J|A|C|S
COMMUNICATIONS

Published on Web 11/10/2006

Unambiguous Determination of the Ionization State of a Glycoside Hydrolase Active Site Lysine by $^{1}H-^{15}N$ Heteronuclear Correlation Spectroscopy

David K. Y. Poon, Mario Schubert, Jason Au, Mark Okon, 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 Laboratories, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z3

Received August 18, 2006; E-mail: mcintosh@chem.ubc.ca

Lysines play key structural and functional roles in proteins, including involvement in ion pairs, as targets of post-translational modifications, and in general acid/base and nucleophilic catalysis. These roles depend intimately upon the ionization states, and hence the pK_a values, of their side chain amines. Unfortunately, unless neutron or atomic resolution X-ray diffraction data are available, lysine protonation states are typically only inferred from crystal or solution state structural studies. Although indirect methods, such as kinetic analyses of pH-dependent chemical modifications, sometimes yield pK_a values, experimental insights into amine protonation equilibria are usually provided by NMR spectroscopy. Due to hydrogen exchange (HX) of the labile $^{1}H^{\zeta}$ with water, this has generally involved measurement of the pH-dependent chemical shifts of the amine $^{15}N^{\zeta}$ by insensitive direct detection[1] or of the neighboring $^{13}C^{\delta/\epsilon}$ by $^{1}H-^{13}C$ heteronuclear correlation experiments.[2] However, such chemical shifts may depend upon factors other than the charge of a given lysine, and it may not be possible to span a pH range sufficient to observe its titration and thereby deduce its ionization state. As demonstrated herein, when performed with care to minimize the effects of HX, ^{1}H-detected ^{15}N heteronuclear correlation experiments provide a sensitive method to directly investigate lysine amines under near-physiological conditions and to unambiguously determine their ionization states, as well as characterize their dynamic behavior.

In an ongoing effort to dissect the electrostatic basis for the catalytic proficiency of glycoside hydrolases, we are using NMR spectroscopy to determine the ionization states of active site residues in the 34 kDa family 10 catalytic domain, CexCD, from *Cellulomonas fimi* β-(1,4)-glycosidase Cex (or CfXyn10A) along its double-displacement reaction pathway.[3] When recorded with an atypical upfield ^{15}N spectral window, the $^{1}H-^{15}N$ HMQC spectrum[4] of uniformly ^{15}N-labeled apo-CexCD yielded a single peak at chemical shifts characteristic of an amine ($^{1}H^{\zeta}$ 8.1 ppm, $^{15}N^{\zeta}$ 34.8 ppm; Figure 1A).[1] On the basis of a 3D ^{15}N-edited NOESY experiment, this signal is assigned to K302.[5] The amine of K302 forms a completely buried ion pair with the carboxylate of D277, as well as hydrogen bonds with the carbonyl oxygens of A291 and L293,[6] thereby accounting for its protection from HX and direct detection by ^{1}H NMR. Upon covalent modification of CexCD by 2,4-dinitrophenyl 2-deoxy-2-fluoro-β-cellobioside to form a long-lived glycosyl−enzyme intermediate (2FCb−CexCD),[3,6] a second amine signal was observed at 8.0 ppm ($^{1}H^{\zeta}$) and 35.6 ppm ($^{15}N^{\zeta}$) (Figure 1B). This signal, also observed for CexCD noncovalently inhibited with xylobiose-derived aza-sugars (Supporting Information), is assigned to the sole lysine residue, K47, in the active site of the enzyme. As revealed by X-ray crystallography (Figure 2), the amine of K47 is hydrogen bonded to both the distal and proximal sugar moieties of the cellobioside, thereby leading to its burial and protection from HX in the trapped intermediate. The importance of these interactions is highlighted by a 75-fold decrease in k_{cat}/K_m of the Cex K47A mutant for hydrolysis of 2,4-dinitrophenyl β-cellobioside relative to the wild-type enzyme.[7]

