Electronic Supplementary Material (ESI) for Green Chemistry.


# Supplementary Information

## Highly efficient and selective biocatalytic production of glucosamine from chitin

Y. M. Lv,[a] P. Laborda,[a] K. Huang,[c] Z. P. Cai,[a] M. Wang,[a] A. M. Lu,[b] C. Doherty,[c] L. Liu,[a] S. L. Flitsch,[c] and J. Voglmeir[a]

[a.]Glycomics and Glycan Bioengineering Research Center (GGBRC), College of Food Science and Technology, Nanjing Agricultural University, Nanjing, Peoples Republic of China.

[b.]College of Sciences, Nanjing Agricultural University, Nanjing, Peoples Republic of China.

[c.]Manchester Institute of Biotechnology, University of Manchester, United Kingdom

**Contents:**

**Supplementary Fig. S6.** $^{1}$H NMR spectra of recycled CmCBDA beads.

**Supplementary Fig. S7.** Colorimetric activity test of CmCBDA towards N-acylglucosamine derivatives and Alloc-, Boc-, Cbz- and Fmoc-protected glucosamine.

**Supplementary Fig. S8.** $^{1}$H and $^{13}$C NMR spectra of N-propanoyl glucosamine.

**Supplementary Fig. S9.** $^{1}$H and $^{13}$C NMR spectra of N-butanoyl glucosamine.

**Supplementary Fig. S10.** $^{1}$H and $^{13}$C NMR spectra of N-hexanoyl glucosamine.

**Supplementary Fig. S11.** $^{1}$H and $^{13}$C NMR spectra of Alloc-protected glucosamine.

**Supplementary Fig. S12.** $^{1}$H and $^{13}$C NMR spectra of Fmoc-protected glucosamine.

**Supplementary Fig. S13.** Activity of CmCBDA towards GlcNAcc and various N-acetyl amino acids.

**Supplementary Fig. S14.** Influence of the Cu(II) ion concentration on CmCBDA activity.

**Supplementary Fig. S15.** Relative activities of recombinant CmCBDA in presence of different concentrations of urea, 2-mercaptoethanol, imidazole and sodium dodecylsulphate (SDS).

**Supplementary Fig. S16.** Temperature stability of CmCBDA.

**Supplementary Fig. S17.** Circular dichroism (CD) spectroscopy of CmCBDA incubated with various concentrations of GlcNAc (Panel A) and of an enzyme free control mixture (Panel B).

**Supplementary Fig. S18.** $^{1}$H NMR spectra of CmCBDA-catalysed deacetylation reaction of GlcNAc to GlcN after 12 h (top panel) and acid hydrolysis of GlcNAc to GlcN hydrochloride (GlcN HCl) after 4 h (bottom panel). **B** Visual comparison of the enzymatic hydrolysis reaction with acid hydrolysis reaction and alkaline hydrolysis reaction.

**Supplementary Fig. S19.** $^{1}$H NMR spectrum of the dialysate obtained after combined ZgβHexN2854 (1.25 U) and CmCBDA (5.5 U) treatment of chitinous mushroom extract.

## 1) Synthesis of N-acyl derivatives of glucosamine as CmCBDA substrates and experimental procedures for screening CmCBDA activity

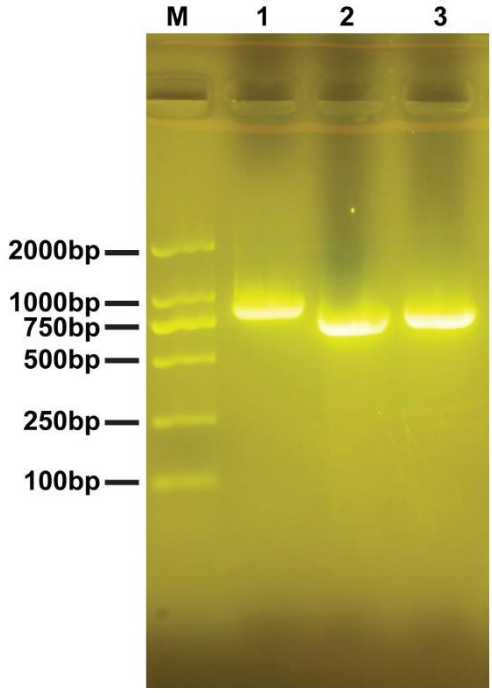
For the synthesis of N-acyl derivatives of glucosamine, a NHS-activated ester solution was prepared by stirring a solution containing the corresponding acid (0.186 mmol) (propanoic acid, butanoic acid and hexanoic acid) with N-hydroxysuccinimide (24 mg, 0.209 mmol) and N,N-dicyclohexylcarbodiimide (40 mg, 0.194 mmol) in AcOEt (200 µL) overnight at room temperature. Then, a solution containing glucosamine (20 mg, 0.093 mmol) and $Et_3N$ (14 µL, 0.102 mmol) in MeOH (1 mL) was added to the NHS-activated ester solution. The resulting suspension was stirred for 6 h at room temperature. After evaporation of the solvent, N-propanoyl, N-butanoyl and N-hexanoylglucosamine were isolated by silica gel column using a mixture AcOEt/MeOH/AcOH 2:1:1 as the eluent with 70%, 98% and 91% yield, respectively (NMR spectra and the corresponding annotation are shown in the Supplementary Fig. S8-S10). The enzymatic reaction (25 µL) containing 100 mM of N-acylglucosamine derivative, 200 mM of $Na_2HPO_4$ buffer (pH 8.0) and recombinant CmCBDA (0.013 U), were incubated for 6 h, and then quenched by adding 25 µL of chloroform. After centrifugation, an aliquot of the upper-phase was used for the determination of the produced GlcN after o-phtalaldehyde (OPA) derivatisation using the colorimetric assay in the article section 'Activity and substrate specificity screening'. Screening Results are shown in the Supplementary Figure S7.

## 2) Alloc-, Boc-, Cbz- and Fmoc-protected glucosamines

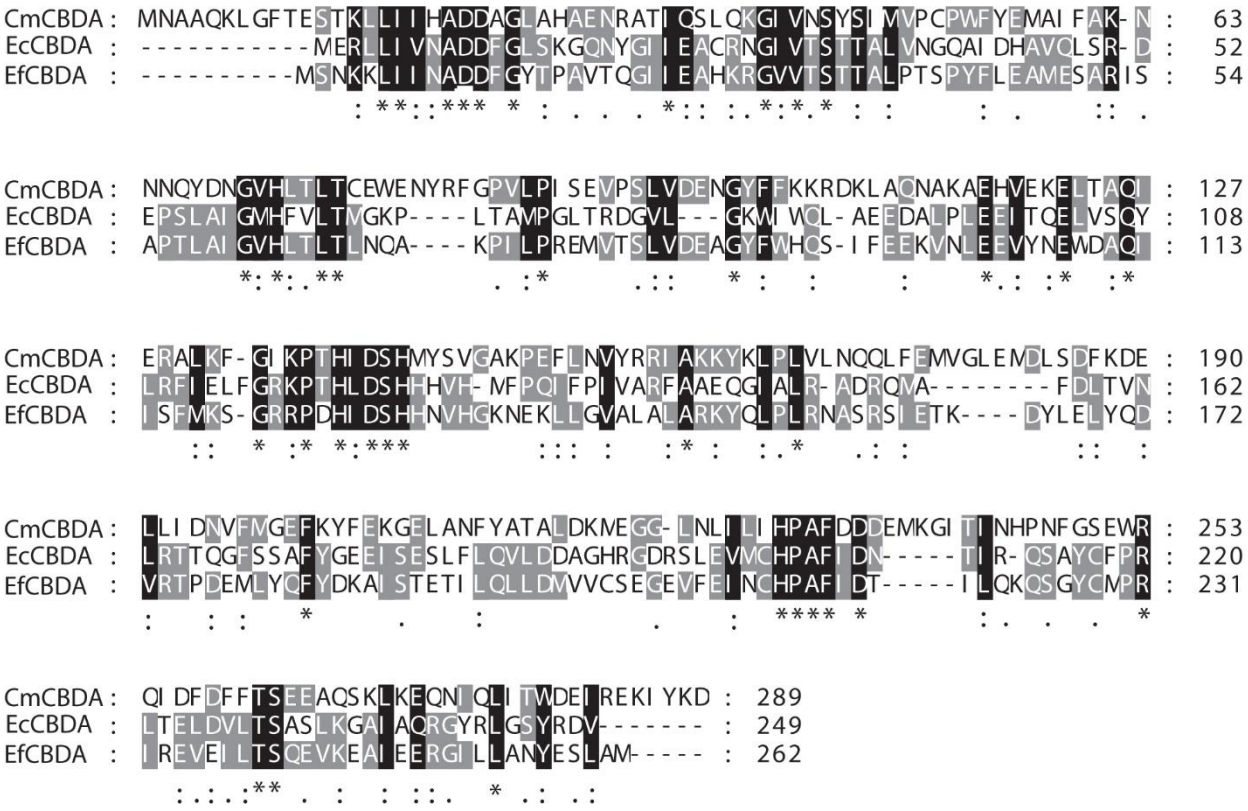
Boc- and Cbz-protected glucosamine were purchased from commercial suppliers. Fmoc- and Alloc-protected glucosamine were synthetized. A solution of glucosamine hydrochloride (20 mg, 0.093 mmol), $NaHCO_3$ (14 mg, 0.186 mmol) and 9-fluorenylmethyl chloroformate or allyl chloroformate (0.140 mmol) in 1 mL $H_2O$/Acetone 1:1 was stirred for 24 h at room temperature. After evaporation of the reaction solvent, the residue was washed several times with diethyl ether. Then, the residue was subjected to purification by silica-gel chromatography using a mixture AcOEt/MeOH 4:1 as the eluent obtaining the desired carbamate (Fmoc-protected glucosamine: 79%, Alloc-protected glucosamine: 72% yield). (NMR spectra and the corresponding annotation are shown in the Supplementary Fig. S11-S12). Reaction mixtures (25 µL) containing 100 mM of N-protected glucosamine, 200 mM of $Na_2HPO_4$ buffer (pH 8.0) and recombinant enzyme CmCBDA (0.013 U), were incubated for 6 h, and then quenched by adding 25 µL of chloroform. After centrifugation, an aliquot of the upper-phase was used for the determination of the produced GlcN after o-phtalaldehyde (OPA) derivatisation using the colorimetric assay in the article section 'Activity and substrate specificity screening'. Screening Results are shown in the Supplementary Figure S7.

