetic methods [3].

The majority of arabinofuranosidases known to date belong to family 51 of the glycosyl hydrolase classification system. This family is exclusively composed of retaining arabinofuranosidases which catalyse hydrolysis via a double displacement mechanism leading to net conservation of the anomeric configuration during the reaction [4,5]. Although no structural information is yet available for family 51 arabinofuranosidases, hydrophobic cluster analysis of amino acid sequences and sequence alignment analysis have unambiguously assigned this family to clan GH-A of the glycosyl hydrolase classification system [6]. Clan GH-A is composed of 16 distinct enzyme families (1, 2, 5, 10, 17, 26, 30, 35, 39, 42, 51, 53, 59, 72, 79 and 86) which together cover a vast array of substrate specificities. Cursory examination of the structural models, which are now available for representative enzymes from half of the GH-A families, confirms the conclusions of previous work which attribute a common $(\beta/\alpha)_8$ fold to the catalytic domain of all GH-A members [7,8]. Moreover, more detailed analyses have revealed the existence of several common sequence motifs which translate into common functional features. The most prominent of these motifs are the two catalytic amino acids which catalyse the hydrolysis of the glycosidic bond via a retaining, double displacement mechanism. All GH-A members possess these two residues, one acting as the catalytic acid-base (contained within the GH-A motif, Asn-Glu) and the other as a nucleophile (Glu in GH-A), which are always localised at the C-terminal ends of β-strands 4 and 7 respectively [8]. In addition to these functionally critical, strictly conserved features, GH-A members often exhibit other less well-conserved motifs which contribute to a generic active site structure. These include a tryptophan (or equivalent aromatic amino acid) which, at the C-terminal extremity of β8, forms a non-prolyl *cis* peptide bond with an adjacent amino acid [8–11] and a tyrosine which is found at the C-terminal extremity of β6 [9]. However, despite such similarities in the catalytic domains of individual GH-A members, these enzymes show different substrate recognition features and are generally classified as either endo- or exo-acting enzymes. Previously, Juers et al. [12] have partly attributed this rich functional diversity to the presence of additional, non-catalytic domains which, when combined with the common $(\beta/\alpha)_8$ scaffold, modify the overall active structure.

Previous studies on the family 51 α-L-arabinofuranosidases from *Thermobacillus xylanilyticus* (AbfD3) [13] and *Geobacil-*

*Corresponding author. Fax: (33)-3-26 355 369.
E-mail address: michael.odonohue@univ-reims.fr (M.J. O'Donohue).

Abbreviations: *p*NP, *p*-nitrophenol; *p*NP-ara*p*, *p*NP-α-L-arabinopyranoside; *p*NP-glc*p*, *p*NP-β-D-glucopyranoside; *p*NP-xyl*p*, *p*NP-β-D-xylopyranoside; *p*NP-ara*f*, *p*NP-α-L-arabinofuranoside; *p*NP-rib*f*, *p*NP-β-D-ribofuranoside; AbfD3, α-L-arabinofuranosidase D3

0014-5793/03/$22.00
doi:10.1016/S0014-5793(03)01061-5

lus stearothermophilus (AbfA T-6) [14,15], which employed site-directed mutagenesis to identify the catalytic couple, have led to the localisation of the two catalytic residues, Glu^{176} (the acid-base catalyst) and Glu^{298} (the nucleophile). Unexpectedly, in our study we also revealed a third glutamate (Glu^{28}) which is not directly involved in catalysis, but which is nevertheless critical for normal hydrolytic activity. Initially, based on our kinetic data we tentatively proposed that Glu^{28} might be involved in the stabilisation of a catalytic transition state. Therefore, the aim of this study was to test this hypothesis and to gain further insight into the role of Glu^{28} in AbfD3. Additionally, we sought to investigate the occurrence of this motif in family 51, and more generally in clan GH-A.

2. Materials and methods

2.1. Bioinformatics analyses

Several analyses were performed using Structure Prediction Meta Server (http://bioinfo.pl/Meta/). Secondary structure predictions for AbfD3 were generated using SAMT99, SAMT02 [16] and Profsec [17] and a list of best potential structural homologues was compiled by 3D Jury [18]. 3D Jigsaw, which also forms part of the aforementioned Meta Server, provided information concerning potential domain structure in AbfD3. Finally, the motif discovery tool MEME (http://meme.sdsc.edu/) [19] was used to identify the five best conserved motifs, having a minimum size of 6 amino acids, within a sequence set composed of 43 full-length family 51 enzyme sequences (gi|17741471, gi|15140932, gi|13473100, gi|1913931, gi|1770022, gi|16079903, gi|25351195, gi|10174479, gi|25989577, gi|23020570, gi|22992223, gi|439852, gi|23326899, gi|23336185, gi|23335773, gi|22992228, gi|9967515, gi|15643050, gi|25351193, gi|557472, gi|29367433, gi|557475, gi|2494817, gi|21107441, gi|21112233, gi|3128393, gi|23023940, gi|1877425, gi|13398412, gi|13398414, gi|15228368, gi|13937191, gi|17380938, gi|22324432, gi|15026526, gi|1168266, gi|21280331, gi|13810441, gi|16417958, gi|6978931, gi|3581838, gi|15212230, gi|7481897).

