Supporting Information

Fluorescence-Activated Cell Sorting and Directed Evolution of α-N-Acetylgalactosaminidases Using a Quenched Activity-Based Probe (qABP)

Kamaladasan Kalidasan, Ying Su, Xiaoyuan Wu, Shao Q. Yao and Mahesh Uttamchandani

Defence Medical and Environmental Research Institute, DSO National Laboratories, 27 Medical Drive, Singapore 117510, Singapore
Department of Chemistry, National University of Singapore, 3 Science Drive 3, Singapore 117543, Singapore
Department of Biological Science, National University of Singapore, 14 Science Drive 4, Singapore 117543, Singapore.

‡ Joint first authors
*Address correspondence to: chmyaosq@nus.edu.sg, mahesh@dso.org.sg

Contents

Page S2-S7: Probe synthesis and characterization of intermediates.
Page S7-S8: General procedures.
Page S9-S11: Results and discussion
Page S12-S20: $^1$H and $^{13}$C NMR
1. Chemistry.

Scheme S1. Detailed synthesis of fluorescently quenched **Probe A**.

(2S,4R,5R,6R)-6-(acetoxymethyl)-3-azidotetrahydro-2H-pyran-2,4,5-triyl triacetate (**2**). Commercial galactosamine (0.54g, 2.5mmol) was dissolved in dry MeOH, followed by addition of K$_2$CO$_3$ (0.485g, 3.5mmol), CuSO$_4$·5H$_2$O (6.25mg, 0.025mmol) and 1H-imidazole-1-sulfonyl azide (0.625g, 3mmol). The above mixture was stirred at r.t. for 18 h and concentrated in vacuo. The residue was redissolved in 5ml of Ac$_2$O with 2ml pyridine and stirred at r.t. for another 3 h, after which the solvent was removed in vacuo and H$_2$O was added to stop the reaction. The residue was extracted with ethyl acetate and separated by flash chromatography (EA/Hex=1/2) to give **2** (0.65g, 70%). $^1$H-NMR (500 MHz, CDCl$_3$) δ 6.26 (d, $J$ = 3.15 Hz, 0.35H), 5.51 (d, $J$ = 8.15 Hz, 0.65H), 5.42 (d, $J$ = 3.15 Hz, 0.35H), 5.32 (d, $J$ = 3.15 Hz, 0.65H), 5.25 (dd, $J_1$ = 11.05 Hz, $J_2$ = 3.45 Hz, 0.35H), 4.87 (dd, $J_1$ = 10.7 Hz, $J_2$ = 3.15 Hz, 0.65H), 4.25-4.22 (m, 0.4H), 4.11-4.02 (m, 2H), 4.01-3.97 (m, 0.6H), 4.75 (dd, $J_1$ = 10.73 Hz, $J_2$ = 3.78 Hz, 0.35H), 3.88 (dd, $J_1$ = 10.73 Hz, $J_2$ = 8.85 Hz, 0.65H), 2.14 (s, 1.8H), 2.11 (s, 1.2H), 2.10 (s, 3H), 2.01 (s, 1.2H), 2.00 (s, 1.8H), 1.97 (s, 1.2H), 1.97 (s, 1.8H). $^{13}$C-NMR (125 MHz, CD$_3$CCl$_3$) δ 170.1, 169.7,
169.6, 169.4, 168.3, 92.6, 90.2, 71.5, 71.1, 68.5, 66.7, 66.1, 60.9, 60.8, 59.5, 56.7, 20.6, 20.6, 20.4, 20.3.