## 3) N-acetyl amino acids as CmCBDA substrates

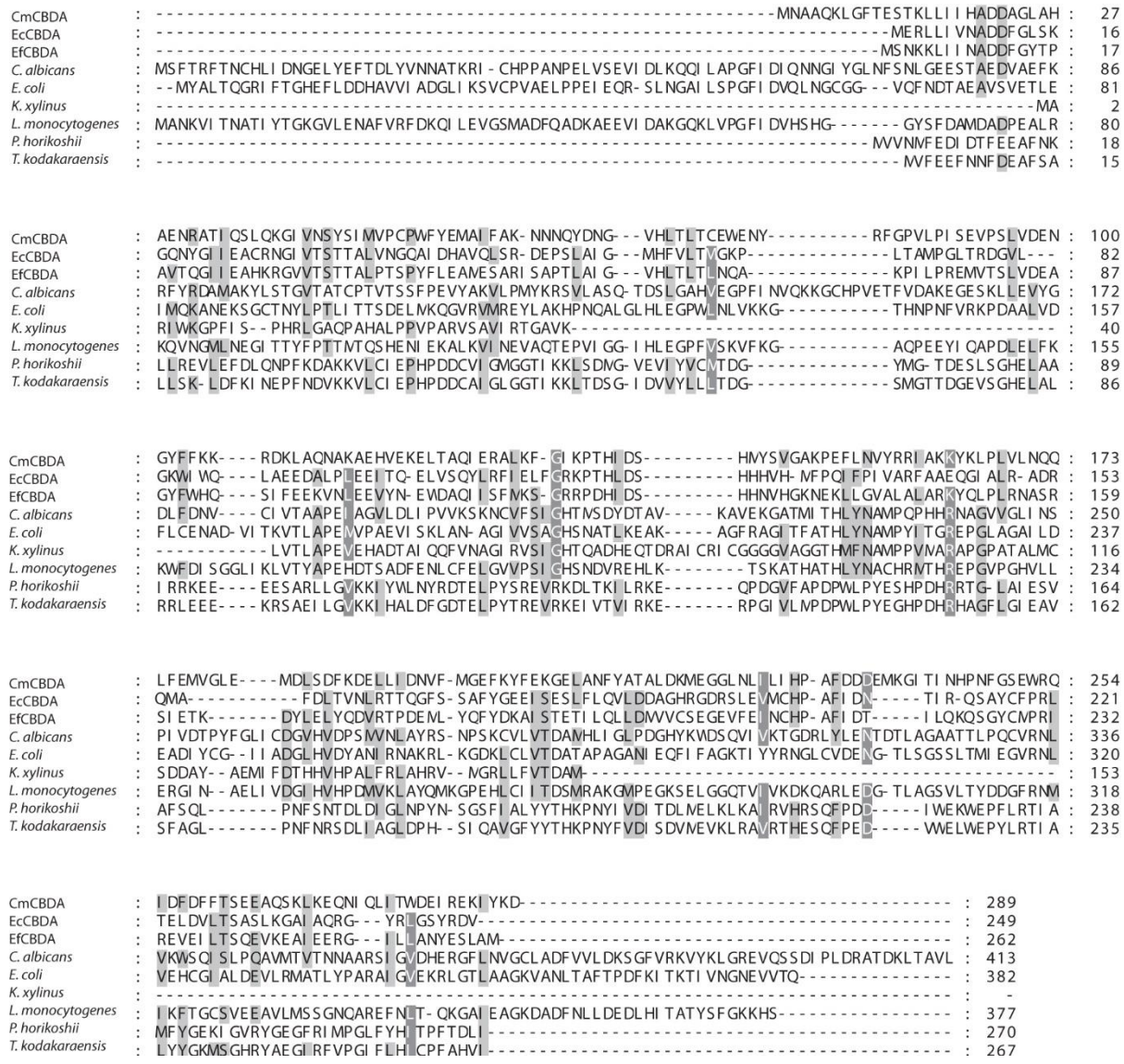
N-acetyl amino acids of all proteinogenic amino acids were purchased from commercial suppliers. Reaction mixtures (25 µL) containing 100 mM of N-acetyl amino acids, 200 mM of $Na_2HPO_4$ buffer (pH 8.0) and recombinant enzyme CmCBDA (0.013 U), were incubated for 6 h, and then quenched by adding 25 µL of chloroform. After centrifugation, an aliquot of the upper-phase was used for the determination of the produced GlcN after o-phtalaldehyde (OPA) derivatisation using the colorimetric assay in the article section 'Activity and substrate specificity screening'. Screening Results are shown in the Supplementary Figure S7.



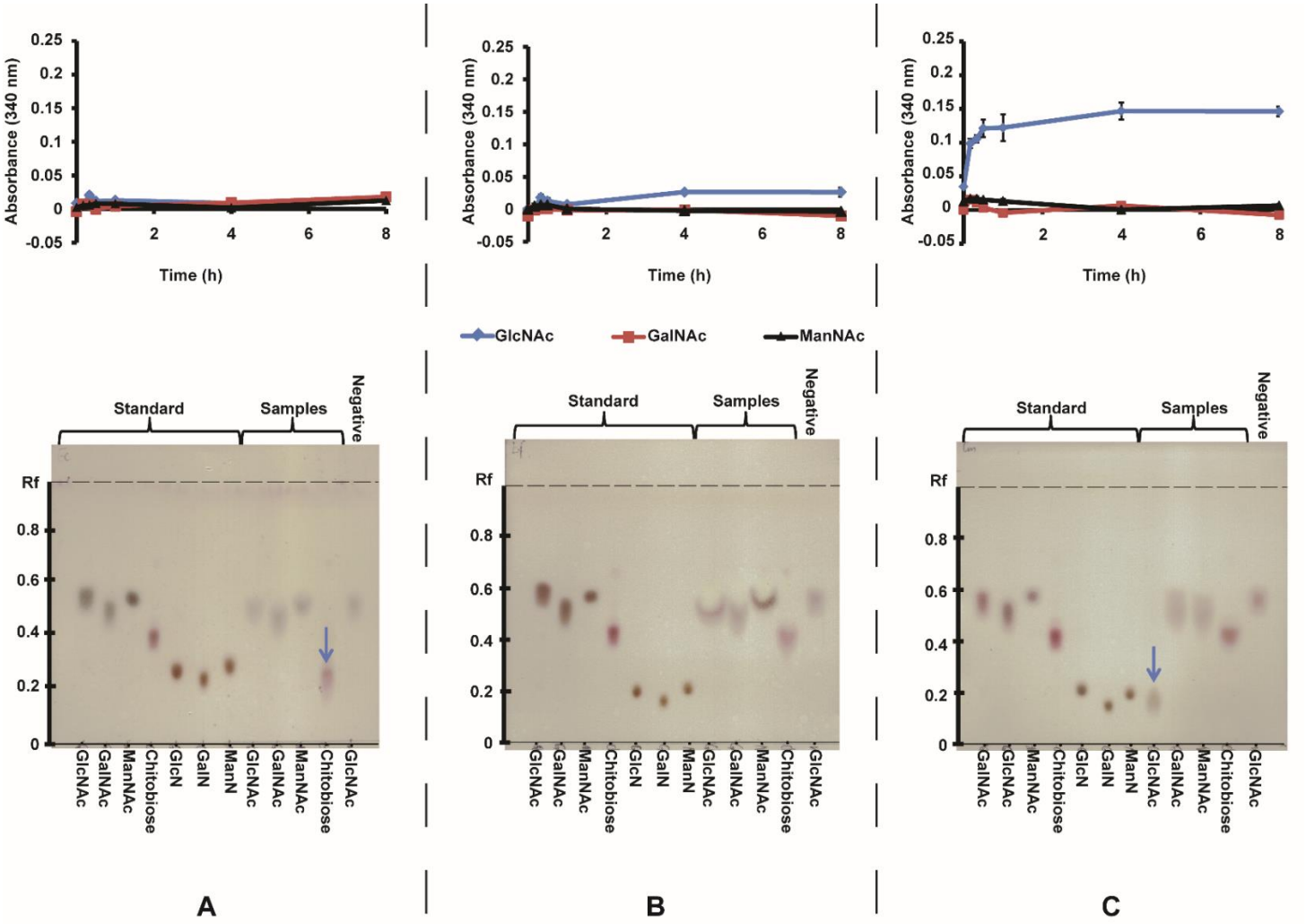

**Supplementary Fig. S1.** Agarose gel electrophoresis of the amplified DNA segments encoding CmCBDA (line-1), EcCBDA (line-2), EfCBDA (line-3); M – DNA marker



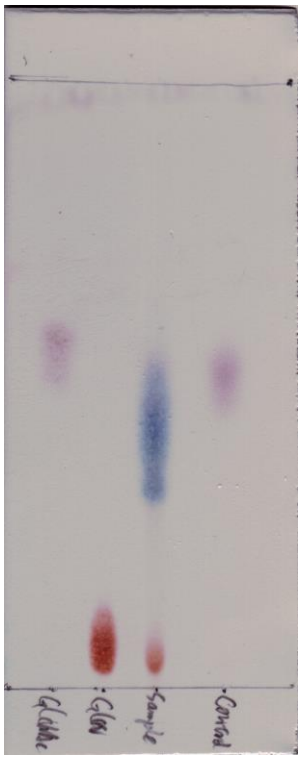

**Supplementary Fig. S2.** Amino acid sequence alignment of CmCBDA, EcCBDA and EfCBDA. Fully conserved amino acids are labelled with asterisk; conserved substitutions with colons, and semi-conserved substitutions with a dot.