2.2. General reagents and substrates

The substrates *p*NP-α-L-arabinopyranoside (*p*NP-ara*p*), *p*NP-β-D-glucopyranoside (*p*NP-glc*p*), *p*NP-β-D-xylopyranoside (*p*NP-xyl*p*) and all other general chemicals were supplied by Sigma-Aldrich (Saint-Quentin Fallavier, France). The substrates *p*NP-α-L-arabinofuranoside (*p*NP-ara*f*) and *p*NP-β-D-ribofuranoside (*p*NP-rib*f*), which were synthesised essentially according to established methods [3], were gifts from Dr R. Plantier (Université de Reims Champagne-Ardenne, France).

2.3. Site-saturation mutagenesis

Site-directed mutagenesis was performed using the QuikChange™ site-directed mutagenesis kit (Stratagene, Amsterdam, The Netherlands). Mutations were created in independent reactions using two complementary mutagenic primers. The forward primers had the same generic sequence, with NNN representing the mutated codon in each case: 5′-CGGCCATTTCTCGNNNCATCTCGGGC-GATGC-3′ Cys, TGC; Lys, AAA; Thr, ACC; Tyr, TAC; Leu, CTG; Ser, AGC; His, CAT; Gly, GGC; Ala, GCC; Asp, GAC; Gln, CAG; Phe, TTC; Trp, TGG; Pro, CCG; Arg, CGG; Ile, ATC; Met, ATG; Asn, AAC; Val, GTG.

After mutagenesis, all mutated genes were sequenced using a MegaBACE 1000 automated DNA sequencer (Amersham Biosciences, Saclay, France) and DYEnamic™ ET dye terminator technology (Amersham Biosciences). For sequencing, the T7 forward, 5′-TAA-TACGACTCACTATAGGG-3′, and T7 reverse, 5′-GCTAGTTAT-TGCTCAGCGG-3′, primers were employed. The wild type and mutated proteins were expressed in *Escherichia coli* cells and purified as previously described [20].

2.4. Enzyme assays

The activities of wild type and mutant enzymes were determined by continuous measurement of *p*-nitrophenol (*p*NP) release. An aliquot (900 μl) of substrate solution (*p*NP-glycoside, 5 mM in 50 mM sodium acetate, pH 5.8) was mixed in a cuvette with 100 μl of enzyme and pre-incubated at 60°C for 5 min. Afterwards, hydrolysis of the *p*NP-glycoside was monitored over a 15 min period at 401 nm. To calculate the specific activity of each mutated enzyme, the molecular extinction coefficient was assumed to be the same as that of the wild type AbfD3 ($\varepsilon_{280nm} = 85\ 230\ M^{-1}\ cm^{-1}$). For determination of K_M and k_{cat}, the same assay was performed using various substrate concentrations usually ranging from 0.001 to 15 mM. Kinetic parameters were automatically derived from Michaelis–Menten representations using the SigmaPlot 2000 software (version 6.1) equipped with the Enzyme Kinetics module 1.0 (SPSS Science, Paris, France).

2.5. High pressure anion exchange chromatography

The progression of transglycosylation reactions was monitored by analysing the products of discontinuous reactions using high pressure anion exchange chromatography equipped with pulsed amperometric detection (HPAEC, Dionex, CA, USA). The reactions were prepared by adding 0.11 IU (approximately 200 ng) of AbfD3 to 280 μl of *p*NP-ara*f* (5 mM in 50 mM sodium acetate, pH 5.8). Reactions were incubated at 60°C for various times and were stopped by adding one volume of universal buffer (28.60 mM citric acid, 22.35 mM KH_2PO_4, 28.60 mM H_3BO_3, 25.54 mM sodium barbital, pH 11). Reactions catalysed by $AbfD3\text{-}Ala^{28}$ were performed using 3.6×10^{-3} IU (approximately 6 μg) of enzyme.

Separation of the reaction species was performed on a CarboPac PA-1 column (250×4.5 mm) equipped with the appropriate guard column. Elution was achieved by applying a gradient of 1 M sodium acetate in 0.1 M NaOH at a flow rate of 1 ml/min [21]. Purified and structurally characterised α-L-arabinofuranosyl-1,2-arabinofuranoside was used as the chromatographic standard.

