

# NMR Spectroscopic Characterization of a $\beta$ -(1,4)-Glycosidase along Its Reaction Pathway: Stabilization upon Formation of the Glycosyl–Enzyme Intermediate<sup>†</sup>

David K. Y. Poon, Martin L. Ludwiczek, Mario Schubert,<sup>‡</sup> Emily M. Kwan, Stephen G. Withers, and Lawrence P. McIntosh\*

Department of Biochemistry and Molecular Biology, Department of Chemistry, The Protein Engineering Network of Centres of Excellence, and The Michael Smith Laboratory, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z3

Received August 18, 2006; Revised Manuscript Received October 17, 2006

**ABSTRACT:** NMR spectroscopy was used to search for mechanistically significant differences between the thermodynamic and dynamic properties of the 34 kDa ( $\alpha/\beta$ )<sub>8</sub>-barrel catalytic domain of  $\beta$ -(1,4)-glycosidase Cex (or CfXyn10A) in its free (apo-CexCD) and trapped glycosyl–enzyme intermediate (2FCb–CexCD) states. The main chain chemical shift perturbations due to the covalent modification of CexCD with the mechanism-based inhibitor 2,4-dinitrophenyl 2-deoxy-2-fluoro- $\beta$ -cellobioside are limited to residues within its active site. Thus, consistent with previous crystallographic studies, formation of the glycosyl–enzyme intermediate leads to only localized structural changes. Furthermore, <sup>15</sup>N relaxation methods demonstrated that the backbone amide and tryptophan side chains of apo-CexCD are very well ordered on both the nanosecond to picosecond and millisecond to microsecond time scales and that these dynamic features also do not change significantly upon formation of the trapped intermediate. However, covalent modification of CexCD led to the increased protection of many amides and indoles, clustered around the active site of the enzyme, against fluctuations leading to hydrogen exchange. Similarly, thermal denaturation studies demonstrated that 2FCb–CexCD has a significantly higher midpoint unfolding temperature than apo-CexCD. The covalently modified protein also exhibited markedly increased resistance to proteolytic degradation by thermolysin relative to apo-CexCD. Thus, the local and global stability of CexCD increase along its reaction pathway upon formation of the glycosyl–enzyme intermediate, while its structure and fast time scale dynamics remain relatively unperturbed. This may reflect thermodynamically favorable interactions with the relatively rigid active site of Cex necessary to bind, distort, and subsequently hydrolyze glycoside substrates.

Glycoside hydrolases play important roles in biological systems, ranging from the degradation of polysaccharides as food sources to the manipulation and modification of glycoconjugates on the surfaces of proteins and cells. These proficient enzymes catalyze the hydrolysis of glycosidic bonds with rate enhancements of up to 10<sup>17</sup>-fold (1, 2). Dissecting the structural and dynamic bases for this remarkable catalytic efficiency not only provides valuable fundamental insights into enzymatic catalysis but also establishes a framework for engineering glycosidases for applications in biotechnological processes and for the development of glycosidase inhibitors as possible therapeutics.

The catalytic domain (CexCD)<sup>1</sup> of the  $\beta$ -(1,4)-glycosidase Cex (or CfXyn10A) from the soil bacterium *Cellulomonas fimi* has emerged as a model system for the analysis of xylanases and cellulases. With an ( $\alpha/\beta$ )<sub>8</sub>-barrel fold, CexCD is a family 10 member of the GH-A clan of glycoside hydrolases (EC 3.2.1.8; CAZY database at <http://afmb.cnrs-mrs.fr/CAZY/index.html>) (3, 4). Using a retaining double-displacement mechanism, this secreted bacterial enzyme catalyzes the cleavage of  $\beta$ -1,4-glycosidic bonds in xylan, as well as in a range of soluble aryl glycosides based upon xylose and glucose (5). Through detailed kinetic, crystallographic, and mutational studies, the nucleophile (Glu233) and the general acid/base catalyst (Glu127) in CexCD were identified, and key structural and energetic determinants for substrate specificity and covalent glycosyl–enzyme intermediate formation were defined (6, 7).

Although significant progress has been made in understanding the structural basis for catalysis by glycoside

<sup>†</sup> This research was funded by grants from NSERC and from the Government of Canada's Network of Centres of Excellence Program, supported by the CIHR and NSERC through the Protein Engineering Network of Centres of Excellence (PENCE, Inc.). Instrument support was provided by CIHR, the Canadian Foundation for Innovation (CFI), the British Columbia Knowledge Development Fund, the University of British Columbia Blusson Fund, and the Michael Smith Foundation for Health Research. M.S. acknowledges a Feodor Lynen Fellowship from the Alexander von Humboldt Foundation.

\* To whom correspondence should be addressed: Department of Biochemistry and Molecular Biology, Life Sciences Center, 2350 Health Sciences Mall, University of British Columbia, Vancouver, BC, Canada V6T 1Z3. E-mail: mcintosh@chem.ubc.ca. Phone: (604) 822-3341. Fax: (604) 822-5227.

<sup>‡</sup> Current address: Molecular Biology and Biophysics, ETH Zürich, 8093 Zürich, Switzerland.

<sup>1</sup> Abbreviations: CD, circular dichroism; CexCD, residues 1–315 of *C. fimi* Cex containing the catalytic domain; 2FCb–CexCD, CexCD covalently modified at Glu233 with 2-deoxy-2-fluorocellobiose; 2FCb-DNP, 2,4-dinitrophenyl 2-deoxy-2-fluoro- $\beta$ -cellobioside; HSQC, heteronuclear single-quantum correlation; HX, hydrogen exchange; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; pH\*, observed pH meter reading without correction for isotope effects; T<sub>m</sub>, midpoint thermal unfolding temperature.

hydrolases, lacking is a description of the changes in the thermodynamic and dynamic properties of these enzymes and their substrates that accompany binding and hydrolysis. To search for such relevant changes in Cex along its reaction pathway, we have characterized the 315-residue catalytic domain in its free (apo-CexCD) and glycosyl-enzyme intermediate (2FCb-CexCD) forms. The latter species was trapped by the covalent modification of Glu233 with the mechanism-based inhibitor (or "slow-substrate") 2,4-dinitrophenyl 2-deoxy-2-fluoro- $\beta$ -cellobioside (7). Using NMR and CD spectroscopy, combined with thermolysin proteolysis, we demonstrate that the local and global stability of CexCD increase along its reaction pathway upon formation of the glycosyl-enzyme intermediate, while its structure and fast time scale dynamics remain relatively unperturbed. Such behavior, seen also with *Bacillus circulans* xylanase (8), may reflect thermodynamically favorable interactions within the relatively rigid active site of CexCD necessary to bind, distort, and subsequently hydrolyze glycoside substrates.

## MATERIALS AND METHODS

**Protein Expression and Purification.** Full-length Cex was expressed as described previously using pUC12-1.1Cex-PTIS in *Escherichia coli* BL21( $\lambda$ DE3) cells (9). This plasmid contains the gene encoding Cex with an endogenous secretory sequence that is cleaved upon transport to the periplasm. The 70%  $^2\text{H}$ -labeled and 99%  $^{13}\text{C}$ - and  $^{15}\text{N}$ -labeled protein was produced starting with a 20 mL overnight culture in LB medium (100%  $\text{H}_2\text{O}$ ) containing 100  $\mu\text{g}/\text{mL}$  carbenicillin. The resultant cells were centrifuged, washed with unlabeled M9 medium, and used to inoculate a 20 mL culture in M9 medium prepared with 6 g of  $\text{Na}_2\text{HPO}_4$ , 3 g of  $\text{KH}_2\text{PO}_4$ , 0.5 g of  $\text{NaCl}$ , 120 mg of  $\text{MgSO}_4$ , 11 mg of  $\text{CaCl}_2$ , 2.7 mg of  $\text{FeCl}_3$ , 100 mg of carbenicillin, 1 mg of thiamine, 3 g of 99% [ $^{13}\text{C}_6$ ]glucose, 1 g of 99%  $^{15}\text{NH}_4\text{Cl}$ , and 1 g of 99% [ $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ ]Celtone (Spectra Stable Isotope Inc.) in 700 mL of  $\text{D}_2\text{O}$  and 300 mL of  $\text{H}_2\text{O}$ . After growth overnight, this culture was used to inoculate the remainder of the labeled M9 medium, which was then grown to an  $\text{OD}_{600}$  of  $\sim 0.5$ , induced with 0.1 mM IPTG, and harvested 48 h later. The temperature was 30  $^\circ\text{C}$  throughout the entire growth and expression period.  $^{15}\text{N}$ -labeled Cex was also expressed by this protocol, using 1 g of 99%  $^{15}\text{NH}_4\text{Cl}$  and 1 g of 99% [ $^{15}\text{N}$ ]Celtone (Spectra Stable Isotope Inc.) in 1 L of M9 medium.

