Understanding functional dynamics and conformational stability of beta-glycosidases

Ben Bdira, F.

Appendix I

```
# Lenght:      602
# Identity:    88/602 (14.6%)
# Similarity: 161/602 (26.7%)
# Gaps:       238/602 (39.5%)

EGCII    1 ------------------------------------------------------------MSG
GBA      1 EFARPCIPKSFGYSSVVCVCNATYCDSFDPPTFPALGTFSRYESTRSGRRMELSMGPIQA

EGCII    4 SGSGSGTALTPSYLKDDDGRSLILRGFNTASSAKSAPDGYPQFTEADLAREYADMGTNFV
GBA     61 NHTGTGLLLTLQ----QPEQKFQKVKGFGGAMTDAAALNILALR-----------------

EGCII   64 RFLISWRSVEPAPGVYDQQYLDRVEDRVGWYAERGYKVMLDMHQDVYSGAITPEGNS---
GBA     99 --------SPPAQNLLKSYFS--EEGIGYNIIRVPMASCDFSIRTYTYADTPDDFQLHN

EGCII  121 ---------------------------------------------GNGACAIGNGAPAWAT
GBA    149 FSLPEEDTKLKIPLIHRALQLAQRPVSLLASPWTSPTWLKTNGAVNGKGSL---------

EGCII  137 YMDGLPVEP-QPRWELYYIQPGVMRAFDNFWNTTGKHPELVEHYAKAWRAVADRFADNDA
GBA    200 --KGQPGDIYHQTWARYFVKF---------LDAY------AEHKLQFW------------

                *
EGCII  196 VVAYDLMNEPFGGSLQGPAEEAGPLAAMYQRTTD--------------AIRQVDQDTWVC
GBA    231 --AVTAENEPSAGLISGYPFQCLGFTPEHQRDFIARDLGPTLANSTHHNVRLLMLDDQRL
                *
EGCII  242 VAEQAIGVNQGLPSGLTKLDDPRAGQQRIA-----YCPHLYPLPLDIGDGHEGLARTIT-
GBA    289 ILEHWAKV---------VITDPEAAKYVHGIAVHWYLDFLAPAKATLGETHRLFPNTMLF

EGCII  296 -------------DVTLDAWRANTA----------------------------HTARVL
GBA    340 ASEACVGSKFWEQSVRLGSWDRGMQYSHSIITNLLYHVVGWTDWNLALNPEGGPNWVRNF
             *
                *
EGCII  314 GDVPIILGEFGLDTTLPGARDYIERVYGTAREMGAGVSYWSSDPGPWGPYIPDGTQTLLV
GBA    400 VDSPIIV-DITKDTFYKQPMF--------------------YHLGHESKFIPEGSQRVGL

EGCII  374 DTLNKPYPRAVAGTPTLWSSTSDRLQLTIEPDAAITAPTEIYLPEAGFPGDVHVE-GADV
GBA    439 VASQKNDLDAVALMHPDGSAVVVVL-----NRSSKDVPLTIKDPAVGFLETISPGYSIHT

EGCII  433 VGWDRQSRDLTVRTPADSGNVTVTVTPAA
GBA    494 YLWHRQLLVDTM-----------------
```

Figure A1.1: EGCII and GBA primary sequences alignment using Needleman-Wunsch algorithm from EMBOSS Needle server (emboss.open-bio.org). Identical residues are highlighted on black background, conservative changes are highlighted in grey. The acid /base and nucleophile catalytic residues are marked with stars.