3. Results and discussion

3.1. Primary sequence analysis

The Structure Prediction Meta Server was used in order to perform a variety of prediction analyses on the entire AbfD3 sequence (495 amino acids). Fig. 1 shows a summary of the major results that were obtained. Nine high-scoring (3D Jury score >100) potential structural homologues were identified by the various remote servers. All of these hits were GH-A enzymes belonging exclusively to families 1 or 5. The major activities represented by this group of enzymes were cellulase, β-glucanase (exo and endo), β-glucosidase and mannanase. The top three hits shown in Fig. 1 (3D Jury scores >115) were two family 1 β-glucosidases from *Bacillus polymyxa* (gi|80132) [22,23] and *Bacillus circulans* sp. *alkalophilus* (Q03506) [24,25] respectively and a family 5 β-(1,3)-exoglucanase from *Candida albicans* (P29717) [11,26]. Closer examination of the results from the Meta Server analysis revealed that a major region of the AbfD3 sequence (approximately residues 40–300) and its corresponding predicted secondary structure align rather well with most of the actual secondary structure elements which form the $(\beta/\alpha)_8$ barrels of all three of these enzymes. Likewise, the two major GH-A motifs for the catalytic acid-base, Asn-Glu (residues 175–176 in AbfD3), and the nucleophile, Glu^{298}, are perfectly aligned, as is the conserved tyrosine (Tyr^{242} in AbfD3) motif at the C-terminal side of β6. However, with regard to the structural arrangement of the β8–α8 region, these data reveal that the AbfD3 sequence is unlikely to adopt a family 1-type arrangement. In contrast, the predicted structure for AbfD3 does resemble that of the family 5 β-(1,3)-exoglucanase from *C. albicans*, although the highly conserved Trp motif (Trp^{363} in exoglucanase) appears to be absent or replaced by a Tyr. Furthermore, the results of 3D Jigsaw analysis indicate that, in AbfD3, α8 is immediately followed by a C-terminal domain (307–495), composed mainly of β-strands, which resembles

