## A potential fortuitous binding of inhibitors of an inverting family GH9 $\beta$ -glycosidase derived from isofagomine.

Solange Moréra\*<sup>a</sup>, Armelle Vigouroux<sup>a</sup>, Keith A. Stubbs\*<sup>b</sup>

<sup>a</sup>Laboratoire d'Enzymologie et Biochimie Structurales (LEBS), CNRS, avenue de la terrasse, 91198-Gif-sur-Yvette, France. E-mail: morera@lebs.cnrs-gif.fr Tel: +33 1 69823470

<sup>b</sup>Chemistry M313, School of Biomedical, Biomolecular and Chemical Sciences, The University of Western Australia, 35 Stirling Hwy, Crawley, WA Australia, 6009. E-mail: kstubbs@cyllene.uwa.edu.au Tel: +61 8 6488 2725

#### Experimental

#### General

<sup>1</sup>H and <sup>13</sup>C NMR spectra were recorded on a Bruker ARX500 (500 MHz for <sup>1</sup>H and 125 MHz for <sup>13</sup>C). Elemental analyses of all synthesized compounds used in enzyme assays were performed at the Australian National University Microanalytical Facility. Flash chromatography was performed on BDH silica gel with the specified solvents. Thin-layer chromatography (TLC) was effected on Merck silica gel 60 F<sub>254</sub> aluminium-backed plates that were stained by heating (>200 °C) with 5% sulfuric acid in EtOH. Percentage yields for chemical reactions as described are quoted only for those compounds that were purified by column chromatography, and purity was assessed by TLC or <sup>1</sup>H NMR spectroscopy.

(3R, 4R, 5R)-3-Acetoxy-N-benzyloxycarbonyl-4-hydroxy-5-(hydroxymethyl)piperidine 6 Isofagomine<sup>1</sup> (1.0 g, 6.8 mmol) was treated with benzyl chloroformate (1.1 ml, 7.4 mmol) and NaHCO<sub>3</sub> (700 mg) in H<sub>2</sub>O/MeOH/THF (2:1:1, 10 ml) and the solution stirred at room temperature (1 h). The mixture was treated with HCl (9 ml of 1 M) and then concentrated. The resulting residue was then treated with PhCH(OMe)<sub>2</sub> (1.1 ml, 7.4 mmol) and CSA (20 mg) in CHCl<sub>3</sub> (20 ml) and was heated under reflux (2 h). The solution was allowed to cool and then treated with pyridine (10 ml) and Ac<sub>2</sub>O (5 ml) and the mixture kept (rt, 2 h). The mixture was guenched with MeOH and concentrated. The residue (2.7 g) was then heated in AcOH-H<sub>2</sub>O (4:1, 20 ml) (70 °C, 1 h). Concentration of the reaction mixture give a pale yellow oil that was subjected to flash chromatography (EtOAc-petrol, 7:3) to yield 6 as a colourless oil (1.8 g, 81%); <sup>1</sup>H NMR (500 MHz, d<sub>6</sub>-DMSO) 1.51-1.55 (m, 1H, H-5), 2.01 (s, 3H, COCH<sub>3</sub>), 2.68-2.73 (m, 2H, H-2,6), 3.33-3.38 (m, 2H, H-2,6), 3.66, 4.05 (2m, 3H, H-4, CH<sub>2</sub>O), 4.44 (m, 1H, OH), 4.56 (m, 1H, OH), 5.04-5.09 (m, 3H, H-3, OCH<sub>2</sub>Ph), 7.30-7.38 (m, 5H, Ph); <sup>13</sup>C NMR (125.8 MHz, d<sub>6</sub>-DMSO) 20.88 (COCH<sub>3</sub>), 40.56, 44.75, 44.78 (C-2,C-5,C-6), 59.68, 66.31, 69.70, 72.93 (C-3,C-4,CH<sub>2</sub>O,OCH<sub>2</sub>Ph), 127.39, 127.81, 128.40, 136.84 (Ph), 154.32 (NCO), 169.97 (COCH<sub>3</sub>); HR-MS m/z (FAB) 324.1468; [M+H]<sup>+</sup> requires 324.1447. Anal. calcd for C<sub>16</sub>H<sub>21</sub>NO<sub>6</sub>: C, 59.43; H, 6.55. Found: C, 59.29; H, 6.48%.