The cells were harvested by centrifugation (6000 rpm for 20 min), resuspended in  $\sim 25$  mL of 50 mM potassium phosphate and 0.02%  $\text{NaN}_3$  at pH 7.2 in  $\text{H}_2\text{O}$  (K-P-7 buffer), and lysed by being passed twice through a French pressure cell at 10 000 psi. After addition of 25 units (1  $\mu\text{L}$ ) of benzonase (Novagen) and 4 mg of phenylmethanesulfonyl fluoride, cell debris was removed by centrifugation (15 000 rpm for 30 min), and the supernatant containing  $^2\text{H}$ -,  $^{13}\text{C}$ -, and  $^{15}\text{N}$ -labeled Cex was loaded at a flow rate of 0.5 mL/min on an  $\sim 150$  mL FPLC column (GE Healthcare) packed with washed, autoclaved, and degassed long fibrous cellulose (Sigma) suspended in K-P-7 buffer. In the case of  $^{15}\text{N}$ -labeled Cex, cellulose was added directly to the pooled supernatants from the medium and the lysed cells, stirred for 48 h at 4  $^\circ\text{C}$ , and then packed in a FPLC column. After the column had been washed with  $\sim 270$  mL of 1 M  $\text{NaCl}$  in K-P-7 buffer, followed by  $\sim 200$  mL of K-P-7 buffer, Cex was

eluted with  $\text{H}_2\text{O}$  at a rate of 1 mL/min while collecting 10 mL fractions. The appropriate fractions were pooled and concentrated to  $\sim 3$  mL using a 10 kDa MWCO polyether-sulfone membrane (Pall Life Sciences) in a stirred ultrafiltration cell (Amicon Corp.) and adjusted to 100 mM Tris, 5 mM cysteine, and 2 mM EDTA at pH 8.0. Washed agarose-immobilized papain (150  $\mu\text{L}$ ) (Pierce) was then added to cleave Cex within the linker region between its catalytic domain and carbohydrate-binding module. After incubation overnight at 37  $^\circ\text{C}$  on a tube roller, the papain was removed by centrifugation at 5000 rpm. The agarose beads were rinsed three times with K-P-7 buffer, and all the supernatants were combined and incubated overnight at 4  $^\circ\text{C}$  on a tube roller with 2 g of washed Avicel (Fluka Biochemika). The Avicel, with bound uncleaved Cex and the isolated carbohydrate-binding module, was removed by centrifugation at 5000 rpm. The supernatant, containing CexCD, was concentrated by ultrafiltration using a 3 kDa MWCO membrane polyether-sulfone (Pall Life Sciences) in a stirred ultrafiltration cell (Amicon Corp.) and exchanged into 20 mM potassium phosphate and 0.02%  $\text{NaN}_3$  at pH 6.5 (NMR buffer). The final yields were 31 and 14 mg of  $^2\text{H}$ -,  $^{13}\text{C}$ -, and  $^{15}\text{N}$ -labeled CexCD and  $^{15}\text{N}$ -labeled CexCD, respectively, per liter of medium. The enzyme concentration was determined spectrophotometrically using a predicted  $\epsilon_{280}$  value of 52 870  $\text{M}^{-1}\text{cm}^{-1}$  (<http://ca.expasy.org/tools/protpar-ref.html>) (10), and its purity was checked by ESI-MS and SDS-PAGE.

2FCb-CexCD was produced by incubating CexCD ( $\sim 1$  mM) with 2,4-dinitrophenyl 2-deoxy-2-fluoro- $\beta$ -cellobioside (2FCb-DNPC) (11) in a molar ratio of 1:3. The glycosyl-enzyme intermediate is stable on the order of months as confirmed by kinetic reactivation studies (12), NMR spectroscopy, and mass spectrometry.

**NMR Spectral Assignments.** NMR spectra were acquired at 30  $^\circ\text{C}$  on a Varian Inova 600 MHz spectrometer equipped with a  $^1\text{H}/^2\text{H}/^{13}\text{C}/^{15}\text{N}$  pulsed field gradient probe. The backbone  $^1\text{H}^{\text{N}}$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}^\alpha$ ,  $^{13}\text{C}^\beta$ , and  $^{13}\text{C}'$  resonances of  $^2\text{H}$ -,  $^{13}\text{C}$ -, and  $^{15}\text{N}$ -labeled apo-CexCD ( $\sim 1$  mM protein in NMR buffer with  $\sim 10\%$   $\text{D}_2\text{O}$  as a lock solvent) were assigned using the following suite of  $^2\text{H}$ -decoupled sensitivity-enhanced TROSY-based three-dimensional triple-resonance experiments with selection against  $^1\text{H}$ - $^{13}\text{C}$  groups: HNCA, HN(CO)CA, HNCO, HN(CA)CO, HN(CA)CB, HN(COCA)-CB, NN-NOESY, and HN-NOESY (13). Due to the similarity of its TROSY-HSQC spectrum to that of apo-CexCD, only the HNCA, HN(CO)CA, HN(CA)CB, and HN(COCA)-CB spectra were employed to assign the  $^1\text{H}^{\text{N}}$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}^\alpha$ , and  $^{13}\text{C}^\beta$  resonances of  $^2\text{H}$ -,  $^{13}\text{C}$ -, and  $^{15}\text{N}$ -labeled 2FCb-CexCD. Tryptophan  $^1\text{H}^{\epsilon 1}$  and  $^{15}\text{N}^{\epsilon 1}$  signals were assigned from HN-NOESY and  $C\beta\text{H}\delta$  experiments (14). Spectra were processed using NMRPipe (15) and analyzed with SPARKY 3.0 (16).  $^1\text{H}$  and  $^{13}\text{C}$  shifts were referenced to external sodium 2,2-dimethyl-2-silapentane-5-sulfonate and  $^{15}\text{N}$  shifts indirectly via gyromagnetic ratios (17). Reported chemical shifts were not corrected for an offset of  $^1J_{\text{NH}}/2$  due to the TROSY selection, or for  $^2\text{H}$  isotope effects. The chemical shift assignments of apo-CexCD and 2FCb-CexCD have been deposited in the BioMagResBank (<http://www.bmrb.wisc.edu/>) as accession numbers 7264 and 7265, respectively.

**$^{15}\text{N}$  Relaxation Measurements.**  $^{15}\text{N}$  relaxation measurements of  $^2\text{H}$ -,  $^{13}\text{C}$ -, and  $^{15}\text{N}$ -labeled CexCD were carried out with a 600 MHz spectrometer at 30  $^\circ\text{C}$  using TROSY

versions of  $T_1$ ,  $T_2$ , and heteronuclear  $^1\text{H}$ – $^{15}\text{N}$  NOE experiments (18). Steady state  $^{15}\text{N}$  heteronuclear NOE values were measured by recording spectra with and without 3 s of proton saturation and a total recycle delay of 5.016 s.  $T_1$  and  $T_2$  values were obtained by nonlinear least-squares fitting of the cross-peak heights to the two-parameter equation for an exponential decay using CurveFit (19). Heteronuclear NOE values were calculated as the ratio of the peak intensities recorded with and without proton saturation. Uncertainties in all spectra were estimated as described previously (18). Analysis of the  $^{15}\text{N}$  relaxation data to obtain parameters describing the anisotropic global tumbling of CexCD, as well as internal dynamics according to the extended model-free formalism of Lipari and Szabo (20–22), was carried out using TENSOR2 (23). Data were fit to the five standard models of internal mobility:  $S^2$ ;  $S^2$  and  $\tau_e$ ;  $S^2$  and  $R_{ex}$ ;  $S^2$ ,  $\tau_e$ , and  $R_{ex}$ ; and  $S^2$ ,  $S_f^2$ , and  $\tau_s$  (18, 22). A chemical shift anisotropy term with a magnitude of 107 ppm was used for the tryptophan  $^{15}\text{N}^{\epsilon 1}$  (24, 25). Amide and indole  $^{15}\text{N}$  relaxation dispersion data were acquired with relaxation-compensated  $^1\text{H}$ – $^{15}\text{N}$  CPMG-TROSY-HSQC experiments using a constant time CPMG  $T_2$  delay of 40 ms (26).

**Proton–Deuterium Exchange.** Samples of  $^{15}\text{N}$ -labeled apo-CexCD and 2FCb–CexCD ( $\sim 0.4$  mM) in 0.5 mL of NMR buffer were lyophilized to dryness. Immediately after rehydration in an equivalent volume of  $\text{D}_2\text{O}$ , a series of sensitivity-enhanced  $^1\text{H}$ – $^{15}\text{N}$  TROSY-HSQC spectra (13) were acquired over the course of 65 h at 30 °C for each sample. To detect the faster-exchanging residues, the first six spectra were recorded with four transients per FID (15 min total), while the latter spectra were recorded with 16 transients per FID for an improved signal-to-noise ratio. After the spectra had been scaled according to acquisition time, the exchange rates ( $k_{ex}$ ) were extracted from nonlinear least-squares fitting of peak intensity versus time to the equation  $I_{(t)} = I_{(0)}e^{-k_{ex}t}$  using SPARKY 3.0 (16). Data were also collected after incubation for 500 days at 4 °C to qualitatively compare the HX behavior of the most slowly exchanging amides in apo-CexCD and 2FCb–CexCD. The latter sample contained a 5:1 molar excess of 2FCb–DNP to help ensure complete inhibition over this time period. The  $\text{pH}^*$  values after exchange were 6.5 for both samples. Control experiments using  $\text{H}_2\text{O}$  buffer demonstrated that CexCD retains its folded structure after rehydration. Predicted amide exchange rates,  $k_{pred}$ , for an unstructured polypeptide with the sequence of CexCD were calculated with Sphere (27) using poly-D,L-alanine reference data corrected for amino acid type, pH, temperature, and isotope effects (28, 29).

**Thermal Denaturation Measurements.** Thermal unfolding of apo-CexCD and 2FCb–CexCD was assessed by circular dichroism (CD) spectroscopy using a 2 mm path length cell in a Jasco J-720 spectropolarimeter equipped with a Neslab RTE-111 heater and circulating water bath. The conformation of each protein sample [ $\sim 0.20$  mg/mL in 10 mM sodium phosphate (pH 6.5)] was monitored at 220 nm in response to heating at a rate of 1 °C/min. The midpoint unfolding temperature,  $T_m$ , was obtained by nonlinear least-squares data fitting to a standard equation describing a two-state conformational equilibrium (30).