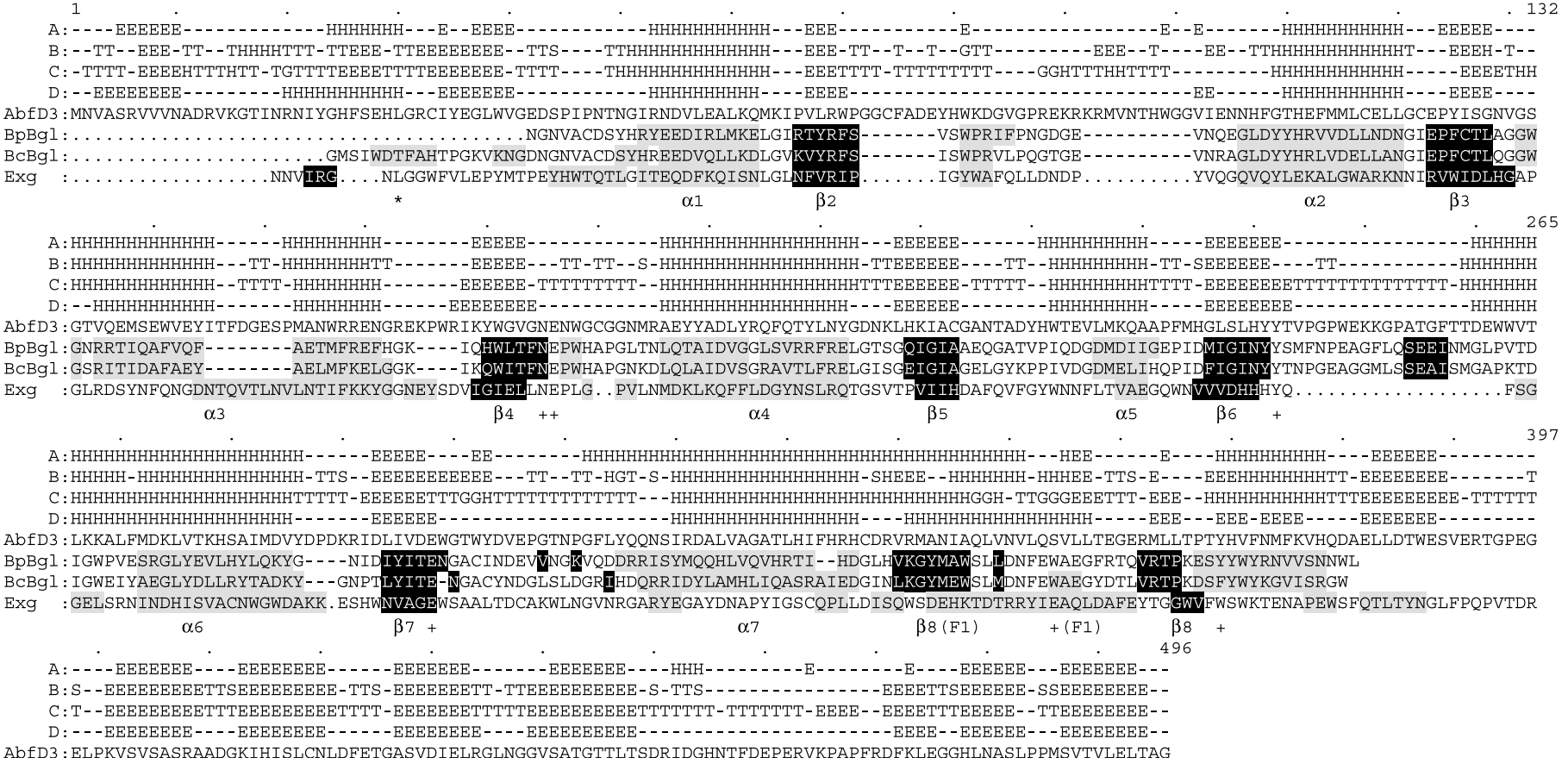

Fig. 1. Secondary structure prediction and structural homology searching. A (SAMT99), B (SAMT02-DSSP), C (SAMT02-STRIDE) and D (Profsec) are the results of secondary structure predictions for AbfD3. The symbols used are: H, α-helix; E, extended strand in β-ladder; G, 3/10 helix; T, H-bonded turn; S, bend. The wild type sequence (AbfD3, O69262) is aligned with three top-scoring structural homologues: β-glucosidases from *B. polymyxa* (gi|80132) and *B. circulans* sp. *alkalophilus* (Q03506) and a β-1,3-exoglucanase from *C. albicans* (P29717). Numbering corresponds to AbfD3. Grey and black shading indicate α-helix and β-strand structure respectively. The β8 configuration for family 1 enzymes is indicated by (F1). + indicates the positions of conserved GH-A motifs and * indicates the position of Glu^{28} in AbfD3.

Table 1
Comparison of the kinetic parameters of the hydrolysis of *p*NP-ara*f* catalysed by wild type AbfD3 or one of the 19 mutant variants

Mutation	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (mM^{-1} s^{-1})
Wild type	0.68 ± 0.01	4060 ± 150.27	5970.588 ± 125.96
His^{28}	0.73 ± 0.03	4500 ± 155.26	6164.384 ± 121.62
Ser^{28}	0.87 ± 0.00	3600 ± 141.01	4137.931 ± 123.52
Tyr^{28}	0.95 ± 0.01	2000 ± 120.51	2105.263 ± 96.82
Asp^{28}	0.46 ± 0.01	1.01 ± 0.02	2.19 ± 0.11
Cys^{28}	1.06 ± 0.00	1.10 ± 0.09	1.04 ± 0.04
Gly^{28}	1.12 ± 0.01	0.82 ± 0.05	0.73 ± 0.01
Val^{28}	1.38 ± 0.08	1.00 ± 0.07	0.72 ± 0.09
Leu^{28}	1.37 ± 0.02	0.95 ± 0.08	0.69 ± 0.05
Ala^{28}	0.98 ± 0.01	0.59 ± 0.04	0.60 ± 0.08
Lys^{28}	1.3 ± 0.00	0.64 ± 0.04	0.49 ± 0.06
Met^{28}	2.07 ± 0.03	0.80 ± 0.03	0.39 ± 0.03
Ile^{28}	2.40 ± 0.07	0.76 ± 0.05	0.32 ± 0.02
Thr^{28a}	32.7 ± 0.40	0.5 ± 0.03	0.02 ± 0.01
Phe^{28a}	26.81 ± 0.51	0.42 ± 0.03	0.02 ± 0.01
Trp^{28a}	22.63 ± 0.43	0.48 ± 0.03	0.02 ± 0.01
Pro^{28a}	30.5 ± 0.68	0.60 ± 0.09	0.02 ± 0.01
Arg^{28a}	20.28 ± 0.32	0.51 ± 0.05	0.03 ± 0.00
Asn^{28a}	33.41 ± 0.062	0.38 ± 0.01	0.01 ± 0.01
Gln^{28a}	32.97 ± 0.78	0.43 ± 0.04	0.01 ± 0.01

[a]Unusually high K_M value, precluded precise determination of kinetic parameters.

the immunoglobulin-like domains which have been identified in many glycosyl hydrolases [27–29] and notably in GH-A β-galactosidases [30]. In the extreme N-terminal region, AbfD3 is much shorter than the other three sequences. Consequently, the homologous sequences were truncated and no alignment was therefore deduced for the N-terminal AbfD3 sequence (region 1–37) which contains the functionally important Glu^{28}. In order to further investigate the significance of these findings, motif discovery analysis in family 51 sequences was performed using the MEME tool. This revealed four motifs which were present in the vast majority of the 43 sequences. Motifs 1 and 4 (AbfD3 residues 51–79 and 157–177 respectively) were found in all the sequences, whereas motifs 2 and 3 (AbfD3 residues 18–39 and 329–357) were present in all arabinosidase sequences (41 sequences), but absent in the sequences of the family 51 endoglucanases from *Fibrobacter succinogenes* [31] and *Alicyclobacillus acidocaldarius*. Interestingly, motif 2 covers the N-terminal region containing Glu^{28} while motif 3 covers the C-terminal region which precedes the putative β8.

