2464 *Biochemistry* **2010,** *49,* 2464–2474
DOI: 10.1021/bi100036f

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Circular Permutation of *Bacillus circulans* Xylanase: A Kinetic and Structural Study[†]

Stephan Reitinger,[‡,∥] Ying Yu,[§] Jacqueline Wicki,[‡,∥] Martin Ludwiczek,[⊥] Igor D'Angelo,[⊥] Simon Baturin,[‡,⊥] Mark Okon,[‡,⊥] Natalie C. J. Strynadka,[⊥,#,◇] Stefan Lutz,[§] Stephen G. Withers,*[,‡,∥,⊥] and Lawrence P. McIntosh*[,‡,∥,⊥,#]

[‡]*Department of Chemistry, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1,* [§]*Department of Chemistry, Emory University, 1515 Dickey Drive, Atlanta, Georgia 30322,* [∥]*Centre for High Throughput Biology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1,* [⊥]*Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z3,* [#]*Michael Smith Laboratory, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4, and* [◇]*Centre for Blood Research, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z3*

Received January 11, 2010; Revised Manuscript Received February 16, 2010

ABSTRACT: The 20 kDa *Bacillus circulans* Bcx is a well-studied endoxylanase with a β-jellyroll fold that places its N- and C-termini in salt bridge contact. Initial experiments verified that Bcx could be circularly permuted by PCR methods to introduce new termini in loop regions while linking its native termini directly or via one or two glycines. Subsequently, a library of circular permutants, generated by random DNase cleavage of the circularized *Bcx* gene, was screened for xylanase activity on xylan in Congo Red-stained agar. Analysis of 35 unique active circular permutants revealed that, while many of the new termini were introduced in external loops as anticipated, a surprising number were also located within β-strands. Furthermore, several permutations placed key catalytic residues at or near the new termini with minimal deleterious effects on activity and, in one case, a 4-fold increase. The structure of one permutant was determined by X-ray crystallography, whereas three others were probed by NMR spectroscopy. These studies revealed that the overall conformation of Bcx changed very little in response to circular permutation, with effects largely being limited to increased local mobility near the new and the linked old termini and to a decrease in global stability against thermal denaturation. This library of circularly permuted xylanases provides an excellent set of new start points for directed evolution of this commercially important enzyme, as well as valuable constructs for intein-mediated replacement of key catalytic residues with unnatural analogues. Such approaches should permit new insights into the mechanism of enzymatic glycoside hydrolysis.

The endoxylanase Bcx[1] from *Bacillus circulans* is a 20 kDa glycoside hydrolase that has served as an excellent model system for understanding enzyme mechanisms in general and glycosidase mechanisms in particular. The insights obtained not only are of academic interest but also have substantial commercial importance since xylanases are widely used in commercial food preparation, as well as in the kraft pulp and paper industry and potentially in biofuel generation. Consequently, Bcx and other closely related GH family 11 xylanases (*1*) have been subjected to numerous studies on the structural and functional level (*2*).

Bcx degrades xylan through a retaining, double-displacement mechanism involving a glycosyl-enzyme intermediate (*3*−*6*). Formation and hydrolysis of this covalent intermediate proceeds via oxocarbenium ion-like transition states, with the assistance of two key catalytic glutamic acids. One residue (Glu78) acts as a nucleophile, attacking the anomeric carbon of the substrate. The other (Glu172) functions as a general acid/base catalyst, assisting in the formation of the intermediate by proton donation, as well as in the subsequent hydrolysis step by deprotonating the attacking nucleophilic water. The active site residues of Bcx, as well as the dynamic and electrostatic properties of the glycosyl-enzyme intermediate, have been characterized extensively by X-ray crystallography and NMR spectroscopy (*5*, *7*−*12*).

Studies of Bcx, along with parallel investigations on another xylanase of a different tertiary fold (the GH family 10 *Cellulomonas fimi* endoxylanase Cex) and continuing collaborative computational and heavy atom isotope effect studies, have provided particularly detailed insights into fundamental mechanisms of carbohydrate degradation (*6*, *13*, *14*). With this same set of mechanistic tools, analyses have been performed on interesting mutants derived from site-directed and random mutagenesis methods, thereby providing additional insights into structure/function relationships. However, further understanding could be

[†]This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada to L.P.M. and S.G.W. and from the National Science Foundation (CBET-0730312) and Petroleum Research Fund of the American Chemical Society (PRF 47135-AC1) to S.L. Instrument support was provided by the Canadian Institutes for Health Research, the Canadian Foundation for Innovation, the British Columbia Knowledge Development Fund, the UBC Blusson Fund, and the Michael Smith Foundation for Health Research. S.R. and M.L. acknowledge Erwin Schrödinger Fellowships from the Austrian Science Fund. I.D. acknowledges postdoctoral support from the Michael Smith Foundation for Health Research. S.G.W. thanks the Canada Research Chairs program for salary support.

*To whom correspondence should be addressed: L.P.M.: e-mail, mcintosh@chem.ubc.ca; phone, 604-822-3341; fax, 604-822-5227. S.G.W.: e-mail, withers@chem.ubc.ca; phone, 604-822-3042; fax, 604-822-8869.

[1]Abbreviations: Bcx, *Bacillus circulans* xylanase; cfu, colony forming unit; cp, cyclic permutant; 2F-DNPX2, 2,4-dinitrophenyl 2-deoxy-2-fluoro-β-xylobioside; HSQC, heteronuclear single-quantum correlation; ONPX2, 2-nitrophenyl β-xylobioside; DNPX2, 2,5-dinitrophenyl β-xylobioside; T_m, midpoint unfolding temperature.

 Published on Web 02/17/2010

obtained if these enzymes were amenable to circular permutation, such that new start and finish positions for the protein sequences could be generated. Not only would this allow interrogation of the role of the dynamics of the "old" and "new" termini in enzyme structure and function but it would also facilitate the specific incorporation of unnatural or labeled amino acids into the protein by semisynthetic approaches, for example, using intein methodologies (*15*). Bcx is particularly well suited to circular permutation since its β-jellyroll structure places its native N- and C-termini next to each other, interacting directly via a salt bridge.

Systematic circular permutation of genes has emerged as a useful tool to conduct studies on polypeptide folding and stability (*16*−*20*). The overall tertiary structure of a protein is usually retained, since, with the exception of possible residues introduced to bridge the native termini, the amino acid sequence is not changed but simply rearranged. Nevertheless, circular permutation can significantly affect stability, dynamics, and function (*21*−*27*). According to the Circular Permutation Database (CPDB), more than 4000 naturally occurring or artificially generated circularly permuted proteins have been identified to date (*28*). In the first reported circular permutation experiment, the termini of bovine pancreatic trypsin inhibitor (BPTI) were linked chemically, and the circular polypeptide chain was subsequently cleaved by proteolysis at a specific reactive site (*29*). In recent years, recombinant DNA technologies have been used to generate circularly permuted variants of proteins such as phosphoribosyl anthranilate isomerase (*30*), β-glucanases (*31*, *32*), aspartate transcarbamoylase (*33*, *34*), dual chain avidin (*35*), T4 lysozyme (*25*), and *Candida antarctica* lipase B (*24*, *27*, *36*). A sortase-catalyzed transpeptidation approach to synthesize circular proteins was also recently reported (*37*).

A particularly attractive approach, especially for the objectives of our mechanistic study, involves the generation of libraries of circular permutants via random cleavage of the circularized gene. Expression of the resultant permuted gene library in *Escherichia coli*, followed by activity-based screening, can lead to the identification of active permuted enzymes. Application of this approach to *C. antarctica* lipase B not only generated a large library of functional permutants but also identified some with substantially (up to 175-fold) increased activity (*36*). In this paper, we present the enzymatic, structural, and dynamic characterization of a set of active, circularly permutated Bcx variants generated by this random cleavage protocol. Most significantly, new termini could be introduced into the active site of the enzyme without dramatically altering its catalytic activity, thus establishing the framework for future mechanistic studies, as proposed above.

EXPERIMENTAL PROCEDURES

PCR Cloning of "Designed" Permutants. Initial clones encoding Bcx permutants with a single glycine linker joining the native ends (Ala1 and Trp185) and with new termini introduced at positions 102 (cpG102^{G1}), 123 (cpA123^{G1}), or 139 (cpG139^{G1}) were generated by a four-primer PCR approach and placed in the pET16b vector (Novagen) using *Nco*I and *Xho*I restriction sites.[2] Subsequent clones with linkers of zero (cpA123^{G0}) and two glycine residues (cpA123^{G2}) were derived from the cpA123^{G1} template by PCR. Primer sequences are summarized in Supporting Information Table S1, and the amino acid sequences of all designed circular permutants are listed in Supporting Information Table S2.

Random Circular Permutation of Bcx. The random circular permutation of *Bcx* was performed as described previously for *C. antarctica* lipase B (*24*). Briefly, the gene encoding cpA123^{G2} (with a two-glycine linker, deletion of Arg122, and the Thr123Ala mutation) was permutated by PCR methodology to position the unique internal *Nsi*I restriction site at both ends of the gene sequence, creating *cpBcx_Nsi*I (Supporting Information Table S1). In preparation for the permutation experiment, *cpBcx_Nsi*I was cloned into the high-copy DNA plasmid pSTBlue (Novagen, Madison, WI). Following plasmid amplification in *E. coli* DH5α, the purified vector was digested with *Nsi*I, and the desired 560-bp DNA fragment was isolated via agarose gel electrophoresis. Next, the linear fragment was circularized by intramolecular ligation (2 μg/mL DNA) with 30 units/mL T4 DNA ligase (Promega, Madison, WI) in the manufacturer's buffer overnight at 16 °C. The reaction mixture was concentrated by ethanol precipitation, and the remaining linear DNA was eliminated by treatment with ExoIII (120 units/μg of DNA; Promega) for 30 min at 37 °C. After heat inactivation of ExoIII (10 min at 72 °C), the circular DNA was recovered via QIAquick purification (Qiagen, Valencia, CA).