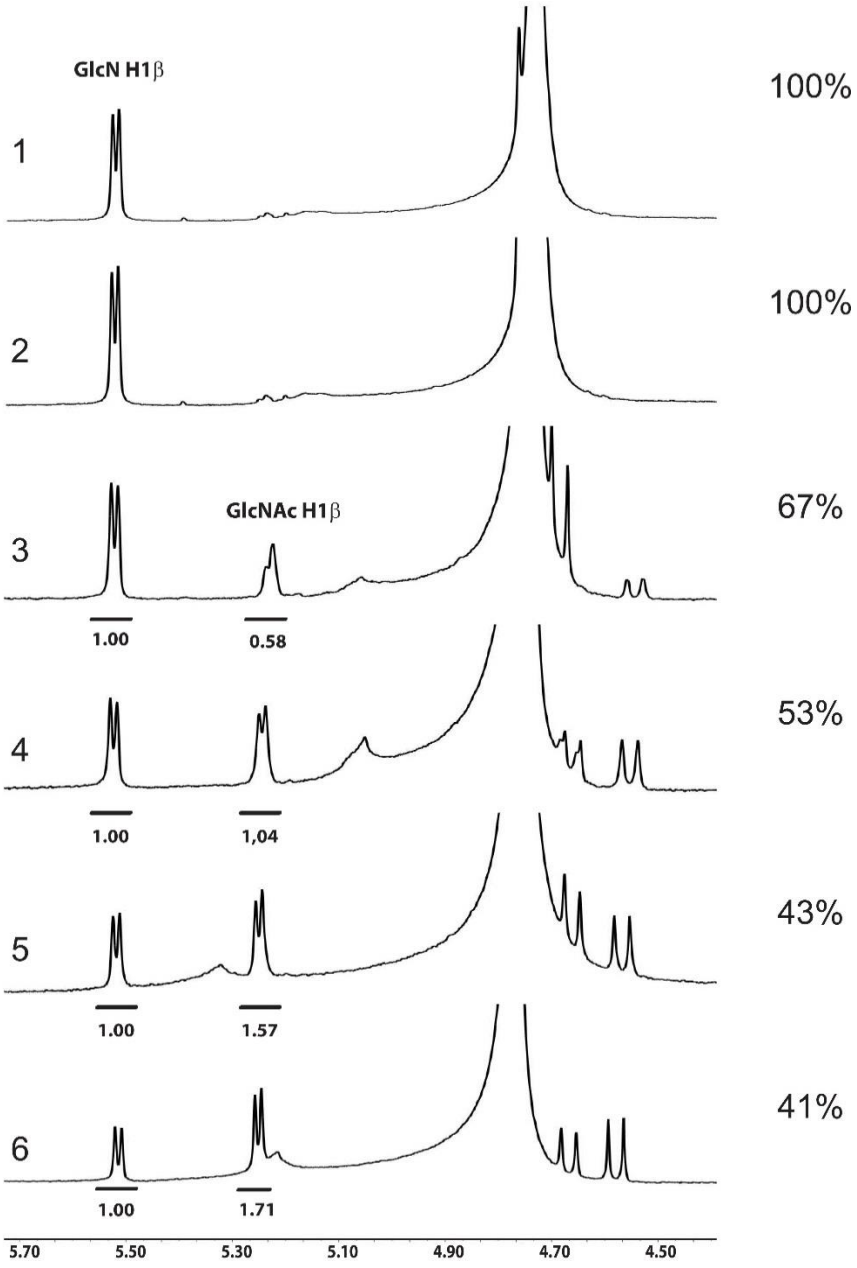

**Supplementary Fig. S3.** Amino acid sequence alignment of *N*-acetylglucosamine deacetylases (CmCBDA, EcCBDA, EfCBDA) and other reported relevant enzymes: *N*-acetylglucosamine 6-phosphate deacetylases from *Candida albicans*, *Escherichia coli*, *Komagataeibacter xylinus*, *Listeria monocytogenes*, *N*,*N'*-diacetylchitobiose deacetylases from *Pyrococcus horikoshii, and Thermococcus kodakaraensis*.



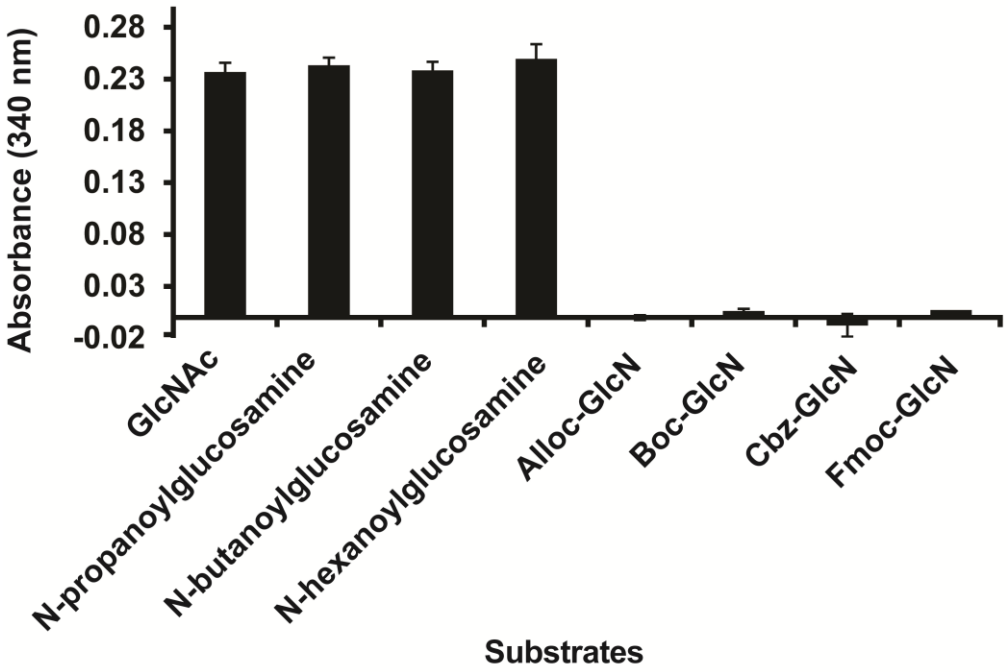

**Supplementary Fig. S4.** Colorimetric and TLC activity tests of **A** EcCDA, **B** EfCDA, or **C** CmCDA for the qualitative analysis of the conversion. Colorimetric tests were done using GlcNAc, GalNAc or ManNAc as substrates after 0 min, 10 min, 20 min, 30 min, 1 h, 4 h, 8 h, 16 h, and 24 h reaction time. TLC activity tests were done using GlcNAc, GalNAc, ManNAc or chitobiose as substrates after 24 h reaction time.