3.2. Site-saturation mutagenesis at position 28

In order to further investigate the specific role of Glu^{28} in AbfD3, site-saturation mutagenesis was performed in order to generate all of the possible amino acid substitutions at position 28. For rapid identification of mutants displaying altered catalytic properties, initial screening of all mutants was performed by measuring their specific activities towards a variety of *p*NP-glycosides. Following this crude selection, Michaelis–Menten parameters were determined for all mutants in the presence of *p*NP-ara*f* (Table 1) and for selected mutants in the presence of other *p*NP-glycosides (Table 2).

With regard to the hydrolysis of *p*NP-ara*f*, three mutants (AbfD3-His^{28}, AbfD3-Ser^{28} and AbfD3-Tyr^{28}) displayed significant activity. The most active of these was AbfD3-His^{28}, whereas the previously reported conservative mutation, $Glu^{28} \rightarrow Asp$ [13], led to a significant loss of activity, despite an improvement in substrate binding. AbfD3-His^{28} displayed wild type values for both K_M and k_{cat}/K_M, while a slight increase in k_{cat} compared to the wild type enzyme indicated that this side chain alteration is favourable for transition state formation and/or stabilisation. For the other two mutants, AbfD3-Ser^{28} and AbfD3-Tyr^{28}, the side chain alterations provoked slightly different effects. The $Glu^{28} \rightarrow Tyr$ substitution caused alterations of similar magnitude in both the K_M and k_{cat} values (40% increase and 50% decrease respectively), whereas the effect of the $Glu^{28} \rightarrow Ser$ substitution had a more marked effect on the k_{cat} value (28% decrease compared to an 11% increase in K_M). Among the remaining, weakly active mutants, two different consequences of mutation could be distinguished. Seven mutations, including the isosteric $Glu \rightarrow Gln^{28}$, provoked severe perturbations in K_M values which precluded accurate determination of the kinetic parameters, whereas others displayed reasonable K_M values, but weak k_{cat} values.

Reactions performed in the presence of *p*NP-glycosides other than *p*NP-ara*f* revealed that the mutant AbfD3-His^{28} displayed a significant increase in activity towards *p*NP-ara*p*. This change in activity towards the conformationally more rigid pyranose form of L-arabinose, in which C4 displays an axial secondary alcohol, was due to a major improvement in the formation of the enzyme–substrate complex (20-fold decrease in K_M). In contrast, during the course of the initial activity screening, none of the mutants appeared to display improved activity towards *p*NP-xyl*p*, which is the C4 epimer of *p*NP-ara*p*. Several mutants showed greater activity towards *p*NP-rib*f* when compared to the wild type enzyme. In the case of AbfD3-Asn^{28}, which displayed the most significant improvement in activity towards *p*NP-rib*f*, this change in stereospecificity was brought about by a major improvement in K_M. Indeed, the affinity of this mutant for *p*NP-rib*f* was higher than that measured for wild type AbfD3 in the presence of *p*NP-ara*f*. However, despite stabilisation of the ground state enzyme–substrate complex, AbfD3-Asn^{28} is nevertheless a

Table 2
Kinetic parameters of reactions catalysed by wild type AbfD3 and selected mutants in the presence of various *p*NP-glycosides

Reaction	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (mM^{-1} s^{-1})
AbfD3+*p*NP-ara*p*	27.4 ± 0.43[a]	0.37 ± 0.03[a]	0.01[a]
AbfD3-His^{28}+*p*NP-ara*p*	1.34 ± 0.02	1300 ± 136.10	970 ± 110.25
AbfD3+*p*NP-rib*f*	1.94 ± 0.09	197 ± 4.21	101.55 ± 2.54
AbfD3-Asn^{28}+*p*NP-rib*f*	0.41 ± 0.08	233 ± 28.17	568.46 ± 106.28
AbfD3-Arg^{28}+*p*NP-rib*f*	0.88 ± 0.08	272 ± 5.13	308.38 ± 17.01
AbfD3-Trp^{28}+*p*NP-rib*f*	0.87 ± 0.05	262 ± 15.95	301.46 ± 4.80
AbfD3-Cys^{28}+*p*NP-rib*f*	1.04 ± 0.12	256 ± 5.00	247.45 ± 27.96
AbfD3+*p*NP-glc*p*	1.82 ± 0.2	169.40 ± 5.46	94.00 ± 12.06
AbfD3-Asp^{28}+*p*NP-glc*p*	3.39 ± 0.13	991.63 ± 9.56	292.46 ± 8.69

[a]Unusually high K_M value, precluded precise determination of kinetic parameters.