#### (3R, 4R, 5R)-3-Acetoxy-5-acetyloxymethyl-N-benzyloxycarbonyl-4-hydroxypiperidine 7

Acetyl chloride (570 µl, 8.0 mmol) was added to the diol **6** (2.0 g, 6.2 mmol) and pyridine (1.3 ml, 16 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (10 ml) at -30 °C and the resulting mixture allowed to warm to room temperature (3 h). The mixture was quenched with MeOH and subjected to a usual workup (CH<sub>2</sub>Cl<sub>2</sub>), followed by flash chromatography (EtOAc-petrol, 2:3), to yield **7** as a gum (1.8 g, 83%); <sup>1</sup>H NMR (500 MHz, d<sub>6</sub>-DMSO) 1.91-1.96 (m, 1H, H-5), 2.01, 2.04 (2s, 6H, COCH<sub>3</sub>), 2.87-3.05 (m, 2H, H-2,6), 3.54-3.61 (m, 1H, H-4), 4.03-4.12 (m, 2H, H-2,6), 4.25-4.34 (m, 1H, CH<sub>2</sub>O), 4.46 (dd, 1H, *J*<sub>5,CH</sub> 3.5, *J*<sub>CH,CH</sub> 11.1 Hz, CH<sub>2</sub>O), 4.53 (ddd, 1H, *J* 5.0, 8.5, 9.7 Hz, H-3), 5.07 (ABq, 2H, *J* 21.0 Hz, CH<sub>2</sub>Ph), 5.43 (d, 1H, *J*<sub>4,OH</sub> 5.4 Hz, OH), 7.29-7.42, (m, 5H, Ph); <sup>13</sup>C NMR (125.8 MHz, d<sub>6</sub>-DMSO) 20.84, 20.91 (COCH<sub>3</sub>), 41.56, 43.97, 44.32 (C-2,C-5,C-6), 63.11, 66.47, 69.06, 72.42 (C-3,C-4,CH<sub>2</sub>O,OCH<sub>2</sub>Ph), 128.48, 129.35, 129.57, 136.69 (Ph), 154.38 (NCO), 168.33, 169.87 (CO); HR-MS m/z (FAB) 366.1530; [M+H]<sup>+</sup> requires 366.1553. Anal. calcd for C<sub>18</sub>H<sub>23</sub>NO<sub>7</sub>: C, 59.17; H, 6.34. Found: C, 59.31; H, 6.21%.

#### General procedure for glycosylation reactions.

A mixture of the appropriate trichloroacetimidate (1.5 mmol), the alcohol 7 (360 mg, 1.0 mmol) and crushed 4A sieves in dry  $CH_2Cl_2$  (5 ml) was cooled to -30 °C. TMSOTf (10 µl) was added and the mixture was stirred for one hour at -30 °C and then allowed to warm to room temperature over one hour. Triethylamine was then added to neutralise the solution and the mixture was then filtered through Celite. Concentration of the filtrate and flash chromatography (EtOAc:hexane 1:1) of the resultant residue using the appropriate solvents gave the desired compounds.

## (*3R*, 4*R*, 5*R*)-3-Acetoxy-5-acetyloxymethyl-*N*-benzyloxycarbonyl-4-[(tetra-*O*-acetyl-β-D-glucopyranosyl)oxy]piperidine 11

Using the trichloroacetimidate **8** (740 mg)<sup>2</sup> and the alcohol **7** gave the title compound **11** (500 mg, 72%);  $[\alpha]_D$  -6.3°; lit.<sup>3</sup> -6.4°; The <sup>1</sup>H and <sup>13</sup>C NMR spectra were consistent with that previously reported<sup>3</sup>; HR-MS m/z (FAB) 696.2511;  $[M+H]^+$  requires 696.2504. Anal. calcd for C<sub>32</sub>H<sub>41</sub>NO<sub>16</sub>: C, 55.25; H, 5.94. Found: C, 55.11; H, 5.87%.