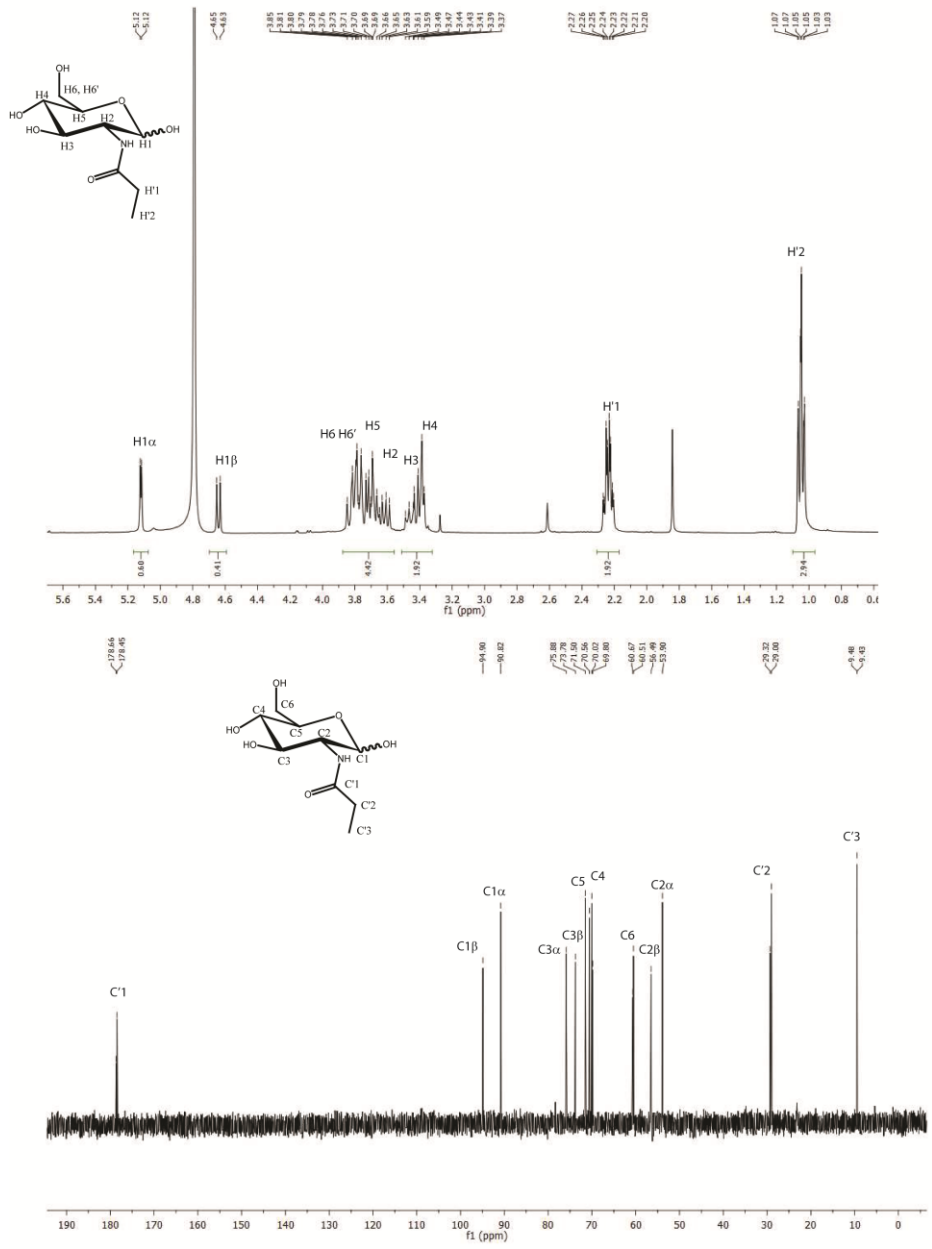

**Supplementary Fig. S5.** TLC result of CmCBDA-catalysed deacetylation of a 50 g/mL solution of GlcNAc. 0.05 U of CmCBDA were used in the reaction. The complete conversion of GlcNAc to GlcN could be observed.



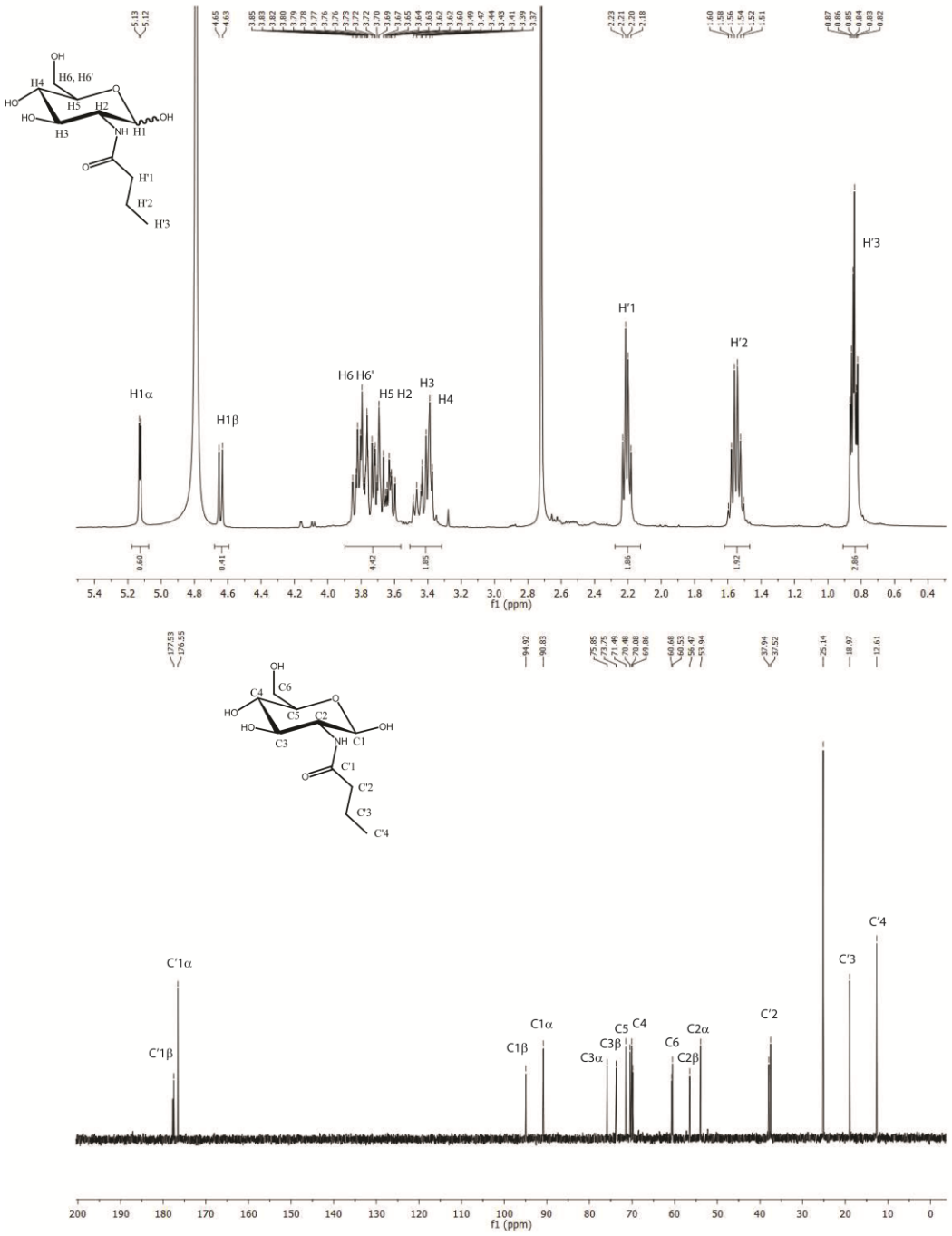

**Supplementary Fig. S6.** $^{1}$H NMR spectra of recycleded CmCBDA beads. The numbers 1 to 6 on the left side indicate the number of 24 h incubation cycles. Bead-immobilized CmCBDA were each time incubated with 100 µL of fresh GlcNAc solution (20 g/mL). The conversion rate (right side) was calculated by integration of the relative areas of the signals corresponding to the anomeric proton of the β anomer of glucosamine (GlcN H1β) and of the proton of the β anomer of GlcNAc (GlcNAc H1β) considering that the α/β ratios of glucosamine and GlcNAc are 0.53:0.47 and 0.61:0.39, respectively.