poor catalyst since its k_{cat} value (233 s^{-1}) in the presence of *p*NP-rib*f* was similar to that of wild type AbfD3 (197 s^{-1}). Other mutants active towards *p*NP-rib*f*, AbfD3-Trp^{28}, AbfD3-Arg^{28} and AbfD3-Cys^{28}, follow this trend, displaying improved K_M values and only very modest improvements in k_{cat}. Interestingly, like AbfD3-Asn^{28}, AbfD3-Trp^{28} and AbfD3-Arg^{28} (but not AbfD3-Cys^{28}) are also among those mutants which display the least affinity (K_M values >20 mM) for *p*NP-ara*f*. Finally, mutant AbfD3-Asp^{28} ($k_{cat}/K_M = 292.46 \pm 8.69$) was significantly more active on *p*NP-glc*p* than wild type AbfD3 (k_{cat}/K_M 94.00 ± 12.06). This result is particularly interesting because inspection of the kinetic parameters reveals that the increase in activity is due to an increase in k_{cat} and not, as one might suppose, a decrease in K_M. In comparison in *p*NP-ara*f* hydrolysis, the Glu$^{28} \rightarrow$ Asp substitution provoked a 29% reduction in K_M and a radical decrease (4000-fold) in k_{cat} [13]. Therefore, it would appear that while the presence of the shorter Asp side chain at position 28 is unfavourable for the binding of the glucosyl moiety, it is favourable with regard to catalysis.

Overall, if one considers the strict conservation of Glu^{28} in family 51 arabinofuranosidases, the success of the Glu$^{28} \rightarrow$ His (and to a lesser extent Ser and Tyr) substitution is rather unexpected. Perhaps similar volume occupation [32] and water-accessible surface areas [33] could partly explain the functional feasibility of the His substitution. However, unlike Glu, His displays a lower propensity for hydrogen bond acceptance and therefore, in AbfD3-His^{28}, one would expect that hydrogen bonding patterns are disturbed. Although such changes are not apparent in the data obtained for *p*NP-ara*f* hydrolysis, the altered specificity towards *p*NP-ara*p* is a clear indication of reorganised hydrogen bond patterns in the -1 subsite which modulate specificity towards the sugar C4-OH. This is further evidenced by the data obtained for AbfD3-Asn^{28}. In this case, it is likely that residue 28 is located in the -1 subsite where it interacts with the ribofuranose moiety. Compared to α-L-arabinofuranose, β-D-ribofuranose differs at both C3 and C4, therefore one may suppose that the presence of Asn at position 28 allows improved substrate binding, perhaps via new hydrogen bonds to the C3 and C4 hydroxyl groups. Significantly, in the structural homologues which were revealed by bioinformatics analysis, analogous situations have been identified. In the family 1 glycosidases the sugar C3- and C4-OH groups are hydrogen-bonded by a strictly conserved glutamine (Gln^{20} and Gln^{39} in β-glucosidases from *B. polymyxa* and *Spodoptera frugiperda* respectively) which is located at the C-terminal side of β1 [23,34]. In family 5 exoglucanases, a similar function is played by a conserved glutamate (Glu^{27} in the exoglucanase from *C. albicans*), which is not present in family 5 cellulases. This residue, which is located between β1 and α1, is orientated towards the active site where it forms vital hydrogen bonds to both the C3- and C4-OH groups of the glucosyl moiety [11].

3.3. Transglycosidase activity of AbfD3-Ala^{28}

Previously, in the presence of *p*NP-ara*f*, AbfD3 has been shown to perform a transglycosylation reaction which produces a mixture of α-L-arabinofuranosyl-(1,2), -(1,3) and -(1,5)-arabinofuranosides [3]. Therefore, in order to complete our analysis of the different AbfD3-Xxx^{28} mutants, each mutant was tested for its ability to catalyse this reaction. Surprisingly, despite the very low catalytic efficiency of many of the mutants, thin layer chromatographic analysis (data not shown) indicated that almost all of the mutants could perform transglycosylation. In most cases, similar R_f values indicated that the reaction products were identical to those produced in the wild type AbfD3-catalysed reaction. Following this observation, the transglycosylation reactions catalysed by wild type AbfD3 and the hydrolytically impaired mutant AbfD3-Ala^{28} (10^4-fold reduction in k_{cat}/K_M) were quantitatively compared by monitoring the production of the major α-L-arabinofuranosyl-1,2-arabinofuranoside product (Fig. 2). Initially, the rate of AbfD3-catalysed reaction was higher than that of AbfD3-Ala^{28} and maximum product yield (6%) was achieved within 30 s. However, beyond this point secondary hydrolysis predominated and the product of the wild type AbfD3-catalysed reaction rapidly disappeared and was no longer detectable after 10 min. In contrast, due to the extremely low level of hydrolytic activity in the AbfD3-Ala^{28}-catalysed reaction, product slowly accumulated over a longer period, with maximum yield (6%) being achieved after 10 min. Over longer incubation periods, the product was hydrolysed but at a very slow rate. After an 8 h incubation period approximately 50% of the product was still present.