The circularized DNA was linearized by limited DNase I digestion (0.0005 units/μg of DNA; Roche, Indianapolis, IN) in 50 mM Tris-HCl (pH 7.5) and 1 mM $MnCl_2$ at room temperature for 15 min. The reaction was quenched by addition of 10 μL of EDTA (0.5 M), followed by treatment with T4 DNA polymerase (1 unit/μg of DNA; Promega), 150 μM dNTPs, and T4 DNA ligase (2 units/μg of DNA; Promega) in T4 ligase buffer at ambient temperature for 1 h to repair DNA nicks and to create blunt ends. The resulting cpBcx library was purified by agarose gel electrophoresis and cloned into pET27-PP, predigested with *Pst*I and *Pac*I. The pET27-PP vector is derived from pET-27b (Novagen) and carries unique *Pst*I and *Pac*I restriction sites (underlined), as well as three stop codons (bold) in all three reading frames in the multiple cloning site. The new cleavage sites were introduced by primer overlap extension, using forward primer 5′CTG CTC CTC GCT GCC CAG CCG GCG ATG GCC TGC AGA TGG ATA TCG GAA TTA ATT CG3′ and reverse primer 5′GAT CTC GAG **TTA** GTT AGT **TAA** TTA AGC GGC CGC AAG CTT GTC GAC3′. In addition, the pET27-PP encodes a *pelB* leader sequence for protein secretion to the periplasmic space.

Library Screening Using a Congo Red Overlay Assay. The circularly permuted *Bcx* gene library was transformed into *E. coli* BL21(λDE3) and spread out on LB agar plates containing 30 μg/mL kanamycin. After 12 h incubation at 37 °C, colonies were replica-plated onto fresh LB agar media with 30 μg/mL kanamycin. While the master plate was stored at 4 °C, the replica plates were incubated for 3 h at 37 °C prior to overlaying them with molten agar containing 0.4% agar, 0.4% birchwood xylan (Sigma, St. Louis, MO), 30 μg/mL kanamycin, and 1 mM IPTG. Once the overlay agar had solidified, plates were incubated for another 8 h at 37 °C. To visualize xylanase activity, the agar plates were stained with 0.5% Congo Red solution for 15 min, followed by destaining with 1 M NaCl for 30 min (*38*). Colonies

[2]Nomenclature of the circular permutants: Termini of circular permutants were either joined directly (G0) or joined via the insertion of one (G1) or two (G2) glycine residues. All permutants and cpWT have the Arg122 deletion and the Thr123Ala mutation except cpG102^{G1} and cpG139^{G1}. Permutants, recloned to remove additional or replace missing residues at the new termini, are indicated by a prime symbol (e.g., cpN35$^{G2'}$).

expressing functional glycoside hydrolases could be identified by formation of a clearing zone surrounding the host cells. The corresponding colonies on the master plate were picked and regrown for DNA sequence analysis. For initial characterization of the resulting Bcx permutants, the pET27-PP constructs were used (Supporting Information Table S3). For subsequent detailed kinetic analysis, genes of selected permutants were subcloned into pET-21b via the *Nde*I and *Hin*dIII restriction sites. The PCR primers for this step were also designed to remove additional or replace missing residues at the new termini of the permutants (Supporting Information Table S4),

Protein Purification. Gene expression and protein purification were carried out as described (*12*, *39*). *E. coli* BL21(λDE3) cells were grown at 30 °C in LB media containing 100 μg/mL carbenicillin (for pET16b and pET21b constructs) or 50 μg/mL kanamycin (for pET27-PP constructs), respectively, or in M9 medium with 1 g/L $^{15}NH_4Cl$ or 1 g/L $^{15}NH_4Cl$ and 3 g/L [$^{13}C_6$]glucose (Spectra Stable Isotope Inc.). Protein expression was induced with 50 μM IPTG at $OD_{600} \sim 0.6$, and cells harvested 20 h later. For preliminary screening, proteins were purified with only SP-Sepharose (HiTrap, 5 mL column; GE Healthcare) ion-exchange chromatography in 20 mM MES buffer, pH 6.0, and eluted with a 0−1 M NaCl gradient. Their concentrations were determined by absorbance spectroscopy with a Unicam UV/vis spectrometer UV4 using the predicted ε_{280} value of 81790 M^{-1} cm^{-1} (*40*) and corrected by purity estimated from SDS−PAGE images taken and analyzed with AlphaImager and AlphaEase FC software (Alpha Innotech, Corp.). For all other experiments, the permutants were first purified by ion-exchange chromatography, followed by Sephacryl S-100 (HiPrep 16/60; GE Healthcare) size-exclusion chromatography using 20 mM MES buffer, pH 6.0, and 50 mM NaCl for activity assays or 10 mM phosphate buffer, pH 6.5, for structural analyses. In these cases, enzyme concentrations were determined by active site titration.

Active Site Titrations Using 2F-DNPX2. Enzyme (final concentration of 5−20 μM) was added to a solution of 0.39 mM 2,4-dinitrophenyl 2-deoxy-2-fluoro-β-xylobioside (2F-DNPX2), 20 mM MES, pH 6.0, and 50 mM NaCl at 22 °C. The 2,4-dinitrophenolate released, due to covalent inactivation of the active enzyme (*41*), was monitored at 400 nm. The increase in the A_{400} observed was corrected for the absorbance due to spontaneous hydrolysis of 2F-DNPX2 and enzyme-catalyzed turnover. From the net ΔA_{400}, which corresponds to the concentration of 2,4-dinitrophenolate released ($\Delta\varepsilon_{400} = 11.40$ mM^{-1} cm^{-1}, where $\Delta\varepsilon$ is the difference in molar absorptivity between 2,4-dinitrophenol and its corresponding xylobioside at pH 6.0), the concentration of active enzyme was determined.

Steady-State Enzyme Kinetics. Two aryl β-xylobiosides were used as substrates in the assays described below: 2,5-dinitrophenyl β-xylobioside (DNPX2), $\Delta\varepsilon_{440nm} = 3.57$ mM^{-1} cm^{-1}, pK_a of 2,5-dinitrophenol = 5.22, and 2-nitrophenyl β-xylobioside (ONPX2), $\Delta\varepsilon_{400nm} = 1.07$ mM^{-1} cm^{-1}, pK_a of 2-nitrophenol = 7.22 (*42*). All substrates were synthesized and characterized according to previously published procedures (*43*, *44*). Spectrophotometric assays were performed using either a Cary 4000 or Unicam UV4 spectrophotometer in 200 μL microblack-walled quartz cuvettes or 1000 μL disposable methacrylate cuvettes (*42*).

The second-order rate constants (k_{cat}/K_m) for the hydrolysis of ONPX2 and DNPX2 were determined from progress curves at low substrate concentrations using the substrate depletion method (*6*). Enzyme was added to reaction mixtures containing substrate concentrations less than $(1/5)K_m$, and the release of nitrophenolate was monitored until substrate depletion was observed. The change in absorbance with respect to time was fitted to a first-order rate equation using the program GraFit 5.0 (*45*).

Michaelis−Menten steady-state parameters for the hydrolysis of DNPX2 were determined at six different substrate concentrations ranging from 0.2 to 2 times the estimated K_m value. The relative insolubility of DNPX2 precluded study at higher substrate concentrations. From the experimental rate versus substrate concentration data, K_m and k_{cat} values were calculated directly using GraFit 5.0 (*45*).

Thermal Denaturation Measurements. Circular dichroism (CD) spectra were measured on a JASCO J-810 spectropolarimeter. A 1 mm path length cell containing protein at a concentration of 8 μM in 20 mM sodium phosphate buffer, pH 7.0, was used. The long-lived glycosyl-enzyme intermediate forms were generated by treatment with excess 2F-DNPX2 for at least 3 h. Thermal denaturation curves were recorded by monitoring the signal at 219 nm as a function of temperature, increasing at a rate of 1 °C/min. The midpoint unfolding temperature, T_m, was calculated by nonlinear least-squares data fitting to a standard equation describing a two-state conformational equilibrium (*46*).

X-ray Crystallography. Crystals of $cpA123^{G1}$ were grown at 4 °C using the hanging drop method by equilibrating 2 mL of purified protein solution (25−30 mg/mL) against an equal volume of 13−20% saturated $(NH_4)_2SO_4$ in 40 mM Tris-HCl, pH 8.0. Prior to data collection, all crystals were serially transferred for a few seconds into cryoprotectant solutions composed of mother liquor supplemented with 5−25% glycerol and then frozen in liquid nitrogen. X-ray data collection was performed under cryogenic conditions (100 K) using an in-house rotating anode X-ray generator (Cu Kα radiation $\lambda = 1.541$ Å). Data were recorded using an image plate detector and processed using HKL2000 (*47*). The wild-type Bcx coordinates 1HV1.pdb were used as a search model for molecular replacement with PHASER (*48*). The correct solution was then used for model rebuilding and structure refinement with CNS (simulated annealing) (*49*) and REFMAC (maximum likelihood functions) (*50*), alternating with manual adjustments using COOT (*51*). A random sample containing roughly 5% of the total number of reflections was excluded from the refinement and used for the calculation of the free *R* factor. Tight noncrystallographic symmetry restraints and geometry were maintained throughout all of the different steps of refinement and then partially relaxed at the final stage. Care was taken to avoid overfitting the structure by reducing the X-ray data weighting term during refinement. Water molecules were assigned at 3σ residual $F_o - F_c$ electron density areas located within 3 Å from the protein and confirmed by visual inspection. All models displayed acceptable stereochemical geometries, with >95% of the residues in the most favorable regions of the Ramachadran plot. Supporting Information Table S5 provides a summary of this crystallographic analysis. The final coordinates of $cpA123^{G1}$ were deposited in the RCSB Protein Data Bank under accession code 3LB9. Structural figures were made using PYMOL (*52*).