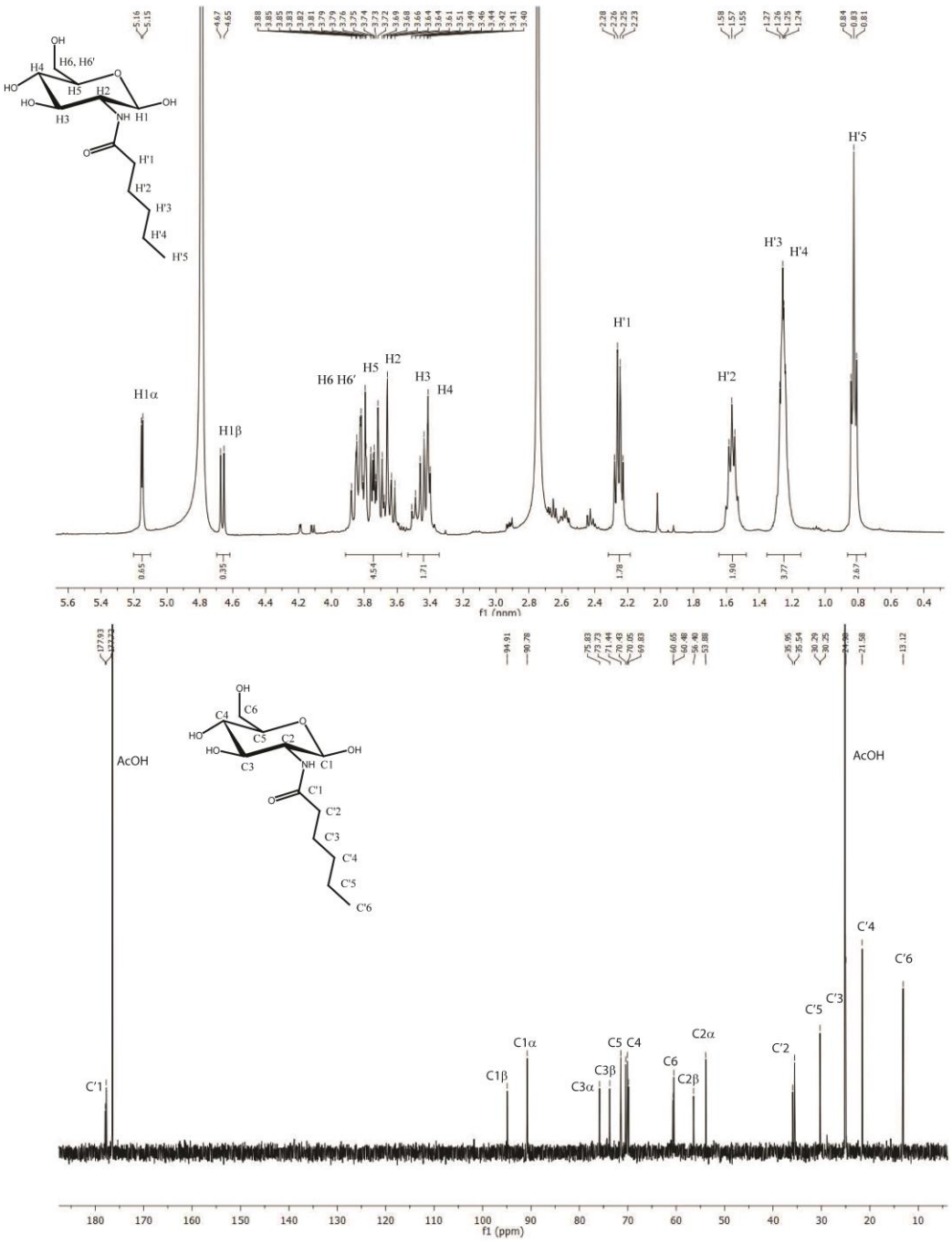

**Supplementary Fig. S7.** Colorimetric activity test of CmCBDA towards N-acylglucosamine derivatives and Alloc-, Boc-, Cbz- and Fmoc-protected glucosamine.




**Supplementary Fig. S8.** $^{1}$H and $^{13}$C NMR spectra of N-propanoyl glucosamine. **$^{1}$H NMR (400 MHz, $D_2O$): δ** (duplicated signals are observed for some protons; asterisks indicate those corresponding to the alpha anomer) 5.12* (d, 1H, *J* = 3.6 Hz), 4.64 (d, 1H, *J* = 8.4 Hz); 3.85-3.58 (m, 4H); 3.49-3.34 (m, 2H); 2.24 (q, 2H, *J* = 7.6 Hz), 2.24* (q, 2H, *J* = 7.6 Hz); 2.23 (q, 2H, *J* = 7.6 Hz); 1.05 (t, 3H, *J* = 7.6 Hz), 1.05* (t, 3H, *J* = 7.6 Hz). **$^{13}$C NMR (100 MHz, $D_2O$): δ** (duplicated signals are observed for some carbons; asterisks indicate those corresponding to the alpha anomer) 178.66, 178.45*; 94.90, 90.82*; 75.88*, 73.78; 71.50*, 70.56; 70.02*, 69.80; 60.67, 60.51*; 56.49, 53.90*; 29.32, 29.00*; 9.48*, 9.43.

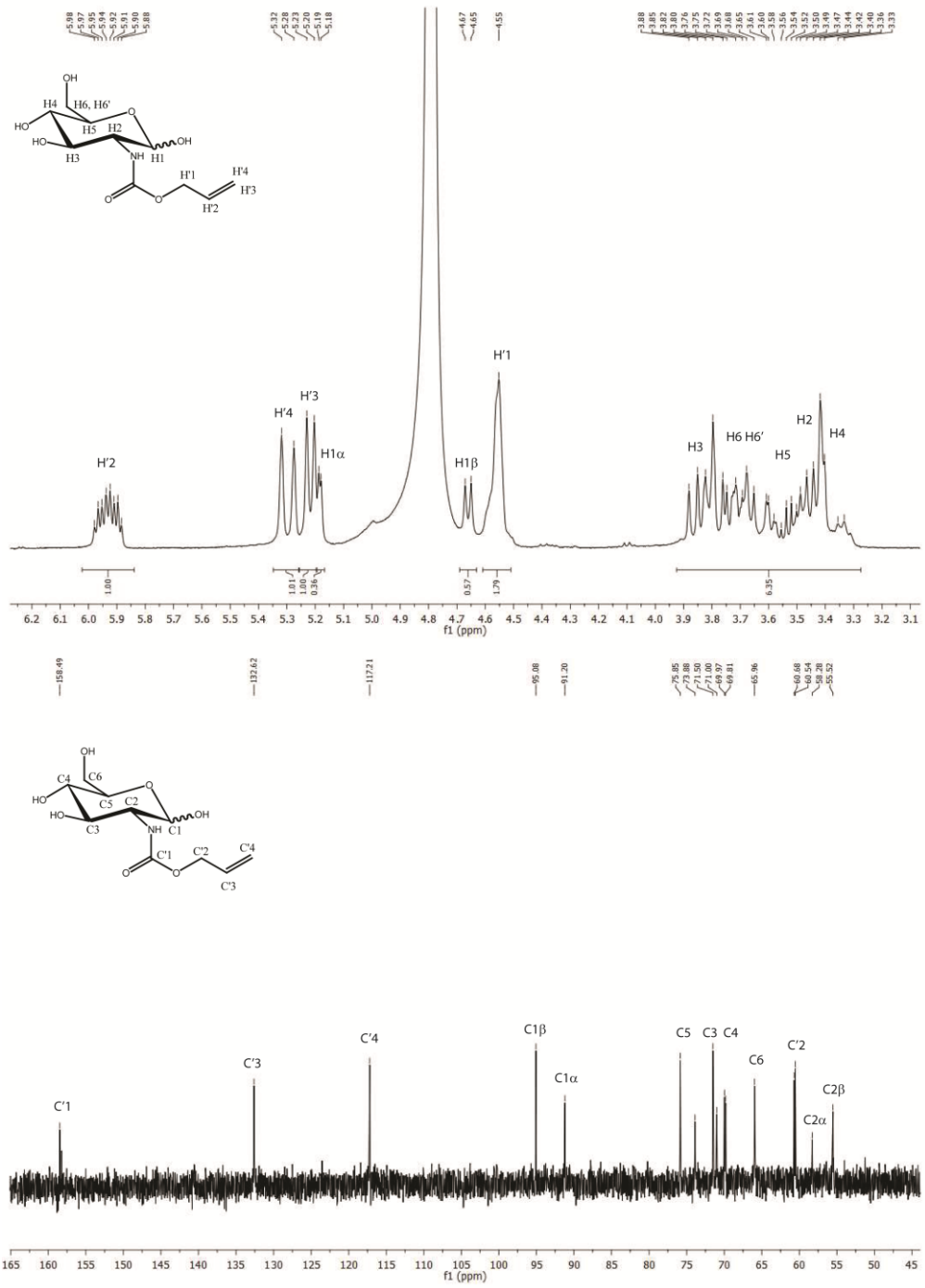



**Supplementary Fig. S9.** $^{1}$H and $^{13}$C NMR spectra of N-butanoyl glucosamine. **$^{1}$H NMR (400 MHz, $D_2O$): δ** (duplicated signals are observed for some protons; asterisks indicate those corresponding to the alpha anomer) 5.12* (d, 1H, *J* = 3.6 Hz), 4.64 (d, 1H, *J* = 8.4 Hz); 3.86-3.59 (m, 4H); 3.49-3.37 (m, 2H); 2.21 (m, 2H); 1.55 (sext, 2H, *J* = 7.2 Hz); 0.85 (t, 3H, *J* = 7.6Hz), 0.84* (t, 3H, *J* = 7.6Hz). **$^{13}$C NMR (100 MHz, $D_2O$): δ** (duplicated signals are observed for some carbons; asterisks indicate those corresponding to the alpha anomer) 177.53, 176.55*; 94.92, 90.83*; 75.85*, 73.75; 71.49, 70.48*; 70.08, 69.86*; 60.68, 60.53*; 56.47, 53.94*; 37.94, 37.52*; 18.97; 12.61.