**Thermolysin-Catalyzed Proteolysis.** Apo-CexCD and 2FCb–CexCD (0.50 mg/mL) were incubated at 50 °C with 0.20 mg/mL thermolysin (Sigma) in 0.10 mM sodium acetate

buffer (pH 5.5) containing 50 mM NaCl and 10 mM  $\text{CaCl}_2$  (31). At designated time points, aliquots were withdrawn and the proteolytic reaction was quenched by adding EDTA to a final concentration of 15 mM. The quenched samples were analyzed using 15% SDS–PAGE gels, stained with Sypro Red fluorescent dye (Sigma), and scanned with a Typhoon 9400 imaging system (GE Healthcare). The intensities of the bands, corresponding to the intact protein, were quantified using ImageQuant (GE Healthcare).

## RESULTS

**NMR Spectral Assignments.** The 34 kDa CexCD yielded excellent quality and well-dispersed NMR spectra, presumably due to its compact, stable  $(\alpha/\beta)_8$ -barrel fold with a mix of secondary structural elements. Similar behavior was seen with the  $(\alpha/\beta)_8$ -barrel proteins tryptophan synthase (32) and malate synthase G (33). NMR spectral assignments for the dimeric triosephosphate isomerase from yeast have also been reported recently (34). The conventional and TROSY  $^1\text{H}$ – $^{15}\text{N}$  HSQC spectra of CexCD, recorded with a 600 MHz spectrometer, were comparable in terms of resolution and signal-to-noise ratio. However, TROSY-based triple-resonance experiments, combined with aliphatic deuteration, were necessary in assigning the signals from the main chain nuclei of this protein via scalar correlations. These experiments, which involve magnetization both originating from and detected on  $^1\text{H}^N$ , require that amide deuterons, incorporated during biosynthesis in  $\text{D}_2\text{O}$ , be exchanged for protons. This is usually accomplished via reversible unfolding and refolding in  $\text{H}_2\text{O}$ -based buffers (35). Unfortunately, efforts to unfold CexCD using chemical (guanidinium hydrochloride and urea), pH, and thermal denaturation followed by an array of renaturing strategies failed to yield the folded active protein. As a result, CexCD was expressed in 70%  $\text{D}_2\text{O}$  to allow for substantial deuteration, while ensuring a minimal 30%  $^1\text{H}$  occupancy of all amides. Additional D–H exchange occurred when the folded protein was purified and then stored in NMR buffer for  $\sim 2$  weeks at 4 °C prior to the commencement of NMR experiments.

Nearly complete assignments of the resonances from the main chain  $^1\text{H}^N$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}^\alpha$ , and  $^{13}\text{C}^\beta$  nuclei in apo-CexCD and 2FCb–CexCD, as well as all seven tryptophan indole  $^{15}\text{N}^{\epsilon 1}\text{H}$  groups, were obtained using a suite of triple-resonance experiments (Figure 1). Of 304 non-proline backbone amides, 93% of the expected  $^1\text{H}^N$  and  $^{15}\text{N}$  signals in both apo-CexCD and 2FCb–CexCD were assigned. Similarly, 94% of the  $^{13}\text{C}^\alpha$  and  $^{13}\text{C}^\beta$  resonances in apo-CexCD and 93% in 2FCb–CexCD were identified. Missing assignments were due to spectral overlap, as well as to weak signals resulting from fast HX with water or from partial  $^1\text{H}^N$  occupancy of slowly exchanging amides. As summarized in Figure S1 of the Supporting Information, these chemical shifts are consistent with the secondary structure of CexCD determined crystallographically and the involvement of its four Cys residues in disulfide bonds (Cys167–Cys199 and Cys261–Cys267).

**Chemical Shift Perturbations upon Formation of the Glycosyl–Enzyme Intermediate.** To qualitatively examine the effects of covalent glycosyl–enzyme intermediate formation on the structure of CexCD, the  $^1\text{H}^N$  and  $^{15}\text{N}$  chemical shifts of corresponding residues in apo-CexCD and 2FCb–CexCD were compared. As summarized in Figure S1 of the Sup-



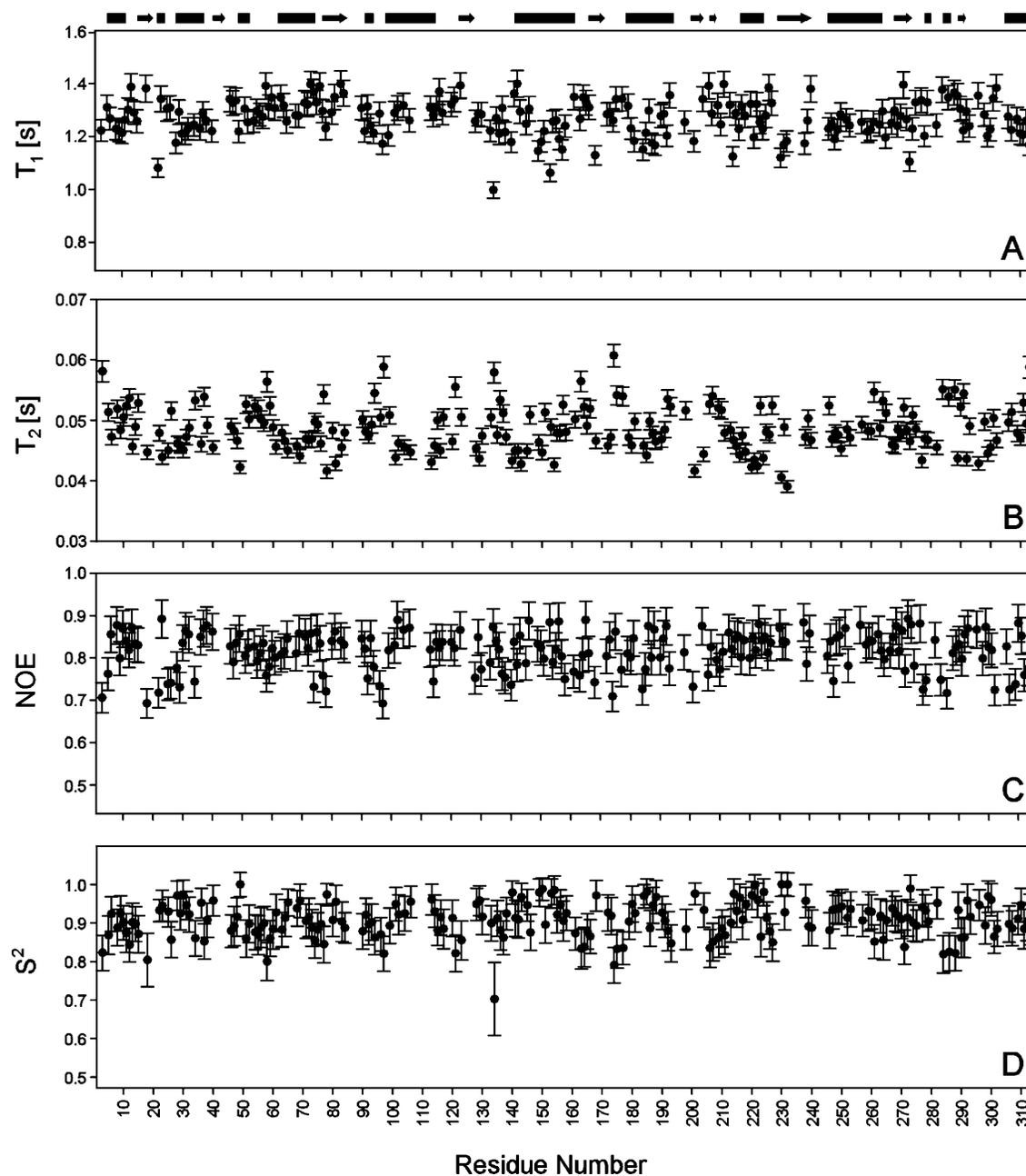


FIGURE 2: Plots of the measured  $^{15}\text{N}$   $T_1$ ,  $T_2$ , and heteronuclear  $^1\text{H}$ – $^{15}\text{N}$  NOE relaxation parameters and the fit anisotropic model-free order parameters  $S^2$  as a function of residue number for apo-CexCD. The secondary structural elements derived from the crystal structure of this protein (PDB entry 2EXO) are shown, with bars and arrows representing  $\alpha$ -helices and  $\beta$ -strands, respectively. Missing data points correspond to prolines and residues with overlapped or unassigned NMR signals. Formation of the trapped glycosyl–enzyme intermediate does not significantly change these relaxation parameters, as seen by inspection of the corresponding plots for 2FCb–CexCD in Figure S2 of the Supporting Information. The data for both proteins are provided in Tables S1 and S2 of the Supporting Information.

group of the proximal bound sugar (5). This residue exhibits the largest amide chemical shift difference between the two forms of CexCD. In addition to Trp281, Trp84 and Trp273 also are involved in direct aromatic–carbohydrate ring interactions with the substrate, and both show significant indole  $^{15}\text{N}^{\text{e}1}$ – $^1\text{H}^{\text{e}1}$  shift perturbations upon formation of the glycosyl–enzyme intermediate (Figure 1).