NMR Spectroscopy. NMR spectra were recorded at 25 °C with Varian Unity 500 MHz and cryoprobe-equipped Inova 600 MHz spectrometers. Data were processed with NMRpipe (*53*) and analyzed using Sparky (*54*). Samples of $^{13}C/^{15}N$-labeled cpWT (0.94 mM), $^{13}C/^{15}N$-labeled $cpN35^{G2'}$ (0.7 mM), and

Table 1: Characterization of Designed and Selected Random Circular Permutants of Bcx

enzyme	k_{cat}/K_m [b] (mM^{-1} s^{-1})	relative activity	T_m (°C) apoenzyme	T_m (°C) inactivated enzyme
WT	0.35[c]	1	62.7 ± 0.2	74.2 ± 0.6
cpG102^{G1}	0.44[c]	1.3	50.0 ± 0.2	72.3 ± 0.6
cpG139^{G1}	0.43[c]	1.2	47.6 ± 0.2	66.2 ± 0.6
cpA123^{G1} [a]	0.10[c]	0.3	53.2 ± 0.2	69.5 ± 0.6
cpA123^{G0} [a]	0.06[c]	0.2	54.8 ± 0.2	nd
cpA123^{G2} [a]	0.08[c]	0.2	55.8 ± 0.2	nd
WT	0.32[d]	1		
cpWT [a]	0.44[d]	1.4	56.0 ± 0.2	66.9 ± 0.6
cpN35$^{G2'}$ [a]	1.44[d]	4.5	50.4 ± 0.2	63.0 ± 0.4
cpY94$^{G2'}$ [a]	0.53[d]	1.7	52.3 ± 0.2	66.3 ± 0.9
cpY174$^{G2'}$ [a]	0.76[d]	2.4	45.9 ± 0.2	56.2 ± 0.4

[a]Enzymes are ΔArg122 and T123A. [b]Average errors in kinetic parameters are ±5−10%. [c]Kinetic parameters for the hydrolysis of ONPX2 were determined at 25 °C in 100 mM sodium phosphate buffer (pH 6.0) containing 0.1% (w/v) BSA. [d]Kinetic parameters for the hydrolysis of ONPX2 were determined at 22 °C in 20 mM MES (pH 6.0) and 50 mM NaCl containing 0.1% (w/v) BSA.

^{15}N-labeled cpY94$^{G2'}$ (0.04 mM) were in 10 mM sodium phosphate buffer, pH 6.5, with 5% D_2O lock solvent. The spectral assignments of main chain nuclei in cpWT and cpN35$^{G2'}$ were obtained from ^{15}N-HSQC, HN(CA)CB, CBCACONH, HNCO, HN(CA)CO, and ^{15}N-editted TOCSY- and NOESY-HSQC spectra (*55*), combined with knowledge of the previously reported assignments for the wild-type protein (*9*, *56*).

Backbone amide ^{15}N T_1 and T_2 relaxation and heteronuclear ^{1}H−^{15}N NOE experiments were performed with ^{13}C/^{15}N-labeled cpWT and cpN35$^{G2'}$ at 25 °C on an Inova 600 MHz spectrometer (*57*). The T_1 and T_2 values were determined using a nonlinear, least-squares fitting of the intensity of the cross-peaks to an exponential decay in the Sparky software. NOE values were obtained by comparing the ratios of cross-peak intensities with and without proton saturation. Anisotropic tumbling and internal dynamics parameters calculated from the relaxation data were processed according to the model-free formalism of Lipari and Szabo (*58*), using TENSOR2 (*59*).

RESULTS

Creation of Rational Design Permutants. As an initial feasibility study, three circular permutants (cpG102^{G1}, cpA123^{G1}, and cpG139^{G1}) were generated by rational design. The 2.7 Å distance between the native N- and C-termini of Bcx was bridged with a single glycine residue, and new termini were introduced at the indicated positions (Supporting Information Table S2). The sites were selected due to their locations in exposed surface loops, minimizing the potential risk for disruption of structure. Following expression and purification of the three permutants, steady-state kinetics revealed only minor differences in k_{cat}/K_m for hydrolysis of ONPX2 (0.3−1.3 fold) compared to wild-type enzyme (Table 1). Although all three variants were stable under laboratory conditions, thermal denaturation experiments demonstrated a substantial reduction in their midpoint temperatures of unfolding (T_m). For the apoproteins, the T_m values dropped by 10−15 °C, whereas for their covalently modified 2-deoxy-2-fluoro-β-xylobiosyl-enzyme adducts, the T_m values were lowered by 2−8 °C, respectively. Crystallographic studies of cpA123^{G1}, discussed below, confirmed the retention of the permutant's secondary and ternary structure relative to wild-type enzyme. Only minimal structural perturbations were observed at the site of the glycine linker, as well as the position of the new termini. These data demonstrated that Bcx is indeed amenable to circular permutation.

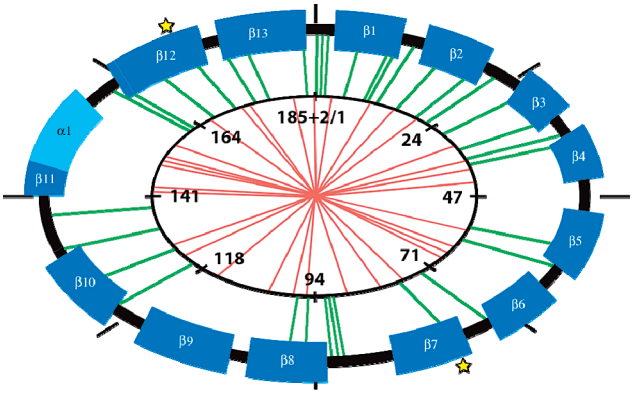

FIGURE 1: Circular permutation of Bcx with a two-glycine linker. The unbiased distribution of permutation sites over the entire length of Bcx in the naïve library was examined by DNA sequencing of 32 randomly selected members (red lines, inner circle). Functional screening of 3000 library members for xylanase activity yielded 35 candidates with termini in unique positions (green lines, outer circle). Secondary structure elements, as defined by Promotif (*65*), of the wild-type protein are marked (β-strands, dark blue; α-helix, light blue). The positions of the catalytic residues (Glu78 and Glu172) are indicated by yellow stars.

Optimization of Linker Length. The length and composition of the peptide linker connecting the native termini can have a significant effect on enzyme activity and stability. To explore the impact of truncation or insertion of additional amino acids in the linker region on the properties of Bcx, we chose cpA123^{G1} as a template for generating two additional variants. In one case, we deleted the linker and directly joined the native termini (cpA123^{G0}), and in the other, we extended the linker by a second glycine residue (cpA123^{G2}). Thermal denaturation curves revealed that cpA123^{G2} had a slightly higher T_m value than cpA123^{G0} and cpA123^{G1}, whereas the enzymatic activities of all three hydrolases measured with ONPX2 were essentially the same (Table 1). Based on its slightly higher stability, cpA123^{G2} was chosen as the template for the subsequent generation of random permutants.

Random Circular Permutation of Bcx. Extending our studies of circular permutated Bcx beyond the three designed variants, we applied a random circular permutation protocol for

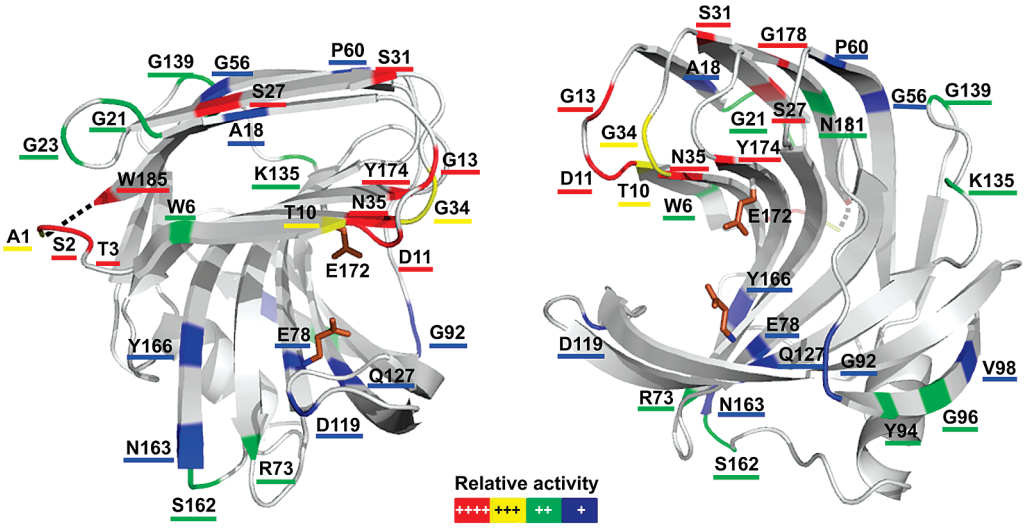

FIGURE 2: "Heat map" summary of the active Bcx circular permutants (Table 2) plotted on rotated structural views of the wild-type protein. The two glycines linking Ala1 and Trp185 are indicated as a dashed line. The positions of new N-termini are labeled, and a color code summarizes relative activity. The catalytic Glu78 and Glu172 are displayed in stick mode (brown).

generating a comprehensive combinatorial library of all possible Bcx permutants. Briefly, we self-ligated the gene encoding for cpA123^{G2} to create circular DNA and then relinearized the sequence by limited digestion with DNase I. The resulting library of permutated genes was cloned into a modified pET expression vector, allowing for blunt-end ligation of the gene inserts next to a start codon and stop codons in all three reading frames. This cloning strategy introduced two mutations: an extra methionine residue at the N-terminus, which is accommodated by deleting Arg122, and a substitution of Thr123 to Ala due to the *Pst*I restriction site. After transformation of the library into *E. coli* BL21 cells (2×10^5 cfu), DNA sequence analysis of 32 randomly selected members indicated a stochastic distribution of new termini across the entire length of the protein sequence (Figure 1, red lines).