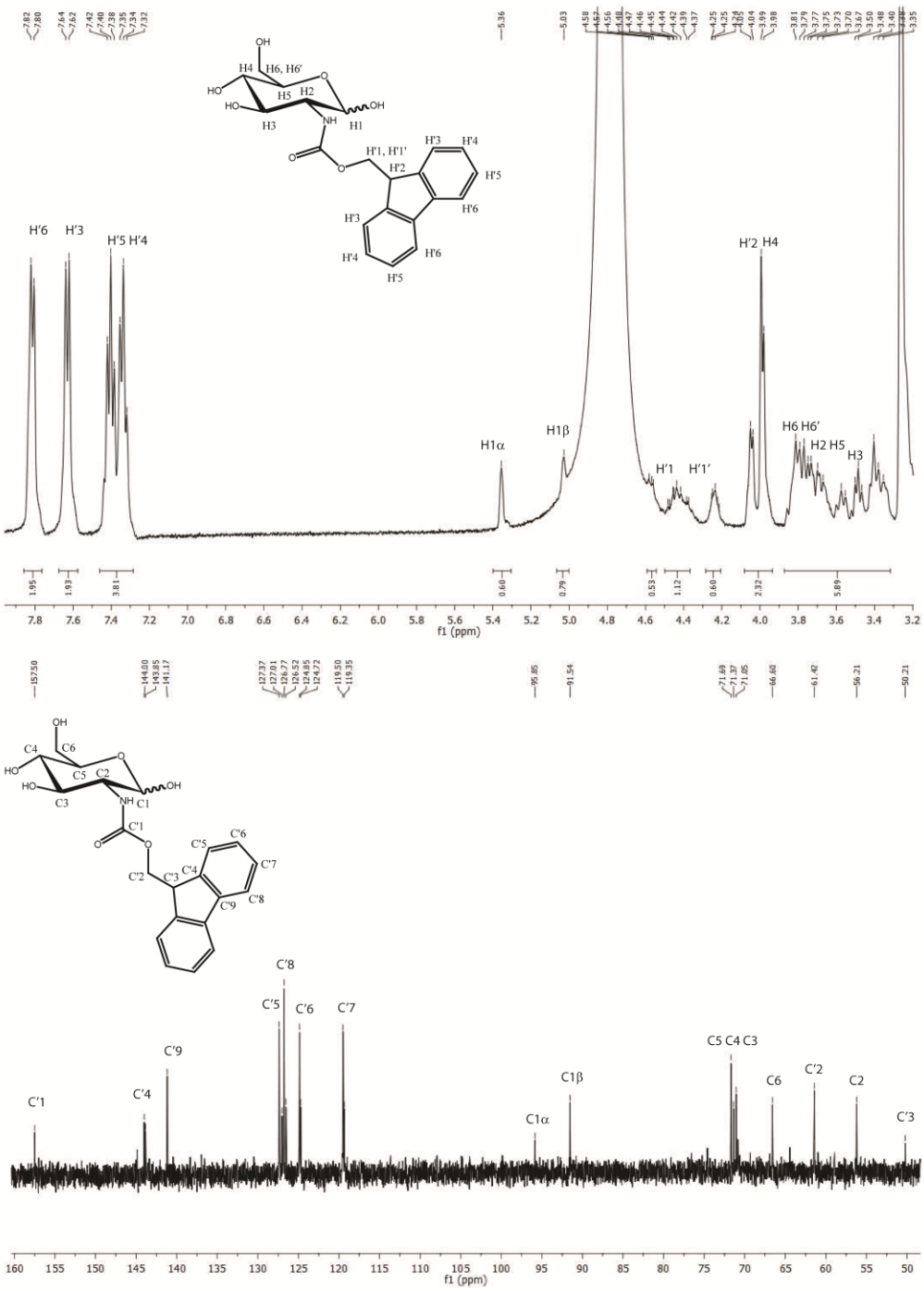



**Supplementary Fig. S10.** $^{1}$H and $^{13}$C NMR spectra of N-hexanoyl glucosamine. **$^{1}$H NMR (400 MHz, $D_2O$): δ** (duplicated signals are observed for some protons; asterisks indicate those corresponding to the alpha anomer) 5.15* (d, 1H, *J* = 3.2 Hz), 4.66 (d, 1H, *J* = 8.4 Hz); 3.89-3.61 (m, 4H); 3.52-3.39 (m, 2H); 2.25 (m, 2H); 1.57 (qn, 2H, *J* = 7.6 Hz); 1.26 (m, 4H); 0.86 (t, 3H, *J* = 6.4 Hz). **$^{13}$C NMR (100 MHz, $D_2O$): δ** (duplicated signals are observed for some carbons; asterisks indicate those corresponding to the alpha anomer) 177.93, 177.71*; 94.91, 90.78*; 75.83*,

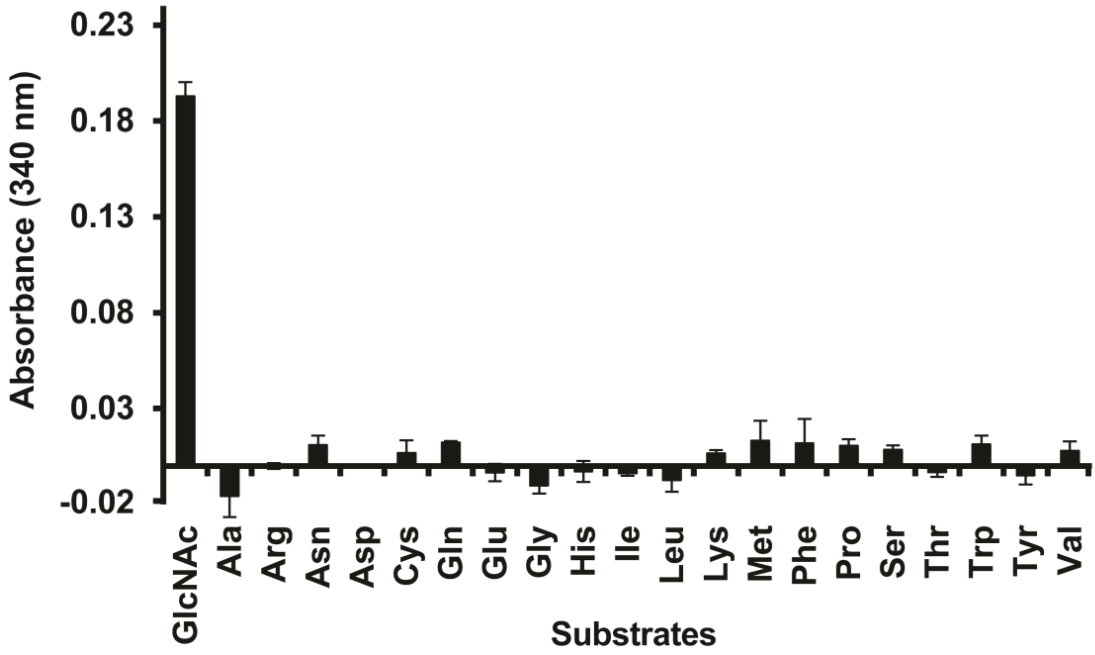
73.73; 71.44*, 70.43; 70.05*, 69.83; 60.65, 60.48*; 56.40, 53.88*; 35.95, 35.54*; 30.29*, 30.25; 24.98; 21.58; 13.12.



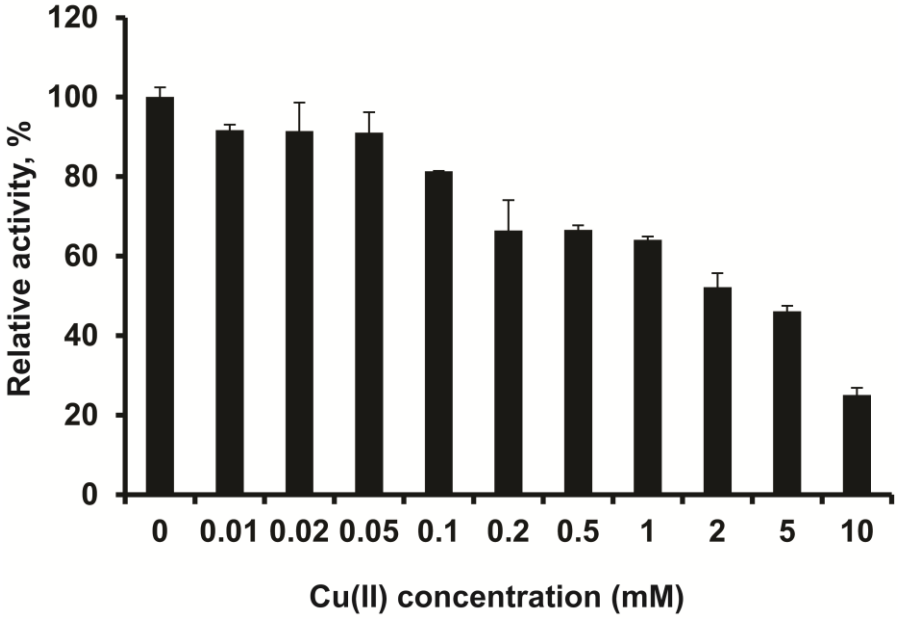

**Supplementary Fig. S11.** $^{1}$H and $^{13}$C NMR spectra of Alloc-protected glucosamine. **$^{1}$H NMR (400 MHz, $D_2O$): δ** (duplicated signals are observed for some protons; asterisks indicate those corresponding to the alpha anomer) 5.93 (m, 1H); 5.30 (d, 1H, *J* = 17.6 Hz); 5.22 (d, 1H, *J* = 10.8 Hz); 5.18* (d, 1H, *J* = 3.2 Hz), 4.66 (d, 1H, *J* = 8.4 Hz); 4.61-4.53 (m, 2H); 3.89-3.29 (m, 6H). **$^{13}$C NMR (100 MHz, $D_2O$): δ** (duplicated signals are observed for some carbons; asterisks indicate those corresponding to the alpha anomer) 198.49; 132.62; 117.21; 95.08, 91.20*; 75.85, 73.88*; 71.50, 71.00*; 69.97, 69.81*; 65.96; 60.68*, 60.54; 58.28*, 55.52.