## (3R, 4R, 5R)-3-Acetoxy-5-acetyloxymethyl-*N*-benzyloxycarbonyl-4-{[(tetra-*O*-acetyl- $\beta$ -D-glucopyranosyl)-(1 $\rightarrow$ 4)-*O*-(tri-*O*-acetyl- $\beta$ -D-glucosyl)]oxy}piperidine 12

Using the trichloroacetimidate **9**  $(1.2 \text{ g})^4$  and the alcohol 7 gave the title compound **12** (605 mg, 61%);  $[\alpha]_D$  -18.8°; lit.<sup>3</sup> -19°; The <sup>1</sup>H and <sup>13</sup>C NMR spectra were consistent with that previously reported<sup>3</sup>; HR-MS m/z (FAB) 984.3361;  $[M+H]^+$  requires 984.3349. Anal. calcd for C<sub>44</sub>H<sub>57</sub>NO<sub>24</sub>: C, 53.71; H, 5.84. Found: C, 53.75; H, 5.92%.

# (3*R*, 4*R*, 5*R*)-3-Acetoxy-5-acetyloxymethyl-*N*-benzyloxycarbonyl-4-{[(tetra-*O*-acetyl- $\beta$ -D-glucosyl)-(1 $\rightarrow$ 4)-*O*-(tri-*O*-acetyl- $\beta$ -D-glucosyl)-(1 $\rightarrow$ 4)-*O*-(tri-*O*-acetyl- $\beta$ -D-glucosyl)]oxy}piperidine 13

Using the trichloroacetimidate **10**  $(1.6 \text{ g})^5$  and the alcohol **7** gave the title compound **13** (710 mg, 56%);  $[\alpha]_D$  -12.9°; lit.<sup>3</sup> -12.6°; The <sup>1</sup>H and <sup>13</sup>C NMR spectra were consistent with that previously reported<sup>3</sup>; HR-MS m/z (FAB) 1272.4181;  $[M+H]^+$  requires 1272.4194. Anal. calcd for C<sub>56</sub>H<sub>73</sub>NO<sub>32</sub>: C, 52.87; H, 5.78. Found: C, 52.80; H, 5.71%.

#### General procedure for deprotection reactions.

Sodium hydroxide (1.0 g) was added to a mixture of the compound (200 mg) in MeOH- $H_2O$  (2:1, 10 ml) and the mixture refluxed (3 h). The mixture was then cooled, concentrated and the resulting residue taken up in  $H_2O$  (10 ml). The resulting solution was then brought to pH 5 by the addition of HCl (2 M). The mixture was applied to a column of cation-exchange resin (Dowex 50W-X2, H<sup>+</sup> form), washed with  $H_2O$  and then eluted with aqueous  $NH_3$  (1.5 M). The eluate was concentrated, taken up in  $H_2O$  (5 ml), and applied to an anion exchange column (Sephadex-DEAE A-25) and eluted with  $H_2O$ . The fractions containing the desired material were pooled and concentrated to give the desired compounds.

#### (3R, 4R, 5R)-4-(β-D-Glucopyranosyl)oxy-3-hydroxy-5-(hydroxymethyl)piperidine 1

Using the acetate **11** (200 mg) gave the title compound **11** (40 mg, 45%);  $[\alpha]_D$  -3.3°; lit.<sup>3</sup> - 3.3°; The <sup>1</sup>H and <sup>13</sup>C NMR spectra were consistent with that previously reported<sup>3</sup>; HR-MS m/z (FAB) 310.1509;  $[M+H]^+$  requires 310.1502. Anal. calcd for C<sub>12</sub>H<sub>23</sub>NO<sub>8</sub>: C, 46.60; H, 7.49. Found: C, 46.49; H, 7.55%.

## (3*R*, 4*R*, 5*R*)-4-[( $\beta$ -D-Glucopyranosyl)-(1 $\rightarrow$ 4)-*O*-( $\beta$ -D-glucosyl)]oxy-3-hydroxy-5-(hydroxymethyl)piperidine 2

Using the acetate **12** (200 mg) gave the title compound **2** (53 mg, 55%);  $[\alpha]_D$  -10.1°; lit.<sup>3</sup> - 10.4°; The <sup>1</sup>H and <sup>13</sup>C NMR spectra were consistent with that previously reported<sup>3</sup>; HR-MS m/z (FAB) 472.2009;  $[M+H]^+$  requires 472.2030. Anal. calcd for C<sub>18</sub>H<sub>33</sub>NO<sub>13</sub>: C, 45.86; H, 7.06. Found: C, 45.70; H, 7.15%.