Enzyme variants with glycoside hydrolase activity were identified by screening the library against xylan using a Congo Red overlay assay. The examination of 3000 cfu yielded 59 active candidates. Among them, DNA sequence analysis found 35 hydrolases with unique termini (Figure 1, green lines, and Supporting Information Table S3). The new N- and C-termini of these variants were distributed broadly across the sequence of Bcx. When mapped onto the tertiary structure of the protein, new termini were found to occur not only in exposed loops but also within β-strands and even surprisingly near several active site residues (Figure 2).

Qualitative Analysis of Randomly Generated Bcx Permutants. A preliminary qualitative analysis of these 35 permutants was performed to assess expression levels, purifiable yields, and specific activity of individual variants for hydrolysis of DNPX2 (Table 2). As a reference, we used "cpWT", one of the functional library members whose new termini coincide with the positions in wild-type Bcx. cpWT is derived from cpA123^{G2} and hence carries the Arg122 deletion and the T123A mutation. These changes lower its T_m value by 7 °C relative to the wild-type enzyme, yet lead to slightly higher activity with the test substrates (Tables 1 and 3). To assess protein expression levels and solubility of all 35 Bcx permutants, culture samples of the *E. coli* expression host were analyzed for total and soluble target protein by ion-exchange chromatography and SDS−PAGE. Relative to cpWT, a majority of permutants showed lower amounts of soluble protein and a tendency for formation of inclusion bodies. The trend likely reflects reduced overall protein stability and possibly impaired folding kinetics. The relative activities of the single-step purified proteins, determined using a colorimetric substrate depletion assay, largely (though not in all cases) correlated with the purified yields and thus likely with stability. As summarized in Figure 2 and Table 2, several permutants exhibited wild-type-like activity, whereas activities of others were significantly impaired.

Detailed Stability and Activity Studies on Three Selected Permutants. Three permutants with new termini within or near the active site (cpN35$^{G2'}$, cpY94$^{G2'}$, and cpY174$^{G2'}$) were selected for detailed characterization. The genes encoding these permutants were corrected for missing or extra residues resulting from the random circular permutation protocol and subsequently recloned into the pET21 vector without an encoded pelB leader. The expressed and purified enzymes were quantified by an active site titration assay using the covalent inactivator 2F-DNPX2.

Initial measures of the effects of the permutations on the structure and stability of Bcx were obtained using CD spectroscopy. cpN35$^{G2'}$, cpY94$^{G2'}$, and cpY174$^{G2'}$ yielded CD spectra that were practically identical to those of both cpWT and wild-type Bcx (not shown). Thus, each variant folded properly, and circular permutation at the three sites had little impact on the overall structures of the enzymes. Nonetheless, the stability of these variants was impaired since thermal denaturation studies revealed a significant drop in T_m values from 56 °C for cpWT to, in the worst case, 46 °C for cpY174$^{G2'}$ (Table 1). The structure and catalytic viability of each permutant were probed by treatment with the mechanism-based inhibitor 2F-DNPX2 to trap them as their covalent glycosyl-enzyme intermediate. Indeed, in each case the T_m value was raised by approximately 10 °C upon trapping of the intermediate, as also found for the wild type and cpWT, therefore indicating correct folding and enzymatic mechanism.

Kinetic parameters for hydrolysis of ONPX2 and DNPX2 by the selected permutants and wild-type Bcx are presented in Tables 1 and 3, respectively. Kinetic properties of wild-type Bcx were consistent with previous reports (*6*, *12*, *42*). Deletion

Table 2: Qualitative Analysis of Randomly Generated Bcx Circular Permutants

enzyme[a]	sequence	relative activity[b]	relative yield[c]	structural comment[d]
cpWT	A1/W185	++++	++++	deletion/mutation in thumb
cpA1^{G2}	A1/V184-CN	+++	++	N-terminal region before strand β1
cpS2^{G2}	S2/A1-S	++++	+++	N-terminal region before β1
cpT3^{G2}	T3/S2-TN	++++	+++	N-terminal region before β1
cpW6^{G2}	W6/Y5-WHN	++	+	start β1 (outer edge sheet A)
cpT10^{G2}	T10/W9-TN	+++	++	end β1 (outer edge sheet A)
cpD11^{G2}	D11/T10-N	++++	++	loop L1−2
cpG13^{G2}	G13/T10-N	++++	++	deletion in loop L1−2
cpA18^{G2}	A18/T10-N	+	+	deletion of loop L1−2 and part of β2 (outer edge sheet B)
cpG21^{G2}	G21/N17-VTN	++	+	within β2 (outer edge sheet B)
cpG23^{G2}	G23/S22	++	+	loop L2−3
cpS27^{G2}	S27/Y26-SVNWSITN	++++	++	insertion in β3 (in sheet B)
cpS31^{G2}	S31/W30-S	++++	++	end β3 (in sheet B)
cpG34^{G2}	G34/T33-GITN	+++	++++	loop L3−4
cpN35^{G2}	N35/N32-ITN	++++	++++	loop L3−4 and start β4 (in sheet A); N35 contacts general acid/base E172
cpG56^{G2}	G56/A55-N	+	+	within β5 (in sheet B)
cpP60^{G2}	P60/A59-P	+	+	within β5 (in sheet B)
cpR73^{G2}	R73/T72-N	++	++	end β6 (in sheet A)
cpE78^{G2}	E78/I77-TN	+	+	start β7 (in sheet A); E78 is the nucleophile
cpG92^{G2}	G92/T91-N	+	+	loop L7−8
cpY94$^{G2}{}_1$	Y94/P90-N	++	+	loop L7−8 and β8 (outer edge sheet A)
cpY94$^{G2}{}_2$	Y94/Y93-YITN	+++	++	start β8 (outer edge sheet A)
cpG96^{G2}	G96/Y94-ITN	++	+	within β8 (outer edge sheet A)
cpV98^{G2}	V98/T93	+	++	deletion of β8 (outer edge sheet A)
cpD119^{G2}	D119/I118-DN	+	++++	thumb (L9−10)
cpQ127^{G2}	Q127/Y108-N	+	+	deletion of thumb and parts β9 and β10 (sheet A); Q127 contacts nucleophile E78
cpK135^{G2}	K135/S134-KLTN	++	++	loop L10−11
cpG139^{G2}	G139/T138	++	++	loop L10−11
cpS162^{G2}	S162/G161	++	+++	loop L11−12
cpN163^{G2}	N163/L160-TN	+	+	loop L11−12 and start β12 (edge sheet A)
cpY166^{G2}	Y166/A165-YHN	+	+	within β12 (edge sheet A)
cpY174^{G2}	Y174/G173-YQ	++++	+++	end β12 (in sheet A); near G172
cpG178^{G2}	G178/S177-GSSN	++++	++++	start β13 (in sheet B)
cpN181^{G2}	N181/S177	++	++	deletion within β13 (in sheet B)
cpW185^{G2}	W185/V184-W	++++	++++	end β13 (in sheet B)
cpLink2^{G2}	G-A1/V184-LTN	++++	++++	loop L1−13

[a]All enzyme variants listed are ΔArg122 and T123A. C-Terminal amino acid extensions are listed. Small variations in chain length of individual permutants are caused by reading frame shifts and staggered ends upon DNase I digestion. Complete sequences are provided in Supporting Information Table S3. [b]Activity, determined using the substrate depletion method for the hydrolysis of DNPX2 in 20 mM MES (pH 6.0), 50 mM NaCl, and 0.1% (w/v) BSA, relative to cpWT: <25% (+), 25−50% (++), 51−75% (+++), and >75% (++++). [c]Yields of ion-exchange chromatographically purified proteins relative to cpWT: <25% (+), 25−50% (++), 51−75% (+++), and >75% (++++). [d]Secondary structure of wild-type Bcx: β1 (sheet A) residues 5−10; β2 (sheet B) 15−20; β3 (B) 25−31; β4 (A) 35−42; β5 (B) 50−61; β6 (A) 64−73; β7 (A) 77−85; β8 (A) 93−100; β9 (A) 103−116; β10 (A) 122−132; β11 (B) 142−145; β12 (A) 163−174; β13 (B) 177−185; α-helix 146−155. Loops are identified by flanking strands.

Table 3: Kinetic Parameters for the Hydrolysis of DNPX2 by Wild-Type Bcx and Selected Circular Permutants

enzyme	k_{cat}/K_m (s^{-1} mM^{-1})[a]	relative activity[a]	k_{cat}/K_m (s^{-1} mM^{-1})[b]	k_{cat} (s^{-1})[b]	K_m (mM)[b]
WT	19 ± 2	1	14 ± 1	34 ± 1	2.4 ± 0.1
cpWT[c]	21 ± 2	1.1	18 ± 2	56 ± 1	3.1 ± 0.1
cpN35$^{G2'}$ [c]	76 ± 8	4	46 ± 5	150 ± 4	3.3 ± 0.2
cpY94$^{G2'}$ [c]	14 ± 1	0.7	9.9 ± 1	21 ± 1	2.1 ± 0.2
cpY174$^{G2'}$ [c]	17 ± 2	0.9	nd	nd	nd

[a]k_{cat}/K_m measured by substrate depletion. [b]Complete Michaelis−Menten kinetic parameters measured at 22 °C. The highest substrate concentration tested was $2K_m$. [c]Enzymes are ΔArg122 and T123A.

of Arg122 and substitution of Thr123 with an alanine residue in cpWT, somewhat surprisingly, resulted in a slight improvement in catalytic efficiency. However, based on numerous mutagenesis studies, the thumb regions of family 11 xylanases are tolerant to extensive mutations, including amino acid deletions (*60*). Even more interestingly, the circular permutants were also comparable or slightly better catalysts than wild-type Bcx, with activities enhanced by up to ∼4-fold in the case of cpN35$^{G2'}$.