(2R,3R,4R,5R)-2-(acetoxyethyl)-5-azido-6-hydroxymethyl-tetrahydro-2H-pyran-3,4-diyl diacetate (3). 2 (0.49g, 1.3mmol) was dissolved in 2ml of dry DMF, followed by addition of hydrazine acetate (0.32g, 3.25mmol). The above mixture was stirred at r.t. for 0.5h and concentrated in vacuo. The residue was redissolved in 10ml of ethyl acetate and washed with H2O, brine, and dried on Na2SO4. The crude product was purified by flash column (EA/Hex=5/1) to obtain 3 (0.31 g, 60%). 1H-NMR (500 MHz, CDCl3) δ 5.44 (d, J = 3.15 Hz, 0.6H), 5.40 (d, J = 3.8 Hz, 0.85H), 5.37 (d, J = 3.8 Hz, 0.27H), 5.32 (d, J = 3.15 Hz, 0.36H), 4.81 (dd, J1 = 10.7 Hz, J2 = 3.15 Hz, 0.37H), 4.69 (d, J = 8.2, 0.38H), 4.45 (t, J = 6.62, 0.63H), 4.12-4.05 (m, 2H), 3.90 (t, J = 6.3, 0.47H), 3.72 (dd, J1 = 11.025, J2 = 3.48, 0.64H), 3.64 (dd, J1 = 10.725, J2 = 7.6, 0.4H), 2.13 (s, 3H), 2.04 (s, 3H), 2.04 (s, 3H).

13C-NMR (125 MHz, CDCl3) δ 170.6, 170.6, 170.1, 169.9, 96.3, 92.3, 71.1, 70.8, 68.3, 67.7, 66.4, 61.9, 61.7, 61.5, 58.0, 20.6, 20.5, 20.5.

(2R,3R,4R,5R,6S)-2-(acetoxyethyl)-5-azido-6-(2,2,2-trichloro-1-iminoethoxy)tetrahydro-2H-pyran-3,4-diyl diacetate (4). 3 (0.31g, 0.92mmol) was dissolved in 5ml dry DCM on ice bath, followed by addition of DBU (135μl, 0.92mmol), and stirred at 0°C for 10 min. Subsequently, 2,2,2-trichloroacetoniitrile (0.93ml, 9.2mmol) was added dropwise, and the mixture was stirred for another 2 h. The solvent was remove in vacuo and the resulting mixture was purified by flash chromatography to obtain 4 (0.32g, 71%). 1H-NMR (500 MHz, CDCl3) δ 8.78 (s, 1H), 6.46 (d, J = 3.15 Hz, 1H), 5.49 (d, J = 1.90 Hz, 1H), 5.33 (dd, J1 = 3.88 Hz, J2 = 3.15 Hz, 1H), 4.37 (t, J = 6.625 Hz, 1H), 4.10 (dd, J1 = 11.35 Hz, J2 = 6.95 Hz, 1H), 4.03-3.98 (m, 2H), 2.12 (s, 3H), 2.02 (s, 3H), 1.95 (s, 3H). 13C-NMR (125 MHz, CDCl3) δ 170.0, 169.7, 169.5, 160.4, 94.3, 90.4, 68.9, 68.5, 66.7, 61.0, 56.9, 20.4, 20.4, 20.3.

(2R,3R,4R,5R,6R)-2-(acetoxyethyl)-6-(4-(2-(allyloxy)2-oxoethyl)amino)-1-hydroxy-2-oxoethylphenoxy)-5-azidotetrahydro-2H-pyran-3,4-diyl diacetate (5). Flame-dried Schlenk flask was filled with 200mg of oven-dried 4Å
MS and 20ml of dry DCM. 4 (0.26 g, 0.55 mmol) was added and the mixture was cooled on ice bath, followed by addition of allyl 2-(2-hydroxy-2-(4-hydroxyphenyl)acetamido)acetate (0.13g, 0.5mmol). The mixture was further stirred at -20°C for 30 min, followed by dropwise addition of TMSOTf (110 µl, 0.62mmol) over a period of 20 min. The reaction was continued on ice-salt bath for another 2h. Finally, the solvent was removed in vacuo and resulting mixture was purified by flash chromatography to give 5 (0.165 g, 53%). 

1H-NMR (500 MHz, CDCl3) δ 7.45 (d, J =8.2 Hz, 2H), 7.12 (d, J = 8.5 Hz, 2H), 5.94-5.88 (m, 1H), 5.76 (d, J = 2.75 Hz, 1H), 5.51 (s, 1H), 5.33-5.20 (m, 2H), 5.05 (s, 1H), 4.61 (d, J = 4.35 Hz, 1H), 4.40 (m, 1H), 4.23-3.93 (m, 6H), 2.15 (s, 3H), 2.05 (s, 3H), 1.91 (s, 3H).