#### (3*R*, 4*R*, 5*R*)-4-[(β-D-Glucopyranosyl)-(1→4)-*O*-(β-D-glucosyl)-(1→4)-*O*-(β-D-glucosyl)]oxy-3-hydroxy-5-(hydroxymethyl)piperidine 3

Using the acetate **13** (200 mg) gave the title compound **3** (60 mg, 58%);  $[\alpha]_D$  -1.5°; lit.<sup>3</sup> - 1.6°; The <sup>1</sup>H and <sup>13</sup>C NMR spectra were consistent with that previously reported<sup>3</sup>; HR-MS m/z (FAB) 634.2541; [M+H]<sup>+</sup> requires 634.2558. Anal. calcd for C<sub>24</sub>H<sub>43</sub>NO<sub>18</sub>: C, 45.50; H, 6.84. Found: C, 45.26; H, 6.97%.

#### Kinetic Analysis of AaCel9a

*Aa*Cel9A kinetics and  $K_i$  determinations were performed using PNP-cellobioside as substrate with release of the dinitrophenolate measured at 400 nm with enzyme prepared as previously described<sup>6</sup>. All reactions were performed using conditions as previously described<sup>7</sup> and  $K_i$  values were determined from Dixon plots using inhibitor concentrations spanning 1/3 to 3 times the  $K_i$  with the PNP-cellobioside concentration near the  $K_M$  value (3.0 mM). Kinetics were fitted using GRAFIT 5.0 (Erithacus Software, Horley, U.K.).

#### Crystallization and data collection

Recombinant cleaved (without His-tag) AaCel9A was expressed, purified and crystallized as described previously.<sup>6</sup> The orthorhombic crystals were manually obtained using home made solutions in hanging drops containing 20% (4S)-2-methyl-2,4-pentanediol (MPD) with 50 mM Hepes pH 7.5 over pits containing 40% MPD and 100 mM Hepes pH 7.5. Crystals were soaked with 25 mM derivative compounds such as cellobiose-like, cellotriose-like or cellotetraose-like isofagomine during 1 minute before being transferred in a loop and flash-frozen in liquid nitrogen. Data collection experiments were carried out at 100 K on the PROXIMA I beamline at SOLEIL (Saint-Aubin, France). Diffraction intensities were integrated with the program XDS.<sup>8</sup> Data collection and processing statistics are given in Supporting Table.

#### Structure determination, refinement and final model

Refinement details of the three structures are shown in Supporting Table. Molecular graphics images were generated using PYMOL (http://www.pymol.org).

As the soaked crystals are isomorphous with those obtained previously<sup>6</sup>, preliminary phases were calculated using the coordinates of the unliganded form (PDB code 3GZK). The resulting electron density map revealed bound glycosyl and isofagomine units for each case. For the cellobiose-like isofagomine soak, one compound binds at -2 and -1 subsites and another at +1 and +2 subsites with an isofagomine moiety in the catalytic -1 and +1 subsites. For the cellotriose-like isofagomine soak, a cellotriose-like isofagomine binds in the -3, -2 and -1 subsites with the isofagomine moiety in the catalytic -1 subsite. A cellobiose occupies the +1 and +2 subsites. For the cellotetraose-like isofagomine was identified from -4 to -1 subsites with again the isofagomine moiety at -1. In this structure, a cellobiose-like isofagomine was also identified at +1 and +2 subsites with the isofagomine unit bound at +2. Refinement was performed using BUSTER<sup>9</sup> and electron density maps were evaluated using COOT.<sup>10</sup>

88 2 ww	AaCel9A-glucosyl	AaCel9A - cellobiosyl	AaCel9A -
	isofagomine 1	isofagomine 2	cellotriosyl
	( <b>3RX8</b> )	(3RX5)	isofagomine 3
			(3RX7)
Precipitant and	MPD	MPD	MPD
Crystallisation method	soaking	soaking	soaking
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2	$P 2_1 2_1 2_1$	P 2 <sub>1</sub> 2 <sub>1</sub> 2
Cell parameters			
a (Å)	85.4	85.2	85.2
b (Å)	129.3	129.5	129.1
c (Å)	49.4	49.3	49.2
Resolution range (Å)	40-2.56 (2.71-2.56)	40-1.99 (2.11-1.99)	40-2.02 (2.14-
Total reflections	54907 (8439)	209732 (30763)	2.02) 151475 (21723)
Unique reflections	16969 (2718)	38137 (5842)	36125 (5663)
Completeness (%)	92 1 (93 6)	99 1 (95 8)	99 1 (97 4)
	92.1(93.0) 8 $4(2.15)$	85 (26)	87(22)
$1/O$ D $(0/)^a$	0.4(2.13)	0.5(2.0)	0.7(2.2)
$\mathbf{K}_{\text{merge}} (70)$	14 (03.3)	15.4 (69)	13.7 (72)
$\mathbf{K}_{\text{cryst}}$ (%)	10.5	10.0	1/.1
K <sub>free</sub> (%)	23.4	20.1	20.5
R.m.s deviation	0.01	0.01	0.01
bond length (A)	0.01	0.01	0.01
bond angles (°)	1.1	1	1
Solvent molecules	70	183	168
Average B ( $Å^2$ )			
protein	32.7	21.9	24
9MR	35.6 and 50.8		37.6
G2I		21.9	
CBI		28.5	
G3I			33.6
calcium ion	26	13.4	15.4
zinc ion	29.3	17.2	19.1
solvent	32.9	26.7	28.4