X-ray Crystallographic Structure of cpA123^{G1}. The structure of the initially designed cpA123^{G1} variant was determined by X-ray crystallography in order to obtain insights into the consequences of permutation. The protein crystallized with three molecules in the asymmetric unit, with each monomer superimposing closely upon one other and upon the structure of the wild-type enzyme (Figure 3, center). Since the overall fold of Bcx is retained, we focus our comments on the localized conformational differences at the termini of the wild-type and permutated protein.

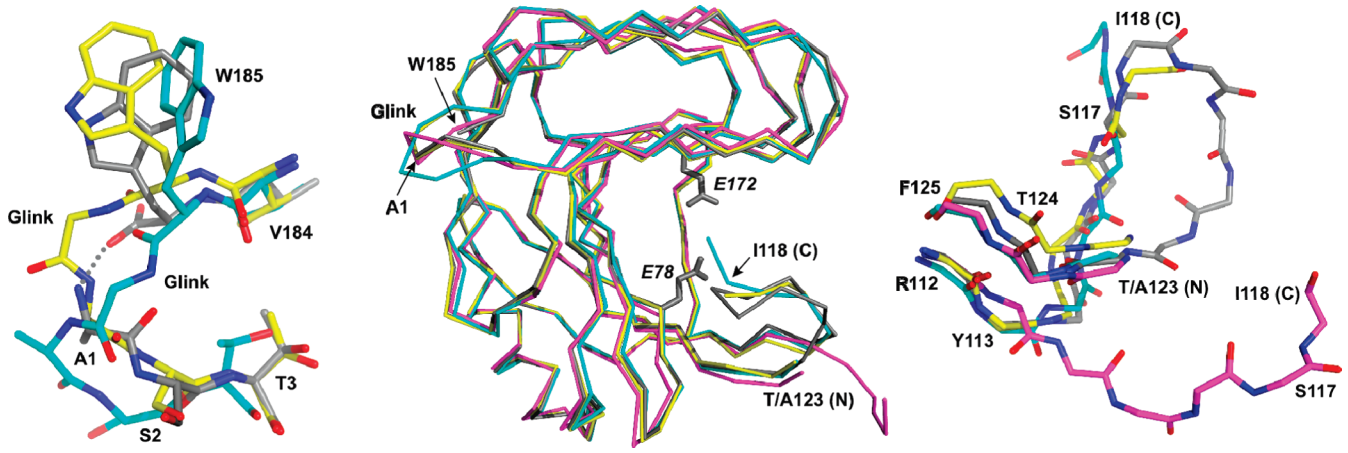

FIGURE 3: X-ray crystallographic structure of cpA123^{G1}. The three chains in the asymmetric unit of the cyclic permutant (cyan, yellow, magenta) are superimposed upon one another and wild-type Bcx (1xnb.pdb; gray). Center: Full C$^{\alpha}$ traces of the proteins, along with side chains of the catalytic Glu78 and Glu172. Left: Expanded view of the N- and C-termini (Ala1 and Trp185, respectively) of the wild-type protein and the single glycine linking these residues in cpA123^{G1}. For clarity, only two chains of cpA123^{G1} are shown. The salt bridge between Ala1 and Trp185 is identified by a dotted line. Right: Expanded main chain trace of the thumb region of wild-type Bcx and the new N- (Ala123) and C- (Ile118) termini of cpA123^{G1}. The final three residues (...Asp119-Gly120-Asp121) of the permutant are not detected and thus are likely disordered. Wild-type numbering is used, and oxygen (red) and nitrogen (blue) atoms are identified by color.

In wild-type Bcx, the positively charged α-amine of Ala1 and negatively charged α-carboxyl of Trp185 are separated by ~2.7 Å and form a salt bridge to one another (Figure 3, left). In cpA123^{G1}, these residues are covalently linked via a single glycine residue that is readily accommodated with only small localized perturbations (< 2 Å). However, differences in conformation between the linker residues in the three monomeric copies of the permutant within the crystal lattice suggest that this region is relatively mobile.

The new termini of cpA123^{G1} are located within the "thumb" of Bcx, which is an exposed *β*-hairpin that lies over the "palm" active site of the enzyme. The new N-terminal residues of cpA123^{G1} superimpose closely upon the corresponding residues (Thr/Ala123-Thr124-Phe125... by wild-type numbering) in the intact thumb of wild-type Bcx (Figure 3, right). By contrast, electron density not observable for the final three C-terminal residues (...Asp119-Gly120-Asp121) of the permutant and quite different conformations are seen for the penultimate residues in the three monomer subunits. Together, this indicates that the new C-terminus of cpA123^{G1}, located in the cleaved thumb of the circularly permuted protein, is flexible.

NMR Spectroscopic Analyses. NMR spectroscopy was also employed to examine the structural and dynamic properties of the Bcx permutants. The $^{1}H-^{15}N$ HSQC spectra of cpWT, cpN35$^{G2'}$, and cpY94$^{G2'}$ are presented in Figure 4. Due to limited solubility, cpY174$^{G2'}$ was not investigated. These HSQC spectra serve as highly sensitive "fingerprints" of the proteins, with each signal corresponding to a backbone or side chain $^{1}H^{N}-^{15}N$ amide or indole group. Thus, on a first level, the similarity of the three spectra confirms that each permutant indeed adopts the same well-folded tertiary structure. More detailed analyses of the structures and dynamic properties were then obtained after assignment of the signals of ^{1}H, ^{13}C, and ^{15}N nuclei in the polypeptide backbone of cpWT and cpN35$^{G2'}$, and the ^{1}H and ^{15}N nuclei in cpY94$^{G2'}$, by standard heteronuclear approaches.

Chemical shift perturbations are an exquisitely sensitive indicator of structural perturbations that occur in a protein due to mutation or ligand binding. Mapping of the $^{1}H^{N}$ and ^{15}N chemical shift differences between cpWT and wild-type Bcx onto the structure of the latter revealed that perturbations due to the Arg122 deletion and T123A mutation are localized to the thumb region of the enzyme (Supporting Information Figure S1). The thumb is known to exhibit modestly enhanced flexibility relative to the well-ordered core of the protein (*9*, *61*) and thus appears able to accommodate these changes without compromising activity. Similarly, differences in the corresponding $^{1}H^{N}$ and ^{15}N chemical shifts of cpN35$^{G2'}$ versus cpWT arise for residues at the new and old termini of the two proteins. Again, this is indicative of only localized structural changes due to the mutation/deletion and the permutation.

The main chain ^{1}H and ^{13}C chemical shifts of a protein are also a sensitive indicator of its secondary structure. Based upon the "secondary structure propensity" (SSP) algorithm (*62*), cpN35$^{G2'}$ retains essentially the same *β*-strand and α-helix conformation, including within the proximity of the old and new termini, as its wild-type parent (Supporting Information Figure S2). However, it is noteworthy that signals from only one of the two linker glycine amides could be identified in the $^{1}H-^{15}N$ HSQC spectrum of cpN35$^{G2'}$. Similarly, signals from the catalytic general acid/base Glu172 as well as Gly173, Tyr174, and Gln175 were not observed, presumably due to conformational exchange broadening (i.e., motions on a millisecond to microsecond time scale that cause chemical shift changes). These residues are structurally adjacent to the new terminus at Asn35 (i.e., the side chains of Asn35 and Glu172 are hydrogen bonded in the wild-type protein), suggestive of dynamic perturbations due to the permutation.

The backbone dynamics of cpN35$^{G2'}$ were investigated by ^{15}N T_1, T_2, and heteronuclear NOE relaxation measurements (Figure 5). Fitting of these data by the Lipari−Szabo model-free formalism yielded a global correlation time of 9.8 ns for the isotropic tumbling of the permutant. This value closely matches that reported for wild-type Bcx (*9*), verifying that cpN35$^{G2'}$ is also monomeric under the conditions used for these NMR measurements. The local backbone motions of cpN35$^{G2'}$ are reflected in the residue-specific model-free order parameter S^2, which decreases from 1 to 0 with increasing mobility of the $^{1}H^{N}-^{15}N$ bond vector on the nanosecond to picosecond time scale. As with wild-type Bcx (*9*), cpN35$^{G2'}$ is a remarkably well ordered protein with little variation in S^2 across its sequence. However, Thr33 and