13C-NMR (125 MHz, CDCl3) δ 174.6, 170.7, 170.5, 170.0, 169.4, 156.0, 135.3, 131.9, 128.4, 116.7, 116.2, 114.7, 99.8, 96.9, 73.5, 71.9, 70.7, 68.3, 67.5, 66.6, 65.4, 61.5, 60.8, 60.1, 57.4, 40.3, 19.3, 19.1.

(2R,3R,4R,5R,6R)-5-acetamido-2-(acetoxymethyl)-6-(4-(2-((2-(allyloxy)-2-oxoethyl)amino)-1-hydroxy-2-oxoethyl)phenoxy)tetrahydro-2H-pyran-3,4-diyl diacetate (6). 5 (0.165 g, 0.31 mmol) was dissolved in AcSH and the reaction was stirred at r.t. for 12 h. Upon solvent removal, the residue was redissolved in ethyl acetate and purified by flash chromatography to give 6 (80 mg, 47%). 

1H-NMR (500 MHz, MeOD) δ 7.44 (d, J =8.5 Hz, 2H), 7.11 (d, J = 8.2 Hz, 2H), 5.95-5.87 (m, 1H), 5.61 (d, J =3.4 Hz, 1H, α form), 5.50 (d, J =5.5 Hz, 1H), 5.38 (dd, J1 =11.55 Hz, J2 =3.15 Hz, 1H), 5.33-5.20 (m, 2H), 5.04 (d, J =1.5 Hz, 1H), 4.61 (d, J =5.65 Hz, 2H), 4.37 (t, J =6.375 Hz, 1H), 4.15-3.99 (m, 5H), 2.16 (s, 3H), 1.99 (s, 3H), 1.95 (s, 3H), 1.90 (d, J =3.4 Hz, 3H).

13C-NMR (125 MHz, CDCl3) δ 174.6, 172.4, 171.5, 170.7, 170.5, 169.3, 156.3, 134.9, 131.9, 128.3, 117.3, 116.7, 96.6, 73.5, 68.1, 67.3, 66.6, 65.4, 61.6, 60.1, 40.4, 29.899, 21.2, 21.1, 19.6, 19.3.

(2R,3R,4R,5R,6R)-5-acetamido-2-(acetoxymethyl)-6-(4-(2-((2-(allyloxy)-2-oxoethyl)amino)-1-(((4-nitrophenoxy)carbonyl)oxy)-2-oxoethyl)phenoxy)tetrahydro-2H-pyran-3,4-diyl diacetate (7). 6 (80 mg, 0.14 mmol) was dissolved in dry DCM, followed by slow addition of pyridine (1ml, 0.87mmol) and p-nitrochloroformate (60mg, 0.29 mmol). The reaction was stirred at r.t. for 6 h. Upon solvent removal, the residue was redissolved in ethyl acetate and purified by flash chromatography to give 7 (22.0 mg, 22%). The unstable product was used immediately in the next step.
(2R,3R,4R,5R,6R)-5-acetamido-2-(acetoxyethyl)-6-(4-((4-((E)-(4-(dimethylamino)phenyl)diazenyl)phenyl)-1,7,10,13-tetraoxo-8,14-dioxao-2,6,11-triazaheptadec-16-en-9-yl)phenoxy)tetrahydro-2H-pyran-3,4-diyl diacetate (8). The reaction was stirred at r.t. for 2 h. Upon solvent removal, the residue was redissolved in DCM and purified by flash chromatography to give 8 (33 mg, 90%). 1H-NMR (500 MHz, MeOD) δ 7.93 (d, J = 8.0Hz, 2H), 7.86-7.81 (m, 4H), 7.49 (d, J = 8.80Hz, 2H), 7.14 (d, J = 8.85Hz, 2H), 6.81 (d, J = 8.85, 2H), 5.96 (s, 1H), 5.92-5.84 (m, 1H), 5.63 (d, J = 3.8Hz, 1H, α form), 5.49 (d, J = 3.15Hz, 1H), 5.38 (dd, J1 =11.35 Hz, J2 =3.15 Hz, 1H), 5.30-5.18 (m, 2H), 4.59 (d, J = 5.70Hz, 2H), 4.33 (t, J = 5.70Hz, 1H), 4.12-4.02 (m, 4H), 3.45 (t, J = 6.95Hz, 2H), 3.33-3.32 (m, 1H), 3.25 (dd, J1 =12.60 Hz, J2 =6.3 Hz, 2H), 3.09 (s, 6H), 2.16 (s, 3H), 2.00 (s, 3H), 1.97 (s, 3H), 1.88 (s, 3H). 13C-NMR (125 MHz, CDCl3) δ 173.7, 172.4, 172.0, 172.0, 171.8, 170.5, 169.5, 158.2, 157.2, 156.3, 154.5, 144.7, 135.6, 133.2, 131.9, 130.4, 130.3, 130.2, 129.2, 126.4, 122.9, 118.6, 118.1, 118.1, 112.6, 97.8, 76.4, 69.4, 68.7, 68.5, 66.7, 66.3, 41.8, 40.3, 39.4, 38.3, 30.5, 22.4, 20.6, 20.5. LC-MS (IT-TOF) calcd for [M+H]+: 946.383, Found 946.352.