**Backbone Amide Dynamics from  $^{15}\text{N}$  Relaxation Measurements.** The backbone dynamic properties of CexCD were probed using  $^{15}\text{N}$  relaxation measurements. Reliable  $T_1$ ,  $T_2$ , and heteronuclear NOE relaxation data were recorded for 199 and 194 of the possible 304 non-proline backbone amides in apo-CexCD and 2FCb–CexCD, respectively

(Figure 2; Figure S2 and Tables S1 and S2 of the Supporting Information). The remaining residues were not analyzed due to spectral overlap or weak signals. The average  $T_1$  and  $T_2$  values were  $1.27 \pm 0.04$  and  $0.048 \pm 0.0013$  s, respectively, for apo-CexCD and  $1.32 \pm 0.05$  and  $0.047 \pm 0.0013$  s for 2FCb–CexCD, respectively, at 600 MHz and 30 °C. Fitting of the resultant  $T_1/T_2$  ratios with TENSOR2 (23) yielded effective correlation times ( $\tau_m$ ) of  $16.1 \pm 0.1$  ns for the global tumbling of apo-CexCD and  $16.8 \pm 0.1$  ns for 2FCb–CexCD. These measured values agree well with the expected  $\tau_m$  of  $\sim 18.5$  ns for a monomeric 315-residue globular protein at 30 °C (36).

The  $^{15}\text{N}$   $T_1$  and  $T_2$  relaxation data were also fit to a fully anisotropic diffusion tensor using the crystallographically determined structures of apo-CexCD and 2FCb-CexCD. For apo-CexCD,  $D_{zz} = (1.18 \pm 0.02) \times 10^7 \text{ s}^{-1}$ ,  $D_{yy} = (0.97 \pm 0.02) \times 10^7 \text{ s}^{-1}$ , and  $D_{xx} = (0.94 \pm 0.01) \times 10^7 \text{ s}^{-1}$ , whereas for 2FCb-CexCD,  $D_{zz} = (1.11 \pm 0.02) \times 10^7 \text{ s}^{-1}$ ,  $D_{yy} = (0.94 \pm 0.02) \times 10^7 \text{ s}^{-1}$ , and  $D_{xx} = (0.91 \pm 0.01) \times 10^7 \text{ s}^{-1}$ . These values approximate the diffusion of an axially symmetric prolate ellipsoid with a  $D_{\parallel}/D_{\perp}$  of  $\sim 1.24 \pm 0.02$  for apo-CexCD and  $1.19 \pm 0.03$  for 2FCb-CexCD. Furthermore, the principal axes of the fit diffusion tensors are roughly colinear with those of the corresponding moments of inertia calculated from the crystallographic coordinates of the two forms of CexCD, and the ratio of these experimental values agrees with those predicted by HYDRONMR (apo-CexCD, observed 1.26:1.03:1.00, predicted 1.27:1.03:1.00; 2FCb-CexCD, observed 1.22:1.03:1.00, predicted 1.26:1.03:1.00) (37). Thus, the rotational diffusion behaviors of apo-CexCD and 2FCb-CexCD derived from  $^{15}\text{N}$  relaxation measurements are consistent with the crystal structures of the prolate-shaped proteins. On the basis of these structures, HYDRONMR predicted a small increase of  $\sim 2\%$  for the  $\tau_m$  value of 2FCb-CexCD versus that of the apoprotein. The slightly larger observed increase of  $\sim 4\%$  may also reflect experimental errors and/or variations in exact sample conditions.

The internal motional properties of the apo-CexCD and 2FCb-CexCD backbone amides on the nanosecond to picosecond time scale can be extracted from the  $^{15}\text{N}$  relaxation data by the extended anisotropic Lipari-Szabo model-free formalism in terms of a generalized order parameter,  $S^2$ , which decreases from 1 to 0 with decreasing spatial restriction of the NH bond vector (Figure 2; Figure S2 and Tables S1 and S2 of the Supporting Information). Of the 199 amides analyzed in apo-CexCD, the relaxation behavior of 175 could be described using solely an  $S^2$  term (and hence for fast internal motions  $\tau_e \ll \tau_m$ ), whereas 22 were best fit using both an  $S^2$  and  $\tau_e$  term. Ala18 fitted to  $S^2$  and an additional term,  $R_{\text{ex}}$ , for conformational exchange, and only Gly134 fitted to  $S^2$  and a second term,  $S_f^2$ , for motions on a time scale near  $\tau_m$ . Furthermore, the average overall  $S^2$  value for apo-CexCD was  $0.91 \pm 0.04$ . Surprisingly, no evidence of enhanced flexibility on this time scale was detected even for the loop regions of the protein, as demonstrated by the average  $S^2$  values of  $0.93 \pm 0.04$ ,  $0.90 \pm 0.04$ , and  $0.89 \pm 0.05$  for helices, strands, and loops, respectively. Upon covalent modification to form 2FCb-CexCD, no noteworthy changes in the fast time scale dynamics of the catalytic domain were detected. Of the 194 amides that were analyzed, the relaxation behavior of 174 fit solely to an  $S^2$  term and 19 to both an  $S^2$  and  $\tau_e$  term, and only Gly97 required an  $S_f^2$  term. Similar to the apoprotein, 2FCb-CexCD exhibited an overall average  $S^2$  of  $0.91 \pm 0.05$ , with the helices, strands, and loops having average  $S^2$  values of  $0.92 \pm 0.04$ ,  $0.92 \pm 0.04$ , and  $0.89 \pm 0.05$ , respectively. Most importantly, no significant changes occurred in the amide  $S^2$  values of the active site residues that could be measured in both forms of the protein. Together, these data indicate that CexCD has a rather rigid backbone with uniformly restricted subnanosecond time scale mobility in both its unmodified and trapped glycosyl-enzyme intermediate forms.

As is frequently the case (38), we do not observe a correlation between the NMR-derived  $S^2$  order parameters and the X-ray crystallographic main chain thermal factors ( $B$  values) of apo-CexCD, which are lowest for residues forming its  $(\alpha/\beta)_8$ -barrel core and increase approximately radially to become highest at the most exposed loop regions (PDB entry 2EXO). Thus, these factors may represent crystal disorder or motions on time scales not detected by  $^{15}\text{N}$  relaxation. More importantly, the thermal factors do not change appreciably upon formation of 2FCb-CexCD (PDB entry 1EXP), indicating that any such main chain motions or crystal disorder are also not influenced by the presence of the covalently bonded glycoside.

To probe for millisecond to microsecond time scale backbone motions of CexCD,  $^{15}\text{N}$  relaxation-dispersion measurements were also performed using a 600 MHz spectrometer. These experiments sensitively detect the contribution of conformational exchange to the effective decay of the transverse  $^{15}\text{N}$  signal (26). No evidence of such conformational exchange was observed for either apo-CexCD or 2FCb-CexCD (data not shown). Thus, the amide backbone of CexCD also appears uniformly rigid on the millisecond to microsecond, as well as nanosecond to picosecond, time scales, and formation of the glycosyl-enzyme intermediate does not measurably alter this behavior.

*Tryptophan Side Chain Dynamics from  $^{15}\text{N}$  Relaxation Measurements.* Similar to the main chain amides, the seven tryptophan  $^{15}\text{N}^{\epsilon 1}-^1\text{H}^{\epsilon 1}$  moieties of CexCD exhibited  $S^2$  values of 0.85–0.93 in both its apo and glycosyl-enzyme intermediate forms (Tables S1 and S2 of the Supporting Information). Furthermore, no evidence of conformational exchange was observed for these groups in relaxation-dispersion experiments. Thus, the indole side chains of apo-CexCD and 2FCb-CexCD are also well ordered on the millisecond to microsecond and nanosecond to picosecond time scales.

*Amide and Indole Hydrogen Exchange.* The proton-deuterium HX rate of an amide or indole depends on its structural and electrostatic environment, as well as local and global conformational fluctuations that allow contact with water. Most important are motions involving the disruption of hydrogen bonds (39).  $^1\text{H}-^{15}\text{N}$  HSQC spectra provide a convenient method for measuring the site-specific HX rates of a protein. Quantitative exchange rates, corresponding to  $t_{1/2}$  values from 100 min to 65 h, were determined for 44 residues with well-resolved amide signals in apo-CexCD (Figure S1 of the Supporting Information). Within this time frame, exchange corresponds to local, rather than global, fluctuations of the protein (39). Fifty-six residues with resolved signals exchanged too rapidly and the remainder too slowly over this time period for their respective HX rates to be determined accurately. In a similar fashion, HX measurements were carried out with 2FCb-CexCD. Quantitative exchange rates were determined for 31 amides in the inhibited protein with clearly resolved  $^1\text{H}-^{15}\text{N}$  HSQC signals. Sixty-six residues with resolved signals exchanged too rapidly and the rest too slowly for accurate analysis. Note that the numbers of residues with rapid or with quantifiable exchange rates differed between apo-CexCD and 2FCb-CexCD in large part due to different patterns of  $^1\text{H}-^{15}\text{N}$  HSQC spectral overlap. As expected, the most rapidly exchanging amides in both forms of CexCD generally

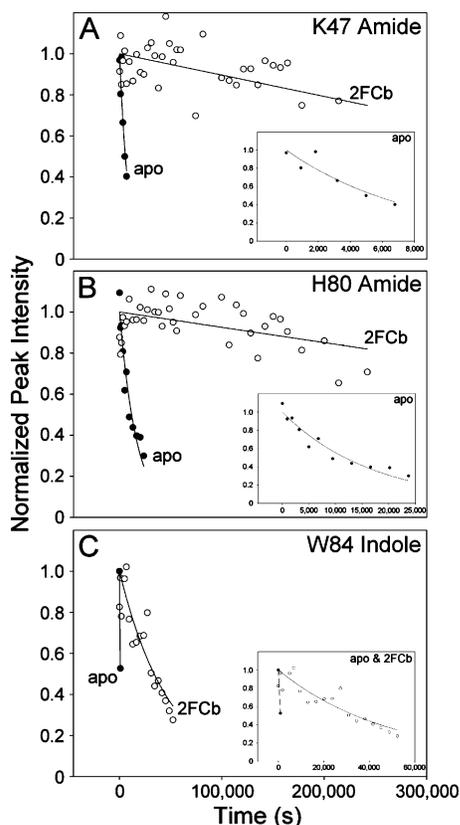


FIGURE 3: Formation of the glycosyl-enzyme intermediate stabilizes the active site of CexCD against local fluctuations leading to proton-deuterium HX. Shown are the normalized peak intensities vs time in D<sub>2</sub>O buffer (pH\* 6.5 and 30 °C) for the amide <sup>1</sup>H<sup>ε</sup>s of Lys47 (A) and His80 (B) and the indole <sup>1</sup>H<sup>ε</sup> of Trp84 (C) in apo-CexCD (●) and 2FCb-CexCD (○). The insets are expanded along the time axis. For the amides of Lys47 and His80 in 2FCb-CexCD and indole of Trp84 in apo-CexCD, exchange was too slow or fast to reliably fit a rate constant, and thus, the lines are drawn to guide the eye.

corresponded to residues on the enzyme's surface, in exposed loops, at the N-termini of helices, or on the outside edges of  $\beta$ -strands, where they are readily accessible to the solvent and not involved in stable intramolecular hydrogen bonds.