Analysis of one-bond $^{1}H-^{15}N$ coupling patterns provides a route to unambiguously determine the ionization state of a slowly exchanging amine without resorting to pH titrations. As shown in the inset of Figure 1B, without ^{1}H-decoupling during the t_1 period of a $^{1}H-^{15}N$ HSQC experiment, the $^{15}N^{\zeta}$ signals of K47 and K302 in 2FCb−CexCD appear as quartets split by $|^{1}J_{NH}| \sim 75$ Hz. Thus, each nitrogen is directly bonded to three protons, and both lysines are in their ionized ($-NH_3^+$) states at pH 6.5. Furthermore, the $^{1}H^{\zeta}$ and $^{15}N^{\zeta}$ chemical shifts of these two residues do not change significantly between pH 3 (below which CexCD aggregates) and pH 9 (above which rapid HX preludes their detection). Given the lack of an observable titration, the pK_a values of K47 and K302 must be >9 in the glycosyl−enzyme intermediate. The same behavior was observed for K302 in unmodified CexCD. As a reference, ^{15}N NMR measurements yielded a pK_a of 11 for the side chain amine of $^{15}N^{\zeta}$-lysine, with $^{15}N^{\zeta}$ chemical shifts of 32.5 and 25.5 ppm in its ionized ($-NH_3^+$) and neutral ($-NH_2$) states, respectively (25 °C; Supporting Information).

Qualitative insights into the side chain dynamics of K47 and K302 are also provided by Figure 1B (inset). In an elegant analysis of the "methyl-TROSY" effect, Kay and co-workers[8] demonstrated that, in the ^{1}H-coupled HSQC spectrum of a $^{13}CH_3$ spin system, a ^{13}C quartet with intensity ratios of 3:1:1:3 is expected in the absence of any differential relaxation, as is closely approximated in the fast tumbling limit. (A $^{13}CH_2$ yields a 1:0:1 "triplet".) In contrast, in the macromolecular limit, the outer lines from a methyl group, which is undergoing fast 3-fold rotation and only $^{13}C-^{1}H$ dipolar relaxation, will decay 9 times faster than the inner lines. Thus, the intensity of the outer lines of the ^{13}C multiplet relative to that of the inner lines will decrease with decreasing amplitude of motion of the methyl 3-fold axis. By analogy for $^{15}NH_3^+$ spin systems, from the observed $^{15}N^{\zeta}$ multiplet intensity ratios of ~1:1:1:1 for K302 and ~0.5:1:1:0.5 for K47 (versus ~2.7:1:1:2.7 for $^{15}N^{\zeta}$-lysine; Supporting Information), the amine of the latter active site lysine appears more rigidly positioned in 2FCb−CexCD than that of the former ion-paired lysine. However, in both cases, as evidenced by single $^{1}H^{\zeta}$ resonances, there is still fast amine 3-fold rotation and hence rapid interconversion of the $^{1}H^{\zeta}$'s donated to the corresponding hydrogen bond acceptors for K47 and K302.

CexCD contains 18 lysines, most of which are exposed on the surface of the protein and thus not detected in a $^{1}H-^{15}N$ HMQC spectrum recorded at 30 °C and pH 6.5. This results from rapid HX, leading to increased proton line widths and decreased efficiency of $^{1}H-^{15}N$ coherence transfer. Strikingly, upon slowing exchange