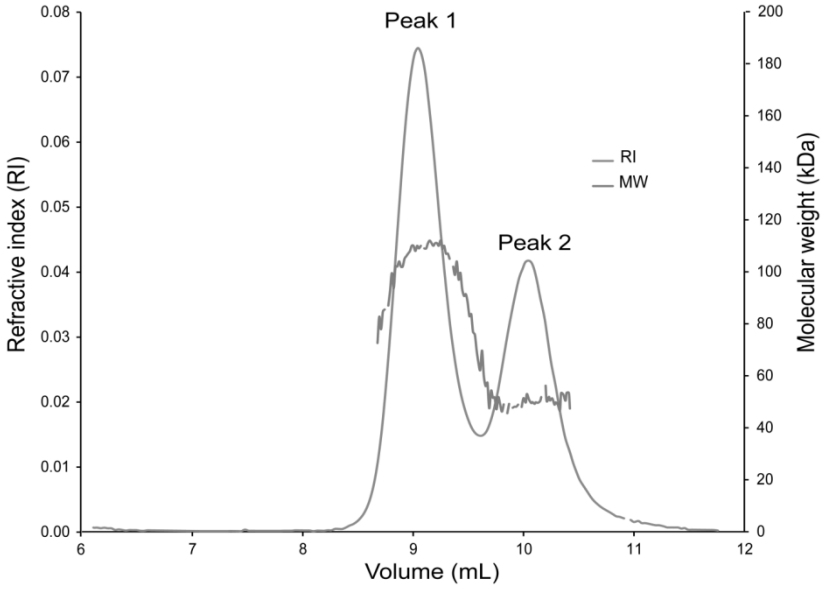
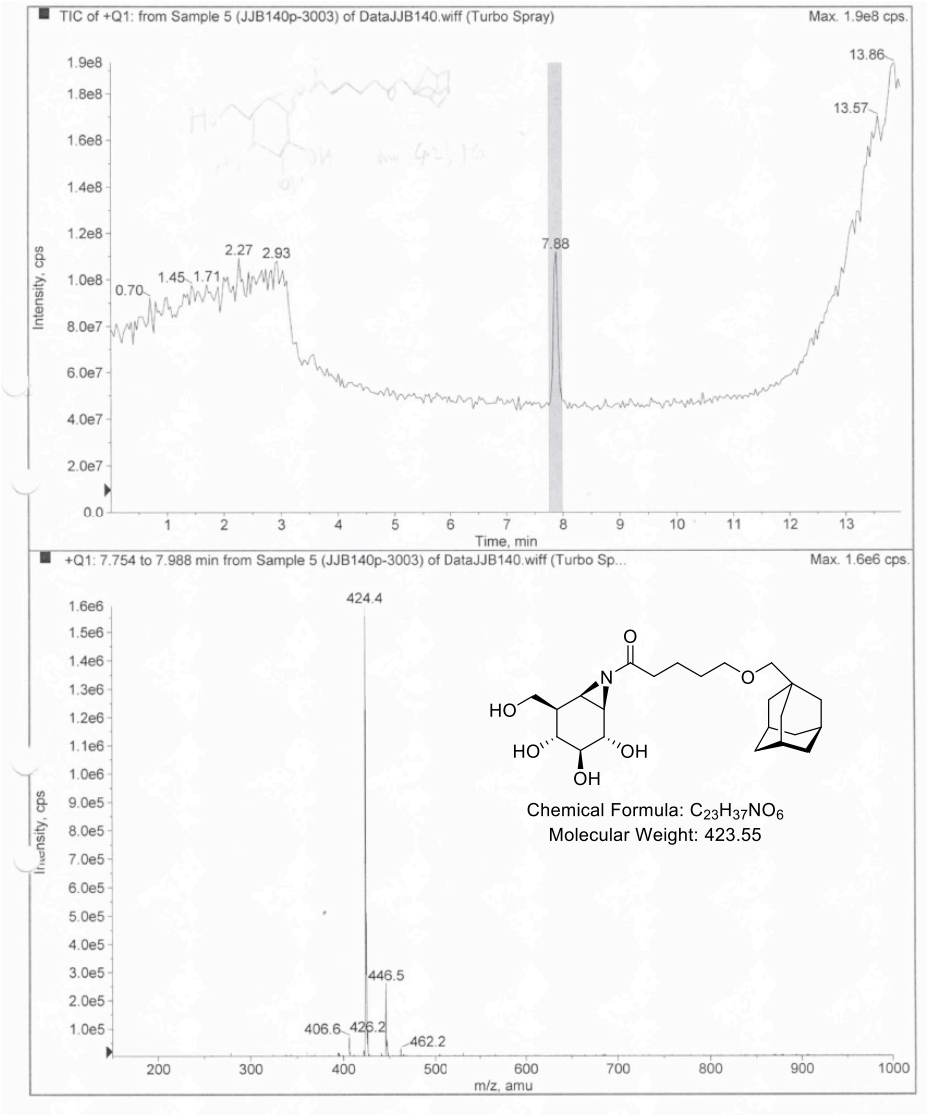


Figure A1.2. EGCII size exclusion chromatogram analyzed by Multi Angle Laser Light Scattering.

Results

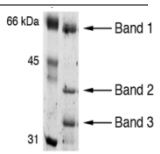
	Peak 1	Peak 2
Volume (mL)	8.75-9.42	9.64-10.42
Injection Mass (g)	$1.0\ 10^{-4}$	$1.0\ 10^{-4}$
Calc. Mass (g)	$3.82\ 10^{-5}$	$2.47\ 10^{-5}$
dn/dc (mL/g)	0.186	0.186
Polydispersity (Mw/Mn)	1.003 +/- 0.033	1.002 +/- 0.086
Molar Mass Moments (g/mol)	$1.062\ 10^{+5}$	$5.0\ 10^{+4}$