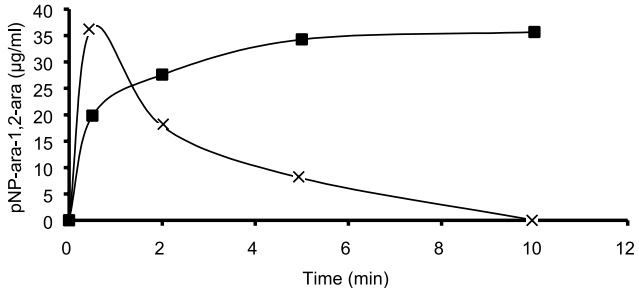

Fig. 2. Time-dependent production of α-L-arabinofuranosyl-1,2-arabinofuranoside. The progress of transglycosylation reactions catalysed by wild type AbfD3 (×) or AbfD3-Ala^{28} (■) was monitored by determining the quantity of α-L-arabinofuranosyl-1,2-arabinofuranoside produced in discontinuous assays using HPAEC-PAD.

Appealingly, these results are reminiscent of the behaviour of mutant forms of the β-glycosidase from *Sulfolobus solfataricus* [35]. When Glu^{432}, a -1 subsite residue which provides hydrogen bonds to the C4- and C6-OH groups, was replaced by other amino acids, the hydrolytically impaired mutants maintained their ability to perform transglycosylation. In contrast to glycosynthases [36,37], enzymes lacking their catalytic nucleophile, synthetically able mutant glycosidases such as those described in this work, or those described by Corbett et al. [35], constitute an interesting mechanistic phenomenon which has not been considered in any great detail. However, glycone sugar distortion during catalysis is now well established [38,39]. Therefore, if Glu^{28} in wild type AbfD3 and Glu^{432} in SSβG mediate substrate recognition via strong hydrogen bonds, these residues might play a vital role in catalysis by inducing a suitable saccharide conformation/orientation in the subsite -1. Such a mechanism has already been invoked to explain another analogous situation which involves the family 13, clan GH-B human pancreatic α-amylase (HPA) [40]. When Asp^{300} in HPA was replaced by Asn, the hydrolytically diminished HPA mutant conserved its ability to perform transglycosylation. Based on comparison with cyclodextrin glycosyltransferase [41], the authors suggested that Asp^{300} might participate in pyranose ring deformation in sub-

site −1 via strong hydrogen bonds to the hydroxyl group at C2. Such a deformation would favour H_2O-mediated nucleophilic attack at the anomeric centre. Therefore it was concluded that in the mutated enzyme, correct ring deformation would not occur and thus deglycosylation would proceed using the superior sugar acceptor.

4. Conclusion

Although a crystallisation note has been published recently [42], no structural models for family 51 are currently available in the Brookhaven database. Fortunately, using bioinformatics analyses and experimental approaches we have revealed that Glu^{28} is contained within an arabinofuranosidase-associated motif and is functionally homologous to a conserved residue found in exo-enzymes from both family 1 and, especially, family 5. In these enzymes, a Gln (family 1) or a Glu (family 5), located between β1 and α1, participates in the −1 subsite structure where it plays a vital role in substrate fixation and specificity. Interestingly, rather like family 5 endo-acting enzymes, family 51 endoglucanases do not possess an equivalent of Glu^{28}. This suggests that Glu^{28} contributes to the overall pocket structure which determines exo-activity in family 51 arabinofuranosidases. Finally, with regard to the interesting glycosynthetic properties of mutated AbfD3, our results open exciting prospects for the creation of novel enzymes able to catalyse the synthesis of arabino- and ribo-oligosaccharides.

Acknowledgements: The authors thank the technical staff of the laboratory for their skilful work and Dr V. Tran for his critical reading of the manuscript. This work was partly funded by the GlycoVal programme of Europol'Agro.