9-(4-(((2R,3R,4R,5R,6R)-3-acetamido-4,5-diacetoxy-6-(acetoxyethyl)tetrahydro-2H-pyran-2-yl)oxy)phenyl)-1-(4-((E)-(4-(dimethylamino)phenyl)diazenyl)phenyl)-1,7,10-trioxo-8-oxa-2,6,11-triazaatridecan-13-oic acid (9). The reaction was stirred at r.t. for 2 h. Upon solvent removal, the residue was redissolved in DCM and purified by flash chromatography to give 9 (25 mg, 90%). 1H-NMR (500 MHz, CD3OD/CDCl3=10/1) δ 7.89 (d, J = 8.10Hz, 2H), 7.80 (d, J = 9.05Hz, 2H), 7.76 (d, J = 7.40Hz, 2H), 7.43 (d, J = 7.95Hz, 2H), 7.06 (d, J = 7.95Hz, 2H), 6.75 (d, J = 8.90, 2H), 5.92 (s, 1H), 5.57 (d, J = 3.8 Hz, 1H), 5.44 (d, J = 4.65Hz, 1H), 5.34 (d, J = 11.4Hz, 1H), 4.61 (d, J = 13.85Hz, 1H), 4.27 (s, 1H), 4.04-3.99 (m, 2H), 3.85-3.70 (m, 2H), 3.41-3.21 (m, 4H), 3.05 (s, 6H), 2.13 (s, 3H), 1.97 (s, 3H), 1.92 (s, 3H), 1.82 (s, 3H). 13C-NMR (125 MHz, MeOD/CDCl3=10/1) δ 173.5, 171.9, 171.9, 171.7, 169.4, 158.0, 157.2, 156.2, 154.4, 144.6, 135.4, 132.2, 130.1, 129.2, 126.3, 122.9, 118.0, 112.51, 97.7, 69.3, 68.79, 68.6, 68.4, 62.8, 40.4, 26.3, 22.5, 20.6, 20.6, 20.5. LC-MS (IT-TOF) calcd for [M+H]+: 906.352, Found 906.323.
(4)5-(((2R,3R,4R,5R,6R)-3-acetamido-4,5-diacetoxy-6-(acetoxyethyl)tetrahydro-2H-pyran-2-yl)oxy)phenyl)-1-(4-((E)-(4-(dimethylamino)phenyl)diazemlphenyl)-1,7,10,13-tetraoxo-8-oxa-2,6,11,14-tetraazaheptadecan-17-yl)carbamoyl)-2-(6-hydroxy-3-oxo-3,10-dihydroanthracen-9-yl)benzoic acid (10). 9 (25.0 mg, 0.028 mmol) was dissolved in 2ml DMF, followed by addition of the fluorescent derivative (14mg, 0.03mmol) and HATU (10.5 mg, 0.028mmol). The reaction was stirred at r.t. overnight. The crude product was purified by flash chromatography to give 10 (22 mg, 61%). $^1$H-NMR (500 MHz, CD$_3$OD) $\delta$ 8.15-8.05 (m, 2H), 7.88 (t, $J$ = 8.2Hz, 2H), 7.81-7.74 (m, 4H), 7.46-7.41 (m, 2H), 7.10 (t, $J$ = 7.55Hz, 2H), 6.78 (t, $J$ = 9.45Hz, 2H), 5.75 (d, $J$ = 3.80Hz, 1H), 5.60 (d, $J$ = 2.55Hz, 1H), 5.45 (s, 1H), 5.35 (d, $J$ = 8.80Hz, 1H), 4.60 (d, $J$ = 12.0 Hz, 1H), 4.29 (s, 1H), 4.08-3.68 (m, 6H), 3.34-3.18 (m, 10H), 3.07 (s, 6H), 2.14 (s, 3H), 1.97 (s, 3H), 1.94 (s, 3H), 1.84 (s, 3H). $^{13}$C-NMR (125 MHz, CD$_3$OD) $\delta$ 173.8, 172.9, 172.8, 172.0, 171.8, 171.5, 171.4, 171.0, 169.6, 168.4, 168.2, 258.2, 157.8, 157.5, 156.3, 154.6, 144.7, 137.6, 135.6, 131.1, 130.5, 130.4, 130.0, 129.2, 126.4, 122.982, 118.2, 118.1, 112.6, 103.7, 97.7, 69.4, 68.7, 68.5, 62.9, 55.8, 40.3, 38.4, 37.9, 30.5, 22.4, 20.6, 20.5, 20.5. LC-MS (ESI) calcd for [M+H]$^+$: 1320.47, Found: 1320.65; calcd for [M+2H]$^{2+}$/2: 660.74, Found: 660.65.