Insights into the effect of formation of the glycosyl-enzyme intermediate on the dynamic properties of CexCD can be obtained from comparison of the HX rates measured for corresponding residues in the unmodified and inhibited forms of this protein. Of the 21 amides for which quantitative exchange rates were measured in both apo-CexCD and 2FCb-CexCD,  $k_{\text{ex}(\text{apo})}/k_{\text{ex}(2\text{FCb})} = 1.1 \pm 0.3$ . Thus, the presence of the bound inhibitor does not significantly alter most local, low-energy fluctuations of CexCD that lead to HX over the relatively short time frame of these measurements. However, it is notable that active site residues Lys47 and His80 are clearly protected from HX upon formation of the glycosyl-enzyme intermediate, as evidenced by  $k_{\text{ex}(\text{apo})}/k_{\text{ex}(2\text{FCb})}$  ratios of  $>50$  and  $>25$ , respectively (Figure 3A,B). These ratios are lower limits because the slow exchange rates of these amides in 2FCb-CexCD could not be determined quantitatively. Similarly, of the three tryptophans in the active site of CexCD, <sup>1</sup>H<sup>ε</sup> of Trp281 exchanged too fast after transfer into D<sub>2</sub>O buffer to be detected in any spectra recorded for either form of the enzyme ( $t_{1/2} < 15$  min). Trp273, which donates a <sup>1</sup>H<sup>ε</sup> hydrogen bond to the bound inhibitor, also exchanged rapidly in apo-CexCD but was

detectable in the initial spectra obtained with 2FCb-CexCD (15 min  $< t_{1/2} < 60$  min). Trp84, which hydrogen bonds via its <sup>1</sup>H<sup>ε</sup> to the general acid/base Glu127 (6, 7), exhibited an  $\sim 35$ -fold decrease in  $k_{\text{ex}}$  from  $\sim 7 \times 10^{-4} \text{ s}^{-1}$  (i.e.,  $t_{1/2} \sim 15$  min) for apo-CexCD to  $2.0 \times 10^{-5} \text{ s}^{-1}$  for 2FCb-CexCD (Figure 3C). The increased level of protection of the amides of Lys47 and His80 and the indole of Trp84 may reflect dampened local fluctuations or altered solvent accessibilities or electrostatic environments due to the bound cellobioside.

Further insights into the dynamic properties of CexCD are provided by longer time scale exchange studies. Quantitative HX measurements were carried out initially over a short time period to avoid any possible reactivation of 2FCb-CexCD. Somewhat surprisingly, after being stored for 500 days at 4 °C in the presence of excess 2FCb-DNP, the protein remained fully inhibited as evidenced by its diagnostic <sup>1</sup>H-<sup>15</sup>N HSQC spectrum. This permitted a qualitative comparison of the slow HX kinetics of CexCD in its unmodified and glycosyl-enzyme intermediate states. After this storage period, 40 and 61 resolved backbone amide residues remained visible in the <sup>1</sup>H-<sup>15</sup>N HSQC spectra of apo-CexCD and 2FCb-CexCD, respectively, and had thus undergone less than  $\sim 30\%$  exchange (Figure 4A,B). More significantly, as illustrated in Figure 4C, the residues that exhibit greater protection against HX in 2FCb-CexCD than in apo-CexCD clustered around the active site region of the protein. Thus, the presence of the bound intermediate also stabilizes the active site amides of CexCD against subglobal or global fluctuations that lead to HX.

**Thermal Denaturation.** To further investigate the effects of formation of the glycosyl-enzyme intermediate, the thermal stabilities of apo-CexCD and 2FCb-CexCD were monitored by CD spectroscopy (Figure 5A). Since this enzyme does not refold reversibly under a wide range of conditions after thermal or chemical-induced denaturation, the data were fitted only to obtain midpoint unfolding temperatures. Although a reliable monitor of relative stability, this approach does not yield reversible thermodynamic parameters. Regardless, formation of the trapped glycosyl-enzyme intermediate significantly stabilizes the catalytic domain, as reflected by the measured  $T_m$  values of 55.0 and 65.5 °C for apo-CexCD and 2FCb-CexCD, respectively.

**Thermolysin-Catalyzed Proteolysis.** As demonstrated rigorously by Park and Marqusee (31), proteolysis offers an alternative method for characterizing the local and global stability of proteins. Proteins in their folded conformations are generally resistant to proteolysis, unless they contain highly flexible polypeptide sequences in exposed loop or linker regions. Thus, to become proteolyzed, residues within structured regions of proteins must undergo conformational changes to higher-energy states where cleavable sites are accessible to proteases.

We have used thermolysin-catalyzed proteolysis to compare the relative stabilities of apo-CexCD and 2FCb-CexCD. Under benign conditions, apo-CexCD is resistant to cleavage by this broad specificity, thermally stable protease (as well as to papain, which is used to separate the catalytic and carbohydrate-binding modules of full-length Cex). However, upon addition of subdenaturing concentrations of urea (not shown) or at elevated temperatures, the unmodified protein is readily degraded by thermolysin [ $t_{1/2} \sim 1$  h at 50 °C (Figure 5B)]. Transient fragments of CexCD were not detected by

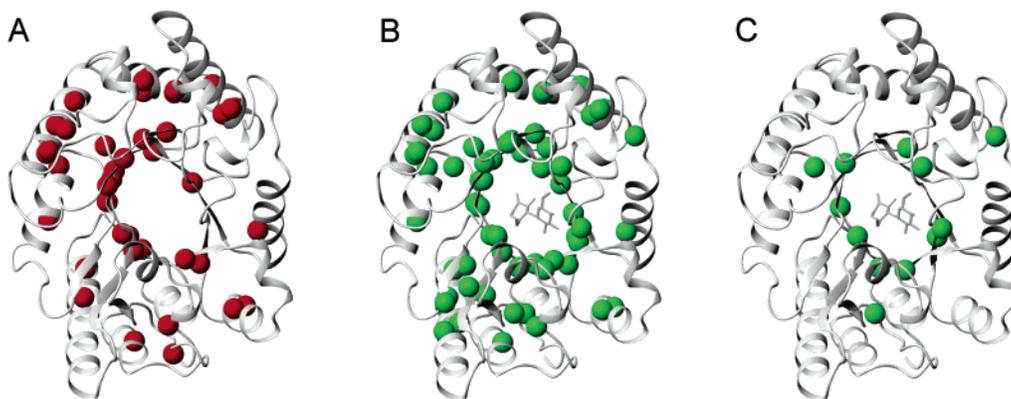


FIGURE 4: Formation of the glycosyl–enzyme intermediate stabilizes amides surrounding the active site of CexCD against higher-energy global or subglobal fluctuations leading to proton–deuterium HX. Shown are ribbon diagrams of (A) apo-CexCD and (B) 2FCb–CexCD with red and green spheres, respectively, identifying the amides most highly protected from HX, i.e., undergoing <30% exchange after storage in D<sub>2</sub>O for 500 days at 4 °C and pH\* 6.5. Only amides with unambiguously resolved HSQC signals are displayed. In panel C, green spheres denote residues showing <30% exchange in 2FCb–CexCD and >90% exchange in apo-CexCD (none exhibited the reverse pattern of slower exchange in the apoprotein). Only amides with unambiguously resolved signals in both forms of the protein are indicated.