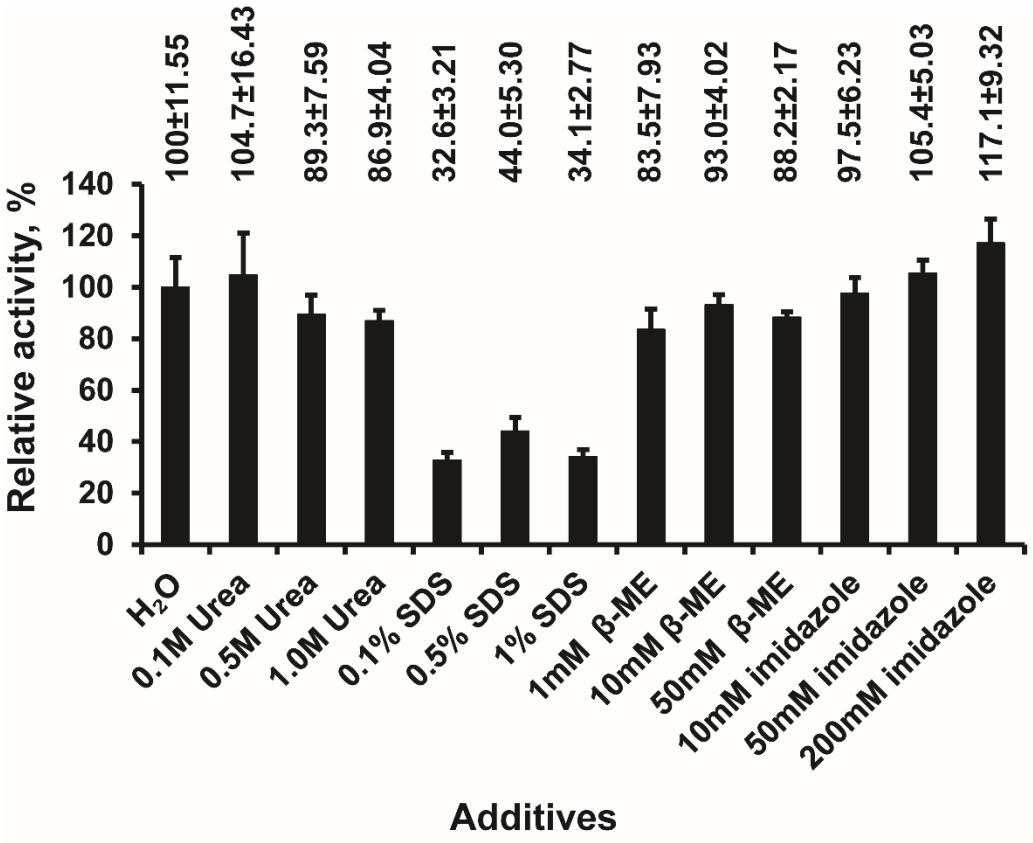

**Supplementary Fig. S12.** $^{1}$H and $^{13}$C NMR spectra of Fmoc-protected glucosamine. **$^{1}$H NMR (400 MHz, $D_2O$/MeOD 1:1): δ** (duplicated signals are observed for some protons; asterisks indicate those corresponding to the alpha anomer) 7.85-7.79 (m, 2H); 7.66-7.61 (m, 2H); 7.45-7.31 (m, 2H); 5.36* (s, 1H), 5.03 (s, 1H); 4.59-4.35 (m, 2H); 4.08-3.96 (m, 2H); 3.87-3.32 (m, 5H). **$^{13}$C NMR (100 MHz, MeOD): δ** (duplicated signals are observed for some carbons; asterisks indicate those corresponding to the alpha anomer) 157.50; 144.00, 143.85*; 141.17; 127.37, 127.01*; 126.77, 126.52*; 124.85, 124.72*; 119.50, 119.35*; 95.85*, 91.54; 71.69; 71.37; 71.05; 66.60; 61.42; 56.21; 50.21.