Electronic Supplementary Material (ESI) for Chemical Communications
This journal is © The Royal Society of Chemistry 2013

Electronic Supplementary Material (ESI) for Chemical Communications
This journal is © The Royal Society of Chemistry 2013

5-(((2R,3R,4R,5R,6R)-3-acetamido-4,5-dihydroxy-6-(hydroxymethyl)tetrahydro-2H-pyran-2-yl)oxy)phenyl)-1-(4-((E)-(4-(dimethylamino)phenyl)diazemlphenyl)-1,7,10,13-tetraoxo-8-oxa-2,6,11,14-tetraazaheptadecan-17-yl)carbamoyl)-2-(6-hydroxy-3-oxo-3,10-dihydroanthracen-9-yl)benzoic acid (Probe A). 10 (10.2 mg, 0.008 mmol) was dissolved in 2ml CH$_3$OH, followed by addition of NaOCH$_3$ (2.5 mg, 0.046mmol). The reaction was stirred at r.t. for 1h. The crude product was purified on semi-preparative HPLC to give Probe A (2.5 mg, 28%). $^1$H-NMR (500 MHz, DMSO/H$_2$O = 3/4) $\delta$ 8.24-8.22 (m, 1H), 7.97-7.93 (m, 2H), 7.75-7.54 (m, 6H), 7.33-7.27 (m, 2H), 7.04-7.01 (m, 2H), 6.74-6.46 (m, 7H), 5.65-5.58 (m, 1H), 5.34 (d, $J$ = 2.00Hz, 1H), 4.13-4.15 (m, 1H), 3.86-3.44 (m, 6H), 3.28-3.02 (m, 10H), 2.96 (s, 6H), 1.83 (s, 3H). Chemical Formula: C$_{61}$H$_{63}$N$_9$O$_{17}$, HR-MS for [M+Na]$^+$ calcd 1216.4234, found 1216.4234.
2. Biology and Screening.