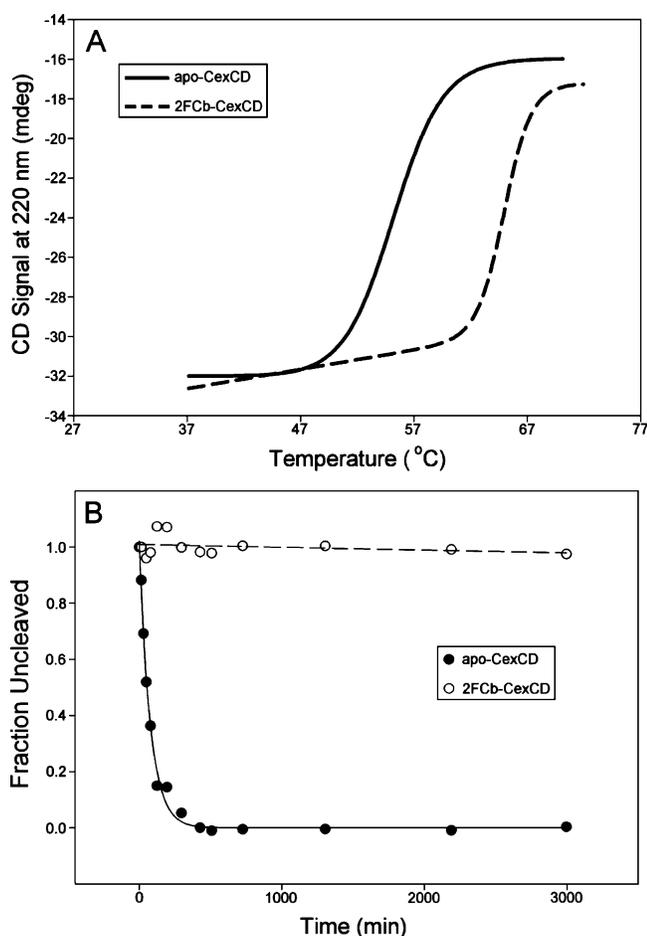


FIGURE 5: CexCD is stabilized against thermal denaturation and proteolytic degradation upon formation of the glycosyl–enzyme intermediate. (A) Thermal denaturation at pH 6.5 monitored by CD spectroscopy. The fit  $T_m$  values for apo-CexCD (—) and 2FCb–CexCD (---) were 55.0 and 65.5 °C, respectively. (B) Thermolysin proteolysis of apo-CexCD (●, solid line) and 2FCb–CexCD (○, dashed line) at 50 °C and pH 5.5 monitored by SDS–PAGE. The data were normalized to the fluorescence intensity of the uncleaved form of each protein, measured at the starting time of the digestion.

SDS–PAGE, indicating that after a rate-limiting initial cleavage, the protein is destabilized to the extent that it is rapidly hydrolyzed to small polypeptides. Given that this

temperature is approaching the  $T_m$  value of apo-CexCD, it is likely that the initial cleavage occurs within the population of globally unfolded protein molecules. In striking contrast, 2FCb–CexCD was not measurably proteolyzed by thermolysin, even after incubation for 50 h at 50 °C (Figure 5B). Thus, consistent with thermal denaturation studies, the native state of CexCD is markedly stabilized upon formation of the glycosyl–enzyme intermediate.

## DISCUSSION

In this study, we have used NMR and CD spectroscopy, along with thermolysin-catalyzed proteolysis, to characterize the catalytic domain of the  $\beta$ -1,4-glycosidase Cex and the consequences of its modification with a “slow substrate”, 2,4-dinitrophenyl 2-deoxy-2-fluoro- $\beta$ -cellobioside. This modification leads to a stable, yet catalytically competent, glycosyl–enzyme intermediate, thereby allowing a comparison of the structural, dynamic, and thermodynamic properties of CexCD at two points along its double-displacement reaction pathway.

*NMR Spectral Assignments and Structural Analyses.* As a requisite first step for these studies, we have assigned the <sup>1</sup>H, <sup>13</sup>C, and <sup>15</sup>N resonances from backbone amide and tryptophan indole nuclei in apo-CexCD and 2FCb–CexCD using TROSY-based NMR experiments, combined with 70% uniform side chain deuteration. The  $\alpha$ -helices and  $\beta$ -strands identified from chemical shifts, as well as global rotational properties derived from <sup>15</sup>N  $T_1$  and  $T_2$  relaxation lifetimes, agree closely with the crystallographically determined structures of this catalytic domain. Furthermore, amide and indole chemical shift perturbations due to the covalent modification of CexCD are localized to residues within its active site region. Recognizing that chemical shifts are exquisitely sensitive to local conformational changes, this is consistent with the essentially identical crystallographic structures reported for CexCD in its apo, covalent intermediate, and noncovalently inhibited states (5–7). Together, these data indicate that CexCD adopts a very well defined active site, poised for complementary interactions with a covalently bound glycosyl–enzyme intermediate. Indeed, the defined conformation of its active site leads to the observed ~40 fold higher activity of Cex toward xylan relative to cellulose,

as the side chains of Gln87 and Trp281 must be displaced outward to accommodate the hydroxymethyl groups present only in the latter substrate (5).

**Backbone Amide and Tryptophan Indole Dynamics.** Amide and indole  $^{15}\text{N}$  relaxation and relaxation–dispersion measurements revealed that the backbone and tryptophan side chains of apo-CexCD are also uniformly rigid on the nanosecond to picosecond and millisecond to microsecond time scales, respectively. Furthermore, as with the main chain crystallographic thermal factors of this domain, these dynamic properties did not change upon formation of the glycosyl–enzyme intermediate. In combination with the above-mentioned structural studies, this supports further the hypothesis that the active site of CexCD is a rigid scaffold, precisely positioning residues needed for the binding and subsequent hydrolysis of glycoside substrates. Interestingly, the primary amide of Gln87, which along with the indole of Trp281 establishes the higher activity of Cex for hydrolysis of xylan relative to cellulose, was reported to have anomalously high thermal factors in apo-CexCD (5). Furthermore, the glutamine side chain became unobservable in the crystal structure of 2FCb–CexCD due to unfavorable interactions with the hydroxymethyl of the distal sugar yet well ordered in that of the 2-fluoroxyllobiosyl–enzyme intermediate. Thus, some limited active site side chain flexibility may be required for hydrolysis of both xylose- and glucose-based substrates by Cex.

It is noteworthy that the  $(\alpha/\beta)_8$ -barrel or TIM barrel is one of the most common protein folds, with ligand or substrate-binding regions generally formed by loops grafted onto the core barrel (40). Although the observation that amides within the  $\alpha$ -helices and  $\beta$ -strands are well ordered is not unexpected, it is somewhat surprising that the same behavior is exhibited by residues within the active site loop regions of CexCD. The  $(\alpha/\beta)_8$ -barrel domain of malate synthase G also exhibits homogeneous backbone mobility in both its apo and pyruvate/acetyl-CoA-bound states (41, 42). However,  $^{13}\text{C}$  and  $^2\text{H}$  relaxation studies did reveal a dampening of methyl side chain dynamics in the active site of this enzyme upon substrate binding (43). In agreement with molecular dynamics simulations, millisecond to microsecond time scale backbone motions of an active site loop, as well as two helices, in the dimeric triosephosphate isomerase saturated with glycerol 3-phosphate were detected by recent NMR relaxation measurements (34). The generality of these observations will require backbone and side chain dynamic studies of additional  $(\alpha/\beta)_8$ -barrel proteins.

**Stabilization of CexCD by Glycosyl–Enzyme Intermediate Formation.** The dynamics and stability of CexCD were monitored by CD spectroscopy, thermolysin proteolysis, and HX studies. Unfortunately, it was not possible to reversibly refold this protein, as required for measurement of the thermodynamic parameters governing its global folding–unfolding equilibrium. Nevertheless, as a qualitative indicator of stability, the  $T_m$  value of CexCD increased by more than 10 °C at pH 6.5 upon formation of the glycosyl–enzyme intermediate. Similarly, 2FCb–CexCD remained resistant to thermolysin under conditions where the apoenzyme was rapidly proteolyzed. Thus, as reported previously for several glycosidases, including *Agrobacterium* sp.  $\beta$ -glucosidase (44) and *B. circulans* xylanase (8), the native structure of CexCD is dramatically stabilized in its glycosyl–enzyme intermedi-

ate state. By thermodynamic linkage, preferential noncovalent binding of a ligand to a folded protein must lead to net stabilization of the resulting complex relative to the unfolded protein and the free ligand. Indeed, irreversible deactivation studies combined with differential scanning calorimetry revealed that the apparent  $T_m$  value of Cex increased by  $\sim 2$  °C upon saturation with the weakly binding substrate, 4-nitrophenyl  $\beta$ -cellobioside (45). Such stabilization need not arise in the case of a glycosyl–enzyme intermediate, since the sugar remains covalently bonded to the enzyme in both its folded and unfolded states. However, inspection of the X-ray crystal structure of 2FCb–CexCD reveals an extensive network of hydrogen bonding interactions between polar and ionic side chains of the –2 and –1 subsites of the enzyme's active site with the hydroxyl groups of the bound fluorocellobioside, augmented by van der Waals contacts primarily involving aromatic residues, including the three previously mentioned tryptophans (5, 7). The former enthalpically favorable interactions (46, 47) likely lead to the observed stabilization of the glycosyl–enzyme intermediate of CexCD. Although less readily measurable and interpretable, entropic factors, including water release associated with the latter aromatic–sugar interactions, may also contribute (47). The complex interplay of such enthalpic and entropic contributions toward the thermodynamics and dynamics of ligand binding and catalysis has been reviewed recently (48).

Quantitative insights into the stability of CexCD are provided by NMR-monitored HX measurements. After 500 days of storage in  $\text{D}_2\text{O}$  buffer (4 °C and pH\* 6.5), numerous amides in both apo-CexCD and 2FCb–CexCD remained highly protiated. According to standard models of protein HX in the pH-dependent EX2 regime, the protection factor for a given amide, defined as the ratio of its predicted exchange rate in a random coil polypeptide,  $k_{\text{pred}}$ , to its measured value in a folded protein,  $k_{\text{ex}}$ , can be interpreted as the inverse of an equilibrium constant describing fluctuations between a closed, nonexchangeable state and a transiently exposed, exchange-competent state. Thus, protection factors provide a measure of the residue-specific free energy changes [ $\Delta G^\circ_{\text{HX}} = RT \ln(k_{\text{pred}}/k_{\text{ex}})$ ] governing local or global conformational equilibria leading to exchange (39). Although not proven herein for CexCD, protein HX typically follows an EX2 mechanism under conditions such as those of this study, whereas the open-limited EX1 mechanism is usually observed only at elevated pH values or with destabilized systems (39). On the basis of these arguments, the protection factors for these slowly exchanging residues are  $> 10^7$ , and thus, the  $\Delta G^\circ$  of unfolding of apo-CexCD under native conditions is  $> 9$  kcal/mol. Given its elevated thermal stability, this free energy change must be even greater for 2FCb–CexCD.