10.1021/ja065766z CCC: $33.50

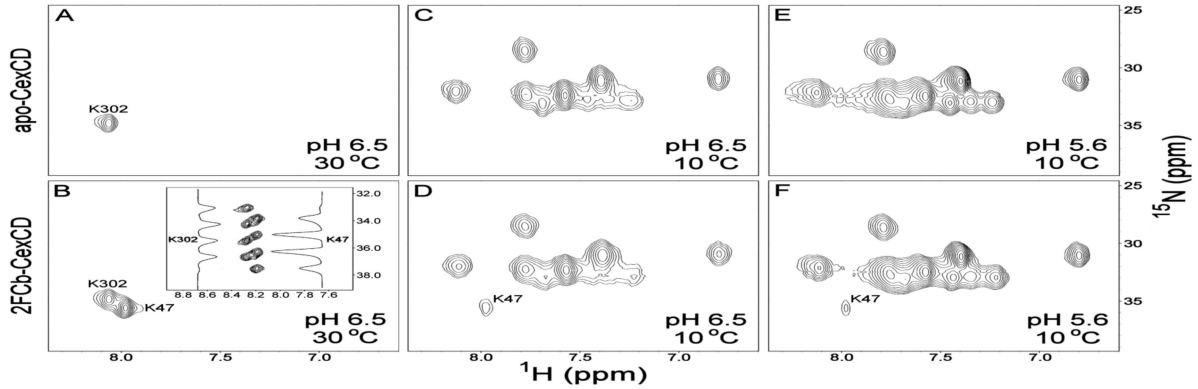

Figure 1. $^{1}H-^{15}N$ HMQC spectra of apo-CexCD (A, C, E) and 2FCb−CexCD (B, D, F) recorded using Varian Unity 500 or Inova 600 spectrometers. Samples contained ~0.4 mM uniformly ^{15}N-labeled protein in 20 mM potassium phosphate and 0.02% NaN_3. The gradient HMQC sequence incorporated selective flipback pulses to minimize water excitation[4] and delays set for $|^{1}J_{NH}| \sim 75$ Hz. Sensitivity-enhanced flipback HSQC experiments optimized for AX_3 spin systems yielded comparable spectra, albeit with marginally less signal-to-noise, for this 34 kDa protein. The high-resolution ^{1}H-coupled $^{1}H-^{15}N$ HSQC spectrum of 2FCb−CexCD at 30 °C and pH 6.5 is shown, along with the $^{15}N^{\zeta}$ traces at the $^{1}H^{\zeta}$ shifts of K47 and K302 (B, inset).

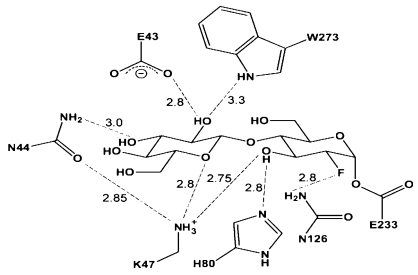

Figure 2. Proposed hydrogen-bonding interactions (dashed lines) within the active site of 2FCb−CexCD at pH 6.5. Adapted from ref 6 (1EXP) with distances (Å) between N and O atoms and experimentally determined charge states for K47, as well as E43 and H80.[5]

by simply reducing the sample temperature to 10 °C, $^{1}H^{\zeta}-^{15}N^{\zeta}$ signals from at least eight additional amines became detectable (Figure 1C,D). In contrast, the peaks from the internal K47 and K302 diminished in intensity due to the slower global tumbling of the enzyme. Furthermore, since this exchange is both specific and general base catalyzed,[1,9] reduction of the sample pH to 5.6 led to the detection of at least 10 amines with increased signal intensities (Figure 1E,F). Each of these yielded a ^{15}N quartet in a ^{1}H-coupled HSQC spectrum (with intensity ratios of ~2:1:1:2, not shown) and must arise from a fully protonated mobile surface lysine or the N-terminal amine. The observed chemical shift dispersion of 35.6−28.5 ppm in ^{15}N and 8.1−6.4 ppm in ^{1}H reflects the local environments of these amines, rather than differences in their charge states. Unfortunately, no resolved signal was present in the spectra of apo-CexCD, yet absent in those of 2FCb−CexCD or the K47A mutant, that could be confidently attributed to K47 in the unmodified protein. This may result from spectral overlap, rapid HX, or other unfavorable relaxation properties of the active site lysine. The latter is likely the case as we were able to assign the full side chain ^{13}C spin systems of surface lysines but not the internal K47 in the C(CO)TOCSY−NH spectra of (70% ^{2}H, 99% $^{13}C/^{15}N$-labeled) CexCD.[5] With their $^{13}C^{\epsilon}$ shifts identified, it should be possible to assign the amine signals of the surface lysines using $^{1}H^{\zeta}-^{15}N^{\zeta}-^{13}C^{\epsilon}$ or $^{1}H^{\epsilon}-^{13}C^{\epsilon}-^{15}N^{\zeta}$ correlation experiments.[10] Regardless, due to the absence of a detectable amine signal, we could not determine the ionization state of K47 in apo-CexCD.