2.1 Cloning and Generation of Random-Mutagenesis Library.

Strains reported in Liu et al., were sourced from ATCC. Specifically the *Elizabethkingia miricola* (ATCC Strain no: 33958) strain was acquired. Genomic DNA was extracted using the bacterial DNA extraction kits (Qiagen). There was some confusion as to the actual source of the enzymes reported by Clausen et al., and which exact strains they came from. This was because there were discrepancies in the sequences deposited by the authors in NCBI, against their actual reported primer sequences and proteins in the supporting information. After going through these details carefully, we realized that the authors had wrongly reported the strain from which they isolated the NAG enzyme, which should have been *Elizabethkingia miricola* instead of *Elizabethkingia meningosepticum*. This was confirmed by our own sequencing results which confirmed the former strain. We thus cloned out the relevant galactosidase sequence using Gateway® Cloning Technology by following the manufacturer’s instructions (Invitrogen). Briefly, the NAG gene was amplified by PCR with designed primers (Table S1), then was ligated into TOPO S/D vector and transformed into Top10 cells. After verification by colony PCR and insert DNA sequencing, LR reaction was performed with pDEST17 vector and then the final construct was transformed into NovaBlue(DE3) competent cells (NovaGen). The construct was subjected to random mutagenesis using a GeneMorph II EZclone Domain mutagenesis kit (Stratagene). 10ng of template plasmid DNA per reaction was used to establish the mutant library, with the expected mutation frequency of 9-16 mutations/kb. The mutant plasmid library was transformed into NovaBlue(DE3) competent cells by electroporation to get the high transformation efficiency estimated at 1×10^8 cfu/ng. Then we randomly sequenced 20 colonies from the electroporated mutant library, each provided diverse sequences confirming the vast diversity generated.

2.2 Hypotonic Loading prior to Flow Cytometric Screening.

1 ml of induced cell culture was spun down and re-suspended in 25μl of hypotonic buffer (50% LB in autoclaved water) and warmed at 37°C for 5 min. 200μM of probes were prepared and warmed at 37°C for 5 min. 25μl of the pre-warmed probe was added to pre-warmed cells to a final probe concentration of 100μM. The cell suspension was re-incubated 37°C for a further 1-5 min. A 10-fold dilution with ice-chilled 100% LB produced a final volume of 500μl with final probe concentration of 10μM and 92.5% LB. The tubes were kept on ice at 4°C for 30 min before FACS analysis.

2.3 Protein Expression and Purification.

Proteins were successfully purified with Ni-NTA agarose using standard affinity purification procedures (Qiagen Ni-NTA handbook). Briefly, 200μl of Ni-NTA agarose bead suspension (Qiagen) was pipetted into a fritted column and washed. 10 ml of the lysate was added and incubated with the Ni-NTA beads for 2 h at 4°C. The column was gently mixed during this incubation. The flow through was then drained. Repeated washing was then performed to remove background proteins. Elution was performed in PBS buffer containing 200mM imidazole. Typical protein purification gels obtained are as shown in Fig. 2A of the maintext.
2.4 Kinetic Analysis.

The enzyme activity assays were carried out in 50μl reactions, containing 1×PBS (pH7.4) with either 2nM and 20nM of the purified NAG protein or approximately 10μg of crude NAG-overexpressing cell lysates. Coumarin Aα (4-Methylumbelliferyl 2-acetamido-2-deoxy-α-D-galactopyranoside; purchased from Carbosynth with Cat #EM04782) or Probe A working stocks in DMSO were added to varying final concentrations. The mixtures were monitored at λ\text{360nm}/420nm and λ\text{490nm}/520nm at 37°C on a BioTek Microplate reader, in a time-dependent manner (Fig. S1). For comparison of the activity of NAG protein and the mutant, assays were performed in triplicate with Coumarin Aα, over 8 different probe concentrations over 5μM and 200μM, along with a 10nM of total enzyme. The initial enzyme kinetics data and plot was monitored by Gen5 software at λ\text{360nm}/420nm. These were plotted on a Michealis-Menten plot, as shown in Fig. 4A of the maintext, using the Graphpad software. Non-linear regression was used to derive the K_M and k_cat values for both the wildtype and variant NAG enzyme, as shown in Fig. 4B of the main text.

2.5 Protein Labeling and In-Gel Fluorescence Scanning.

After enzyme activity assay, the samples were separated by SDS-PAGE, and the gels were scanned on the Typhoon fluorescence gel scanner for detection of fluorescently labeled protein bands under the FITC channel (Fig. S2).
3. Results and Discussion.

Table S1. Primer sequences used/derived in the current study.