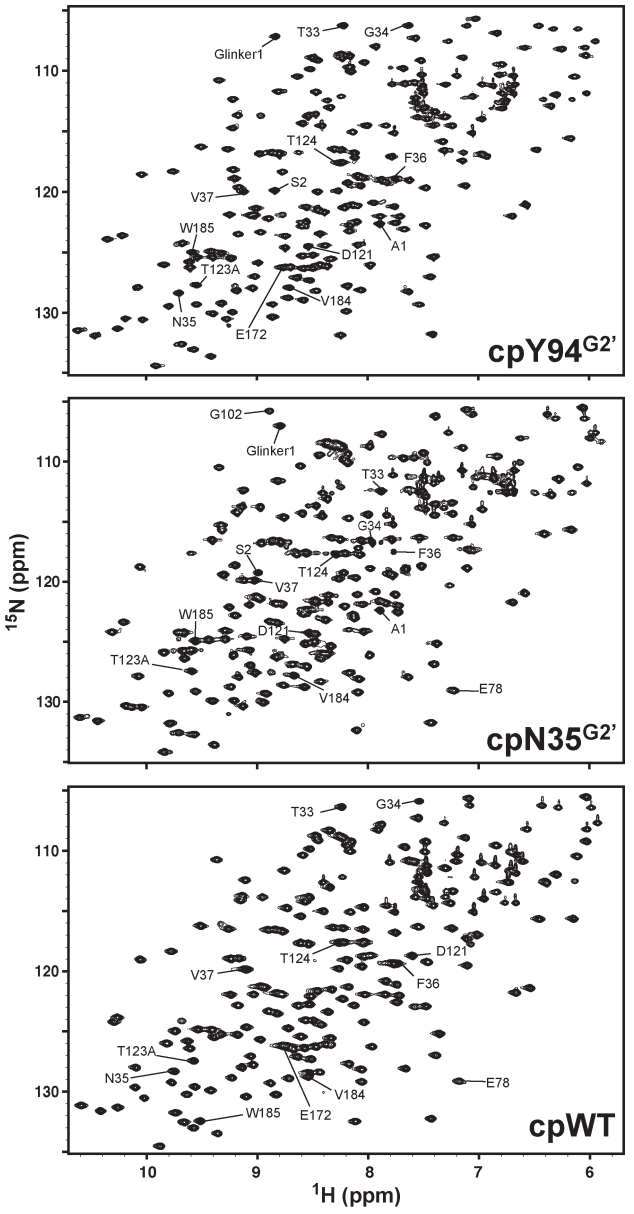

FIGURE 4: $^{1}H-^{15}N$ HSQC spectra of (top) cpY94$^{G2'}$, (middle) cpN35$^{G2'}$, and (bottom) cpWT confirm that all three proteins are well folded with similar structures. Signals from selected residues are identified using wild-type numbering.

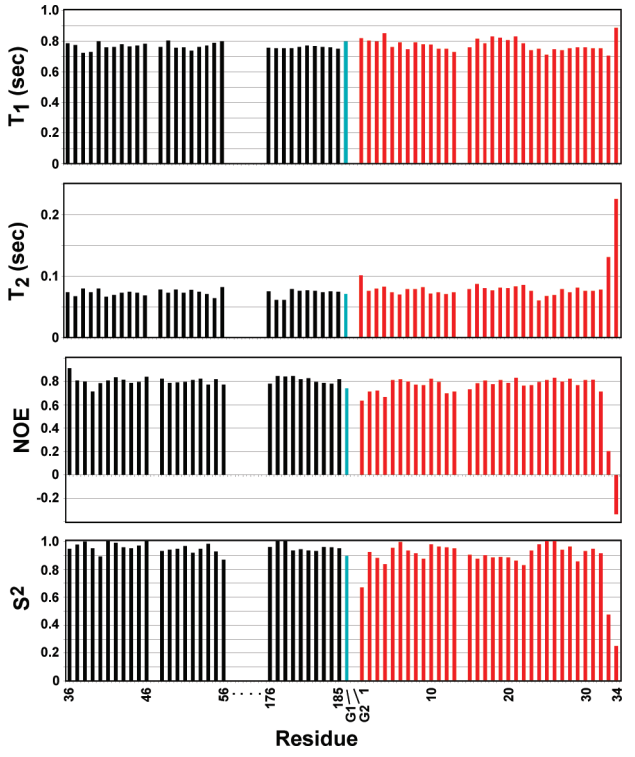

FIGURE 5: The backbone dynamics of cpN35$^{G2'}$ were characterized by amide ^{15}N T_1, T_2, and $^{1}H\{^{15}N\}$-NOE measurements, along with fit Lipari–Szabo model-free order parameters S^2. The NOE and S^2 values decrease and T_2 values increase with increasing mobility of the $^{1}H^{N}-^{15}N$ bond vector on the nanosecond to picosecond time scale. Only data for residues near the old and new termini are shown as the remainder of the protein is highly ordered with similar relaxation behavior to that reported previously for the wild-type protein (*9*). Wild-type numbering is used, but the sequence is shown in the correct continuous order for cpN35$^{G2'}$ (i.e., red-colored residues normally precede black-colored residues in Bcx). The first linker glycine (G1) is in cyan. Missing data points correspond to proline or residues with unassigned, overlapping, or weak NMR signals.

Gly34 at the new C-terminus of cpN35$^{G2'}$ do show significantly reduced S^2 values indicative of enhanced conformational mobility relative to the remainder of the protein. In contrast, the new N-terminus is relatively rigid. In addition, residues at the linked termini, including the first glycine, are well ordered with only Ala1 showing modestly reduced S^2 values.

DISCUSSION

In this study, we have generated and characterized a library of designed and randomly generated circular permutants of Bcx. The observation that variants of this xylanase can be created with new N- and C-termini within exposed loops is not in itself surprising since other proteins with a jellyroll fold have been circularly permuted (*16*, *31*, *32*). What is, however, of great interest is the fact that there are so many sites around the protein structure at which the polypeptide chain can start and still yield an active enzyme. This behavior is presumably a reflection of the stability of Bcx and a very promising way of providing greater diversity for future engineering and directed evolution efforts. It also provides a very valuable library of "start points" for the semisynthetic incorporation of unnatural amino acids to probe further the enzymatic mechanism of this model GH family 11 glycoside hydrolase.

Overall Features. Inspection of Figures 1 and 2 reveals that, as would be expected, a large number of the new start sites in the Bcx circular permutants are indeed located in exposed loops. Examples include cpG23^{G2} in loop L2−3, cpD119^{G2} in L9−10 (the "thumb"), cpG139 ^{G2} in L10−11, and cpS162^{G2} in L11−12. Due to the cloning protocol, the generated permutants also frequently incorporated deletions of wild-type residues or acquired non-native residues at their new termini. As a notable example, in cpA18^{G2} residues 11−17, which in wild-type Bcx form loop L1−2 and part of strand β2 on the outer edge of β-sheet B, were absent.

Surprisingly, numerous permutants involved cleavage of β-strands that were otherwise present in the wild-type structure (Table 2 and Figures 1 and 2). Even more remarkably, several of the resulting new termini were located within the interior and not on the outer edges of the two large β-sheets in Bcx. For example, cpS27^{G2} both cleaved strand β3 within sheet B and introduced eight new residues. Although these appended C-terminal amino acids are likely disordered, some could be involved in unexpected

non-native interactions. Further X-ray crystallographic or NMR spectroscopic studies would be required to resolve such structural questions. Additional permutants of interest include cpN181^{G2}, with three residues deleted in strand β13 within sheet B, and cpV98^{G2}, lacking strand β8 on the edge of sheet A. Presumably the hydrogen-bonding interactions between the remaining strands of the β-sheet are sufficient to correctly localize the clipped strand.

The most notable regions in which active permutants are not found are the single α-helix of Bcx and β-strand 9, which forms part of the active site platform. This may simply be a result of limited sample size or could reflect differences in the local interactions of residues within this helix versus those in loops and β-strands.

Finally, it is interesting that the most active of the permutants tended to be located at the ends of β-strands or in loop regions near the "knuckles" of Bcx (Figure 2). The resilience of this region may reflect a need to vary local structure to optimize binding and catalysis.

Biophysical Characterization. Considerable structural insights into the consequences of relocation of the N- and C-termini were obtained by CD and NMR spectroscopy and X-ray crystallography. These studies showed that, at least for cpN35$^{G2'}$, cpY94$^{G2'}$, and cpA123^{G1}, the secondary and tertiary structures were essentially identical to that of the wild-type parent. Therefore, the ~10 °C lower T_m values of these permutants likely result from localized destabilizing interactions around the new N- and C-termini or at the site of the newly fused original termini. Interestingly, the trapping of these permutants as their covalent 2-fluoroglycosyl-enzyme intermediates resulted in substantial stabilization of their structures, often to a greater degree than was seen for the wild-type enzyme. Indeed, in some cases, the additional stabilization afforded was sufficient that the trapped permutants unfolded at essentially the same temperature as the trapped cpWT parent.

In wild-type Bcx, the native N- and C-terminal residues are well ordered, interacting directly via a main chain salt bridge (*9*). However, in the X-ray crystal structure of cpA123^{G1}, the single glycine linker and these adjacent "old" terminal residues adopted multiple conformations, indicative of local flexibility. Similarly, NMR relaxation studies revealed that the second of the two glycines in the linker of cpN35$^{G2'}$ is dynamic, undergoing conformational exchange broadening on a time scale (millisecond to microsecond) leading to the loss of a detectable signal. The first glycine appeared well ordered, at least in the faster nanosecond to picosecond regime. The conformational plasticity of the fused termini likely leads to the similar stability and activity measured for the three cpA123 permutants with linkers ranging from zero to two glycines.

The newly introduced termini in the characterized permutants also appeared to be accommodated with only localized structural and dynamic perturbations. In the crystal structure of cpA123^{G1}, the opened "thumb" adopted a wild-type-like position, albeit with six exposed residues at the new C-terminus of this variant adopting multiple conformations or being unobservable due to disorder. A similar conclusion was reached through NMR relaxation studies of cpN35$^{G2'}$. In particular, only Thr33 and Gly34 at the new C-terminus of this permutant exhibited reduced order parameters indicative of fast time scale motions. However, the absence of detectable NMR signals from active site residues that are structurally adjacent to position 35 (including the catalytic Glu172) is suggestive of localized conformational exchange. Overall, however, any such structural and dynamic perturbations are relatively small and localized. This may result from the uniformly high degree of conformational stability exhibited throughout the entire backbone of the wild-type protein (*9*).