In summary, using $^{1}H-^{15}N$ correlation spectroscopy, we have investigated the lysine amines in CexCD and its trapped glycosyl−enzyme intermediate, demonstrating unambiguously that K47 interacts with the cellobiosyl moiety via its positively charged side chain. It is likely that amines often remain undetected in NMR studies of proteins simply due to the optimization of ^{15}N spectral parameters for amides. Indeed, in a cursory survey, well-resolved amine signals were also observed in the $^{1}H-^{15}N$ HMQC spectra of ubiquitin, RanGAP, and a xylanase. By using conditions of reduced temperature, slightly acidic pH, and low general base concentrations to minimize the effects of HX, ^{1}H-detected ^{15}N correlation experiments provide a sensitive route to investigate directly the structural, dynamic, and functional roles of N-terminal and lysine side chain amines in proteins and protein complexes. Methyl TROSY approaches[8] and background deuteration should facilitate the analysis of amines in systems larger than the 34 kDa catalytic domain of Cex.

Acknowledgment. We are grateful to Lewis Kay for insightful discussions. This research was funded by NSERC with instrument support from PENCE, CIHR, MSFHR, CFI, BCKDF, and the UBC Blusson Fund. M.S. was a von Humboldt Foundation Feodor Lynen Fellow.

Supporting Information Available: Reference spectral data for $^{15}N^{\zeta}$-lysine and spectra of CexCD inhibited with xylobiose aza-sugars. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Leipert, T. K.; Noggle, J. H. *J. Am. Chem. Soc*. **1975**, *97*, 269−272. (b) Blomberg, F.; Mauer, W.; Rüterjans, H. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 1409−1413. (c) Knoblauch, H.; Rüterjans, H.; Bloemhoff, W.; Kerling, K. E. T. *Eur. J. Biochem.* **1988**, *172*, 485−497.
(2) Gao, G.; Prasad, R.; Lodwig, S. N.; Unkefer, C. J.; Beard, W. A.; Wilson, S. H.; London, R. E. *J. Am. Chem. Soc.* **2006**, *128*, 8104−8105.
(3) Tull, D.; Withers, S. G.; Gilkes, N. R.; Kilburn, D. G.; Warren, R. A.; Aebersold, R. *J. Biol. Chem*. **1991**, *266*, 15621−15625.
(4) (a) Griffey, R. H.; Redfield, A. G.; Loomis, R. E.; Dahlquist, F. W. *Biochemistry* **1984**, *24*, 817−822. (b) Grzesiek, S.; Bax, A. *J. Am. Chem. Soc.* **1993**, *115*, 12593−12594.
(5) Poon, D. K. Y.; Ludwiczek, M. L.; Schubert, M.; Kwan, E. M.; Withers, S. G.; McIntosh, L. P. *Biochemistry* **2006**, in press.
(6) White, A.; Tull, D.; Johns, K.; Withers, S. G.; Rose, D. R. *Nat. Struct. Biol.* **1996**, *3*, 149−154.
(7) Wicki, J.; Withers, S. G. Unpublished data.
(8) Tugarinov, V.; Hwang, P. M.; Ollerenshaw, J. E.; Kay, L. E. *J. Am. Chem. Soc.* **2003**, *125*, 10420−10428.
(9) (a) Englander, S. W.; Kallenbach, N. R. *Q. Rev. Biophys.* **1983**, *16*, 521−655. (b) Henry, G. D.; Sykes, B. D. *J. Biomol. NMR* **1995**, *5*, 59−71. (c) Liepinsh, E.; Otting, G. *Magn. Reson. Med.* **1996**, *35*, 30−42.
(10) Farmer, B. T., II; Venters, R. A. *J. Biomol. NMR* **1996**, *7*, 59−71.

JA065766Z