HX measurements also demonstrate the stabilization of CexCD upon formation of the glycosyl–enzyme intermediate. Although most weakly protected amides ( $t_{1/2} < 65$  h) exhibited comparable exchange rates in the apoprotein and covalently modified protein, it is notable that two active site residues, Lys47 and His80, in this category were  $> 50$ - and  $> 25$ -fold more protected in 2FCb–CexCD, respectively. Consistent with these effects, the labile side chain amine  $\text{H}^{\delta 1}$  of Lys47 and the imidazole ring  $\text{H}^{\delta 1}$  of His80 are also protected from HX in 2FCb–CexCD (ref 49 and unpublished observations). Similarly, the indole of Trp84 was  $\sim 35$ -fold

more protected in the trapped glycosyl–enzyme intermediate. In the crystal structures of both forms of the enzyme, the amide of His80 is hydrogen bonded to a bound water molecule, the amide of Lys47 is weakly hydrogen bonded to the carbonyl oxygen of N44 within a reverse turn, and the indole of Trp84 is hydrogen bonded to the general acid/base Glu127. Given the lack of any clear structural perturbations or altered solvent accessibility for these residues, their increased level of HX protection in 2FCb–CexCD likely results from dampened local fluctuations (on time scales slower than those that can be detected by  $^{15}\text{N}$  relaxation) due to the presence of the nearby bound inhibitor. Alternatively, changes in the electrostatic environment of the active site due to glycosylation, and hence neutralization, of the nucleophile Glu233, as well as the predicted deprotonation of the general acid Glu127 (50), could affect the rate of  $\text{OH}^-$ -catalyzed HX.

Even more pronounced differences between the two forms of CexCD are seen with the HX behavior of the most slowly exchanging amides. Although only qualitatively measured, 12 amides in apo-CexCD that exchanged extensively between 65 h and 500 days of storage in  $\text{D}_2\text{O}$  buffer remained well-protected in 2FCb–CexCD. Note that this number corresponds to amides with unambiguously resolved HSQC signals in both forms of the protein and is likely a lower limit of the number of residues actually exhibiting greater HX protection due to the covalent modification of CexCD. As illustrated in Figure 4C, these amides, which include Glu43, Asp123, Asn126, Asn169, Trp273, and the nucleophile Glu233, cluster around the active site of the enzyme. Thus, formation of the glycosyl–enzyme intermediate also stabilizes this region of CexCD against higher-energy sub-global or global fluctuations that lead to the slow HX of well-protected active site amide groups. A more detailed analysis of these fluctuations would require approaches such as “native state HX” measurements, in which exchange is measured as a function of increasing denaturant concentrations to progressively alter the free energy folding landscape of a protein (51).

**Implications for Catalysis.** Building upon previous enzymatic and X-ray crystallographic studies, the characterization of CexCD by NMR spectroscopy provides further insight into the catalytic mechanism of this model glycoside hydrolase. It is reasonable to assume that a marked stabilization, similar to that observed with 2FCb–CexCD, also results upon formation of the glycosyl–enzyme intermediate with natural substrates. The enzyme–substrate interactions responsible for this stabilization presumably facilitate catalysis by also lowering the energy of the transition state leading to this intermediate. Importantly, the covalent modification of CexCD does not measurably alter the already well-ordered backbone structure and restricted fast time scale mobility of this enzyme. A similar conclusion was reported for the family 11  $\beta$ -1,4-xylanase from *B. circulans* (8). Such behavior is not exclusive to these glycoside hydrolases, as comprehensively reviewed by Boehr et al. (38) and exemplified by a recent study showing that the backbone of the TEM-1  $\beta$ -lactamase is also highly ordered in its apo state (52). In contrast, several enzymes do exhibit significant NMR-detectable dynamic changes along their reaction pathways that are often associated with rate-limiting conformational transitions (38). The restricted backbone flexibility of these

two glycoside hydrolases may reflect the high stability of both enzymes, which, when secreted, are exposed to potentially harsh extracellular environments. Also, neither enzyme is subjected to allosteric regulation, and thus, neither requires conformational flexibility to respond to an allosteric effector (53).

Overall, these data support a hypothesis that the active sites of glycoside hydrolases such as Cex are rigidly positioned in a predominantly “lock-and-key” manner to bind and subsequently hydrolyze their substrates. To date, high-resolution crystallographic studies have been reported for CexCD in its apo, noncovalently inhibited, and covalent glycosyl–enzyme states (5–7, 54, 55); lacking are descriptions of its Michaelis enzyme–substrate and –product complexes. Nevertheless, on the basis of detailed analyses of several related glycoside hydrolases, it is likely that the substrate must be distorted for the aglycone to bind within the +1 subsite of CexCD (56). Such substrate distortion from a chair to a postulated skew-boat conformation may facilitate direct, in-line nucleophilic attack by Glu233 at the anomeric center, with concomitant aglycone leaving group departure. Of course, some conformational mobility of CexCD must occur along the reaction coordinate, including substrate binding and distortion, oxocarbenium ion-like transition state formation, and product release, yet this may require excursions to induced or preexisting higher-energy states on time scales or amplitudes not sampled in this study. For example, activated cellobiosides are hydrolyzed by CexCD with rate constants in the range of  $10^2$ – $10^3$   $\text{s}^{-1}$  for the glycosylation step and  $10$   $\text{s}^{-1}$  for the rate-limiting deglycosylation step, which correspond to microsecond to second time scale events (12). Characterization of such transient conformational changes will require additional dynamic studies, including the investigation of methyl side chain motions of Cex in its apo, noncovalently inhibited, trapped glycosyl–enzyme intermediate, and product complexes by  $^2\text{H}$  and  $^{13}\text{C}$  NMR relaxation methods and other complementary biophysical approaches. The spectral assignments of its catalytic domain open the door to such measurements, as well as to the study of active site electrostatic interactions, which are necessary to understand further the catalytic proficiency of this model glycosidase.

## ACKNOWLEDGMENT

We thank Lewis Kay, Ranjith Muhandiram, and Mark Okon for assistance with NMR spectroscopy, Lloyd Mackenzie for synthesizing 2FCb–DNP, Shouming He for mass spectrometric analyses within the University of British Columbia Proteomics Core Facility, and Federico Rosell for help with CD spectroscopy in the University of British Columbia Laboratory for Molecular Biophysics.

## SUPPORTING INFORMATION AVAILABLE

HX behavior, secondary chemical shifts, and chemical shift perturbations of CexCD upon formation of the glycosyl–enzyme intermediate (Figure S1), observed and fit  $^{15}\text{N}$  relaxation parameters for 2FCb–CexCD (Figure S2), and  $^{15}\text{N}$  relaxation of apo-CexCD and 2FCb–CexCD (Tables S1 and S2, respectively). This material is available free of charge via the Internet at <http://pubs.acs.org>.