References

[1] Saha, B.C. (2003) J. Ind. Microbiol. Biotechnol. 30, 279–291.
[2] Saha, B.C. (2000) Biotechnol. Adv. 18, 403–423.
[3] Rémond, C., Ferchichi, M., Aubry, N., Plantier-Royon, R., Portella, C. and O'Donohue, M.J. (2002) Tetrahedron Lett. 43, 9653–9655.
[4] Zechel, D.L. and Withers, S.G. (2000) Acc. Chem. Res. 33, 11–18.
[5] Sinnott, M.L. (1990) Chem. Rev. 90, 1171–1202.
[6] Henrissat, B. and Bairoch, A. (1996) Biochem. J. 316, 695–696.
[7] Henrissat, B., Callebaut, I., Fabrega, S., Lehns, P., Mornon, J.-P. and Davies, G. (1995) Proc. Natl. Acad. Sci. USA 92, 7090–7094.
[8] Jenkins, J., Lo Leggio, L., Harris, G. and Pickersgill, R. (1995) FEBS Lett. 362, 281–285.
[9] Sakon, J., Adney, W.S., Himmel, M.E., Thomas, S.R. and Karplus, P.A. (1996) Biochemistry 35, 10648–10660.
[10] Hogg, D., Woo, E.J., Bolam, D.N., McKie, V.A., Gilbert, H.J. and Pickersgill, R.W. (2001) J. Biol. Chem. 276, 31186–31192.
[11] Cutfield, S.M., Davies, G.J., Murshudov, G., Anderson, B.F., Moody, P.C.E., Sullivan, P.A. and Cutfield, J.F. (1999) J. Mol. Biol. 294, 771–783.
[12] Juers, D.H., Huber, R.E. and Matthews, B.W. (1999) Protein Sci. 8, 122–136.
[13] Debeche, T., Bliard, C., Debeire, P. and O'Donohue, M.J. (2002) Protein Eng. 15, 21–28.
[14] Shallom, D., Belakhov, V., Solomon, D., Shoham, G., Baasov, T. and Shoham, Y. (2002) J. Biol. Chem. 277, 43667–43673.
[15] Shallom, D., Belakhov, V., Solomon, D., Gilead-Gropper, S., Baasov, T., Shoham, G. and Shoham, Y. (2002) FEBS Lett. 514, 163–167.
[16] Karplus, K., Barrett, C. and Hughey, R. (1999) Bioinformatics 14, 846–856.
[17] Rost, B. (2000).
[18] Ginalski, K., Elofsson, A., Fischer, D. and Rychlewski, L. (2003) Bioinformatics 19, 1015–1018.
[19] Bailey, T.L. and Elkan, C. (1994) in: Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology (Altman, R., Brutlag, D., Karp, P., Lathrop, R. and Searls, D., Eds.), pp. 28–36, AAAI Press, Menlo Park, CA.
[20] Debeche, T., Cummings, N., Connerton, I., Debeire, P. and O'Donohue, M.J. (2000) Appl. Environ. Microbiol. 66, 1734–1736.
[21] Beldman, G., Searle-van Leeuwen, M.J.F., De Ruiter, G.A., Siliha, H.A. and Voragen, A.G.J. (1993) Carbohydr. Pol. 20, 159–168.
[22] Gonzalez-Candelas, L., Ramon, D. and Polaina, J. (1990) Gene 95, 31–38.
[23] Sanz-Aparicio, J., Hermoso, J.A., Martinez-Ripoll, M., Lequerica, J.L. and Polaina, J. (1999) J. Mol. Biol. 275, 491–502.
[24] Paavilainen, S., Hellman, J. and Korpela, T. (1993) Appl. Environ. Microbiol. 59, 927–932.
[25] Hakulinen, N., Paavilainen, S., Korpela, T. and Rouvinen, J. (2000) J. Struct. Biol. 126, 69–79.
[26] Chambers, R.S., Broughton, M.J., Cannon, R.D., Carne, A., Emerson, G.W. and Sullivan, P.A. (1993) J. Gen. Microbiol. 139, 325–334.
[27] Little, E., Bork, P. and Doolittle, R.F. (1994) J. Mol. Evol. 39, 631–643.
[28] Perrakis, A., Ouzounis, C. and Wilson, K.S. (1995) Fold. Des. 2, 291–294.
[29] Dauter, Z., Dauter, M., Brzozowski, A.M., Christensen, S., Borchert, T.V., Beier, L., Wilson, K.S. and Davies, G.J. (1999) Biochemistry 38, 8385–8392.
[30] Jacobson, R.H., Zhang, X.J., DuBose, R.F. and Matthews, B.W. (1994) Nature 369, 761–766.
[31] Malburg, S., Malburg Jr., L., Liu, T., Iyo, A. and Forsberg, C. (1997) Appl. Environ. Microbiol. 63, 2449–2453.
[32] Zamyatnin, A.A. (1972) Prog. Biophys. Mol. Biol. 24, 109–123.
[33] Chothia, C. (1976) J. Mol. Biol. 105, 1–12.
[34] Marana, S.R., Terra, W.R. and Ferreira, C. (2002) Eur. J. Biochem. 269, 3705–3714.
[35] Corbett, K., Fordham-Skelton, A.P., Gatehouse, J.A. and Davis, B.G. (2001) FEBS Lett. 509, 355–360.
[36] Mayer, C., Zechel, D.L., Reid, S.P., Warren, R.A. and Withers, S.G. (2000) FEBS Lett. 466, 40–44.
[37] Mackenzie, L.F., Wang, Q., Warren, R.A.J. and Withers, S.G. (1998) J. Am. Chem. Soc. 120, 5583–5584.
[38] White, A. and Rose, D.R. (1997) Curr. Opin. Struct. Biol. 7, 645–651.
[39] Vasella, A., Davies, G.J. and Böhm, M. (2002) Curr. Opin. Chem. Biol. 6, 619–629.
[40] Rydberg, E.H., Li, C., Maurus, R., Overall, C.M., Brayer, G.D. and Withers, S.G. (2002) Biochemistry 41, 4492–4502.
[41] Uitdehaag, J.C.M., Mosi, R., Kalk, K.H., van der Veen, B.A., Dijkhuizen, L., Withers, S.G. and Dijkstra, B.W. (1999) Nat. Struct. Biol. 6, 432–436.
[42] Hövel, K., Shallom, D., Niefind, K., Baasov, T., Shoham, G., Shoham, Y. and Schomburg, D. (2003) Acta Crystallogr. D Biol. Crystallogr. 59, 913–915.