LC/MS spectra of compound **8**



Appendix II

Table A2.1: LC-MS identification of peptides in Band 1



Start-End	m/z (exp)	z	ppm	Score	Peptide sequence
117-145	9.655.467	3	3	30	VKGFGGAMTDAAALNILALSPPAQNLLLK
119-145	13.422.294	2	1	80	GFGGAMTDAAALNILALSPPAQNLLLK*
146-159	8.244.048	2	1	81	SYFSEEGIGYNIIR
160-170	6.417.996	2	0	53	VPMASCDFSIR
195-202	3.305.582	3	-2	48	LKIPLIHR
197-202	3.747.444	2	-2	48	IPLIHR
203-225	8.551.525	3	1	22	ALQLAQRPVSLLASPWTSPTWLK
234-250	6.386.585	3	0	36	GSLKGQPGDIYHQTWAR
238-250	7.648.692	2	1	54	GQPGDIYHQTWAR
251-263	5.442.807	3	1	44	YFVKFLDAYAEHK
255-263	3.651.819	3	0	43	FLDAYAEHK
264-296	12.509.374	3	-1	138	LQFWAVTAENEPSAGLLSGYPFQCLGFTPEHQR
297-301	3.111.708	2	-2	22	DFIAR
317-324	5.022.659	2	0	63	LLMLDDQR
325-332	4.893.011	2	2	31	LLLPHWAK
325-342	6.677.226	3	0	38	LLLPHWAKVVLTDPEAAK
333-342	5.217.918	2	-2	10	VVLTDPEAAK
333-360	7.816.724	4	0	37	VVLTDPEAAKYVHGIAVHWYLDFLAPAK
343-360	10.505.557	2	-1	86	YVHGIAVHWYLDFLAPAK
343-368	7.421.416	4	-1	28	YVHGIAVHWYLDFLAPAKATLGETHR
369-385	9.364.572	2	2	85	LFPNTMLFASEACVGSK
386-392	4.762.372	2	-1	39	FWEQSVR
393-398	3.671.845	2	-1	40	LGSWDR
399-434	13.853.512	3	1	116	GMQYSHSIITNLLYHVVGWTDWNLALNPEGGPNWVR*
435-447	7.309.018	2	-2	89	NFVDSPIIVDITK
435-452	7.057.049	3	0	24	NFVDSPIIVDITKDTFYK
448-452	3.371.629	2	-1	15	DTFYK
453-464	4.979.123	3	1	24	QPMFYHLGHFSK
465-472	4.672.421	2	-2	15	FIPEGSQR
473-480	4.012.456	2	1	51	VGLVASQK
481-502	7.744.033	3	2	84	NDLDAVALMHPDGSAVVVVLNR*
503-512	3.632.164	3	-1	16	SSKDVPLTIK
506-512	3.932.421	2	0	14	DVPLTIK

Table A2.2: LC-MS identification of peptides in Band 3

Start-End	m/z (exp)	z	ppm	Score	Peptide sequence
234-250	6.386.578	3	-1	42	**G**SLKGQPGDIYHQTWAR
238-250	7.648.693	2	1	64	GQPGDIYHQTWAR
238-250	7.648.694	2	2	69	GQPGDIYHQTWAR
251-263	5.442.804	3	0	45	YFVKFLDAYAEHK
255-263	5.472.696	2	1	49	FLDAYAEHK
264-296	12.509.392	3	1	92	LQFWAVTAENEPSAGLLSGYPFQCLGFTPEHQR
302-316	5.446.115	3	-3	41	DLGPTLANSTHHNVR
317-324	5.022.657	2	0	63	LLMLDDQR
325-332	4.893.012	2	2	48	LLLPHWAK
325-342	6.677.218	3	-1	56	LLLPHWAKVVLTDPEAAK
343-360	10.505.557	2	-1	97	YVHGIAVHWYLDFLAPAK
343-368	7.421.419	4	-1	56	YVHGIAVHWYLDFLAPAKATLGETHR
369-385	9.364.542	2	-1	105	LFPNTMLFASEACVGSK
386-392	4.762.373	2	-1	52	FWEQSVR
399-434	13.800.196	3	1	117	GMQYSHSIITNLLYHVVGWTDWNLALNPEGGPNWVR
435-447	7.309.021	2	-2	92	NFVDSPIIVDITK
453-464	7.463.647	2	1	63	QPMFYHLGHFSK
473-480	4.012.456	2	1	60	VGLVASQK
473-502	7.726.692	4	-2	46	VGLVASQKNDLDAVALMHPDGSAVVVVLNR
473-505	8.482.103	4	0	60	VGLVASQKNDLDAVALMHPDGSAVVVVLNRSSK
481-502	11.611.033	2	3	90	NDLDAVALMHPDGSAVVVVLNR*

G noting the N-terminus residue of Band 3; asterisk * (Asterisk) noting Methionine oxidation