Active Site Permutants. Intriguingly, several of the randomly generated active permutants placed new start sites within the active site of Bcx. Of particular interest is cpE78^{G2}, which has the nucleophile Glu78 as its new N-terminus. Unfortunately, this permutant exhibited only low activity in qualitative tests and proved to be unstable and thus was not examined in detail. However, it is in many ways remarkable that it was active at all, given the necessity of precise localization of the catalytic nucleophile and the need to shield it from solvent to avoid esterolytic cleavage of the glycosyl-enzyme intermediate.

The permutant with the greatest activity, some 4-fold higher than the wild type, is cpN35$^{G2'}$ with Asn35 at its N-terminus. This is a particularly interesting residue, since it forms hydrogen bonds with the general acid/base Glu172 and is thought to play a role in controlling the orientation and ionization state of the acid/base catalyst (*10*). Indeed, mutation of Asn35 in Bcx into an aspartic acid shifts the pH optimum from 5.7 to 4.6 with a 20% increase in maximal activity. An equivalent mutation is seen in naturally occurring "low pH" xylanases. Detailed site-specific pK_a measurements by ^{13}C NMR revealed that this pH behavior arose from a "reverse protonation" mechanism, with a particularly strong hydrogen bond between the two residues forming in the covalent glycosyl-enzyme intermediate and contributing substantially to catalysis. In that regard, it was interesting to find that only the new C-terminus of cpN35$^{G2'}$ was significantly disordered as measured by S^2 values, while the new N-terminus appears to be stably held, consistent with the high activity of this permutant. Nonetheless, an indication of the changed environment is provided by the fact that signals from the acid/base Glu172 were not observed, presumably due to conformational exchange broadening arising from a lower level of ordering. Indeed, it was not just Glu172 that was unobservable, but a large segment of loop from Glu172 to Gln175 and including Tyr174, itself the "start site" of an active permutant with high activity.

The most striking randomly generated permutant is cpQ127^{G2}, in which the entire thumb between Tyr108 and Gln127 is absent. The β-hairpin thumb loop helps to form the active site cleft of Bcx, with Arg112 and Pro116 hydrogen bonding to the substrate in the −1 subsite. Furthermore, Gln127 hydrogen bonds to the nucleophile Glu78. Qualitatively, cpQ127^{G2} showed low levels of expression and significantly reduced activity but nonetheless was active, both in the original Congo Red assay and with an aryl xylobioside substrate (Table 2). Deletion of nine thumb residues (corresponding to Tyr111 to Asp119 of Bcx) from the GH family 11 *Thermobacillus xylanilyticus* xylanase was also reported to lead to a folded enzyme with substantially impaired, albeit measurable, activity toward xylan (*60*). Unfortunately, efforts to express and purify cpQ127$^{G2'}$ in sufficient amounts for more quantitative kinetic, structural, thermodynamic measurements of this surprisingly active thumb-severed permutant were not successful.

Perspective. The ability to generate a large number of active, circularly permuted variants of Bcx is a testament to its inherent stability and perhaps to its β-jellyroll fold. Permutation sites are most commonly located at exterior loops, yet somewhat surprisingly are also found within β-sheet regions. Presumably there are

sufficient hydrogen-bonding interactions surrounding these "cleavage sites" to stabilize the new termini. Structural studies, both by X-ray crystallography and by NMR spectroscopy, confirmed the integrity of the core and the occurrence of only localized perturbations in these native-like permutants.

These results open up several new directions of study. For example, it would be interesting to carry out comparative directed evolution studies on some of the permutants of Bcx that start at Asn35 and Tyr174, both to see if activities can be evolved more readily than is the case with the wild-type enzyme and to see if variants of higher stability can be identified (*63*, *64*). Likewise, it would be particularly interesting to carry out directed evolution on permutants such as cpE78^{G2} and cpQ127^{G2} to try to improve their stabilities such that they could then be expressed, purified, and studied in detail.

This library also provides excellent potential starting points for simplified intein-based approaches to chemobiological syntheses of variants in which key catalytic residues have been substituted with labeled or modified amino acids. Candidate permutants for such studies to substitute the catalytic nucleophile E78 include cpR73^{G2} and cpE78^{G2} for N-terminal substitution or possibly cpY94^{G2} for C-terminal substitution, though these may depend on prior stabilization through directed evolution. Likewise, variants containing unnatural acid/base residues (Glu172) could be generated starting from cpS162^{G2} or cpY174^{G2} for N- and C-terminal substitution, respectively. Analysis of variants generated in this way, using the full repertoire of structural and mechanistic tools that have been assembled, in conjunction with computational approaches, should provide particularly deep insights into the fundamentals of enzymatic catalysis.

SUPPORTING INFORMATION AVAILABLE

Additional tables summarizing cloning primers, permutant sequences, and X-ray crystallographic statistics, as well as figures of NMR chemical shift perturbations and secondary structural predictions. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res. 37*, D233–D238.
2. Collins, T., Gerday, C., and Feller, G. (2005) Xylanases, xylanase families and extremophilic xylanases. *FEMS Microbiol. Rev. 29*, 3–23.
3. Gebler, J., Gilkes, N. R., Claeyssens, M., Wilson, D. B., Beguin, P., Wakarchuk, W. W., Kilburn, D. G., Miller, R. C., Jr., Warren, R. A., and Withers, S. G. (1992) Stereoselective hydrolysis catalyzed by related beta-1,4-glucanases and beta-1,4-xylanases. *J. Biol. Chem. 267*, 12559–12561.
4. Miao, S., Ziser, L., Aebersold, R., and Withers, S. G. (1994) Identification of glutamic acid 78 as the active site nucleophile in *Bacillus subtilis* xylanase using electrospray tandem mass spectrometry. *Biochemistry 33*, 7027–7032.
5. Wakarchuk, W. W., Campbell, R. L., Sung, W. L., Davoodi, J., and Yaguchi, M. (1994) Mutational and crystallographic analyses of the active site residues of the *Bacillus circulans* xylanase. *Protein Sci. 3*, 467–475.
6. Wicki, J., Schloegl, J., Tarling, C. A., and Withers, S. G. (2007) Recruitment of both uniform and differential binding energy in enzymatic catalysis: xylanases from families 10 and 11. *Biochemistry 46*, 6996–7005.
7. McIntosh, L. P., Hand, G., Johnson, P. E., Joshi, M. D., Korner, M., Plesniak, L. A., Ziser, L., Wakarchuk, W. W., and Withers, S. G. (1996) The pKa of the general acid/base carboxyl group of a glycosidase cycles during catalysis: a ^{13}C-NMR study of *Bacillus circulans* xylanase. *Biochemistry 35*, 9958–9966.
8. Sidhu, G., Withers, S. G., Nguyen, N. T., McIntosh, L. P., Ziser, L., and Brayer, G. D. (1999) Sugar ring distortion in the glycosyl-enzyme intermediate of a family G/11 xylanase. *Biochemistry 38*, 5346–5354.
9. Connelly, G. P., Withers, S. G., and McIntosh, L. P. (2000) Analysis of the dynamic properties of *Bacillus circulans* xylanase upon formation of a covalent glycosyl-enzyme intermediate. *Protein Sci. 9*, 512–524.
10. Joshi, M. D., Sidhu, G., Pot, I., Brayer, G. D., Withers, S. G., and McIntosh, L. P. (2000) Hydrogen bonding and catalysis: a novel explanation for how a single amino acid substitution can change the pH optimum of a glycosidase. *J. Mol. Biol. 299*, 255–279.
11. Joshi, M. D., Sidhu, G., Nielsen, J. E., Brayer, G. D., Withers, S. G., and McIntosh, L. P. (2001) Dissecting the electrostatic interactions and pH-dependent activity of a family 11 glycosidase. *Biochemistry 40*, 10115–10139.
12. Ludwiczek, M. L., Heller, M., Kantner, T., and McIntosh, L. P. (2007) A secondary xylan-binding site enhances the catalytic activity of a single-domain family 11 glycoside hydrolase. *J. Mol. Biol. 373*, 337–354.
13. White, A., Tull, D., Johns, K., Withers, S. G., and Rose, D. R. (1996) Crystallographic observation of a covalent catalytic intermediate in a beta-glycosidase. *Nat. Struct. Biol. 3*, 149–154.
14. Poon, D. K., Ludwiczek, M. L., Schubert, M., Kwan, E. M., Withers, S. G., and McIntosh, L. P. (2007) NMR spectroscopic characterization of a beta-(1,4)-glycosidase along its reaction pathway: stabilization upon formation of the glycosyl-enzyme intermediate. *Biochemistry 46*, 1759–1770.
15. Kwon, J. S., Bal, J., Hwang, H. M., and Kim, J. Y. (2008) A circularly permuted beta-lactamase as a novel reporter for evaluation of protein cyclization efficiency. *J. Microbiol. 46*, 456–461.
16. Heinemann, U., and Hahn, M. (1995) Circular permutation of polypeptide chains: implications for protein folding and stability. *Prog. Biophys. Mol. Biol. 64*, 121–143.
17. Hennecke, J., Sebbel, P., and Glockshuber, R. (1999) Random circular permutation of DsbA reveals segments that are essential for protein folding and stability. *J. Mol. Biol. 286*, 1197–1215.
18. Iwakura, M., Nakamura, T., Yamane, C., and Maki, K. (2000) Systematic circular permutation of an entire protein reveals essential folding elements. *Nat. Struct. Biol. 7*, 580–585.
19. Cellitti, J., Llinas, M., Echols, N., Shank, E. A., Gillespie, B., Kwon, E., Crowder, S. M., Dahlquist, F. W., Alber, T., and Marqusee, S. (2007) Exploring subdomain cooperativity in T4 lysozyme I: structural and energetic studies of a circular permutant and protein fragment. *Protein Sci. 16*, 842–851.
20. Haglund, E., Lindberg, M. O., and Oliveberg, M. (2008) Changes of protein folding pathways by circular permutation. Overlapping nuclei promote global cooperativity. *J. Biol. Chem. 283*, 27904–27915.
21. Fishburn, A. L., Keeffe, J. R., Lissounov, A. V., Peyton, D. H., and Anthony-Cahill, S. J. (2002) A circularly permuted myoglobin possesses a folded structure and ligand binding similar to those of the wild-type protein but with a reduced thermodynamic stability. *Biochemistry 41*, 13318–13327.
22. Pieper, U., Hayakawa, K., Li, Z., and Herzberg, O. (1997) Circularly permuted beta-lactamase from *Staphylococcus aureus* PC1. *Biochemistry 36*, 8767–8774.
23. Piervincenzi, R. T., and Chilkoti, A. (2004) Effect of genetic circular permutation near the active site on the activity and stability of an enzyme inhibitor. *Biomol. Eng. 21*, 33–42.
24. Qian, Z., and Lutz, S. (2005) Improving the catalytic activity of *Candida antarctica* lipase B by circular permutation. *J. Am. Chem. Soc. 127*, 13466–13467.
25. Zhang, T., Bertelsen, E., Benvegnu, D., and Alber, T. (1993) Circular permutation of T4 lysozyme. *Biochemistry 32*, 12311–12318.
26. Qian, Z., Horton, J. R., Cheng, X., and Lutz, S. (2009) Structural redesign of Lipase B from *Candida antarctica* by circular permutation and incremental truncation. *J. Mol. Biol. 393*, 191–201.
27. Yu, Y., and Lutz, S. (2010) Improved triglyceride transesterification by circular permuted *Candida antarctica* lipase B. *Biotechnol. Bioeng. 105*, 44–50.
28. Lo, W. C., Lee, C. C., Lee, C. Y., and Lyu, P. C. (2009) CPDB: a database of circular permutation in proteins. *Nucleic Acids Res. 37*, D328–332.
29. Goldenberg, D. P., and Creighton, T. E. (1983) Circular and circularly permuted forms of bovine pancreatic trypsin inhibitor. *J. Mol. Biol. 165*, 407–413.
30. Luger, K., Hommel, U., Herold, M., Hofsteenge, J., and Kirschner, K. (1989) Correct folding of circularly permuted variants of a beta alpha barrel enzyme in vivo. *Science 243*, 206–210.