#### Supporting Table. Data collection and refinement statistics

9MR (cellobiose-like isofagomine), G2I (cellotriose-like isofagomine), CBI (cellobiose), G3I (cellotetraose-like isofagomine),

Values in parentheses are for outer resolution shell <sup>a</sup>  $R_{merge} = \Sigma_{hkl}\Sigma_i |I_i(hkl) - \langle I(hkl) \rangle | / \Sigma_{hkl} \Sigma I_i(hkl)$ , where  $I_i(hkl)$  is the *i* th observed amplitude of reflection hkl and  $\langle I(hkl) \rangle$  is the mean amplitude for all observations *i* of reflection hkl. <sup>b</sup> Rcryst =  $\Sigma ||Fobs| - |Fcalc|| / \Sigma |Fobs|$ 

<sup>c</sup> 5% of the data were set aside for free R-factor calculation.

#### Data bank accession codes

The atomic coordinates and structure factors of the *Aa*Cel9A in complex with **1**,**2** and **3** have been deposited in the Protein Data Bank (http://www.rcsb.org) under accession codes 3RX8, 3RX5 and 3RX7 respectively.

#### **Supporting Figure 1**



Cellotriose-like isofagomine **2** bound to the active site cleft of AaCel9A in a F<sub>o</sub>-F<sub>c</sub> omit map contoured at 2.5 $\sigma$ . Binding sites from -3 to +2 are indicated.

#### **Supporting Figure 2**



Cellotetraose-like isofagomine **3** bound to the active site cleft of *Aa*Cel9A in a  $F_0$ - $F_c$  omit map contoured at 2.5 $\sigma$ . Binding sites from -4 to +2 are indicated.

#### **Supporting Figure 3**



Overlay of the three AaCe9A complexes. Cellobiose-like isofagomine 1 depicted in green, cellotetriose-like isofagomine 2 in magenta and cellotetraose-like isofagomine 3 in cyan. Binding sites from -4 to +2 are indicated.

#### References

- 1. E. D. Goddard-Borger and R. V. Stick, Aust. J. Chem, 2007, 60, 211.
- 2. R. R. Schmidt, J. Michel and M. Roos, Liebigs Ann. Chem., 1984, 1343.
- 3. J. M. Macdonald, R. V. Stick, D. M. G. Tilbrook and S. G. Withers, Aust. J. Chem., 2002, 55, 747.
- 4. F. J. Urban, B. S. Moore and R. Breitenbach, Tetrahedron Lett., 1990, 31, 4421.
- 5. T. K. Ritter, K. T. Mong, H. Liu, T. Nakatani and C.-H. Wong, *Angew. Chem. Int. Ed.*, 2003, **42**, 4657.
- 6. K. Eckert, A. Vigouroux, L. Leggio and S. Moréra, J. Mol. Biol., 2009, 394, 61.
- 7. K. Eckert, F. Zielinski, L. L. Leggio and E. Schneider, *Appl. Microbiol. Biotechnol.*, 2002, **60**, 428.
- 8. W. Kabsch, J. Appl. Cryst., 1993, 26, 795.
- 9. E. Blanc, P. Roversi, C. Vonrhein, C. Flensburg, S. M. Lea and G. Bricogne, Acta Crystallogr. D Biol. Crystallogr., 2004, 60, 2210.
- 10. P. Emsley and K. Cowtan, Acta Crystallogr. D Biol. Crystallogr., 2004, 60, 2126.