## REFERENCES

- Rye, C. S., and Withers, S. G. (2000) Glycosidase mechanisms, *Curr. Opin. Chem. Biol.* 4, 573–580.
- Zechel, D. L., and Withers, S. G. (2001) Dissection of nucleophilic and acid-base catalysis in glycosidases, *Curr. Opin. Chem. Biol.* 5, 643–649.
- Henrissat, B., and Davies, G. (1997) Structural and sequence-based classification of glycoside hydrolases, *Curr. Opin. Struct. Biol.* 7, 637–644.
- Coutinho, P. M., and Henrissat, B. (1999) Carbohydrate-active enzymes: An integrated database approach, in *Recent advances in carbohydrate bioengineering* (Gilbert, H. J., Davies, G., Henrissat, B., and Svensson, B., Eds.) pp 3–12, The Royal Society of Chemistry, Cambridge, U.K.
- Notenboom, V., Birsan, C., Warren, R. A. J., Withers, S. G., and Rose, D. R. (1998) Exploring the cellulose/xylan specificity of the  $\beta$ -1,4-glycanase cex from *Cellulomonas fimi* through crystallography and mutation, *Biochemistry* 37, 4751–4758.
- White, A., Withers, S. G., Gilkes, N. R., and Rose, D. R. (1994) Crystal structure of the catalytic domain of the  $\beta$ -1,4-glycanase cex from *Cellulomonas fimi*, *Biochemistry* 33, 12546–12552.
- 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.
- 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.
- O'Neill, G. P., Kilburn, D. G., Warren, R. A., and Miller, R. C., Jr. (1986) Overproduction from a cellulase gene with a high guanosine-plus-cytosine content in *Escherichia coli*, *Appl. Environ. Microbiol.* 52, 737–743.
- Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein, *Protein Sci.* 4, 2411–2423.
- McCarter, J. D., Adam, M. J., Braun, C., Namchuk, M., Tull, D., and Withers, S. G. (1993) Syntheses of 2-deoxy-2-fluoro monosaccharide and oligosaccharide glycosides from glycals and evaluation as glycosidase inhibitors, *Carbohydr. Res.* 249, 77–90.
- Tull, D., and Withers, S. G. (1994) Mechanisms of cellulases and xylanases: A detailed kinetic study of the exo- $\beta$ -1,4-glycanase from *Cellulomonas fimi*, *Biochemistry* 33, 6363–6370.
- Yang, D. W., and Kay, L. E. (1999) Improved  $^1\text{H}^{\text{N}}$ -detected triple resonance TROSY-based experiments, *J. Biomol. NMR* 13, 3–10.
- Yamazaki, T., Forman-Kay, J. D., and Kay, L. E. (1993) 2-dimensional NMR experiments for correlating  $^{13}\text{C}\beta$  and  $^1\text{H}\delta/\epsilon$  chemical-shifts of aromatic residues in  $^{13}\text{C}$ -labeled proteins via scalar couplings, *J. Am. Chem. Soc.* 115, 11054–11055.
- 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.
- Goddard, T. D., and Kneeler, D. G. (1999) *Sparky* 3, University of California, San Francisco.
- Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995)  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$  chemical-shift referencing in biomolecular NMR, *J. Biomol. NMR* 6, 135–140.
- Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by  $^{15}\text{N}$  NMR relaxation, *Biochemistry* 33, 5984–6003.
- Palmer, A. G., III (1998) *Curvfit*, Columbia University, New York.
- 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.
- Kay, L. E., Torchia, D. A., and Bax, A. (1989) Backbone dynamics of proteins as studied by  $^{15}\text{N}$  inverse detected heteronuclear NMR spectroscopy: Application to staphylococcal nuclease, *Biochemistry* 28, 8972–8979.
- Mandel, A. M., Akke, M., and Palmer, A. G., III (1995) Backbone dynamics of *Escherichia coli* ribonuclease H1: Correlations with structure and function in an active enzyme, *J. Mol. Biol.* 246, 144–163.
- Dosset, P., Hus, J. C., Marion, D., and Blackledge, M. (2001) A novel interactive tool for rigid-body modeling of multi-domain macromolecules using residual dipolar couplings, *J. Biomol. NMR* 20, 223–231.
- Hu, W., Lee, K. C., and Cross, T. A. (1993) Tryptophans in membrane-proteins: Indole ring orientations and functional implications in the gramicidin channel, *Biochemistry* 32, 7035–7047.
- Goldman, M. (1984) Interference effects in the relaxation of a pair of unlike spin-1/2 nuclei, *J. Magn. Reson.* 60, 437–452.
- Mulder, F. A. A., Hon, B., Mittermaier, A., Dahlquist, F. W., and Kay, L. E. (2002) Slow internal dynamics in proteins: Application of NMR relaxation dispersion spectroscopy to methyl groups in a cavity mutant of T4 lysozyme, *J. Am. Chem. Soc.* 124, 1443–1451.
- Zhang, Y. Z. (2006) Protein and peptide structure and interactions studied by hydrogen exchange and NMR, Ph.D. Thesis, University of Pennsylvania, Philadelphia.
- Bai, Y. W., Milne, J. S., Mayne, L., and Englander, S. W. (1993) Primary structure effects on peptide group hydrogen-exchange, *Proteins* 17, 75–86.
- Connelly, G. P., Bai, Y. W., Jeng, M. F., and Englander, S. W. (1993) Isotope effects in peptide group hydrogen-exchange, *Proteins* 17, 87–92.
- 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.
- Park, C., and Marqusee, S. (2004) Probing the high energy states in proteins by proteolysis, *J. Mol. Biol.* 343, 1467–1476.
- Vadrevu, R., Falzone, C. J., and Matthews, C. R. (2003) Partial NMR assignments and secondary structure mapping of the isolated ot subunit of *Escherichia coli* tryptophan synthase, a 29-kD TIM barrel protein, *Protein Sci.* 12, 185–191.
- Tugarinov, V., Muhandiram, R., Ayed, A., and Kay, L. E. (2002) Four-dimensional NMR spectroscopy of a 723-residue protein: Chemical shift assignments and secondary structure of malate synthase G, *J. Am. Chem. Soc.* 124, 10025–10035.
- Massi, F., Wang, C., and Palmer, A. G., III (2006) Solution NMR and computer simulation studies of active site loop motion in triosephosphate isomerase, *Biochemistry* 45, 10787–10794.
- Gardner, K. H., and Kay, L. E. (1998) The use of  $^2\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$  multidimensional NMR to study the structure and dynamics of proteins, *Annu. Rev. Biophys. Biomol. Struct.* 27, 357–406.
- Daragan, V. A., and Mayo, K. H. (1997) Motional model analyses of protein and peptide dynamics using  $^{13}\text{C}$  and  $^{15}\text{N}$  NMR relaxation, *Prog. Nucl. Magn. Reson. Spectrosc.* 31, 63–105.
- de la Torre, J. G., Huertas, M. L., and Carrasco, B. (2000) HYDRONMR: Prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations, *J. Magn. Reson.* 147, 138–146.
- Boehr, D. D., Dyson, H. J., and Wright, P. E. (2006) An NMR perspective on enzyme dynamics, *Chem. Rev.* 106, 3055–3079.
- Englander, S. W., Sosnick, T. R., Englander, J. J., and Mayne, L. (1996) Mechanisms and uses of hydrogen exchange, *Curr. Opin. Struct. Biol.* 6, 18–23.
- Nagano, N., Orenge, C. A., and Thornton, J. M. (2002) One fold with many functions: The evolutionary relationships between TIM barrel families based on their sequences, structures and functions, *J. Mol. Biol.* 321, 741–765.
- Tugarinov, V., and Kay, L. E. (2003) Quantitative NMR studies of high molecular weight proteins: Application to domain orientation and ligand binding in the 723 residue enzyme malate synthase G, *J. Mol. Biol.* 327, 1121–1133.
- Tugarinov, V., Hwang, P. M., and Kay, L. E. (2004) Nuclear magnetic resonance spectroscopy of high-molecular-weight proteins, *Annu. Rev. Biochem.* 73, 107–146.
- Tugarinov, V., and Kay, L. E. (2005) Quantitative  $^{13}\text{C}$  and  $^2\text{H}$  NMR relaxation studies of the 723-residue enzyme malate synthase G reveal a dynamic binding interface, *Biochemistry* 44, 15970–15977.
- Street, I. P., Kempton, J. B., and Withers, S. G. (1992) Inactivation of a  $\beta$ -glucosidase through the accumulation of a stable 2-deoxy-2-fluoro- $\alpha$ -D-glucopyranosyl enzyme intermediate: A detailed investigation, *Biochemistry* 31, 9970–9978.
- Nikolova, P. V., Creagh, A. L., Duff, S. J. B., and Haynes, C. A. (1997) Thermostability and irreversible activity loss of exoglucanase/xylanase Cex from *Cellulomonas fimi*, *Biochemistry* 36, 1381–1388.

46. Tomme, P., Creagh, A. L., Kilburn, D. G., and Haynes, C. A. (1996) Interaction of polysaccharides with the N-terminal cellulose-binding domain of *Cellulomonas fimi* CenC. 1. Binding specificity and calorimetric analysis, *Biochemistry* 35, 13885–13894.
47. Zolotnitsky, G., Cogan, U., Adir, N., Solomon, V., Shoham, G., and Shoham, Y. (2004) Mapping glycoside hydrolase substrate subsites by isothermal titration calorimetry, *Proc. Natl. Acad. Sci. U.S.A.* 101, 11275–11280.
48. Williams, D. H., Stephens, E., O'Brien, D. P., and Zhou, M. (2004) Understanding noncovalent interactions: Ligand binding energy and catalytic efficiency from ligand-induced reductions in motion within receptors and enzymes, *Angew. Chem., Int. Ed.* 43, 6596–6616.
49. Poon, D. K. Y., Schubert, M., Au, J., Okon, M., Withers, S. G., and McIntosh, L. P. (2006) Unambiguous determination of the ionization state of a glycoside hydrolase active site lysine by <sup>1</sup>H-<sup>15</sup>N heteronuclear correlation spectroscopy, *J. Am. Chem. Soc.* 128, 15388–15389.
50. 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 <sup>13</sup>C-NMR study of *Bacillus circulans* xylanase, *Biochemistry* 35, 9958–9966.
51. Bai, Y. W. (2006) Protein folding pathways studied by pulsed- and native-state hydrogen exchange, *Chem. Rev.* 106, 1757–1768.
52. Savard, P. Y., and Gagné, S. M. (2006) Backbone dynamics of TEM-1 determined by NMR: Evidence for a highly ordered protein, *Biochemistry* 45, 11414–11424.
53. Kern, D., and Zuiderweg, E. R. P. (2003) The role of dynamics in allosteric regulation, *Curr. Opin. Struct. Biol.* 13, 748–757.
54. Notenboom, V., Birsan, C., Nitz, M., Rose, D. R., Warren, R. A. J., and Withers, S. G. (1998) Insights into transition state stabilization of the  $\beta$ -1,4-glycosidase Cex by covalent intermediate accumulation in active site mutants, *Nat. Struct. Biol.* 5, 812–818.
55. Notenboom, V., Williams, S. J., Hoos, R., Withers, S. G., and Rose, D. R. (2000) Detailed structural analysis of glycosidase/inhibitor interactions: Complexes of Cex from *Cellulomonas fimi* with xylobiose-derived aza-sugars, *Biochemistry* 39, 11553–11563.
56. Davies, G. J., Mackenzie, L., Varrot, A., Dauter, M., Brzozowski, A. M., Schulein, M., and Withers, S. G. (1998) Snapshots along an enzymatic reaction coordinate: Analysis of a retaining  $\beta$ -glycoside hydrolase, *Biochemistry* 37, 11707–11713.

BI061694C