31. Hahn, M., Piotukh, K., Borriss, R., and Heinemann, U. (1994) Native-like in vivo folding of a circularly permuted jellyroll protein shown by crystal structure analysis. *Proc. Natl. Acad. Sci. U.S.A. 91*, 10417–10421.
32. Ay, J., Hahn, M., Decanniere, K., Piotukh, K., Borriss, R., and Heinemann, U. (1998) Crystal structures and properties of de novo circularly permuted 1,3−1,4-beta-glucanases. *Proteins: Struct., Funct., Genet. 30*, 155–167.
33. Zhang, P., and Schachman, H. K. (1996) In vivo formation of allosteric aspartate transcarbamoylase containing circularly permuted catalytic polypeptide chains: implications for protein folding and assembly. *Protein Sci. 5*, 1290–1300.
34. Graf, R., and Schachman, H. K. (1996) Random circular permutation of genes and expressed polypeptide chains: application of the method to the catalytic chains of aspartate transcarbamoylase. *Proc. Natl. Acad. Sci. U.S.A. 93*, 11591–11596.
35. Hytönen, V. P., Hörhä, J., Airenne, T. T., Niskanen, E. A., Helttunen, K. J., Johnson, M. S., Salminen, T. A., Kulomaa, M. S., and Nordlund, H. R. (2006) Controlling quaternary structure assembly: subunit interface engineering and crystal structure of dual chain avidin. *J. Mol. Biol. 359*, 1352–1363.
36. Qian, Z., Fields, C. J., and Lutz, S. (2007) Investigating the structural and functional consequences of circular permutation on lipase B from *Candida antarctica*. *ChemBioChem 8*, 1989–1996.
37. Antos, J. M., Popp, M. W., Ernst, R., Chew, G. L., Spooner, E., and Ploegh, H. L. (2009) A straight path to circular proteins. *J. Biol. Chem. 284*, 16028–16036.
38. Teather, R. M., and Wood, P. J. (1982) Use of Congo Red polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. *Appl. Environ. Microbiol. 43*, 777–780.
39. Sung, W. L., Luk, C. K., Zahab, D. M., and Wakarchuk, W. (1993) Overexpression of the *Bacillus subtilis* and *circulans* xylanases in *Escherichia coli*. *Protein Expression Purif. 4*, 200–206.
40. Gasteiger, E., Hoogland, C., A., G., S., D., R., W., M., D., A., R., and A., B. (2005) Protein identification and analysis tools on the ExPASy Server, in The Proteomics Protocols Handbook (Walker, J. M., Ed.) pp 571−607, Humana Press, Totowa, NJ.
41. Withers, S. G., Rupitz, K., and Street, I. P. (1988) 2-Deoxy-2-fluoro-D-glycosyl fluorides—a new class of specific mechanism-based glycosidase inhibitors. *J. Biol. Chem. 263*, 7929–7932.
42. Lawson, S. L., Wakarchuk, W. W., and Withers, S. G. (1997) Positioning the acid/base catalyst in a glycosidase: studies with *Bacillus circulans* xylanase. *Biochemistry 36*, 2257–2265.
43. Ziser, L., and Withers, S. G. (1994) A short synthesis of [beta]-xylobiosides. *Carbohydr. Res. 265*, 9–17.
44. Ziser, L., Setyawati, I., and Withers, S. G. (1995) Syntheses and testing of substrates and mechanism-based inactivators for xylanases. *Carbohydr. Res. 274*, 137–153.
45. Leatherbarrow, R. J. (2001) GraFit Version 5, 5.0 ed., Erithacus Software Limited, Horley, U.K.
46. Pace, C. N., Shirley, B. A., and Thomson, J. A. (1989) Measuring the conformational stability of a protein, in Protein Structure: A Practical Approach (Creighton, T. E., Ed.) pp 311−330, IRL Press, Oxford, U.K.
47. Otwinowski, Z., Minor, W., and Carter, Charles W., Jr. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307–326.
48. Read, R. (2001) Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystallogr., Sect. D 57*, 1373–1382.
49. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR System: a new software suite for macromolecular structure determination. *Acta Crystallogr., Sect. D 54*, 905–921.
50. Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F., and Murshudov, G. N. (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr., Sect. D 60*, 2184–2195.
51. Emsley, P., and Cowtan, K. (2004) COOT: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D 60*, 2126–2132.
52. DeLano, W. (2002) PyMOL Release 0.99, DeLano Scientific LLC, Palo Alto, CA.
53. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277–293.
54. Goddard, T. D., and Kneeler, D. G. (1999) Sparky 3, 3rd ed., University of California, San Francisco, CA.
55. Sattler, M., Schleucher, J., and Griesinger, C. (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Res. Spectrosc. 34*, 93–158.
56. Plesniak, L. A., Wakarchuk, W. W., and McIntosh, L. P. (1996) Secondary structure and NMR assignments of *Bacillus circulans* xylanase. *Protein Sci. 5*, 1118–1135.
57. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Formankay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and a phosphopeptide-complexed Src homology-2 domain atudied by N-15 NMR relaxation. *Biochemistry 33*, 5984–6003.
58. Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc. 104*, 4546–4559.
59. Dosset, P., Hus, J.-C., Blackledge, M., and Marion, D. (2000) Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J. Biomol. NMR 16*, 23–28.
60. Paes, G., Tran, V., Takahashi, M., Boukari, I., and O'Donohue, M. J. (2007) New insights into the role of the thumb-like loop in GH-II xylanases. *Protein Eng. Des. Select. 20*, 15–23.
61. Pollet, A., Vandermarliere, E., Lammertyn, J., Strelkov, S. V., Delcour, J. A., and Courtin, C. M. (2009) Crystallographic and activity-based evidence for thumb flexibility and its relevance in glycoside hydrolase family 11 xylanases. *Proteins: Struct., Funct., Genet. 77*, 395–403.
62. Marsh, J. A., Singh, V. K., Jia, Z. C., and Forman-Kay, J. D. (2006) Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: implications for fibrillation. *Protein Sci. 15*, 2795–2804.
63. Palackal, N., Brennan, Y., Callen, W. N., Dupree, P., Frey, G., Goubet, F., Hazlewood, G. P., Healey, S., Kang, Y. E., Kretz, K. A., Lee, E., Tan, X. Q., Tomlinson, G. L., Verruto, J., Wong, V. W. K., Mathur, E. J., Short, J. M., Robertson, D. E., and Steer, B. A. (2004) An evolutionary route to xylanase process fitness. *Protein Sci. 13*, 494–503.
64. Belien, T., Van Campenhout, S., Bosch, A. V., Bourgois, T. M., Rombouts, S., Robben, J., Courtin, C. M., Delcour, J. A., and Volckaert, G. (2007) Engineering molecular recognition of endoxylanase enzymes and their inhibitors through phage display. *J. Mol. Recognit. 20*, 103–112.
65. Hutchinson, E. G., and Thornton, J. M. (1996) PROMOTIF—a program to identify and analyze structural motifs in proteins. *Protein Sci. 5*, 212–220.