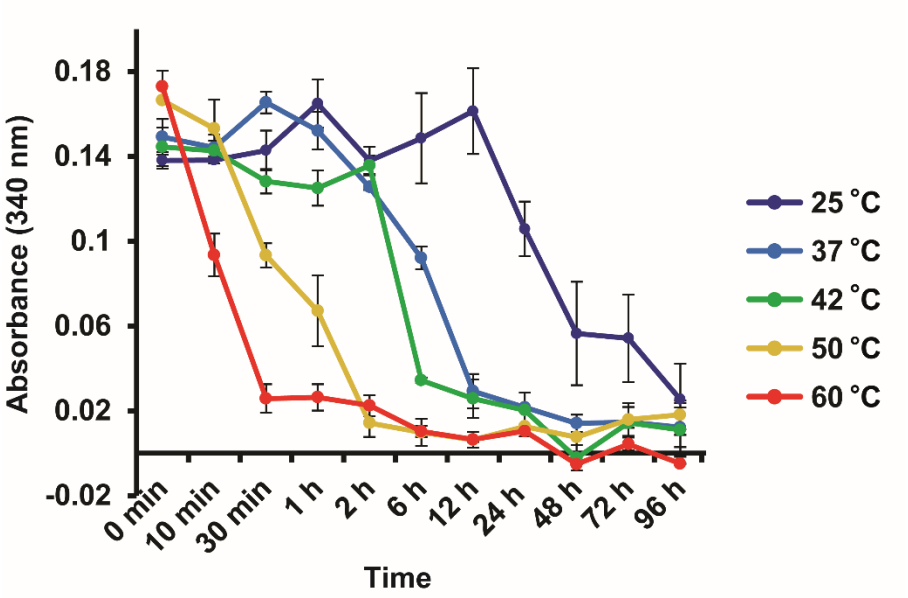

**Supplementary Fig. S13.** Activity of CmCBDA towards GlcNAc and various N-acetyl amino acids.



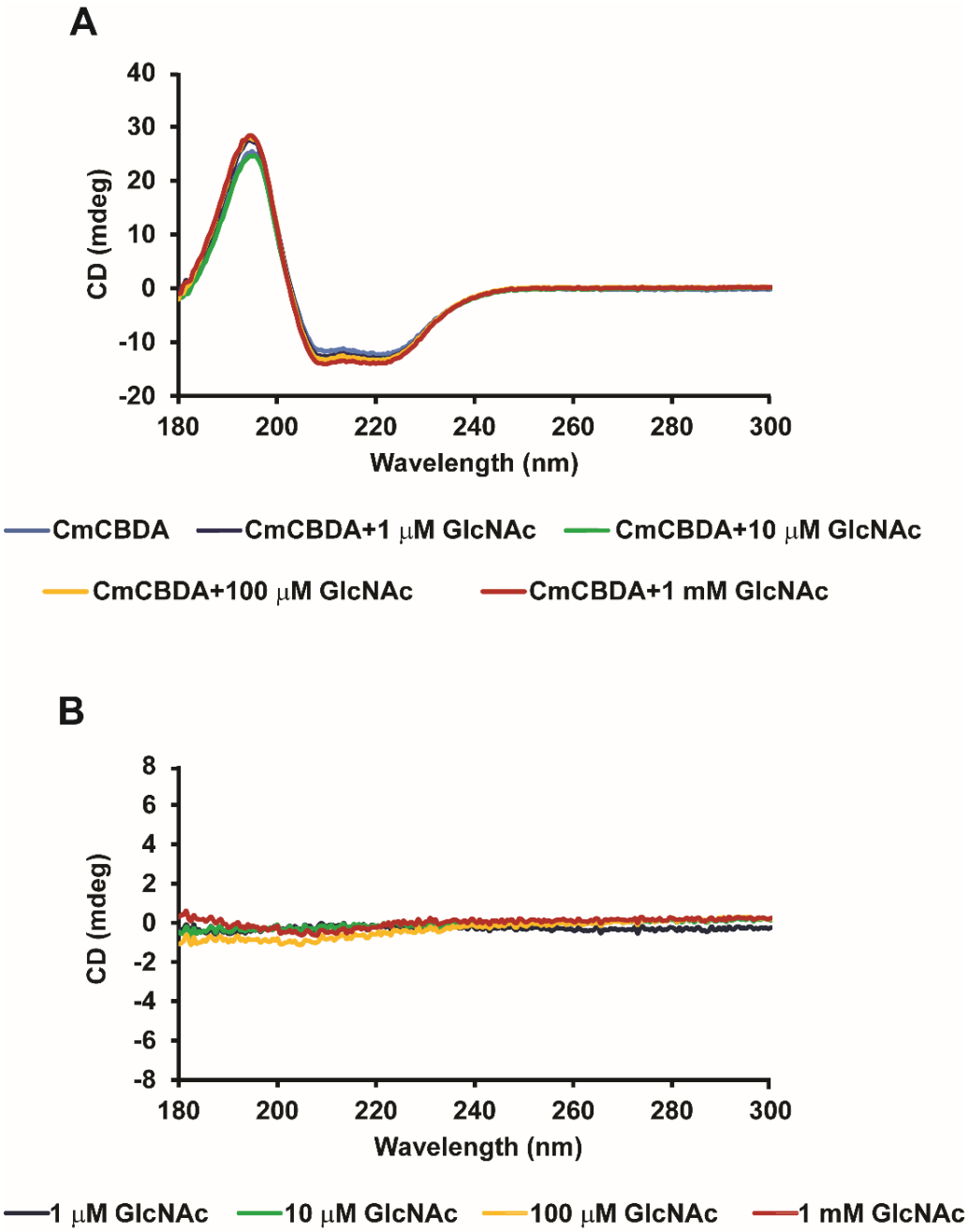

**Supplementary Fig. S14.** Influence of the Cu(II) ion concentration on CmCBDA activity.



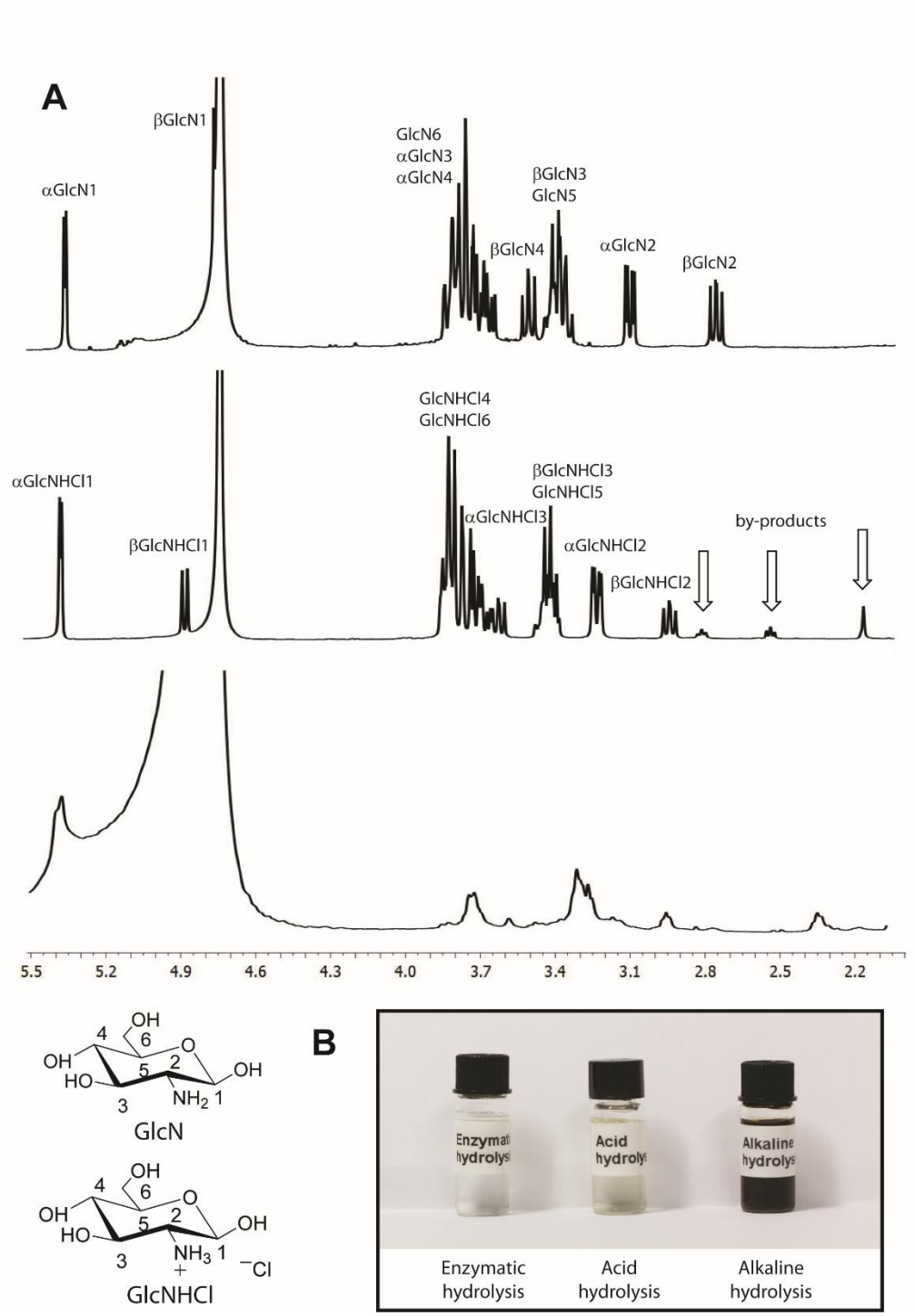

**Supplementary Fig. S15.** Relative activities of recombinant CmCBDA in presence of different concentrations of urea, 2-mercaptoethanol, imidazole and sodium dodecylsulphate (SDS).



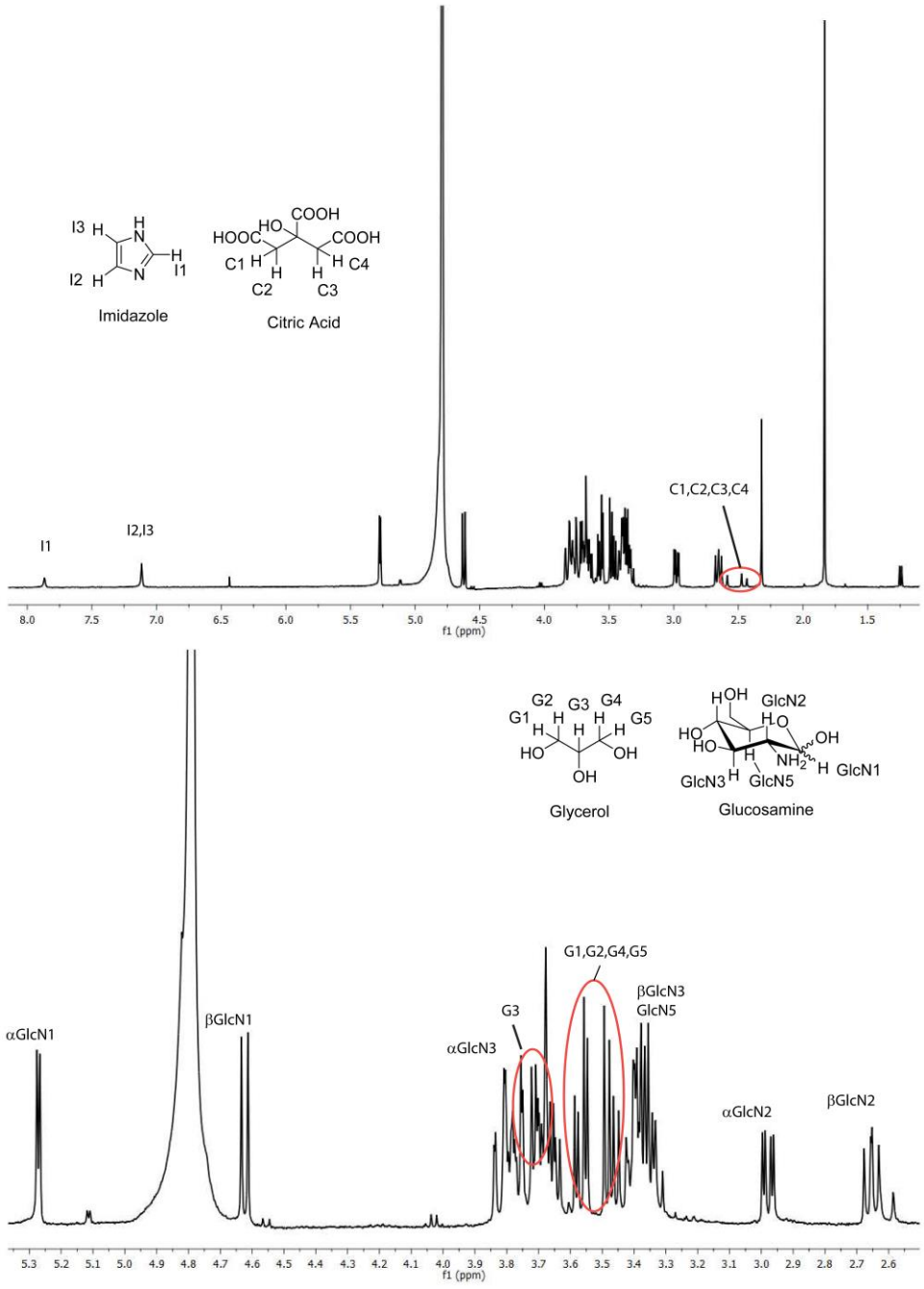

**Supplementary Fig. S16.** Temperature stability of CmCBDA.




**Supplementary Fig. S17.** Circular dichroism (CD) spectroscopy of CmCBDA incubated with various concentrations of GlcNAc (Panel A) and of an enzyme free control mixture (Panel B). Samples were analysed at a constant CmCBDA concentration of 1 mg/ml (in Panel A, Panel B had no CmCBDA added) in a 0.01 cm path length cuvette with different concentrations of GlcNAc ranging from 1 μM to 10 mM. Each sample was recorded 5 times repeat. Sodium phosphate buffer(0.1 M, pH 8.0) was analysed under the same condition and subtracted, which were subsequently averaged to obtain the final result. The secondary structure of protein with different concentrations of GlcNAc was analysed by using CDNN software (version 2.0.3.188). No obvious changes in the secondary structure of CmCBDA were observed with increasing amounts of GlcNAc.




**Supplementary Fig. S18. A** $^{1}$H NMR spectra of CmCBDA-catalysed deacetylation reaction of GlcNAc to GlcN after 12 h (top panel) and acid hydrolysis of GlcNAc to GlcN hydrochloride (GlcN HCl) after 4 h (bottom panel). **B** Visual comparison of the enzymatic hydrolysis reaction with acid hydrolysis reaction and alkaline hydrolysis reaction.




**Supplementary Fig. S19.** $^{1}$H NMR spectrum of the dialysate obtained after combined ZgβHexN2854 (1.25 U) and CmCBDA (5.5 U) treatment of chitinous mushroom extract. Enzymes were added in a dialysis tubing containing 20 mL of the mushrooms extract. The reaction mixture was incubated in 100 mL of water over 72 h. Characteristic signals corresponding to GlcN are indicated. Signals corresponding to citrate, glycerol and imidazole could also be found. Signals corresponding with the hydrogens linked to C-4 and C-6 of GlcN overlap with the G3 signal of glycerol.