<table>
<thead>
<tr>
<th>Primer Type</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAG Forward Primer</td>
<td>CACCCCTAAAAAGGTAAGAATAGC</td>
</tr>
<tr>
<td>NAG Reverse Primer</td>
<td>GGGTTAGTAGTCGTCATTTATTGC</td>
</tr>
<tr>
<td>Forward Sequencing Primer (T7-F)</td>
<td>TAATACGACTCACTATAGGG</td>
</tr>
<tr>
<td>Reverse Sequencing Primer (T7-R)</td>
<td>TATGCTAGTTATTGCTCAG</td>
</tr>
</tbody>
</table>

Fig. S1 Continuous microplate monitoring of enzymatic activity. (A) Structure of commercial Coumarin Aα fluorogenic substrate. Time-dependent protein activity of NAG using (B) Coumarin Aα substrate and (C) Probe A. Left panels are assays conducted with purified wildtype NAG protein and right panels are assays with induced NAG cell lysates. Red curves depict assays run enzyme, blue curves depict blanks with probe only. The activity was monitored at $\lambda_{360\text{nm}/420\text{nm}}$ and $\lambda_{490\text{nm}/520\text{nm}}$ respectively, using 20 nM of enzyme/approximately 10μg of cell lysate and 20 μM of probe, buffered in PBS (pH 7.4).
Fig. S2 In-gel fluorescence scanning of NAG-overexpressing bacterial cell lysates upon labeling with **Probe A**. MW marker; Lane 1: **Probe A** only; Lane 2: NovaBlue(DE3) cell lysates labeled with **Probe A**; Lane 3: NAG-overexpressing NovaBlue(DE3) cell lysates labeled with **Probe A**. (Left - Fluorescence image, Right - Coomassie Blue stained image of the same gel). Results indicated that many proteins, including NAG, in the NAG-overexpressing lysate was labeled by the probe, and the labeling was NAG activity-dependent (as no labeling was observed in lane 2).

**Table S2.** Mutations in identified NAG variant

<table>
<thead>
<tr>
<th>Reference Position</th>
<th>Nucleotide Change</th>
<th>Amino Acid Change</th>
<th>Mutation</th>
<th>Location from active site</th>
</tr>
</thead>
<tbody>
<tr>
<td>311</td>
<td>A&gt;T</td>
<td>H104L</td>
<td>Yes</td>
<td>Within 10 angstroms</td>
</tr>
<tr>
<td>528</td>
<td>A&gt;T</td>
<td>T176T</td>
<td>Silent</td>
<td></td>
</tr>
<tr>
<td>1041</td>
<td>T&gt;A</td>
<td>S347R</td>
<td>Yes</td>
<td>Within 20 angstroms</td>
</tr>
</tbody>
</table>

**Fig. S3** Sequence information for identified variant. The chromatograms indicate the bi-directional reads with the highlighted yellow bands indicating each of the non-synonymous mutations identified.
**Table S3.** List of 5 representative variants sequenced at every round.

<table>
<thead>
<tr>
<th>Cell sorting Round</th>
<th>Variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>G191S, I192S, N193G</td>
</tr>
<tr>
<td>0</td>
<td>I192V, N193S</td>
</tr>
<tr>
<td>0</td>
<td>N17T, A24D, I26T, A252V</td>
</tr>
<tr>
<td>0</td>
<td>L395S</td>
</tr>
<tr>
<td>0</td>
<td>K366A</td>
</tr>
<tr>
<td>1</td>
<td>D82Y, Y83H</td>
</tr>
<tr>
<td>1</td>
<td>L319R</td>
</tr>
<tr>
<td>1</td>
<td>N390D</td>
</tr>
<tr>
<td>1</td>
<td>V21I, T176K</td>
</tr>
<tr>
<td>1</td>
<td>K366R</td>
</tr>
<tr>
<td>2</td>
<td>H341Y *</td>
</tr>
<tr>
<td>2</td>
<td>G267E *</td>
</tr>
<tr>
<td>2</td>
<td><strong>H104L, S347R</strong></td>
</tr>
<tr>
<td>2</td>
<td>K20E *</td>
</tr>
<tr>
<td>2</td>
<td>L145P, M146V *</td>
</tr>
</tbody>
</table>

*The other round 2 variants sequenced did not show elevated activity when tested and had difficulties being expressed under native conditions.

4. $^1$H and $^{13}$C NMR.
Electronic Supplementary Material (ESI) for Chemical Communications
This journal is © The Royal Society of Chemistry 2013
5. References.